EMC® Networked Storage

Topology Guide

P/N 300-004-456
REV A14

EMC Corporation
Corporate Headquarters:
Hopkinton, MA 01748-9103
1-508-435-1000
www.EMC.com

Copyright © 2001-2007 EMC Corporation. All rights reserved.

Published December, 2007

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

All other trademarks used herein are the property of their respective owners.

Contents

Preface 31

Chapter 1 Fibre Channel SAN Topologies

Fibre Channel topology overview 36
Simple Fibre Channel SAN topologies 50
Single switch fabrics 50
Two switch fabrics 79
Blade switch with direct attached storage 129
Complex Fibre Channel SAN topologies 146
Four switch full mesh 147
Compound core edge switches 199
Heterogeneous switch interoperability 256
FICON connectivity 365
Overview 365
General information 366

Chapter 2 iSCSI SAN Topologies

TCP/IP technology overview 386
IPv6 386
TCP terminology 397
TCP error recovery 400
Network congestion 403
IP Security 404
iSCSI technology overview 405
iSCSI discovery 405
Digests 406
iSCSI error recovery 406
iSCSI security 407

iSCSI solutions 410
General best practices 410
General supported configurations 411
EMC native iSCSI targets 412
Configuring iSCSI targets 414
Bridged solutions 416
Summary 429

Chapter 3 Extended Distance

Distance extension technologies overview 432
Early implementations of SAN environments 432
DWDM 433
CWDM 436
SONET 437
GbE 439
TCP/IP 440
TCP error recovery 444
Network congestion 447
IP security 447
Tunneling and IPSec 448
Distance extension consideration 450
Link speed 450
Data buffering and flow control 451
TCP/IP 461
Active and passive devices 462
FC-SONET/GbE/IP 469
IP-based distance extension solutions 471
General best practices 471
EMC-Brocade distance extension solution 472
Configuring IPSec 478
Fast Write and tape pipelining 480
EMC-Cisco MDS distance extension solution 482
EMC-Brocade M-Series distance extension solution 484
EMC-QLogic distance extension solution 489
Summary 493

Chapter 4 Secure SANs

Understanding how to secure your SANs 496
Security attacks against SANs 498
Snooping (Eavesdropping) 498

Spoofing (Modification / Fabrication) 500
Denial-of-service (Disruption) 500
Secure SAN architecture, components, and mechanisms 502
SAN architectures 503
Security components 504
Security mechanisms 504
Building secure SANs 514
Design considerations 514
Security implementation examples 516
EMC Celerra iSCSI data security 516
Brocade 518
Brocade M-Series 524
Cisco 529
EMC Select security appliances 535
EMC Select Decru DataFort FC-Series security appliance 537
EMC Select Neoscale CryptoStore security appliance 550

Chapter 5 Backup and Recovery

Overview of backups 558
Direct-attached backups 559
LAN-based backups 561
LAN-free backups 566
Serverless backups 571
Backup over long distances using FCIP and routers 574
FCIP tape acceleration with EMC Connectrix MDS switches 576
Enabling FCIP tape acceleration on Connectrix MDS switches 578
Brocade 7500 and CDL over IP with Fastwrite acceleration and tape pipelining 579
Case study notes 579
NAS backups 582
Backup to disk 583
Backup media: Tape libraries 584
Mirrored fabric backup solution 586
Tapes and fabrics 591
SCSI tape 591
Fibre Channel tape 591
Sharing tape and disk on the same HBA 593
Tape library 595

Chapter 6 Clusters

Requirements for Oracle 10*g* R1 RAC with ASM on AIX 5L 598
Setup and configuration summary 599
Oracle 10*g* RAC supported configuration 603
Oracle 10*g* release 2 (10.2) RAC installation notes 604
SLES9 Oracle 10g with EMC arrays and EMC PowerPath 607
Configuration procedure 607

Chapter 7 SAN Management

Overview of SAN Management 610
Define your needs 610
Enterprise storage management 613
EMC ControlCenter 613
Fabric management 619
Interface options 620
Discovery 621
Zoning 629
Alerting 638
Monitoring 645

Chapter 8 Storage Virtualization and Replication

Invista 654
Benefits of virtualization 654
Invista architecture 655
Invista components 657
Invista disaster recovery 657
Related Invista documentation 658
Intelligent switches 658
Logical topology 659
Traffic flow 663
RecoverPoint 664
Benefits of virtualization 664
RecoverPoint architecture 665
RecoverPoint components 666
Related RecoverPoint documentation 667
Intelligent switches 668
Logical topology 668
SAN connectivity 671
Simple SAN topology 671

Complex SAN topologies 675
Interswitch link hops 685

Appendix A Storage Networking Concepts and Protocols

Overview 694
Channels versus networks versus SANs 695
Channels 695
Networks 695
Fibre Channel: Channels with network characteristics 696
Storage Area Networks (SANs) 697
Physical versus logical topologies 702
Physical topology 702
Logical topology 706
Combing physical and logical topologies 711
Fabric design considerations and recommendations 720
Fabric design considerations 721
Common fabric topologies 726
Nonstandard topologies 739
Layout management 739
Switch interoperability 740
Multisite fabrics 740
Tape connectivity 743
General fabric design recommendations 746
Determining customer requirements 759
Scalability 759
Choosing a switch type 763
FC SAN concepts 765
Fibre Channel standards 766
Hosts 778
HBAs 783
ASICs (ASIC or Application Specific Integrated Circuits) 783
8b / 10b encoding and decoding 785
SERDES 791
Optics 791
Fiber 798
Data transfer rates 800
Fibre Channel port types 804
FC-AL (Fibre Channel Arbitrated Loop) 806
Hubs 806
FC-SW (Fibre Channel switched fabric) 806

Flow control 901
In order delivery 904
Inter-Switch Link (ISL) 908
Frame services in Fibre Channel 908
Frame structure in Fibre Channel 913
Class of Service (C.O.S.) 925
Buffer-to-buffer credit (BB_Credit) 928
Zoning 929
Storage 929
NPIV 930
Fibre Channel Routing 937
DWDM 953
CWDM 953
Fast Write 954
Vendor-specific features 957
Port fencing 960
Threshold alerts 960
Management 961
IP SAN concepts 962
Internetworks, TCP, and routing 962
IP addressing 965
IP over Symmetrix directors 968
SRDF over GigE remote director 968
iSCSI using Symmetrix multiprotocol channel director 971

Appendix B EMC Connectrix Products Data

Overview 974
Connectrix B-Series directors and switches 976
Switch and blade ID numbers 976
Connectrix AP-7600B 978
Connectrix AP-7420B 989
Connectrix ED-48000B 998
Connectrix ED-24000B 1009
Connectrix ED-12000B 1015
Connectrix MP-7500B 1019
Connectrix DS-5000B 1033
Connectrix DS-4900B 1042
Connectrix DS-16B3 1051
Connectrix DS-8B3 1054
Connectrix DS-32B2 1058
Connectrix DS-16B2 1060
Connectrix DS-16B and DS-8B 1065

Connectrix MDS-Series directors and switches 1068
Connectrix MDS Storage Service Module (SSM) 1068
Connectrix MDS 9513 1069
Connectrix MDS 9509 1076
Connectrix MDS 9506 1080
Connetrix MDS 9216 1091
Connectrix M-Series directors and switches 1095
Switch model numbers 1095
Connectrix ED-10000M 1096
Connectrix ED-140M 1108
Connectrix M Quad Port Module (QPM) 1111
XPM port card 1119
Connectrix ED-64M 1122
Connectrix ED-1032 1124
Connectrix DS-4700M 1127
Connectrix DS-4400M 1134
Connectrix DS-24M2 1141
Connectrix DS-32M, DS-32M2 1143
Connectrix DS-16M, DS-16M2 1145
Connectrix M Series (Brocade M-Series) Open Trunking feature 1147
Connectrix EC-1700 cabinet system 1153
Guidelines for installing EC-1700 components 1155
Connectrix EC-1500 cabinet system 1158
Connectrix EC-1200 cabinet system 1166

Appendix C Non-EMC Products Data

CNT/Inrange switches and directors 1174
CNT/Inrange FC/9000 1174
Nortel OPTera Metro 1180
Available Optera Metro topologies 1181
Power budget calculations 1183
Diagnostics and maintenance 1184
Ciena ONLINE7000 1185
ONLINE*7000* card types 1188
ONLINE*7000* topologies 1191
ONLINE*7000* protection scheme 1192
ONLINE*7000* power budget calculations 1194
ONLINE*7000* diagnostics and maintenance 1194
Ciena CN 4200/CN4200 MC 1195
Available modules 1198
Common optical filters 1203

Service protection schemes 1203
Supported network topologies 1205
Power budget calculations 1208
Diagnostics and maintenance 1209
Cisco Metro 1500 1210
Metro 1500 power budget calculations 1211
Supported Metro 1500 topologies 1212
Metro 1500 diagnostics and maintenance 1213
Cisco ONS 15540 1214
ONS 15540 system components 1215
Available ONS 15540 topologies 1219
ONS 15540 protection scheme 1221
ONS 15540 power budget calculations 1223
ONS 15540 overall optical link loss budget 1223
ONS 15540 diagnostics and maintenance 1223
Cisco ONS 15454 MSTP 1224
Intended audience 1224
Description 1224
Key features 1225
System architecture 1228
MSTP SFP support 1242
SFP technical details 1253
MSTP management 1259
MSTP further reading 1260
Finisar FLX-2000 link extender 1261
SONET overview 1263
CIENA CN 2000 1266
Lucent OptiStar EdgeSwitch 1271
Blade server switch modules 1273
DELL SW3014 Brocade 4-port 1/2 Gb FC switch module 1274
DELL SW4016 Brocade 4-port 1/2/4 Gb FC switch module 1276
HP A7535A Brocade 4-port 1/2/4Gb FC switch module 1277
IBM PN 26K5601 Brocade 2-port entry-level switch module 1279
IBM PN 90P0165 Brocade 2-port enterprise-level switch module 1280
IBM PN 32R1812 Brocade 6-port SAN switch module 1281
IBM PN 32R1813 Brocade 3-port SAN switch module 1283
DELL SW4314 Brocade M-Series 6-port SAN switch module 1285

DELL SW4416 Brocade M-Series 6-port 1/2/4 Gb FC switch module 1286
IBM PN 32R1790 Brocade M-Series 6-port SAN switch module 1288
IBM PN 32R1833 Brocade M-Series 6-port SAN switch module 1291
IBM PN 32R1905 Brocade M-Series 3-port SAN switch module 1293
Blade servers 1296
Blade server architecture 1296
Blade server chassis 1296
Server blades 1299
Management and KVM modules 1302
I/O modules 1304
Ethernet pass-through modules 1306
Power supplies/cooling 1308
Blade servers supported by EMC 1309
Modes supported on the switch modules 1312
Further reading 1313
Brocade Access Gateway 1314
Hardware and software requirements for Access Gateway 1314
Access Gateway theory of operation 1315
Access Gateway CLI commands 1318
Advantages of Access Gateway 1320
Case studies: How to set up an Access Gateway fabric and its benefits 1321
Access Gateway qualification plan 1347
Qualification results 1348

Appendix D Legacy Information

Fibre Channel arbitrated loop (FC-AL) 1352
Overview 1352
Symmetrix and Fibre Channel connectivity 1359
Arbitrated loop addressing 1361
Primitive signals and sequences 1365
Loop initialization 1369
Arbitration process 1379
Alternate Buffer-to-Buffer Credit (BB_Credit) 1388
QuickLoop 1391
Need for QuickLoop 1391
Defining QuickLoop 1391

Configuration examples 1393
Bridges 1396
SCSI-to-Fibre Channel bridges 1396
Crosspoint 4200 SCSI-to-fabric configuration 1399
ADIC SAN Gateway SCSI-to-fabric configuration 1402
ADIC SAN Gateway loop-to-fabric configuration 1404
Interfacing arbitrated loop to switched fabric 1408
Operating modes 1408
Connectivity devices that support FC-AL 1408
Storage area network management 1417
Distance topology 1417
Capacity topology in the loop environment 1418
Consolidation topology in the arbitrated loop environment 1419
Combined topologies 1420
CNT (Inrange) 1421

Glossary 1423

Index 1445

Figures

	Title	Page
1	Single switch topology example	50
2	Single switch SAN using DS-4100B and EZSwitchSetup with zoning	51
3	Connectrix MDS 9506 using VSANs	55
4	Single switch fabric built with a DS-4700M	63
5	Single switch fabric built with QLogic SB5602	72
6	Two switch fabric example	79
7	Two switch SAN with DS-4100B and DS-4900B using CLI to configure SAN	81
8	Port settings	88
9	Two Connectrix MDS 9506s using VSANs	92
10	Two ED-10000Ms using virtual switches	102
11	Connection Description dialog box	106
12	Two SB5602 switches using Quicktools	118
13	(a) Blade server direct attached to storage (b) two-tier FC SAN	130
14	IBM blade server directly attached to storage	136
15	IBM Brocade 4 GB SAN switch module (32R1812) directly attached to storage	137
16	Four switch full mesh fabric	147
17	Four switch fabric with ED-48000Bs	149
18	Four Connectrix MDS 9506s full mesh configuration	159
19	Four ED-140Ms full mesh configuration	171
20	Four SB9000s in full mesh configuration	188
21	Four switch compound core edge switches	199
22	Four ED-48000Bs in full mesh configuration with edge switches attached	201
23	Four MDS 9513s in full mesh configuration with edge switches attached	218
24	Four ED-10000Ms in full mesh configuration with edge switches attached	234

25 Connection description dialog box 238
26 Phase 1: Basic configuration topology 262
27 Phase 2: Adding Connectrix MDS 9513 264
28 Phase 3: Moving half of host and storage ports 268
29 Phase 4: Complete moving host and storage ports 269
30 Phase 5: Adding Connectrix MDS 9216 270
31 Phase 6: Moving hosts and storage to new edge 271
32 Phase 7: Adding Connectrix MDS switch to the core 272
33 Phase 1: Base configuration 274
34 Phase 2: Adding Connectrix MDS 9513 to the core of the fabric 280
35 Phase 3: Moving half the hosts and storage ports 283
36 Phase 4: Completely moving host and storage ports 284
37 Phase 5: Adding Connectrix MDS 9216 285
38 Phase 6 topology 287
39 Phase 7 topology 288
40 Phase 1: Basic configuration 291
41 Phase 2: Adding Connectrix MDS 9513 293
42 Phase 3: Moving half the host and storage ports 296
43 Phase 4: Completely moving host and storage ports 297
44 Phase 5: Adding Connectrix MDS 9216 298
45 Phase 6: Moving hosts and storage to new edge 300
46 Phase 7: Adding Connectrix M switch to core example 301
47 Phase 1: Basic configuration 303
48 Phase 2: Adding Connectrix MDS 9513 306
49 Phase 3: Moving half the host and storage ports 309
50 Phase 4: Completely moving host and storage ports 310
51 Phase 5: Adding Connectrix MDS 9216 311
52 Phase 6: Moving hosts and storage to new edge 313
53 Phase 7: Adding Connectrix M switch to core 314
54 Phase 1: Basic configuration 316
55 Phase 2: Adding ED-24000B 318
56 Phase 3: Moving half the host and storage ports 320
57 Phase 4: Completely moving host and storage ports 321
58 Phase 5: Adding DS-16B2 and DS-220B 322
59 Phase 6: Moving hosts and storage to the new edge switches 324
60 Phase 7: Adding Connectrix B switch 325
61 Donor and target topology 328
62 Phase 1: Basic configuration of Connectrix M fabric 330
63 Phase 2: Adding Connectrix B fabric and SAN router 332
64 Phase 3: Configuring Connectrix B SAN routing 334
65 Phase 4: Moving half of the host and storage ports 347
66 Phase 5: Complete moving host and storage ports in the core 350
67 Cascaded IOCP 369

68 IPv6 packet header structure 394
69 Slow start and congestion avoidance 402
70 Fast retransmit 403
71 Direct connect example 411
72 Network connect example 411
73 Celerra Network server 413
74 iSCSI gateway service basic implementation 417
75 Supportable configuration example 420
76 Brocade M-Series multiprotocol switch 426
77 DWDM example 433
78 Fibre Channel link extension 435
79 STS-1 organization 438
80 Slow start and congestion avoidance 445
81 Fast retransmit 447
82 BB_Credit mechanism 452
83 Flow control managed by Fibre Channel switch (without buffering from distance extension devices) 463
84 Flow control (with buffering from distance extension devices) 463
85 Normal write command process 464
86 Write command with write acceleration 465
87 RDF JO Single Round Trip 466
88 RF ports benefit only 467
89 All F_Ports will benefit 468
90 Link initialization (More than 100 ms R_T_TOV) 469
91 Brocade Silkworm AP7420 distance extension example 474
92 Brocade SilkWorm 7500 configuration example 478
93 Single tunnel, Fast Write and tape pipelining enabled 481
94 Multiple tunnels to multiple ports 482
95 Cisco MDS 9000 distance extension example 483
96 Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment example 1 485
97 Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment example 2 486
98 Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment example 3 486
99 SANbox 6142 Intelligent Router 490
100 SAN security bridges 497
101 Snooping in a SAN 499
102 Spoofing in a SAN 500
103 Denial-of-service attack 501
104 Security zones 503
105 LUN masking 517
106 CHAP authentication 518

107 Single-ID mode (virtualization disabled) 538
108 Multi-ID mode (virtualization enabled) 539
109 DataFort port mapping (storage side virtualization enabled) 540
110 DataFort LUN mapping (storage side virtualization enabled) 541
111 Recommended configuration of Decru DataFort with EMC storage..... 544
112 Topology without DataFort example 548
113 Virtualization enabled with port mapping 549
114 Recommended CryptoStor 2002 configuration for EMC storage products 552
115 CryptoStor Tape 700 deployment example 554
116 Data flow for direct-attached backups 559
117 LAN backup infrastructures 561
118 LAN backup: Additional steps 563
119 Bandwidth used in LAN backups 564
120 LAN-free backup 566
121 Process flow for SAN backups 568
122 Bandwidth used in LAN-free backups 569
123 Serverless backup 571
124 Backup over long distances 574
125 FCIP tape acceleration 576
126 Brocade tape acceleration — Host remote backup 580
127 Brocade tape a cceleration — NDMP remote backup 581
128 Example of backup software solution 588
129 Core/edge fabric example with recommended component placement 589
130 Oracle 10g RAC installation and configuration flowchart 602
131 Typical three-node fabric topology with IBM AIX hosts 603
132 Search for Connectivity dialog box 622
133 Topology view 623
134 Fabric Manager main view 624
135 Core/edge topology with core switch designation 624
136 Circular and tree designs 625
137 Cisco Fabric Manager main view 626
138 Device Manager default view 627
139 Connectrix Manager basic setup 628
140 Selecting switches to be managed 628
141 Connectrix Manager main view 629
142 Zoning option 631
143 Manage Zone option 632
144 Manage Zone Set option 633
145 Broade Fabric Manager, Zone tab 634
146 Cisco Fabric Manager, Switches tab 635

147 Fabric Manager VSANs 636
148 Zone library 637
149 Active Zone Set tab 638
150 Storage allocation alerts 640
151 At A Glance view 641
152 Fabric Manager Alerts tab 642
153 Fabric Manager, Select switch 643
154 Device Manager 644
155 View All 645
156 Real-time Device Performance 646
157 Fabric Manager Event Monitoring view 647
158 Performance graph example 648
159 Device Manager 649
160 Performance graph 650
161 Performance tab: Statistics values 651
162 Control path functions 656
163 Logical topology 659
164 List of WWNs 661
165 WWNs of 9 VIs from a single SSM blade 662
166 Intelligent switch running RecoverPoint 666
167 Logical topology 668
168 Example Back End and Front End VSAN 669
169 Zoning with VTs and VIs 670
170 Connectrix B AP-7420B simple SAN topology 672
171 Connectrix MDS SSM blade in simple SAN topology 673
172 MDS 9513 with a 24-port line card and an SSM blade 674
173 Core/edge design 677
174 Core/edge design with a Connectrix B 678
175 Storage on a single switch 680
176 Core/edge/core design with Brocade 681
177 Storage on multiple switches 683
178 Storage on multiple swithces with Connectrix B 684
179 Four-switch mesh design using MDS 9506s (Example 1) 686
180 Four-switch mesh design using MDS 9506s (Example 2) 687
181 Four-switch mesh of ED-48000Bs with an AP-7420B as the intelligent switch 689
182 Four-switch mesh of MDS 9506s 691
183 SAN example 697
184 FCIP-MTU 1500 699
185 Physical and logical topologies 702
186 Tiered fabrics 703
187 Single-core fabric 705
188 Dual-core fabric 705

189 Multiple tiers in the same fabric 707
190 FC-Switched fabric distance topology example 709
191 FC-Switched fabric capacity topology example 710
192 FC-Switched fabric consolidation topology example 711
193 FC-Switched fabric combined topologies example 711
194 Three-tier physical and logical fabric 712
195 Balanced fabric 714
196 Mirrored fabrics 715
197 Logical fabric segregation 718
198 Single switch topology example 727
199 Two switch fabric example 728
200 Examples of a full mesh 730
201 Compound core/edge fabric example 732
202 Partial-mesh example 734
203 Partial-mesh migration to core/edge fabric 736
204 Connectivity tier fabric example 737
205 DWDM across mirrored fabric 742
206 RDF extended fabric 743
207 Tape drive-sharing model 744
208 Traffic-reduction model 745
209 Zone set activation 748
210 Domain add 749
211 Fibre Channel levels 766
212 Fibre Channel architectural layers 768
213 Examples of SOF and EOF delimiters 770
214 Example of primitive signals 770
215 Link initialization handshake 772
216 Exchange sequence and frame relationship 773
217 The Fibre Channel exchange 774
218 Sequences in an Exchange 776
219 Before hardware failure, outage 780
220 After any one hardware failure, outage 781
221 Example of fan-in 782
222 Example of fan-out 782
223 ASIC 784
224 8b/10b Encoder/decorder 785
225 Example 786
226 Transmitting 10 bit encoded bytes 787
227 Transmission words 788
228 Data character “F4” 788
229 Sending encoded data 789
230 SERDES example 791
231 Optics 792

232 Measured values 795
233 Fibre construction 799
234 Single and multimode optical fibre 800
235 Standard Fibre Channel port types 804
236 Switched Fabric example 807
237 Generic switch construct 813
238 Generic switch construct with services 814
239 Switched fabric example 816
240 Fabric Configuration process 817
241 Switch port initialization 818
242 ELP Exchange Link Parameters process 822
243 Normal FC Frame and a VFT_Header Frame 824
244 EFP request payload 826
245 Build fabric 827
246 Principle switch selection process 1 829
247 Principle switch selection process 2 831
248 Domain Identifier Assigned (DIA) 833
249 Request Domain Identifier (RDI) 834
250 Exchange of zoning information 836
251 Path selection topology 838
252 Hello frame format and payload 839
253 Link State Update frame format and payload 841
254 Sample LSR database and fabric topology 844
255 Switch B determining the shortest path 846
256 Switch B is settled 847
257 LSR database 848
258 Switch A is settled 849
259 LSR database 850
260 Shortest path 851
261 LSR database 852
262 Adding hosts and storage ports and setting up routes 855
263 Routing table example 856
264 Routing tab 859
265 Route properties 863
266 Add Preferred Path 865
267 Topology window 866
268 Route properties dialog box 867
269 Newly configured Preferred Path windows 868
270 Configuration example 870
271 Port assignment example 870
272 Configuration after initiator 2 is rebooted 871
273 Before frame-based trunking 873
274 After frame-based trunking 874

275 Backpressure example using mass transit system 878
276 Increase in consumption of buffer 879
277 Buffer filled causing overflow 879
278 Topology example 880
279 Login and credit initialization process example 881
280 Login and credit initialization 887
281 Switch receives the FLOGI 888
282 Switch processes the FLOGI 889
283 Queues 890
284 Queue example 891
285 Transmit example 891
286 Released credit 892
287 Slow drain example 893
288 Uncongested environment 894
289 Impact of a slow drain port 895
290 Virtual Output Queue for port 7 continues to grow 896
291 Virtual Output Queue for port 7 consumes entire Queue on Switch A, port 1 896
292 ISL failure 899
293 Nodes 901
294 End-to-End Credit 902
295 Summary of flow control 902
296 ED_TOV 904
297 Host and storage separated by two hops 905
298 New ISL added. Shorter path introduced 906
299 New frames queued for transmit down shorter path 906
300 Sequence ID "D" received before Sequcne ID "5" 907
301 IOD is set. Sequence ID "D" is held and will be received in proper order 908
302 Extended link services 910
303 FLOGI and Accept frame payload 911
304 PLOGI and accept payload 912
305 Process Login and Accept frame payload 913
306 Frame format 913
307 Frame types 914
308 Frame header 915
309 Destination ID 916
310 Source ID 917
311 Class specific control 917
312 Frame type 917
313 Frame Control field 918
314 Frame Control (F_CTL) field 918
315 Sequence ID 920

316 Data Field Control 920
317 Sequence Count 921
318 Originiator Exchange ID 921
319 Responder Exchange ID 921
320 Offset / parameter field 922
321 Payload 922
322 CRC protection 923
323 Frame definition 924
324 Class 1 Dedicated connection 925
325 Class 2 operation 926
326 Class 3 Flow control 927
327 Class of service summary 928
328 Traditional NPIV initionalization 931
329 NPIV-capable N_Port initialization 932
330 Blade Servers 932
331 NPIV Gateways 933
332 Normal operation with FC-AL 934
333 Engine A failure 935
334 Failover to Engine B 935
335 Using NPIV 936
336 Fibre Channel Routing 938
337 Interoperarability in Fibre Channel Routing 939
338 Resource consolidation in Fibre Channel Routing 941
339 Distance extension in Fibre Channel Routing 942
340 Scalability in Fibre Channel Routing 943
341 MetaSAN with edge-to-edge and backbone fabrics 945
342 MetaSAN with interfabric links (IFLs) 946
343 Edge fabrics connected through a backbone fabric 949
344 MetaSAN with imported devices 951
345 Typical SCSI write 954
346 SCSI write over distance without Fast Write 955
347 Fast Write over distance with appliance 956
348 Internetwork example 962
349 TCP and UDP 964
350 iSCSI implementation example 967
351 TCP connection example: One TCP connection 969
352 TCP connection example: Four TCP connections 969
353 Connectivity: Symmetrix DMX series to Symmetrix 8000 series 970
354 Media conversion 971
355 Direct connection of host NIC (from multiple media types) to iSCSI MPCD 972
356 Switched layer 2 (single subnet) to iSCSI MPCD 972
357 Port side view 979

358 Port numbering in the switch 980
359 Nonport side of the switch 983
360 Connectrix B AP-7420 990
361 AP-7420B (Port Side) 991
362 AP-7420B Multilprotocol Router (Nonport Side) 993
363 Port side of the ED-48000B Director 1000
364 Nonport side of the ED-48000B director 1001
365 Fully populated ED-48000B with ports numbered 1004
366 ED-24000B, front view 1010
367 ED-12000B, front view 1015
368 ED-12000B, rear view 1016
369 Port side view of the MP-7500B 1021
370 Port numbering in the MP-7500B 1022
371 Nonport side of a MP-7500B 1023
372 Port side view of the DS-5000B 1034
373 Port numbering in the DS-5000B switch 1035
374 DS-4900B Port side view 1044
375 DS-4900B Non-port side view 1045
376 Connectrix DS-16B3 1051
377 Connectrix DS-8B3 1054
378 Connectrix DS-32B2 1058
379 Connectrix DS-16B2 1061
380 ISLs with no trunking 1063
381 ISLs with trunking 1064
382 Connectrix DS-16B 1065
383 Connectrix MDS storage service module (SSM) 1068
384 MDS 9513 chassis front panel view1 1072
385 MDS 9513 chassis rear panel view 1073
386 MDS 9509 chassis front panel view1 1078
387 MDS 9509 chassis rear panel view 1079
388 MDS 9506 chassis front panel view1 1083
389 MDS 9506 chassis rear panel view 1084
390 Connectrix MDS 9216 1091
391 MDS 9216 supervisor module 1091
392 MDS 9216 interface module 1092
393 MDS 9216 16-port line card 1092
394 MDS 9216 32-port line card 1092
395 Connectrix ED-10000M 1097
396 ED-10000M quadrants 1101
397 FRUs on front of ED-10000M 1105
398 FRUs on rear of ED-10000M 1106
399 Connectrix ED-140M, front view 1108
400 Connectrix ED-140M, rear view 1109

401 QPM 4 G faceplate 1113
402 QPM 4 G 1113
403 QPM configuration 1 1116
404 QPM configuration 2 1117
405 QPM configuration 3 and 4 1118
406 XPM card 1119
407 Connectrix ED-64M 1122
408 Connectrix ED-1032 1125
409 DS-4700M (front view) 1128
410 DS-4700M (rear view) 1128
411 DS-4400M (front view) 1135
412 DS-4400M (rear view) 1135
413 Connectrix DS-24M2 1141
414 Connectrix DS-32M, D32M2 1143
415 Connectrix DS-16M, DS-16M2 1145
416 Open Trunking on single fabric 1148
417 Open Trunking on balanced fabrics 1149
418 Open Trunking on full mesh 1149
419 Open Trunking on three-hop core / edge; one hop logical (Example 1) 1150
420 Open Trunking on three-hop core / edge; one hop logical (Example 2) 1151
421 Open Trunking on three-hop core / edge; three-hop logical 1152
422 EC-1700 cabinet front view 1153
423 EC-1700-B cabinet configuration examples 1156
424 EC-1700-C cabinet configuration examples 1 – 2 1157
425 EC-1500 cabinet configuration: M series examples 1 – 3 1160
426 EC-1500 cabinet configuration: M series examples 4 – 6 1161
427 EC-1500 cabinet configuration: M series examples 7 – 9 1161
428 EC-1500 cabinet configuration: M series examples 10 – 12 1162
429 EC-1500 cabinet configuration: B series examples 1163
430 EC-1500 cabinet configuration: MDS series examples 1 – 3 1164
431 EC-1500 cabinet configuration: MDS series examples 4 – 6 1164
432 EC-1500 cabinet configuration: MDS series examples 7 – 8 1165
433 EC-1200 cabinet configuration examples 1 – 4 1167
434 EC-1200 cabinet configuration examples 5 – 8 1168
435 EC-1200 cabinet configuration examples 9 – 12 1168
436 EC-1200 cabinet configuration examples 13 – 16 1169
437 EC-1200 cabinet configuration examples 17 – 20 1170
438 EC-1200 cabinet configuration examples 21 – 24 1170
439 CNT / Inrange FC / 9000 1175
440 Nortel OPTera shelf diagram 1181
441 Point-to-point protected topology 1182

442 Hubbed ring topology 1182
443 OPTera protection scheme 1183
444 ONLINE*7000* UPSR protection diagram for main shelf 1186
445 I/O flowchart: Point-to-point/ring configuration over extended distance..... 1187
446 Point-to-point topology 1191
447 Line drop mode 1191
448 Ring network 1191
449 O-UPSR ring configuration in normal mode 1192
450 O-UPSR ring configuration in failure mode 1193
451 Ciena CN 4200 FlexSelect Advanced Services Platform 1195
452 CN 4200 block diagram 1196
453 CN 4200 MC block diagram 1197
454 M6S module 1198
455 F10-T module 1199
456 F10-A module 1199
457 FC4-T module 1200
458 OPS-1 module 1200
459 OPS-2 module 1201
460 OPS-2 850 module 1201
461 Fixed-gain Optical Amplifier module 1202
462 Optical Supervisory Channel (OSC) module 1203
463 Redundant path line card protection 1204
464 CN 4200/CN 4200 MC optical protection switching 1205
465 Unprotected point-to-point configuration 1206
466 Protected point-to-point configuration 1206
467 Linear Add/Drop configuration 1206
468 Hubbed ring configuration 1207
469 Meshed ring configuration 1208
470 Point-to-point (two-site) configuration 1212
471 Multi hop (three-site) configuration 1213
472 Cisco 15540 1214
473 Client/network signal transmission in Cisco ONS 15540 1215
474 Unprotected point-to-point topology 1219
475 Protected point-to-point topology 1219
476 Bus topology 1220
477 Hubbed-ring topology 1220
478 Meshed ring topology 1221
479 Cisco 15540 splitter protection diagram 1222
480 Cisco ONS 15454 MSTP ANSI and ETSI mechanics 1225
481 2.5 Gb/s Data muxponder 1232
482 2.5 Gb/s data muxponder (protected and unprotected) 1233
483 8-Port enhanced data muxponder 1235

484 8-Port data muxponder (unprotected) 1236
485 8-Port buffer credit spoofing flow chart 1238
486 10 Gb/s Multi-rate enhanced transponder 1239
487 10 Gb/s Multi-rate enhanced transponder block daigram 1240
488 STS-1 organization 1263
489 CIENA CN 2000, OUSP 2048 model (rear view) 1267
490 CIENA CN 2000, OUSP 2048E model (rear view) 1267
491 Typical CIENA CN 2000 distance extension setup 1270
492 OptiStar switches over SONET network 1271
493 Basic blade server architecture 1297
494 Front plane of a Dell blade server with 10 server blades 1298
495 Back plane of a Dell blade server with 10 server blades 1299
496 Server blade example 1300
497 Management module example 1302
498 KVM module example 1303
499 Pass-through example 1305
500 FC blade server – Switch module connectivity mechanism 1306
501 Ethernet switch module example 1307
502 Power module example 1308
503 Fan module example 1309
504 Blade servers using Fibre Channel switch modules 1315
505 Blade servers using Acess Gateway 1316
506 Topology 1322
507 Brocade-based blade server chassis attached to a 4-switch full mesh Connectrix B fabric 1331
508 Connectrix MDS fabric 1332
509 Completed migration 1333
510 Topology 1340
511 Arbitrated loop example 1353
512 Arbitrated loop 1354
513 Arbitrated loop with 4 nodes 1355
514 Arbitrated loop with a hub 1356
515 Hub port bypass 1357
516 Arbitrated loop (FC-AL) 1358
517 Loop after arbitration won and ports opened 1358
518 First implementation on Symmetrix 1359
519 Expanded connectivity 1360
520 Dual loop disk drives 1360
521 Highly available dual port disk solution 1361
522 Loop addressing 1362
523 AL_PA priority 1363
524 AL_PA to loop ID chart 1364
525 Primitive signals 1365

526 Primitive sequences 1366
527 Loop Port State Machine (LPSM) 1368
528 Initialization procedure 1370
529 LISM frame format 1371
530 LIFA / LIPA / LIHA / LISA frames 1373
531 LIRP / LILP frame format 1376
532 FLOGI and Accept 1377
533 PLOGI and Accept 1378
534 PRLI and Accept 1378
535 Arbitration Step 1 1380
536 Arbitration Step 2 1380
537 Arbitration Step 3 1381
538 Arbitration Step 4 1381
539 Arbitration Step 5 1382
540 Arbitration Step 6 1383
541 Access fairness window 1384
542 Opening a loop circuit 1385
543 Open primitive signals 1385
544 Close loop: Step 1 1387
545 Close loop: Step 2 1387
546 Close loop: Step 3 1388
547 Alternate credit 1389
548 Single QuickLoop configuration 1393
549 Dual QuickLoop configuration 1394
550 Fabric / QuickLoop mixed-mode configuration 1394
551 Simple storage mode configuration 1397
552 Examples of tape pools 1411
553 FC-AL high-availability distance topology example 1417
554 FC-AL capacity expansion topology example 1418
555 FC-AL high-availability capacity expansion topology example 1418
556 FC-AL consolidation topology example 1419
557 FC-AL high-availability consolidation topology example 1419
558 FC-AL combined hub topologies example 1420

Tables

	Title	Page
1	172.23.199.22 through 172.23.199.27	223
2	172.23.200.22 through 172.23.200.27	223
3	Domain ID correlation	276
4	EMC-tested modes	362
5	FICON compared to ESCON	365
6	EMC CUP models	372
7	FICON port numbering in the MDS 9000 family	382
8	iSCSI solution features comparison table	429
9	SONET/Synchronous Digital Hierarchy (SDH)	438
10	STS-1s and optical carrier rates	450
11	Switch ASIC and model numbers	453
12	Bloom and Bloom II ASICs	455
13	Condor ASIC	456
14	Goldeneye ASIC	456
15	Brocade M-Series credit chart	460
16	Cisco MDS credit chart	461
17	Symmetrix fibre adapter credit chart	461
18	Network impact on effective throughput example	472
19	Fast Write and tape pipelining comparison	481
20	Distance extension comparison table for TCP/IP products	493
21	Distance extension comparison table for non TCP/IP products	494
22	Secure SANs component mechanisms	505
23	Brocade security features	519
24	Implement Role Based Access Control (RBAC	522
25	Brocade M security features	525
26	Cisco security features	530
27	Fibre Channel topology solutions	708
28	Single switch subjective rating	727
29	Two switch subjective rating	728
30	Full-mesh subjective rating	729

31 Compound core/edge subject rating 731
32 Partial-mesh subjective rating 734
33 Subjective rating for connectivity tier 736
34 Full mesh with two ISLs 759
35 Full mesh with four ISLs 760
36 Using 64 port core switches 761
37 Using 140 port core switches 761
38 Using 256 port core switches 762
39 Number of ISLs increased between cores and edge switches 763
40 Bit rates 801
41 200 MB/s link with max 60 BB_Credits 802
42 400 MB/s link with maximum of 200 BB_Credit 803
43 Standard Fibre Channel port types 804
44 Hold timers 897
45 Oversubscription 975
46 Switch ID numbers 977
47 Blade ID numbers 977
48 Serial cable pinouts 982
49 Environmental requirements 987
50 Power supply specifications 987
51 General specifications 988
52 Serial cable pinouts 1024
53 Environmental requirements 1029
54 Power supply specifications 1030
55 General specifications 1031
56 Environmental requirements 1039
57 Power supply specifications 1040
58 General specifications 1041
59 Chassis capabilities 1085
60 EMC and Brocade models and ASICs 1095
61 ED-10000M Oversubscription 1099
62 QPM model numbers: New sales 1118
63 QPM model numbers: Upgrades 1119
64 EC-1700-B cabinet models 1154
65 EC-1700-C cabinet models 1154
66 Protocol support per service card 1229
67 Supported modules on ONS 15454 MSTP 1230
68 Data muxponder receiver trunk side specifications 1234
69 8 Gb/s data muxponder receiver trunk side specifications 1237
70 Buffer credits supported per port on 8-port enhanced data muxponder 1238
71 10 Gb/s multi-rate enhanced transponder receiver trunk side specifications 1241

72 ONS15454 MSTP SFP matrix 1242
73 SONET/SDH SFPs 1244
74 Data SFPs 1245
75 DWDM SFPs 1247
76 CWDM SFPs 1249
77 Grey XFPs 1250
78 DWDM XFPs 1251
79 SONET/SDH SFPs optical specification 1254
80 Ethernet pluggables optical specification 1255
81 ESCON SFPs optical specification 1255
82 FC/FICON pluggables optical specification 1256
83 CWDM SFPs optical specification 1257
84 DWDM SFPs optical specification 1257
85 DWDM SFPs optical performances 1257
86 XFPs optical specification 1258
87 DWDM XFPs optical specification 1258
88 DWDM XFPs optical performances 1258
89 STS-1s and optical carrier rates 1264
90 Maximum number of Fibre Channel circuits on the OUSP 1268
91 Switch features 1273
92 Indicator codes for blade servers 1301
93 Indicator codes for management modules 1303
94 Indicator codes for I/O modules 1305
95 Processor information 1310
96 Mezzanine cards 1310
97 FC switch modules 1311
98 Supported Interoperable configurations 1312
99 New CLI commands 1319
100 Fibre Channel Access Gateway support limitations 1349
101 Supported SCSI-to-Fibre Channel fabric bridges 1397
102 Interfacing arbitrated loop to switched fabric: Summary 1415

Preface

The EMC Networked Topology Guide provides a top-down view of networked storage and assists the network designer in designing a suitable networked storage instrustructure. This guide provides best practices for Storage Networking Technologies that are created through E-Lab interoperability and integration testing and authored by E-Lab engineers and product content experts.

E-Lab would like to thank all the contributors to this document, including EMC engineers, EMC field personnel, and partners. Your contributions are invaluable.

As part of an effort to improve and enhance the performance and capabilities of its product lines, EMC periodically releases revisions of its hardware and software. Therefore, some functions described in this document may not be supported by all versions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes. If a product does not function properly or does not function as described in this document, please contact your EMC representative.

EMC Support Matrix and E-Lab Interoperability Navigator

For the most up-to-date information, always consult the *EMC Support Matrix* (ESM), available through E-Lab Interoperability Navigator (ELN) at: http://elabnavigator.EMC.com, under the **PDFs and Guides** tab.

The *EMC Support Matrix* links within this topology guide will take you to Powerlink where you are asked to log in to the E-Lab Interoperability Navigator. Instructions on how to best use the ELN (tutorial, queries, wizards) are provided below this **Log in** window. If you are unfamiliar with finding information on this site, please read these instructions before proceeding any further.

Under the **PDFs and Guides** tab resides a collection of printable resources for reference or download. All of the matrices, including the ESM (which does not include most software), are subsets of the E-Lab Interoperability Navigator database. Included under this tab are:

- The *EMC Support Matrix*, a complete guide to interoperable, and supportable, configurations .
- Subset matrices for specific storage families, server families, operating systems or software product.
- Host connectivity guides for complete, authoritative information on how to configure hosts effectively for various storage environments.

Under the **PDFs and Guides** tab, consult the *Internet Protocol* pdf under the "Miscellaneous" heading for EMC's policies and requirements for the *EMC Support Matrix*.

Audience

This guide is intended for EMC field personnel, including technology consultants, and for the storage architect, administrator, and operator involved in acquiring, managing, operating, or designing a networked storage environment that contains EMC and host devices.

Related documentation

Related documents include:

- *EMC Support Matrix*, available through E-Lab Interoperability Navigator at http://elabnavigator.EMC.com >**PDFs and Guides**
- RSA security solutions documentation can be found at http://RSA.com > **Content Library**

All of the following documentation and release notes can be found on http://Powerlink.EMC.com. From the toolbar, select **Support > Technical Documentation and Advisories**, then choose the appropriate Hardware/Platforms, Software, or Host Connectivity/HBAs documentation links.

Hardware documents and release notes include:

- Connectrix B series
- Connectrix M series
- Connectrix MDS (release notes only)
- CLARiiON
- Celerra
- EMC Symmetrix

Software documents include:

- EMC ControlCenter
- RecoverPoint
- Invista
- TimeFinder
- PowerPath

The following E-Lab documentation is also available:

- Host Connectivity Guides
- HBA Guides

For vendor-specific documentation, refer to the vendor's website.

Conventions used in this document

EMC uses the following conventions for special notices.

Note: A note presents information that is important, but not hazard-related.

IMPORTANT

An important notice contains information essential to operation of the software. The important notice applies only to software.

CAUTION

A caution contains information essential to avoid data loss or damage to the system or equipment. The caution may apply to hardware or software.

Typographical conventions

EMC uses the following type style conventions in this document:

Normal	Used in running (nonprocedural) text for: • Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus) • Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, filenames, functions, utilities • URLs, pathnames, filenames, directory names, computer names, links, groups, service keys, file systems, notifications
Bold:	Used in running (nonprocedural) text for: • Names of commands, daemons, options, programs, processes, services, applications, utilities, kernels, notifications, system call, man pages

	Used in procedures for: • Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus) • What user specifically selects, clicks, presses, or types
Italic:	Used in all text (including procedures) for: • Full titles of publications referenced in text • Emphasis (for example a new term) • Variables
`Courier:`	Used for: • System output, such as an error message or script • URLs, complete paths, filenames, prompts, and syntax when shown outside of running text.
`Courier bold:`	Used for: • Specific user input (such as commands)
`Courier italic:`	Used in procedures for: • Variables on command line • User input variables
< >	Angle brackets enclose parameter or variable values supplied by the user
[ ]	Square brackets enclose optional values
|	Vertical bar indicates alternate selections - the bar means “or”
{ }	Braces indicate content that you must specify (that is, x or y or z)
...	Ellipses indicate nonessential information omitted from the example

Where to get help

EMC support, product, and licensing information can be obtained as follows.

Product information — For documentation, release notes, software updates, or for information about EMC products, licensing, and service, go to the EMC Powerlink website (registration required) at:

http://Powerlink.EMC.com

Technical support — For technical support, go to EMC Customer Service on Powerlink. To open a service request through Powerlink, you must have a valid support agreement. Please contact your EMC sales representative for details about obtaining a valid support agreement or to answer any questions about your account.

Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Please send your opinion of this document to: techpubcomments@EMC.com.

1

Fibre Channel SAN Topologies

This chapter provides the following information on Fibre Channel SAN topologies.

- Fibre Channel topology overview 36
- Simple Fibre Channel SAN topologies 50
- Complex Fibre Channel SAN topologies..... 146
- FICON connectivity 365

Fibre Channel topology overview

This chapter provides information on Fibre Channel SAN topologies.

For valuable information that may be helpful prior to building a SAN, refer to Appendix A, "Storage Networking Concepts and Protocols."

For purposes of this document, E-Lab™ uses the following definitions for simple and complex SANs:

- A *simple* Fibre Channel SAN consists of less than four Directors and switches connected by ISLs and has no more than two hops.
- A *complex* Fibre Channel SAN consists of four or more Directors and switches connected by ISLs and has any number of hops.

Instructions for using this chapter

This section was written so that it could be read from front to back or used as a quick reference guide. The concept of *inheritance* was borrowed from the software development community to avoid duplicate best practices, host and storage layout, switch and fabric management, and security information.

This section has been broken into four levels:

- The top level is the "Fibre Channel topology overview" and it contains overall best practices, host and storage layout, switch and fabric management, and security information for all Fibre Channel SANs.
- The next level consists of two sections: "Simple Fibre Channel SAN topologies" on page 50 and "Complex Fibre Channel SAN topologies" on page 146. Each of these sections has best practices, host and storage layout, switch and fabric management, and security information specific to each configuration and it also inherits all of the information listed in "“Fibre Channel topology overview”.
- Next in the hierarchy is the actual topology (that is, "Single switch fabrics" on page 50, "Two switch fabrics" on page 79, "Blade switch with direct attached storage" on page 129, "Four switch full mesh" on page 147, "Compound core edge switches" on page 199, and "Heterogeneous switch interoperability" on page 256). Each of these sections has best practices, host and storage layout, switch and fabric management, and security information specific to each topology, and also inherits all of the best practices, host and storage layout, switch and fabric

management, and security information from "Simple Fibre Channel SAN topologies" on page 50 or "Complex Fibre Channel SAN topologies" on page 146.

- Finally, the actual case studies contain best practices, host and storage layout, switch and fabric management, and security information specific to the case study and each case study inherits all of the best practices above it. For easy reference, each section contains a link to the best practices section from which it inherits.

The inheritance approach is meant to enhance the readability and supportability of the document.

General information is provided in this section for the following SAN topologies:

- "Simple Fibre Channel SAN topologies" on page 50
- "Complex Fibre Channel SAN topologies" on page 146

Examples are provided using Connectrix® B, Connectrix MDS, Connectrix M, and QLogic switches. Information specific to these examples is detailed as needed.

All of the configurations shown throughout this chapter use the same host and storage ports. This was done intentionally to better expose the differences between switch implementations and fabric topologies rather than attempt to address all possible host and storage combinations.

For a complete list of supported host/switch/storage combinations, and the most up-to-date supported configurations, refer to the *EMC Support Matrix*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com.

Best practices

Consider the following best practices:

- Plan for failures.
  - Connect the host and storage ports in such a way as to prevent a single point of failure from affecting redundant paths. For example, if you have a dual-attached host and each HBA accesses its storage through a different storage port, do *not* place both storage ports for the same server on the same Line Card or ASIC.
  - Use two power sources.

- For host and storage layout:
  - To reduce the possibility of congestion, and maximize ease of management, connect hosts and storage port pairs to the same switch where possible.

Note: Refer to "Host and storage layout" on page 43 for information on host and storage layout.

- For security:

Note: Refer to "Security" on page 47 for information on security.

- Use single initiator zoning.

  Special consideration for SRDF ports:
  - If desired, the administrator can use *SRDF Single Zoning* when multiple sources and targets are contained in a single zone.
  - An SRDF zone should only contain RF ports.
  - If multiple zones are used, design zoning must be designed so that functionality such as Dynamic SRDF meets customer requirements. For example, if Dynamic SRDF is in use, zoning requirements can change. With Dynamic SRDF, any RFs that have Dynamic SRDF connectivity established to one another must be in the same zone.
  - A maximum of 10 RFs per switch is recommended, for example when you have a two site configuration with four DMXs in each site. Each DMX contributes four SRDF ports. No single switch should have more than ten RFs connected to it. In this example, a minimum of two switches need to be deployed at each site. The main reason for restricting a switch to ten RFs is due to NameServer traffic. NameServer traffic is an important consideration and needs to be kept to a minimum in order to minimize link recovery times when RSCNs occur. By distributing across multiple switches, processing of NameServer traffic is also able to scale.
- Use dual management networks whenever two or more FC switches are being used.
- Before building the SAN, gather the following customer-supplied information that will be needed during the implementation:
  - IP addresses, subnet mask, and gateway for each switch

- Information for each switch
  - Switch names
  - Port names
  - Passwords
  - Number of HBAs and storage arrays to be connected to the switch

These values are used in both manual and GUI-based setup methods.

- Use a port fencing policy.

  For more information on Port fencing, refer to "Port fencing" on page 960.

- Use the available performance monitoring tools.

  For more information on threshold alerts, refer to "Threshold alerts" on page 960.

- Use the latest supported firmware version and ensure that the same version of firmware is used throughout the fabric. In homogeneous switch vendor environments, all switch firmware versions inside each fabric should be equivalent, except during the firmware upgrade process.

  Note: Refer to the *EMC Support Matrix* for the most up-to-date information.

- Periodically (or following any changes) back up switch configurations.
- Utilize a syslog server.

  Note: It is also recommended to install an NTP client on the syslog server.

- Use persistent Domain IDs.
- Estimate light budget.

The following are switch-specific best practices:

Connectrix B

Firmware download

Preparing for a firmware download

Before executing a firmware download, performing the following tasks is recommended. This information will not only help validate that the firmwaredownload did not disrupt the existing configuration, but also provide the support team with sufficient information in an unlikely event that something goes wrong after a fimrwaredownload.

1. Establish a Telnet session and log in to the Connectrix B switch.
2. It is advisable to have session logging enabled so that the output of your commands is saved to a file.
3. Enter the **switchShow** command to display switch information.
4. Enter the **portCfgShow** command to display current switch port settings.
5. Enter the **nsShow** command to display the total number of devices in the switch local Name Server database.
6. Enter the **nsAllShow** command to display the total number of devices in the switch global Name Server database.
7. Enter the **fabricShow** command to display the total number of domains in the fabric. If you are changing a dual domain Connectrix B director to a single domain Connectrix B director, this value will be one domain less after the operation.
8. Display the MIB configuration information using the **snmpMibCabShow** or **agtCfgShow** command.
9. Upload the switch configuration settings to an FTP server using the **configUpload** command.
10. Enter the **supportShow** and **supportSave** commands (these commands are version dependent) to provide baseline information for advanced support.

Note: If you are upgrading a Connectrix B director that is configured with two logical domains, perform these steps for *both* logical switches.

Before you enter the **firmwareDownload** command, read the release notes for the new firmware to find out if there are any firmware download issues.

After the firmwareDownload on a Connectrix B director switch, it is recommended you validate that the firmware versions have been synchronized on both CPs.

Zoning

Note the following best practices:

- A zoneset can be managed and activated from any switch in the fabric, but it is recommended that it be managed from a single entry switch within a fabric to avoid any complications with multiple users accessing different switches in a fabric to make concurrent zone changes.
- The system administrators should coordinate zoning configuration activity to avoid running into a situation where two administrators are making changes simultaneously.
- To avoid any lengthy outages due to errors in Connectrix B SAN configurations it is recommended to backup the existing configuration before making any changes.
- In order to avoid the high risk involved in adding a new unauthorized switch to a Connectrix B fabric, it is advisable to limit the creation of switch-to-switch ports. This can be done by locking the already connected switch-to-switch ports in the SAN using the **portCfgEport** command. Such locking down of E_Ports is persistent across reboots. A **portCfgEport <port number>,0 <disable>** must be run on ports that are not connected to other switches in the fabric to block them from forming ISLs between switches.

ISL trunking

More than a best practice, the administrator configuring a Connectrix B SAN must be aware that the frame-level trunking for Connectrix B switches requires all ports in a given ISL trunk to reside within an ASIC group on each end of the link.

- On 2 Gb/s switches, port groups are built on contiguous 4-port groups called *quads*. For example, on a Connectrix DS-8B2 , there are two quads: ports 0-3 and ports 4-7.

- On 4 Gb/s switches like the Connectrix DS-4100B, trunking port groups are built on contiguous 8-port groups called *octets*. In this product, there are four octets: ports 0-7, 8-15, 16-23, and 24-31.

The administrator must use the ports within a group specified above to form an ISL trunk. It is also possible to configure multiple trunks within a port group.

Connectrix MDS

The following are requirements and guidelines for using IVR NAT:

IVR NAT port login (PLOGI) requests received from hosts are delayed for a few seconds to perform the rewrite on the FC ID address. If the host's PLOGI timeout value is set to a value less than five seconds, it may result in the PLOGI being unnecessarily aborted and the host being unable to access the target. EMC recommends that you configure the host bus adapter for a timeout of at least ten seconds (most HBAs default to a value of 10 or 20 seconds).

Note: IVR NAT requires Cisco MDS SAN-OS Release 2.1(1a) or later on all switches in the fabric performing IVR. If you have isolated switches with an earlier release that are involved in IVR, you must remove any isolated fabrics from being monitored by Fabric Manager server and then re-open the fabric to use IVR NAT.

Connectrix M

- When using Connectrix Manager, persist the fabric after setup is complete.
- Where possible, configure the SAN to use Open Fabric Mode 1.0.

QLogic

- A zoneset can be managed and activated from any switch in the fabric, but it is recommended that it be managed from a single entry switch within a fabric.
- When configuring a switch using the Enterprise Fabric Suite, users are presented with several GUI wizard screens. The defaults are usually the switch best practice settings. User must verify the following default zoning-based settings:
  - **Autosave**
    - **True**: All zoneset merges from the fabric are automatically saved to NVRAM and maintained as a *configured* copy.
    - **False**: A cached RAM copy of the merged database will be available for viewing and editing.

  - **Default setting**: True.
- **DiscardInactive Zone**
  - **True**: If this parameter is set to True, all inactive zoning objects on a switch will be discarded. This may prevent the zone database from growing too large and consuming all the allowed space reserved on a switch for the zoning database. It may also prevent a switch from isolating unnecessarily.
  - **False**: If this parameter is set to False, all inactive zoning objects will remain in the configured zoning information on a switch.
  - **Default setting**: False.
- **DefaultZone**
  - **Allow**: All devices connected to the switched fabric can see each other if they are not specified as part of a zone.
  - **Deny**: Devices connected to the switch cannot see each other unless a zone is activated.
  - **Default setting**:

    Allow for all edge switches.
    Deny for the SB9000.

- The system administrators should coordinate zoning configuration activity to avoid running into a situation where two administrators are making changes simultaneously.
- If using FCID or Domain port zone member types, EMC recommends that the Domain ID of each switch in the fabric be locked.
- When a new switch is installed in a fabric, EMC recommends not to have a configured zoning database or an Active zoneset. Run the **reset zoning** command to clear zone.

Host and storage layout

The correct way to attach hosts and storage to a SAN is completely dependent upon customer environments. Historically, the best practice placed hosts on edge switches and high-use storage ports on core switches. This was recommended because high-use storage ports are sometimes accessed by many different hosts on different parts of the fabric. If this is the case in your environment, this configuration would still be the best option. However, if you have high-use storage ports that are only accessed by a couple of hosts and it is possible to locate them all on the same switch, this is the preferred configuration instead of forcing the use of ISLs. ISLs are a valuable and limited

resource and should be reserved for providing connectivity between ports that are unable to be placed on the same switch.

With this in mind, the following information provides helpful general guidelines.

- Whenever practical, locate HBAs and the storage ports they will access on the same switch. If it is not practical to do this, minimize the number of ISLs the host and storage need to traverse.
- Some of the switch class products being produced today only contain a single ASIC. If this is the case, then the positioning of the host and storage ports is strictly a matter of personal preference. However, if the switch being used contains multiple ASICs, try to connect host and storage pairs to the same ASIC. This prevents using the shared internal data transfer bus and reduces switch latency. In addition to performance concerns, consider fault tolerance as well. For example, if a host has two HBAs, each one accessing its own storage port, do not attach both HBAs, both storage ports, or all of the HBA and storage ports to the same ASIC.

  Note: If you are unsure of the ASIC layout for the switch you are working with, refer to Appendix B, "EMC Connectrix Products Data".

- When working with hosts that have more than one connection to more than one storage port, always connect the HBAs and, if possible, the storage ports that it accesses to different FC switches. If a completely separate fabric is available, connect each HBA and storage port pair to different fabrics. Please refer to the "Methodology 1: Balanced fabrics" on page 713 section for more information.

Switch and fabric management

The following are management interfaces common to all platforms:

- CLI

  The command line interface can be accessed through Tenet. On some platforms, it is also possible to access the CLI using a serial cable.

- SNMP

  Simple Network Management Protocol is a TCP/IP protocol that generally uses the User Datagram Protocol (UDP) to exchange messages between a management device and a network management system.

Specific switch and fabric management information follows:

Connectrix B

All switches in the Connectrix B family can be managed by CLI, Web Tools, and Fabric Manager. EMC supports the use of EZSwitchSetup for the Connectrix DS-220B and the DS-4100B Fibre Channel switch products.

- Web Tools

  Web Tools provides switch / fabric management through a web browser pointed to the IP address of the switch. Brocade uses a built-in web server for this function.

- Fabric Manager

  Fabric Manager is a licensed management application used to manage a single switch or a large SAN in real-time.

- EZSwitchSetup

  EZSwitchSetup is an easy-to-use web-based management application that allows you to assign the IP address, switch name, time, switch password, and zoning configuration.

  EZSwitchSetup assists in the setup of a single switch SAN using either the Connectrix DS-4100B or the DS-220B. EZSwitchSetup uses the admin account as the default login.

  EZSwitchSetup can assist with the following functions in a single switch SAN:

  - Monitor the switch, including port and FRU status
  - Manage basic zoning
  - Perform basic switch configurations

  EZSwitchSetup gives the SAN administrator the ability to choose from three different zoning options:

  - Typical Zoning

    By selecting the **Typical Zoning** option, user-selectable ports are zoned for storage devices, and the remaining are used for hosts. EZSwitchSetup automatically creates a zone with the storage ports and host ports using the port zoning method. No device WWN is needed for this setup, and the devices do not have to be connected. The GUI shows you where to hook both the host and storage devices when finished.

Select **Typical Zoning** if:

- You want to set up and manage a switch without knowing about domains.
- You do not know how zoning works.
- You do not know where to connect devices.

- Custom Zoning

  By selecting the **Custom Zoning** option, the user is given a matrix of all storage and host ports connected to the switch. This allows the user the option of customizing which hosts can access which storage by selecting the crossing point of each storage and host.

  Select **Custom Zoning** if:

  - You want to customized which initiators access which targets, but do *not* want to configure the zones yourself.

- Advanced Zoning

  By selecting the **Advanced Zoning** option, the advanced zoning tool used in Web Tools is launched and you have the option of using mixed zoning (port and WWN) on the switch.

  Select **Advanced Zoning** if:

  - You are very familiar with zoning and zoning practices.

Connectrix MDS

- Device Manager
- Fabric Manager
- Fabric Manager server

Connectrix M

- Connectrix Manager Basic

  Connectrix Manager Basic software, formerly SANpilot, is a web-based management tool for small and medium businesses. It allows for management of fabrics with up to six switches. It is included at no extra cost on every Brocade M-Series Sphereon™ fabric switch.

- Connectrix Manager

  Connectrix Manager is a stand-alone Java-based application that utilizes a client/server architecture to provide a management interface. The Connectrix Manager server is a headless application that runs on the service processor.

Note: The service processor provides a dial-home facility and data backup function.

The Connectrix Manager server also provides an interface for Connectrix Manager clients, an SNMP Agent, alerting using SNMP traps/email, and an ECCAPI to allow for management through EMC ControlCenter®.

The Connectrix Manager client runs on both the Service Processor and on remote workstations, and it is the visible user interface to the server.

QLogic

At the time of this publication, the following applications can be used to manage a QLogic fabric:

- Enterprise Fabric Suite 2007

  The Enterprise Fabric Suite 2007 is a combined application for management of multiple QLogic fabrics, for performance monitoring, fabric monitoring, and extended distance. The Enterprise Fabric Suite 2007 replaces the SANsurfer Switch Manager that QLogic supported before the 6.2.x firmware code.

- Quicktools

  Quicktools is an embedded application for basic discovery, setup, configuration, and zoning management for a QLogic switch. It replaces the embedded SANsurfer Switch Manager that QLogic supported before the 6.2.x firmware code.

Security

It is important to secure your fabric. General security best practices for an FC SAN are as follows:

- Implement some form of zoning
- Change default password
- Disable unused or infrequently used Management Interfaces
- Use SSL or SSH if available
- Limit physical access to FC switches

Specific switch and fabric management information follows:

Connectrix B

Fabric Operating Software (FOS) versions 5.2.x and above now include an access control list (ACL) feature which gives the SAN administrator the ability to restrict both device and switch login throughout the fabric.

The following two ACL policies offer the base Fabric Operating System (FOS):

- Device Connection Control (DCC) policies
- Switch Connection Control (SCC) policy

FOS versions 5.2.x and above introduced Role Based Access Control (RBAC).

Also introduced in FOS versions 5.2.x and above was the concept of Admin Domains (AD).

To maintain a secure network, you should avoid using Telnet (you can use secTelnet if you are using Fabric OS v2.6 or later, or v3.1 or later), or any other unprotected application when you are working on the switch. For example, if you use Telnet to connect to a machine, and then start an SSH or secure Telnet session from that machine to the switch, the communication to the switch is in clear text and therefore *not* secure.

The FTP protocol is also not secure. When you use FTP to copy files to or from the switch, the contents are in clear text. This includes the remote FTP server login and password. This limitation affects the following commands: **saveCore**, **configUpload**, **configDownload**, and **firmwareDownload**.

Connectrix MDS

The MDS supports SSH and SFTP protocols.

Connectrix M

After the environment has been configured and all hosts see the appropriate volumes, configure the following security features.

- For homogeneous Brocade M-Series fabrics:
  - If Enterprise Fabric mode is available, enable it.
  - If Enterprise Fabric mode is not available, enable:
    - Fabric binding
    - Switch binding

  - Port binding

- For heterogeneous fabrics containing Brocade M-Series switches:
  - Enable switch binding and port binding.

QLogic

After building the QLogic fabric, it is important to secure the fabric.

- QLogic switches offer a mix of fabric security protection features which include user security, connection security, and device security.
- The data path for switch management communication is encrypted using secure shell (SSH) for the CLI and Secure Socket Layer (SSL) for Enterprise Fabric Suite 2007 and Quicktools.
- Users can set up device connection security to control what devices have access to the switch. ISL and port authentication is achieved using Fibre Channel Security Protocol (FC-SP) and DH-CHAP. Additional device authentication is performed using FC-GS-4 Ct.
- The Port Binding security feature was introduced in the latest firmware v6.2.1 supported at the time of this publication. As a part of the QLogic security feature, port binding enables the specification of WWNs that are allowed to connect to a switch port. Up to 32 WWNs may be specified. Port binding is available through the CLI **set config security port** command.

For more information on these security features, refer to Chapter 4, ”Secure SANs.”

The *SANbox 5602 Switch Management User’s Guide, Managing Fabrics* is available at: http://download.qlogic.com/manual discusses how to enable security features using switch management interfaces.

Simple Fibre Channel SAN topologies

This section provides examples of simple Fibre Channel SAN topologies.

- "Single switch fabrics," next
- "Two switch fabrics" on page 79
- "Blade switch with direct attached storage" on page 129

Single switch fabrics

This section provides examples of single switch fabrics.

Overview of Fabric Design considerations

General layout

A single switch fabric, consists of only a single switch (Figure 1). The switch is also connected to a single management LAN through IP.

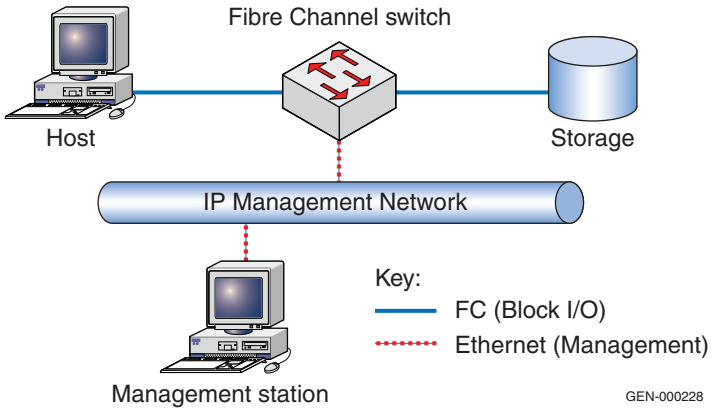


Figure 1 **Single switch topology example**

Best practices

For general best practices in single switch fabrics, refer to "Best practices" on page 37.

Host and storage layout

Note: The correct way to attach hosts and storage to a SAN is completely dependent upon the customers' environment, but the following information may be helpful.

In this example, both the host and storage ports are located on the same switch. There is also the opportunity to locate the host and

storage pairs so that they are in the same quad or octet of ports which are controlled by the same switch ASIC. This connection scenario helps eliminate frames traveling over the backplane of the switch and increases the speed of frame routing. Many switch class products being sold today are switch-on-a-chip architecture, and do not contain many discrete ASICs. For these architectures, port placement for performance or HA concerns does not need to be considered.

For general information on host and storage layout in single switch fabrics, refer to "Host and storage layout" on page 43.

Switch and fabric management

All switch and fabric management tools can be used to configure this environment. For general information on switch and fabric management, refer to "Switch and fabric management" on page 44.

Security

Even when dealing with single switch fabrics, it is important to think about security. Typically, in a single switch fabric, if physical access to the switch can be controlled, default passwords are changed, and unused interfaces are disabled, zoning will be enough to secure the fabric. It is also recommended to implement some form of port binding. This can be as simple as enabling port binding if the feature exists, or as complicated as disabling all unused ports and hard-setting port type to F_Port.

For more information on security, refer to "Security" on page 47.

Connectrix B example

General layout

Figure 2 shows a single switch SAN using DS-4100B, and EZSwitchSetup with typical zoning.

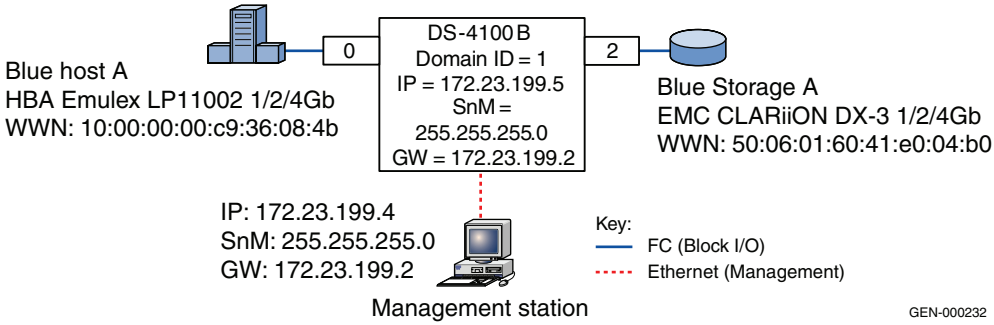


Figure 2 **Single switch SAN using DS-4100B and EZSwitchSetup with zoning**

Supported switches

Connectrix DS-220B and DS-4100B are supported with EZSwitchSetup.

Best practices

For general best practices in single switch fabrics, refer to "Best practices" on page 50.

For Connectrix B specific best practices, refer to "Connectrix B" on page 40.

Host and storage layout

When using Typical Zoning in EZSwitchSetup, the GUI guides the SAN administrator on recommended placement of both host and storage.

In a single switch SAN design, both hosts and storage can be placed anywhere on the switch. Connectrix B switches group ASIC-controlled port groups, either by quad or octet. In some circumstances, it you should connect both host and storage to these groups which eliminates frames from traveling through the backplane of the switch to reach their destination.

For general information on host and storage layout in single switch fabrics, refer to "Host and storage layout" on page 50.

Switch and fabric management

For this case study, the EZSwitchSetup tool is used to setup a Connectrix DS-4100B switch with one host and one storage array. EZSwitchSetup offers three zoning options. The Typical Zoning option is used in this case study. For more information on EZSwitchSetup, refer to "EZSwitchSetup" on page 45.

For general information on host and storage layout in single switch fabrics, refer to "Switch and fabric management" on page 51.

Security

For general information on security in single switch fabrics, refer to "Security" on page 51.

Setting up this topology

Assumptions specific to this case study:

- The switch is not powered on.
- Network drop, IP addresses, subnet mask, and gateway are provided by the customer.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.

- The customer has provided a server or laptop with CD drive, serial DB-9 connector, and a NIC card to be used to configure the switch.
- Configuration will be done using the *EZSwitchsetup CD*.

Configure a single switch SAN

To configure a single switch SAN, follow these steps:

Note: Host and storage ports do not have to be connected until Step 8.

1. Power-on the switch.
2. Connect to the switch management port using RS-232 serial cable. Set the workstation serial port to use 9600 baud rate, 8 databits, no parity, 1 stop bit, and no flow control.
3. Connect RJ-45 network cable from the workstation to the network management port of switch.
4. Launch the EZSwitchSetup CD.
5. Issue the switch IP address (**172.23.199.5**), subnet mask (**255.255.255.0**), and default gateway (**172.23.199.2**) and then click **Next**.
6. Supply a password for the Admin account, issue a switch name, change the date, and then click **Next**.
7. Select **Typical Zoning**, on the **Select Zoning** screen, and then click **Next**.
8. Select the ports as host or storageon the **Configure Typical Zoning** screen, , and then click **Next**.
   Each time you click, you toggle the port between a desired host port, designated by the color blue, the letter **H**, and a desired storage port, designated by the color green, and the letter **S**.
9. Connect host and storage using the **Configure Typical Zoning** GUI as a guide.
10. Use the pull-down menu selections on the **Specify Devices** screen to select the number of HBA connections and the number of storage connections planned for the switch, and then click **Next**.

11. The **Connect Devices** screen provides a graphical view of recommended connections for both host and storage on the switch. Connection status is represented by colored lines.

    - A green line indicates a good connection
    - A red line indicates an invalid connection
    - A blue dashed line indicates a missing connection

    Click **Next**.

12. The **Finish** screen supplies a summary of the switch configuration and allows you to print the configuration if needed. Click **Finish**.

13. Validate zoning configuration by ensuring that each HBA has access to the storage device(s).

    Each HBA should see every storage device when using **Typical Zoning**.

    The **Validate** link in the task panel of the **Switch Manager** page checks for devices that are not zoned properly, allows you to delete the devices from the zoning database, and display a matrix of which HBA can see which storage device.

Connectrix MDS example

General layout

Figure 3 illustrates an Connectrix MDS 9506 using VSANs.

Note: VSANs will be configured and used in the following example. Refer to Appendix B, "EMC Connectrix Products Data" for more information on the Connectrix MDS VSAN feature.

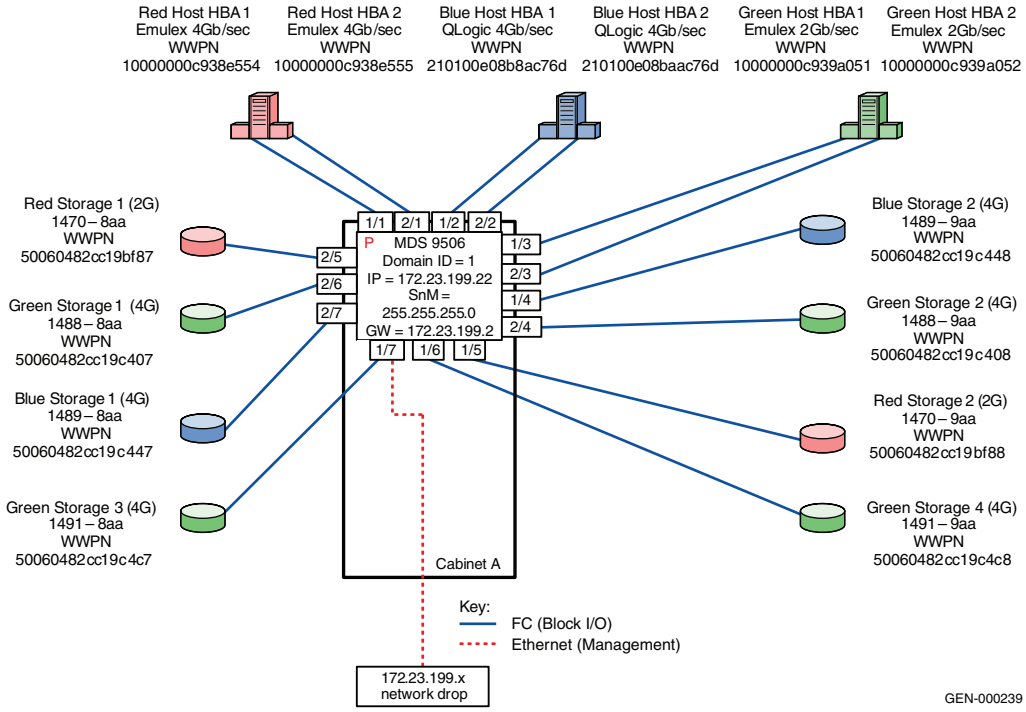


Figure 3 **Connectrix MDS 9506 using VSANs**

Best practices

For general information on best practices in single switch fabrics, refer to "Best practices" on page 50. Specific information for this example follows.

Port fencing is on by default.

Host and storage layout

For general information on host and storage layout in single switch fabrics, refer to "Host and storage layout" on page 50. Specific information for this example follows.

There are no host or storage restrictions for Line Rate Mode cards . Oversubscribed cards should be used for hosts only.

Switch and fabric management

For general information on switch and fabric management in single switch fabrics, refer to "Switch and fabric management" on page 51. Specific information for this example follows.

Cisco Fabric Manager can be used for this example.

Security

For general information on security in single switch fabrics, refer to "Security" on page 51. Specific information for this example follows.

Port binding can be used for security.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and Setup Manual*, which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data".
  - For Cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and Setup Manual*, which can be accessed from Powerlink.
- The switches have *not* been connected to the power source and are *not* powered on.
- Network drops, IP addresses, subnet mask and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix B, "EMC Connectrix Products Data".

Note: Connectrix MDS switches can be placed on either a public or private network. There are advantages to both configurations. For more information, refer to "Public versus private" on page 961.

In this example it is assumed that The customer has provided one Ethernet cable and an IP of 172.23.199.22.

- The proper number of line cards have been installed in each chassis. In this case, two line cards in each chassis are required and installed in slots 1 and 2.
  - For help in determining how many ports are required, see "Determining customer requirements" on page 759.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- A laptop, supplied by the installer, will be used to configure the IP addresses of the switches, and this laptop has a serial DB-9 connector.
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- Fabric Manager will be used for VSAN setup.

Configure the IP address

To configure the IP address:

1. Power up the cabinet by connecting power cords to the power receptacles provided by the customer.
2. Select one of the switches to configure and set the IP to 172.23.199.22.
3. Supply a network connection to the appropriate subnet.
4. Connect to the serial port of the switch, using an RS232 serial cable, with a baud rate of 9600, 8 data bits, no parity, 1 stop bit and no flow control.

   The **login** prompt appears.
5. Log in the first time with a user name of *admin* and a password of *admin*.

   You are prompted to supply a new strong password for CLI user admin.
6. For this example, select **no** when asked if you want to run setup.

   Note: This example starts with the switch that has a Domain ID of **1** and an IP address of **172.23.199.22**.

7. Repeat the above steps for each switch, supplying the appropriate IP.

CLI commands to configure the IP and gateway

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***interface mgmt 0***

  Switch(config-if)#***ip address 172.23.199.22 255.255.255.0***

  End with **CNTL/Z**.

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***ip default-gateway 172.23.199.2***

  End with **CNTL/Z**.

To authorize access on a switch for Device and Fabric Manager, run this command on every switch while supplying a user name (nnn) and password (ppp):

- Switch#***conf t***

  Switch(config)# ***snmp-server user nnn network-admin auth md5 ppp***

  Switch(config)#***end***

  Switch# ***copy running-config startup-config***

  Switch# ***exit***

Install Fabric Manager and Device Manager

To install Fabric Manager and Device Manager:

1. Open your web browser.
2. Enter the IP address of the switch into the address bar.
3. Follow the prompts and accept all defaults to install both Fabric Manager and Device Manager.

Fabric Manager and Device Manager can be started using the configured SNMP server user name and password in "CLI commands to configure the IP and gateway" on page 162.

Configure a VSAN

To configure a VSAN:

1. Open the Device Manager for the switch with an IP address of **172.23.199.22.**

2. Open the **VSAN** dialog by selecting **VSAN**.
3. Click **Create**.
4. Enter the value of **100** into the **VSAN ID** field.
5. Set the **VSAN Name** to be **"Red_VSAN_100"**.
6. Use the default interop mode.
7. Click **Create**.

Configure the other VSAN in this physical switch

1. Repeat Step 2 through Step 7, above, for VSAN 200 and 300 noting that:
   - For Virtual switch 200, use a VSAN name of **"Green_VSAN_200"**.
   - For Virtual switch 300, use a VSAN name of **"Blue_VSAN_300"**.
2. Using the following table, assign and enable the ports to the proper VSAN using Device Manager for the switch with the IP address **172.23.199.22**.

Slot #	Port #	Name	VSAN ID
1	1	Red Host HBA 1	100
1	2	Blue Host HBA 1	300
1	3	Green Host HBA 1	200
1	4	Blue Storage 2	300
1	5	Red Storage 2	100
1	6	Green Storage 4	200
1	7	Green Storage 3	200
1	8		
2	1	Red Host HBA 2	100
2	2	Blue Host HBA 2	300
2	3	Green Host HBA 2	200
2	4	Green Storage 2	200
2	5	Red Storage 1	100

Slot #	Port #	Name	VSAN ID
2	6	Green Storage 1	200
2	7	BlueStorage 1	300
2	8		

Connect cables

To connect the cables:

1. Connect host and storage ports.
2. Attach fiber cable between switches and N_Ports.

Configure domains

To configure domains:

1. Open **Fabric Manager**. A topology of two switches appears.
2. From **Fabric Manager** open the **"Red_VSAN_100"** folder.
3. Select **Domain Manager**.
4. Select the **Configuration** menu.
5. Set a Domain ID, **1** for switch **172.23.199.22** and **2** for **172.23.200.22**.
6. To set a principal switch, set the priority to **1** in the domain menu.
7. Repeat Step 2 through Step 6 for **"Green_VSAN_200"** and "**Blue_VSAN_300"**.

Zone hosts and storage

To zone hosts and storage:

1. From **Fabric Manager,** select **"Red_VSAN_100"**.
2. Select **Edit Full Zone Database**.
3. Create a zone by selecting the **Zone** menu and clicking the **Add** button.
4. Provide a descriptive name for the zone. This example zones "Red host HBA 1" and "Red Storage 1", so **"RedHBA1_1470_8aa"** will be used. Press **OK**.
5. Locate, then single-click, **"Red Host HBA 1"** (WWPN 10000000c938e554) in the **Potential zone members list**.

6. Click the right-pointing arrow on the divider between the **Potential members list** and the zones list to add the HBA to the zone. Select **Add to zone or alias**.
7. Locate, then single click, **"Red Storage 1"** (WWPN 50060482cc19bf87) in the **Potential zone members list**.
8. Click the right-pointing arrow on the divider between the **Potential members list** and the **zones list** to add the Storage port to the zone.
9. Repeat Step 2 through Step 8 for all host and storage pairs in the environment.
10. Create a zone set by selecting the **Zonesets** menu, and then clicking on the **Add** button.
11. Provide a descriptive name for the zone set. This example uses the name **"RED Fabric 1"**.

Add only those zones that are necessary on Red_Fabric_1. In this case, the two red zones listed below, "**RedHBA1_1470_8aa**" and "**RedHBA2_1470_9aa**" are used. Repeat for other Fabrics, when completed, you should have three zone sets as shown below.

```
Zone set name = "Red_Fabric_1"

          Zone name = "RedHBA1_1470_8aa"
             Zone Member = "10000000c938e554"
             Zone Member = "50060482cc19bf87"

                 "Red_Fabric_2"

          Zone name = "RedHBA2_1470_9aa"
             Zone Member = "10000000c938e555"
             Zone Member = "50060482cc19bf88"

Zone set name = "Blue_Fabric_1"

          Zone name = "BlueHBA1_1489_8aa"
             Zone Member = "210100e08b8ac76d"
             Zone Member = "50060482cc19c447"

              "Blue_Fabric_1"

          Zone name = "BlueHBA2_1489_9aa"
             Zone Member = "210100e08baac76d"
             Zone Member = "50060482cc19c448"
```

```
Zone set name = "Green_Fabric"

          Zone name = "GreenHBA1_AllGreenStorage"
             Zone Member = "10000000c939a051"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"

          Zone name = "GreenHBA2_AllGreenStorage"
             Zone Member = "10000000c939a052"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this time, the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the *Fabric OS Configuration Guide* for more details.

Configure security

- Enable switch binding.
- Enable port binding.

Configure proactive monitoring and countermeasures

ISL thresholds are 80 % by default.

Configure port fencing

Port fencing is on by default.

Connectrix M example

General layout

Figure 4 illustrates a single switch fabric built with a DS-4700M.

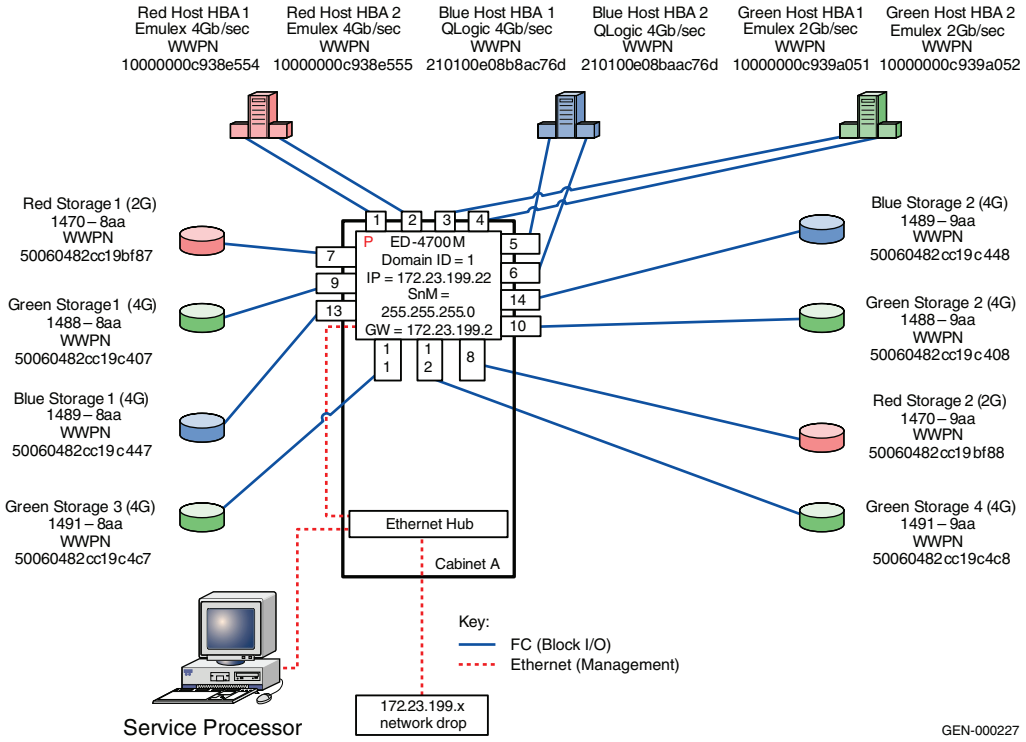


Figure 4 **Single switch fabric built with a DS-4700M**

All Connectrix Mswitches can be used in this configuration.

Best practices

- Configure threshold alerts:

  Even though this is a single switch fabric and ISLs are not used, some benefits can still be realized from configuring threshold alerts, especially in configurations with large fan in and fan out ratios.

- Enable a fencing policy:

  In this example, a fencing policy is enabled to help protect the switch from any misbehaving N_Ports. This is usually more of a concern on large port count switches where many misbehaving N_Ports can create an issue by tying up the switch processor with needless login requests. Even though this problem occurs in single switch fabrics, so we still recommend it as a best practice.

For general information on best practices in single switch fabrics, refer to "Best practices" on page 50.

Host and storage layout

In this case study, a DS-4700M is utilized. Since the DS-4700M is a switch-on-a-chip type of architecture, the host and storage ports can be placed anywhere. No special considerations need to be made in regards to port placement for performance or HA related issues.

For general information on host and storage layout in single switch fabrics, refer to "Host and storage layout" on page 50.

Switch and fabric management

Any management interface could be utilized to set up this environment, but in this case study Connectrix Manager Basic is used.

For general information on switch and fabric management in single switch fabrics, refer to "Switch and fabric management" on page 51.

Security

In this example, unused interfaces are disabled and Web SSL, fabric binding, switch binding, and port binding are configured. See "Configure security" on page 70.

For general information on security in single switch fabrics, refer to "Security" on page 51. For more information on security, refer to Chapter 4, "Secure SANs."

Setting up this topology

Assumptions specific to this case study:

- The switch is installed in a customer-supplied cabinet.
- The customer has provided the proper power receptacles.

  Note: For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data."

- The switch is *not* connected to the power source and is *not* powered on.

- The serial port is not used and the IP addresses will be programmed through the network interface.
- The customer has provided the network drop, IP addresses, subnet mask, and gateway.
- License keys have been obtained.

  Note: To obtain a license key, go to the URL listed on the transaction code certificate that shipped with the product.

- The customer has provided a server to be used to configure the switches.
- One interface from the customer-supplied server is directly connected to the FC switch through a crossover cable, and this interface has an IP address of 10.1.1.11.
- The customer has provided a temporary password used as the default password when configuring the IP address.
- Connectrix Manager is not used; instead the switch is managed through Connectrix Manager Basic (EWS).

Configure the IP address

Note: Connectrix M switches ship with a default IP address of 10.1.1.10.

To configure the IP address:

1. Power up the switch by connecting the power cords to the power receptacles provided by the customer.
2. Open a command prompt and ping the IP address of **10.1.1.10**.
   - If you get a response, continue to step 2.
   - If you do not get a response, either IML the switch by pushing the gray button on the CTP, or configure the IP address using the serial port.

     Note: See the *Installation and Service Guide* for more information about configuring the IP address using the serial port.

3. Launch Internet Explorer and type the IP address **10.1.1.10** in the **Address** bar, and press **Enter**. The **Login** banner appears.
4. Click **Accept** and the **Login** dialog appears.

5. Enter the default username and password, *Administrator* and *password*, and press **Enter**.

IMPORTANT

It is strongly recommended that when prompted to change the password, change it to the password provided by the customer.

6. Click the **Details** button below the switch icon in the **Connectrix Manager Basic Topology** view.
7. Navigate to the **Network Configuration** dialog by clicking the **Configure** pull-down menu and selecting **switch,** then **Network.**
8. Enter the IP address (**172.23.199.22**), subnet mask (**255.255.255.0**) and gateway (**172.23.199.2**), then click **OK**.

Note: The error message *Unable to display…* appears. This appears because the IP address of the switch just changed and you are no longer able to connect to the switch at its old IP address.

9. Launch Internet Explorer and type the IP address **172.23.199.22** in the **Address** bar, and press **Enter**. The **Login** banner displays.
10. Click **Accept** and the **Login** dialog displays.
11. Enter the username and password that were configured at the end of Step 5.

Configure the FC switches

To configure the FC switches:

1. Set the switch name and fabric parameters.
   a. From the **Product > Hardware** screen in Connectrix Manager Basic, under **configure / Switch / Parameters**, ensure that the preferred Domain ID is set to the appropriate number (**1** in this example).
   b. Check the **Insistent Domain ID** checkbox to enable insistent Domain IDs.

Note: If the preferred Domain ID and the active Domain ID do not match, you cannot set the preferred Domain ID and enable insistent Domain IDs at the same time. Instead change the preferred Domain ID, click **OK**, re-open the dialog, check the **Insistent Domain ID** checkbox, and then click **OK**.

c. Click **OK**.

d. From the **Product > Hardware** screen in Connectrix Manager Basic, under **configure / Switch / Fabric Parameters**, ensure that:

– **Interop Mode** is set to **Open Fabric 1.0**

– **Switch Priority** is set to **Principal**.

e. Click **OK**

f. From the **Product > Hardware** screen in Connectrix Manager Basic, under **configure / Switch / Identification**, enter the switch name into the **Name** field.

g. Click **OK**.

2. Configure the switch ports.

Note: In this example all of the switch ports are internally connected to the same ASIC. As a result, port layout is strictly a matter of user preference. In addition, outside of a hardware requirement (that is, Connectrix M QPM) or a known problem with auto-negotiation existing between two port types, it is recommended that you leave ports at auto-negotiate for both port type and speed.

a. From the **Product > Hardware** screen in Connectrix Manager Basic, under **configure / Ports / Basic Information**, enter the port name shown in the following table:

ASIC #	Port #	Name	Speed
1	Port 0	Red Host HBA 1	Negotiate
	Port 1	Red Host HBA 2	Negotiate
	Port 2	Green Host HBA 1	Negotiate
	Port 3	Green Host HBA 2	Negotiate
	Port 4	Blue Host HBA 1	Negotiate
	Port 5	Blue Host HBA 2	Negotiate
	Port 6	Red Storage 1	Negotiate
	Port 7	Red Storage 2	Negotiate
	Port 8	Green Storage 1	Negotiate
	Port 9	Green Storage 2	Negotiate
	Port 10	Green Storage 3	Negotiate

ASIC #	Port #	Name	Speed
1	Port 11	Green Storage 4	Negotiate
	Port 12	Blue Storage 1	Negotiate
	Port 13	Blue Storage 2	Negotiate

Connect cables

To connect cables:

1. Connect ISLs.
   a. Attach fiber cable between the switches as shown in Figure 4 on page 63.
2. Connect host and storage ports.
   a. Attach fiber cable between switches and N_Ports.
3. Verify port login status.
   a. After all cables are connected, use Connectrix Manager Basic to verify the all of the ports logged into the switch.

Note: The login status can be verified by referring to the node list, which is under the product pull-down menu.

Zone host and storage

To zone host and storage:

1. In Connectrix Manager Basic, under **Configure / Zoning,** open the **Zoning** dialog.
2. Create a zone by entering a descriptive name for the zone in the **Zone Name** field. This example zones “Red host HBA 1” and “Red Storage 1”, so “RedHBA1_1470_8aa” will be used.
3. Locate, then single click **Red Host HBA 1** (WWPN 10000000c938e554) in the **potential zone members** list.
4. To add the HBA to the zone, click the right-pointing arrow on the divider between the **potential members list** and the **Zone Name** list.
5. Locate, then single click, **Red Storage 1** (WWPN 50060482cc19bf87) in the **potential zone members** list.
6. To add the storage port to the zone, click the right-pointing arrow on the divider between the **potential members list** and the **Zone Name** list.

7. Create a zone set by entering a descriptive name for the zone set in the **Zone Set** field. This case will use the date of "Oct_31_06_1140".

   Note: Entering a descriptive Zone set name is only required when creating the first zone. Skip this step when adding additional zones.

8. Click the right-pointing arrow between the **Zone Name** and **Zone Set** lists.

9. Repeat Step 2 through Step 8 for all host and storage pairs in the environment.

Upon completion, the zone set should be similar to the following:

```
Zone set name = "Oct_31_06_1140"

            Zone name = "RedHBA1_1470_8aa"
               Zone Member = "10000000c938e554"
               Zone Member = "50060482cc19bf87"

            Zone name = "RedHBA2_1470_9aa"
               Zone Member = "10000000c938e555"
               Zone Member = "50060482cc19bf88"

            Zone name = "BlueHBA1_1489_8aa"
               Zone Member = "210100e08b8ac76d"
               Zone Member = "50060482cc19c447"

            Zone name = "BlueHBA2_1489_9aa"
               Zone Member = "210100e08baac76d"
               Zone Member = "50060482cc19c448"

            Zone name = "GreenHBA1_AllGreenStorage"
               Zone Member = "10000000c939a051"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"

            Zone name = "GreenHBA2_AllGreenStorage"
               Zone Member = "10000000c939a052"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking, and modifying host configuration files are required before the SAN setup is complete. Refer to the OS configuration guide for more detail.

Configure security

Once all of the hosts see their volumes, security can be configured.

Disable unused interfaces:

1. From Connectrix Manager Basic, uncheck **Enable Software** and **Enable CLI** options. This prevents users from managing the switch through Connectrix Manager or Telnet.

Enable Web SSL:

1. Click **SSL** from the **Configure** menu.
2. Click **Enable** for **Web SSL**.
3. Log in to the switch again.

Note: It is advisable to change the password now.

Enable Fabric Binding:

1. Click **security,** and then select **Fabric Binding**.
2. Click **Load Current Fabric**.
3. Click **Activate**.
4. Click **Close**.

Enable Switch Binding:

1. Click **security,** and then select **Switch Binding**.
2. Select **Enabled, Restrict E_Ports** from the **Switch Binding State** pull-down list.
3. Locate the switch WWPNs in the **Attached Device WWN** list and add them to the **Switch binding Member** list one at a time by highlighting each switch entry, and then clicking on the arrow pointing to the right.

   Note: All switch entries start with 1000080088.

4. Click **Update**.

5. Click **Close**.

Enable Port Binding:

1. Click **Security,** and then select **Port Binding**.
2. Ensure that every box in the **Attached** column has a check in it.

Note: As soon as this is done, devices that are not currently connected to each port cannot connect. This holds true for ports that are currently connected to the switch but on another port. This means that swapping ports for troubleshooting does *not* work unless port binding is disabled on both ports being swapped.

---

3. Click **OK**.

Configure proactive monitoring and counter-measures

Configure threshold alerts:

1. Open a command prompt and Telnet to the switch by typing **Telnet 172.23.199.22**.
2. Log in with the username and password of the switch.
3. Enter the **Perf** command.
4. Enter the **thresh** command.
5. Enter the **throughput** command.
6. Enter the **addalert eightyPercent** command.
7. Enter the **addport eightyPercent all** command.
8. Enter the **setParams eightyPercent 1 5** command.
9. Enter the **setUtilPercentage eightyPercent 80** command.
10. Enter the **setutiltype eightyPercent 3** command.
11. Enter the **..** command.
12. Enter the **setstate eightyPercent 1** command.
13. Enter the **logo** command.

Configure Port Fencing:

1. Click **Configure** and select **Port Fencing** from the list.
2. Enable the **Default Security Policy** by highlighting the entry and clicking **Enable**.

3. Enable the **Default Link Level Policy** by highlighting the entry and clicking **Enable**.

QLogic example

General layout

Figure 5 shows a single switch fabric built with a QLogic SB5602

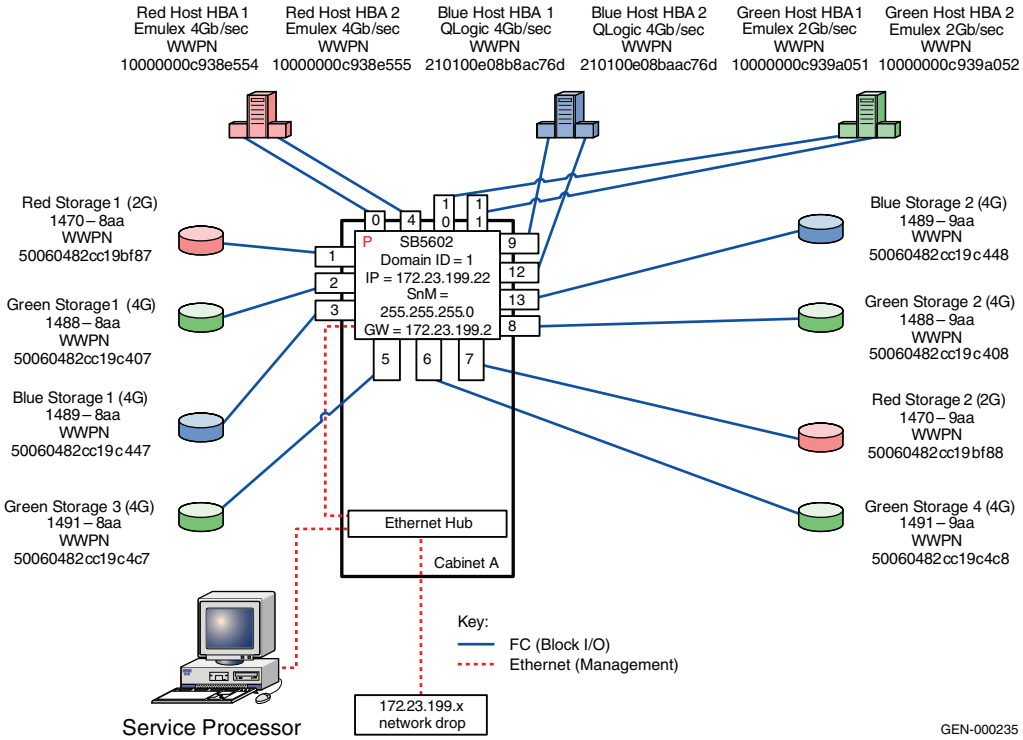


Figure 5 **Single switch fabric built with QLogic SB5602**

Best practices

For general information on best practices in single switch fabrics, refer to "Best practices" on page 50. For QLogic-specific best practices, refer to "QLogic" on page 42.

Host and storage layout

In this case study, a SANbox 5602 is utilized. Since the SANbox 5602 is a switch-on-a-chip (single Agate 2 ASIC) type of architecture, the host and storage ports can be placed anywhere. No special considerations need to be made in regards to port placement for performance or HA related issues.

For general information on host and storage layout in single switch fabrics, refer to "Host and storage layout" on page 50.

Switch and fabric management

For general information on switch and fabric management in single switch fabrics, refer to "Switch and fabric management" on page 51. For QLogic-specific information, refer to "QLogic" on page 47.

Security

For QLogic security information that may be applicable, refer to "Security" on page 119.

For general information on security in single switch fabrics, refer to "Security" on page 51.

Setting up this topology

Assumptions specific to this case study:

- The QLogic SB5602 switches are installed in a customer-supplied cabinet.

  For switch installation guidelines, refer to the Sections 3 and 4: Planning and Installation for an SB5602 switch in the following document:
  http://download.qlogic.com/manual/27144/59096-01_A.pdf

- The customer has provided the proper power receptacles.

  The power requirements for the SB5602 are 1 Amp at 100VAC or 0.5 Amp at 240VAC. For cabinet power requirements, refer to the customer-supplied cabinet setup manual, if any.

- The switch has *not* been connected to the power source and is *not* powered on.
- Do not use the serial port. Program IP addresses with the network interface.
- Network drop, IP addresses, subnet mask, and gateway, have been provided by the customer.
- License keys have been obtained.

  Note: To obtain the license key, go to the URL listed on the transaction code certificate that shipped with the product.

- The customer has provided a server to be used to configure the switches.
- One interface from the customer-supplied server is directly connected to the FC switch using a crossover cable and this interface has an IP address of 10.0.0.253.
- The customer has provided a temporary password to be used as the default password when configuring the IP address.

- The QLogic CLI commands is used to configure this switch.

Note: The SANsurfer switch manager will be used in the QLogic case study for configuring a two switch fabric (refer to "QLogic example" on page 118).

Configure the IP address

Note: QLogic switches ship with a default IP address of 10.0.0.1.

To configure the IP address:

1. Power up the switch by connecting the power cords to the power receptacles provided by the customer.
2. Open a command prompt and ping the IP address of **10.0.0.1**.
   - If you get a response, continue at step 2.
   - If you do not get a response, you can either hard reset the switch, (verify that the input power LED is illuminated, the heartbeat LED is blinking, and that the system fault LED is not on), or configure the IP address using the serial port.

   Note: See the *Installation and Service Guide* for more information about configuring the IP address using the serial port.

3. Log on to a switch using Telnet. Open a command line window on the workstation, and enter the Telnet command followed by the switch default IP address:

   # telnet 10.0.0.1

4. A Telnet window for the switch opens, prompting for a login.
5. Enter the default username, password and authority (if prompted), as **admin, password,** and **Admin**, and then click **Enter.**

IMPORTANT

It is strongly recommended that when prompted to change the password, you change it to a password that was provided by the customer.

6. Commands related to switch configuration are available only within the admin session. Start an admin session by issuing the **admin start** command at the switch prompt.

7. Run the **set setup system** command at the switch prompt to change the system configuration settings. The system configuration fields display.

8. For each parameter, a new value can be entered as specified in Step 9, or press the **Enter** key to accept the current or default value shown in brackets. If you wish to terminate this process before reaching the end of the list, press **q** or **Q**, and then **Enter**.

9. Enter the following new values for the parameter fields listed on issuing the **set setup system** command in order to configure the switch IP address for remote logging:

   Eth0Network Discovery [Static]

   Eth0NetworkAddress [172.23.199.22]

   Eth0NetworkMask [255.255.255.0]

   Eth0GatewayAddress [172.23.199.2]

   Press **q** or **Q**, then **Enter**.

   Note: The switch Telnet window might close since the IP address of the switch has just changed and you are unable to connect to the switch at its old IP address any longer.

Configure FC switches

To configure the FC switches:

1. Set the switch name and fabric parameters.

   a. Enter the admin session by issuing the **admin start** command at switch prompt.

   b. Issue the **set switch config** command.

   c. A list of attributes with formatting and default values follows. A new value can be entered as specified in Step d, or the current value can be accepted by pressing the **Enter** key. To terminate this process before reaching the end of the list, press **q** or **Q**, then **Enter**.

   d. Enter the following new values for the parameter fields listed on issuing the **set switch config** command, in order to configure the switch name, a fixed Domain ID (set “Domain ID lock” to **true** to hardest the Domain ID to a fixed value), and the interopmode:

```
DefaultDomain ID    [1          ]
DomainIDLock        [True       ]
SymbolicName        [SB5602     ]
InteropMode         [Standard   ]
```

Press **Enter** for all the remaining fields to accept the default values.

2. Configure the switch ports.

Note: Outside of a hardware requirement or if a known problem with auto-negotiation exists between two port types. We recommend that you always leave ports at auto-negotiate for both port type and speed.

a. The **set configure port <portnumber>** command must be issued to configure the speed, port type, and port name for the different portnumbers that will be used in the topology. There are 16 ports on this switch.

b. Upon issuing this command, a list of port parameters appears. Only the **LinkSpeed**, **PortType,** and the **SymPortName** fields must be filled with the appropriate values as per the following table.

Port #	Symbolic port name	Port type	Port speed
0	Red Host HBA 1	F_Port	AutoNeg.
1	Red Storage 1	F_Port	AutoNeg.
2	Green Storage 1	F_Port	AutoNeg.
3	Blue Storage 1	F_Port	AutoNeg.
4	Red Host HBA 2	F_Port	AutoNeg.
5	Green Storage 3	F_Port	AutoNeg.
6	Green Storage 4	F_Port	AutoNeg.
7	Red Storage 2	F_Port	AutoNeg.
8	Green Storage 2	F_Port	AutoNeg.
9	Blue Host HBA 1	F_Port	AutoNeg.
10	Green Host HBA 1	F_Port	AutoNeg.

Port #	Symbolic port name	Port type	Port speed
11	Green Host HBA 2	F_Port	AutoNeg.
12	Blue Host HBA 2	F_Port	AutoNeg.
13	Blue Storage 2	F_Port	AutoNeg.

Connect cables

To connect cables:

1. Connect ISLs.
   a. Attach fiber cable between the switches as shown in Figure 5 on page 72.
2. Connect host and storage ports.
   a. Attach fiber cable between switches and N_Ports.
3. Verify port login status.
   a. After all cables are connected, issue the **show port** command to verify the all of the ports logged into the switch.

Zone host and storage

To zone host and storage:

1. Open the admin session by issuing the **admin start** command, and open the zoning session by issuing the **zoning edit** command.
2. Create a zone by issuing the **zone create [zone]** command with a descriptive name [zone] for the zone. This example zones “Red host HBA 1” and “Red Storage 1”, so “RedHBA1_1470_8aa” are used.
3. Add the zone members to the respective zones as described following Step 8. For example, add WWPN 10000000c938e554 ("Red host HBA 1") and WWPN 50060482cc19bf87 ("Red Storage 1") to the zone "RedHBA1_1470_8aa" created in the Step 2. This can be done by issuing the **zone add [zone] [member_list]** command, which in this case will be replaced by "zone add RedHBA1_1470_8aa 10000000c938e554 50060482cc19bf87". The member list can also take other formats as discussed in the *QLogic SB5602 Switch Configuration Guide*.
4. Similarly, create the other zones listed after Step 8.

5. Create a zone set by issuing the **zoneset create [zone_set]** with a descriptive name [zone_set] for the zoneset. This case uses the date of "Oct_31_06_1140".

6. Add the zones created in Step 3 and Step 4 to the zone set created in Step 5 by issuing the **zoneset add [zone_set] [zone_list]** command.

   In this case, it is replaced by "zoneset add Oct_31_06_1140 RedHBA1_1470_8aa RedHBA2_1470_9aa GreenHBA1_1489_8aa GreenHBA2_1489_9aa BlueHBA1_AllBlueStorage BlueHBA2_AllBlueStorage". The zone names in the zone list are separated using a <space>.

7. Close the zoning edit session by running a **zoning cancel** command at the prompt, before activating the zone.

8. Issue the **zoneset activate [zone_set]** command (n this case the "zoneset activate Oct_31_06_1140" command) to activate the zoneset.

Uponcompletion, the zone set should be similar to the following:

```
Zone set name = "Oct_31_06_1140"

            Zone name = "RedHBA1_1470_8aa"
               Zone Member = "10000000c938e554"
               Zone Member = "50060482cc19bf87"

            Zone name = "RedHBA2_1470_9aa"
               Zone Member = "10000000c938e555"
               Zone Member = "50060482cc19bf88"

            Zone name = "BlueHBA1_1489_8aa"
               Zone Member = "210100e08b8ac76d"
               Zone Member = "50060482cc19c447"

            Zone name = "BlueHBA2_1489_9aa"
               Zone Member = "210100e08baac76d"
               Zone Member = "50060482cc19c448"

            Zone name = "GreenHBA1_AllGreenStorage"
               Zone Member = "10000000c939a051"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"

            Zone name = "GreenHBA2_AllGreenStorage"
               Zone Member = "10000000c939a052"
               Zone Member = "50060482cc19c407"
```

```
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point, the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN setup is complete. Refer to the OS configuration guide for more detail.

Two switch fabrics

This section contains information on two switches in a Fibre Channel SAN topology.

Overview of fabric design considerations

General layout

Figure 6 shows an example of a two switch fabric. The switches are connected using two ISLs. Both switches are connected to Management Network A.

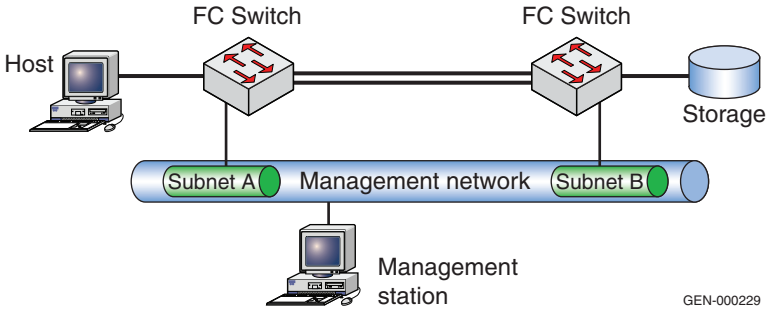


Figure 6 **Two switch fabric example**

Every switch type can be used in any position in this configuration.

Best practices

The following best practices are specific for two switch fabrics.

- ISL subscription best practice — While planning the SAN, keep track of how many host and storage pairs utilize the ISLs between domains. As a general best practice, if two switches are connected by ISLs, ensure that there is a minimum of two ISLs between them and that there are no more than six initiator and target pairs per ISL. For example, if 14 initiators access a total of 14 targets

between two domains, a total of three ISLs would be necessary. This best practice should not be applied blindly when setting up a configuration. Consider the applications that will use the ISLs.

- One of the use cases for a two switch fabric is distance extension. In these configurations, it is essential to monitor the ISLs for oversubscription conditions (utilization > 80%) which may lead to back pressure and any errors that are incrementing, especially bit errors or invalid transmission words, as these can lead to credit starvation. For more information on credit starvation, refer to "Buffer-to-Buffer credit information" on page 454. See the individual case studies below for information on how to configure in each environment.

For general information on best practices in all SANs, refer to "Best practices" on page 37.

Host and storage layout

Specific information for two switch fabrics follows.

- In the two switch fabric examples used in this section, hosts and storage can be connected to either switch, but should be connected to the *same* switch when possible. A notable exception to this is in a distance extension environment when the two switches are used to aggregate many different connections over an ISL and provide additional BB_Credit (buffer-to-buffer credit). In this configuration, the whole point of having two switches is to use the ISL.

For general information on host and storage layout in all SANs, refer to "Host and storage layout" on page 43.

Switch and fabric management

Specific information for two switch fabrics follows.

- All management applications can be used to monitor this environment.

For general information on switch and fabric management in all SANs, refer to "Switch and fabric management" on page 44.

Security

For general information on security in all SANs, refer to "Security" on page 47.

Connectrix B example

General layout

Figure 7 illustrates a two switch SAN with a D-4100B and a DS-4900B using CLI to configure the SAN.

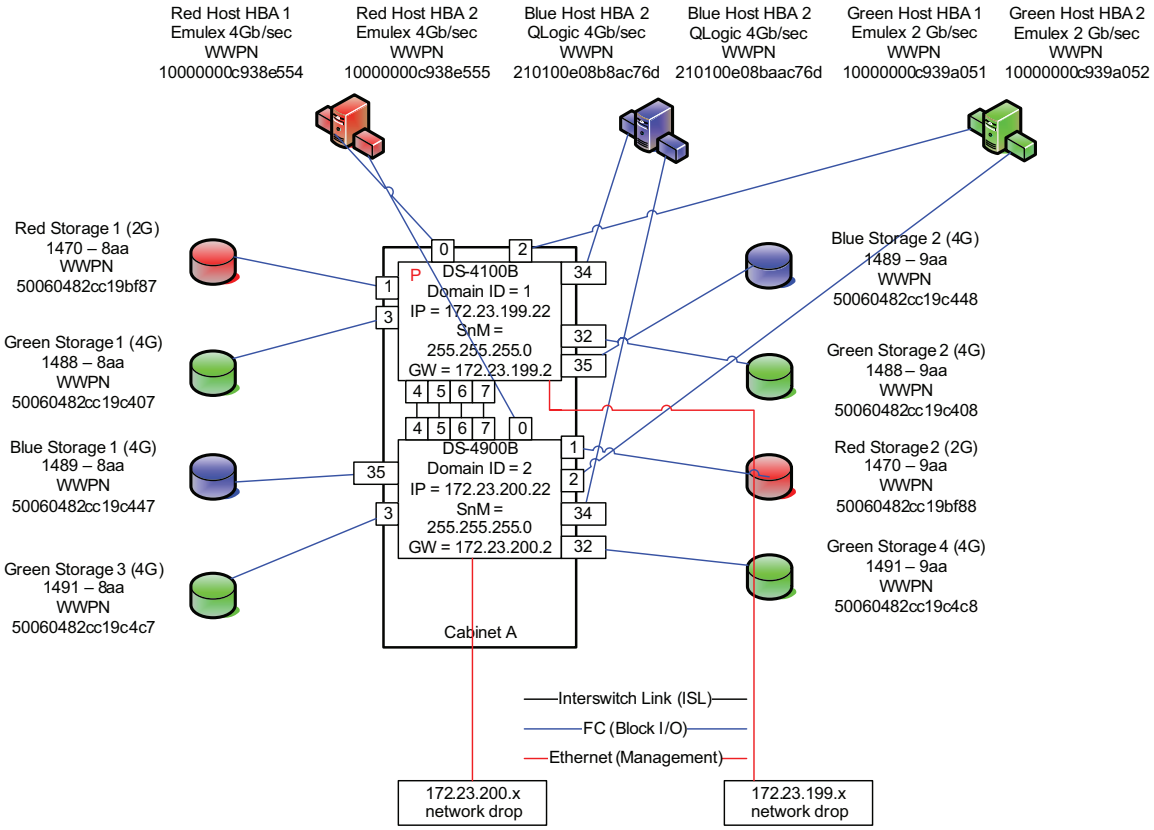


Figure 7 **Two switch SAN with DS-4100B and DS-4900B using CLI to configure SAN**

Best practices

For general information on best practices for two switch fabrics, refer to "Best practices" on page 79.

Specific information for this example follows.

Even when trunking is not being deployed immediately, it is recommended that you locate ISLs on the same quad or octet to facilitate a smooth transition to a trunking configuration at a later time.

For more Connectrix B specific best practices, refer to "Connectrix B" on page 40.

Host and storage layout

Both hosts and storage can be placed anywhere on the SAN. Connectrix B groups ASIC controlled port groups by octet on the DS-4900B. In some circumstances, it may be recommended to connect both host and storage to these groupings. This eliminates frames traveling through the backplane of the switch to reach their destination.

For general information on host and storage layout for two switch fabrics, refer to "Host and storage layout" on page 80.

Switch and fabric management

In this example, the CLI (Command Line Interface) is used to create a two switch fabric consisting of a DS-4900B and a DS-4100B. Once the two switch fabric has been connected, three hosts and their associated storage will be attached and properly configured.

For general information on switch and fabric management for two switch fabrics, refer to "Switch and fabric management" on page 80.

Security

For general information on security for two switch fabrics, refer to "Security" on page 80. For more information on security, refer to Chapter 4, "Secure SANs."

Setting up this topology

Assumptions specific to this case study:

- The Fibre Channel switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- Redundant power sources are available.
  - For switch power requirements, refer to Appendix C, "Non-EMC Products Data."
- A laptop is available with the following specifications:
  - Running some version of Windows
  - HyperTerminal is installed
  - Laptop serial ports are DB-9 connections and COM1 will be used to configure the switch IP addresses.
- A serial cable (straight through) and Ethernet cable (crossover) are available.
- Network drops, IP addresses, subnet mask and gateway, have been provided by the customer.

- SFP transceivers and compatible fiber cables are available as required.
- Access to an FTP server, for backing up (uploading) or downloading the switch configuration is available.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- Trunking licenses have been purchased by the Customer for both switches and are available for installation.

Configure the IP address

To configure the IP address:

Note: The DS-4900B uses an RJ-45 connector for the serial port instead of a DB-9 connector. Because of this, an RJ-45 serial cable (10-ft (3 m) long) is shipped with switch in addition to an RJ-45 to DB-9 adaptor. Use these to configure the IP address on the DS-4900B. The DS-4100B provides an older style DB-9 connector for the serial port interface.

---

1. Attach the straight through (DB-9 to DB-9) serial cable between the serial port on the DS-4100B and the RS-232 serial port on the management PC (COM1). The serial cable is wired with only pins 2, 3, and 5 wired straight through.
2. Power up the switch by connecting the power cords to the power receptacles provided by the customer. Make sure that all power cords are connected to the switch for maximum redundancy.

   Note: Both power supplies must be connected in order to bring the switch online.

   ---

3. Configure HyperTerminal:
   a. Open HyperTerminal by clicking on the **Start / Programs / Accessories / Communications / HyperTerminal.**

      The **Connection Description** dialog box displays

   b. Type a descriptive director name in the **Name** field and click **OK**.

      The **Connect To** dialog box appears.

c. Ensure the **Connect using** field displays **COM1** and click **OK**.

The **COM1 Properties** dialog box appears.

d. Ensure the following port settings parameters have been properly selected:

- Bits per second:9600
- Databits: 8
- Parity: None
- Stop bits: 1
- Flow control: None

e. Press **Return** to get a prompt.

4. Log in to the switch using the default values: Username: admin and Password: *password.*

IMPORTANT

It is strongly recommended that when prompted to change the password, you change it to the password that was provided by the customer. This can also be done using the passwd command from the command prompt at any time.

5. At the prompt, enter **ipaddrset** and press **Return**.

6. When prompted, supply the IP address (**172.23.199.22**), Subnet Mask (**255.255.255.0**) and Gateway address (**172.23.199.2**). Enter **Y** for all other values to leave them at their defaults.

Note: The Fibre Channel addresses and DHCP are not used for this example.

```
switch:admin> ipaddrset
Ethernet IP Address [10.77.77.77]:10.32.53.47
Ethernet Subnetmask [255.0.0.0]:255.255.240.0
Fibre Channel IP Address [0.0.0.0]:
Fibre Channel Subnetmask [0.0.0.0]:
Gateway IP Address [0.0.0.0]:10.32.48.1
IP address is being changed...Done.
Committing configuration...Done.
```

7. Verify IP address change using **ipaddrshow**.

```
switch:admin> ipaddrshow
SWITCH
Ethernet IP Address: 10.32.53.47
Ethernet Subnetmask: 255.255.240.0
Fibre Channel IP Address: none
```

```
Fibre Channel Subnetmask: none
Gateway IP Address: 10.32.48.1
DHCP: Off
```

8. Power down the switch and disconnect the serial cable.
9. Connect the switch to the 10/100BaseT Ethernet connection for the 172.23.199.x network which was provided by the Customer.
10. Power up the switch.

    The switch can now be accessed via an IP-based management tool.

11. Repeat Step 1 through Step 10 on the DS-4900B using an IP address of **172.23.200.22**, Subnet mask of **255.255.255.0**, and a Gateway of **172.23.200.2**.

Note: The RJ-45 serial cable and DB-9 adapter will need to be used to configure the switch IP address. Connect the DB-9 / RJ-45 adapter to COM1 and the RJ-45 serial cable between the adapter and the switch.

Configure FC switches

To configure FC switches:

1. Set the switch name for the DS-4100B.

   a. From the switch prompt, enter **switchname DS-4100B**.

Note: The following configurations need to be done with the switch *disabled*.

---

2. Configure the fabric parameters.

   a. From the switch prompt, enter **switchdisable** to disable the switch.

   b. From the switch prompt, enter **configure** to enter the configuration parameter menu.

   c. Enter **Y** at the **Fabric Parameters** prompt.

   d. Enter **1** for desired Domain ID at the Domain prompt and press **Enter**.

   e. The R_A_TOV should be automatically set to **10000**. If it is not, enter **10000** at the prompt and press **Enter**.

f. The E_D_TOV should be automatically set to **2000**. If it is not, enter **2000** at the prompt and press **Enter**.

g. Accept the following defaults for the rest of the fields under the **Fabric Parameters** menu by pressing **Enter** after each prompt:

- WAN_TOV = 0
- MAX_HOPS = 7
- Data field size = 2112
- Sequence Level Switching = 0
- Disable Device Probing = 0
- Suppress Class F Traffic = 0
- Switch PID Format = 1
- Per-frame Route Priority = 0
- Long Distance Fabric = 0
- BB_Credit = 16

Note: For this case study, there is no long distance between the DS-4900B switches. The ISLs connecting the two are less than 10 km.

h. At the **Insistent Domain ID Mode** prompt, enter **y** to accept the **Insistent Domain ID** setting.

Note: When this mode is set, the switch attempts to acquire the domain number programmed in its **Switch Fabric Settings** from the fabric.

i. Accept the default values from the remaining **Fabric Parameter Menu** items by pressing **Enter** after each prompt:

- Virtual Channel parameters (yes, y, no, n): [**no**]
- F_Port login parameters (yes, y, no, n): [**no**]
- Zoning Operation parameters (yes, y, no, n): [**no**]
- RSCN Transmission Mode (yes, y, no, n): [**no**]
- Arbitrated Loop parameters (yes, y, no, n): [**no**]
- System services (yes, y, no, n): [**no**]
- Portlog events enable (yes, y, no, n): [**no**]
- ssl attributes (yes, y, no, n): [**no**]
- http attributes (yes, y, no, n): [**no**]
- snmp attributes (yes, y, no, n): [**no**]
- rpcd attributes (yes, y, no, n): [**no**]
- cfgload attributes (yes, y, no, n): [**no**]
- web tools attributes (yes, y, no, n): [**no**]

3. Install the necessary licenses.

   Note: In this case, the only license to be installed will be the trunking license.

   a. From the switch prompt, enter **licenseadd** *<license for this switch>*.

      Example:

      ```
      switch:admin> licenseadd byeSRbdzyQkzfTS0
      ```

      Note: To enable trunking on the switch after unlocking the license, you need to reinitialize the ports. To reinitialize the ports, you can either disable, and then reenable, the switch or disable, and then reenable, the affected ISL ports. Since the switch is already disabled and will be re-enabled in the next step, the explicit disabling and re-enabling of the switch and / or switch ports will not be necessary.

4. From the switch prompt, enter **switchenable** to enable the switch.

5. Repeat Step 1 through Step 4 for the DS-4900B using the switch name of DS-4900B and **2** for the desired Domain ID.

Verify the firmware version loaded on the swtiches

Run the **firmwareshow** command to verify the firmware version.

```
switch:admin> firmwareshow
Primary partition:   v5.2.1a
Secondary Partition: v5.2.1a
switch:admin>
```

Verify port settings

Use command **PortCfgShow** to show current configuration of a port. All ports should be set to Auto-negotiate speed and the port type should not be locked to either L_Port or G_Port. See Figure 8 on page 88.

```
Ports of Slot 0    0  1  2  3    4  5  6  7    8  9 10 11   12 13 14 15
-----------------+--+--+--+--+----+--+--+--+----+--+--+--+----+--+--+--
Speed              AN AN AN AN   AN AN AN AN   AN AN AN AN   AN AN AN AN
Trunk Port         ON ON ON ON   ON ON ON ON   ON ON ON ON   ON ON ON ON
Long Distance      .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
VC Link Init       .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
Locked L_Port      .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
Locked G_Port      .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
Disabled E_Port    .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
ISL R_RDY Mode     .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
RSCN Suppressed    .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
Persistent Disable.. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
NPIV capability    ON ON ON ON   ON ON ON ON   ON ON ON ON   ON ON ON ON
Mirror Port        .. .. .. ..   .. .. .. ..   .. .. .. ..   .. .. .. ..
```

Figure 8 **Port settings**

Set the swtich date and time

Note: To assist with the review of support logs should the need arise, it is recommended that you sync up switch time to real time, ideally via an NTP server. You can synchronize the local time of the principal or primary fabric configuration server (FCS) switch to that of an external Network Time Protocol (NTP) server. In this example, the date and time will be set manually.

To set the date and time of a switch manually:

1. Via telnet, log in to the switch as admin.
2. Enter the date command at the command line using the following syntax:

   `date "MMDDhhmm[CC]YY"`

   where:

   - MM is the month (01-12)
   - DD is the date (01-31)
   - hh is the hour (00-23)
   - mm is minutes (00-59)
   - CC is the century (19-20)
   - YY is the year (00-99)

   Note: Year values greater than 69 are interpreted as 1970-1999; year values less than 70 are interpreted as 2000-2069. The date function does not support Daylight Savings Time or time zones, so changes will have to be reset manually.

Example:

```
switch:admin> date
Fri May 5 21:50:00 UTC 1989
switch:admin>
switch:admin> date "0624165203"
Tue Jun 24 16:52:30 UTC 2003
switch:admin>
```

Install SFP transceivers and connect cables

To install SFP transceivers and connect the cables:

1. Install the SFP transceivers in the Fibre Channel ports.

   Note: The transceivers are keyed to ensure correct orientation. If a transceiver does not install easily, ensure that it is correctly oriented.

2. Starting with the ISLs, connect the fiber cables one at a time, verfying the login status of each as they are attached. If a connectivity issue is encountered during this phase, it is easier to troubleshoot it now rather than after all cables have been attached.

   a. Connect ISLs between ports 4-7 as shown in Figure 7 on page 81.

3. Connect host and storage ports.

   a. Attach fiber cable between switches and N_Ports, as shown in Figure 7 on page 81.

   b. Verify port login status using the **switchshow** and **nsshow** commands.

Zone hosts and storage

To zone hosts and storage:

1. Create zones using the **zonecreate** commands below:

```
zonecreate "RedHBA1_1470_8aa", "10000000c938e554; 50060482cc19bf87"
zonecreate "RedHBA2_1470_9aa", "10000000c938e555; 50060482cc19bf88"
zonecreate "BlueHBA1_1489_8aa", "210100e08b8ac76d; 50060482cc19c447"
zonecreate "BlueHBA2_1489_9aa", "210100e08baac76d; 50060482cc19c448"
zonecreate  "GreenHBA1_AllGreenStorage", "10000000c939a051; 50060482cc19c407;
   50060482cc19c408; 50060482cc19c4c7; 50060482cc19c4c8"
zonecreate "GreenHBA2_AllGreenStorage", "10000000c939a052; 50060482cc19c407;
   50060482cc19c408; 50060482cc19c4c7; 50060482cc19c4c8"
```

2. Create the configuration by using the **cfgcreate** command.

```
cfgcreate "Oct_31_06_1140" , "RedHBA1_1470_8aa; RedHBA2_1470_9aa;
   BlueHBA1_1489_8aa; BlueHBA2_1489_9aa; GreenHBA1_AllGreenStorage;
   GreenHBA2_AllGreenStorage"
```

3. Enable the configuration by using the **cfgenable** command.

```
cfgenable "Oct_31_06_1140"
```

4. Enter **Y** at the confirmation prompt.
5. Enter **cfgshow** to display zoning info.

When completed, the zone information should be similar to what is shown below.

Defined configuration:

```
cfg: Oct_31_06_1140
   RedHBA1_1470_8aa; RedHBA2_1470_9aa; BlueHBA1_1489_8aa; BlueHBA2_1489_9aa;
   GreenHBA1_AllGreenStorage; GreenHBA2_AllGreenStorage"
zone: RedHBA1_1470_8aa
      10000000c938e554; 50060482cc19bf87
zone: RedHBA2_1470_9aa
      10000000c938e555; 50060482cc19bf88
zone: BlueHBA1_1489_8aa
      210100e08b8ac76d; 50060482cc19c447
zone: BlueHBA2_1489_9aa
      210100e08baac76d; 50060482cc19c448
zone: GreenHBA1_AllGreenStorage
      10000000c939a051; 50060482cc19c407;
      50060482cc19c408; 50060482cc19c4c7;
      50060482cc19c4c8
zone: GreenHBA2_AllGreenStorage
      10000000c939a052; 50060482cc19c407;
      50060482cc19c408; 50060482cc19c4c7;
      50060482cc19c4c8
```

Effective configuration:

```
CFG: Oct_31_06_1140
Zone: RedHBA1_1470_8aa
      10000000c938e554
      50060482cc19bf87
Zone: RedHBA2_1470_9aa
      10000000c938e555
      50060482cc19bf88
Zone: BlueHBA1_1489_8aa
      210100e08b8ac76d
      50060482cc19c447
Zone: BlueHBA2_1489_9aa
      210100e08baac76d
      50060482cc19c448
Zone: GreenHBA1_AllGreenStorage
```

```
      10000000c939a051
      50060482cc19c407
      50060482cc19c408
      50060482cc19c4c7
      50060482cc19c4c8
Zone name = "GreenHBA2_AllGreenStorage
      10000000c939a052
      50060482cc19c407
      50060482cc19c408
      50060482cc19c4c7
      50060482cc19c4c8
```

Save configuration

In case the configuration is lost, or unintentional changes are made, keep a backup copy of the configuration file on a host computer.

To upload a configuration file:

1. Verify that the FTP service is running on the host computer. The host must have an FTP server application running.
2. Connect to the switch through the Telnet and log in as admin.
3. Enter the **configUpload** command.

   The command becomes interactive and you are prompted for the required information.

   Example:

```
switch:admin> configupload
Protocol (scp or ftp) [ftp]: ftp
Server Name or IP Address [host]: 192.1.2.3
User Name [user]: JohnDoe
File Name [config.txt]: /pub/configurations/config.txt
Password: xxxxx
configUpload complete: All config parameters are uploaded.
switch:admin>
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps,such as configuring LUN Masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the OS configuration guide for more detail.

Connectrix MDS example

Note: VSANs will be configured and used in this example. Refer to Appendix B, "EMC Connectrix Products Data" for more information on the Connectrix MDS VSAN feature.

General layout

Figure 9 shows two MDS 9506s using VSANs.

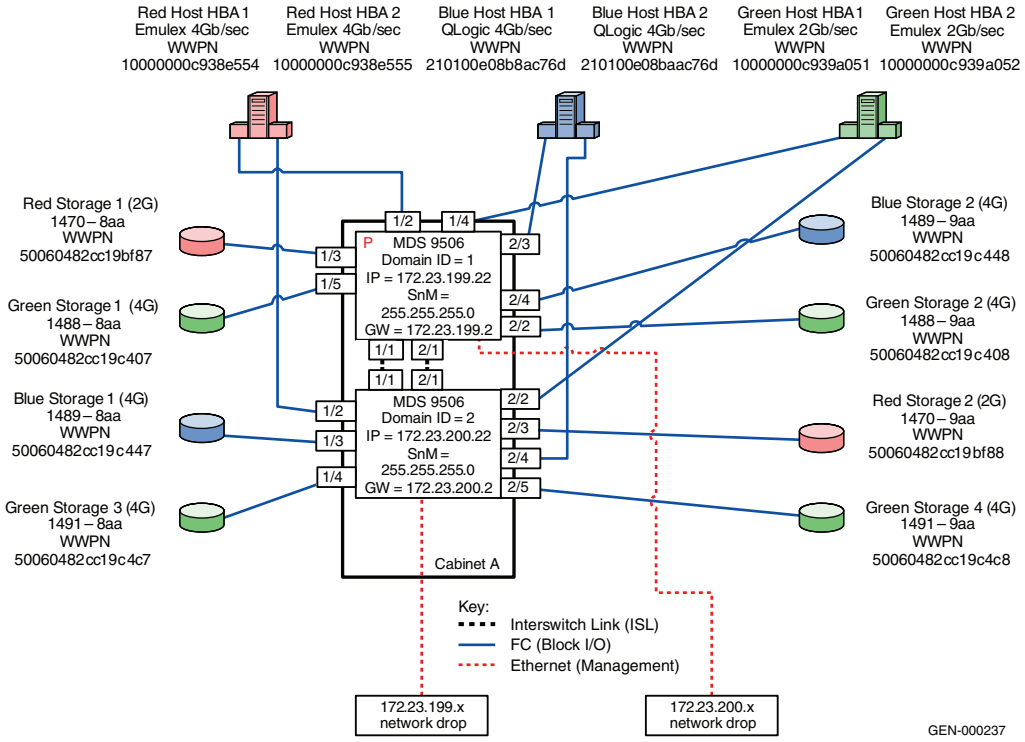


Figure 9 **Two Connectrix MDS 9506s using VSANs**

Best practices

For general information on best practices for two switch fabrics, refer to "Best practices" on page 79. Specific information for this example follows.

By default thresholds are set to 80% utilization.

Host and storage layout

For general information on host and storage layout for two switch fabrics, refer to "Host and storage layout" on page 80. Specific information for this example follows.

Line Rate Mode cards have no special restrictions. Over-subscribed cards should be used for hosts only.

Switch and fabric management

For general information on switch and fabric management for two switch fabrics, refer to "Switch and fabric management" on page 80. Specific information for this example follows.

Cisco Fabric Manager may be used for management.

Security

For general information on security for two switch fabrics, refer to "Security" on page 80. Specific information for this example follows.

Use switch binding and port binding for security.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix C, "Non-EMC Products Data."
  - For Cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The switches have *not* been connected to the power source and are *not* powered on.
- Network drops, IP addresses, Subnet mask and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix B, "EMC Connectrix Products Data".

Note: Connectrix MDS switches can be placed on either a public or private network. There are advantages to both configurations.For more information, refer to "Public versus private" on page 961.

In this example, it is assumed that The customer has provided two Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network.

- The correct number of line cards have been installed into each chassis. In this case, two line cards in each chassis are required and installed in slots 1 and 2.
  - For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- Use the laptop supplied by the installer to configure the IP addresses of the switches; this laptop has a serial DB-9 connector.
- Use the temporary password provided by the customer as the default password when configuring the IP address.
- Use Fabric Manager for VSAN setup.

Configure the IP address

To configure the IP address:

1. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer.
2. Select one of the switches to configure and set the IP to 172.23.199.22.
3. Supply a network connection to the appropriate subnet.
4. Connect to the serial port of the switch, using an RS232 serial cable, with a baud rate of 9600, 8 data bits, no parity, 1 stop bit and no flow control.

   The **login** prompt should display.
5. Log in the first time with user name *admin* and password *admin*.

   You should be prompted to supply a new strong password for CLI user admin.
6. For this example, select **no** when asked if you want to run setup.

   Note: This example will start with the switch that will have a Domain ID of **1** and an IP address of **172.23.199.22**.
7. Repeat above steps for each switch, supplying the appropriate IP.

CLI commands to configure the IP and gateway

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***interface mgmt 0***
  Switch(config-if)#***ip address 172.23.199.22 255.255.255.0***
  End with **CNTL/Z**.

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***ip default-gateway 172.23.199.2***
  End with **CNTL/Z**.

To authorize access on a switch for Device and Fabric Manager, run this command on every switch while supplying a user name (nnn) and password (ppp):

- Switch#***conf t***

  Switch(config)# ***snmp-server user nnn network-admin auth md5 ppp***
  Switch(config)#***end***
  Switch# ***copy running-config startup-config***
  Switch# ***exit***

Installing Fabric Manager and Device Manager

To install Fabric Manager and Device Manager:

1. Open your web browser.
2. Enter the IP address of the switch in the address bar.
3. Follow the prompts and accept all defaults to install both Fabric Manager and Device Manager.

Fabric Manager and Device Manager can be started using the configured snmp-server user name and password provided in "CLI commands to configure the IP and gateway" on page 162.

Creating a VSAN

To create a VSAN:

1. Open the Device Manager for the switch with an IP address of **172.23.199.22**.
2. Select **FC** from top toolbar.

3. Select **VSAN**.
4. Select **Create VSAN**.
5. Enter the value of **100** into the **VSAN ID** field.
6. Set the **VSAN Name** to be **"Red_VSAN_100"**.
7. Use the default interop mode.
8. Click **Create**.
9. Configure the other VSAN in this physical switch:

Example 1: IP Address 172.23.199.22

For the switch with the IP address **172.23.199.22:**

a. Repeat Step 2 through Step 8 for VSAN 200 and 300.

– For Virtual switch 200, use a VSAN name of **"Green_VSAN_200"**.

– For Virtual switch 300, use a VSAN name of **"Blue_VSAN_300"**.

b. Following the tables below, assign and enable the ports to the proper VSAN using Device Manager.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to SW 2	1
1	2	Red Host HBA 1	100
1	3	Red Storage 1	100
1	4	Green Host HBA 1	200
1	5	Green Storage 1	200
1	6	Blue Host HBA 1	300
1	7		
1	8		
2	1	TE ISL to SW 2	1
2	2	Green Storage 2	200
2	3	Blue Host HBA 1	300
2	4	Blue Storage 2	300
2	5		
2	6		
2	7		
2	8		

Example 2: IP Address 172.23.200.22

For the switch with the IP address **172.23.200.22:**

c. Repeat Step 2 through Step 8 for VSAN 100, 200 and 300 on this switch.

– For VSAN 100 use a Name **"Red_VSAN_100"**.

– For VSAN 200 use a VS Name **"Green_VSAN_200"**.

– For VSAN 300, use a VS Name **"Blue_VSAN_300"**.

d. Assign the ports to the proper VSAN using Device Manager, using the following table:.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to SW 1	1
1	2	Red Host HBA 2	100
1	3	Blue Storage 1	300
1	4	Green Storage 1	200
1	5		
1	6		
1	7		
1	8		
2	1	TE ISL to SW 1	1
2	2	Green HBA 2	200
2	3	Red Storage 2	100
2	4	Blue Host HBA 2	300
2	5	Green Storage 4	200
2	6		
2	7		
2	8		

Connecting cables

To connect the cables:

1. Connect ISLs.

   a. Attach Fiber cable between switches as shown in Figure 9 on page 92.

   b. After all cables are connected, use Fabric Manager to verify that all ISL connections are up.

   c. Re-arrange icons to accurately reflect the switch configuration.

2. Connect host and storage ports.
   a. Attach fiber cable between switches and N_Ports.

Configure domains

To configure domains:

1. Open **Fabric Manager**. It should show a topology of two switches.
2. Open the **"Red_VSAN_100"** folder from **Fabric Manager** open the **"Red_VSAN_100"** folder.
3. Select **Domain Manager**.
4. Select the **Configuration** menu.
5. Set a Domain ID. **1** for switch **172.23.199.22** and **2** for **172.23.200.22**.
6. Set the priority to **1** in the domain menu, to set a principal switch.
7. Repeat Step 2 through Step 6 for **"Green_VSAN_200"** and "**Blue_VSAN_300"**.

Zone hosts and storage

To zone hosts and storage:

1. From **Fabric Manager** select **"Red Vsan 100"**.
2. Select **Edit Full Zone Database**.
3. Create a zone by clicking **Zone** button under the **Zones Tree**.
4. Provide a descriptive name for the zone. This example will zone "Red host HBA 1" and "Red Storage 1", so **"RedHBA1_1470_8aa"** will be used. **Press Enter**.
5. Locate, then single click, **"Red Host HBA 1"** (WWPN 10000000c938e554) in the **Potential zone members list**.
6. Click the right-pointing arrow on the divider between the **Potential members list** and the zones list to add the HBA to the zone.
7. Locate, then single click, **"Red Storage 1"** (WWPN 50060482cc19bf87) in the **Potential zone members list**.
8. Click the right-pointing arrow on the divider between the **Potential members list** and the **zones list** to add the Storage port to the zone.

9. Repeat Step 2 through Step 7 for all host and storage pairs in the environment.
10. Create a zone set by clicking **New Set** under the **Zone sets Tree**.
11. Provide a descriptive name for the zone set. This example uses the name **"RED Fabric 1"**.

Add only those zones that will be necessary on Red_Fabric_1. In this case only the zone named "RedHBA1_1470_8aa" should be added to the Red_Fabric_1 zone set. Repeat for other Fabrics. When completed, you should have 5 zone sets, as shown below.

```
Zone set name = "Red_Fabric_1"

          Zone name = "RedHBA1_1470_8aa"
             Zone Member = "10000000c938e554"
             Zone Member = "50060482cc19bf87"


          Zone name = "RedHBA2_1470_9aa"
             Zone Member = "10000000c938e555"
             Zone Member = "50060482cc19bf88"

Zone set name = "Blue_Fabric_1"

          Zone name = "BlueHBA1_1489_8aa"
             Zone Member = "210100e08b8ac76d"
             Zone Member = "50060482cc19c447"


          Zone name = "BlueHBA2_1489_9aa"
             Zone Member = "210100e08baac76d"
             Zone Member = "50060482cc19c448"

Zone set name = "Green_Fabric"

          Zone name = "GreenHBA1_AllGreenStorage"
             Zone Member = "10000000c939a051"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"

          Zone name = "GreenHBA2_AllGreenStorage"
             Zone Member = "10000000c939a052"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"
```

Optional: Configure IVR (Inter-VSAN Routing)

The configuration of VSANs on a fabric allows for security, scalability and availability. However, this isolation of traffic between VSANs prevents users from accessing resources, such as tape libraries, located in other VSANs. The solution to this limitation is Cisco's Inter-VSAN Routing feature, which allows initiators in one VSAN to access targets in other VSANs without merging the VSANs. Perhaps a host in the Red VSAN needs to access storage in the Blue VSAN. Configuring IVR zones and IVR zone sets containing the allowed initiators and targets allows communication between these resources.

1. In Fabric Manager:
   a. Click the **Zone** tab in upper tool bar.
   b. Click the **IVR** tab.
   c. Select **Wizard**.
2. Select the VSANs that will participate in IVR in the fabric.

   Select **VSAN 100**, **200** and **300**.
3. Select the end devices that you want to communicate over IVR.

   Select the following:

   VSAN 100: 10000000c938e554
   VSAN 200: 50060482cc19c407
   VSAN300: 50060482cc19c447
   VSAN 300: 210100e08b8ac76d
4. Enter the VSAN ID of the VSAN you want to use as the transit VSAN between the VSANs selected for the IVR zone.

   Select **VSAN 1** as the transit VSAN.

   Note: VSAN 1 connects both switches and all trunking traffic will pass over this link to communitcate with VSANs in other switches.
5. Set the IVR zone and IVR zone set.

   IVR NAME = **IVRZONE1**

   IVR ZONENET NAME = **IVRZONESET1**
6. Verify all steps that Fabric Manager will take to configure IVR in the fabric.

7. Click **Finish** if you want to enable IVR NAT and IVR topology and to create the associated IVR zones and IVR zone set.

   or

   Click **Cancel** to exit the IVR Wizard without saving any changes.

8. The **Save Configuration** dialog displays. You can save the configuration of the master witch to be copied to other IVR-enabled switches.

   Click either **Continue Activation** or **Cancel**.

9. Click **Finish**.

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the Fabric OS configuration guide for more detail.

Connectrix M example

General layout

Figure 10 shows an example of two ED-10000Ms (Brocade M-Series Intrepid 10000) using virtual switches.

Note: Virtual switches are configured and used in the following examples. Refer to the Appendix B, "EMC Connectrix Products Data" for more information on the Connectrix M virtual switch feature.

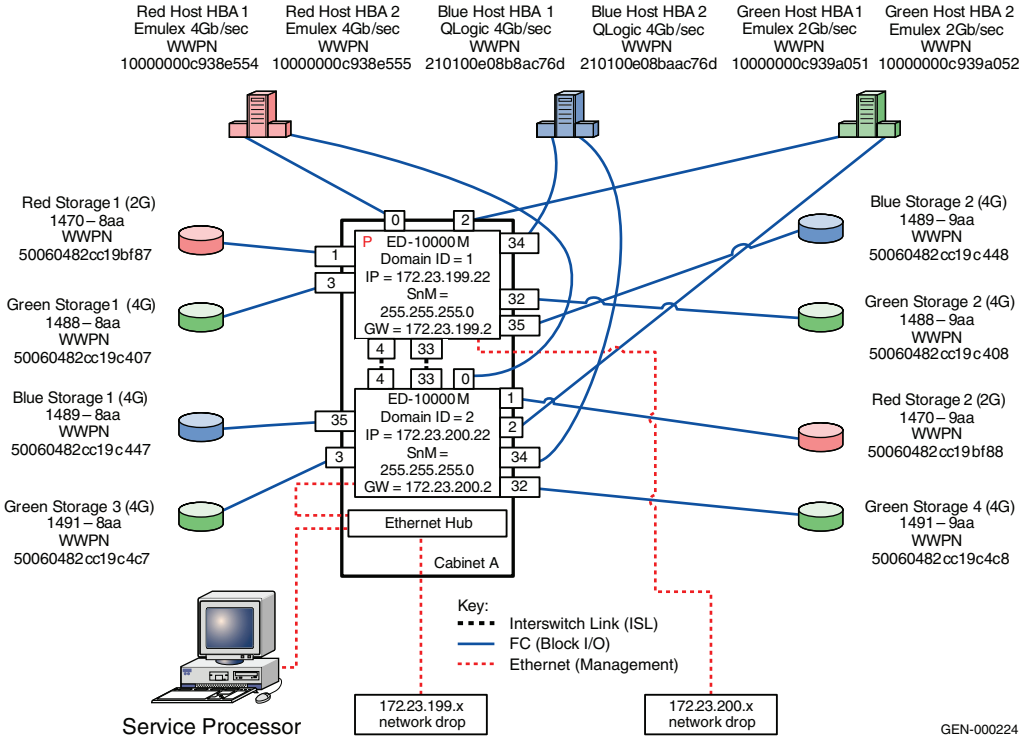


Figure 10 **Two ED-10000Ms using virtual switches**

Although all Connectrix M switch types can be used in a two switch fabric configuration, only the ED-10000M supports virtual switches.

Best practices

Specific information for this example follows.

- Configure threshold alerts to monitor for utilization in excess of 80%.

- Configure counter threshold alerts to monitor for invalid transmission words, encoding errors, and CRC errors. Any of these can indicate a problem with the physical link that may eventually lead to bleeding BB_Credit. If any ISLs will be used for long distance, monitoring for errors is a necessity.
- Configure a fencing policy.

For general information on best practices for two switch fabrics, refer to "Best practices" on page 79.

Host and storage layout

ED-10000Ms are used in this configuration. Due to the small number of host and storage pairs as well as ISLs in this configuration, special considerations need not be made to ensure fairness. However, when a larger number of ports are present, host and storage layout can become important. Refer to ED-10000M fairness considerations in Appendix B, "EMC Connectrix Products Data".

For general information on host and storage layout for two switch fabrics, refer to "Host and storage layout" on page 80.

Switch and fabric management

Connectrix Manager is used to configure this environment.

For general information on switch and fabric management for two switch fabrics, refer to "Switch and fabric management" on page 80.

Security

In "Configure security" on page 115 default passwords are changed, Enterprise Fabric mode is enabled, and a Port binding configuration will be put into effect.

For general information on security for two switch fabrics, refer to "Security" on page 80. For further information on security, refer to Chapter 4, "Secure SANs."

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data".

- For Cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.

- The switches have *not* been connected to the power source and are *not* powered on.
- Network drops, IP addresses, Subnet mask, and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix C, "Non-EMC Products Data."

Note: Connectrix M switches can be either attached to the Ethernet hub that comes with the cabinet, or directly connected to the customer's LAN. In both cases, the switches can be placed on either a public or private network. There are advantages to both configurations. For more information, refer to "Public versus private" on page 961.

This example assumes that The customer has provided us with 2 Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network.

- The correct number of LMQs have been in installed in each chassis. In this case, two LMQs in each chassis are required.
  - For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- A Service Processor is installed in at least one of the cabinets.
- A laptop supplied by the installer will be used to configure the IP addressees of the switches. This laptop has a serial DB-9 connector.
- A null modem cable is available.
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- Use Connectrix Manager. The Connectrix Manager installation kit is available.
- Connectrix Manager is not already installed on the Service Processor.

Configure the IP address

To configure the IP address:

1. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer. This causes both the directors and the Service Processor to power up.
2. Select one of the switches to configure.

Note: This example will start with the switch with a Domain ID of 1 and an IP address of **172.23.199.22**.

Change a product IP address, subnet mask, or gateway address

Perform the following steps to change a product IP address, subnet mask, or gateway address. An asynchronous RS-232 modem cable and maintenance terminal (desktop or notebook PC) with a Windows-based operating system and RS-232 serial communication software (such as ProComm Plus or HyperTerminal) are required.

1. Connect one end of the RS-232 modem cable to the 9-pin maintenance port on the active CTP card. The active CTP card is identified by the green Role LED.
2. Connect the other cable end to a 9-pin serial communication port (**COM1** or **COM2**) at the rear of the maintenance terminal PC.
3. Power on the maintenance terminal. At the Windows desktop, click **Start** at the left side of the task bar. The **Windows Workstation** menu appears.

4. From the **Windows Workstation** menu, sequentially select the **Programs, Accessories, Communications**, and **HyperTerminal** options. The **Connection Description** dialog box displays (Figure 11).

Figure 11 **Connection Description dialog box**

5. Type a descriptive director name in the **Name** field and click **OK**.

   The **Connect To** dialog box appears.

6. Ensure the **Connect using** field appears COM1 or COM2 (depending on the port connection to the director), and click **OK**.

   The **COMn Properties** dialog box appears, where n is 1 or 2.

7. Configure Port Settings parameters:

   - Bits per second—**115200**.
   - Data bits—**8**.
   - Parity—**None**.
   - Stop bits—**1**.
   - Flow control—**Hardware**.

   Click **OK**. The **New Connection - HyperTerminal** window displays.

8. At the login prompt type **Administrator** with a password of *password*.

9. At the EOS# prompt, type the system ip command and press **Enter**. The **New Connection - HyperTerminal** window appears with configuration information listed:
   - *IP address* (default is **10.1.1.10**).

     Note: IP address is the preinstalled IP address but may not be correct for the director.

   - *Subnet mask* (default is **250.0.0.0**).
   - *Gateway address (*default is **0.0.0.0***)*
10. Change the IP address, subnet mask, and gateway address as directed by the customer. To change the addresses, type the following and press **Enter.**

    system ip 172.23.199.22 255.255.255.0 172.23.199.2

    The IP address is xxx.xxx.xxx.xxx, the subnet mask is yyy.yyy.yyy.yyy, and the gateway address is zzz.zzz.zzz.zzz, where the octets xxx, yyy, and zzz are decimals from zero through 255. If an address is to remain unchanged, type the current address in the respective field.

To configure the switch IP address of the other switch in the cabinet, follow Step 3 through Step 10, above and use the IP address of **172.23.200.22**.

Install Connectrix Manager

Note: Refer to the *Connectrix Manager User Guide* for complete installation instructions.

To install Connectrix Manager:

1. Insert the Connectrix Manager installation CD into the CD-RW drive.
2. Browse to the CD-RW drive in windows explorer.
3. Locate, then double-click, the **install.exe** file. The installer is displayed. Typically, this file is located under the **CtxMgr <***version number***>** folder.
4. Accept all the defaults by clicking **Next** and **OK**, when appropriate.
5. Click **Done** when the **Installation Complete** dialog appears.

6. Click **Next** when the **Welcome** dialog appears.
7. Accept the EULA terms and click **Next**.
8. At the **Copy data** and **Settings** dialog, select **No,** then **Next**.

   Note: "**No**" is selected in this step since this section assumes this is a new installation.

9. Assign a Connectrix Manager 9.0 server name, and then click **Next**.

   This case uses **CMServer.**

10. Enter the serial number and license key in the provided fields in the **Connectrix Manager 9.0 Server License** dialog.
    - The serial number is located on the back of the Connectrix Manager installation CD case.
    - To obtain a license key, locate the transaction code certificate and go to the URL listed.
      a. Enter the serial number located on the back of the Connectrix Manager installation CD case.
      b. In the **Transaction Code** fields, type the transaction code(s) shipped with the software.
      c. Click **Next**.
      d. Confirm the existing and new features to be enabled.
      e. Click **Next**.

         The license key and all enabled features appears.

      f. Retain a copy for your records.
      g. Enter this key into the **License key** field.
      h. Click **OK**.
11. Once the installation is complete, log in to Connectrix Manager with the username of *administrator* and the password of *password*.

    Note: Change this default as soon as possible. For this example, the username *nnn* and password *nnn* are used.

Manage the switches

To manage the switches:

1. From Connectrix Manager topology view, select the **Setup** from the **Discover** pull-down menu. The **Discover Setup** dialog is displayed.
2. Click the **Out-of-Band** tab.
3. Under **Available Addresses**, click **Add...** .
4. **In the IP address field,** enter **172.23.199.22**.
5. In the **Subnet Mask** field, enter **255.255.255.0**.
6. Click **OK**.
7. Click **Add**.
8. In the **IP address** field, enter **172.23.200.22**.
9. In the **Subnet Mask** field, enter **255.255.255.0**.
10. Click **OK**.
11. Ensure that the switches are highlighted in the **Available Addresses** list, and then click the bottom right-pointing arrow. The IP addresses of **172.23.199.22** and **172.23.200.22** appear in the **Selected Individual Addresses** list.
12. Click **OK**.

Configure the first virtual switch

To configure the first virtual switch:

1. Open the Element Manager for the switch with an IP address of **172.23.199.22**.
2. Open the **Virtual Switches** dialog by selecting the **virtual switches** menu item under the **Configure** pull-down menu.
3. Create a new virtual switch by clicking the **New** button.
4. Enter the value of **100** into the **VF ID** field.
5. Set the **VS Name** to be **Red Virtual Switch**.
6. Set the switch **Priority** to be **Principal**.

Note: In later steps, the red and blue virtual switches do not have any ISLs associated with them, so setting the priority to be principal is not necessary at this time. However, setting it now can prevent problems in the future.

7. Set the **Interop Mode** to be **Open Fabric 1.0**.
8. Click the **Domain** tab.
9. Set the preferred **Domain ID** to be **1**.
10. Ensure that the **Insistent** Domain ID checkbox is checked.
11. Click **OK**.

Configure other virtual switches on this physical switch

To configure virtual switch 200 and 300 on this physical switch:

Follow Step 1 through Step 11 under "Configure the first virtual switch" on page 109, and then continue with the following steps:

12. For virtual switch **200** use a **VS Name** of **Green Virtual Switch** and a **Domain ID** of **2**.

    For virtual switch **300**, use a **VS Name** of **Blue Virtual Switch** and a Domain ID of **3**.

13. Using the table below, assign the ports to the proper virtual switch by selecting the port in the **All Ports** list, selecting the appropriate virtual switch in the virtual switches list, and then clicking the right arrow button.

Slot #	SPPe #	Port #	Name	Virtual Switch ID
0	0	Port 0	Red Host HBA 1	VS 100
		Port 1	Red Storage 1	VS 100
		Port 2	Green Host HBA 1	VS 200
		Port 3	Green Storage 1	VS 200
		Port 4	ISL to Domain 2 (Port 4)	VS 200
		Port 5		
		Port 6		
		Port 7		
1	0	Port 32	Green Storage 2	VS 200
		Port 33	ISL to Domain 2 (Port 33)	VS 200
		Port 34	Blue Host HBA 1	VS 300

		Port 35	Blue Storage 1	VS 300
		Port 36		
		Port 37		
		Port 38		
		Port 39		

14. Click **OK**.

Configure virtual switches on the physical switch with an IP address of 172.23.200.22

To configure virtual switches 100, 200, and 300:

Follow Step 1 through Step 11 under "Configure the first virtual switch" on page 109, and then continue with the following steps:

12. For virtual switch **100** use a **VS Nam**e of **Red Virtual Switch** and a **Domain ID** of **11** and leave switch priority at default.

    For virtual switch **200** use a **VS Nam**e of **Green Virtual Switch** and a **Domain ID** of **12**and leave switch priority at default.

    For virtual switch **300**, use a **VS Name** of **Blue Virtual Switch** and a Domain ID of **13** and leave switch priority at default.

13. Using the table below, assign the ports to the proper virtual switch by selecting the port in the **All Ports** list, selecting the appropriate virtual switch in the virtual switches list, and then clicking the right arrow button.

Slot #	SPPe #	Port #	Name	Virtual Switch ID
0	0	Port 0	Red Host HBA 2	VS 100
		Port 1	Red Storage 2	VS 100
		Port 2	Green Host HBA 2	VS 200
		Port 3	Green Storage 3	VS 200
		Port 4	ISL to Domain 1 (Port 4)	VS 200
		Port 5		
		Port 6		
		Port 7		
1	0	Port 32	Green Storage 4	VS 200
		Port 33	ISL to Domain 1 (Port 33)	VS 200
		Port 34	Blue Host HBA 2	VS 300

		Port 35	Blue Storage 2	VS 300
		Port 36		
		Port 37		
		Port 38		
		Port 39		

14. Click **OK**.

Configure FC switches

To configure FC switches:

1. Configure the switch ports.

   a. Open the Element Manager for the core switch that has an IP address of **172.23.199.22** by double clicking its icon.

   Note: Core switches are located in the Core Switch Group. If you cannot locate the core switch, right-click a virtual switch icon and select **Core Switch**.

   b. From the **Configure** menu, select the **ports menu** item.

   c. Configure the port names as shown in the table below.

Example 1: IP Address 172.23.199.22

Slot #	SPPe #	Port #	Name	Virtual Switch ID
0	0	Port 0	Red Host HBA 1	VS 100
		Port 1	Red Storage 1	VS 100
		Port 2	Green Host HBA 1	VS 200
		Port 3	Green Storage 1	VS 200
		Port 4	ISL to Domain 2 (Port 4)	VS 200
		Port 5		
		Port 6		
		Port 7		
1	0	Port 32	Green Storage 2	VS 200
		Port 33	ISL to Domain 2 (Port 33)	VS 200
		Port 34	Blue Host HBA 1	VS 300
		Port 35	Blue Storage 1	VS 300
		Port 36		

		Port 37		
		Port 38		
		Port 39		

Example 2: IP Address 172.23.200.22

Configure the port names of the switch with an IP address of **172.23.200.22**, as shown in the table below:

Slot #	SPPe #	Port #	Name	Virtual Switch ID
0	0	Port 0	Red Host HBA 2	VS 100
		Port 1	Red Storage 2	VS 100
		Port 2	Green Host HBA 2	VS 200
		Port 3	Green Storage 3	VS 200
		Port 4	ISL to Domain 1 (Port 4)	VS 200
		Port 5		
		Port 6		
		Port 7		
1	0	Port 32	Green Storage 4	VS 200
		Port 33	ISL to Domain 1 (Port 33)	VS 200
		Port 34	Blue Host HBA 2	VS 300
		Port 35	Blue Storage 2	VS 300
		Port 36		
		Port 37		
		Port 38		
		Port 39		

Connect cables

To connect cables:

1. Connect ISLs.
   a. Attach fiber cable between switches as shown in Figure 10 on page 102.
   b. After all cables are connected, use Connectrix Manager to verify that all ISL connections are up.
   c. Rearrange icons to accurately reflect the switch configuration and then persist the fabric.

Note: After persisting the fabric, when looking at the topology view you can immediately detect if something has changed in the environment. If an ISL or device disappeared, yellow alert icons display. Because of this, it is a recommended to *always* persist the fabric *after* changes have been made.

2. Connect the host and storage ports.
   a. Attach fiber cable between switches and N_Ports.

Zone hosts and storage

To zone hosts and storage:

1. Open the **Zoning** dialog in Connectrix Manager by right-clicking the appropriate fabric topology and selecting **zoning menu**.
2. Create a zone by clicking the **New Zone** button under the zones set **Tree**.
3. Provide a descriptive name for the zone. This case zones *Red host HBA* and *Red Storage 1*, so **RedHBA1_1470_8aa** is entered. Press **Enter**.
4. Locate, then click **Red Host HBA 1** (WWPN 10000000c938e554) in the **potential zone members** list.
5. Click the right-pointing arrow on the divider between the **potential members** list and the **zones** list to add the HBA to the zone.
6. Locate, then click, **Red Storage 1** (WWPN 50060482cc19bf87) in the **potential zone members** list.
7. Click the right-pointing arrow on the divider between the **potential members** list and the **zones** list to add the Storage port to the zone.
8. Repeat steps 2 through 7 for all host and storage pairs in the environment.
9. Create a zone set by clicking **New Set** under the zone sets **Tree**.
10. Provide a descriptive name for the zone set. This case uses the date of “Oct_31_06_1140”.
11. Add only those zones that are necessary on **Red_Fabric_**1. In this case, only the zone named **RedHBA1_1470_8aa** should be added to the **Red_Fabric_1** zone set. Repeat for other fabrics. When completed, you should have 5 zone sets as shown next.

```
Zone set name = "Red_Fabric_1"

          Zone name = "RedHBA1_1470_8aa"
             Zone Member = "10000000c938e554"
             Zone Member = "50060482cc19bf87"

Zone set name = "Red_Fabric_2"

          Zone name = "RedHBA2_1470_9aa"
             Zone Member = "10000000c938e555"
             Zone Member = "50060482cc19bf88"

Zone set name = "Blue_Fabric_1"

          Zone name = "BlueHBA1_1489_8aa"
             Zone Member = "210100e08b8ac76d"
             Zone Member = "50060482cc19c447"

Zone set name = "Blue_Fabric_1"

          Zone name = "BlueHBA2_1489_9aa"
             Zone Member = "210100e08baac76d"
             Zone Member = "50060482cc19c448"

Zone set name = "Green_Fabric_1"

          Zone name = "GreenHBA1_AllGreenStorage"
             Zone Member = "10000000c939a051"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"

          Zone name = "GreenHBA2_AllBGreenStorage"
             Zone Member = "10000000c939a052"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the Fabric OS configuration guide for more detail.

Configure security

Once the hosts and storage ports have been properly configured, the fabric should be secured.

Enable Enterprise Fabric Mode:

1. From the **Connectrix Manage**r topology view, select a switch in the fabric that needs to be secured.

   Note: Do *not* double-click and open the Element Manager.

2. Select **Enterprise Fabric Mode** under the **Configure** pull-down menu.
3. Click **Activate**.
4. Click **Close**.

Disable unused interfaces:

1. From the **Connectrix Manager** topology view, select the **Security** tab.
2. Ensure that the appropriate fabric is selected in the fabrics list.
3. Under the **Authentication** tab, select a switch in the **Product Configuration** list.
4. Under the **Users** tab uncheck the **Enable Web Server** and **Enable Telnet** checkboxes.
5. Click **Apply to** and the **Apply to Other Products** dialog is appears.
6. Click **Select All**, and then **OK**.
7. Review the changes and ensure that they are correct, and then click **Start**.
8. Ensure that the **Processing complete - all changes applied successfully** message appears in the **status** field.
9. Click **Close**.

Configure proactive monitoring and countermeasures

1. Telnet into the switch with an IP address of 172.23.199.22 by entering the following command:

   telnet 172.23.199.22

2. Enter the username and password when prompted.
3. Enter the command **FC Performance**.

Configure a througphput threshold alert:

1. Enter the command:

   config throughput 1 TTA portlist 5 5 both 80 all

Configure counter threshold alerts:

1. Enter the command:

   config counter 2 CTA1 portlist 5 CountBBCreditZero 1000000 all

2. Enter the command:

   config counter 3 CTA2 portlist 5 CountClass3Discards 100 all

3. Enter the command:

   config counter 4 CTA3 portlist 5 CountInvalidTxWords 40 all

4. Enter the command:

   config counter 5 CTA4 portlist 1440 CountInvalidTxWords 100 all

Activate all of the alerts:

1. Enter the command **config activate 1**.
1. Enter the command **config activate 2**.
2. Enter the command **config activate 3**.
3. Enter the command **config activate 4**.
4. Enter the command **config activate 5**.

Configure fencing:

By default there are three fencing policies defined that need to be enabled.

1. Enter the command **fc portfencing enable "Default Security Policy"**.
2. Enter the command **fc portfencing enable "Default Link Level Policy"**.
3. Enter the command **fc portfencing enable "Default Protocol Error P"**.

Configure the other switch:

Repeat all the steps in the"Configure proactive monitoring and countermeasures" on page 116 to configure the other switch with the IP address **172.23.200.22**.

QLogic example

General layout

Figure 12 shows two SB5602 switches using Quicktools.

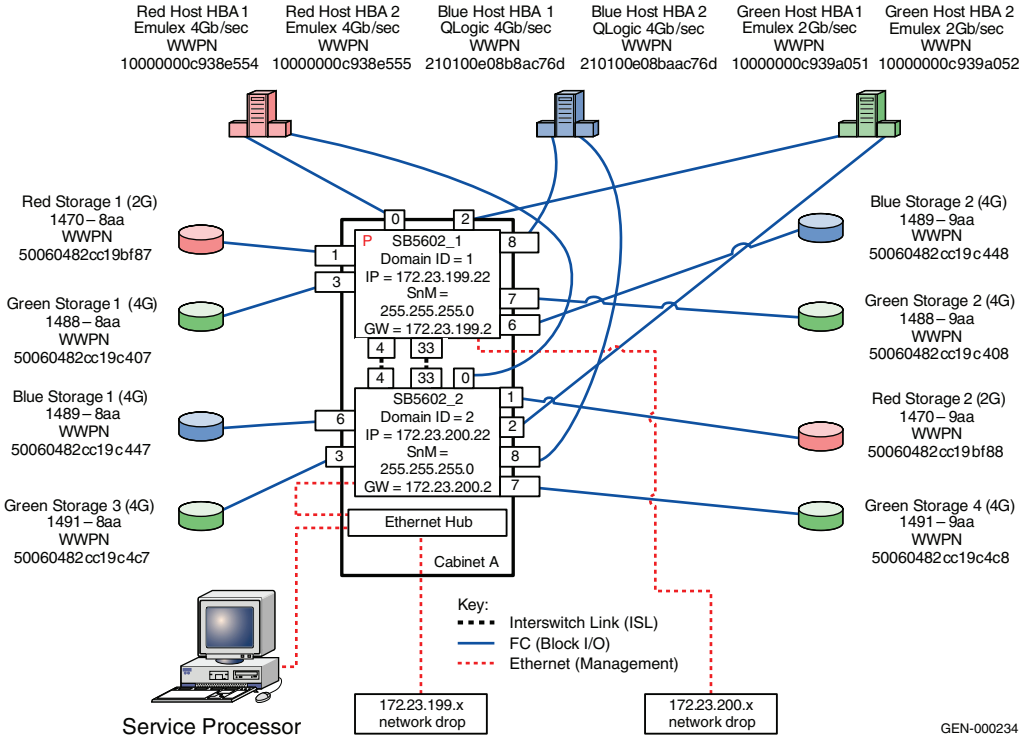


Figure 12 **Two SB5602 switches using Quicktools**

Best practices

For general information on best practices for two switch fabrics, refer to "Best practices" on page 79. For QLogic-specific best practices, refer to "QLogic" on page 42.

Host and storage layout

The SANbox 9000s are used in this configuration. Due to the small number of host and storage pairs as well as ISLs in this configuration, special considerations need not be made to ensure fairness.

For general information on host and storage layout for two switch fabrics, refer to "Host and storage layout" on page 80.

Switch and fabric management

For general information on switch and fabric management for two switch fabrics, refer to "Switch and fabric management" on page 80. For QLogic-specific information, refer to "QLogic" on page 47.

Security

Enable/disable system services

The system services dialog using Quicktools provides a central location for the user to enable or disable any of the external services such as SNMP, SSL, SSH, Embedded Web Applet, Telnet sessions, and so on. This feature can help the user select the desired switch interface, enable some security features, and disable unused interfaces.

- To display the **System Services** dialog, open the **Switch** menu and click **Services**. Check/uncheck the features displayed in the pop-up window based on user requirements.
- With CLI, the above features can be enabled using the **Set Setup Services** command. Please follow the directions displayed on running this command to enable/disable the system services.

Note: When enabling SSL, verify that the date/time settings on this switch, and the workstation from where the SSL connection will be started, match. A new SSL certificate may need to be created to ensure a secure connection to this switch.

If .all services to the switch are disabled, the connection to the switch may be lost

---

Enable fabric binding

The Fabric Binding feature can be enabled using CLI with the **Set Config Security** command. In addition, this command provides the Autosave feature, which allows the user to save any changes to the switch security settings (security sets) in the permanent memory.

To enable Fabric Binding on switches in a fabric, run the following commands at the switch command prompt:

```
SANbox #> admin start
SANbox (admin) #> config edit
The config named default is being edited.
SANbox (admin-config)#> set config security
A list of attributes with formatting and current values will follow.
Enter a new value or simply press the ENTER key to accept the current value.
If you wish to terminate this process before reaching the end of the list
press 'q' or 'Q' and the ENTER key to do so.

FabricBindingEnabled (True / False) [True]
AutoSave (True / False) [True ]

Finished configuring attributes.
This configuration must be saved (see config save command) and
```

```
activated (see config activate command) before it can take effect.
To discard this configuration use the config cancel command.

SANbox (admin-config)#> config save
SANbox (admin)#> config activate
SANbox (admin)#> admin end
```

Fabric Binding associates switch worldwide names with a Domain ID in the creation of ISL security groups. Please refer to "Section 10: Device Security Configuration" of the QLogic reference document, *SANbox 9000 Fibre Channel Switch: Command Line Interface*, to learn more about QLogic security policies and settings, located at http://download.qlogic.com/manual.

Note: The Fabric Binding feature is *not* supported on the SB9000 and has *not* been added to the Quicktools interface.

Enable port binding

The Port Binding feature can be enabled using CLI with the **Set Config Security Port [port-number]** command. In addition, this command asks for the WWN of the port/device that is allowed to connect to the port given by [port-number].

To enable the port binding feature run the following commands at the switch prompt:

```
SANbox #> admin start
SANbox (admin) #> config edit
SANbox (admin-config) #> set config security port 1
A list of attributes with formatting and current values will follow.
Enter a new value or simply press the ENTER key to accept the current value.
If you wish to terminate this process before reaching the end of the list
press 'q' or 'Q' and the ENTER key to do so.

PortBindingEnabled (True / False)[False] true
WWN (N=None / WWN)[None ] 50:06:04:82:cc:f9:bf:87 // The red storage WWN
WWN (N=None / WWN)[None ] 10:00:00:c0:dd:00:b9:f8 // replace with SB5602_1 switch WWN
WWN (N=None / WWN)[None ] n

Finished configuring attributes.
This configuration must be saved (see config save command) and
activated (see config activate command) before it can take effect.
To discard this configuration use the config cancel command.

SANbox (admin-config)#> config save
SANbox (admin)#> config activate
SANbox (admin)#> admin end
```

Set threshold alerts

The **Set Config Threshold** CLI command is used to set the port alarm threshold parameters by which the switch monitors port performance and generates alarms. The command initiates a configuration session to generate and log alarms for selected events such as CRC errors, decode errors, ISL connection, count variation, device login errors, device logout errors, and loss of signal errors.

To set the port alarm thresholds, run the following commands at the switch prompt:

```
SANbox #> admin start
SANbox (admin) #> config edit
SANbox (admin-config) #> set config threshold
A list of attributes with formatting and current values will follow.
Enter a new value or simply press the ENTER key to accept the current value.
If you wish to terminate this process before reaching the end of the list
press 'q' or 'Q' and the ENTER key to do so.

ThresholdMonitoringEnabled (True / False) [False ]
CRCErrorsMonitoringEnabled (True / False) [True ]
RisingTrigger (decimal value, 1-1000) [25 ]
FallingTrigger (decimal value, 0-1000) [1 ]
SampleWindow (decimal value, 1-1000 sec) [10 ]
DecodeErrorsMonitoringEnabled (True / False) [True ]
RisingTrigger (decimal value, 1-1000) [25 ]
FallingTrigger (decimal value, 0-1000) [0 ]
SampleWindow (decimal value, 1-1000 sec) [10 ]
ISLMonitoringEnabled (True / False) [True ]
RisingTrigger (decimal value, 1-1000) [2 ]
FallingTrigger (decimal value, 0-1000) [0 ]
SampleWindow (decimal value, 1-1000 sec) [10 ]
LoginMonitoringEnabled (True / False) [True ]
RisingTrigger (decimal value, 1-1000) [5 ]
FallingTrigger (decimal value, 0-1000) [1 ]
SampleWindow (decimal value, 1-1000 sec) [10 ]
LogoutMonitoringEnabled (True / False) [True ]
RisingTrigger (decimal value, 1-1000) [5 ]
FallingTrigger (decimal value, 0-1000) [1 ]
SampleWindow (decimal value, 1-1000 sec) [10 ]
LOSMonitoringEnabled (True / False) [True ]
RisingTrigger (decimal value, 1-1000) [100 ]
FallingTrigger (decimal value, 0-1000) [5 ]
SampleWindow (decimal value, 1-1000 sec) [10 ]

Finished configuring attributes.
This configuration must be saved (see config save command) and activated (see
config activate command) before it can take effect.
To discard this configuration use the config cancel command.
```

```
SANbox (admin-config)#> config save
SANbox (admin)#> config activate
SANbox (admin)#> admin end
```

Enable RADIUS

The **Set Setup Radius** CLI command is used to configure a RADIUS server on the switch. Refer to "Section 11: RADIUS Server Configuration" of the QLogic reference document, *SANbox 9000 Fibre Channel Switch: Command Line Interface*, to learn more about enabling RADIUS, located at http://download.qlogic.com/manual.

```
SANbox (admin) #> set setup radius

A list of attributes with formatting and current values will follow.
Enter a new value or simply press the ENTER key to accept the current value.
If you wish to terminate this process before reaching the end of the attributes
for the server being processed, press 'q' or 'Q' and the ENTER key to do so.
If you wish to terminate the configuration process completely, press 'qq' or
'QQ' and the ENTER key to so do.

DeviceAuthOrder (1=Local, 2=Radius, 3=RadiusLocal) [Local]
UserAuthOrder (1=Local, 2=Radius, 3=RadiusLocal) [Local]
TotalServers (decimal value, 0-5) [1 ]

Server: 1
ServerIPAddress (dot-notated IP Address) [10.20.11.8]
ServerUDPPort (decimal value) [1812 ]
DeviceAuthServer (True / False) [True ]
UserAuthServer (True / False) [True ]
AccountingServer (True / False) [False ]
Timeout (decimal value, 10-30 secs) [10 ]
Retries (decimal value, 1-3, 0=None) [0 ]
SignPackets (True / False) [False ]
Secret (32 hex or 16 ASCII char value) [******** ]

Do you want to save and activate this radius setup? (y/n): [n]
```

Reference

For more information, refer to the *SANbox 9000 Fibre Channel Switch Command Line Interface Guide* located at http://www..qlogic.com.

For general information on security for two switch fabrics, refer to "Security" on page 80.

Setting up this topology

Assumptions specific to this case study:

- The QLogic SB5602 switches are installed in a customer-supplied cabinet.

For switch installation guidelines, refer to the Sections 3 and 4: Planning and Installation for an SB5602 switch in the following document:
http://download.qlogic.com/manual/27144/59096-01_A.pdf

- The customer has provided the proper power receptacles.

  The power requirements for the SB5602 are 1 Amp at 100VAC or 0.5 Amp at 240VAC. For cabinet power requirements, refer to the customer-supplied cabinet setup manual, if any.

- The switches have *not* been connected to the power source and are *not* powered on.

- Network drops, IP addresses, Subnet mask, and gateway, and other information as specified in the checklist on page 38 have been provided by the customer.

  For switch network requirements refer to Section 2: General Description and Section 4: Installation for an SB5602 switch in the following document:

  http://download.qlogic.com/manual/27144/59096-01_A.pdf

  Note: The switches can be placed on either a public or private network. There are advantages to both configurations. For more information, refer to "Public versus private" on page 961.

  In this example, it is assumed that the customer has provided two Ethernet cables and that one of them is on the 172.23.199.x network, and that the other is connected to the 172.23.200.x network.

- License keys have been obtained.
  - To obtain a license key, go to the URL listed on the transaction code certificate that shipped with the product.

- A laptop, supplied by the installer, is used to configure the IP addresses of the switches. One interface from this laptop is directly connected to the FC switch through a crossover cable (one switch at a time), and this interface has an IP address of 10.0.0.253.

- A null modem cable is available.

- The customer has provided a temporary password to be used as the default password when configuring the IP address.

- Quicktools will be used and the Quicktools installation kit is available.

- Quicktools is already installed on the management laptop as per the Quicktools installation guidelines supplied to the customer.

Configure the IP address

Note: QLogic switches ship with a default IP address of 10.0.0.1.

1. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer. This causes both the switches to power up.
2. Select one of the switches to configure.

   Note: In this example, the switch that will have a Domain ID of **1** and an IP address of **172.23.199.22** will be the first switch to be configured.

3. Considering this switch is operational, open the **Quicktools** web applet by entering the default IP address (**10.0.01**) of this switch in the internet browser.
4. Enter the default login name of *admin* and password of *password* in the **Add a new fabric** dialog, and then click **Add Fabric**.

   IMPORTANT

   If you are prompted to change the password, change it to a password that was provided by the customer.

5. Quicktools detects a first time use, and presents an **Initial Start** dialog from which the Configuration Wizard can be launched. Open the **Wizards** menu and select **Configuration Wizard**. The Configuration Wizard is a series of dialogs used to configure the IP address and other basic parameters on new or replacement switches. In this example, the **Network Properties** dialog configures the IP for this switch.
6. Use the **Network Properties** dialog to configure the IP address:
   a. Open the faceplate display for the switch to be configured.
   b. Open the switch menu and select **Network Properties**.
   c. Using the **Network Properties** dialog, change the IP address as follows:

      Network Discovery: static
      IP address: 172.23.199.22

Gateway: 172.23.199.2
Subnet mask: 255.255.255.0
Active Ethernet Port: (Generally CPU back port)

d. Click **OK**.

Note: At this point, you can lose access to the switch. This is because the IP address of the switch just changed and you are no longer able to connect to the switch at its old IP address. However, remote login is now enabled for this switch.

Configure the other switch IP addresses in the cabinet

To configure the next switch in the cabinet, re-connect the cross-over network cable between the other SB5602 in the cabinet and the Windows laptop and then follow Step 3 through Step 6 in the previous "Configure the IP address" section. Use the IP address of (172.23.200.22), subnet mask (255.255.255.0) and gateway (172.23.200.2).

Configure the first FC switch

1. Set the switch name and fabric parameters.

   a. Open the faceplate display of the first switch that was configured in Step 3 through Step 6 in the "Configure the IP address" section. Open the **Switch** menu, and select **Switch Properties**.

   b. Use the **Switch Properties** dialog to change the following parameters as specified for our case study:

      – Domain ID: **1**
      – Domain ID Lock: (Check to lock the Domain ID assigned so that it does not change after a reboot and another switch with the same Domain ID cannot be added)
      – Symbolic Name: **SB5602_1**
      – Administrative state: Online
      – Leave the other parameters as default.

   c. Use the **Advanced Switch Properties** dialog to set the interopmode and time-out values. Open the **Switch** menu and select **Advanced Switch Properties**. Change/verify the following parameters as specified:

      – Interopmode: Standard

The switch is automatically goes offline, and is restored once the changes (if any) are completed.

2. Configure the switch ports.

Note: It is recommended that you always leave ports at auto-negotiate for both port type and speed unless there is a specific hardware requirement, or if a known problem with auto-negotiation exists between two port types.

a. The port settings can be configured using the **Port Properties** dialog. To open the **Port Properties** dialog, select one or more ports, open the **Port** menu, and then click **Port Properties**.

b. The drop-down list in **Port Properties** can be used to change the following parameters:

– Port Symbolic Name
– Port States
– Port Types
– Port Speeds

The port settings are configured as per the following tables for this case study:

Switch 1: SB5602_1:

Port #	Symbolic port name	Port type	Port speed
0	Red Host HBA 1	F_port	AutoNeg.
1	Red Storage 1	F_port	AutoNeg.
2	Green Storage 1	F_port	AutoNeg.
3	Blue Storage 1	F_port	AutoNeg.
4	ISL to SB5602_2	E_port	AutoNeg.
5	ISL to SB5602_2	E_port	AutoNeg.
6	Blue Storage 2	F_port	AutoNeg.
7	Green Storage 2	F_port	AutoNeg.
8	Blue Host HBA 2	F_port	AutoNeg.

Switch 2: SB5602_2:

Port #	Symbolic port name	Port type	Port speed
0	Red Host HBA 2	F_port	AutoNeg.
1	Red Storage 2	F_port	AutoNeg.

2	Green Host HBA 2	F_port	AutoNeg.
3	Green Storage 31	F_port	AutoNeg.
4	ISL to SB5602_2	E_port	AutoNeg.
5	ISL to SB5602_2	E_port	AutoNeg.
6	Blue Storage 1	F_port	AutoNeg.
7	Green Storage 4	F_port	AutoNeg.
8	Blue Host HBA 1	F_port	AutoNeg.

Repeat Step 1 and Step 2 for the other configured switch.

Connect cables

1. Connect ISLs.
   a. Attach the fiber cable between switches as shown in Figure 12 on page 118.
   b. After all cables are connected, use Quicktools to verify that all ISL connections are up by reviewing the fabric tree and topology view.
   c. Re-arrange icons to accurately reflect the switch configuration.
2. Connect host and storage ports.
   a. Attach fiber cable between switches and N_Ports.

Zone hosts and storage

1. Before creating any zones and enabling zoning, use the **Edit Zoning config** dialog to change the **Interop Auto save**, **Default Zone**, and **Discard inactive parameters**.
   - The **Interop Auto save** parameter (which determines whether any active zone changes that a switch receives from any other switch in the fabric must be saved to the zoning database of the switch) must be set to **False**.
   - The **Default Zone** (which allows all N_Ports to see each other in the absence of any active zones) must be set to **False**.
   - The **Discard inactive parameter (**which automatically removes or pulls out inactive zones from an active zoneset) must be set to **False**.
2. Open the **Edit Zoning** toolbar from the **Zoning** menu on the faceplate display.
3. Create a zone by clicking **Create Zone**.

4. Provide a descriptive name for the zone. This example zones “Red host HBA 1” and “Red Storage 1”.
   Type “**RedHBA1_1470_8aa**" and then press **Enter**.
5. Locate, then click, **“Red Host HBA 1”** (WWPN 10000000c938e554) in the **potential zone members** list.
6. Click **Add Member** to add the selected HBA to the zone created in Step 4.
7. Locate, then click, **“Red Storage 1”** (WWPN 50060482cc19bf87) in the **potential zone members** list.
8. Repeat Step 3 through Step 7 for all host and storage pairs in the environment.
9. Create a zone set by clicking **Create Zone Set**.
10. Provide a descriptive name for the zone set. This example will use the date of “Oct_31_06_1140”.
11. Add all of the zones created in Step 4 through Step 8 to the zone set by dragging the zones listed to the left into the zoneset created in Step 10.
12. To activate this zoneset, select **Activate Zone Set** from the **Zoning Edit** menu. Select a zoneset from the **Select zoneset** pull-down menu and click **Activate**.

When completed, the active zone set should be similar to what is shown next.

```
Zone set name = "Oct_31_06_1140"

            Zone name = "RedHBA1_1470_8aa"
               Zone Member = "10000000c938e554"
               Zone Member = "50060482cc19bf87"

            Zone name = "RedHBA2_1470_9aa"
               Zone Member = "10000000c938e555"
               Zone Member = "50060482cc19bf88"


            Zone name = "BlueHBA1_1489_8aa"
               Zone Member = "210100e08b8ac76d"
               Zone Member = "50060482cc19c447"

            Zone name = "BlueHBA2_1489_9aa"
               Zone Member = "210100e08baac76d"
               Zone Member = "50060482cc19c448"

            Zone name = "GreenHBA1_AllGreenStorage"
```

```
   Zone Member = "10000000c939a051"
   Zone Member = "50060482cc19c407"
   Zone Member = "50060482cc19c408"
   Zone Member = "50060482cc19c4c7"
   Zone Member = "50060482cc19c4c8"

Zone name = "GreenHBA2_AllGreenStorage"
   Zone Member = "10000000c939a052"
   Zone Member = "50060482cc19c407"
   Zone Member = "50060482cc19c408"
   Zone Member = "50060482cc19c4c7"
   Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point, the SAN is ready to pass I/O from the host to storage. Other steps, such as configuring LUN Masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the OS configuration guide for more detail.

Blade switch with direct attached storage

A blade server with two embedded FC switch modules directly attached to storage complies with the definition of a simple two-switch, one-hop SAN model. Before discussing the general layout of this topology type, it is interesting to note how the use of a blade server can simplify a desired two switch-one hop SAN design.

Consider the following scenario:

An end-user needs to hook up ten independent servers to storage using two departmental switches. Each of the ten servers needs its own power supply, cooling mechanism, and cabling. These aspects must also be considered for the departmental switches. This whole two-tier set up can be replaced by a 7U blade server chassis which can house at least ten independent servers, depending on the vendor type (IBM, Dell, HP, and so on). The chassis contains the power, cooling, and cabling components, and often incorporates a network switch. The blades within a single chassis can still run different applications and play independent roles.

The chassis also provides the ability to embed a pair of FC switch modules with a maximum of six external ports, depending on the vendor (Brocade, Brocade M-Series, or QLogic), which can be attached to an external fabric, or in this case, the storage. Each switch module is internally connected to each of the blade servers in the blade server chassis. The behavior, features, and management of the

switch modules are similar to those of departmental stand-alone edge switches.

The detailed blade server concept, its value to the FC SAN world, the basic architecture, and the various EMC-supported blade switch modules are further discussed in Chapter 3, "Extended Distance."

With this in mind, this example examines the general layout of a blade server switch module directly attached to storage. Figure 13 on page 130 is divided into two parts. The left side (a) represents the blade server chassis directly attached to storage through the switch modules, while the right side (b) represents the two-tier FC SAN to which it is equivalent.

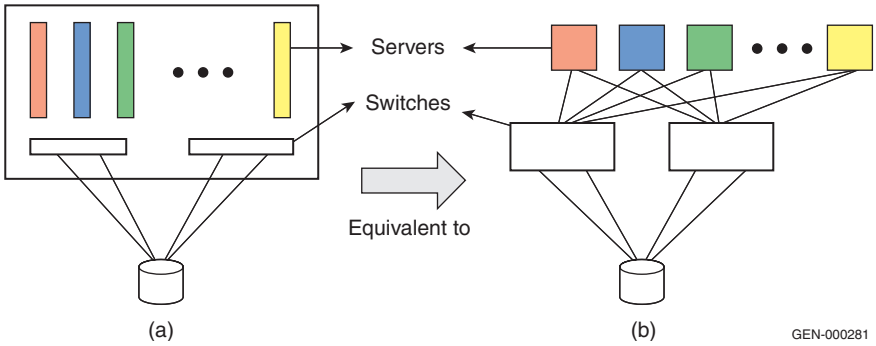


Figure 13 **(a) Blade server direct attached to storage (b) two-tier FC SAN**

This design occupies about 7U of cabinet space, and offers fully redundant fabrics. Most switch modules today have 4 G ports. Considering a switch module has at the most 6 ports (depending on the vendor type), a blade server is capable of up to 48 GB/s throughput.

General steps to set up a blade server

To set up a blade server:

1. Configure the management module.

   a. Physically install the blade server chassis and all its components in a customer-supplied rack as per the vendor-specific blade server hardware setup documentation.

   b. Power up the switch by connecting the power cords to the power receptacles provided by the customer.

   c. After powering it up, it is essential to configure the management module. The management module is generally located on the back face of the blade server chassis and has a serial console port and an Ethernet port.

   Note: At this stage:
   Refer to the "Best practices" on page 134 to study the two options available to configure the management modules and then select the preferred option between the two.

   It is also advisable to refer to the specific blade server vendor documentation to obtain the configuration details for assigning an IP to the management module. There are also different options to consider while doing this at data centers (for example, whether it is advisable to have the payload LAN and management LAN on the same LAN segment, and so on).

   d. In general, if configuring the management module through the *serial console* using the *Hyperterminal* application, it is essential to gather the following information:

      – Connection type: COM 1 or COM 3
      – Other settings:
        – Bits per second
        – Data bits
        – Parity bits
        – Stop bits

      If configuring the management through the *Ethernet port*, that is, through Telnet, the following information is required:

      – Default IP of management module
      – Default username
      – Default password

e. Once configured, the management module can be used to configure the other modules including the I/O modules.

2. Configure the other modules and the server blades.

   a. To configure other modules:

      Once the management module is configured with an IP address, it can be accessed through CLI or a Web interface in most cases, and can be used to configure, or rather assign, IPs to the other I/O modules, such as fibre channel switch modules (FCSM), and the Ethernet module if required.

   b. To configure the server blades:

      Generally, when supplied from the vendor, the server blades are configured for a network internal to the blade server chassis. Most blade servers today have VGA and USB connectors on the server blade which can be used to access and configure the server blades. The desired OS can be installed on the blade using these means of access, and on completing this installation, the other features, such as IP address, etc., can be set up for the server blades.

3. Check the following switch module configuration settings:

   a. The mode of operation: The switch modules must be in their respective native mode for this kind of a configuration.

      – Desired mode on Brocade switch module: Brocade native mode (interopmode 0)
      – Desired mode on Brocade M-Series switch module: Brocade M-Series Fabric mode 1.0
      – Desired mode on QLogic switch module: Standard mode

   b. Domain ID settings:

      – Domain ID range for Brocade switch modules: 1-239
      – Domain ID range for Brocade M-Series switch modules: 1-31
      – Domain ID range for QLogic switch modules: 97-127

   c. Ensure that the switch firmware revision and the switch management application are supported versions.

   d. If default zoning enabled does not allow the host ports to communicate with the storage ports, then the appropriate WWNN or WWPN zoning needs to be performed to establish

connectivity. As a result, the server blades can log into the storage and read from or write to the respective storage devices it is mapped to.

- Default zoning on the Brocade module restricts the initiator (host) ports to see or communicate with the target (storage) ports. Therefore, zoning must be configured on the Brocade modules.
- Default zoning on the Brocade M-Series module restricts the initiator (host) ports to see or communicate with the target (storage) ports. Therefore, zoning must be configured on the Brocade M-Series modules.
- Default zoning on the QLogic module allows the initiator (host) ports to see or communicate with the target (storage) ports. Therefore, zoning need not be configured on the QLogic modules.

4. Ensure switch module connectivity (physical).

   The switch modules must be connected directly to the storage using FC cables. EMC recommends at least two connections between any two FC components in a SAN. In this case, at least four cables must be connected in total: two from each switch module to the respective storage ports.

   The switch modules, as previously stated, are internally connected to each of the blade servers. Thus, a specific number of ports on each switch, depending on the number of server blades the blade server chassis can house, are dedicated F_ports, while the other external ports are G_ports. For this topology, these ports can be configured as F_ports since they will be hooked up to storage.

5. Use CLI / Web GUI to verify connectivity.

   The switch module CLI or switch management application can be used to verify that the physical connectivity was successful and that the switch can see the storage ports. The name server information can be obtained at this time so as to confirm whether all the N_Ports: HBA ports from the server blades and the storage ports, show up.

Best practices

For general best practices, refer to "Best practices" on page 37. The following best practices are recommended for configuring the blade server and setting up the direct attach to storage:

- There are two options available to access the management functions of the FC switch modules through the link to the management module. Option 1 is recommended for easy access and management of the switch module.
  - Option 1: If the IP address assigned to the switch is within the subnet of the management module, then the switch management functions can be accessed without launching the Management Module GUI.
  - Option 2: If the IP address assigned to the switch is not within the subnet of the management module, then the switch management functions can be accessed only by launching the Management Module GUI and then navigating to the specific functions.
- Connect each switch module to storage ports that are mapped to the same set of storage devices. Thus, each server blade can access the storage ports it is zoned with using two paths, each path passing through a different switch module. If either of the switch modules goes down, the server continues to have access to the storage port it is zoned with and there is no disruption in traffic between the server and storage.
- Always check supported switch module firmware release notes for caveats on direct attach to storage, if any.

Host and storage layout

For general host and storage layout information, refer to "Host and storage layout" on page 43. The following information is specific to blade servers.

There is little an end-user can do with the host and storage layout in the case of a blade server since the architecture allows every server blade in the blade server chassis to access both the switch modules. As a result, even if a storage port is hooked up to one of the switch modules, all the server blades can access the storage port through the switch module unless, and until, any of the internal or external ports on the switch modules are blocked to prevent access between the server and storage in this case. However, zoning can take care of the access settings.

Switch and fabric management

For general switch and fabric information, refer to "Switch and fabric management" on page 44. The following information is specific to blade switches.

It is recommended that end-users use the supported vendor specific web application and CLI for managing the switch.

- The Brocade modules can be managed using Brocade CLI or Web Tools.
- The Brocade M-Series modules can be managed using Brocade M-Series SAN browser.
- The QLogic modules can be managed using the QLogic CLI or QLogic SAN surfer.

Security

For general information on security, refer to "Security" on page 47. Specific information for this example follows. For further information on security, refer to Chapter 4, "Secure SANs."

- Blade servers usually exist in the "access layer" of a data center architecture and have specific requirements for security. These mainly deal with establishing secure access to switch, protection against attacks, and identifying users and other servers that are accessing the network. For easy configuration, management, and control, many capabilities, such as firewalls, are consolidated in the embedded switches in this example.
- The access to the embedded switch must be secured. This requires specific features, such as SSHv2 or RADIUS. Both protocols are means of securing connectivity to access the management interface of the switch. Without these capabilities, it would be considerably easier for an unauthorized user to change the configuration of the switch thus affecting the blade server's performance.
- The switch must also provide capabilities to prevent, as well as defend against network attacks. This is enabled by the following features: Port security (limited number of MAC addresses per port), MAC address notification (whether MAC address has moved), DARP inspection (ties a port to an ARP request, helping to ensure that a default gateway cannot be spoofed), IP source guard (Protects against IP being spoofed).
- An access layer switch must make sure that the network allows only authorized devices to connect to it.

Security features on blade server FC switch modules include:

- ISL and ELS authentication, as defined in FC-SP, provide a means to authenticate the identity of a connected switch, host or target and / or authorize a list of devices to join a fabric.
- ISL security is supported on E_Ports. ELS security is supported on F_Ports.
- Fabric Binding, port binding, and switch binding are introduced as a means to control the switch composition of a fabric. They may be enabled on the respective vendor switch modules using any of the switch management applications, just as they are for equivalent standalone switches. This has been explained in steps in the previous case study examples for configuring a two-switch Brocade (refer to "Connectrix B example" on page 81), Brocade M-Series (refer to "Connectrix M example" on page 102) or QLogic (refer to "QLogic example" on page 118) fabric.
- Security configuration management is similar to zoning configuration management.

IBM Brocade example

General layout

This example will show an IBM Brocade 4 GB SAN switch module (32R1812) directly attached to storage.

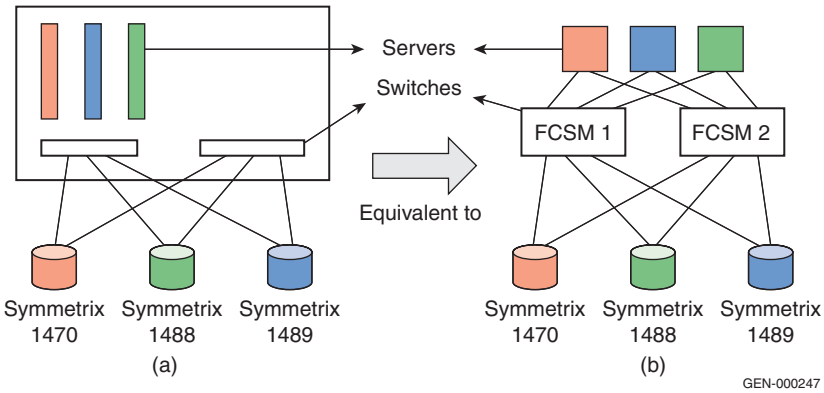


Figure 14 **IBM blade server directly attached to storage**

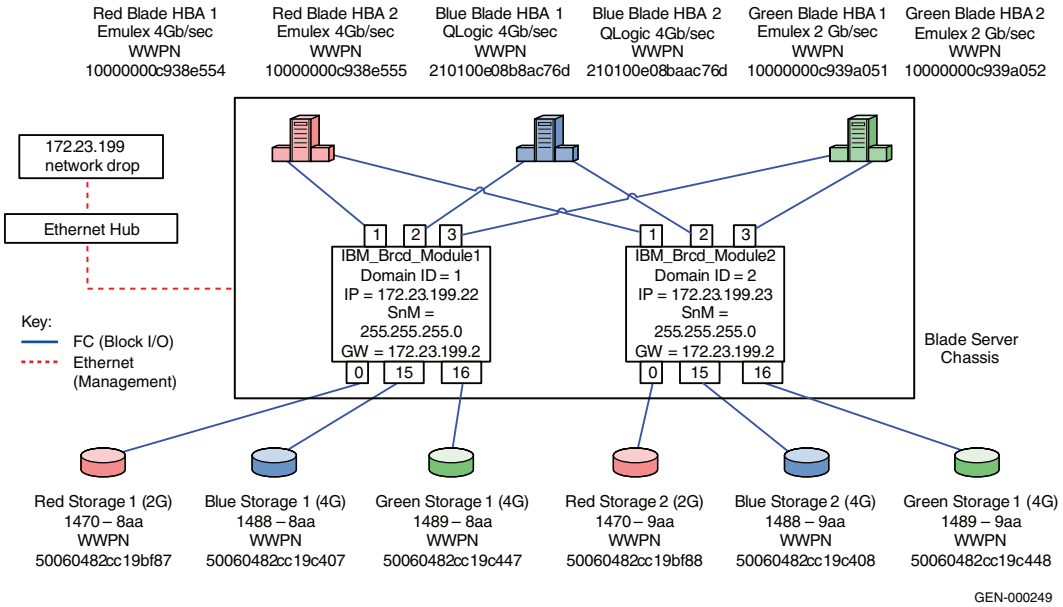


Figure 15 **IBM Brocade 4 GB SAN switch module (32R1812) directly attached to storage**

Figure 15 on page 137 represents an IBM blade server chassis with three server blades placed in slots 1, 2 and 3 of the ten slot chassis. Each of the server blades has a dual port HBA. One of the HBA ports on each server blade is internally connected to one of the Brocade switch modules, while the other HBA port on each server blade is connected to the other Brocade switch module.

Each switch module has 14 internal ports allocated for the 14 server blades the chassis can house. There are 6 external ports on the switch module. In this case, 3 of the 6 ports on each module are hooked up to a port on three different Symmetrix systems. For this example, assume that Port 0, 15 and 16 on Brocade Switch Module 1 (FCSM 1) are hooked to Symmetrix systems 1470 8AA, 1488 8AA and 1489 8AA, while the Ports 0, 15 and 16 on FCSM 2 are hooked to Symmetrix systems 1470 9AA, 1488 9AA and 1489 9AA, respectively.

Setting up this topology

Assumptions before configuring the components of this topology:

- The IBM Blade server chassis and its components have been installed in the cabinet and powered up as per the IBM hardware installation guide for blade servers. (http://www.redbooks.ibm.com/redbooks/pdfs/sg246342.pdf)
- The following components are used for this example:
  - Two Ethernet modules
  - Brocade switch I/O modules
  - Management module
  - 3 server blades
- The IP addresses and operating systems on the server blades can be configured by referring to the following document from IBM: http://www.redbooks.ibm.com/redbooks/pdfs/sg247313.pdf.

Set up IBM Brocade 4 GB SAN

To set up the IBM Brocade 4 GB SAN switch module direct attached to a Symmetrix:

1. Configure the Management Module for the IBM Blade server:

   The primary setup task for the Management Module would be assigning IP addresses, which are necessary to communicate with the Management Module Web GUI and command line interfaces.

   The IBM Management Module has two network interfaces:

   - An external interface (eth 0), which is accessible using the 10/100/1000Base T connector on the Management Module.
   - An internal interface (eth 1), which is connected to the management interfaces of all the installed I/O modules including the IBM Brocade 4 GB SAN switch module in this case.

   Note: The default static IP address of the external interface of the Management module is 192.168.170.125 with a default subnet mask of 255.255.255.0. The default IP address for the internal interface is statically assigned to be 192.168.70.126.

The steps for configuring the IP address on the Management Module are as follows:

a. Prepare a checklist with configuration details:

   Decide on the IP addresses, subnet masks, and gateway addresses to be assigned to the external and internal interfaces of the Management Module. They all must belong to the same subnet. In this example, the IP address on the external interface is changed to 172.23.199.60 and the IP address on the internal interface is changed to 172.23.199.61, which are both on the 199 subnet.

b. Connect the Management Module to a workstation using a cross over network cable.

c. Configure a static IP address for the workstation that is in the same subnet as the Management Module default IP addresses. In this case, a static address of 192.168.70.100 with a subnet mask of 255.255.255.0 was used for the workstation. IBM recommends not using addresses in the range of 192.168.70.125 through 192.168.70.130 since they conflict with the default addresses assigned by the management module.

d. Connect to the Management Module GUI or Web Interface by pointing the Web browser on the workstation to: http://192.168.70.125.

e. Enter a valid user ID and password to log into the Module management interface. The factory default configuration of a Management Module defines a user ID named USERID with a password of PASSW0RD.

Note: The number 0 is between the W and the R in PASSW0RD.

---

f. Select the **Network interfaces** option under the **Management Module (MM) Control** menu.

g. Enter the desired external and internal IP addresses, subnet masks, and default Gateway for the Management Module. In this example:

   External IP address: 172.23.199.60
   Subnet Mask: 255.255.255.0
   Gateway Address: 172.23.199.2

   Internal IP address: 172.23.199.61

Subnet Mask: 255.255.255.0
Gateway Address: 172.23.199.2

Click **Save** to store these new IP addresses.

h. Restart the Management Module.

i. Pull out the cross over cable from the Ethernet port of the Management Module and connect it with an Ethernet cable to a hub on the 199 subnet.

One can now connect to the Management Module Web interface and Telnet into its Command Line Interface using the 172.23.199.60 IP address assigned to its external network interface.

2. Configure the Brocade SAN switch modules:

The primary set up tasks for the Brocade modules are to:

- Assign IP addresses to the Brocade switch (I/O module) Management interfaces.
- Enable the external Ethernet port on the Brocade modules.

Note: When a new switch I/O module is first installed, the Management Module assigns a default IP address to the management interface of the I/O module. The default IP address is chosen based on the I/O module bay on the back plane of the chassis where the I/O module is installed. Those I/O modules installed in I/O module bays 1, 2, 3, and 4 are assigned IP addresses 192.168.70.127, 192.168.70.128, 192.168.70.129, and 192.168.70.130, respectively.

The steps to configure the IP addresses on the Brocade modules are as follows:

a. Prepare a checklist with configuration details:

Decide on the IP addresses, subnet masks, and gateway addresses to be assigned to the external interfaces of the two Brocade Modules. They must belong to the same subnet as the Management module. This example configures the IP address of one Brocade module as 172.23.199.22 and the other as 172.23.199.23.

b. Connect to the Management Module by pointing the web browser to: http://172.23.199.60

c. From the Management Module interface select **I/O Module Tasks->Management**.

d. Select the specific I/O module based on the I/O Bay it is pushed into. It is a general practice to place the Brocade modules in I/O bays 3 and 4. The IP address fields are updated as follows:

– For Bay 3: Brocade module 1

IP address: 172.23.199.22
Subnet mask: 255.255.255.0
Gateway Address: 172.23.199.2

– For Bay 4: Brocade module 2

IP address: 172.23.199.23
Subnet mask: 255.255.255.0
Gateway Address: 172.23.199.2

e. Click **Save** to activate the new IP address.

f. Select the **Advance Management** link for each of the Brocade modules, set the **External Ports** field to **Enable**, and click **Save**. Leave everything else at default settings.

At this point both the Brocade switch modules have an IP address. A Brocade Web Tools browser can now be pointed to this IP or one can Telnet into this IP address to use the Brocade CLI to manage the switch.

3. Configure the switches.

a. Telnet into one of the Brocade switch modules by issuing a **Telnet 172.23.199.22** command.

b. Enter the default username and password of *USERID* and *PASSW0RD*.

IMPORTANT

It is strongly recommended that when prompted to change the password, you change it to a password that was provided by the customer.

c. Verify mode of operation. On issuing an **interopmode** (CLI) command, the switch must return **interopmode: Off** which implies that the switch is running in its native Brocade mode. If not, disable the switch by issuing a **switchdisable** command followed by the **interopmode 0** command. A reboot is required to restore normal configuration on the switch.

d. Assign a Domain ID: The switch modules automatically take a Domain ID in the range of 1-239. If a specific Domain ID is desired, as in this case,

   – The Brocade module needs to be disabled by issuing **switchdisable**.
   – A new Domain ID can be assigned by running **configure**.
   – When prompted to choose *yes* or *no* for **Configure.Fabric parameters**, type **y**.
   – For the Domain ID setting, type **1**.
   – To have Domain ID fixed to **1**, press **Enter** until you reach the **Insistent Domain ID Mode** field and type **y**.
   – Press **Enter** for the rest of the values to accept default settings.

e. Assign a switch name to the Brocade module by issuing the **switchname *IBM_brcd_module1*** command.

f. Ensure that a supported firmware version is running on the switch module by running **version** on the switch CLI. If not, download the latest firmware version supported by EMC and IBM.

g. Verify that no zoning is enabled or is active on the switch modules by running a **cfgshow** command.

h. Configure the ports:

   To configure port speed to auto negotiate issue the **portcfgspeed <portnumber> 0** command.

   Note: Outside of a hardware requirement or if a known problem with auto-negotiation exists between two port types. We recommend that you leave ports at auto-negotiate for both port type and speed.

   To name the port, issue a **portName <portnumber> [name]** command. For example, in this case the internal ports on FCSM1connected to the blade server HBAs can be named as "Red Host HBA 1", "Green Host HBA 1", and "Blue Host HBA 1" as specified in the following tables:

Switch 1, FCSM 1:

Port #	Symbolic port name	Port type	Port speed
0	Red Storage 1	F_port	AutoNeg.
1	Red Blade HBA 1	F_port	AutoNeg.
2	Blue Blade HBA 1	F_port	AutoNeg.
3	Green Blade HBA 1	F_port	AutoNeg.
15	Blue Storage 1	F_port	AutoNeg.
16	Green Storage 1	F_port	AutoNeg.

Switch 2, FCSM 2:

Port #	Symbolic port name	Port type	Port speed
0	Red Storage 2	F_port	AutoNeg.
1	Red Blade HBA 2	F_port	AutoNeg.
2	Blue Blade HBA 2	F_port	AutoNeg.
3	Green Blade HBA 2	F_port	AutoNeg.
15	Blue Storage 2	F_port	AutoNeg.
16	Green Storage 2	F_port	AutoNeg.

Repeat steps Step a through Step f for the other switch module (172.23.199.23, switch Name: IBM_brcd_module2”).

4. Connect the FC switch modules and verify connectivity.

   Connect FC cables from the first two or left-most external port on the switch modules to the two storage devices. To verify connectivity:

   a. A **switchshow** (CLI) on any of the above configured Brocade switch modules with this configuration will show ports 1, 2, and 3 as F_ports with the respective Server Blade HBA port WWNs. Ports 0, 11 on both the Brocade modules will again show up as F_ports with the respective Symm port WWNs.

      The red and green colors indicate what initiator is mapped to which target. The red server blade must log into Symmetrix dir 3A and needs to be zoned accordingly while the green server blades are meant to log into Symmetrix dir 3C and needs to be zoned accordingly.

b. A **fabricshow** (CLI) on any of the modules must show two domains, each switch module representing an independent domain. A nameserver query must list ten ports: six ports coming from the three servers with dual port HBAs and four storage ports.

As shown in Figure 15 on page 137, each server blade has two paths to access the respective Symmetrix director it is zoned with. The two paths go through the two switch modules respectively.

5. Configure zoning on the Brocade switch modules:

Zoning is configured on the Brocade modules using either the Brocade CLI or Web Tools. This example uses Brocade Web Tools to do the zoning.

a. Point the web browser to the Brocade module (172.23.199.22), which opens Web Tools.

b. Select the **Zoning** icon on the bottom left of the interface.

c. Enter username and password as *USERID* and *PASSW0RD*.

d. On the **Zoning** interface screen, select the **Zone** tab and then select **Create**.

e. On the **Create New Zone** dialog box, provide a descriptive zone name. This example will zone “Red Host HBA 1” and “Red Storage 1”, so **“RedHBA1_1470_8aa”** will be entered.

f. Verify that the **Zone Name** shows in the drop-down menu. Enter the respective zone members from the **Member Selection list** of N_Ports: Host and storage WWPNs from the left side of the zoning dialog box into the “Zone Members” area on the right.

g. Repeat Step d through Step f for all host and storage pairs in the environment.

h. Create a zone set by clicking on the **Config** tab and then clicking **New**.

i. In the **Create New Config** dialog box, provide a descriptive name for the configuration (zone set). In this case, the date of “Oct_31_06_1140” will be used.

j. Add the zones created in Step d through Step g into the configuration created in Step i.

k. Select **Actions** from the menu bar, and then select **Enable Config** to activate and save the desired zoning configuration; in this case “Oct_31_06_1140”.

When completed, the active or effective configuration displayed should be similar to what is shown below:

```
Zone set name = "Oct_31_06_1140"

            Zone name = "RedHBA1_1470_8aa"
               Zone Member = "10000000c938e554"
               Zone Member = "50060482cc19bf87"

            Zone name = "RedHBA2_1470_9aa"
               Zone Member = "10000000c938e555"
               Zone Member = "50060482cc19bf88"

            Zone name = "GreenHBA1_1489_8aa"
               Zone Member = "10000000c939a051"
               Zone Member = "50060482cc19c447"

            Zone name = "GreenHBA2_1489_9aa"
               Zone Member = "10000000c939a052"
               Zone Member = "50060482cc19c448"

            Zone name = "BlueHBA1_AllBlueStorage"
               Zone Member = "210100e08b8ac76d"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"

            Zone name = "BlueHBA2_AllBlueStorage"
               Zone Member = "210100e08baac76d"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
```

Complete the SAN setup

At this point, the SAN is ready to pass I/O from the host to storage. Other steps, such as configuring LUN Masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the OS configuration guide for more detail.

Complex Fibre Channel SAN topologies

This section contains information on complex Fibre Channel SAN topologies, including:

- "Four switch full mesh" on page 147
- "Compound core edge switches" on page 199
- "Heterogeneous switch interoperability" on page 256

Best practices

General best practices for simple and complex Fibre Channel SAN topologies are described in "Best practices" on page 37. The following information is specific to complex Fibre Channel SAN topologies only.

- ISL subscription best practice — While planning the SAN, keep track of how many host and storage pairs will be utilizing the ISLs between domains. As a general best practice, if two switches will be connected by ISLs, ensure that there is a minimum of two ISLs between them and that there are no more than six initiator and target pairs per ISL. For example, if 14 initiators access a total of 14 targets between two domains, a total of three ISLs would be necessary. Consider the applications that will use the ISLs before applying this best practice when setting up a configuration.

Host and storage layout

For host and storage layout information for both simple and complex Fibre Channel SAN topologies, refer to "Host and storage layout" on page 43.

Switch and fabric management

For switch and fabric management information for both simple and complex Fibre Channel SAN topologies, refer to "Switch and fabric management" on page 44.

Security

It is important to secure your fabric. For general information on security, refer to "Security" on page 47. For more information on security, refer to Chapter 4, "Secure SANs."

Four switch full mesh

This section contains information on four switch full mesh topologies.

Overview of Fabric Design considerations

General layout

In the four switch full mesh fabric shown in Figure 16, each switch is connected to every other switch with a minimum of two ISLs. This prevents any ISL from becoming a single point of failure. Each switch is also connected to a management LAN through IP.

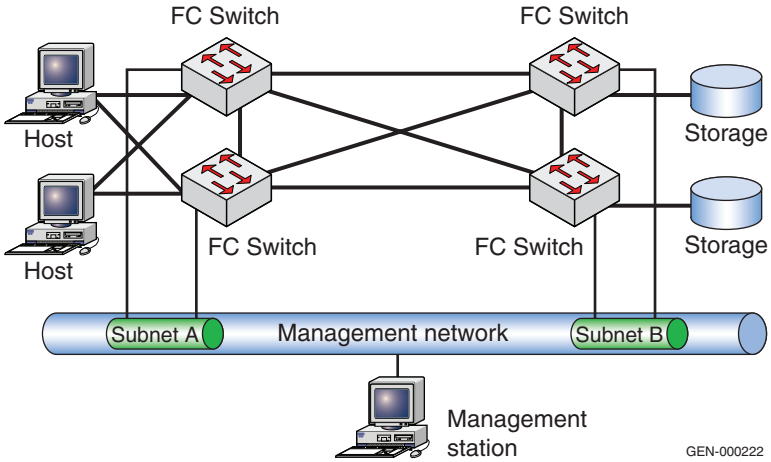


Figure 16 **Four switch full mesh fabric**

Each switch type can be used in any position but this configuration is not recommended for switches with less than 16 ports.

Best practices

For general best practices for all SANs, refer to "Best practices" on page 37. For specific best practices to complex Fibre Channel SAN topologies, refer to "Best practices" on page 146. Specific information for four switch full mesh fabrics follows.

- One of the use cases for a four switch full mesh fabric is distance extension. In these configurations, it is essential to monitor the ISLs for oversubscription conditions (utilization > 80%) which may lead to back pressure and any errors that are incrementing, especially bit errors or invalid transmission words, as these may lead to credit starvation. For more information on credit

starvation, refer to "Buffer-to-Buffer credit information" on page 454. See the individual case studies below for information on how to configure in each environment.

Host and storage layout

Specific information for four switch full mesh fabrics follows.

- In the four switch fabric examples used in this section, hosts and storage can be connected to any switch, but should be connected to the same switch when possible. A notable exception to this is in a distance extension environment when the switches are used to aggregate many different connections over an ISL and provide additional BB_Credit. In this configuration, the whole point of having multiple switches is to use the ISLs.

For general information on host and storage layout or all SANs, refer to "Host and storage layout" on page 43.

Switch and fabric management

For general information on switch and fabric management or all SANs, refer to "Switch and fabric management" on page 44.

Security

For general information regarding security or all SANs, refer to "Security" on page 47. For more information on security, refer to Chapter 4, "Secure SANs."

Connectrix B example

General layout

As shown in Figure 17 on page 149, a four switch full mesh fabric has been created by connecting each switch to every other switch in the fabric with two ISLs. Each switch is also connected to a Management LAN via Ethernet cables. There are three hosts (Red, Blue, and Green) as well as three sets of storage ports (also Red, Blue, or Green). The colors are intended to indicate which hosts access which storage ports. In addition to the CTPs, each ED-48000B contains the following three blades:

- Slot 1 — FC4-16
- Slot 2 — FC4-32
- Slot 3 — FR4-18i

To enhance the continuity of this document, this configuration has been created with the intention of reusing it in other sections. As a result, not all of the capabilities of the FR4-18i will be utilized in this section. Refer to "Case study #6:" on page 327 for more information.

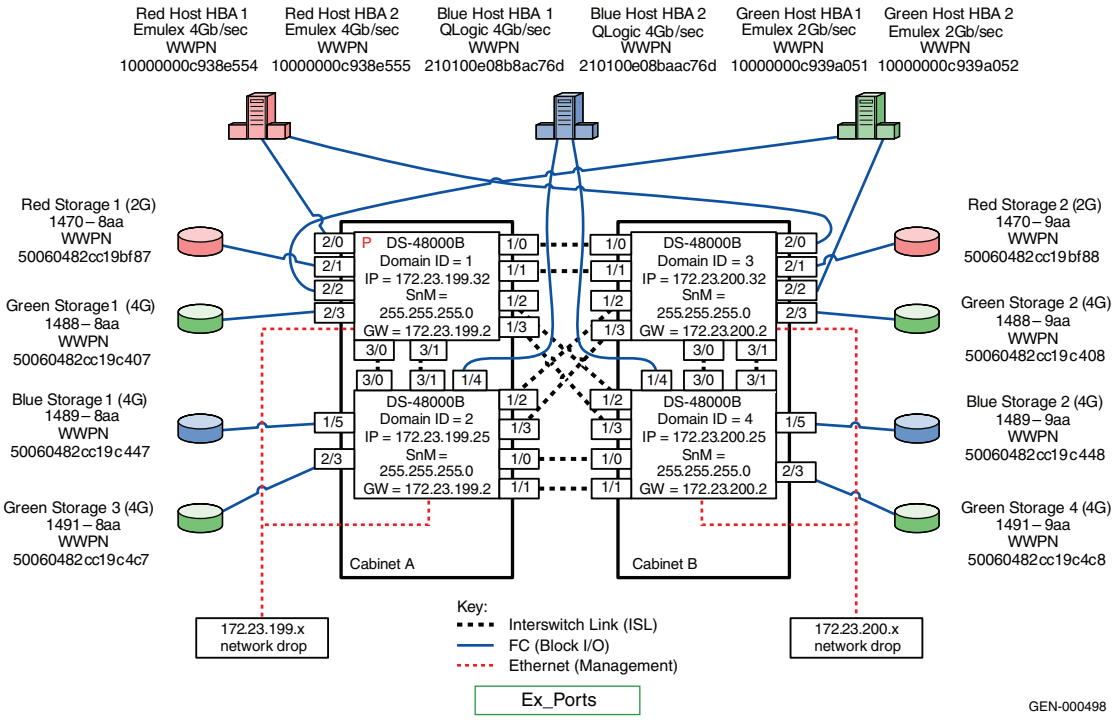


Figure 17 **Four switch fabric with ED-48000Bs**

All Connectrix B series switches supporting v5.2.x Fabric Operating software and higher are supported using this topology. In this example, the Connectrix ED-48000B director is used.

Best practices

For general information on best practices for four switch fabrics, refer to "Best practices" on page 146.

For Connectrix B specific best practices, refer to "Connectrix B" on page 40.

ISL layout

In this example, the ISLs are connected to ports on Slot 1 (FC4-16) and on Slot 3 (FR4-18i) and not to Slot 2 (FC4-32). The reasoning behind this layout is that the FC4-32 is oversubscribed (16:8), and because of this, utilizing these ports for ISLs may make the environment more susceptible to congestion and backpressure. The ISLs are spread out over two different blades to enhance fault tolerance and

serviceability. ISLs destined to the same Domain are kept within the same port octet to take advantage of trunking.

Host and storage layout

Both hosts and storage can be placed anywhere on the SAN. However in this example, we will try to conserve the ports on the FC4-16 and only use them for ISLs or host and storage pairs which could potentially utilize the full 4 Gb / sec bandwidth. Host and storage pairs that cannot fully utilize the full 4 Gb / sec bandwidth are placed onto the FC4-32 blade. For example, while the Red host HBAs are 4 Gb / sec, the storage ports they access are only 2 Gb / sec. Because of this, they are placed onto the FC4-32. Similar reasoning is applied to the Green host since the HBAs are 2 Gb / sec even though the storage ports are 4 Gb / sec. The Blue host and storage ports are all 4 Gb / sec so they are placed on the FC4-16.

Note: It is not essential to configure the environment this way, but it does ensure that all resources are being efficiently utilized.

For general information on host and storage layout for four switch fabrics, refer to "Host and storage layout" on page 146.

Switch and fabric management

In this example, CLI will be used to configure the environment, but Fabric Manager or Web Tools could also have been used.

For general information on switch and fabric management for four switch fabrics, refer to "Switch and fabric management" on page 146.

Security

The connectivity and device discovery for a large Connectrix–B switch fabric may be secured by appointing the following binding techniques. All these are Connectrix-B specific features.

- Fabric binding is a security method for restricting switches within a multiple-switch fabric. The SCC policy prevents unauthorized switches from joining a fabric. Switches are authenticated using digital certificates and unique private keys provided to the Switch Link Authentication Protocol (SLAP).
- Switch binding is a security method for restricting devices that connect to a particular switch. If the device is another switch, this is handled by the SCC policy. If the device is a host or storage device, the Device Connection Control (DCC) policy binds those devices to a particular switch. Policies range from completely restrictive to reasonably flexible, based upon customer needs.

- Port binding is a security method for restricting host or storage devices that connect to particular switch ports. The DCC policy also binds device ports to switch ports. Policies range from completely restrictive to reasonably flexible, based on customer needs. For switches running Fabric OS v5.2.0 and later, the SCC ACL with strict fabric-wide consistency can also be used for switch binding in addition to the Secure Fabric OS mechanism.

The method to enable SCC and DCC policies has been provided in "Enabling the Switch Connection Policy (SCC)" on page 216).

For general information on security for four switch fabrics, refer to "Security" on page 146. Specific information for this example follows.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data".
- The switches have *not* been connected to the power source and are *not* powered on.
- Network drops, IP addresses, subnet mask and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix B, "EMC Connectrix Products Data".

  Note: In this example, we are going to assume that the Customer provided us with 8 Ethernet cables and that four of them are on the 172.23.199.x network and that the other four are connected to the 172.23.200.x network.

- The correct number of line cards have been installed into the ED-48000Bs.
  - For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.

- Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.

- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- FOS v5.2.0a or greater is installed on all switches in the fabric.

Configure the IP address

To configure the IP address:

Note: Connectrix B switches may ship with a default IP address not on the 172.23.199.x subnet. The ED-48000B director uses a maximum of three IPs per unit: one IP for the switch and one IP for each control processor.

1. Attach the provided serial cable between the serial port on Domain 1 and an RS-232 serial port on the management PC. The serial cable is wired with only pins 2, 3, and 5 wired straight through.
2. Power up the switch by connecting the power cords to the power receptacles provided by the customer.
3. Run a terminal emulation program, such as Hyperterm, on Windows hosts, or TERM in a UNIX environment.
4. Configure the terminal for 9600 Baud, 8 Data Bits, No Parity, 1 stop bit, and no flow control.
5. Press **Return** to get a prompt.

IMPORTANT

It is strongly recommended that when prompted to change the password, you change it to a password that was provided by the customer. This can also be done using the passwd command from the prompt at any time.

6. Log in using the default values: Username: *admin* Password: *password.*
7. At the prompt, enter **ipaddrset -sw1** and press **Return**.
8. When prompted, supply IP address (172.23.199.22), subnet mask (255.255.255.0), and gateway address (172.23.199.2).

Note: The Fibre Channel Addresses will not be used for this example.

9. At the prompt, enter **ipaddrset –cp 0** and press **Return**.
10. When prompted, supply hostname (**48K_1_23**), IP address (**172.23.199.23**), subnet mask (**255.255.255.0**), and gateway address (**172.23.199.2**).
11. At the prompt, enter **ipaddrset –cp 1** and press **Return**.
12. When prompted, supply hostname (**48K_1_24**), IP address (**172.23.199.24**), subnet mask (**255.255.255.0**), and gateway address (**172.23.199.2**).
13. Power down the switch and disconnect the serial cable.
14. Connect the switch to a 10/100BaseT Ethernet connection.
15. Power up the switch.

    The switch can now be accessed with IP-based management.
16. Repeat these steps for Domain 2, 3, and 4 using the following information:

	Domain ID	Switch IP	CP0 IP	CP0 Name	CP1 IP	CP1 name
Domain 1	4	172.23.199.22	172.23.199.23	48K_1_23	172.23.199.24	48K_1_24
Domain 2	7	172.23.199.25	172.23.199.26	48K_2_26	172.23.199.27	48K_2_27
Domain 3	10	172.23.200.22	172.23.200.23	48K_3_23	172.23.200.24	48K_3_24
Domain 4	13	172.23.200.25	172.23.200.26	48K_4_26	172.23.200.27	48K_4_27

Note: For Domains 3 and 4, use the gateway address of **172.23.200.2**.

Configure FC switches

To configure FC switches:

1. Set the switch name and fabric parameters for the switch with the Domain ID of 1.

   Note: The following configurations need to be done with the switch *disabled*.

2. Configure the fabric parameters.

   a. From the switch prompt, enter **switchdisable** to disable the switch.

b. From the switch prompt, enter **configure** to enter the configuration parameter menu.

c. Enter **Y** at the **Fabric Parameters** prompt.

d. Enter **1** for desired Domain ID at the Domain prompt and press **Enter**.

e. The R_A_TOV should be automatically set to 10000. If it is not, enter **10000** at the prompt and press **Enter**.

f. The E_D_TOV should be automatically set to 2000. If it is not, enter **2000** at the prompt and press **Enter**.

g. Accept the following defaults for the rest of the fields under the **Fabric Parameters** menu by pressing **Enter** after each prompt:

- WAN_TOV = 0
- MAX_HOPS = 7
- Data field size = 2112
- Sequence Level Switching = 0
- Disable Device Probing = 0
- Suppress Class F Traffic = 0
- Switch PID Format = 1
- Per-frame Route Priority = 0
- Long Distance Fabric = 0
- BB_Credit = 16

Note: For this case study, there is no long distance between the DS-48000B switches. The ISLs connecting the two are less than 10 km.

h. At the **Insistent Domain ID Mode** prompt, enter **y** to accept the **Insistent Domain ID** setting. When this mode is set, the switch attempts to acquire the domain number programmed in its **Switch Fabric Settings** from the fabric.

i. Accept the default values from the remaining **Fabric Parameter Menu** items by pressing **Enter** after each prompt:

- Virtual Channel parameters (yes, y, no, n): [**no**]
- F_Port login parameters (yes, y, no, n): [**no**]
- Zoning Operation parameters (yes, y, no, n): [**no**]
- RSCN Transmission Mode (yes, y, no, n): [**no**]
- Arbitrated Loop parameters (yes, y, no, n): [**no**]
- System services (yes, y, no, n): [**no**]
- Portlog events enable (yes, y, no, n): [**no**]
- ssl attributes (yes, y, no, n): [**no**]

– http attributes (yes, y, no, n): [**no**]
– snmp attributes (yes, y, no, n): [**no**]
– rpcd attributes (yes, y, no, n): [**no**]
– cfgload attributes (yes, y, no, n): [**no**]
– web tools attributes (yes, y, no, n): [**no**]

Note: You may also press **CNTRL D** after making the last change in the menu to exit and save the changes. This will eliminate the need to accept the default values for the rest of the menu items.

j. Repeat from Step a for switches with the Domain IDs 2, 3, and 4.

Connect cables

To connect the cables:

1. Connect ISLs.
   a. Attach Fiber cable between switches as shown in Figure 17 on page 149.
   b. After all cables are connected, use **switchshow** and **topologyshow** commands to ensure all links.
2. Connect host and storage ports.
   a. Attach fiber cable between switches and N_Ports.
3. Verify port login status.
   a. After all cables are connected, use the **switchshow** CLI command to verify the all of the ports logged into the switch.

Zone hosts and storage

To zone hosts and storage, telnet into one of the switches in the fabric and using the following zoning commands:

1. :Create zones using the **zonecreate** commands below:

```
zonecreate "RedHBA1_1470_8aa", "10000000c938e554; 50060482cc19bf87"
zonecreate "RedHBA2_1470_9aa", "10000000c938e555; 50060482cc19bf88"
zonecreate "BlueHBA1_1489_8aa", "210100e08b8ac76d; 50060482cc19c447"
zonecreate "BlueHBA2_1489_9aa", "210100e08baac76d; 50060482cc19c448"
zonecreate  "GreenHBA1_AllGreenStorage", "10000000c939a051; 50060482cc19c407;
   50060482cc19c408; 50060482cc19c4c7; 50060482cc19c4c8"
zonecreate "GreenHBA2_AllGreenStorage", "10000000c939a052; 50060482cc19c407;
   50060482cc19c408; 50060482cc19c4c7; 50060482cc19c4c8"
```

2. Create the configuration by using the **cfgcreate** command.

```
cfgcreate "Oct_31_06_1140" , "RedHBA1_1470_8aa; RedHBA2_1470_9aa;
   BlueHBA1_1489_8aa; BlueHBA2_1489_9aa; GreenHBA1_AllGreenStorage;
   GreenHBA2_AllGreenStorage"
```

3. Enable the configuration by using the **cfgenable** command.

   ```
   cfgenable "Oct_31_06_1140"
   ```

4. Enter **Y** at the confirmation prompt.
5. Enter **cfgshow** to display zoning info.

When completed, the zone information should be similar to what is shown below.

Defined configuration:

```
cfg: Oct_31_06_1140
   RedHBA1_1470_8aa; RedHBA2_1470_9aa; BlueHBA1_1489_8aa; BlueHBA2_1489_9aa;
   GreenHBA1_AllGreenStorage; GreenHBA2_AllGreenStorage"
zone: RedHBA1_1470_8aa
      10000000c938e554; 50060482cc19bf87
zone: RedHBA2_1470_9aa
      10000000c938e555; 50060482cc19bf88
zone: BlueHBA1_1489_8aa
      210100e08b8ac76d; 50060482cc19c447
zone: BlueHBA2_1489_9aa
      210100e08baac76d; 50060482cc19c448
zone: GreenHBA1_AllGreenStorage
      10000000c939a051; 50060482cc19c407;
      50060482cc19c408; 50060482cc19c4c7;
      50060482cc19c4c8
zone: GreenHBA2_AllGreenStorage
      10000000c939a052; 50060482cc19c407;
      50060482cc19c408; 50060482cc19c4c7;
      50060482cc19c4c8
```

Effective configuration:

```
CFG: Oct_31_06_1140
Zone: RedHBA1_1470_8aa
      10000000c938e554
      50060482cc19bf87
Zone: RedHBA2_1470_9aa
      10000000c938e555
      50060482cc19bf88
Zone: BlueHBA1_1489_8aa
      210100e08b8ac76d
      50060482cc19c447
Zone: BlueHBA2_1489_9aa
      210100e08baac76d
      50060482cc19c448
Zone: GreenHBA1_AllGreenStorage
      10000000c939a051
```

```
      50060482cc19c407
      50060482cc19c408
      50060482cc19c4c7
      50060482cc19c4c8
Zone name = "GreenHBA2_AllGreenStorage
      10000000c939a052
      50060482cc19c407
      50060482cc19c408
      50060482cc19c4c7
      50060482cc19c4c8
```

Save configuration

In case the configuration is lost, or unintentional changes are made, keep a backup copy of the configuration file on a host computer.

To upload a configuration file:

1. Verify that the FTP service is running on the host computer. The host must have an FTP server application running.
2. Connect to the switch through the Telnet and log in as admin.
3. Enter the **configUpload** command.

   The command becomes interactive and you are prompted for the required information.

   Example:

```
switch:admin> configupload
Protocol (scp or ftp) [ftp]: ftp
Server Name or IP Address [host]: 192.1.2.3
User Name [user]: JohnDoe
File Name [config.txt]: /pub/configurations/config.txt
Password: xxxxx
configUpload complete: All config parameters are uploaded.
switch:admin>
```

Enable the Switch Connection Policy (SCC)

1. At the switch prompt, enter **fddcfg –fabwideset "SCC:S;DCC;"**
2. Press **Enter**.

   This command will set a strict SCC and tolerant DCC fabric-wide consistency policy.

Note: When a switch is joined to a fabric with a strict Switch Connection Control (SCC) or Device Connection Control (DCC) fabric-wide consistency policy, the joining switch must have a matching fabric-wide consistency policy. If the strict SCC or DCC fabric-wide consistency policies do not match, the switch cannot join the fabric and the

neighboring E_ports will be disabled. If the strict SCC and DCC fabric-wide consistency policies match, the corresponding SCC and DCC Access Control List (ACL) policies are compared.

3. To verify that the policy has been set, the **fddcfg –showall** command can be run on any switch in the fabric. Any switch on the fabric should show output similar to:

```
switch:admin> fddcfg --showall
Local Switch Configuration for all Databases:-
DATABASE - Accept/Reject
------------------------
SCC - accept
DCC - accept
PWD - accept
Fabric Wide Consistency Policy:- "SCC:S;DCC"
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN Masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the OS configuration guide for more detail.

Connectrix MDS example

General layout

Figure 18 illustrates four Connectrix MDS 9506s full mesh configuration.

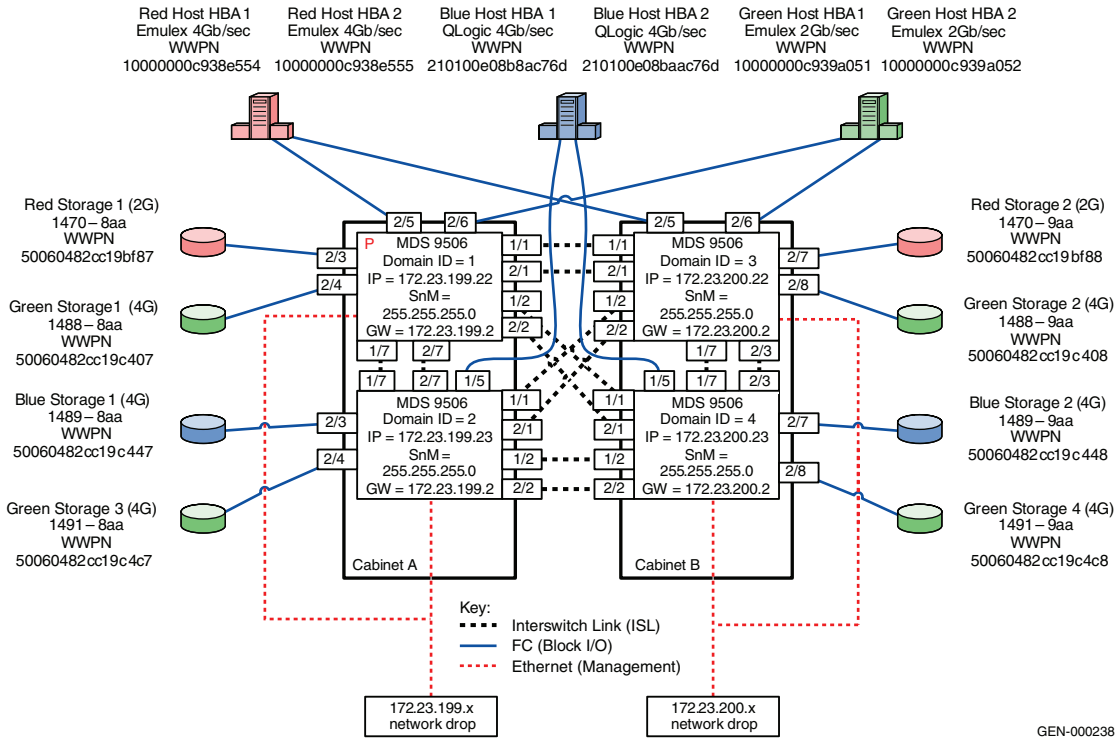


Figure 18 **Four Connectrix MDS 9506s full mesh configuration**

Best practices

For general information on best practices for four switch fabrics, refer to "Best practices" on page 146. Specific information for this example follows.

By default thresholds are set to 80% utilization.

Host and storage layout

For general information on host and storage layout for four switch fabrics, refer to "Host and storage layout" on page 146. Specific information for this example follows.

Line Rate Mode cards have no special restrictions. Over-subscribed cards should be used for hosts only..

Switch and fabric management

For general information on host and storage layout for four switch fabrics, refer to "Switch and fabric management" on page 146. Specific information for this example follows.

For this topology Fabric Manager can be used.

Security

For general information on security, for four switch fabrics, refer to "Security" on page 146. Specific information for this example follows.

Use switch and port binding for security.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data".
  - For Cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The switches have *not* been connected to the power source and are *not* powered on.
- Network drops, IP addresses, subnet mask and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix B, "EMC Connectrix Products Data".

Note: Connectrix MDS switches can be directly connected to the customer's LAN. The switches can be placed on either a public or private network. There are advantages to both configurations. For more information, refer to "Public versus private" on page 961.

This example assumes that The customer has provided us with two Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network.

- The proper number of line cards have been installed into the Connectrix MDS 9513s.
  - For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- A laptop, connected to a Connectrix MDS serial port, will be used to configure the IP addresses of the switches.
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- Cisco CLI, Fabric Manager, and Device Manager will be used.

Configure the IP address

To configure the IP address:

1. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer.
2. Select one of the switches to configure and set the IP to 172.23.199.22.
3. Supply a network connection to the appropriate subnet.
4. Using an RS232 serial cable, connect to the serial port of the switch with a baud rate of 9600, 8 data bits, no parity, 1 stop bit and no flow control.

   The **login** prompt should display.
5. Log in the first time with user name *admin* and password *admin*.

   You should be prompted to supply a new strong password for CLI user admin.
6. For this example, select **no** when asked if you want to run setup.

Note: This example will start with the switch that will have a Domain ID of **1** and an IP address of **172.23.199.22**.

CLI commands to configure the IP and gateway

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***interface mgmt 0***

  Switch(config-if)#***ip address 172.23.199.22 255.255.255.0***

  End with **CNTL/Z**.

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***ip default-gateway 172.23.199.2***

  End with **CNTL/Z**.

To authorize access on a switch for Device and Fabric Manager, run this command on every switch while supplying a user name (nnn) and password (ppp):

- Switch#***conf t***

  Switch(config)# ***snmp-server user nnn network-admin auth md5 ppp***

  Switch(config)#***end***

  Switch# ***copy running-config startup-config***

  Switch# ***exit***

Install Fabric Manager and Device Manager

To install Fabric Manager and Device Manager:

1. Open your web browser.
2. Enter the IP address of the switch into the address bar.
3. Follow the prompts and accept all defaults to install both Fabric Manager and Device Manager.

Fabric Manager and Device Manager can be started using the configured snmp-server user name and password in "CLI commands to configure the IP and gateway" on page 162.

Configure a VSAN

To configure a VSAN:

1. Open the Device Manager for the switch with an IP address of **172.23.199.22**.
2. Open the **VSAN** dialog by selecting the **VSAN** menu item.
3. Click **Create**.
4. Enter the value of **100** into the **VSAN ID** field.
5. Set the **VSAN Name** to be **"Red_VSAN_100"**.
6. Use the default interop mode.
7. Click **Create**.

Configure the other VSANs in this physical switch

1. Repeat Step 2 through Step 7 above, for VSAN 200 and VSAN 300 nothing that:
   a. For virtual switch 200, use **VSAN Name** "**Green_VSAN_200**".
   b. For virtual switch 300, use **VSAN Name** "**Blue_VSAN_300**".
2. Assign and enable the ports to the proper VSAN using Device Manager for the switch.
   a. Following Step 1 and Step 2, configure the ports of the switch with an IP address of 172.23.199.22 as shown in the tables below.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to SW 3	1
1	2	TE ISL to SW 4	1
1	3		
1	4		
1	5		
1	6		
1	7	TE ISL to SW 2	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to SW 3	1
2	2	TE ISL to SW 4	1
2	3	Red Storage 1	100
2	4	Green Storage 1	200
2	5	Red Host HBA 1	100
2	6	Green Host HBA 1	200
2	7	TE ISL to SW 2	1
2	8		

b. Following Step 1 and Step 2, configure the ports of the switch with an IP address of 172.23.199.23 as shown in the tables below.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to SW 3	1
1	2	TE ISL to SW 4	1
1	3	Blue Storage 1	300
1	4	Green Storage 3	200
1	5	Blue Host HBA 2	300
1	6		
1	7	TE ISL to SW 1	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to SW 3	1
2	2	TE ISL to SW 4	1
2	3	Blue Storage 1	300
2	4	Green Storage 3	200
2	5	Green Storage 4	200
2	6		
2	7	TE ISL to SW 1	1
2	8		

c. Following Step 1 and Step 2, configure the ports of the switch with an IP address of 172.23.200.22 as shown in the tables below.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to SW 1	1
1	2	TE ISL to SW 2	1
1	3		
1	4		
1	5		
1	6		
1	7	TE ISL to SW 4	1
1	8		

Slot #	Port #	Name	VSAN ID
2	2	TE ISL to SW 2	1
2	3	TE ISL to SW 4	1
2	4		
2	5	Red Host HBA 2	100
2	6	Green Host HBA 2	200
2	7	Red Storage 2	100
2	8	Green Storage 2	200

d. Following Step 1 and Step 2, configure the ports of the switch with an IP address of 172.23.200.23 as shown in the table below.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to SW 1	1
1	2	TE ISL to SW 2	1
1	3		
1	4		
1	5	Blue Host HBA 2	300
1	6		
1	7	TE ISL to SW 3	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to SW 1	1
2	2	TE ISL to SW 2	1
2	3	TE ISL to SW 3	1
2	4		
2	5		
2	6		
2	7	Blue Storage 2	300
2	8	Green Storage 4	200

Connect cables

To connect the cables:

1. Connect ISLs.

   a. Attach fiber cable between switches as shown in Figure 18 on page 159.

   b. After all cables are connected, use Fabric Manager to verify that all ISL connections are up.

   c. Re-arrange icons to accurately reflect the switch configuration.

Note: After persisting the fabric when a user looks at the topology view, they can immediately detect if something has changed in the environment. If an ISL or device was to disappear, yellow alert icons will be displayed. Because of this, it is a recommended to always persist the fabric after changes have been made.

2. Connect host and storage ports.

   a. Attach fibre cable between the switches and N_Ports.

Zone hosts and storage

To zone hosts and storage:

1. Open the **Zoning** menu in Fabric Manager for the desired VSAN.
2. Create a zone by clicking **New Zone** under **Zones**.
3. Provide a descriptive name for the zone. This example will zone “Red host HBA 1” and “Red Storage 1”. Type **“RedHBA1_1470_8aa”** and press **Enter**.
4. In zone "**RedHBA1_1470_8aa**" select **(WWPN 10000000c938e554)** select **add to zone**.
5. Select **“Red Storage 1” (WWPN 50060482cc19bf87)** in the **potential zone members list**.
6. Create the VSAN zoneset, add the zones, then activate the zoneset.
7. Repeat Step 2 through Step 6 for all host and storage pairs in the environment.

```
Zone set name = "VSAN Red 100"

          Zone name = "RedHBA1_1470_8aa"
             Zone Member = "10000000c938e554"
             Zone Member = "50060482cc19bf87"

          Zone name = "RedHBA2_1470_9aa"
             Zone Member = "10000000c938e555"
             Zone Member = "50060482cc19bf88"

Zone set name = "VSAN Green 200"

          Zone name = "GreenHBA1_1489_8aa"
             Zone Member = "210100e08b8ac76d"
             Zone Member = "50060482cc19c447"
```

```
            Zone name = "GreenHBA2_1489_9aa"
               Zone Member = "210100e08baac76d"
               Zone Member = "50060482cc19c448"

Zone set name = "VSAN Blue 300"

            Zone name = "BlueHBA1_AllBlueStorage"
               Zone Member = "10000000c939a051"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"

            Zone name = "BlueHBA2_AllBlueStorage"

               Zone Member = "10000000c939a052"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the Fabric OS configuration guide for more detail.

Configure IVR with Network Address Translation (NAT)

The configuration of VSANs on a fabric allows for security, scalability and availability. However, this isolation of traffic between VSANs prevents users from accessing resources, such as tape libraries, located in other VSANs. The solution to this limitation is Cisco's Inter-VSAN Routing feature, which allows initiators in one VSAN to access targets in other VSANs without merging the VSANs. Perhaps a host in the Red VSAN needs to access storage in the Blue VSAN. Configuring IVR zone and an IVR zone set containing the allowed initiators and targets allows communication between these resources.

This procedure also includes the configuration of IVR NAT, which allows duplicate Domain IDs to exist in the same fabric.

Without Network Address Translation (NAT), IVR requires unique Domain IDs for all switches in the fabric. You can enable IVR NAT to allow non-unique Domain IDs. This feature simplifies the deployment of IVR in an existing fabric where non-unique Domain IDs may be present.

Note: To use IVR NAT, it must be enabled in all IVR-enabled switches in the fabric IVR configuration distribution. By default, IVR NAT and IVR configuration distribution are disabled in all switches fin the Cisco MDS 9000 Family.

1. In Fabric Manager:
   a. Click the **Zone** tab in upper tool bar.
   b. Click the **IVR** tab.
   c. Select **Wizard**.

   To migrate to IVR NAT mode click **Yes**; otherwise click **No**. You see the **IVR Zone Wizard**.

   Note: If you are not using IVR NAT, Fabric Manager may display an error message if all the switches participating in IVR do not have unique Domain IDs. You must reconfigure those switches before configuring IVR.

2. Select the VSANs that will participate in IVR in the fabric.

   Select **VSAN 100**, **200** and **300**.

3. Select the end devices that you want to communicate over IVR.

   Select the following:

   VSAN 100: 10000000c938e554

   VSAN 200: 50060482cc19c407

   VSAN300: 50060482cc19c447

   VSAN 300: 210100e08b8ac76d

4. Enter the VSAN ID of the VSAN you want to use as the transit VSAN between the VSANs selected for the IVR zone.

   Select **VSAN 1** as the transit VSAN.

   Note: VSAN 1 connects both switches and all trunking traffic will pass over this link to communitcate with VSANs in other switches.

5. Set the IVR zone and IVR zone set.

   IVR NAME = **IVRZONE1**

   IVR ZONENET NAME = **IVRZONESET1**

6. Verify all steps that Fabric Manager will take to configure IVR in the fabric.
7. Click **Finish** if you want to enable IVR NAT and IVR topology and to create the associated IVR zones and IVR zone set.

   or

   Click **Cancel** to exit the IVR Wizard without saving any changes.
8. The **Save Configuration** dialog displays. You can save the configuration of the master witch to be copied to other IVR-enabled switches.

   Click either **Continue Activation** or **Cancel**.
9. Click **Finish**.

Connectrix M example

General layout

Figure 19 shows a four ED-140M full mesh configuration. The colors indicate which hosts access which storage.

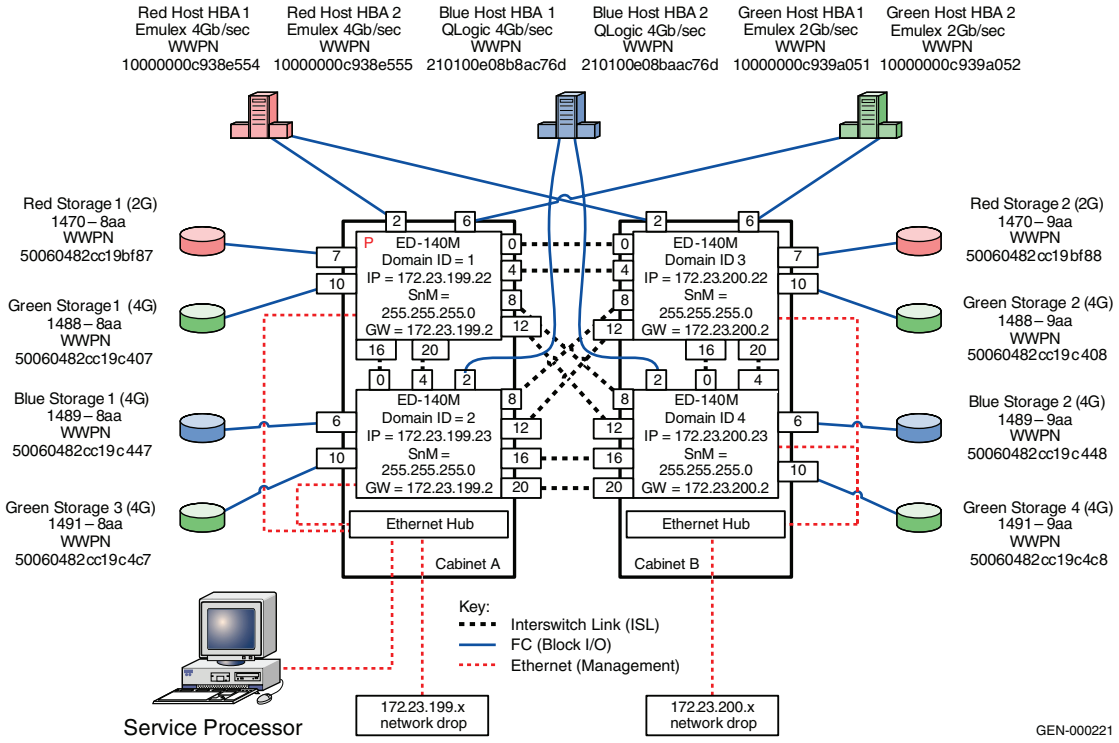


Figure 19 **Four ED-140Ms full mesh configuration**

All switches can be used in this configuration. However, due to the large number of ports consumed by ISLs, it is not recommended to use switches with 16 ports or less in this topology.

Best practices

For general information on best practices, refer to "Best practices" on page 37.

Host and storage layout

In this example, ED-140Ms are being used. As with all Connectrix M directors, the switch does not support cut-through routing. Because of this, there is no performance benefit to arranging the host and storage ports such that they reside on the same port card. However, if all host and storage pairs were one-to-one, there could be a serviceability benefit realized from this arrangement. This is due to

the fact that if it is arranged this way and a port card needed to be replaced, only two host and storage pairs would be impacted, whereas if each port card has connections from a different host and storage pair, up to four pairs could be impacted. This point is raised only for the purposes of having the reader consider their particular situation to see if there are serviceability benefits from a particular arrangement, even if there may be no performance benefits realized.

For general information on host and storage layout, refer to "Host and storage layout" on page 43.

Switch and fabric management

CLI, Connectrix Manager Basic, and Connectrix Manager can all be used to set up this configuration. In this example, Connectrix Manager will be used. For general information on switch and fabric management, refer to "Switch and fabric management" on page 44.

Security

For general information on security, refer to "Security" on page 47.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual*, which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix C, "Non-EMC Products Data."
  - For Cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and setup Manual*, which can be accessed from Powerlink.
- The switches have *not* been connected to the power source and are *not* powered on.

Note: Connectrix M switches ship with a default IP address of 10.1.1.10. If all of the switches are powered on and connected to the hub in the switch cabinet, communication with the switch using the network interfaces will *not* be possible and you will have to program the IP addresses using the serial port.

This section assumes that the serial port will not be used and the IP addresses will be programmed through the network interface.

- Network drops, IP addresses, Subnet mask and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix C, "Non-EMC Products Data."

  Note: Connectrix M switches can be either attached to the Ethernet hub that comes with the cabinet, or directly connected to the customer's LAN. In both cases, the switches can be placed on either a public or private network. There are advantages to both configurations. For more information, refer to "Public versus private" on page 961.

  This example assumes that The customer has provided us with 2 Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network.

- The proper number of port cards have been installed into the ED-140Ms.
  - For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- A Service Processor is installed in at least one of the cabinets.
- The Service Processor will be used to configure the IP addresses of the switches.
- One interface from the Service Processor is connected to the Ethernet hub and this interface has an IP address of 10.1.1.11. (Check **Default** on the SP image and use that value.)
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- Connectrix Manager will be used and the Connectrix Manager installation kit is available.
- Connectrix Manager is not already installed on the Service Processor.

Configure the IP address

To configure the IP address:

1. In the cabinet that contains the Service Processor, select one of the switches to configure.

Note: This example will start with the switch that will have a Domain ID of **1** and an IP address of **172.23.199.2**2.

2. Disconnect the network cables from the other switches in the cabinet.

CAUTION

Failure to do so will cause an IP address conflict when the cabinet is powered up.

Note: The other cabinet should not be powered up yet, but if it is and a network cable is connected between the two hubs, the network cables from the other switches *must be disconnected* as well to avoid running into duplicate IP address types of problems.

3. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer. This will cause both the Directors and the Service Processor to power up.
4. Log in to the Service Processor once it has powered up. The default username and password is *administrator* and *password*.
5. Open a command prompt and ping the IP address of 10.1.1.10. If you get a response, continue to Step 6. If you do not get a response, you can either IML the switch by pushing the little gray button on the CTP or configure the IP address using the serial port.

Note: See the *Installation and Service Guide* for more information about configuring the IP address using the serial port.

6. Launch Internet Explorer and type the IP address of **10.1.1.1**0 into the address bar and press **Enter**. The **Login** banner will be displayed.
7. Click **Accept** and the **Login** dialog will be displayed.
8. Enter the default username and password of *Administrator* and *password*, then press **Enter**.

IMPORTANT

It is strongly recommended that when prompted to change the password, you change it to a password that was provided by the customer.

9. Select the **Details** button underneath the switch icon in the Connectrix Manager Basic Topology view.
10. Navigate to the **Network Configuration** dialog by clicking the **Configure** pull-down menu and selecting **switch,** then **Network**.
11. Enter the IP address (**172.23.199.22**), subnet mask (**255.255.255.0**) and gateway (**172.23.199.2**), then click **OK**.

Note: The error message *unable to display* appears. This appears because the IP address of the switch has just changed and you are no longer able to connect to the switch at its old IP address.

Configure the rest of the switch IP addresses in the cabinet

To configure the next switch in the cabinet, reconnect the network cables to the CTPs and then follow steps 5 – 11 in "Configure the IP address" on page 173. Use the IP address of (172.23.199.23), subnet mask (255.255.255.0) and gateway (172.23.199.2).

Configure the IP addresses of the switches in the other cabinet

To configure the switch in the other cabinet, a cable can be connected between the two hubs to allow for communication between the service processor in one cabinet (cabinet A in our example) and the switches in the other cabinet (cabinet B in our example). This technique will be used in this section.

1. Connect a straight through cable to port 24 on the cabinet B. Ensure that MDIX is set on port 24 by pushing the recessed MDI / MDIX button in.
2. Connect the other end of cable connected to port 24 in step 1 to any port other than 24 on the Ethernet hub in cabinet A.
3. Follow Step 5 through Step 11 in "Configure the IP address" on page 173.
   a. For the top switch in cabinet B, use the IP address of (**172.23.200.22**), subnet mask (**255.255.255.0**) and gateway (**172.23.200.2**).

b. For the bottom switch in cabinet B, use the IP address of (**172.23.200.23**), subnet mask (**255.255.255.0**) and gateway (**172.23.200.2**).

Install Connectrix Manager

Note: Refer to the *Connectrix Manager User Guide* for complete installation instructions.

To install Connectrix Manager:

1. Insert the Connectrix Manager installation CD into the CD-RW drive.
2. Browse to the CD-RW drive in windows explorer.
3. Locate and double-click, the **install.exe** file. The installer is displayed. Typically, this file is located under the **CtxMgr** <*version number*> folder.
4. Accept all the defaults by clicking **next** and **OK**, when appropriate.
5. Click **Done** when the **Installation Complete** dialog appears.
6. Click **Next** when the **Welcome** dialog appears.
7. Accept the EULA terms and then click **Next**.
8. At the **Copy data** and **Settings** dialog, select **No,** then **Next**.

   Note: "**No**" is selected in this step since this section assumes that this is a new installation.

9. Assign a Connectrix Manager 9.0 Server Name and click **Next**.

   This case will use **CMServer.**

10. Enter the serial number and license key in the provided fields in the **Connectrix Manager 9.0 Server License** dialog.
    - The serial number is located on the back of the Connectrix Manager installation CD case.
    - To obtain a license key, locate the transaction code certificate and go to the URL listed.
      a. Enter the serial number located on the back of the Connectrix Manager installation CD case.

b. In the **Transaction Code** fields, type the transaction code(s) shipped with the software.

c. Click **Next**.

d. Confirm the existing and new features to be enabled.

e. Click **Next**.

The license key and all enabled features appears.

f. Retain a copy for your records.

g. Enter this key into the **License key** field.

h. Click **OK**.

11. Once the installation is complete, log in to Connectrix Manager with the username of *administrator* and the password of *password*.

Note: Change this default as soon as possible. For this example, the username *nnn* and password *nnn* will be used.

Manage the switches

To manage the switches:

1. From Connectrix Manager topology view, select the **Setup** from the **Discover** pull-down menu. The **Discover Setup** dialog is displayed.
2. Click the **Out-of-Band** tab.
3. Under **Available Addresses**, click **Add...** .
4. **In the IP Address field,** enter **172.23.199.22**.
5. In the **Subnet Mask** field, enter **255.255.255.0**.
6. Click **OK**.
7. Repeat steps Step 4 through Step 6 for the IP addresses of **172.23.199.23**, **172.23.200.22**, and **172.23.200.23**.
8. Ensure that the switches are highlighted in the **Available Addresses** list and then click the bottom right-pointing arrow. The IP addresses of **172.23.199.22, 172.23.200.22,** and **172.23.200.23** should be shown in the **Selected Individual Addresses** list.
9. Click **OK**.

Configure Fibre Channel switches

To configure Fibre Channel switches:

1. Set the switch name and Domain ID.

   a. From Connectrix Manager, double click the switch icon for the switch with an IP address of **172.23.199.22**. The **Element Manager** is displayed.

   b. Under **Configure**, select **Operating parameters menu**.

   Note: Each switch must have its ports configured in a certain way. Step 2 provides four examples.

   c. Select the **Domain** tab and ensure that the preferred Domain ID is set to the appropriate number (**1** in this example).

   d. Check the **insistent Domain ID** checkbox to enable insistent Domain IDs.

   Note: If the preferred Domain ID and the active Domain ID do not match, you will be *not* be able to set the preferred Domain ID and enable insistent Domain IDs at the same time. Change the **preferred Domain ID** first, click **OK**, re-open the dialog, and then check the **insistent Domain ID** checkbox and click **OK**.

   e. Select the **Identification** tab.

   f. Enter the switch name in the **Name** field.

   g. Select the **Copy** button to make the switch name, the switch nickname.

   h. Select the **Fabric** tab.

   i. Ensure that the Interop Mode is set to **Open Fabric 1.0**.

   j. *Follow this step only for the switch that has been selected to be the Principal switch* (**172.23.199.22** in this example).

      Set **Switch Priority** to **principle.**

   k. Click **OK**.

   l. Repeat Step b through Step i for each switch being configured.

2. Configure the switch ports.

Note: In the following examples all of the port cards are QPMs, which means they are port cards capable of running at 4 GB/s line rates. Due to the hardware architecture of the QPM, only two ports on any QPM are capable of sustaining the full 4 GB/s rate. Due to FSPF concerns, only ports configured for 4 GB/s sustained can be used for 4 GB/s ISLs. See Appendix C, "Non-EMC Products Data." for additional detail on the QPM architecture, its limitations, and applications.

Example 1: IP Address 172.23.199.22

Note: In these examples, 172.23.199.22 is the **principal** switch.

To configure the switch port with an IP address of **172.23.199.22**:

a. Open the Element Manager of the switch with an IP address of **172.23.199.22** by double-clicking its icon.

b. From the **Configure** menu, select **ports menu**.

c. Configure the ports as shown in the table below.

Port	Name	Speed
Port 0	ISL to Port 0 on Domain 3	Negotiate Sustained
Port 1	Inactive due to Negotiate Sustained setting on Port 0	
Port 2	Red HBA 1	Negotiate Sustained
Port 3	Inactive due to Negotiate Sustained setting on Port 2	
Port 4	ISL2 to Port 4 on Domain 3	Negotiate Sustained
Port 5	Inactive due to Negotiate Sustained setting on Port 4	
Port 6	Green HBA 1	Negotiate 2 Max
Port 7	Red Storage 1	Negotiate 2 Max
Port 8	ISL1 to Port 0 on Domain 4	Negotiate Sustained
Port 9	Inactive due to Negotiate Sustained setting on Port 8	
Port 10	Green Storage 1	Negotiate Sustained
Port 11	Inactive due to Negotiate Sustained setting on Port 10	
Port 12	ISL2 to Port 4 on Domain 4	Negotiate Sustained
Port 13	Inactive due to Negotiate Sustained setting on Port 12	
Port 14		

Port	Name	Speed
Port 15		
Port 16	ISL1 to Port 0 on Domain 2	Negotiate Sustained
Port 17	Inactive due to Negotiate Sustained setting on Port 16	
Port 18		
Port 19		
Port 20	ISL2 to Port 4 on Domain 2	Negotiate Sustained
Port 21	Inactive due to Negotiate Sustained setting on Port 20	

Example 2: IP Address 172.23.199.23

To configure the switch port with an IP address of **172.23.199.23**:

a. Open the Element Manager of the switch with an IP address of **172.23.199.23** by double-clicking its icon.

b. From the **Configure** menu, select **ports menu**.

c. Configure the ports of the switch with an IP address of **172.23.199.23** as shown in the table below.

Port	Name	Speed
Port 0	ISL to Port 16 on Domain 1	Negotiate Sustained
Port 1	Inactive due to Negotiate Sustained setting on Port 0	
Port 2	Blue HBA 1	Negotiate Sustained
Port 3	Inactive due to Negotiate Sustained setting on Port 2	
Port 4	ISL2 to Port 20 on Domain 1	Negotiate Sustained
Port 5	Inactive due to Negotiate Sustained setting on Port 4	
Port 6	Blue Storage 1	Negotiate Sustained
Port 7	Inactive due to Negotiate Sustained setting on Port 6	
Port 8	ISL1 to Port 8 on Domain 3	Negotiate Sustained
Port 9	Inactive due to Negotiate Sustained setting on Port 8	
Port 10	Green Storage 3	Negotiate Sustained
Port 11	Inactive due to Negotiate Sustained setting on Port 10	

Port	Name	Speed
Port 12	ISL2 to Port 12 on Domain 3	Negotiate Sustained
Port 13	Inactive due to Negotiate Sustained setting on Port 12	
Port 14		
Port 15		
Port 16	ISL1 to Port 16 on Domain 4	Negotiate Sustained
Port 17	Inactive due to Negotiate Sustained setting on Port 16	
Port 18		
Port 19		
Port 20	ISL2 to Port 20 on Domain 4	Negotiate Sustained
Port 21	Inactive due to Negotiate Sustained setting on Port 20	

Example 3: IP Address 172.23.200.22

To configure the switch port with an IP address of **172.23.200.22**:

a. Open the Element Manager of the switch with an IP address of **172.23.199.23** by double-clicking its icon.

b. From the **Configure** menu, select **ports menu**.

c. Configure the ports of the switch with an IP address of **172.23.200.22** as shown in the table below.

Port	Name	Speed
Port 0	ISL to Port 0 on Domain 1	Negotiate Sustained
Port 1	Inactive due to Negotiate Sustained setting on Port 0	
Port 2	Red HBA 2	Negotiate Sustained
Port 3	Inactive due to Negotiate Sustained setting on Port 2	
Port 4	ISL2 to Port 4 on Domain 1	Negotiate Sustained
Port 5	Inactive due to Negotiate Sustained setting on Port 4	
Port 6	Green HBA 2	Negotiate 2 Max
Port 7	Red Storage 2	Negotiate 2 Max
Port 8	ISL1 to Port 8 on Domain 2	Negotiate Sustained

Port	Name	Speed
Port 9	Inactive due to Negotiate Sustained setting on Port 8	
Port 10	Green Storage 2	Negotiate Sustained
Port 11	Inactive due to Negotiate Sustained setting on Port 10	
Port 12	ISL2 to Port 12 on Domain 2	Negotiate Sustained
Port 13	Inactive due to Negotiate Sustained setting on Port 8	
Port 14		
Port 15		
Port 16	ISL1 to Port 16 on Domain 4	Negotiate Sustained
Port 17	Inactive due to Negotiate Sustained setting on Port 16	
Port 18		
Port 19		
Port 20	ISL2 to Port 12 on Domain 4	Negotiate Sustained
Port 21	Inactive due to Negotiate Sustained setting on Port 20	

Example 4: IP Address 172.23.200.23

To configure the switch port with an IP address of **172.23.200.23**:

a. Open the Element Manager of the switch with an IP address of **172.23.199.23** by double-clicking its icon.

b. From the **Configure** menu, select **ports menu**.

c. Configure the ports of the switch with an IP address of **172.23.200.23** as shown in the table below.

Port	Name	Speed
Port 0	ISL to Port 8 on Domain 1	Negotiate Sustained
Port 1	Inactive due to Negotiate Sustained setting on Port 0	
Port 2	Blue HBA 2	Negotiate Sustained
Port 3	Inactive due to Negotiate Sustained setting on Port 2	
Port 4	ISL2 to Port 12 on Domain 1	Negotiate Sustained
Port 5	Inactive due to Negotiate Sustained setting on Port 4	

Port	Name	Speed
Port 6	Blue Storage 2	Negotiate Sustained
Port 7	Inactive due to Negotiate Sustained setting on Port 6	
Port 8	ISL1 to Port 16 on Domain 3	Negotiate Sustained
Port 9	Inactive due to Negotiate Sustained setting on Port 8	
Port 10	Green Storage 4	Negotiate Sustained
Port 11	Inactive due to Negotiate Sustained setting on Port 10	
Port 12	ISL2 to Port 20 on Domain 3	Negotiate Sustained
Port 13	Inactive due to Negotiate Sustained setting on Port 8	
Port 14		
Port 15		
Port 16	ISL1 to Port 16 on Domain 4	Negotiate Sustained
Port 17	Inactive due to Negotiate Sustained setting on Port 16	
Port 18		
Port 19		
Port 20	ISL2 to Port 20 on Domain 2	Negotiate Sustained
Port 21	Inactive due to Negotiate Sustained setting on Port 20	

Connect cables

To connect cables:

1. Connect ISLs.
   a. Attach fiber cable between switches as shown in Figure 19 on page 171.
   b. After all cables are connected, use Connectrix Manager to verify that all ISL connections are up.
   c. Rearrange icons to accurately reflect the switch configuration and then persist the fabric.

Note: After persisting the fabric, when a user looks at the topology view they can immediately detect if something has changed in the environment. If an ISL or device was to disappear, yellow alert icons appears. Because of this, it is a recommended to *always* persist the fabric *after* changes have been made.

2. Connect the host and storage ports.
   a. Attach fiber cable between switches and N_Ports.

Zone hosts and storage

To zone hosts and storage:

1. Open the **Zoning** dialog in Connectrix Manager by right-clicking the appropriate fabric topology and selecting **zoning menu**.
2. Create a zone by clicking the **New Zone** button under the zones set **Tree**.
3. Provide a descriptive name for the zone. This case will zone *Red host HBA* and *Red Storage 1*, so **RedHBA1_1470_8aa** will be entered. Press **Enter**.
4. Locate, then click, **Red Host HBA 1** (WWPN 10000000c938e554) in the **potential zone members** list.
5. Click the right-pointing arrow on the divider between the **potential members** list and the **zones** list to add the HBA to the zone.
6. Locate, then click. **Red Storage 1** (WWPN 50060482cc19bf87) in the **potential zone members** list.
7. Click the right-pointing arrow on the divider between the **potential members** list and the **zones** list to add the Storage port to the zone.
8. Repeat steps 2 through 7 for all host and storage pairs in the environment.
9. Create a zone set by clicking **New Set** under the zone sets **Tree**.
10. Provide a descriptive name for the zone set. This case will use the date of “Oct_31_06_1140”.
11. Add all of the new zones to the zone set. When completed, the zone set should be similar to what is shown next.

```
Zone set name = "Oct_31_06_1140"
       Zone name = "RedHBA1_1470_8aa"
          Zone Member = "10000000c938e554"
          Zone Member = "50060482cc19bf87"

       Zone name = "RedHBA2_1470_9aa"
          Zone Member = "10000000c938e555"
          Zone Member = "50060482cc19bf88"

       Zone name = "BlueHBA1_1489_8aa"
          Zone Member = "210100e08b8ac76d"
          Zone Member = "50060482cc19c447"

       Zone name = "BlueHBA2_1489_9aa"
          Zone Member = "210100e08baac76d"
          Zone Member = "50060482cc19c448"

       Zone name = "GreenHBA1_AllGreenStorage"
          Zone Member = "10000000c939a051"
          Zone Member = "50060482cc19c407"
          Zone Member = "50060482cc19c408"
          Zone Member = "50060482cc19c4c7"
          Zone Member = "50060482cc19c4c8"

       Zone name = "GreenHBA2_GreenStorage"
          Zone Member = "10000000c939a052"
          Zone Member = "50060482cc19c407"
          Zone Member = "50060482cc19c408"
          Zone Member = "50060482cc19c4c7"
          Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the Fabric OS configuration guide for more detail.

Configure security

Once the hosts and storage ports have been properly configured, the fabric should be secured.

Enable Enterprise Fabric Mode:

1. From the **Connectrix Manage**r topology view, single-click a switch in the fabric that needs to be secured.

   Note: Do *not* double-click and open the Element Manager.

2. Select **Enterprise Fabric Mode** under the **Configure** pull-down menu.
3. Click **Activate**.
4. Click **Close**.

Enable Port binding:

1. From Connectrix Manager, double-click the switch with an IP address of **172.23.199.22** and its Element Manager will be displayed.
2. From the **Configure** pull-down menu, select **Ports**.
3. Right-click in the **Port Binding** column and select **Bind All Ports To Attached WWN**.
4. Click **OK**.
5. Repeat for other switches in the fabric.

Configure proactive monitoring and countermeasures

Configure throughput threshold alerts:

To configure threshold alerts:

1. Open a command prompt and Telnet to the switch by typing **telnet 172.23.199.22**.
2. Log in with the username and password of the switch.
3. Enter the **Perf** command.
4. Enter the **thresh** command.
5. Enter the **throughput** command.
6. Enter the **addalert eightyPercent** command.
7. Enter the **addport eightyPercent all** command.
8. Enter the **setParams eightyPercent 1 5** command.
9. Enter the **setUtilPercentage eightyPercent 80** command.
10. Enter the **setutiltype eightyPercent 3** command.
11. Enter the **..** command.
12. Enter the **setstate eightyPercent 1** command.
13. Enter the **logo** command.

Configure counter threshold alerts

Telnet into each switch and enter the following commands:

1. Enter the **addAlert CTA1** command.
2. Enter the **addPort CTA1 all** command.
3. Enter the **setCounter CTA1 18** command.
4. Enter the **setParams CTA1 10000 30** command.
5. Enter the **addAlert CTA2** command.
6. Enter the **addPort CTA2 all** command.
7. Enter the **setCounter CTA2 10** command.
8. Enter the **setParams CTA2 100 300** command.
9. Enter the **addAlert CTA3** command.
10. Enter the **addPort CTA3 all** command.
11. Enter the **setCounter CTA3 9** command.
12. Enter the **setParams CTA3 40 300** command.
13. Enter the **addAlert CTA4** command.
14. Enter the **addPort CTA4 all** command.
15. Enter the **setCounter CTA4 9** command.
16. Enter the **setParams CTA4 100 70560** command.

Enable the Counter threshold alerts:

1. Enter the **..** command.
2. Enter the **setState CTA1 1** command.
3. Enter the **setState CTA2 1** command.
4. Enter the **setState CTA3 1** command.
5. Enter the **setState CTA4 1** command.

To configure Port Fencing:

1. Click **Configure** and select **Port Fencing** from the list.
2. Enable the **Default Security Policy** by highlighting the entry and clicking **Enable**.
3. Enable the **Default Link Level Policy** by highlighting the entry and clicking **Enable**.

Disable unused interfaces:

1. From the **Connectrix Manager** topology view, select the **Security** tab.
2. Ensure that the appropriate fabric is selected in the fabrics list.
3. Under the **Authentication** tab, select a switch in the **Product Configuration** list.
4. Under the **Users** tab uncheck the **Enable Telnet and Enable Web Sever** checkbox.

Configure the other switch:

Repeat all the steps in the "Configure proactive monitoring and countermeasures" on page 186 to configure the other switch with the IP address **172.23.199.23**, **172172.23.200.22**, and **172.23.200.23**.

QLogic example

General layout

Figure 20 uses four SB9000s in a full mesh configuration.

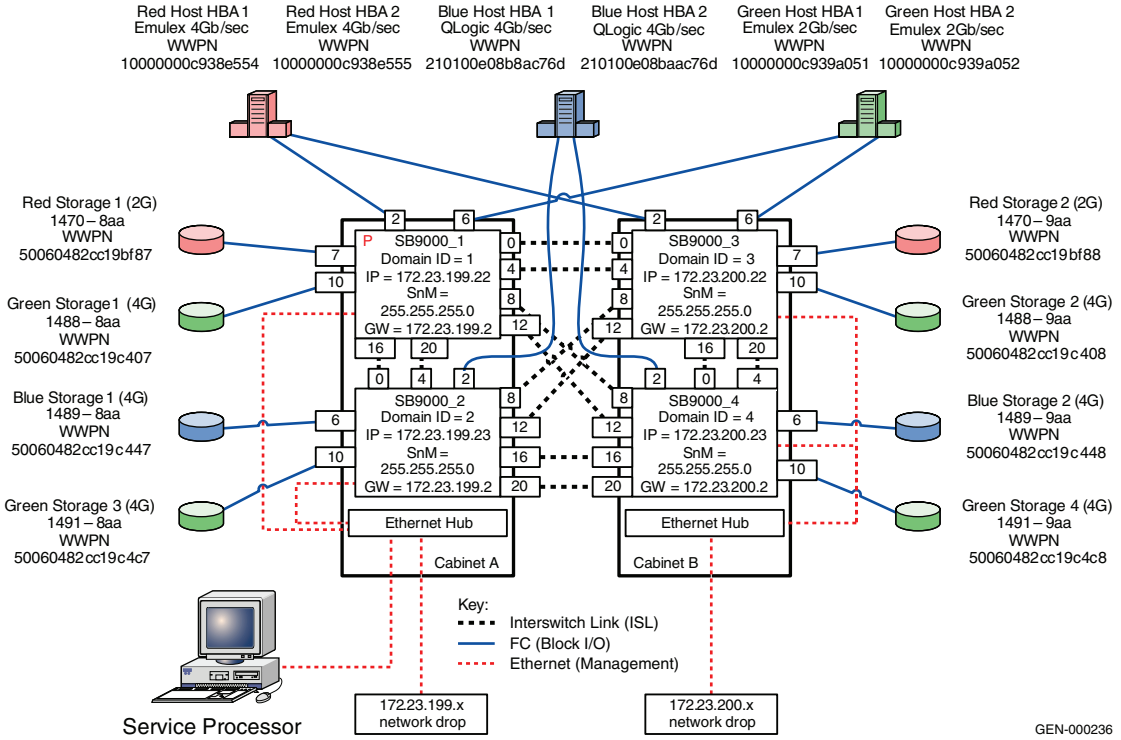


Figure 20 Four SB9000s in full mesh configuration

Best practices

For general information on best practices, refer to "Best practices" on page 37.

Host and storage layout

In this example, SANbox 9000s are being used. There is no *performance* benefit to arranging the host and storage ports so that they reside on the same I/O card. However, if all host and storage pairs were one-to-one (that is, each zoned pair is on the same I/O card, and the different zoned pairs are on different I/O cards) there could be a *serviceability* benefit. Arranged this way, if a port card needs to be replaced, only *two* host and storage pairs will be impacted, whereas if each port card has connections from a different host and storage pair, up to *four* pairs will be impacted. Consider your particular situation to see if there are serviceability benefits from any particular arrangement, even if no performance benefits are realized.

For general information on host and storage layout, refer to "Host and storage layout" on page 43.

Switch and fabric management

For general information on switch and fabric management, refer to "Switch and fabric management" on page 44.

Security

For QLogic security information that may be applicable, refer to "Security" on page 119.

For general information on security, refer to "Security" on page 47.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet
  - For installation instructions, refer to the *Connectrix EC-1500 Cabinet Installation and setup Manual*, available at Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data".
  - For Cabinet power requirements, refer to the *Connectrix EC-1500 Cabinet Installation and setup Manual*, available at Powerlink
- The switches have *not* been connected to the power source and are *not* powered on.

Note: QLogic switches ship with a default IP address of 10.0.0.1. If all of the switches are powered on and connected to the hub in the switch cabinet, communication with the switch using the network interfaces will not be possible and you will have to program the IP Addresses using the serial port. This example assumes that the serial port will not be used and the IP Addresses will be programmed using the network interface.

- Network drops, IP Addresses, Subnet mask, and gateway, have been provided by the customer.
  - For switch or cabinet network requirements, refer to Appendix B, "EMC Connectrix Products Data."

Note: QLogic switches can be either attached to the Ethernet hub that comes with the cabinet, or directly connected to the customer's LAN. In both cases, the switches can be placed on either a public or private network. There are advantages to both configurations. For a complete discussion, refer to "Public versus private" on page 961.

In this example, it is assumed that the customer has provided two Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network.

- License keys have been obtained.
  - To obtain a license key, go to the URL listed on the transaction code certificate that shipped with the product.
- A Windows workstation will be used to configure the IP Addresses of the switches. The Workstation must be configured for an IP Address of 10.1.1.1.253.
- A null modem cable is available.
- The customer has provided a temporary password that will be used as the default password when configuring the IP Address.
- Enterprise Fabric Suite 2007 (EFS 2007) will be used to manage this fabric and the installation kit for this management application is available on a disk that comes with the SB90000 product.
- EFS 2007 is not already installed on the workstation.

Install the Enterprise Fabric Suite 2007

To install the Enterprise Fabric Suite 2007:

1. Close all the applications currently running on the management workstation and insert the **Enterprise Fabric Suite 2007 Installation Disk** into the CD-ROM drive.
2. In the upper left corner of the **Product information** screen, click **Management Software**.
3. Locate your platform in the table and click **Install**.

Configure the IP address, Domain ID, and switch names

To configure the IP address, Domain ID, and switch names:

1. Select one of the switches to configure.

   Note: For this example, the switch that will have a Domain ID of **1** and an IP Address of **172.23.199.22** will be the first switch to be configured.

2. Disconnect the network cables from the other switches in the cabinet. Failure to do so will cause an IP Address conflict when the cabinet is powered up.

   Note: The other cabinet should not be powered up, but if it is and a network cable is connected between the two hubs, remember that the network cables from the other switches must be disconnected as well to avoid running into duplicate IP Address problems.

3. Verify that one interface from the customer-supplied workstation is directly connected to the FC switch using a crossover cable.
4. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer. This will cause both SB9000 switches in that cabinet to power up.
5. Log in to the workstation once it has powered up.
6. Open a command prompt and ping the IP Address**10.0.0.10**.
   - If you get a response, continue at Step 7.
   - If you do not get a response, you can either hard reset the switch and verify the SB9000 chassis status by looking at the maintenance panel or configure the IP Address using the serial port.

Note: See the *QLogic SB9000 Installation and Service Guide* for more information about configuring the IP Address using the serial port.

7. Double click the EFS2007 shortcut, once installed.
8. Enter the default username and password of *admin* and *password*, then click **Enter**.

IMPORTANT

It is strongly recommended that when prompted to change the password, you change it to a password that was provided by the Customer.

9. In the **Initial start** dialog, click **Open Configuration Wizard**. The **Configuration Wizard** recognizes the switch and leads the end user into the configuration process.
10. Upon clicking **Proceed**, the end user is prompted to enter the following configuration information, which is filled as specified based on the desired topology:
    - Switch Domain ID (1-239): 1
    - Domain ID lock (locked / unlocked): locked
    - Switch name: SB9000_1
    - Permanent IP address: 172.23.199.22
    - Permanent subnet mask: 255.255.255.0
    - Permanent gateway address: 172.23.199.2
    - Admin account password: <this can be changed if required>

Note: At this point, you may lose access to the switch. This is because the IP Address of the switch just changed and you are no longer able to connect to the switch at its old IP Address. However, remote login is now enabled for this switch.

Configure the rest of the switch IP addresses in the cabinet

To configure the next switch in the cabinet, re-connect the cross-over network cable between the other SB9000 in the cabinet and the Windows workstation and then follow Step 5 through Step 10 in the previous "Configure the IP address, Domain ID, and switch names"" section. Use the IP Address of (172.23.199.23), subnet mask (255.255.255.0), gateway (172.23.199.2), Domain ID (2) and switch name (SB9000_2).

Configure the IP addresses, Domain IDs, and switch names of the switches in the other cabinet

Note: To configure the switch in the other cabinet, a cable can be connected between the two hubs to allow for communication between the service processor in one cabinet (cabinet A in this example) and the switches in the other cabinet (cabinet B in this example). This example uses this technique.

1. Connect a straight through cable to port 24 on the cabinet B. Ensure that MDIX is set on port 24. By pushing the recessed MDI / MDIX button in.
2. Connect the other end of cable connected to port 24 in Step 1 to any port, other than 24, on the Ethernet hub in cabinet A.
3. Follow Step 5 through Step 10 in the previous "Configure the IP address, Domain ID, and switch names" on page 191".
   - For the top switch in cabinet B, use the IP Address of (172.23.200.22), subnet mask (255.255.255.0), gateway (172.23.200.2), Domain ID (3) and switch name (SB9000_3).
   - For the bottom switch in cabinet B, use the IP Address of (172.23.200.23), subnet mask (255.255.255.0), gateway (172.23.200.2), Domain ID (4) and switch name (SB9000_4).

Configure FC switches interopmode and priority settings

To configure Fibre Channel switches interopmode and priority settings:

1. Select a workstation on the local network that can remotely access all the above configured switches.
2. Install EFS2007 on this management workstation, as explained in "Install the Enterprise Fabric Suite 2007" on page 191.
3. Add the IPs of the 4 switches configured above. These switches will be displayed in a column on the left pane of the GUI.
4. Click on the switch configured first, with a Domain ID of **1**.
5. When the selected switch shows up as a faceplate display:
   a. Open the **Switch** menu on the top of the screen.
   b. Select **Advanced Switch Properties**.
   c. When the **Advanced Switch Properties** dialog displays, set the **Interop Mode** to **Standard**.

The switch will automatically be taken offline and restored once the changes (if any) are completed.

d. Follow this step only for the switch that has been selected to be the Principal switch (172.23.199.22 in this example).

– Set **Switch Principal Priority** to **1** using the CLI command **set switch config**.
– Repeat Step b and Step c for each switch being configured.

6. Configure the switch ports.

a. From the faceplate display, follow the steps below to change the symbolic port name:

– Open the faceplate display and select the respective port.
– Open the **Port** menu and select **Port Symbolic Name**.
– In the **Port symbolic name** dialog, enter a new name for the port in the **Set Port Symbolic Name** field as given in the tables shown in Step d.

b. The characteristics or port settings for the 1 Gb / 2 Gb / 4 Gb and 10 Gb ports are configured using the **Port Properties** setting. To open the corresponding type of **Port Properties** dialog, select one or more ports, open the **Port** menu and select **Port Properties**.

c. The **Port Properties** dialog can be used to change the following settings:

– Port state
– Port type
– Port speed

d. For this case study, the ports can be configured as per the following tables:

Switch 1: SB9000_1:

Port #	Symbolic port name	Port type	Port speed
0	ISL to SB9000_3	E_Port	AutoNeg.
2	Red Host HBA 1	F_Port	AutoNeg.
4	ISL to SB9000_3	E_Port	AutoNeg.
6	Green Host HBA 1	F_Port	AutoNeg.
7	Red Storage 1	F_Port	AutoNeg.
8	ISL to SB9000_4	E_Port	AutoNeg.

10	Green Storage 1	F_Port	AutoNeg.
12	ISL to SB9000_4	E_Port	AutoNeg.
16	ISL to SB9000_2	E_Port	AutoNeg.
20	ISL to SB9000_2	E_Port	AutoNeg.

Switch 2: SB9000_2:

Port #	Symbolic port name	Port type	Port speed
0	ISL to SB9000_1	E_Port	AutoNeg.
2	Blue Host HBA 1	F_Port	AutoNeg.
4	ISL to SB9000_1	E_Port	AutoNeg.
6	Blue Storage 1	F_Port	AutoNeg.
8	ISL to SB9000_3	E_Port	AutoNeg.
10	Green Storage 3	F_Port	AutoNeg.
12	ISL to SB9000_3	E_Port	AutoNeg.
16	ISL to SB9000_4	E_Port	AutoNeg.
20	ISL to SB9000_4	E_Port	AutoNeg.

Switch 3: SB9000_3:

Port #	Symbolic port name	Port type	Port speed
0	ISL to SB9000_1	E_Port	AutoNeg.
2	Blue Host HBA 2	F_Port	AutoNeg.
4	ISL to SB9000_1	E_Port	AutoNeg.
6	Green Host HBA 2	F_Port	AutoNeg.
7	Red Storage 2	F_Port	AutoNeg.
8	ISL to SB9000_2	E_Port	AutoNeg.
10	Green Storage 2	F_Port	AutoNeg.
12	ISL to SB9000_2	E_Port	AutoNeg.
16	ISL to SB9000_4	E_Port	AutoNeg.
20	ISL to SB9000_4	E_Port	AutoNeg.

Switch 4: SB9000_4:

Port #	Symbolic port name	Port type	Port speed
0	ISL to SB9000_3	E_Port	AutoNeg.
2	Blue Host HBA 2	F_Port	AutoNeg.
4	ISL to SB9000_3	E_Port	AutoNeg.
6	Blue Storage 2	F_Port	AutoNeg.
8	ISL to SB9000_1	E_Port	AutoNeg.
10	Green Storage 4	F_Port	AutoNeg.
12	ISL to SB9000_1	E_Port	AutoNeg.
16	ISL to SB9000_2	E_Port	AutoNeg.
20	ISL to SB9000_2	E_Port	AutoNeg.

Repeat Step 1 through Step 6 for the other three configured switches.

Connect cables

1. Connect ISLs.
   a. Attach the fiber cable between switches as shown in Figure 20 on page 188.
   b. After all cables are connected, use the EFS 2007 faceplate data maanger on teh lower portion of the interface to verify that all ISL connections are up .
   c. Re-arrange icons to accurately reflect the switch configuration.
2. Connect host and storage ports.
   a. Attach fiber cable between switches and N_Ports.

Zone hosts and storage

To zone hosts and storage:

1. Before creating any zones and enabling zoning, use the **Edit Zoning config** dialog to change the **Interop Auto save**, **Default Zone**, and **Discard inactive parameters**.
   - The **Interop Auto save** parameter (which determines whether any active zone changes that a switch receives from any other switch in the fabric must be saved to the zoning database of the switch) must be set to **False**.

- The **Default Zone** (which allows all N_ports to see each other in the absence of any active zones) must be set to **False**.
- The **Discard inactive parameter (**which automatically removes or pulls out inactive zones from an active zoneset) must be set to **False**.

2. Open the **Edit Zoning** toolbar from the **Zoning** menu on the faceplate display.
3. Create a zone by clicking **Create Zone**.
4. Provide a descriptive name for the zone. This example will zone “Red host HBA 1” and “Red Storage 1”. Type “**RedHBA1_1470_8aa**" and press **Enter**.
5. Locate and click **“Red Host HBA 1”** (WWPN 10000000c938e554) in the **Potential zone members** list.
6. Click **Add Member** to add the selected HBA to the zone created in Step 4.
7. Locate and click **“Red Storage 1”** (WWPN 50060482cc19bf87) in the **Potential zone members** list.
8. Repeat Step 3 through Step 7 for all host and storage pairs in the environment.
9. Create a zone set by clicking **Create Zone Set**.
10. Provide a descriptive name for the zone set. This example will use the date of “Oct_31_06_1140”.
11. Add all of the zones created in Step 4 through Step 8 to the zone set by dragging the zones listed to the left into the zoneset created in Step 10.
12. To activate this zoneset, select **Activate Zone Set** from the **Zoning Edit** menu. Select a zoneset from the **Select zoneset** pull-down menu and click **Activate**.

When completed, the active zone set should be similar to what is shown below.

```
Zone set name = "Oct_31_06_1140"
            Zone name = "RedHBA1_1470_8aa"
               Zone Member = "10000000c938e554"
               Zone Member = "50060482cc19bf87"

            Zone name = "RedHBA2_1470_9aa"
               Zone Member = "10000000c938e555"
               Zone Member = "50060482cc19bf88"

            Zone name = "BlueHBA1_1489_8aa"
               Zone Member = "210100e08b8ac76d"
               Zone Member = "50060482cc19c447"

            Zone name = "BlueHBA2_1489_9aa"
               Zone Member = "210100e08baac76d"
               Zone Member = "50060482cc19c448"

            Zone name = "GreenHBA1_AllGreenStorage"
               Zone Member = "10000000c939a051"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"

            Zone name = "GreenHBA2_AllGreenStorage"
               Zone Member = "10000000c939a052"
               Zone Member = "50060482cc19c407"
               Zone Member = "50060482cc19c408"
               Zone Member = "50060482cc19c4c7"
               Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

At this point, the SAN is ready to pass I/O from the host to storage. Other steps, such as configuring LUN Masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the OS configuration guide for more detail.

Compound core edge switches

This section provides examples of compound core edge switch topologies.

Overview of fabric design considerations

General layout

Figure 21 shows an example of a four switch compound core edge switches.

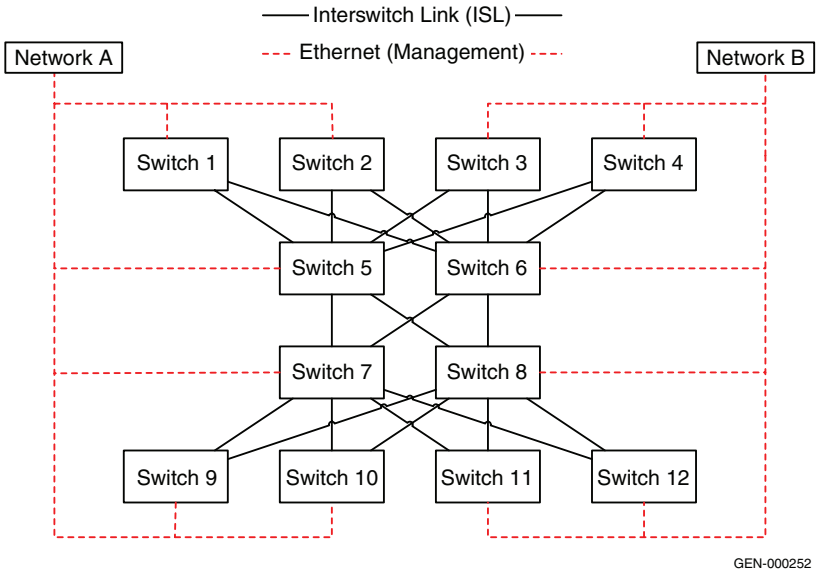


Figure 21 **Four switch compound core edge switches**

In the fabric shown in Figure 21, every core switch is connected to every other core switch while edge switches are only connected to two of the four cores. This is a classic core/edge design. Half of the switches are connected to Management Network A, and the other half are connected to Management Network B.

Only director class products should be used in the core of the fabric. This is due to the high number of ports that are consumed by ISLs and not due to other resource limitations (such as., CPU).

Both director class products and departmental class switches can be used in the edge position.

Best practices

Specific information on a compound core edge switch follows.

- Layout the host and storage connectivity such that if a switch fails, not all of a particular hosts storage becomes inaccessible.
- The use of two separate management networks is more common with balanced fabrics, but it can still be employed when only one fabric is used.
- ISL subscription best practice — While planning the SAN, keep track of the number of host and storage pairs that would be utilizing the ISLs between domains. As a general best practice, if two switches are connected by ISLs, ensure that there is a minimum of two ISLs between them, and that there are no more than six initiator and target pairs per ISL. For example, if 14 initiators access a total of 14 targets between two domains, a total of three ISLs are necessary. This best practice should not be applied blindly when setting up a configuration. Consider the applications that will use the ISLs.

For general information on best practices for all SANs, refer to "Best practices" on page 37.

Host and storage layout

Specific information a compound core edge switch follows.

In the examples that follow, host and storage pairs are located in the following locations:

- Host on the edge and storage on the edge (Red)
- Host on the core and storage on the core (Blue)
- Host on the core and storage on the edge (Green)

The decision on where to place these host and storage pairs was not arbitrary. These were deliberately placed in areas that may not typically be thought of as "good" places to attach host and storage ports. This was done to stress the point that the only good place to attach host and storage ports is where it makes the most sense, given the customer's environment.

For general information on host and storage layout for all SANs, refer to "Host and storage layout" on page 43.

Switch and fabric management

For general information on switch and fabric management for all SANs, refer to "Switch and fabric management" on page 44.

Security

For general information on security for all SANs, refer to "Security" on page 47.

Connectrix B example

Figure 22 shows four ED-48000Bs in a full mesh configuration with edge switches attached.

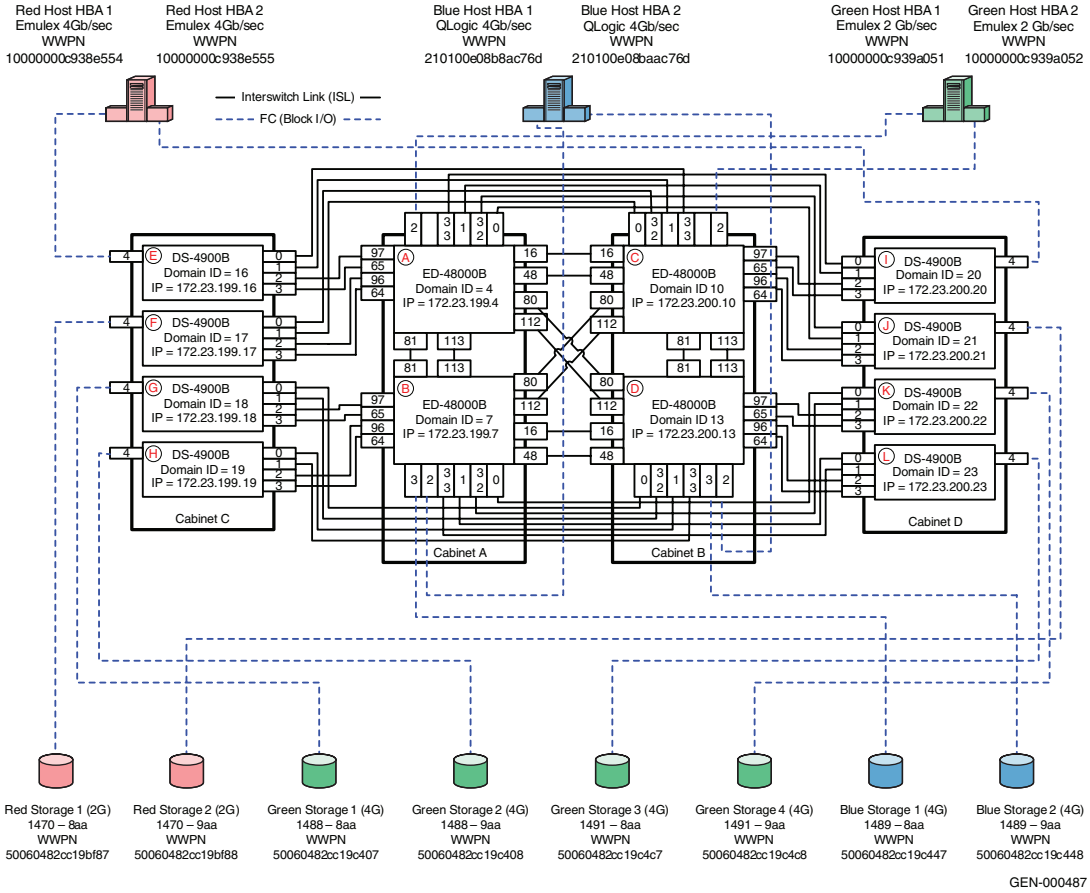


Figure 22 **Four ED-48000Bs in full mesh configuration with edge switches attached**

Note: Any director class Connectrix B product such as the ED-24000B or ED-48000B can be used in the core of this fabric. All EMC-supported Connectrix B director and switch class products can be used as edge switches.

Best practices

Specific information for this example is as follows:

While connecting the edge switches to the ED-48000B director switch it is recommended to be conversant with the architecture of the

ED-48000B, its local switching capability, and the over-subscription management of the different switch port blades (i.e., the supported 16-port, 32-port and 48-port blades) while designing scalable fabrics. An understanding of the product design will help utilizie the ports on the switch blades for switch or host/target connectivity, enabling the simultaneous uncongested ports on all ports as long as simple best practices are followed.

Please refer to the following link for further information on this subject:
http://www.brocade.com/products/competitive/directors.jsp

For general information on best practices for a compound core edge switch, refer to "Best practices" on page 200. For Connectrix B specific best practices, refer to "Connectrix B" on page 40.

Host and storage layout

To ensure fairness and increase the reliability of the SAN by eliminating any single point of overall failure and minimizing the impact that any single hardware failure would have, the N_Port and E_Port connections have been spread out on the ED-48000B over as many switch blade ports as possible.

The N_Port and E_Port connections on the DS-4900B are attached without any special considerations since there are no benefits to performance or reliability by spreading out the connections on the DS-4900B. In case the user wants to configure a trunk on the Connectrix B series switches, it is necessary that all ports in a given ISL trunk reside within an ASIC group on each end of the link. On 2 Gb/s switches, port groups are built on contiguous 4-port groups, called *quads*. On 4 Gb/s switches, like the Connectrix DS-48000B and the DS-4900B in this example, trunking port groups are built on contiguous 8-port groups, called *octets*. In these products, there are four octets: ports 0-7, 8-15, 16-23, and 24-31. The user must use the ports within a group as specified above to form an ISL trunk. It is also possible to configure multiple trunks within a port group.

For this case study, an attempt is made to connect ISLs to the same number ports on both switches to help assist with troubleshooting should the need arise.

For general information on host and storage layout for a compound core edge switch, refer to "Host and storage layout" on page 200.

Switch and fabric management

In this example, both CLI and Web Tools will be used to set up the compound core edge SAN topology, with the four Connectrix ED-48000Bs at the core and the eight DS-4900Bs at the edge.

For general information on switch and fabric management for a compound core edge switch, refer to "Switch and fabric management" on page 200. Specific information for this example follows.

Security

The connectivity and device discovery for a large Connectrix–B switch fabric may be secured by appointing the following binding techniques. All these are Connectrix-B specific features.

- Fabric binding is a security method for restricting switches within a multiple-switch fabric. The SCC policy prevents unauthorized switches from joining a fabric. Switches are authenticated using digital certificates and unique private keys provided to the Switch Link Authentication Protocol (SLAP).
- Switch binding is a security method for restricting devices that connect to a particular switch. If the device is another switch, this is handled by the SCC policy. If the device is a host or storage device, the Device Connection Control (DCC) policy binds those devices to a particular switch. Policies range from completely restrictive to reasonably flexible, based upon customer needs.
- Port binding is a security method for restricting host or storage devices that connect to particular switch ports. The DCC policy also binds device ports to switch ports. Policies range from completely restrictive to reasonably flexible, based on customer needs. For switches running Fabric OS v5.2.0 and later, the SCC ACL with strict fabric-wide consistency can also be used for switch binding in addition to the Secure Fabric OS mechanism.

The method to enable SCC and DCC policies has been provided at the end of the fabric configuration steps for this case study (refer to "Enabling the Switch Connection Policy (SCC)" on page 216).

For general information on security for a compound core edge switch, refer to "Security" on page 200. Specific information for this example follows.

Setting up this topology

Assumption specific to this case study:

- The ED-48000B director class and the DS-4900B edge switches are installed in an EMC-supplied cabinet.

  For installation instructions, refer to the *Connectrix EC-1500 Cabinet Installation and Setup Manual*, accessible from Powerlink.
- The proper power receptacles have been provided by the customer.

For switch power requirements, refer to "Connectrix B-Series directors and switches" on page 976..

- The switches have not been connected to the power source and are not powered on.
- Network drops, IP addresses, subnet mask, and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to "Connectrix B-Series directors and switches" on page 976.
- The proper number of line cards have been installed into the ED-48000Bs.

  For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.

  Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- An EMC-supported Brocade FOS is installed on all switches in the fabric.

  Refer to the *EMC Support Matrix* for the most current support information.
- A 32-port blade is installed in the first slot, a 16-port blade is installed in the second slot, a 48-port blade is installed in the fourth slot and the 32-port blade is installed in the fifth slot of all the ED-48000B director switches.

Configure the IP address

To confure the IP address:

Note: Connectrix B switches may ship with a default IP address that is not on the desired subnet. The ED-48000B director uses a maximum of three IPs per unit: one IP for the switch and one IP for each control processor. The DS-4900B uses only one IP.

---

1. Attach the provided serial cable between the serial port on one of the ED-48000B switches (switch A) and an RS-232 serial port on the management PC. The serial cable is wired with only pins 2, 3, and 5 wired straight through.

2. Power up the switch by connecting the power cords to the power receptacles provided by the customer.
3. Run a terminal emulation program, such as Hyperterm on Windows hosts or TERM in a UNIX environment.
4. Configure the terminal for 9600 Baud, 8 Data Bits, No Parity, 1 stop bit, and no flow control.
5. Press **Return** to get a prompt.

   Note: It is strongly recommended that when prompted to change the password, you change it to a password that was provided by the customer. This can also be done using the passwd command from the prompt at any time.

6. Log in using the default values: Username: *admin*; Password: *password*.
7. At the prompt, enter **ipaddrset -sw1** and press **Return**.
8. When prompted, supply IP address (172.23.199.4), Subnet Mask (255.255.255.0), and Gateway Address (172.23.199.2).

   Note: The Fibre Channel Addresses will not be used for this example.

9. At the prompt, enter **ipaddrset -cp 0** and press **Return**.
10. When prompted, supply hostname (**48K_1_5**), IP address (**172.23.199.5**), subnet mask (**255.255.255.0**), and gateway address (**172.23.199.2**).
11. At the prompt, **enter ipaddrset -cp 1** and press **Return**.
12. When prompted, supply hostname (48K_1_6), IP address (172.23.199.6), subnet mask (255.255.255.0), and gateway address (172.23.199.2).
13. Power down the switch and disconnect the serial cable.
14. Connect the switch to a 10/100BaseT Ethernet connection.
15. Power up the switch. The switch can now be accessed with IP-based management.

16. Repeat these steps for the other three switches B, C, and D using the following information:

	Domain ID	Switch IP	CP0 IP	CP0 Name	CP1 IP	CP1 Name
Switch A	4	172.23.199.4	172.23.199.5	48K_1_5	172.23.199.6	48K_1_6
Switch B	7	172.23.199.7	172.23.199.8	48K_1_8	172.23.199.9	48K_1_9
Switch C	10	172.23.200.10	172.23.200.11	48K_1_11	172.23.200.12	48K_1_12
Switch D	13	172.23.200.13	172.23.200.14	48K_1_14	172.23.200.15	48K_1_15

17. Repeat Step 1 through Step 6 for all of the DS-4900Bs: Switch E, F, G, H, I, J, K and L one-by-one, and then execute the following steps on switch E:
18. At the prompt, enter **ipaddrset** and press **Return**.
19. When prompted, supply IP address (**172.23.199.16**), Subnet Mask (**255.255.255.0**), and Gateway Address (**172.23.199.2**).
20. Power down the switch and disconnect the serial cable.
21. Connect the switch to a 10/100BaseT Ethernet connection.
22. Power up the switch. The switch can now be accessed with IP-based management.
23. Repeat steps Step 17 through Step 21 for the other seven switches F, G, H, I, J, K and L using the following information:

	Domain ID	Switch IP	Switch Name
Switch E	16	172.23.199.16	4900_16
Switch F	17	172.23.199.17	4900_17
Switch G	18	172.23.199.18	4900_18
Switch H	19	172.23.199.19	4900_19
Switch I	20	172.23.200.20	4900_20
Switch J	21	172.23.200.21	4900_21
Switch K	22	172.23.200.22	4900_22
Switch L	23	172.23.200.23	4900_23

Configure FC switches

To configure FC switches:

1. Set the switch name and fabric parameters for switch A.

Note: The following configurations need to be done with the switch *disabled*.

2. Configure the fabric parameters.

   a. From the switch prompt, enter **switchdisable** to disable the switch.

   b. From the switch prompt, enter **configure** to enter the configuration parameter menu.

   c. Enter **Y** at the **Fabric Parameters** prompt.

   d. Enter **4** for desired domain ID (as per the table above) at the Domain prompt and press **Enter**.

   e. The R_A_TOV should be automatically set to 10000. If it is not, enter **10000** at the prompt and press **Enter**.

   f. The E_D_TOV should be automatically set to 2000. If it is not, enter **2000** at the prompt and press **Enter**.

   g. Accept the following defaults for the rest of the fields under the **Fabric Parameters** menu by pressing **Enter** after each prompt:

      – WAN_TOV = 0
      – MAX_HOPS = 7
      – Data field size = 2112
      – Sequence Level Switching = 0
      – Disable Device Probing = 0
      – Suppress Class F Traffic = 0
      – Switch PID Format = 1
      – Per-frame Route Priority = 0
      – Long Distance Fabric = 0
      – BB_Credit = 16

Note: For this case study, there is no long distance between any of the switches. The ISLs connecting the two are less than 10 km.

h. At the **Insistent Domain ID Mode** prompt, enter **y** to accept the **Insistent domain ID** setting. When this mode is set, the switch attempts to acquire the domain number programmed in its **Switch Fabric Settings** from the fabric.

i. Accept the default values from the remaining **Fabric Parameter Menu** items by pressing **Enter** after each prompt:

- Virtual Channel parameters (yes, y, no, n): [**no**]
- F_Port login parameters (yes, y, no, n): [**no**]
- Zoning Operation parameters (yes, y, no, n): [**no**]
- RSCN Transmission Mode (yes, y, no, n): [**no**]
- Arbitrated Loop parameters (yes, y, no, n): [**no**]
- System services (yes, y, no, n): [**no**]
- Portlog events enable (yes, y, no, n): [**no**]
- ssl attributes (yes, y, no, n): [**no**]
- http attributes (yes, y, no, n): [**no**]
- snmp attributes (yes, y, no, n): [**no**]
- rpcd attributes (yes, y, no, n): [**no**]
- cfgload attributes (yes, y, no, n): [**no**]
- web tools attributes (yes, y, no, n): [**no**]

Note: You may also press **CNTRL D** after making the last change in the menu to exit and save the changes. This will eliminate the need to accept the default values for the rest of the menu items.

j. Repeat from Step a for the ED-48000B switches B, C, and D, and for the DS-4900B switches E, F, G, H, I, J, K and L using the values shown in the tables above, especially while entering the Domain IDs in Step d.

3. In this case study, we are setting the Switch A as the principal switch. We have a choice of setting any of the core ED-48000B switches in this configuration as the principal switch. In order to do this:

   a. Telnet into switch A.

   b. At switch prompt, enter **fabricprincipal 1.** This will set the switch A as the principal switch. Verify the “switch mode” by running the **switchshow** command.

4. Ports on the switch may be configured if desired. By default the port type and port speed is set to auto.

a. Issue the **portcfgeport <port number> 1** to configure the port type to E_Port for an ISL connection and to lock it as an E_Port.

b. Issue the **portcfggport <port number> 1** to configure as a generic G_Port.

c. Issue the **portcfgspeed** command to configure the port speed.

Connect cables

To connect the cables:

1. Connect ISLs.

   Attach Fiber cable between switches as shown in Figure 22 on page 201. (also refer to the port tables below for each individual switch. The index denotes the index number that is seen on running a **switchshow** on the brocade switch).

2. Connect host and storage ports.

   Attach fiber cable between switches and N_Ports. (Refer to the port tables listed in this step for each individual switch. The index denotes the index number that is seen on running a **switchshow** on the Brocade switch).

   The port connections need to be made as follows:

   - For Switch A:

Index	Slot/Port#	Name
0	1/0	ISL to domain 21
1	1/1	ISL to domain 20
2	1/2	Green Host HBA 1
16	1/16	ISL to domain 10
32	2/0	ISL to domain 21
33	2/1	ISL to domain 20
48	4/0	ISL to domain 10
64	4/17	ISL to domain 17
65	4/18	ISL to domain 16
80	4/33	ISL to domain 13
81	4/34	ISL to domain 7

96	5/0	ISL to domain 17
97	5/1	ISL to domain 16
112	5/17	ISL to domain 13
113	5/18	ISL to domain 7

- For Switch B:

Index	Slot/Port#	Name
0	1/0	ISL to domain 22
1	1/1	ISL to domain 23
2	1/2	Blue Host HBA 1
3	1/3	Blue storage 1
16	1/16	ISL to domain 13
32	2/0	ISL to domain 22
33	2/1	ISL to domain 23
48	4/0	ISL to domain 13
64	4/17	ISL to domain 19
65	4/18	ISL to domain 18
80	4/33	ISL to domain 10
81	4/34	ISL to domain 4
96	5/0	ISL to domain 19
97	5/1	ISL to domain 18
112	5/17	ISL to domain 10
113	5/18	ISL to domain 4

- For Switch C:

Index	Slot/Port#	Name
0	1/0	ISL to domain 17
1	1/1	ISL to domain 16
2	1/2	Green Host HBA 2
16	1/16	ISL to domain 4

32	2/0	ISL to domain 17
33	2/1	ISL to domain 16
48	4/0	ISL to domain 4
64	4/17	ISL to domain 21
65	4/18	ISL to domain 20
80	4/33	ISL to domain 7
81	4/34	ISL to domain 13
96	5/0	ISL to domain 21
97	5/1	ISL to domain 20
112	5/17	ISL to domain 13
113	5/18	ISL to domain 7

- For Switch D:

Index	Slot/Port#	Name
0	1/0	ISL to domain 18
1	1/1	ISL to domain 19
2	1/2	Blue Host HBA 2
3	1/3	Blue storage 2
16	1/16	ISL to domain 7
32	2/0	ISL to domain 18
33	2/1	ISL to domain 19
48	4/0	ISL to domain 7
64	4/17	ISL to domain 23
65	4/18	ISL to domain 22
80	4/33	ISL to domain 4
81	4/34	ISL to domain 10
96	5/0	ISL to domain 23
97	5/1	ISL to domain 22
112	5/17	ISL to domain 4
113	5/18	ISL to domain 10

- For Switch E:

Port#	Name
0	ISL to domain 10
1	ISL to domain 10
2	ISL to domain 4
3	ISL to domain 4
4	Red Host HBA 1

- For Switch F:

Port#	Name
0	ISL to domain 10
1	ISL to domain 10
2	ISL to domain 4
3	ISL to domain 4
4	Red Storage 1

- For Switch G:

Port#	Name
0	ISL to domain 13
1	ISL to domain 13
2	ISL to domain 7
3	ISL to domain 7
4	Green Storage 1

- For Switch H:

Port#	Name
0	ISL to domain 13
1	ISL to domain 13
2	ISL to domain 7
3	ISL to domain 7
4	Green Storage 2

- For Switch I:

Port#	Name
0	ISL to domain 4
1	ISL to domain 4
2	ISL to domain 10
3	ISL to domain 10
4	Red Host HBA 2

- For Switch J:

Port#	Name
0	ISL to domain 4
1	ISL to domain 4
2	ISL to domain 10
3	ISL to domain 10
4	Red Storage 2

- For Switch K:

Port#	Name
0	ISL to domain 7
1	ISL to domain 7
2	ISL to domain 13
3	ISL to domain 13
4	Green Storage 4

- For Switch L:

Port#	Name
0	ISL to domain 7
1	ISL to domain 7
2	ISL to domain 13
3	ISL to domain 13
4	Green Storage 3

3. Verify port login status.
   a. After all cables are connected, use the **switchshow** CLI command to verify all of the ports logged into the switch. The ports with ISLs must log in as E_Ports while the ports connected to the HBA or storage must log in as F_Ports.

Zone hosts and storage

To zone hosts and storage:

1. Using a web browser, enter the IP address of Switch A (**172.23.199.4**) in the URL / address field.
2. Enter username and password at the prompt.
3. Click the **Zone Admin** icon in the lower left corner at bottom of web page.
4. The **Zone Admin** pop-up window should appear. You may be asked for login credentials again.
5. Under the **Alias** tab, click **New Alias**. A **Create New Alias** window appears. Enter **Red_HBA1** as the alias name.

6. Expand the WWNs folder (if needed) by clicking on the **+** sign. Click on the WWN of Red_HBA1 (**10:00:00:00:c9:38:e5:54)** and click **Add Member** to move the WWN to **Alias Members** column.

   This will create an alias for Blue_HBA1 using one of the WWNs of the host. All zoning created with this alias will now use its WWN for zoning.

7. Click **New Alias** again and in the **Create New Alias** window, enter **Red_Storage1** (50:06:04:82:cc:19:bf:87) as the alias name.

8. Expand the WWNs folder (if needed) by clicking on the **+** sign. Select the WWN of Blue_Storage_A (**50:06:04:82:cc:19:bf:87**) and then select **Add Member** to move the WWN to the **Alias Members** column.

   This will create an alias for Red_Storage1 using the WWN of the storage device. All zoning created with this alias will now use its WWN for zoning.

9. Under the **Zone** tab, click **New Zone**. A **Create New Zone** window appears. Enter "**RedHBA1_1470_8aa**" for Zone Name. This case zones Red Host and Red Storage 1.

10. Expand the Alias folder (if needed) by clicking on the **+** sign. Select the "**Red_HBA1**" alias and then click **Add Member** to move alias to the **Zone Members** column. Select the "**Red_Storage1**" alias and then select **Add Member** to move alias to the **Zone Members** column.

    This will create a zone for the Red_HBA1 to allow access to the Red_Storage1.

11. Under the **Zone Config** tab, click **New Zone Config**. A **Create New Config** window appears. This case uses the date as the name. Enter "**Oct_31_06_1140**" as the config name.

12. Expand the Zones folder (if needed) by clicking on the **+** sign. Select the "**RedHBA1_1470_8aa**" zone and click **Add Member** to move zone to the **Zone Config Members** column.

    This will create a zone with the "**RedHBA1_1470_8aa**" zone as a zoneset member. This zoneset will need to be enabled before becoming effective.

13. Repeat Step 5 through Step 12 to create aliases and a zoneset for the other Red HBA and Red Storage port, Green HBAs, Green Storages, Blue HBAs and Blue Storage as stated below.

14. Click the **Zoning Actions** pull-down menu on top of window. Click **Enable Config** and select "**Oct_31_06_1140**" zone and then click **OK**. This will push the new zone out to the fabric and make it effective.

15. When completed, the zone set, when running the **cfgactvshow** command from CLI, should be similar to what is shown below:

    Effective configuration:

```
Cfg: Oct_31_06_1140
Zone: RedHBA1_1470_8aa
   10:00:00:00:c9:38:e5:54
   50:06:04:82:cc:19:bf:87

Zone: RedHBA2_1470_9aa
   10:00:00:00:c9:38:e5:55
   50:06:04:82:cc:19:bf:88

Zone: BlueHBA1_1489_8aa
   21:01:00:e0:8b:8a:c7:6d
   50:06:04:82:cc:19:c4:47

Zone: BlueHBA2_1489_9aa
   21:01:00:e0:8b:aa:c7:6d
   50:06:04:82:cc:19:c4:48

Zone: GreenHBA1_AllGreenStorage
   10:00:00:00:c9:39:a0:51
   50:06:04:82:cc:19:c4:07
   50:06:04:82:cc:19:c4:08
   50:06:04:82:cc:19:c4:c7
   50:06:04:82:cc:19:c4:c8

Zone: GreenHBA2_GreenStorage
   10:00:00:00:c9:39:a0:52
   50:06:04:82:cc:19:c4:07
   50:06:04:82:cc:19:c4:08
   50:06:04:82:cc:19:c4:c7
   50:06:04:82:cc:19:c4:c8
```

Enabling the Switch Connection Policy (SCC)

To enable the Switch Connection Policy:

1. At the switch prompt, enter **fddcfg -fabwideset** "**SCC:S;DCC**".

2. Press **Enter**.

   This command will set a strict SCC and tolerant DCC fabric-wide consistency policy.

Note: When a switch is joined to a fabric with a strict Switch Connection Control (SCC) or Device Connection Control (DCC) fabric-wide consistency policy, the joining switch must have a matching fabric-wide consistency policy. If the strict SCC or DCC fabric-wide consistency policies do not match, the switch cannot join the fabric and the neighboring E_Ports will be disabled. If the strict SCC and DCC fabric-wide consistency policies match, the corresponding SCC and DCC Access Control List (ACL) policies are compared.

3. To verify that the policy has been set, the **fddcfg –showall** command can be run on any switch in the fabric. Any switch on the fabric should show output similar to:

```
switch:admin> fddcfg --showall
Local Switch Configuration for all Databases:-
DATABASE - Accept/Reject
------------------------
SCC - accept
DCC - accept
PWD - accept
Fabric Wide Consistency Policy:- "SCC:S;DCC"
```

Completing the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete.

Refer to the OS configuration guide for more detail.

Connectrix MDS example

General layout

Figure 23 shows four MDS 9513s in a full mesh configuration with edge switches attached.

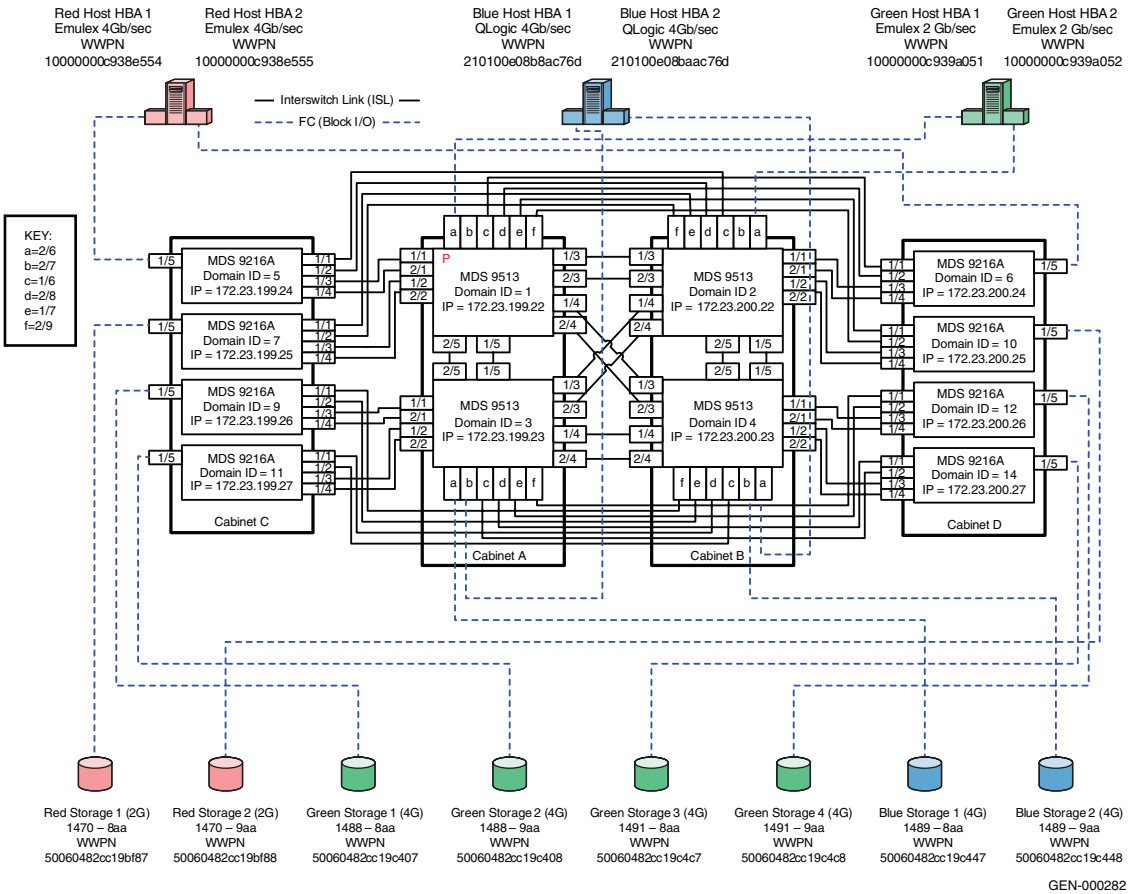


Figure 23 **Four MDS 9513s in full mesh configuration with edge switches attached**

Best practices

For general information on best practices for a compound core edge switch, refer to "Best practices" on page 200. Specific information for this example follows.

By default, utilization is set to 80%.

Host and storage layout

For general information on host and storage layout a compound core edge switch, refer to "Host and storage layout" on page 200. Specific information for this example follows.

Line Rate Cards have no special considerations. Over-subscribed cards should be used for hosts only.

Switch and fabric management

For general information on switch and fabric management a compound core edge switch, refer to "Switch and fabric management" on page 200. Specific information for this example follows.

Cisco Fabric Manager can be used for complex fabrics.

Security

For general information on security a compound core edge switch, refer to "Security" on page 200. Specific information for this example follows.

Enable switch and port binding for security.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix B, "EMC Connectrix Products Data".
  - For Cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and setup Manual* which can be accessed from Powerlink.
- The switches have *not* been connected to the power source and are *not* powered on.
- Network drops, IP addresses, subnet mask and gateway, have been provided by the customer.

  For switch or cabinet network requirements, refer to Appendix B, "EMC Connectrix Products Data".

Note: Connectrix MDS switches can be directly connected to the customer's LAN. The switches can be placed on either a public or private network. There are advantages in both configurations. For more information, refer to "Public versus private" on page 961.
This example assumes that The customer has provided us with two Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network.

- The proper number of line cards have been installed into the Connectrix MDS 9513s.
  - For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.
- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- A laptop, connected to a Connectrix MDS serial port, is used to configure the IP addresses of the switches.
- The customer has provided a temporary password that is used as the default password when configuring the IP address.
- Cisco CLI, Fabric Manager, and Device Manager are used.

Configure the IP address

To configure the IP address:

1. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer.
2. Select one of the switches to configure and set the IP to 172.23.199.22.
3. Supply a network connection to the appropriate subnet.
4. Using an RS232 serial cable, connect to the serial port of the switch with a baud rate of 9600, 8 data bits, no parity, 1 stop bit and no flow control.

   The **login** prompt should display.
5. Log in the first time with a user name *admin* and password *admin*.

   You should be prompted to supply a new strong password for CLI user admin.
6. For this example, select **no** when prompted to run setup.

Note: This example starts with the switch that will have a Domain ID of **1** and an IP address of **172.23.199.22**.

CLI commands to configure the IP and gateway

- Switch# ***config terminal***

  Enter configuration commands, one per line.

Switch(config)# ***interface mgmt 0***

Switch(config-if)#***ip address 172.23.199.22 255.255.255.0***

End with **CNTL/Z**.

- Switch# ***config terminal***

  Enter configuration commands, one per line.

  Switch(config)# ***ip default-gateway 172.23.199.2***

  End with **CNTL/Z**.

To authorize access on a switch for Device and Fabric Manager, run this command on every switch while supplying a user name (nnn) and password (ppp):

- Switch#***conf t***

  Switch(config)# ***snmp-server user nnn network-admin auth md5 ppp***

  Switch(config)#***end***

  Switch# ***copy running-config startup-config***

  Switch# ***exit***

Configure the IP addresses of the switches in the cabinet

To configure the IP address of the switches in the other cabinet:

Follow Step 4 through Step 6 in "Configure the IP address" on page 161.

Use the IP addresses found in Figure 18 on page 159 with subnet mask (255.255.255.0) and gateway (172.23.199.2 and 172.23.200.2).

Configure the rest of the switch IP addresses in the cabinet

To configure the next switch, follow steps 4 – 11 in the "Configure the IP address" section. Use the IP address of (172.23.199.23), subnet mask (255.255.255.0) and gateway (172.23.199.2).

Install Fabric Manager and Device Manager

To install Fabric Manager and Device Manager:

1. Open your web browser.
2. Enter the IP address of the switch into the address bar.

3. Follow the prompts and accept all defaults to install both Fabric Manager and Device Manager.

Configure a VSAN followed by a domain

To configure a VSAN:

1. Open the Device Manager for the switch with an IP address of **172.23.199.22**.
2. Open the **VSAN** dialog by selecting the **VSAN** menu item.
3. Click **Create**.
4. Enter the value **100** in the **VSAN ID** field.
5. Set the **VSAN Name** to "**Red_Vsan_100**".
6. Use the default interop mode.
7. Click **Create**.
8. Enter the value **200** in the **VSAN ID** field.
9. Set the next **VSAN Name** to be "**Green_Vsan_200**".
10. Click **Create**.
11. Enter the value **300** in the **VSAN ID** field.
12. Set the next **VSAN Name** to "**Blue_VSAN_300**".
13. Click **Create**, then click **Close**.
14. From the **Device Manager** menu, select **FC/Domain Manager/Configuration** and set a static Domain ID for the switches as shown in Table 1 and Table 2 on page 223.

Table 1 **172.23.199.22 through 172.23.199.27**

IP	Domain	VSAN_ID	VSAN_ID	VSAN_ID
172.23.199.22	1	100	200	300
172.23.199.23	3	100	200	300
172.23.199.24	5	100	200	300
172.23.199.25	7	100	200	300
172.23.199.26	9	100	200	300
172.23.199.27	11	100	200	300

Table 2 **172.23.200.22 through 172.23.200.27**

IP	Domain	VSAN_ID	VSAN_ID	VSAN_ID
172.23.200.22	2	100	200	300
172.23.200.23	4	100	200	300
172.23.200.24	6	100	200	300
172.23.200.25	10	100	200	300
172.23.200.26	12	100	200	300
172.23.200.27	14	100	200	300

Configure FC switches

To configure FC switches:

1. Configure the switch ports.
   a. Open the Device Manager of the switch with an IP address of 172.23.199.22, or the next switch, by double clicking its icon in Fabric Manager.
   b. From the **Configure** menu, select the **Device menu** item.

c. Admin up and configure the ports as shown in the following tables:

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 5	1
1	2	TE ISL to Domain 7	1
1	3	TE ISL to Domain 2	1
1	4	TE ISL to Domain 4	1
1	5	TE ISL to Domain 3	1
1	6	TE ISL to Domain 6	1
1	7	TE ISL to Domain 10	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to Domain 5	1
2	2	TE ISL to Domain 7	1
2	3	TE ISL to Domain 2	1
2	4	TE ISL to Domain 4	1
2	5	TE ISL to Domain 3	1
2	6	Green Host HBA 1	200
2	7		
2	8	TE ISL to Domain 6	1
2	9	TE ISL to Domain 10	1

d. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.23 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 9	1
1	2	TE ISL to Domain 11	1
1	3	TE ISL to Domain 2	1
1	4	TE ISL to Domain 4	1
1	5	TE ISL to Domain 1	1
1	6	TE ISL to Domain 14	1
1	7	TE ISL to Domain 12	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to Domain 9	1
2	2	TE ISL to Domain 11	1
2	3	TE ISL to Domain 2	1
2	4	TE ISL to Domain 4	1
2	5	TE ISL to Domain 1	1
2	6	Blue Storage 1	300
2	7	Blue Host HBA 1	300
2	8	TE ISL to Domain 14	1
2	9	TE ISL to Domain 12	1

e. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.22 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 6	1
1	2	TE ISL to Domain 10	1
1	3	TE ISL to Domain 1	1
1	4	TE ISL to Domain 3	1
1	5	TE ISL to Domain 4	1
1	6	TE ISL to Domain 5	1
1	7	TE ISL to Domain 7	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to Domain 6	1
2	2	TE ISL to Domain 10	1
2	3	TE ISL to Domain 1	1
2	4	TE ISL to Domain 3	1
2	5	TE ISL to Domain 4	1
2	6	Green Host HBA 2	200
2	7		
2	8	TE ISL to Domain 5	1
2	9	TE ISL to Domain 7	1

f. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.23 as shown in the following tables:

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 1	1
1	2	TE ISL to Domain 2	1
1	3	TE ISL to Domain	1
1	4	TE ISL to Domain	1
1	5	TE ISL to Domain 2	1
1	6	TE ISL to Domain 11	1
1	7	TE ISL to Domain 9	1
1	8		

Slot #	Port #	Name	VSAN ID
2	1	TE ISL to Domain 12	1
2	2	TE ISL to Domain 14	1
2	3	TE ISL to Domain 1	1
2	4	TE ISL to Domain 3	1
2	5	TE ISL to Domain 2	1
2	6	Blue Host HBA 2	300
2	7	Blue Storage 2	300
2	8	TE ISL to Domain 9	1
		TE ISL to Domain 11	1

g. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.24 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 2	1
1	2	TE ISL to Domain 2	1
1	3	TE ISL to Domain 1	1
1	4	TE ISL to Domain 1	1
1	5	Red Host HBA 1	100
1	6		
1	7		
1	8		

h. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.24 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 1	1
1	2	TE ISL to Domain 1	1
1	3	TE ISL to Domain 2	1
1	4	TE ISL to Domain 2	1
1	5	Red Host HBA 2	100
1	6		
1	7		
1	8		

i. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.25 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 2	1
1	2	TE ISL to Domain 2	1
1	3	TE ISL to Domain 1	1
1	4	TE ISL to Domain 1	1
1	5	Red Storage 1	100
1	6		
1	7		
1	8		

j. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.25 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 1	1
1	2	TE ISL to Domain 1	1
1	3	TE ISL to Domain 2	1
1	4	TE ISL to Domain 2	1
1	5	Red Storage 2	100
1	6		
1	7		
1	8		

k. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.26 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 4	1
1	2	TE ISL to Domain 4	1
1	3	TE ISL to Domain 3	1
1	4	TE ISL to Domain 4	1
1	5	Green Storage 1	200
1	6		
1	7		
1	8		

l. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.26 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 3	1
1	2	TE ISL to Domain 3	1
1	3	TE ISL to Domain 4	1
1	4	TE ISL to Domain 4	1
1	5	Green Storage 4	1
1	6		
1	7		
1	8		

m. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.27 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 4	1
1	2	TE ISL to Domain 4	1
1	3	TE ISL to Domain 3	1
1	4	TE ISL to Domain 3	1
1	5	Green Storage 2	200
1	6		
1	7		
1	8		

n. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.27 as shown in the following tables.

Slot #	Port #	Name	VSAN ID
1	1	TE ISL to Domain 3	1
1	2	TE ISL to Domain 3	1
1	3	TE ISL to Domain 4	1
1	4	TE ISL to Domain 4	1
1	5	Green Storage 3	200
1	6		
1	7		
1	8		

Connect cables

To connect the cables:

1. Connect ISLs.

   a. Attach Fiber cable between switches as shown in Figure 18 on page 159.

b. After all cables are connected, use Fabric Manager to verify that all ISL connections are up.

c. Re-arrange icons to accurately reflect the switch configuration.

2. Connect host and storage ports.

   a. Attach fiber cable between the switches and N_Ports.

Configure domains

Using Fabric Manager set the static Domain IDs for each VSAN and switch.

Zone hosts and storage

To zone hosts and storage:

1. Open the **Zoning** dialog in Connectrix Manager by right -clicking the appropriate fabric topology and selecting the **Zoning** menu item.
2. Create a zone by clicking **New Zone** under the **Zones Tree**.
3. Provide a descriptive name for the zone. This example zones “Red host HBA 1” and “Red Storage 1”. Type **“RedHBA1_1470_8aa”** and press **Enter**.
4. Select **“Red Host HBA 1**” (WWPN 10000000c938e554) in the **potential zone members list**.
5. Click the right-pointing arrow on the divider between the **potential members list,** and the **zones list** to add the HBA to the zone.
6. Select **“Red Storage 1”** (WWPN 50060482cc19bf87) in the **potential zone members list**.
7. Click the right- pointing arrow on the divider between the **potential members list** and the **zones list** to add the storage port to the zone.
8. Repeat Step 2 through Step 6 for all host and storage pairs in the environment.
9. Create a zone set by clicking **New Set** under the **Zone Sets Tree**.
10. Provide a descriptive name for the zone set as shown in the example following Step 12.

11. Add all of the new zones to the zone set of the proper VSAN. When completed, the zone sets should be similar to what is shown in the example following these steps.

12. Activate the **VSAN Zone Set**.

```
Zone set name = "Red_Oct_31_06_1140"

          Zone name = "RedHBA1_1470_8aa"
             Zone Member = "10000000c938e554"
             Zone Member = "50060482cc19bf87"

          Zone name = "RedHBA2_1470_9aa"
             Zone Member = "10000000c938e555"
             Zone Member = "50060482cc19bf88"

Zone set name = "Green_Oct_31_06_1140"

          Zone name = "GreenHBA1_AllGreenStorage"
             Zone Member = "10000000c939a051"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"

          Zone name = "GreenHBA2_AllGreenStorage"
             Zone Member = "10000000c939a052"
             Zone Member = "50060482cc19c407"
             Zone Member = "50060482cc19c408"
             Zone Member = "50060482cc19c4c7"
             Zone Member = "50060482cc19c4c8"

Zone set name = "Blue_Oct_31_06_1140"

          Zone name = "BlueHBA1_1489_8aa"
             Zone Member = "210100e08b8ac76d"
             Zone Member = "50060482cc19c447"

          Zone name = "BlueHBA2_1489_9aa"
             Zone Member = "210100e08baac76d"
             Zone Member = "50060482cc19c448"
```

Complete the SAN setup

At this point the SAN is ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN setup is complete. Refer to the OS configuration guide for more detail.

Connectrix M example

General layout

Figure 24 shows four ED-10000Ms in a full mesh configuration with edge switches attached.

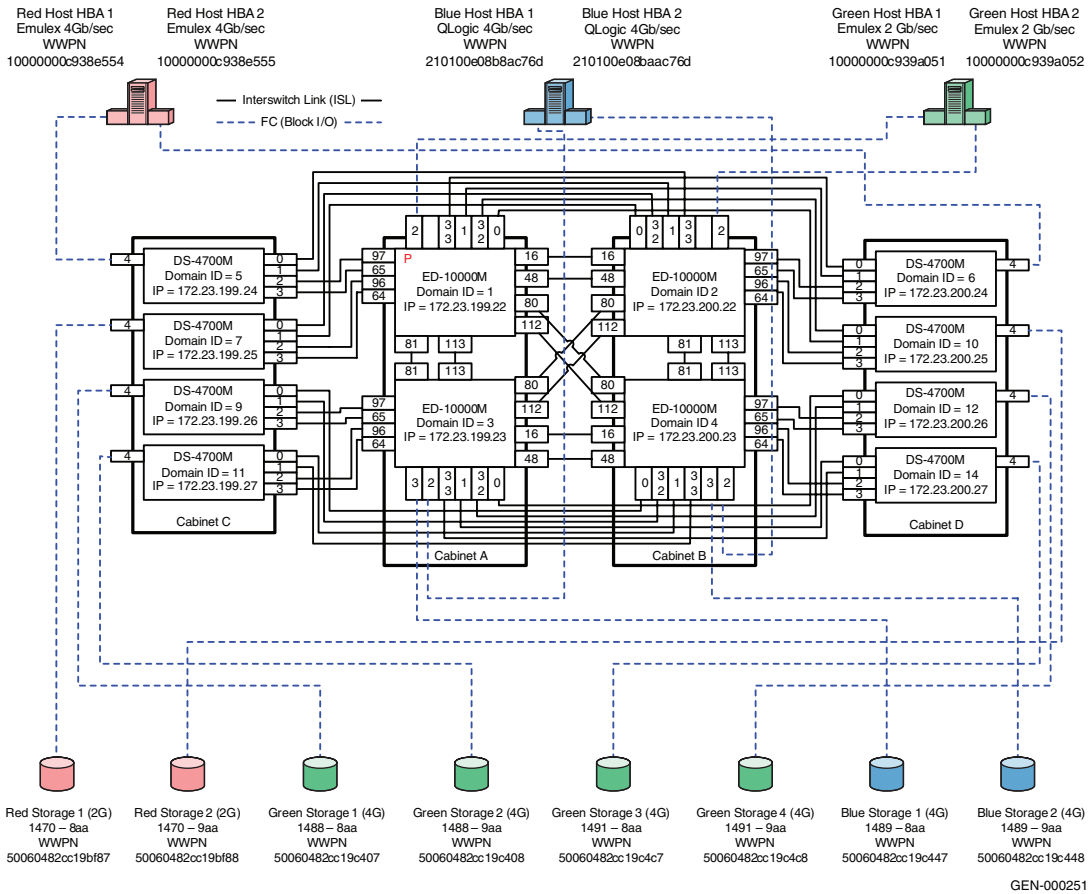


Figure 24 **Four ED-10000Ms in full mesh configuration with edge switches attached**

Note: Any Director class product such as the ED-10000M, ED-140M, or ED-64M can be used in the core of this fabric. All Director and switch class products can be used as edge switches.

Best practices

For general information on best practices for a compound core edge switch, refer to "Best practices" on page 200. Specific information for this example follows.

Host and storage layout

The N_Port and E_Port connections on the ED-10000M were spread out over as many SPPe's as possible. This was done to ensure fairness (refer to the"Connectrix ED-10000M" on page 1096) and increase the reliability of the network by minimizing the impact that any single hardware failure would have. The N_Port and E_Port connections on the DS-4700Ms are attached to the chassis in order since there are no benefits to performance or reliability by spreading out the connections.

In both cases, an attempt has been made to connect ISLs to the same number port on both switch to help assist with troubleshooting should the need arise.

For general information on host and storage layout a compound core edge switch, refer to "Host and storage layout" on page 200.

Switch and fabric management

Connectrix Manager will be used to configure the environment in this example. Using Connectrix Manager Basic or CLI in this example is not practical.

For general information on switch and fabric management a compound core edge switch, refer to "Switch and fabric management" on page 200.

Security

For general information on security on a compound core edge switch, refer to "Security" on page 200. Specific information for this example follows.

At the end of this example, fabric binding, and switch binding are enabled by turning on Enterprise Fabric Mode. For more information on fabric binding, refer to "Fabric binding" on page 508. For more information on switch binding, refer to page 378.

Setting up this topology

Assumptions specific to this case study:

- The switches are installed in an EMC-supplied cabinet.
  - For installation instructions, see *Connectrix EC-1500 Cabinet Installation and setup Manual,* which can be accessed from Powerlink.
- The proper power receptacles have been provided by the customer.
  - For switch power requirements, refer to Appendix C, "Non-EMC Products Data."
  - For cabinet power requirements, refer to *Connectrix EC-1500 Cabinet Installation and setup Manual,* which can be accessed from Powerlink.
- The switches have *not* been connected to the power source and are *not* powered on.

Note: Connectrix M switches ship with a default IP address of 10.1.1.10. If all of the switches are powered on and connected to the hub in the switch cabinet, communication with the switch using the network interfaces is *not* possible and you must program the IP addresses using the serial port.

This section assumes that the serial port is not used and the IP addresses is programmed through the network interface.

- Network drops, IP addresses, Subnet mask and gateway, were provided by the customer.

  For switch or cabinet network requirements, refer to Appendix C, "Non-EMC Products Data."

Note: Connectrix M switches can be either attached to the Ethernet hub that comes with the cabinet, or directly connected to the customer LAN. In both cases, the switches can be placed on either a public or private network. There are advantages to both configurations. For more information**,** refer to "Public versus private" on page 961.

This example assumes that The customer has provided us with two Ethernet cables and that one of them is on the 172.23.199.x network and that the other is connected to the 172.23.200.x network. All switches are connected to the Ethernet hubs in the cabinet.

- The proper number of LIMs have been installed into the ED-10000Ms.

- For help in determining how many ports are required, refer to "Determining customer requirements" on page 759.

- License keys have been obtained.
  - Go to the URL listed on the transaction code certificate that shipped with the product to get the license key.
- A Service Processor is installed in at least one of the cabinets.
- The Service Processor is used to configure the IP addresses of the switches.
- One interface from the Service Processor is connected to the Ethernet hub and this interface has an IP address of 10.1.1.11. (Check **Default** on the SP image and use that value.)
- The customer has provided a temporary password that will be used as the default password when configuring the IP address.
- Connectrix Manager is used and the Connectrix Manager installation kit is available.
- Connectrix Manager is not already installed on the Service Processor.
- The customer has purchased the Security Center License.

Configure the IP addresses of the ED-10000Ms in cabinet A

To configure the IP address of the ED-10000Ms in cabinet A:

1. In the cabinet that contains the Service Processor, select one of the switches to configure.

   Note: This example starts with the switch that has a Domain ID of 1 and an IP address of 172.23.199.22.

2. Power up the cabinet by connecting the power cords to the power receptacles provided by the customer. This causes both the directors and the Service Processor to power up.
3. Log in to the Service Processor once it has powered up. The default username and password is *administrator* and *password*.
4. Perform the following steps to change a product IP address, subnet mask, or gateway address. An asynchronous RS-232 modem cable and maintenance terminal (desktop or notebook PC) with a Windows-based operating system and RS-232 serial communication software (such as ProComm Plus or HyperTerminal) are required.

Note: The Service Processor can be used for configuring the IP addresses through a serial cable.

a. Using a Phillips screwdriver, remove the protective cap from the 9-pin maintenance port. Connect one end of the RS-232 modem cable to the port.

b. Connect the other cable end to a 9-pin serial communication port (**COM1** or **COM2**) at the rear of the maintenance terminal PC.

c. Power-on the maintenance terminal. At the Windows desktop, click **Start**. The **Windows Workstation** menu appears.

Note: The following steps describe changing network addresses using HyperTerminal serial communication software.

d. From the **Windows Workstation** menu, select **Programs, Accessories, Communications**, and **HyperTerminal**. The **Connection Description** dialog box displays (Figure 25).

Figure 25 Connection description dialog box

e. Type a descriptive director name in the **Name** field and click **OK**. The **Connect To** dialog box appears.

f. Ensure that the **Connect using** field appears **COM1** or **COM2** (depending on the port connection to the director), and click **OK**. The **COMn Properties** dialog box appears, where *n* is **1** or **2**.

g. Configure **Port Settings** parameters:

   – *Bits per second* - **57600**.
   – *Data bits* - **8**.
   – *Parity* - **None**.
   – *Stop bits* - **1**.
   – *Flow control* - **Hardware** or **None**.

   Click **OK**. The **New Connection - HyperTerminal** window appears.

h. At the **EOS#** prompt, type the user password (default is **password**) and press **Enter**. The password is case-sensitive. The **New Connection - HyperTerminal** window appears with software and hardware version information for the director.

i. At the **EOS#** prompt, type the **system ip** command and press **Enter**. The **New Connection - HyperTerminal** window appears with configuration information listed:

   – *MAC Address*.
   – *IP Address* (default is **10.1.1.10**).

   Note: IP address is the preinstalled IP address but may not be correct for your director. Consult your OEM provider for precise information about the IP address.

   – *Subnet Mask* (default is 255.0.0.0).
   – *Gateway Address* (default is 0.0.0.0).

   Only the **IP Address**, **Subnet Mask**, and **Gateway Address** fields are configurable.

j. Change the IP address, subnet mask, and gateway address as directed by the customer. To change the addresses, type the following and press **Enter**.

   system ip *xxx.xxx.xxx.xxx yyy.yyy.yyy.yyy zzz.zzz.zzz.zzz*

The IP address is ***xxx.xxx.xxx.xxx***, the subnet mask is ***yyy.yyy.yyy.yyy***, and the gateway address is ***zzz.zzz.zzz.zzz***, where the octets ***xxx***, ***yyy***, and ***zzz*** are decimals from zero through 255. If an address is to remain unchanged, type the current address in the respective field.

Configure the rest of the switch IP addresses in cabinet A

To configure the next switch in the cabinet, follow Step 4a through Step 4j, using the IP address of 172.23.199.23

Configure the IP addresses of the ED-10000Ms in cabinet B

To configure the ED-10000Ms in the other cabinet, follow Step 4a through Step 4j, using the IP address of 172.23.200.22 for the top switch (Domain ID 2) and 172.23.200.23 for the bottom switch (Domain ID 4).

Configure the IP addresses of the ED-4700Ms in cabinets C and D

Note: To configure the switch in the other cabinet, a cable can be connected between the two hubs to allow for communication between the service processor in one cabinet (cabinet A in our example), and the switches in the other cabinet (cabinet B in our example). This technique is used in this section.

1. Connect a straight through cable to port 24 on the cabinet B. Ensure that MDIX is set on port 24 by pushing the recessed MDI/MDIX button.
2. Connect the other end of cable connected to port 24 in Step 1 to any port other than 24 on the Ethernet hub in cabinet A.

   Note: This example configures the switch with the IP address of 172.23.199.24 first.

3. Disconnect the network cables from the other switches in the cabinet. Failure to do so causes an IP address conflict when the cabinet is powered up.

   Note: Cabinet D should not be powered up yet, but if it is, and a network cable is connected between the two hubs, remember that the network cables from the other switches must be disconnected to avoid running into duplicate IP address problems.

4. Power up cabinet C by connecting the power cords to the power receptacles provided by the customer. This causes the switches to power up.
5. Open a command prompt and ping the IP address of **10.1.1.10**. If you get a response, continue at Step 6. If you do not get a response, you can either IML the switch by selecting the gray button on the CTP, or configure the IP address using the serial port.

   Note: See the Installation and Service guide for more information about configuring the IP address using the serial port.

6. Launch Internet Explorer and enter the IP address of **10.1.1.10**, and then select **Enter**. The Login banner appears.
7. Click **Accept** to display the **Login** dialog box.
8. Enter the default username and password of *Administrator* and *password*, and then click **Enter**.

   Note: It is strongly recommended that when prompted to change the password, you change it to a password provided by the customer.

9. Click **Details** in the **Connectrix Manager Basic Topology** view.
10. Navigate to the **Network Configuration** dialog by clicking the **Configure** pull-down menu, and selecting **switch** and then **Network**.
11. Enter the IP address (**172.23.199.24**), subnet mask (**255.255.255.0**) and gateway (**172.23.199.2**), then click **OK**.

Note: An error message appears stating *Unable to display*…. This appears because the IP address of the switch changed and you can no longer connect to the switch at its old IP address.

Configure the rest of the switch IP addresses in cabinet C

1. To configure the next switch in the cabinet, re-connect the network cables to the management port of the switch that will eventually have the IP address of 172.23.199.25, and then follow steps Step 5 through Step 11 in the previous section. Use the IP address of (**172.23.199.25**), subnet mask (**255.255.255.0**) and gateway (**172.23.199.2**).

2. Repeat this process for the other switches in cabinet C using IP addresses **172.23.199.26** and **172.23.199.27**.

Configure the switch IP addresses in cabinet D

Follow the same process for configuring the switch IP addresses in cabinet D for cabinet C, but use the IP addresses of **172.23.200.24**, **172.23.200.25**, **172.23.200.26**, and **172.23.200.27**.

Install Connectrix Manager

Note: Refer to the *Connectrix Manager User Guide* for complete installation instructions.

To install Connectrix Manager:

1. Insert the Connectrix Manager installation CD into the CD-RW drive.
2. Browse to the CD-RW drive in Windows Explorer.
3. Locate, then double-click, the **install.exe** file. The installer appears. Typically, this file is located under the **CtxMgr <***version number***>** folder.
4. Accept all of the defaults by clicking **Next** and **OK**, when appropriate.
5. Click **Done** when the **Installation Complete** dialog appears.
6. Click **Next** when the **Welcome** dialog appears.
7. Accept the EULA terms, and then click **Next**.
8. At the **Copy data** and **Settings** dialog, select **No,** then **Next**.

   Note: "**No**" is selected in this step since this section assumes that this is a new installation.

9. Assign a Connectrix Manager 9.0 Server Name, and click **Next**.

   This case uses **CMServer.**

10. Enter the serial number and license key in the fields provided in the **Connectrix Manager 9.0 Server License** dialog box.
    - The serial number appears on the back of the Connectrix Manager installation CD case.
    - To obtain a license key, locate the transaction code certificate and go to the URL listed.

a. Enter the serial number located on the back of the Connectrix Manager installation CD case.

b. In the **Transaction Code** fields, enter the transaction code(s) shipped with the software.

c. Click **Next**.

d. Confirm the existing and new features to be enabled.

e. Click **Next**.

The license key and all enabled features appear.

f. Retain a copy for your records.

g. Enter this key into the **License key** field.

h. Click **OK**.

11. Once the installation is complete, log in to Connectrix Manager with the username of *administrator* and the password of *password*.

Note: Change this default as soon as possible. For this example, the username *nnn* and password *nnn* will be used.

Configure FC switches

To configure FC switches:

1. Set the switch name and Domain ID.

   a. From Connectrix Manager, double-click the switch icon for the switch with an IP address of **172.23.199.22**. The **Element Manager** appears.

   b. Under **Configure**, select the **Operating parameters** menu item.

   c. Select the **Domain** tab and ensure that the preferred Domain ID is set to the appropriate number (1 in this example).

   d. Check the **insistent Domain ID** checkbox to enable insistent Domain IDs.

Note: If the preferred Domain ID and the active Domain ID do *not* match, you cannot set the preferred Domain ID and enable insistent Domain IDs at the same time. Instead, change the preferred Domain ID first, click **OK**, re-open the dialog, check the insistent Domain ID checkbox, and then click **OK**.

e. Click the **Identification** tab.

f. Enter the switch name in the **Name** field.

g. Click **Copy** to make the switch name the switch nickname.

h. Click the **Fabric** tab.

i. Ensure that the Interop Mode is set to **Open Fabric 1.0**.

j. Follow this step only for the switch that has been selected to be the **Principal** switch (172.23.199.22 in this example).

   1. Set **Switch Priority** to principal.

k. Click **OK**.

l. Repeat steps Step b through Step i for each switch being configured.

2. Configure the switch ports.

   a. Open the Element Manager of the switch with an IP address of 172.23.199.22 by double-clicking its icon.

   b. From the **Configure** menu, select **ports**.

   c. Configure the ports as shown in the table below.

Slot #	SPPe #	Port #	Name
0	0	Port 0	ISL to Domain 10 (Port 1)
		Port 1	ISL to Domain 6 (Port 1)
		Port 2	Green Host HBA 1
0	1	Port 16	ISL to Domain 2 (Port 16)
1	0	Port 32	ISL to Domain 10 (Port 0)
		Port 33	ISL to Domain 6 (Port 0)
		Port 48	ISL to Domain 2 (Port 48)
2	0	Port 64	ISL to Domain 7 (Port 3)
		Port 65	ISL to Domain 5 (Port 3)
2	1	Port 80	ISL to Domain 4 (Port 80)
		Port 81	ISL to Domain 3 (Port 81)
3	0	Port 96	ISL to Domain 7 (Port 2)
		Port 97	ISL to Domain 5 (Port 2)

Slot #	SPPe #	Port #	Name
3	1	Port 112	ISL to Domain 4 (Port 112)
		Port 113	ISL to Domain 3 (Port 113)

d. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.23 as shown in the table below.

Slot #	SPPe #	Port #	Name
0	0	Port 0	ISL to Domain 12 (Port 0)
		Port 1	ISL to Domain 14 (Port 0)
		Port 2	Blue Host HBA 1
		Port 3	Blue Storage 1
0	1	Port 16	ISL to Domain 4 (Port 16)
1	0	Port 32	ISL to Domain 12 (Port 1)
		Port 33	ISL to Domain 14 (Port 1)
1	1	Port 48	ISL to Domain 4 (Port 48)
2	0	Port 64	ISL to Domain 11 (Port 3)
		Port 65	ISL to Domain 9 (Port 3)
2	1	Port 80	ISL to Domain 2 (Port 80)
		Port 81	ISL to Domain 1 (Port 81)
3	0	Port 96	ISL to Domain 11 (Port 2)
		Port 97	ISL to Domain 9 (Port 2)
3	1	Port 112	ISL to Domain 2 (Port 112)
		Port 113	ISL to Domain 1 (Port 113)

e. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.22 as shown in the table below.

Slot #	SPPe #	Port #	Name
0	0	Port 0	ISL to Domain 7 (Port 1)
		Port 1	ISL to Domain 5 (Port 1)
		Port 2	Green Host HBA 2
0	1	Port 16	ISL to Domain 1 (Port 16)
1	0	Port 32	ISL to Domain 7 (Port 0)
		Port 33	ISL to Domain 5 (Port 0)
1	1	Port 48	ISL to Domain 1 (Port 48)
2	0	Port 64	ISL to Domain 10 (Port 3)
		Port 65	ISL to Domain 6 (Port 3)
2	1	Port 80	ISL to Domain 3 (Port 80)
		Port 81	ISL to Domain 4 (Port 81)
3	0	Port 96	ISL to Domain 10 (Port 2)
		Port 97	ISL to Domain 6 (Port 2)
3	1	Port 112	ISL to Domain 3 (Port 112)
		Port 113	ISL to Domain 4 (Port 113)

f. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.200.23 as shown in the table below.

Slot #	SPPe #	Port #	Name
0	0	Port 0	ISL to Domain 9 (Port 0)
		Port 1	ISL to Domain 11 (Port 0)
		Port 2	Blue Host HBA 2
		Port 3	Blue Storage 2

Slot #	SPPe #	Port #	Name
0	1	Port 16	ISL to Domain 3 (Port 16)
1	0	Port 32	ISL to Domain 9 (Port 1)
		Port 33	ISL to Domain 11 (Port 1)
1	1	Port 48	ISL to Domain 3 (Port 48)
2	0	Port 64	ISL to Domain 14 (Port 3)
		Port 65	ISL to Domain 12 (Port 3)
2	1	Port 80	ISL to Domain 1 (Port 80)
		Port 81	ISL to Domain 2 (Port 81)
3	0	Port 96	ISL to Domain 14 (Port 2)
		Port 97	ISL to Domain 12 (Port 2)
3	1	Port 112	ISL to Domain 1 (Port 112)
		Port 113	ISL to Domain 2 (Port 113)

g. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.24 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 2 (Port 33)
	Port 1	ISL to Domain 2 (Port 1)
	Port 2	ISL to Domain 1 (Port 97)
	Port 3	ISL to Domain 1 (Port 65)
	Port 4	Red Host HBA 1

h. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.25 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 2 (Port 32)
	Port 1	ISL to Domain 2 (Port 0)
	Port 2	ISL to Domain 1 (Port 96)
	Port 3	ISL to Domain 1 (Port 64)
	Port 4	Red Storage 1

i. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.26 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 4 (Port 0)
	Port 1	ISL to Domain 4 (Port 32)
	Port 2	ISL to Domain 3 (Port 97)
	Port 3	ISL to Domain 3 (Port 65)
	Port 4	Green Storage 1

j. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.199.27 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 4 (Port 1)
	Port 1	ISL to Domain 4 (Port 33)
	Port 2	ISL to Domain 3 (Port 96)
	Port 3	ISL to Domain 3 (Port 64)
	Port 4	Green Storage 2

k. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.2009.24 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 1 (Port 33)
	Port 1	ISL to Domain 1 (Port 1)
	Port 2	ISL to Domain 2 (Port 97)
	Port 3	ISL to Domain 2 (Port 65)
	Port 4	Red Host HBA 2

l. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.2009.25 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 1 (Port 32)
	Port 1	ISL to Domain 1 (Port 0)
	Port 2	ISL to Domain 2 (Port 96)
	Port 3	ISL to Domain 2 (Port 64)
	Port 4	Red Storage 2

m. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.2009.26 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 3 (Port 0)
	Port 1	ISL to Domain 3 (Port 32)
	Port 2	ISL to Domain 4 (Port 97)
	Port 3	ISL to Domain 4 (Port 65)
	Port 4	Green Storage 3

n. Following Step a through Step c, configure the ports of the switch with an IP address of 172.23.2009.27 as shown in the table below.

ASIC	Port #	Name
0	Port 0	ISL to Domain 3 (Port 1)
	Port 1	ISL to Domain 3 (Port 33)
	Port 2	ISL to Domain 4 (Port 96)
	Port 3	ISL to Domain 4 (Port 64)
	Port 4	Green Storage 4

Connect cables

To connect cables:

1. Connect ISLs.
   a. Attach fiber cable between switches as shown in Figure 24 on page 234.
   b. After all cables are connected, use Connectrix Manager to verify that all ISL connections are up.
   c. Rearrange icons to accurately reflect the switch configuration and then persist the fabric.

   Note: After persisting the fabric, you can immediately detect if something has changed in the environment by looking at the topology view. If an ISL or device disappears, yellow alert icons display. Therefore, it is a recommended to *always* persist the fabric *after* changes have been made.

2. Connect the host and storage ports.
   a. Attach fiber cable between switches and N_Ports.

Zone hosts and storage

To zone hosts and storage:

1. To open the **Zoning** dialog in Connectrix Manager, right-click the appropriate fabric topology and then select the **zoning menu**.
2. Create a zone by clicking the **New Zone** button under the **Zones Tree**.

3. Provide a descriptive name for the zone. This case zones *Red host HBA* and *Red Storage 1*, so enter **RedHBA1_1470_8aa**, and then press **Enter**.
4. Locate, then click, **Red Host HBA 1** (WWPN 10000000c938e554) in the **potential zone members** list.
5. Click the right-pointing arrow on the divider between the **potential members** list and the **zones** list to add the HBA to the zone.
6. Select **Red Storage 1** (WWPN 50060482cc19bf87) in the **potential zone members** list.
7. Click the right-pointing arrow on the divider between the **potential members** list and the **zones** list to add the Storage port to the zone.
8. Repeat steps Step 2 through Step 7 for all host and storage pairs in the environment.
9. Create a zone set by clicking **New Set** under the zone sets **Tree**.
10. Provide a descriptive name for the zone set. This case uses the date of “Oct_31_06_1140”.
11. Add all of the new zones to the zone set. When completed, the zone set should be similar to the following:

```
Zone set name = "Oct_31_06_1140"
       Zone name = "RedHBA1_1470_8aa"
          Zone Member = "10000000c938e554"
          Zone Member = "50060482cc19bf87"

       Zone name = "RedHBA2_1470_9aa"
          Zone Member = "10000000c938e555"
          Zone Member = "50060482cc19bf88"

       Zone name = "BlueHBA1_1489_8aa"
          Zone Member = "210100e08b8ac76d"
          Zone Member = "50060482cc19c447"

       Zone name = "BlueHBA2_1489_9aa"
          Zone Member = "210100e08baac76d"
          Zone Member = "50060482cc19c448"

       Zone name = "GreenHBA1_AllGreenStorage"
          Zone Member = "10000000c939a051"
          Zone Member = "50060482cc19c407"
          Zone Member = "50060482cc19c408"
          Zone Member = "50060482cc19c4c7"
          Zone Member = "50060482cc19c4c8"
```

```
Zone name = "GreenHBA2_GreenStorage"
   Zone Member = "10000000c939a052"
   Zone Member = "50060482cc19c407"
   Zone Member = "50060482cc19c408"
   Zone Member = "50060482cc19c4c7"
   Zone Member = "50060482cc19c4c8"
```

Complete the SAN setup

The SAN is now ready to pass I/O from host to storage. Other steps, such as configuring LUN masking and modification of host configuration files, are required before the SAN Setup is complete. Refer to the Fabric OS configuration guide for more detail.

Configure security

Once the hosts and storage ports are properly configured, the fabric should be secured.

To enable Enterprise Fabric Mode:

1. From the **Connectrix Manage**r topology view, double-click on a switch in the fabric that needs to be secured.

   Note: Do *not* double-click and open the Element Manager.

2. Select **Enterprise Fabric Mode** from the **Configure** pull-down menu.
3. Click **Activate**.
4. Click **Close**.

To enable port binding:

1. From Connectrix Manager, double-click the switch with an IP address of **172.23.199.22** to open Element Manager.
2. From the **Configure** pull-down menu, select **Ports**.
3. Right-click in the **Port Binding** column and select **Bind All Ports To Attached WWN**.
4. Click **OK**.
5. Repeat this process for the other switches in the fabric.

Configure proactive monitoring and countermeasures

Configure the ED-10000Ms:

1. Telnet into the switch with an IP address of 172.23.199.22 by entering the following command:

   telnet 172.23.199.22

2. Enter the username and password when prompted.
3. Enter the **FC Performance** command.

Configure a throughput threshold alert:

1. Enter the command:

   config throughput 1 TTA portlist 5 5 both 80 all

Configure counter threshold alerts:

1. Enter the command:

   config counter 2 CTA1 portlist 5 CountBBCreditZero 1000000 all

2. Enter the command:

   config counter 3 CTA2 portlist 5 CountClass3Discards 100 all

3. Enter the command:

   config counter 4 CTA3 portlist 5 CountInvalidTxWords 40 all

4. Enter the command:

   config counter 5 CTA4 portlist 1440 CountInvalidTxWords 100 all

Activate all of the alerts:

1. Enter the **config activate 1** command.
2. Enter the **config activate 2** command.
3. Enter the **config activate 3** command.
4. Enter the **config activate 4** command.
5. Enter the **config activate 5** command.

Configure fencing:

By default there are three fencing policies defined; to enable them follow these steps.

1. Enter the command:

   fc portfencing enable "Default Security Policy".

2. Enter the command:

   fc portfencing enable "Default Link Level Policy".

3. Enter the command:

   fc portfencing enable "Default Protocol Error P".

Configure the other switches

Repeat all the steps in the"Configure proactive monitoring and countermeasures" on page 253 on the other ED-10000Ms in the fabric.

Configure the DS-4700Ms

Configure throughput threshold alerts:

To configure threshold alerts:

1. Open a command prompt and Telnet to the switch by typing **telnet 172.23.199.22**.
2. Log in with the username and password of the switch.
3. Enter the **Perf** command.
4. Enter the **thresh** command.
5. Enter the **throughput** command.
6. Enter the **addalert eightyPercent** command.
7. Enter the **addport eightyPercent all** command.
8. Enter the **setParams eightyPercent 1 5** command.
9. Enter the **setUtilPercentage eightyPercent 80** command.
10. Enter the **setutiltype eightyPercent 3** command.
11. Enter the **..** command.
12. Enter the **setstate eightyPercent 1** command.
13. Enter the **logo** command.

Configure counter threshold alerts:

Telnet into each DS-4700M and enter the following commands:

1. Enter **addAlert CTA1** command.
2. Enter **addPort CTA1 all** command.
3. Enter **setCounter CTA1 18** command.
4. Enter **setParams CTA1 10000 30** command.
5. Enter **addAlert CTA2** command.
6. Enter **addPort CTA2 all** command.
7. Enter **setCounter CTA2 10** command.
8. Enter **setParams CTA2 100 300** command.
9. Enter **addAlert CTA3** command.
10. Enter **addPort CTA3 all** command.
11. Enter **setCounter CTA3 9** command.
12. Enter **setParams CTA3 40 300** command.
13. Enter **addAlert CTA4** command.
14. Enter **addPort CTA4 all** command.
15. Enter **setCounter CTA4 9** command.
16. Enter **setParams CTA4 100 70560** command.

Enable the Counter threshold alerts:

1. Enter the **..** command.
2. Enter the **setState CTA1 1** command.
3. Enter the **setState CTA2 1** command.
4. Enter the **setState CTA3 1** command.
5. Enter the **setState CTA4 1** command.

Configure Port Fencing:

1. Click **Configure** and select **Port Fencing** from the list.
2. Enable the **Default Security Policy** by highlighting the entry and clicking **Enable**.
3. Enable the **Default Link Level Policy** by highlighting the entry and clicking **Enable**.

To disable unused interfaces:

1. From the **Connectrix Manager** topology view, select **Security**.
2. Ensure that the appropriate fabric is selected in the fabric list.
3. Under the **Authentication** tab, select a switch in the **Product Configuration** list.
4. Under the **Users** tab uncheck the **Enable Telnet** checkbox.

Heterogeneous switch interoperability

As SANs become larger and topologically more complex, it is increasingly advantageous to use switches from different vendors that work together. *Interoperability* is the term used to describe a Fibre Channel fabric that contains switches from more than one vendor. Interoperability has countless definitions in the IT connectivity space. For the context of this chapter, it refers to FC switch interoperability, which increases the ability and flexibility to design complex SANs. It takes into account the various features that different vendor switches provide as well as the specific features that end users look for while designing SANs.

This section provides an in-depth description for setting up an EMC-supported Fibre Channel SAN comprising of FC switches (director class, distribution class, departmental switches and blade server FC switch modules) from different switch vendors, such as Brocade, Cisco, Brocade M-Series, and QLogic. Switch migration procedures to move customers' existing SAN topology from one EMC-supported switch vendor type to another EMC-supported vendor type are also provided.

The information provided in this section has been obtained from vendor product documents and testing experiences in the EMC E-Lab qualification labs.

Interoperability overview

This section explains the concept and significance of heterogeneous interoperability in the FC SAN world.

E_Port interoperability

The approval of the FC-SW-2 standard has created an open standard for switch-to-switch communication, allowing end-users to select best-in-class products with assurance that these products will work together. Currently, most SAN switch vendors offer FC-SW-2

compliant switches including, but not limited to, Brocade, Cisco, Brocade M-Series, and QLogic.

Multi-vendor FC switch interoperability is also referred to as E_Port interoperability, since two switches are connected to each other using their respective E_Ports per FC standards.

Heterogeneous interoperability in EMC context

EMC supports the following vendor switches in a SAN deploying switch interoperability: Brocade, Cisco, Brocade M-Series, and QLogic. Most of the E_Port switch connectivity-based testing between these multi-vendor switches, for different kinds of switch hardware and firmware, is conducted in the EMC E-Lab qualification labs.

The testing mainly involves the validation of:

- link initialization between switches from different vendors
- name server distribution
- zoning changes
- the routing protocols
- fabric management application.

Interoperability support for switches that are compatible with the FC-SW standard are listed in the "Switched Fabric Topology Parameters" section of the*EMC Support Matrix*.

This section comprises of the following attributes (or columns):

- name of the switch
- name of the switches it is interoperable with
- firmware versions tested on the switch and the interoperable switches
- switch (fabric) management application revision
- maximum number of domains per fabric
- maximum number of hops
- maximum domain-to-domain ISLs supported by EMC.

This information is based on the qualifications conducted at E-Lab.

Although the support matrix can only directly represent a two-way interop, it may also be used to represent a three-way interop. Consider there are three switches, each from a different vendor, running firmware versions A, B, and C, respectively.

If the following entries are present in the support matrix:

- A is interoperable with B
- A is interoperable with C
- B is interoperable with C

then one can infer that switches running firmware versions A, B, and C are interoperable and can co-exist in the same fabric.

Significance of a multi-vendor switch configuration

There are at least three user scenarios that indicate the significance of a multi-vendor switch environment:

User scenario 1: Complex SANs

As SANs become larger and topologically more complex, it is getting increasingly useful to get switches from different vendors to work together. As discussed above, there are various vendors in the FC industry, designing and manufacturing switch hardware with different hardware features, protocol related-features, and performance (2 G / 4 G / 10 G). In such a case, end users must be able to select the best in-class products to design a SAN that meets their needs in terms of features and performance.

User scenario 2: Blade servers

Blade server technology is becoming widely accepted, but it is still evolving. At the time of this publication, only some switch vendors, such as Brocade, Brocade M-Series, and QLogic, are manufacturing FC switch modules that can be plugged into the backend of a blade server. A blade server accesses a switched fabric or storage it is being hooked up to using FC switch modules (as discussed in detail in "Blade servers" on page 1296). Hence, if end users have a switch-vendor fabric which is neither of the above switch module manufacturing vendors, interoperability between multi-vendor switches is an important aspect to consider before building such a fabric.

User scenario 3: Switch migration

Depending on their requirements, some end users move their SAN from one switch vendor type to another switch vendor type. This switch migration procedure needs to be executed in phases. One of the transitional phases requires the vendor type switches (one type before and the other after the migration) to be in the same fabric, making interoperability between multi-vendor switches an aspect to be considered before starting a switch migration procedure.

Heterogeneous SAN design

This section provides details on setting up an interoperable SAN environment, and studies different combinations of some of the EMC-supported heterogeneous SANs.

In addition to providing a detailed step-by-step approach to design some specific type of SANs, this section is intended to give you an insight into the kind of configurations previously tested and supported by E-Lab, and to provide some of the best practices and caveats that must be considered while designing these SANs.

Components needed to create a heterogeneous design

The concept of simple SANs and Complex SANs has been discussed in detail in "Fibre Channel topology overview" on page 36. Most of the discussed topologies comprising of at least two switches are applicable to designing heterogeneous SANs. It is important to note that:

- In the heterogeneous SANs, the switched fabric must be comprised of switches from two or more different vendors.
- It is essential to set the appropriate operational mode on the switches so that they can communicate with any other switch from a different vendor in the same fabric.

How to set up an interoperable switched fabric topology

In order to address the individual vendor switch settings for supported switch interoperability configurations, a cookbook approach is used to explain a seven phase switch vendor migration process. The assumptions for this seven phase migration process follows.

Assumptions for the 7 phase migration process

- In all cases, the case studies start with a homogeneous single vendor fabric.
- Compound core edge topology is used for these case studies.
- The switch migration process involves a step to move the edge switches from the same vendor type core switches to a different vendor type core switches.

  Note: This process is helpful to end users who want to move their fabrics from one vendor switch core type to another vendor switch core type.

- One of the steps in the migration involves a stage when different vendor switches co-exist in the same fabric.

Note: This can help end users set up a supported interoperability fabric based on the switch configuration settings from different vendors that can co-exist in a stable fabric.

Phases

The seven phase process that has been appointed across all the case studies is explained in detail in the following sections. This example shows a transition from a switched fabric topology using Vendor A switches, to a topology with Vendor B switches.

Phase 1:

This is the base configuration phase with the pre-existing Vendor A switched fabric topology. Some default settings on these switches need to be changed before introducing a Vendor B switch in this fabric.

Phase 2:

In this phase, a Vendor B core switch is added to the homogeneous Vendor A fabric. This implies that all the Vendor A edge switches now have ISLs to this core Vendor B switch.

Phase 3:

Half of the zoned initiator (hosts) and target (storage) pairs from a Vendor A core switch are moved to the Vendor B core switch.

Phase 4:

All the other initiator-target pairs from the Vendor A core switches are completely moved to the Vendor B core switch.

Note: Phase 3 and 4 are executed in steps to avoid any disruption of traffic between the host and storage during the transition and to avoid any downtime.

Phase 5:

One or more Vendor B switches are added to the edge of the fabric.

Phase 6:

Hosts and storage are moved from any Vendor A edge switches to the new Vendor B edge switches.

Phase 7:

More Vendor B switches are added, if required, to the core of the fabric.

Case studies

Based on the switch vendor types and the interoperability modes currently tested and supported by EMC, five case studies are described in this section.

Case study #1: Migrating from a Connectrix M homogeneous fabric in McDATA fabric mode to a Connectrix MDS fabric.

Assumptions specific to this case study:

- Interoperability mode settings on the switches:
  - ***Brocade M-Series Fabric 1.0*** mode on the Connectrix M switches in the fabric.
  - **Cisco *Interop-4*** mode on the Connectrix MDS switches in the fabric.

  Note: At the time of this release, EMC supports Interop-4 using RPQ for migration purposes only. Please refer to the *EMC Support Matrix* for the most up-to-date support.

- Fabric management applications used for managing the fabric:
  - Cisco Fabric Manager, Connectrix Manager, Connectrix Manager, Connectrix Manager Basic, or Connectrix Manager Basic
  - Phase 1: Base configuration — Pre-existing Connectrix M-Series Core-Edge Fabric

Phase 1: Basic configuration topology

Topology

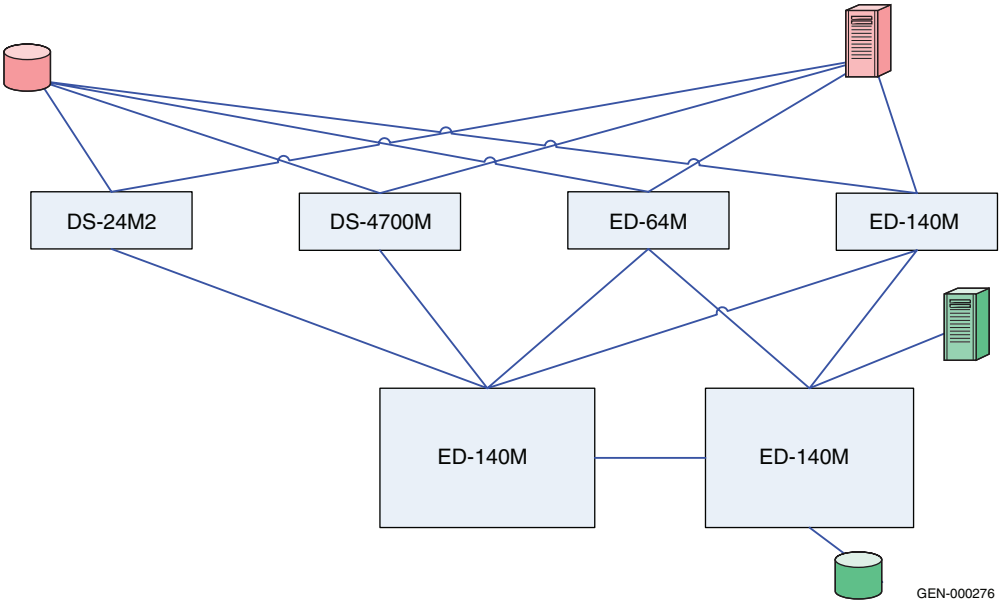


Figure 26 **Phase 1: Basic configuration topology**

As illustrated in Figure 26 on page 262, the two ED-140M director switches are at the core of the fabric while the departmental switches, DS-24M2 and DS-4700M) and director switches, ED-64M and ED-140M, are at the edge. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* for a list of all the other Brocade M-Series / EMC Connectrix M switches that can be supported in a heterogeneous setup, and for the operating modes specified above for this case study.

Note: At the time of this publication, the ED-10000M (or the Brocade M-Series 10000) and the Blade server Brocade M-Series modules (QLogic modules in Brocade M-Series mode) are not supported in a Connectrix MDS (Interop-4) - Brocade M-Series (Brocade M-Series mode 1.0) environment, and thus cannot participate in the switch migration procedure discussed in this case study.

At the time of this release, EMC supports Interop-4 using RPQ for migrating purposes only. Please refer to the *EMC Support Matrix* for the most up-to-date support.

The specific configuration settings, best practices, host and storage layouts, and topology design to withstand failures for a Connectrix M core-edge homogeneous fabric discussed in the "Connectrix M example" on page 63 applies to this base topology.

Checkpoints

Before adding the Connectrix MDS core MDS 9513 director to the homogeneous Brocade M-Series fabric, it is advisable to verify the following fabric characteristics using either Connectrix Manager, Connectrix Manager Basic, or the Brocade M-Series CLI:

- **Proper distribution of the name server information**.

  Verify that the all the host HBA ports and storage ports logged into the fabric are listed in the name server.

- **Proper distribution of zoning information.**

  Verify that the active zoneset comprises zones that contain the desired mapping of host and storage ports.

- **Proper display of fabric and N_Port elements on the management applications**.

  Verify that you have the desired physical topology and domain count of the fabric.

- **No disruption in data transfer**.

  Verify that the data traffic is running appropriately through the SAN.

Phase 2: Adding Connectrix MDS 9513 to the core of the fabric

Topology

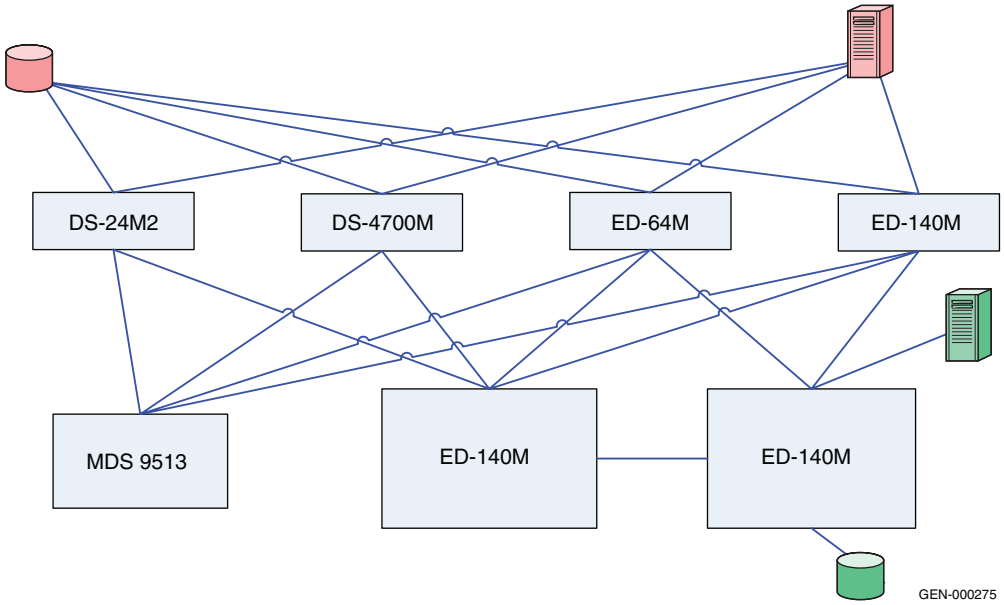


Figure 27 **Phase 2: Adding Connectrix MDS 9513**

As illustrated in Figure 27, the Connectrix MDS 9513 director is added to the core of the fabric. It is important to note that there is no ISL between the Brocade M-Series cores and the Connectrix MDS 9513. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* to obtain a list of all the other Connectrix MDS series switch that can be supported in a heterogeneous set up with the Brocade M-Series switches, and for the operating modes specified above for this case study.

Note: At the time of this release, EMC supports Interop-4 usingRPQ for migration purposes only. Refer to the *EMC Support Matrix* for the most up-to-date support.

Before adding the Connectrix MDS 9513 to the fabric, perform these steps on the Connectrix MDS switch:

1. Create a VSAN on the Connectrix MDS switch.

   A virtual SAN can be used to create multiple logical SANs over the same physical infrastructure.

a. Click **Create VSAN** on the Cisco Fabric Manager.

b. When the **Create VSAN** window appears, check the desired switches to included in the VSAN. In this case, only the MDS 9513.

c. Fill in the **VSAN ID** field with an unused ID number and the VSAN **name** filed with an appropriate name. In this case study, **VSAN ID number = 801**.

d. For VSN attributes, leave everything other than the **InterOperValue** and the **AdminState** at default.

e. Set **InterOperValue** to **Interop-4** and the **AdminState** to **Suspended**.

f. Statically assign VSAN membership for an interface using Fabric Manager, by selecting **Interfaces > FC Physical** from the **Physical Attributes** pane.

   The interface configuration in the **Information** pane appears. Select the **General** tab and double-click to complete the **PortVSAN** field by entering the VSAN ID number (**801**) for every desired port for this fabric.

   For more details on the VSAN settings for Connectrix MDS 9000 family switches refer to the Cisco document: http://www.cisco.com/en/US/products/ps5989/products_configuration_guide_chapter09186a0080662ce3.html#wp1094055

   For configuring Interop-4 specific settings on the Cisco CLI refer to the following Cisco document: http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_configuration_guide_chapter09186a008065cc48.html

   Interop mode 4, also known as the legacy switch Interop mode 4, provides easy integration between VSANs and Brocade M-Series switches running in Brocade M-Series Fabric 1.0 interop mode. It is initially available with Connectrix MDS 9000 SAN-OS v. 3.0(1). Thus, there is no need to reset or restart Brocade M-Series switches in this case study since they can continue to operate in Brocade M-Series fabric mode 1.0 which avoids fabric disruption.

2. To set up interoperability configuration compliant with Connectrix M switches operating in Brocade M-Series Fabric Mode 1.0:

a. Select **VSAN 801** from the **Logical Domains** menu on the Fabric Manager.

b. Select **VSAN** attributes. VSAN attributes can be seen in the information pane.

c. Select **Interop-4** from the interop drop-down menu.

d. Click **Insert Row**.

3. To set up interoperability configuration compliant with Connectrix M switches operating in Brocade M-Series Fabric Mode 1.0:

   a. Replace the local switch WWN with the Brocade M-Series OUI.

      A VSAN running in Interop mode 4 must be configured to use a WWN that follows a specific set of rules regarding the following Brocade M-Series OUI format:

      The changed WWN with the Brocade M-Series OUI must appear as xy:yy:08:00:88:zz:zz:zz.

      – x = 1 or 2
      – y = 0:00 or hex representation of the VSAN number. (For example, 0:0a for VSAN 10.)
      – 3rd, 4th and 5th bytes of the WWN must be 08:00:88 (Brocade M-Series OUI).
      – Recommended: z = last three bytes of the switch WWN.
      – The chosen WWN must be unique across the fabric.

      Once completed, the WWN has a Brocade M-Series OUI and the WWN can conflict only with Brocade M-Series switches or other Connectrix MDS switches with VSANs running in Interop mode 4.

   b. Set a static Domain ID:

      If you configure a static Domain ID using **Domain Manager** on the **Logical Domains** menu of the Connectrix MDS switch Fabric Manager, it must be in the 1 to 31 range for the Interop mode 4 VSAN. However, persistent FC IDs must be configured in the range of 97 to 127. A Connectrix M switch refers to Domain IDs in the range of 1 to 31, while FC IDs are used within the range of 97 to 127, with an offset of 96. The Interop mode 4 VSAN emulates the same behavior as the Connectrix M switch. Therefore, when configuring static Domain IDs on an Connectrix MDS switch, the range of 1 to 31

should be used; however, devices that log into the VSAN (even physically located on the Connectrix MDS switch) receive an FC ID whose Domain ID is 97-127. The DID range for Connectrix MDS in interop-4 is thus reduced from 1-239 to 1-31.

4. To set ports on the Device Manager for the Connectrix MDS 9513:
   a. Select the desired switch, in this case Connectrix MDS 9513. The **Device Manager** appears.
   b. Select the ports participating in this fabric, and set the port speed to **automax 2G**. Leave the other settings as default.
   c. Set the **Admin** option to **up**.
5. Unsuspend the VSAN.

   After configuring all the settings as stated in Step 4, go back to the **VSAN attributes** and change the **AdminState** to **Active**.

IMPORTANT

Before connecting the Connectrix MDS switch to the Connectrix M switches through an ISL, it is important to verify that all the Connectrix M switches and the Connectrix MDS switches are running the supported firmware versions for the respective interop modes. To verify the versions, refer to the *EMC Support Matrix*.

6. Create an ISL between the Connectrix MDS 9513 and all the Connectrix M edge switches as per Figure 27 on page 264.
7. Verify connectivity of the fabric by validating the same set of *checkpoints* as listed for "Checkpoints" on page 263. In this case, Cisco Fabric Manager and Cisco CLI can be used to verify the fabric topology and name server information.

IMPORTANT

At the end of this phase, it is highly recommended you verify that zoning information was distributed appropriately to the Connectrix MDS core. You must compare the active zone set on the Connectrix M and Connectrix MDS fabric management tools to verify that there are no differences.

On Connectrix Manager, a Connectrix MDS switch appear as a Connectrix M switch. However, by clicking **Discover > Setup** on

Connectrix Manager, and adding the IP of the Connectrix MDS switch, the Connectrix M switch is discovered as a Connectrix MDS switch, thus avoiding any kind of confusion that can be caused by seeing other switches marked as Connectrix M.

Phase 3: Moving half of the host and storage ports from Connectrix M core to the Connectrix MDS 9513

Topology

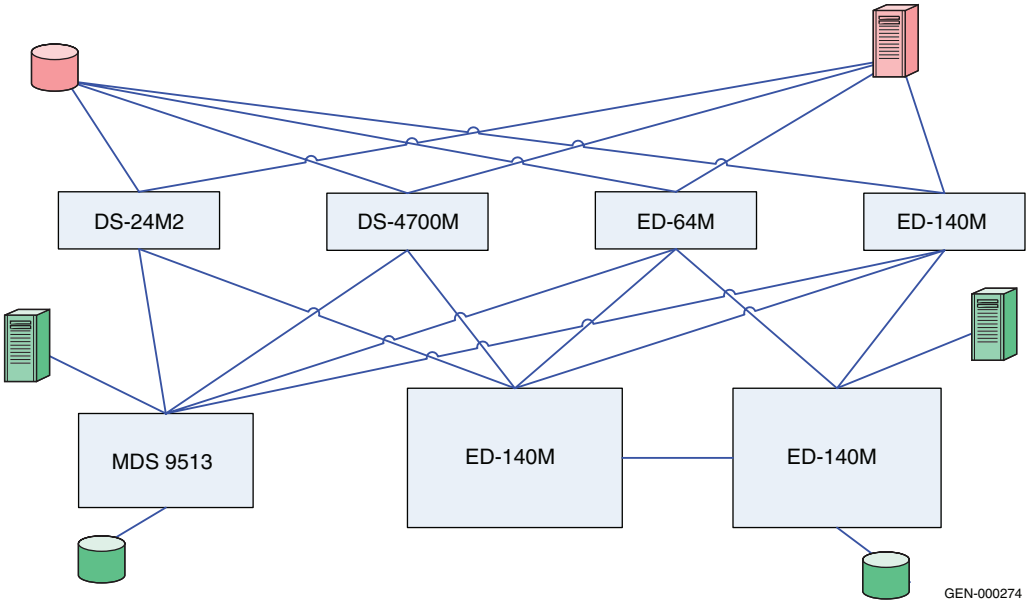


Figure 28 **Phase 3: Moving half of host and storage ports**

As illustrated in Figure 28, the same host-target pair is now connected to both core switches: the ED-140M and the Connectrix MDS 9513.

This phase is an intermediate phase to completely moving the host and storage ports to the Connectrix MDS 9513. This phase ensures that the zones are appropriately pushed from the Connectrix M to the Connectrix MDS switch and that the traffic between the host and storage zoned together is not disrupted. The host and storage ports can be moved successfully to another vendor core switch without any downtime.

There are not any specific steps to be executed in this phase other than physically pulling cables from one switch and plugging them

into the other core switch. However, it is recommended to go through "Checkpoints" on page 263 and validate that this transition did not affect the connectivity and functioning of the fabric.

Phase 4: Completely moving the host and storage ports from Connectrix M core to the Connectrix MDS 9513

Topology

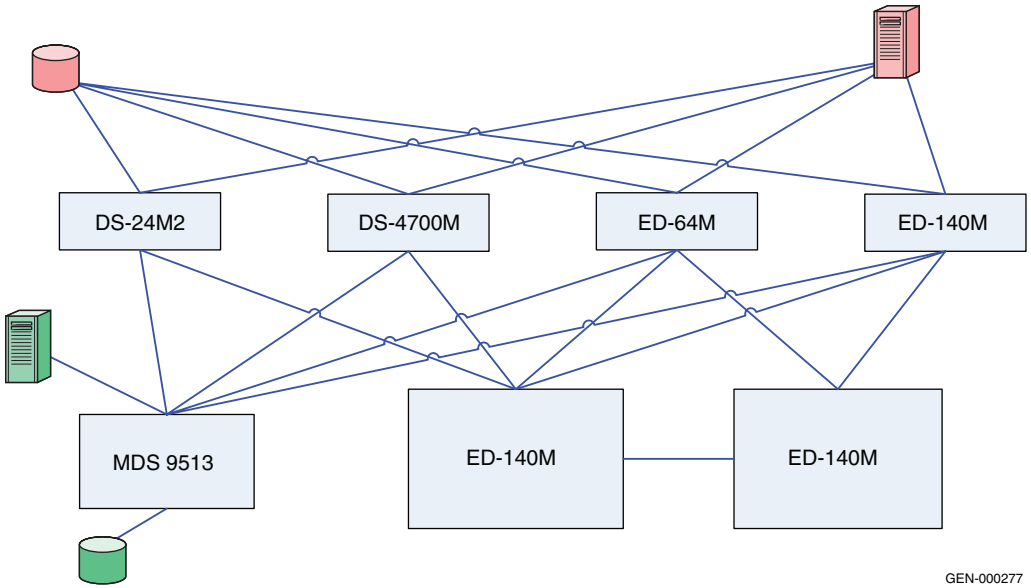


Figure 29 **Phase 4: Complete moving host and storage ports**

This is an extension of "Phase 3: Moving half of the host and storage ports from Connectrix M core to the Connectrix MDS 9513" on page 268. Once a stable fabric is validated at the completion of Phase 3, Phase 4 can be executed by pulling the remaining host-storage pairs from the Connectrix ED-140M switch and transferring them to the Connectrix MDS 9513 core.

Again, it is recommended to go through the "Checkpoints" on page 263 using both the Connectrix M and Connectrix MDS fabric management tools.

Phase 5: Adding a Connectrix MDS 9216 to the edge

Topology

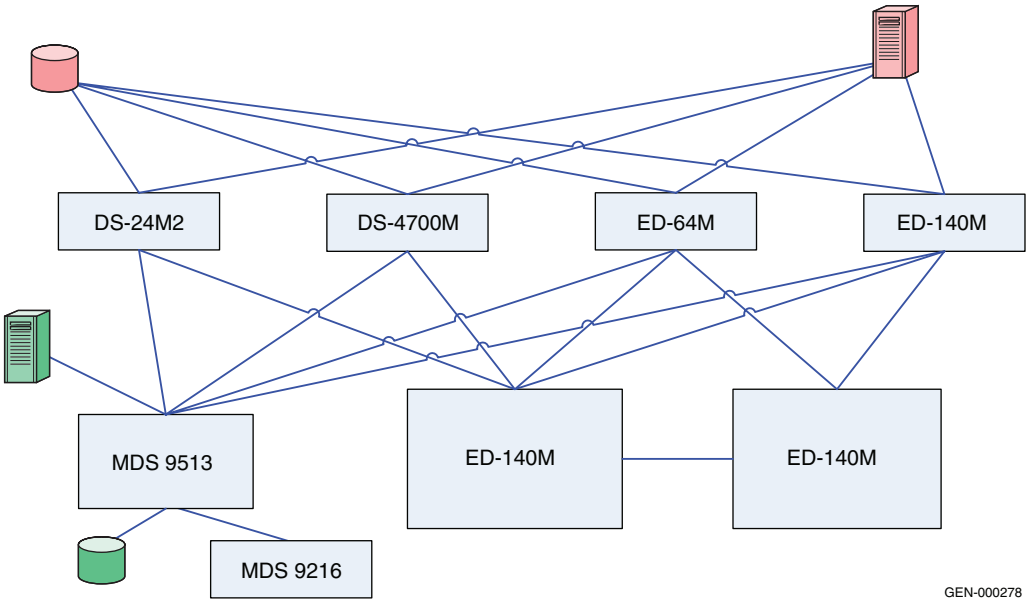


Figure 30 **Phase 5: Adding Connectrix MDS 9216**

To add the Connectrix MDS 9216:

Step 1 through Step 5 in "Phase 2: Adding Connectrix MDS 9513 to the core of the fabric" on page 264 for configuring the Connectrix MDS 9513, must be executed *exactly* as stated. Upon connecting these switches with matching VSAN IDs, the zoning information merges. After executing these steps, you must create an ISL between the edge switch and both the Connectrix MDS and Connectrix B core switches.

Phase 6: Moving hosts and storage to new edge

Topology

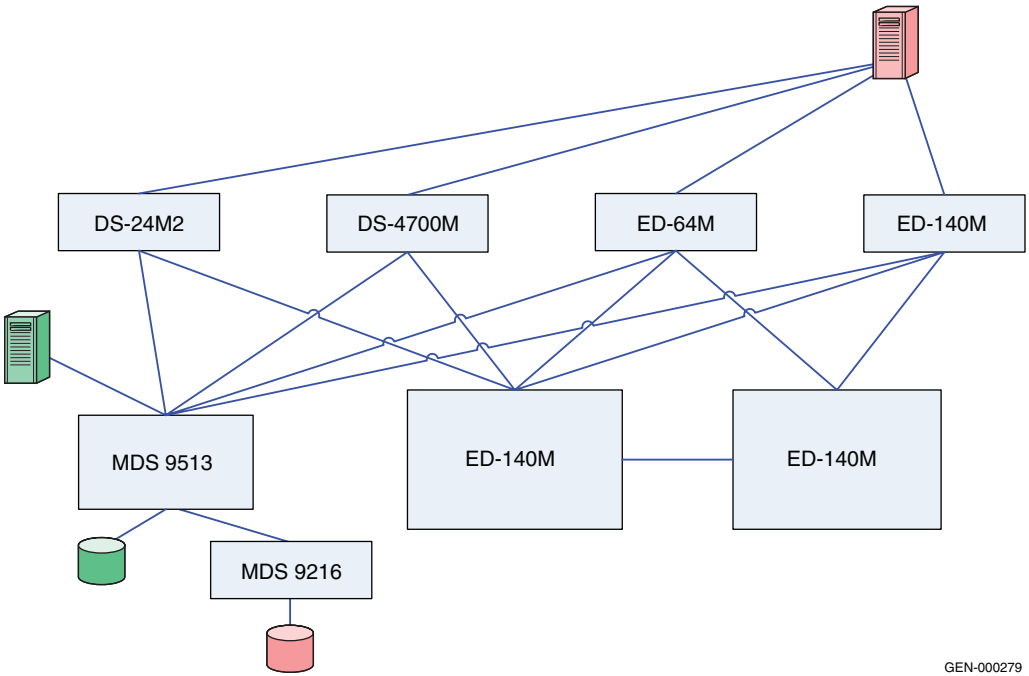


Figure 31 **Phase 6: Moving hosts and storage to new edge**

Hosts and storage ports shared by the other Connectrix M edge switches in the fabric can be completely or partially moved to the Connectrix MDS 9216 edge switch. It is evident in this transitional phase that both switches, Connectrix M and Connectrix MDS, can co-exist in a stable fabric, with the Connectrix M operating in McDATA fabric mode 1.0 and the Connectrix MDS operating in Interop-4 mode.

Note: At the time of this release, EMC supports Interop-4 using RPQ for migration purposes only. Please refer to the *EMC Support Matrix* for the most up-to-date support.

The settings on the switches in this phase can be used to configure a Connectrix M – Connectrix MDS interoperability fabric. All the non-default settings discussed in "Phase 2: Adding Connectrix MDS 9513 to the core of the fabric" on page 264 for the Connectrix MDS

9513, Connectrix MDS 9216, and the Connectrix M switches apply to setting up a heterogeneous Connectrix MDS-Connectrix M fabric.

Phase 7: Adding a Connectrix MDS switch to the core Topology

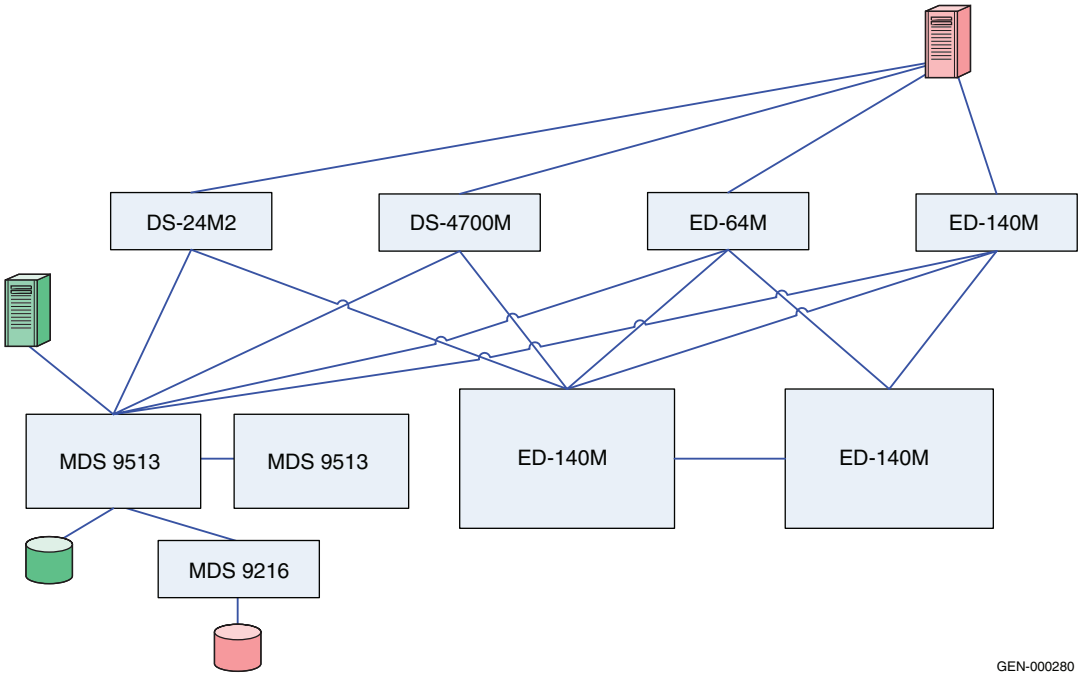


Figure 32 **Phase 7: Adding Connectrix MDS switch to the core**

As illustrated in Figure 32, another Connectrix MDS switch, the Connectrix MDS 9513 can be added to the core with similar settings as the previous Connectrix MDS 9513; you can create an ISL to the existing core Connectrix MDS 9513 director and the other edge switches in the fabric.

Note: To ensure a clean fabric merge, the VSAN must be configured with the same VSAN ID and attributes as the existing VSAN.

Complete migration to Connectrix MDS

At the end of case study 1, a migration from a Connectrix M-only fabric to a Connectrix M-Connectrix MDS fabric with Connectrix MDS switches at the core, connected through ISLs to every edge

switch in the fabric, occurred. The host and storage ports (except for the blade server host ports) can be completely moved over to the Connectrix MDS switches as specified in Phase 3, 4 and 6. The Connectrix M edge switches (except for the blade server Brocade M-Series switch modules), and the Connectrix M core switches can then be pulled out from the fabric. A fully operational Connectrix MDS-only fabric exists when completed. This is a complete migration from one switch type (Connectrix M) to another type (Connectrix MDS).

IMPORTANT

In a Connectrix MDS-only fabric, it is recommended that all Connectrix MDS switches operate in (default) mode.
At the end of the migration explained in this case study, Connectrix MDS switches are in interop mode.

In order to change interop mode on the Connectrix MDS switches, do not reboot the switches. The **Interop mode** attribute on currently active VSANs must be changed from **Interop-4** to **Default**.

Note: At the time of this release, EMC supports Interop-4 using RPQ for migration purposes only. Please refer to the *EMC Support Matrix* for the most up-to-date support.

However, active zones that were pushed to the Connectrix MDS from the Connectrix M cannot be modified once interop mode on the Connectrix MDS is changed. Zoning must be reconfigured. It is highly recommended that you back up the configuration to avoid losing all zoning information in the active zoneset on Connectrix MDS switches.

Warnings or caveats

Refer to EMC Knowledgebase solution emc149735 for all interop issues.

Case study #2 Setting up a heterogeneous switched fabric with Connectrix M switches in Open Fabric mode and the Connectrix MDS switches in Interop-1 mode.

Assumptions specific to this case study:

- Interoperability mode settings on the switches:
  - ***Brocade M-Series Open Fabric 1.0*** mode on Brocade M-Series switches in the fabric.

- **Cisco *Interop-1*** mode on Cisco switches in the fabric.

◆ Fabric management applications used for managing the fabric:

- Cisco Fabric Manager, Connectrix Manager, or Connectrix Manager Basic

If you migrate a Connectrix M fabric to a Connectrix MDS fabric (when the ED-10000M is not a part of the fabric), keep the fabric in Brocade M-Series Mode and follow the migration process as explained in "Case study #1:" on page 261. In this case study, when migrating from a Connectrix M fabric to a Cisco fabric, the non-default settings must be configured on Connectrix M and the Connectrix MDS switches before they are linked with an ISL.

Phase 1: Base configuration — Pre-existing Connectrix M core-edge fabric

Topology

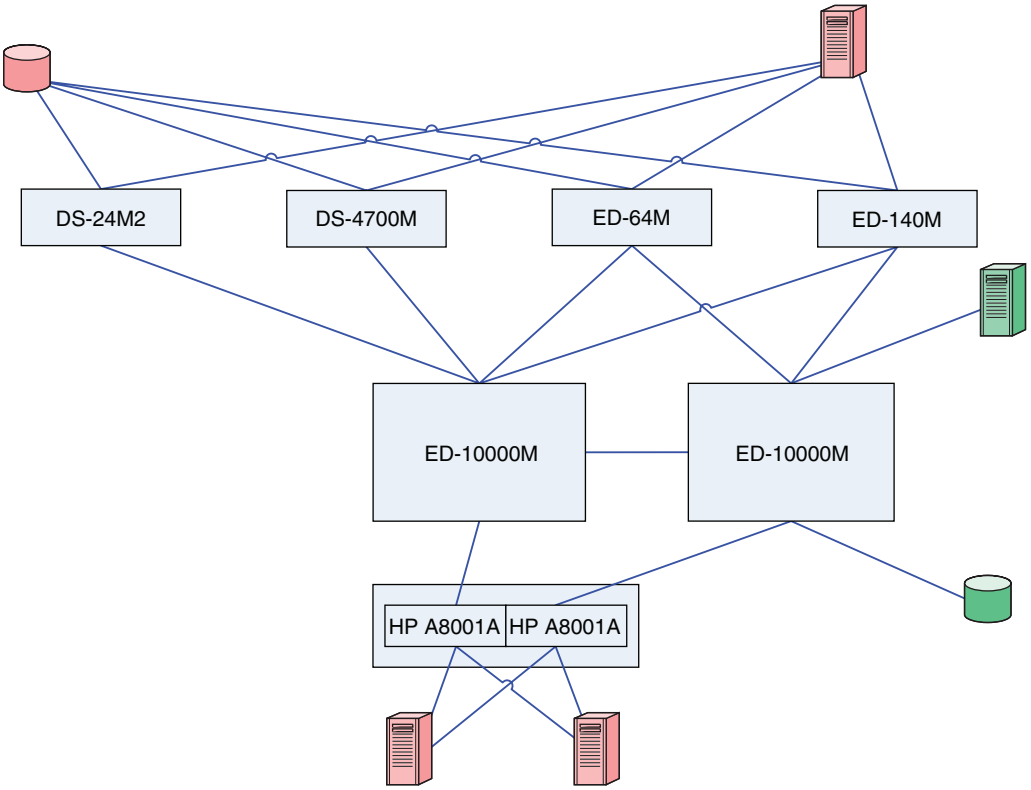


Figure 33 **Phase 1: Base configuration**

As can be seen in Figure 33 on page 274, the two ED-10000M director switches are at the core of the fabric while the departmental switches: the DS-4500M, DS-4700M, and director switches: ED-64M and ED-140M, are at the edge. The Blade server Brocade M-Series switch modules are also included as edge switches in this topology. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* to obtain a list of all the other Brocade M-Series/EMC Connectrix M switches that can be supported in a heterogeneous setup, and for the operating modes specified above for this case study.

Specific best practices, host and storage layouts, and topology design to withstand failures for a Brocade M-Series core-edge homogeneous fabric discussed in the "Connectrix M example" on page 63 applies to this base topology.

In this case study, unlike "Case study #1:" on page 261, Connectrix M switches will be set in Open Fabric Mode or interop mode, ISLs will be created to link the Connectrix M switches to the Connectrix MDS switches, which will be operating in their supported Interop-1 mode, and then the Connectrix M cores will be pulled out of the fabric. Considering this is a Connectrix M-only fabric with all switches operating in the Brocade M-Series Fabric Mode by default, the following non-default settings must be configured on the Connectrix M switches using the Connectrix Manager:

Before adding the Connectrix MDS 9513 to the fabric, the following steps need to be executed on the Connectrix MDS switch:

1. Verify firmware versions.

IMPORTANT

Before merging the Connectrix M and Connectrix MDS switches, verify that the firmware levels running on these switches are supported in the *EMC Support Matrix* for the respective interop modes. It is also recommended to upgrade the switches to the latest supported firmware versions listed in the *EMC Support Matrix*. You must also check that the correct version of the Connectrix Manager application is installed for fabric management.

2. Avoid duplicate Domain IDs in fabric.

Before merging the fabrics, if the Connectrix MDS switch Domain IDs are known, it must be ensured that no duplicate Domain IDs exist on the Connectrix M switches. Connectrix M and Connectrix MDS use different numbering schemes for their Domain IDs. Table 3 provides the inter-company Domain ID correlation scheme:

Table 3 Domain ID correlation

Brocade M-Series Domain ID	Cisco Domain ID
1	97
2	98
3	99
4	100
5	101
6	102
7	103
8	104
9	105
10	106
11	107
12	108
13	109
14	110
15	111
16	112
17	113
18	114
19	115
20	116
21	117
22	118
23	119
24	120

Table 3 Domain ID correlation (continued)

Brocade M-Series Domain ID	Cisco Domain ID
25	121
26	122
27	123
28	124
29	125
30	126
31	127

If there is a Connectrix M fabric that consists of Connectrix M switches with Domain IDs of 1 and 2, and another Connectrix MDS fabric that consists of Connectrix MDS switches with Domain IDs of 97(1) and 98(2), an attempted fabric merge will fail because duplicate Domain IDs will cause the fabrics to segment.

How to configure the preferred Domain ID and set it as a persistent ID using Connectrix Manager is addressed in "Configure the first virtual switch" on page 109.

3. Verify Interop Mode.

   The correct interoperation mode: **Open Fabric Mode 1.0** must be set on the Connectrix M switches in the fabric before merging them with Connectrix MDS. The Interop mode can be verified and changed using Connectrix Manager as follows:

   a. Before changing the interop mode for a Brocade M-Series switch, the switch must be set "offline".

   b. At the **Connectrix Manager** window, right click the appropriate product icon at the physical map. A pop-up menu appears. Select **Element Manager** from the menu.

      When the **Element Manager** application opens, the last view (tab) accessed by an end-user opens by default.

   c. At the **Configure** menu in the **Element Manager**, select **Operating Parameters**, then **Fabric Parameters**. The **Configure fabric Parameters** dialog box appears.

   d. At the **Interop Mode** drop-down list, select **Open Fabric 1.0** and click **Activate**.

4. Disable Rerouting Delay for the Connectrix M switches.

   In an interoperability scenario, the **Rerouting Delay** on Brocade M-Series switches must be disabled using the Connectrix Manager management application as follows:

   a. At the **Connectrix Manager** window, right-click the appropriate product icon at the physical map. A pop-up menu appears. Select **Element Manager** from the menu.

      When the **Element Manager** application opens, the last view (tab) accessed by an end-user opens by default.

   b. At the **Configure** menu in the **Element Manager**, select **Operating Parameters**, then **Switch Parameters**. The **Configure Switch Parameters** box will be displayed.

   c. Ensure the **Rerouting delay** checkbox is *not* selected (checked), then click **Activate**.

5. Set Connectrix M port speed and port type for Connectrix MDS.

   Ensure that the port speeds and port types are configured equally between the Connectrix M switches and the Connectrix MDS switches to be merged. The port speed may be set to **auto-negotiate** and the ports to be linked with ISLs to each other can be locked as **E_Ports**. The port configuration steps using the Connectrix Manager management GUI has been explained in "Configure the first virtual switch" on page 109.

For more details on the settings to be applied to the Connectrix M switches before merging them with a Connectrix MDS fabric are provided in the following document:

```
http://www.cisco.com/en/US/products/hw/ps4159/ps4358/
   products_configuration_guide_chapter09186a00805c7a40.
   html#wp1040185
```

Checkpoints

Before adding the Connectrix MDS core Connectrix MDS 9513 director to the homogeneous Connectrix M fabric, it is advisable to verify the following fabric characteristics, using either the Connectrix Manager or the Brocade M-Series CLI.

- **Proper distribution of the Name Server information**.

  Verify that the all the host HBA ports and storage ports logged into the fabric are listed in the name server.

- **Proper distribution of Zoning information**.

  Verify that the active zoneset comprises of the zones that contain the desired mapping of host and storage ports.

- **Proper display of fabric and N_Port elements on the management applications**.

  Verify that the physical topology and domain count of the fabric is as desired.

- **No disruption in data transfer**.

  Verify that the data traffic is running appropriately through the SAN.

Phase 2: Adding the Connectrix MDS 9513 to the core of the fabric

Topology

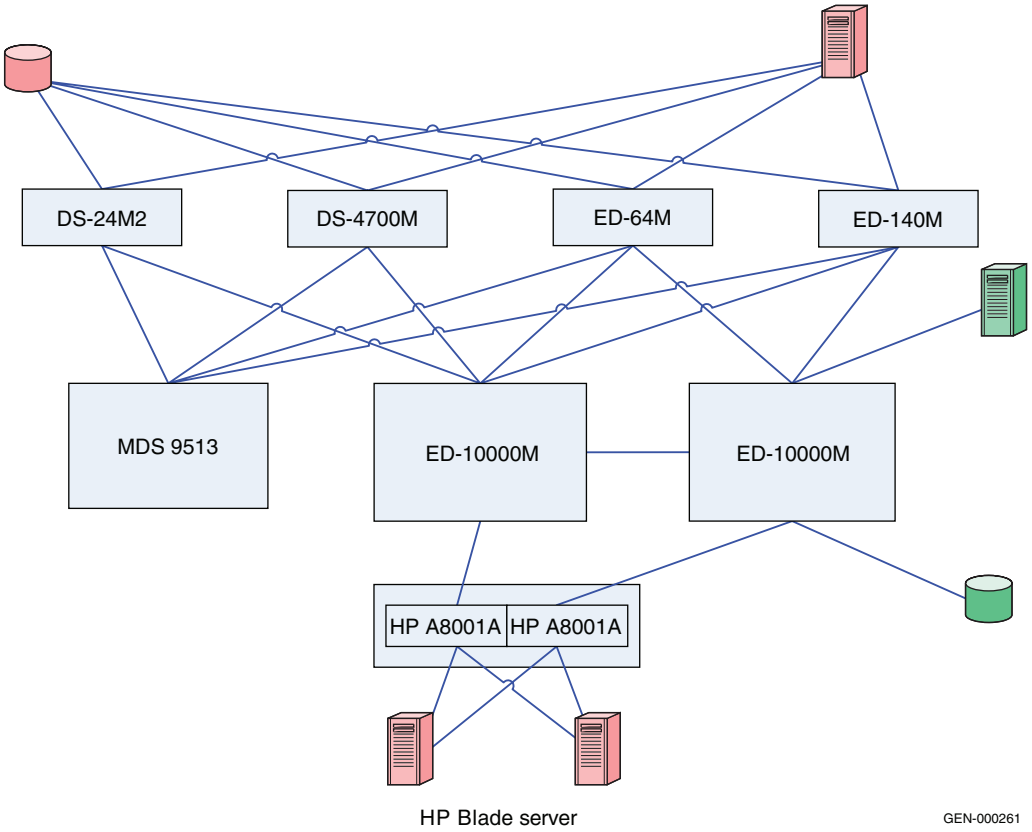


Figure 34 **Phase 2: Adding Connectrix MDS 9513 to the core of the fabric**

As can be seen in Figure 34, the Connectrix MDS 9513 director is added to the core of the fabric. It is important to note that there is no ISL between the ED-10000M core and the Connectrix MDS 9513. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* to obtain a list of all the other Connectrix MDS-series switches that can be supported in a heterogeneous set up with the Connectrix M switches, and for the operating modes specified above for this case study.

Before adding the Connectrix MDS 9513 to the fabric, the following steps need to be executed on the Connectrix MDS switch:

1. Create a VSAN on the Connectrix MDS switch.

   A Virtual SAN can be used to create multiple logical SANs over the same physical infrastructure.

   a. Click **Create VSAN** on Cisco **Fabric Manager**.

   b. Once the **Create VSAN** window appears, select the switches to include in the VSAN. In this case, only the Connectrix MDS 9513 is included.

   c. Fill in the **VSAN ID** field with an unused ID number and the **VSAN name** filed with an appropriate name. In this case study, **VSAN ID number = 801** is assigned.

   d. For **VSAN** attributes, leave everything other than the **InterOperValue** and the **AdminState** at default. VSAN attributes can be seen in the information pane.

   e. Set **InterOperValue** to **Interop-1** and the **AdminState** to **Suspended**.

   f. Statically assign VSAN membership for an interface using Fabric Manager, by choosing **Interfaces > FC Physical** from the **Physical Attributes** pane. The interface configuration appears in the **Information** pane. Select the **General** tab on this window, and double-click and complete the **PortVSAN** field by entering the **VSAN ID** number (801) for every port desired to be used by this fabric.

      For more details on the VSAN settings and for configuring the Interop-1 specific settings for Connectrix MDS 9000 family switches refer to the following Cisco document: http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_configuration_guide_chapter09186a00805c7a40.html#wp1040185

2. To set port settings on the Device Manager for the Connectrix MDS 9513:

   a. Select the desired switch in this case, the Connectrix MDS 9513. The **Device Manager** appears.

   b. Select the ports participating in this fabric and set the port speed to **automax 2 G**. Leave the other settings as default.

   c. Set the **Admin** option to **up**.

3. Unsuspend the VSAN.

   After configuring all the settings as stated above, go back to the **VSAN attributes** and change the **AdminState** to **Active**.

4. Link, using an ISL, the Connectrix MDS 9513 to all the Connectrix M edge switches as shown in Figure 34 on page 280.

5. Verify connectivity of the fabric by validating the same set of *checkpoints* as listed for "Checkpoints" on page 278 of this case study. In this case, even the Cisco Fabric Manager and Cisco CLI can be used to verify the fabric topology and name server information.

IMPORTANT

At the end of this phase, it is highly recommended to check that the zoning information has been distributed appropriately to the Connectrix MDS core. The active zoneset on the Connectrix M and Connectrix MDS fabric management tools must be compared to verify that there are no differences.

On Connectrix Manager, the Connectrix MDS appears as a Connectrix M switch. If you select **Discover > Setup** on Connectrix Manager and add the IP of the Connectrix MDS switch, the Connectrix M switch is discovered as a Connectrix MDS switch, thus avoiding confusion.

Phase 3: Moving half of the host and storage ports from Connectrix M core to the Connectrix MDS 9513

Topology

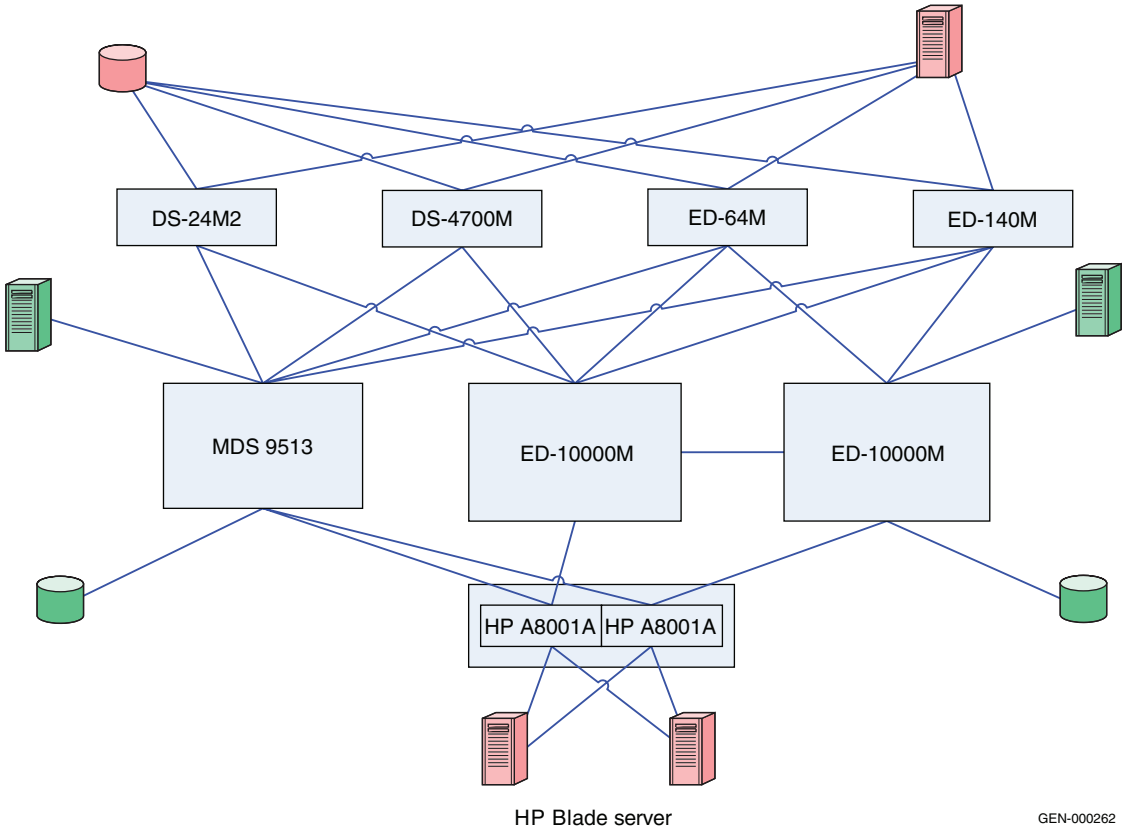


Figure 35 **Phase 3: Moving half the hosts and storage ports**

As shown Figure 35, the same host-target pair is now connected to both core switches: the ED-10000M and the Connectrix MDS 9513.

This intermediate phase moves the host and storage ports to the Connectrix MDS 9513. This ensures that zones are appropriately pushed from Connectrix M switches to Connectrix MDS switches, and that traffic between the host and storage zone is not disrupted. Without any downtime, host and storage ports can be moved to another core switch.

There are no specific steps to be executed at this phase other than physically pulling cables from one switch and plugging them into the other core switch. Follow the "Checkpoints" on page 278, and validate that this transition did not affect the connectivity and functionality of the fabric.

Phase 4: Completely moving the host and storage ports from Connectrix M core to the Connectrix MDS 9513

Topology

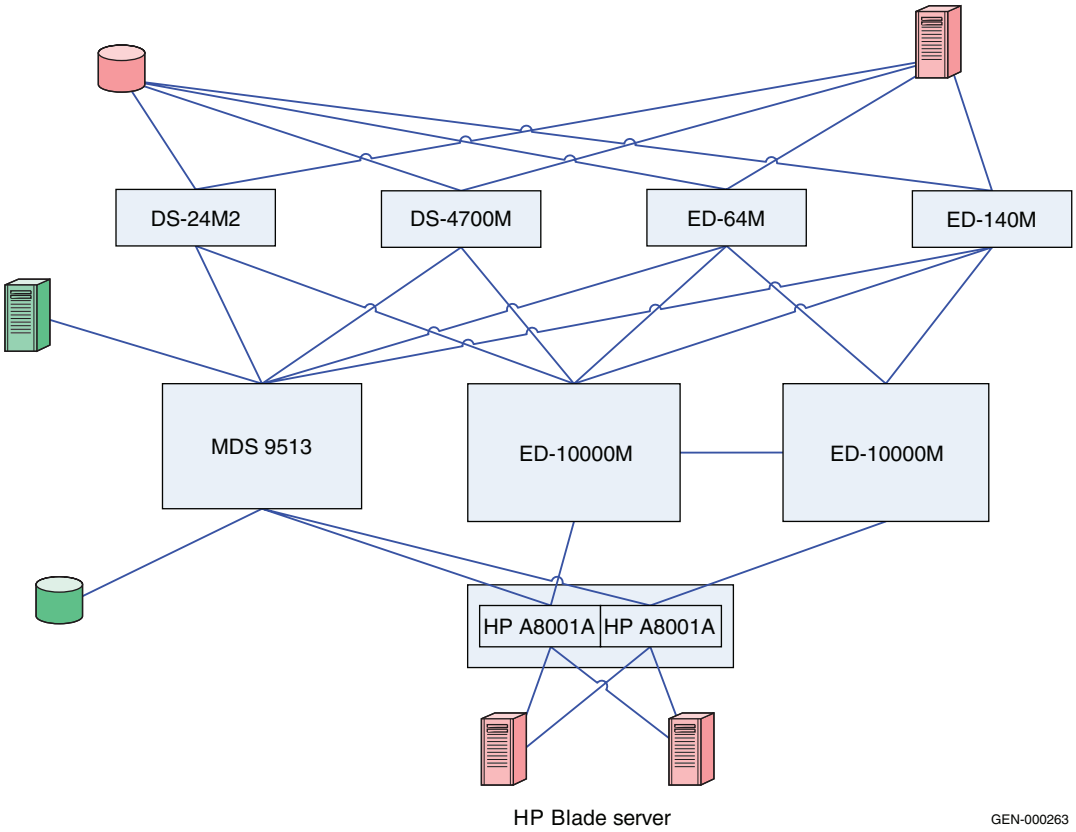


Figure 36 **Phase 4: Completely moving host and storage ports**

This is an extension of "Phase 3: Moving half of the host and storage ports from Connectrix M core to the Connectrix MDS 9513" on page 283. After the completion of Phase 3, validate that a stable fabric exists, and then execute Phase 4 by pulling the remaining

host-storage pairs from the ED-10000M switch, and transferring them to the Connectrix MDS 9513 core.

Follow the "Checkpoints" on page 278, using both Connectrix M and Connectrix MDS fabric management tools.

Phase 5: Adding the Connectrix MDS 9216

Topology

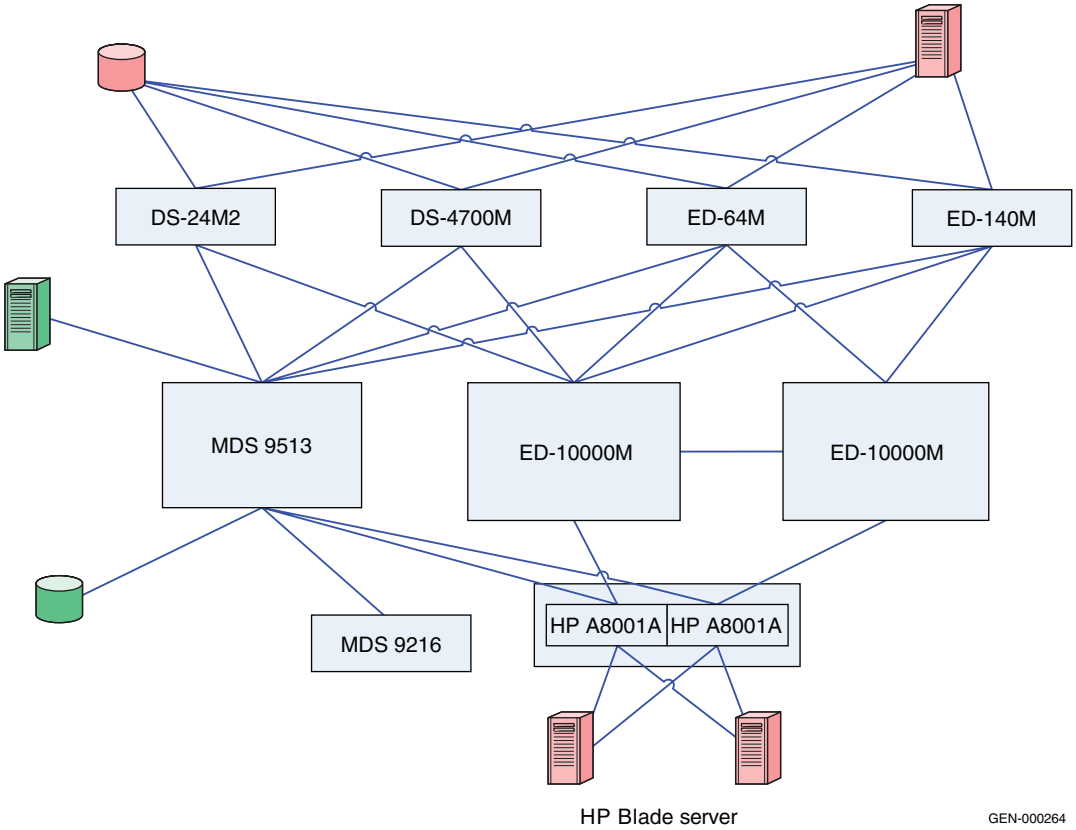


Figure 37 **Phase 5: Adding Connectrix MDS 9216**

To configure the Connectrix MDS 9216:

1. Create a VSAN on the Connectrix MDS switch.

   You can use a Virtual SAN to create multiple logical SANs over the same physical infrastructure.

a. Click **Create VSAN** on Cisco **Fabric Manager**.

b. When the **Create VSAN** window appears, check the switches you want to include in this VSAN. In this case, only the Connectrix MDS 9513 is included.

c. Enter the **VSAN ID** with an unused ID number, and the **VSAN name** with an appropriate name. In this case study, **VSAN ID number = 801** is assigned.

d. For **VSAN** attributes, leave all other field at default, except **InterOperValue** and **AdminState**. You can view VSAN attributes in the **Information** pane.

e. Set **InterOperValue** to **Interop-1** and the **AdminState** to **Suspended**.

f. To statically assign VSAN membership for an interface using Fabric Manager select **Interfaces > FC Physical** from the **Physical Attributes** pane. The interface configuration appears in the **Information** pane. Select the **General** tab, double-click and then complete the **PortVSAN** field by entering the **VSAN ID** number (801) for every port you want the fabric to use.

   For more details on VSAN settings and configuring the Interop-1 specific settings for Connectrix MDS 9000 family switches, refer to the following document: on the Cisco website:
   http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_configuration_guide_chapter09186a00805c7a40.html#wp1040185

2. To set port settings on the Device Manager for the Connectrix MDS 9513:

   a. Select the desired switch, in this case, the Connectrix MDS 9513. The **Device Manager** for this switch appears.

   b. Select the ports participating in this fabric and set the port speed to **automax 2 G**. Leave the other settings as default.

   c. Set the **Admin** option to **up**.

3. Unsuspend the VSAN.

   After configuring all the settings as stated above, go back to the **VSAN attributes** and change the **AdminState** to **Active**.

4. Link, using an ISL, the Connectrix MDS 9513 to all the Connectrix M edge switches as shown in Figure 37 on page 285.

Upon connecting these switches with matching VSAN IDs, the zoning information will merge. After executing the first four steps above, this edge switch must be linked with an ISL to both the Connectrix MDS and Connectrix M core switches.

Phase 6: Moving hosts and storage to new edge

Topology

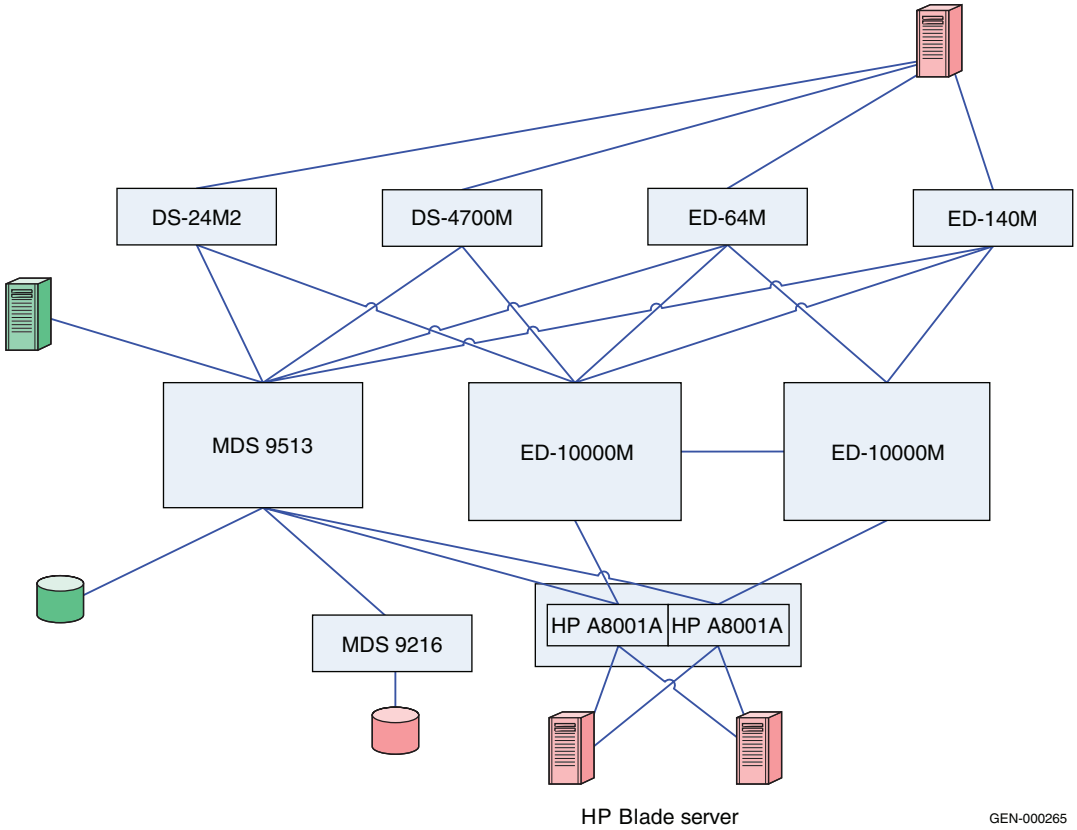


Figure 38 **Phase 6 topology**

Hosts and storage ports shared by the other Connectrix M edge switches in the fabric can be completely, or partially, moved to the Connectrix MDS 9216 edge switch. In this transitional phase both switches, Connectrix M and Connectrix MDS, can co-exist in a stable fabric, with the Connectrix M operating in Open Fabric Mode 1.0 and the Connectrix MDS operating in Interop-1 mode.

The settings on the switches in this phase can be used to dictate the configuration settings to set up a Connectrix M-Connectrix MDS interop fabric. All the non-default settings discussed in "Phase 2: Adding the Connectrix MDS 9513 to the core of the fabric" on page 280 for the Connectrix MDS 9513, Connectrix MDS 9216, and the Connectrix M switches apply to setting up a heterogeneous Connectrix MDS-Connectrix M fabric from the ground up.

Phase 7: Adding a Connectrix MDS switch to the core

Topology

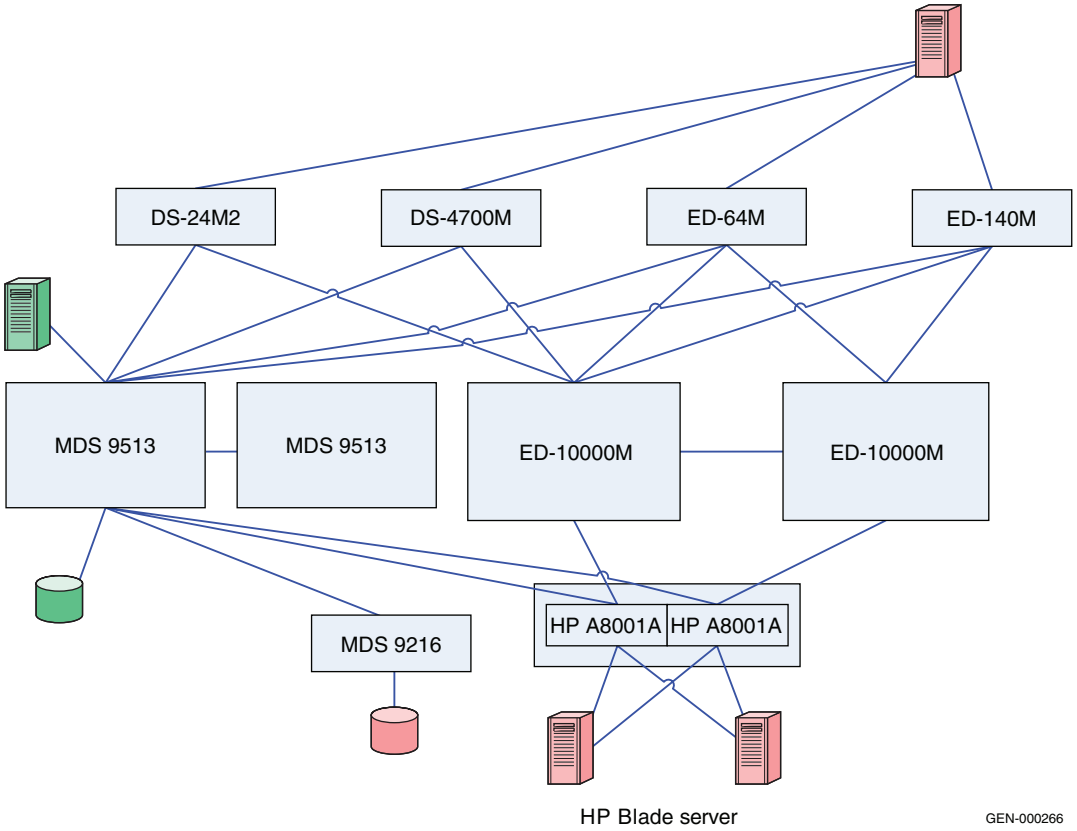


Figure 39 **Phase 7 topology**

As shown in Figure 39, another Connectrix MDS switch, the Connectrix MDS 9513, can be added to the core with similar settings as the previous Connectrix MDS 9513, and can be connected through an ISL to the existing core Connectrix MDS 9513 director and the

other edge switches in the fabric. It is important to note that the VSAN must be configured with the same VSAN ID and attributes as the existing VSAN to ensure a clean fabric merge.

Complete migration to Connectrix MDS

At the end of case study #2, a migration from a Connectrix M-only fabric to a Connectrix M-Connectrix MDS fabric with Connectrix MDS at the core, connected through ISLs to every edge switch in the fabric, occurred. The host and storage ports (except for the blade server host ports) can be completely moved over to the Connectrix MDS switches as specified in Phase 3, 4 and 6. The Connectrix M edge switches (except for the blade server Brocade M-Series switch modules) and the Connectrix M core switches can then be pulled out from the fabric. A fully operational Connectrix MDS-only fabric exists when completed. This is a complete migration from one switch vendor type (Brocade M-Series) to another vendor type (Cisco).

IMPORTANT

In a Connectrix MDS-only fabric, it is recommended to have all the Connectrix MDS switches operating in their native (default) mode. The Connectrix MDS switches at the end of the migration are in their interopmode.

In order to change the interopmode on the Connectrix MDS switches, it is not necessary to reboot the switches. The **Interop mode** attribute on the currently active VSANs needs to be changed from **Interop-1** to **Default**.

The active zones that were pushed on to the Connectrix MDS switches from the Connectrix M switches, cannot be modified once this is done and the zoning needs to be reconfigured. It is highly recommended to take a backup of the configuration to avoid losing all the zoning information that was present in the active zoneset on the Connectrix MDS switches.

Warnings or caveats

Please refer to EMC Knowledgebase solution emc149735 for all interop issues.

Case study #3 Migrating from a Connectrix B homogeneous fabric in Connectrix B native mode to a Connectrix MDS fabric

Assumptions specific to this case study:

- Interoperability mode settings on the switches are:
  - ***Brocade native mode (interopmode 0)*** mode on the Connectrix B switches in the fabric.
  - ***Cisco Interop-3*** mode on the Connectrix MDS switches in the fabric.
- Fabric management applications used for managing the fabric:
  - Cisco Fabric Manager, Brocade Web Tools.

Note: The Brocade Fabric Manager is not supported for management of a Connectrix B-Connectrix MDS heterogeneous fabric.

Phase 1: Base configuration – Pre-existing Connectrix B core-edge fabric

Topology

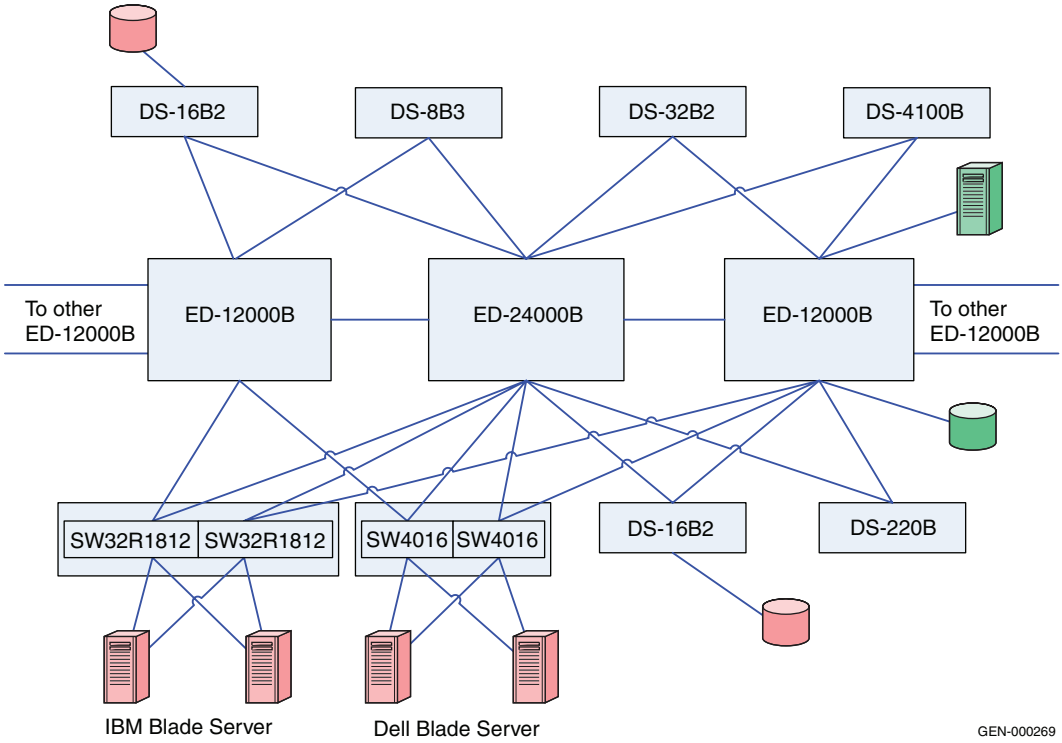


Figure 40 **Phase 1: Basic configuration**

As shown in Figure 40, the two ED-12000Bs and the ED-24000B director switches are at the core of the fabric while the departmental switches, the DS-4100B, DS-220B, DS-16B3, DS-8B2, and Connectrix B-based blade server switch modules from IBM (IBM 32R1813) and Dell (SW4016), are at the edge. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* to obtain a list of all the other Brocade/EMC Connectrix B switches that can be supported in a heterogeneous set up as well as the operating modes specified above for this case study.

The specific configuration settings, best practices, host and storage layouts, and topology design to withstand failures for a Connectrix B

core-edge homogeneous fabric discussed in the "Connectrix B example" on page 81 applies to this base topology.

Checkpoints

Before adding the Cisco core Connectrix MDS 9513 director to the homogeneous Connectrix B fabric, verify the following fabric characteristics using either the Connectrix B switch management application, Web Tools, or the Brocade CLI.

- **Proper distribution of the Name Server information.**

  Verify that the all the host HBA ports and storage ports logged into the fabric are listed in the name server.

- **Proper distribution of Zoning information**.

  Verify that the active zoneset comprises of the zones that contain the desired mapping of host and storage ports.

- **Proper display of fabric and N_Port elements on the management applications**.

  Verify that the physical topology and domain count of the fabric is as desired.

- **No disruption in data transfer**.

  Verify that the data traffic is running appropriately through the SAN.

Phase 2: Adding the Connectrix MDS 9513 to the core of the fabric

Topology

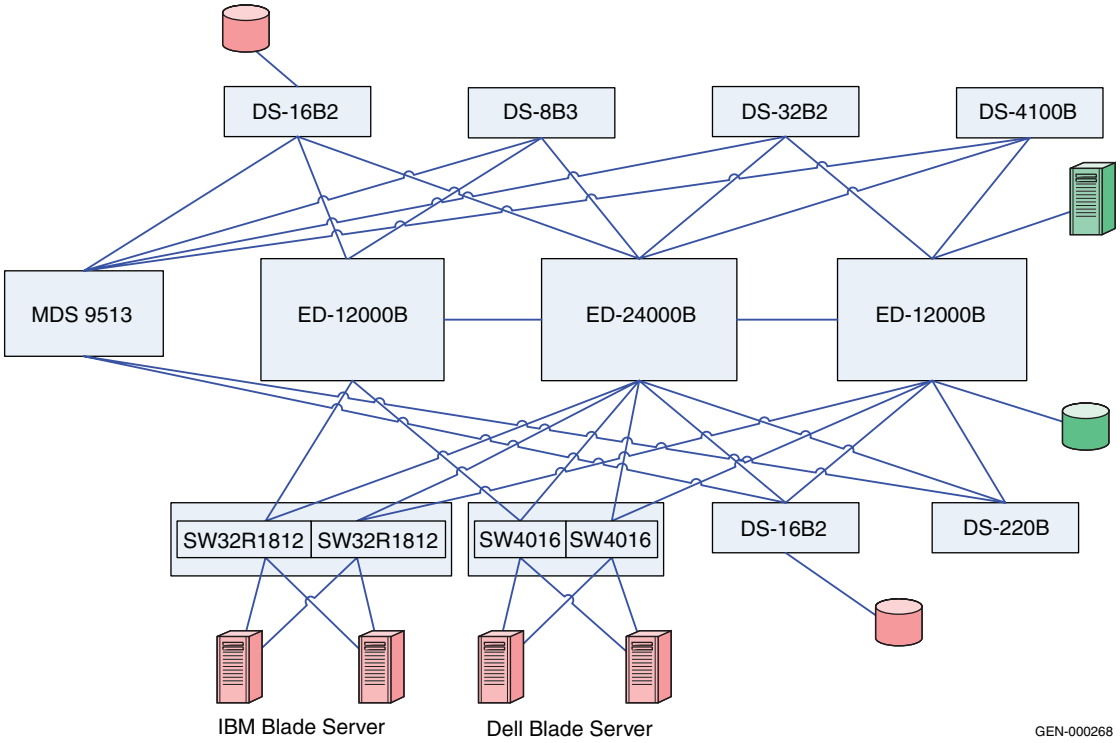


Figure 41 **Phase 2: Adding Connectrix MDS 9513**

As illustrated in Figure 41, the Connectrix MDS 9513 director is added to the core of the fabric. Also it is important to note that there is no ISL between the Connectrix B cores and the Connectrix MDS 9513. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* to obtain a list of all the Connectrix MDS series switches that can be supported in a heterogeneous fabric with the Connectrix B switches, as well as the operating modes specified above for this case study.

Before adding the MDS 9513 to the fabric, the following steps need to be executed on the Connectrix MDS switch:

1. Create a VSAN on the Connectrix MDS switch.

   A Virtual SAN can be used to create multiple logical SANs over the same physical infrastructure.

   a. Click **Create VSAN** on the Cisco Fabric Manager.

   b. When the **Create VSAN** window appears, check the switches that must be included in this VSAN. In this case, only the Connectrix MDS 9513 would be included.

   c. Fill in the **VSAN ID** field with an unused ID number and the VSAN name filed with an appropriate name. In this case study, the **VSAN ID number = 150** is assigned.

   d. For the VSN attributes leave everything other than the **InterOperValue** and the **AdminState** at default.

   e. Set **InterOperValue** to **Interop-3** and the **AdminState** to **Suspended**.

   f. Assign interfaces to the legacy switch interop mode 2 VSAN, i.e., statically assign VSAN membership for an interface using Fabric Manager, by choosing **Interfaces > FC Physical** from the **Physical Attributes** pane. The interface configuration in the **Information** pane appears. Select the **General** tab on this window and double-click and complete the **PortVSAN** field by entering the VSAN ID number (**150**) for every port desired to be used by this fabric.

      For more details on the VSAN settings for Connectrix MDS 9000 Family switches refer to the Cisco document:

      http://www.cisco.com/en/US/products/ps5989/products_configuration_guide_chapter09186a0080662ce3.html#wp1094055

      For configuring Interop-3 specific settings on the Cisco CLI refer to the following Cisco document:

      http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_configuration_guide_chapter09186a00805c77fe.html

      The legacy switch interoperability mode 3 for Connectrix B switches with more than 16 ports (and a core PID =1) was introduced with Connectrix MDS SAN-OS Release 1.3. With

this VSAN-based interop mode, Connectrix B switches do not have to be altered from their native mode and can be seamlessly added to a new or existing Connectrix MDS SAN-OS VSAN.

2. Lower the Connectrix MDS 9000 switch ISL buffer-to-buffer credits (BB_Credits) to match the Brocade BB_Credits, because the Connectrix B switch cannot handle more than 16 BB_Credits on an ISL.

3. Configure the Connectrix B switches.

   a. Now that the Connectrix MDS 9513 director is configured, there will be no configuration or disruption to the Connectrix B fabric. All that is required is to enable the new ISL ports. A **portcfgislmode** command must be run against all the ports on the Connectrix B edge switches that will be linked with an ISL to the Connectrix MDS 9513 core switch.

   b. Brocade's **msplmgmtdeactivate** command must also explicitly be run prior to connecting from a Connectrix B switch to an Connectrix MDS 9000 family switch.

4. Before connecting the Connectrix MDS switch to the Connectrix B switches through an ISL, it is important to verify that all the Connectrix B switches and the Connectrix MDS switch are running the supported firmware versions for the respective interop modes. This can be checked by referring to the *EMC Support Matrix* entries for these switches

5. Create an ISL between the Connectrix MDS 9513 and all the Connectrix B edge switches as per Figure 41 on page 293.

6. Verify connectivity of the fabric by validating the same set of "Checkpoints" on page 292. In this case, the Cisco Fabric Manager and Cisco CLI can be used to verify the fabric topology and name server information, in addition to the Brocade CLI and Brocade Web Tools.

IMPORTANT

At the end of this phase, it is highly recommended to check that the zoning information has been distributed appropriately to the Connectrix MDS core. The active zoneset on the Connectrix B and Connectrix MDS fabric management tools must be compared to verify that there are no differences.

Phase 3: Moving half of the host and storage ports from Connectrix B core to the Connectrix MDS 9513

Topology

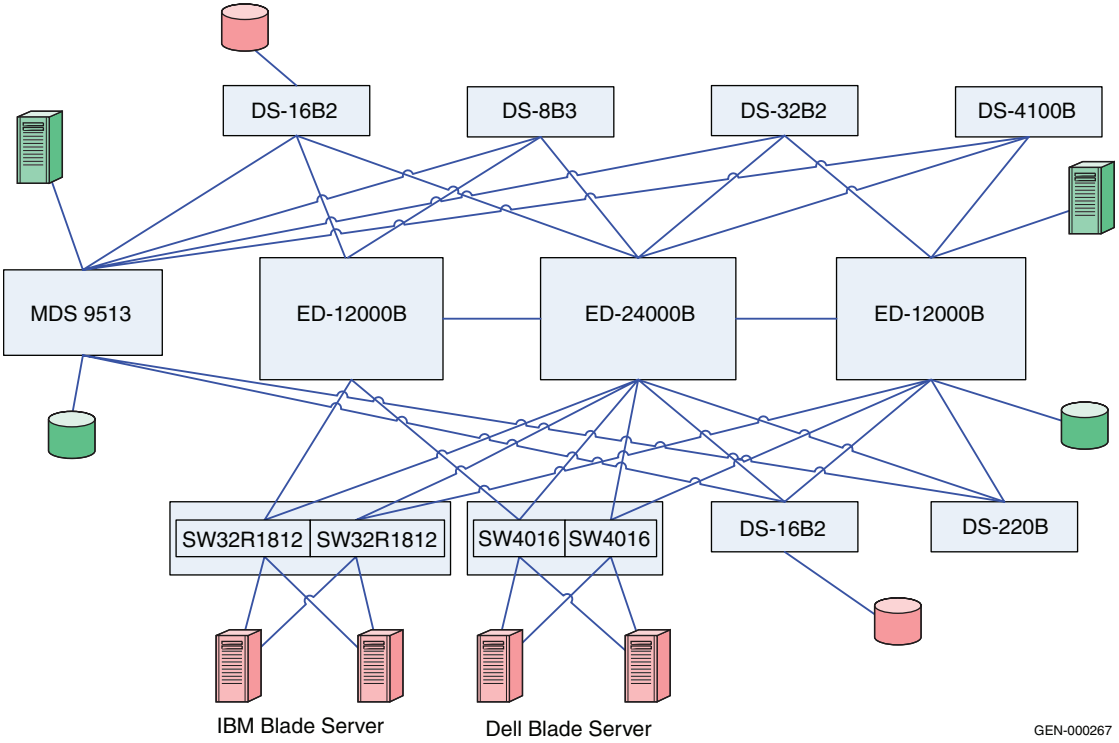


Figure 42 **Phase 3: Moving half the host and storage ports**

As illustrated in Figure 42, the same host-target pair is now connected to both core switches: the ED-12000B and the Connectrix MDS 9513.

This phase is an intermediate phase to completely moving the host and storage ports to the Connectrix MDS9513. This phase ensures that the zones are appropriately pushed from the Connectrix B to the Connectrix MDS switches and that the traffic between the host and storage zoned together is not disrupted. The host and storage ports can be moved successfully to another core switch without any downtime.

There are no specific steps to be executed at this phase other than physically pulling cables from one switch and plugging them into the

other core switch. However, it is recommended to follow the "Checkpoints" on page 292 and validate that this transition did not affect the connectivity and functioning of the fabric.

Phase 4: Completely moving the host and storage ports from Connectrix B core to the Connectrix MDS 9513

Topology

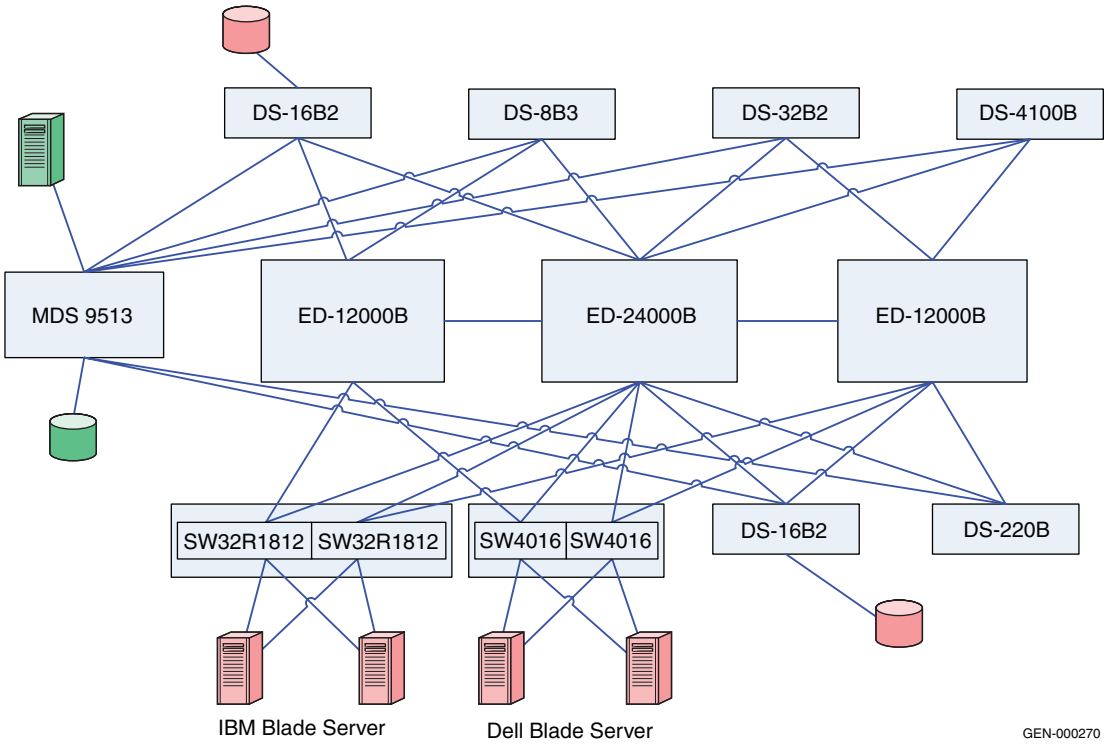


Figure 43 **Phase 4: Completely moving host and storage ports**

This is an extension of "Phase 3: Moving half of the host and storage ports from Connectrix B core to the Connectrix MDS 9513". On validating that a stable fabric exists at the completion of Phase 3, Phase 4 can be executed by pulling the remaining host-storage pairs from the ED-12000B switch and transferring them to the Connectrix MDS 9513 core.

Again, it is recommended to go follow the "Checkpoints" on page 292 using both Connectrix B and Connectrix MDS supported fabric management tools.

Phase 5: Adding an Connectrix MDS 9216 to the edge

Topology

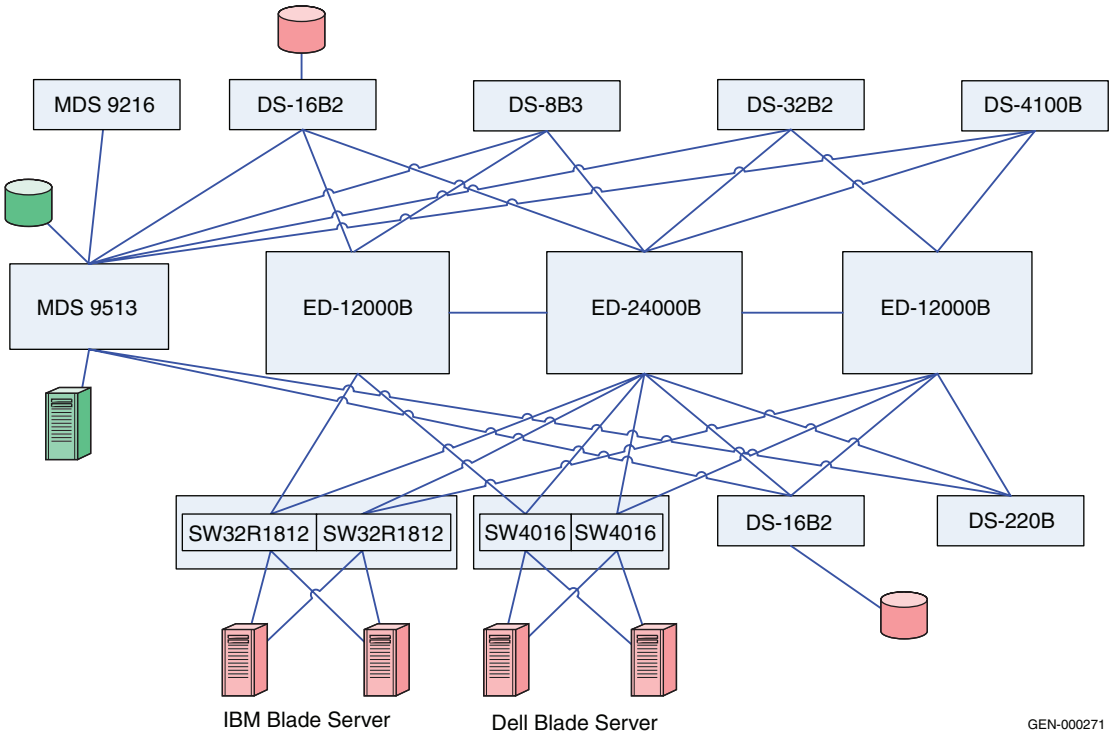


Figure 44 **Phase 5: Adding Connectrix MDS 9216**

To configure the Connectrix MDS 9216:

1. Create a VSAN on the Connectrix MDS switch.

   A Virtual SAN can be used to create multiple logical SANs over the same physical infrastructure.

   a. Select **Create VSAN** on the Cisco **Fabric Manager**.

   b. Once the **Create VSAN** window appears, check the switches that must be included in this VSAN. In this case, only the Connectrix MDS 9513 will be included.

   c. Fill in the **VSAN ID** field with an unused ID number and the **VSAN name** filed with an appropriate name. In this case study, **VSAN ID number = 801** is assigned.

d. For **VSAN** attributes, leave everything other than the **InterOperValue** and the **AdminState** at default. VSAN attributes can be seen in the information pane.

e. Set **InterOperValue** to **Interop-1** and the **AdminState** to **Suspended**.

f. Statically assign VSAN membership for an interface using Fabric Manager, by choosing **Interfaces > FC Physical** from the **Physical Attributes** pane. The interface configuration appears in the **Information** pane. Select the **General** tab on this window, and double-click and complete the **PortVSAN** field by entering the **VSAN ID** number (801) for every port desired to be used by this fabric.

   For more details on the VSAN settings and for configuring the Interop-1 specific settings for Connectrix MDS 9000 family switches refer to the following Cisco document: http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_configuration_guide_chapter09186a00805c7a40.html#wp1040185.

2. To set port settings on the Device Manager for the Connectrix MDS 9513:

   a. Select the desired switch, in this case, the Connectrix MDS 9513. The **Device Manager** for this switch appears.

   b. Select the ports participating in this fabric, and then set the port speed to **automax 2 G**. Leave the other settings as default.

   c. Set the **Admin** option to **up**.

3. Unsuspend the VSAN.

   After configuring all the settings as stated above, go back to the **VSAN attributes,** and change the **AdminState** to **Active**.

4. Link, using an ISL, the Connectrix MDS 9513 to all the Connectrix M edge switches as shown in Figure 37 on page 285.

When connecting these switches with matching VSAN IDs, the zoning information merges. After executing the steps above, you must link this edge switch with an ISL to both the Connectrix MDS and Connectrix M core switches.

Phase 6: Moving hosts and storage to new edge

Topology

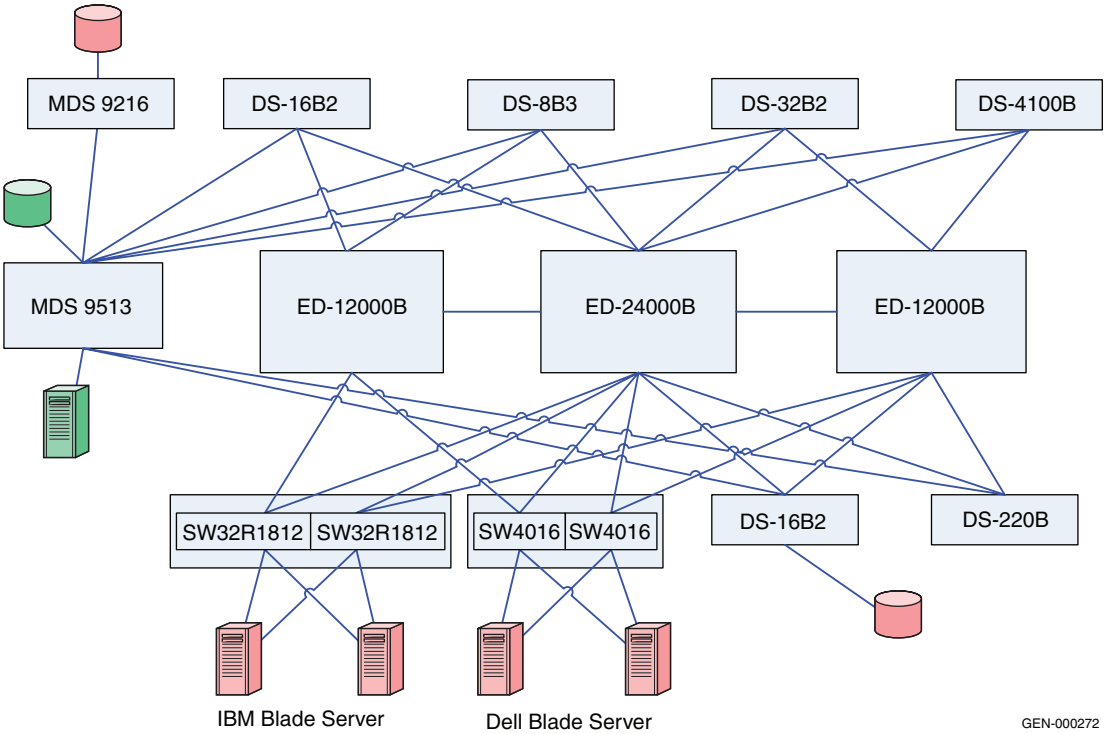


Figure 45 **Phase 6: Moving hosts and storage to new edge**

Hosts and storage ports shared by the other Connectrix B edge switches in the fabric (except for the Server Blade N_Ports logged into the Brocade blade server switch modules) can be completely or partially moved to the Connectrix MDS 9216 edge switch. It is evident in this transitional phase that both Connectrix B and Connectrix MDS switches, can co-exist in a stable fabric with Connectrix B operating in Connectrix B native mode and Connectrix MDS operating in Interop-3 mode.

The settings on the switches in this phase can be used to implement configuration settings for a Connectrix B–Connectrix MDS interop fabric. All the non-default settings discussed in "Phase 2: Adding the Connectrix MDS 9513 to the core of the fabric" on page 293 for Connectrix MDS 9513, Connectrix MDS 9216 and Connectrix B

switches apply to setting up a heterogeneous Connectrix MDS-Connectrix MDS fabric from the ground up.

Phase 7: Adding a Connectrix M switch to the core

Topology

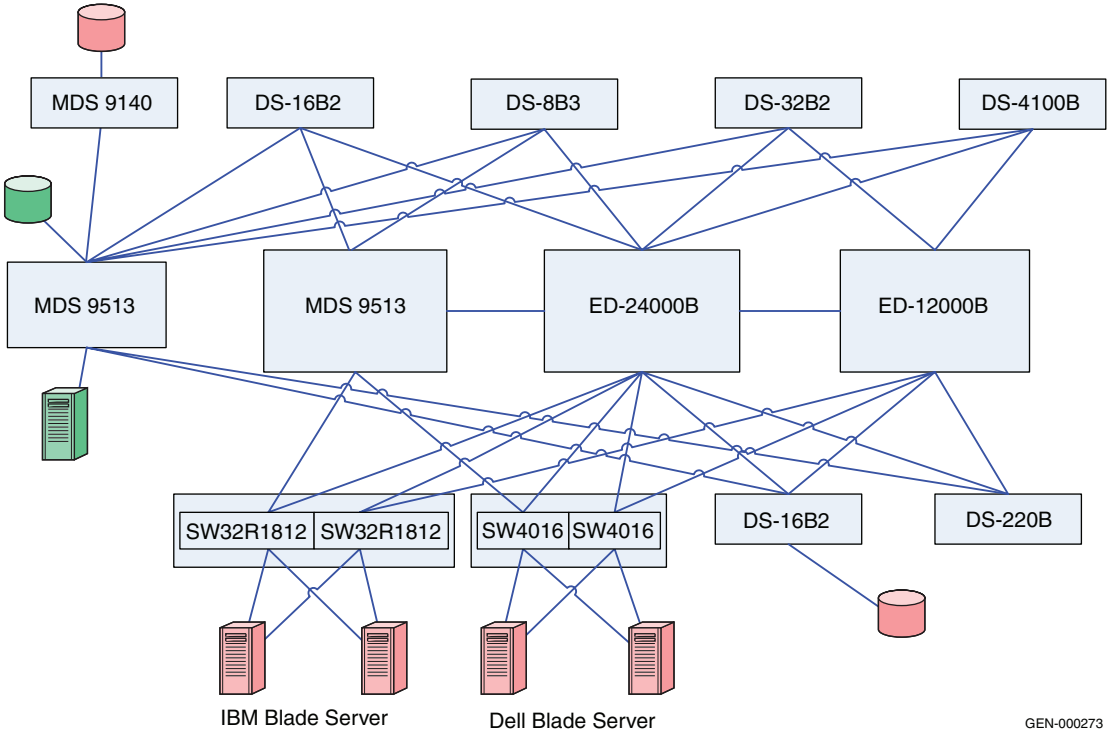


Figure 46 **Phase 7: Adding Connectrix M switch to core example**

As illustrated in Figure 46, you can add a Connectrix MDS 9513 to the core with similar settings as the previous Connectrix MDS 9513, and you can also create an ISL to the existing core Connectrix MDS 9513 director and other edge switches in the fabric. It is important that you configure the new VSAN with the same VSAN ID and attributes as the existing VSAN to ensure a clean fabric merge.

Warnings or caveats

Refer to EMC Knowledgebase solution emc149735 for all interop issues.

Complete migration to Connectrix MDS

At the end of case study #3 you have completed a migration from a Connectrix B-only fabric to a Connectrix B-Connectrix MDS fabric with Connectrix MDS switches at the core, connected through ISLs to every edge switch in the fabric. The host and storage ports (except for the blade server host ports) can be completely moved to the Connectrix MDS switches as specified in Phase 3, 4 and 6. The Connectrix B edge switches (except for the blade server Brocade switch modules) and the Connectrix B core switches can now be pulled out of the fabric. A fully operational Connectrix MDS-only fabric now exists. This is a complete migration from one switch type (Connectrix B) to another type (Connectrix MDS).

IMPORTANT

In a Connectrix MDS-only fabric, it is recommended that all Connectrix MDS switches operate in their native (default) mode. Connectrix MDS switches are in interopmode at the end of the migration.

You need not reboot the switch to change the interopmode on a Connectrix MDS. The **Interop mode** attribute on currently active VSANs must be changed from **Interop-3** to **Default**.

You cannot modify active zones that were pushed to Connectrix MDS switches from Connectrix B switches; zoning must be reconfigured. It is highly recommended that you backup the configuration to avoid losing all zoning information in the active zoneset on the Connectrix MDS switches.

Warnings or caveats

Refer to EMC Knowledgebase solution emc149735 for all interop issues.

Case Study #4

Setting up a heterogeneous switched fabric with Connectrix B switches in the interop mode (interopmode 1), and the Connectrix MDS switches in Interop-1 mode

Assumptions specific to this case study:

- Interoperability mode settings on the switches:
  - Interop mode: ***interopmode 1*** on the Connectrix B switches in the fabric.
  - Cisco ***Interop-1*** mode on the Connectrix MDS switches in the fabric.

- Fabric management applications used for managing the fabric:
  - Cisco Fabric Manager
  - Web Tools only for Connectrix B switches

When migrating from a Connectrix B to a Connectrix MDS fabric, keep the fabric in Connectrix B native mode and follow the migration process explained in "Case study #3" on page 290. This case study shows the settings that you must configure on Connectrix B and Connectrix MDS switches before they are linked with an ISL. This is a required step when migrating from a Connectrix B to a Connectrix MDS fabric.

Phase 1: Base configuration – Pre-existing Connectrix B core-edge fabric

Topology

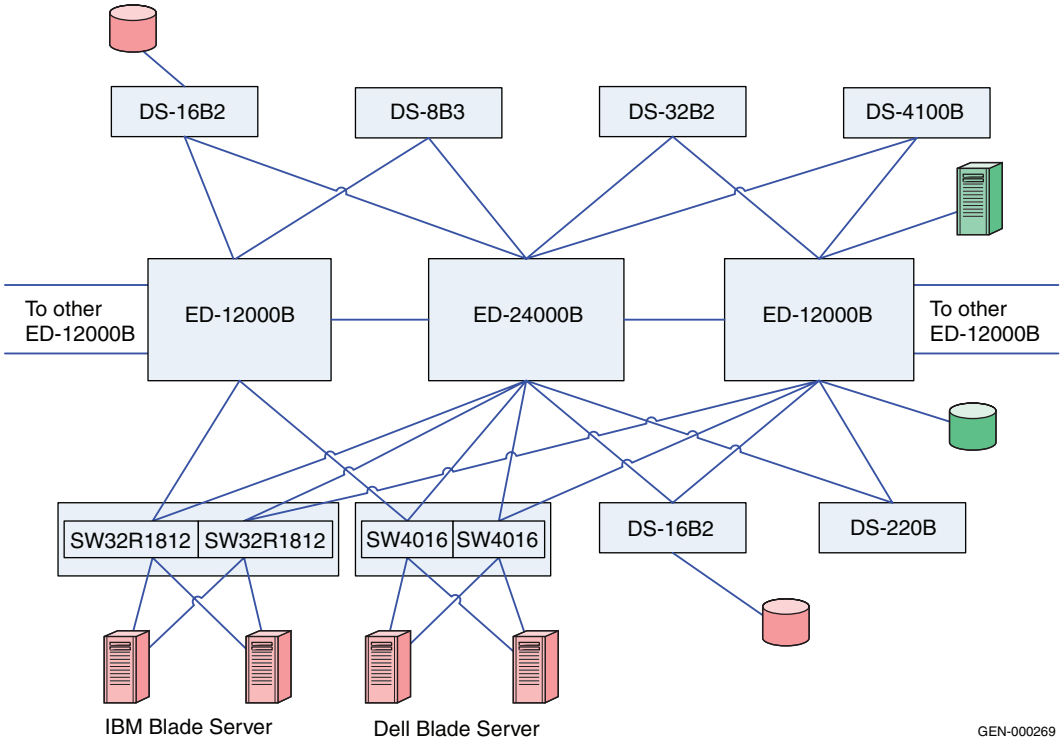


Figure 47 **Phase 1: Basic configuration**

As shown in Figure 47 on page 303, the two ED-12000B, and ED-24000B director switches are at the core of the fabric. Edge hardware includes DS-4100B, DS-220B, DS-16B3, DS-8B2, Brocade-based blade server switch modules from IBM (IBM 32R1813), and Dell (SW4016). Refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* for a list of other Brocade/EMC Connectrix switches supported in a heterogeneous setup, and for relevant operating modes.

The following information discussed in the "Connectrix B example" on page 81 applies to this base topology:

- Specific best practices
- Host and storage layouts
- Design to withstand failures for a Connectrix B core-edge homogeneous fabric

This case study, unlike "Case study #3" on page 290, sets Connectrix B switches to interop mode, links them with an ISL to Connectrix MDS switches (operating in their supported Interop-1 mode), and then pulls out the Connectrix B cores from the fabric.

Since this is a Connectrix B-only fabric with all switches operating in Connectrix B native mode by default, the following non-default settings must be configured on the Connectrix B switches using the Brocade CLI:

1. Telnet into all the Connectrix B switches in the fabric; one at a time.
2. Run the **interopmode** command at the switch command prompt.
3. If interopmode is **On**, the switch can become part of a fabric that can accept the addition of a Connectrix MDS. If it is **Off**, perform Step 4 through Step 8.
4. Disable the switch by running the **switchdisable** command at the switch prompt.
5. Issue the **interopmode 1** command.
6. Enter a confirmation yes: **y** to switch the interopmode to **on**. Enter **y** in the warning box.
7. Reboot the switch to activate configuration settings.
8. After reboot, repeat Step 1 through Step 3 to verify that interopmode has been set to **On**.

9. Run the **version** command to verify that the switch is running the supported firmware version for interop with Connectrix MDS.
10. The **msplmgmtdeactivate** command must explicitly be run prior to connecting from a Connectrix B switch to a Connectrix MDS 9000 Family switch.

Checkpoints

Before adding the Connectrix MDS core Connectrix MDS 9513 director to the Connectrix B fabric, verify the following fabric characteristics using either the Brocade switch management application, Web Tools, or the Brocade CLI.

- **Proper distribution of the Name Server information**.

  Verify that the all host HBA ports and storage ports that are logged into the fabric are listed in the name server.

- **Proper distribution of Zoning information**.

  Verify that the active zoneset comprises zones that contain the desired mapping of host and storage ports.

- **Proper display of fabric and N_Port elements on the management applications**.

  Verify that the physical topology and domain count of the fabric are correct.

- **No disruption in data transfer**.

  Verify that the data traffic is running appropriately through the SAN.

Phase 2: Adding the Connectrix MDS 9513 to the core of the fabric

Topology

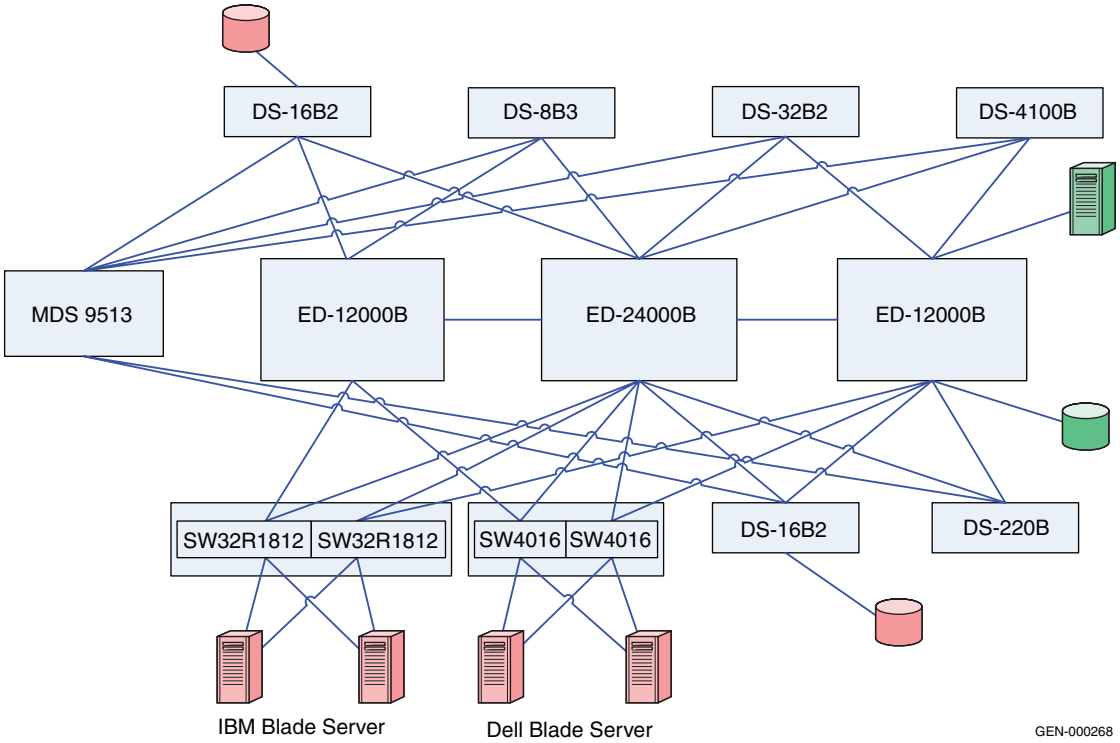


Figure 48 **Phase 2: Adding Connectrix MDS 9513**

As shown in Figure 48, the Connectrix MDS 9513 director is added to the core of the fabric. It is important to note that there is no ISL between the Connectrix B cores and the Connectrix MDS 9513. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* for a list of Cisco / Connectrix MDS switches that are supported in a heterogeneous configuration with Connectrix B switches. In addition, operating modes are listed.

Before adding a Connectrix MDS 9513 to the fabric, follow these steps for a Connectrix MDS switch:

1. Create a VSAN on the Connectrix MDS.

   A Virtual SAN can be used to create multiple logical SANs over the same physical infrastructure.

   a. Click **Create VSAN** on the **Cisco Fabric Manager**.

   b. When the **Create VSAN** window appears, check the switches you want to include in the VSAN. In this case, only Connectrix MDS 9513 switches are included.

   c. Fill in the **VSAN ID** field with an unused ID number, and the **VSAN name** field with an appropriate name. In this case study, the **VSAN ID number = 801** is assigned.

   d. For VSAN attributes, leave everything other than the **InterOperValue** and the **AdminState** at default. VSAN attributes can be seen in the **information** pane.

   e. Set **InterOperValue** to **Interop-1** and the **AdminState** to **Suspended**.

   f. Statically assign VSAN membership for an interface using Fabric Manager, by choosing **Interfaces > FC Physical** from the **Physical Attributes** pane.
   The interface configuration appears in the **Information** pane. Select the **General** tab.

   g. Double-click and complete the **PortVSAN** field by entering the **VSAN ID number** (**801**) for every port that will be used by this fabric.

For more details on VSAN settings and for configuring Interop-1 specific settings for Connectrix MDS 9000 Family switches, refer to the following document on the Cisco website:

```
http://www.cisco.com/en/US/products/hw/ps4159/ps4358/pro
   ducts_configuration_guide_chapter09186a00805c77de.html
```

2. Set port settings on the Device Manager for the Connectrix MDS 9513:

   a. Select the desired switch, in this case the Connectrix MDS 9513.
   The **Device Manager** for this switch displays.

b. Select the ports participating in this fabric and set the port speed to **automax 2 G**. Leave the other settings as **default**.

c. Set the **Admin** option to **up**.

3. Unsuspend the VSAN.

   After configuring all the settings as stated above, go back to the **VSAN attributes** and change the **AdminState** to **Active**.

4. Using an ISL, link the Connectrix MDS 9513 to all of the Connectrix B edge switches that appear in the topology diagram.

5. Verify connectivity of the fabric by validating the same set of "Checkpoints" on page 305. In this case, the Cisco Fabric Manager and Cisco CLI can be used to verify the fabric topology and name server information.

IMPORTANT

At the end of this procedure, make sure the zoning information was distributed appropriately to the Connectrix MDS core. Compare the active zoneset on the Connectrix B and Connectrix MDS fabric management tools to verify that no differences exist.

Phase 3: Moving half of the host and storage ports from Connectrix B core to the Connectrix MDS 9513

Topology

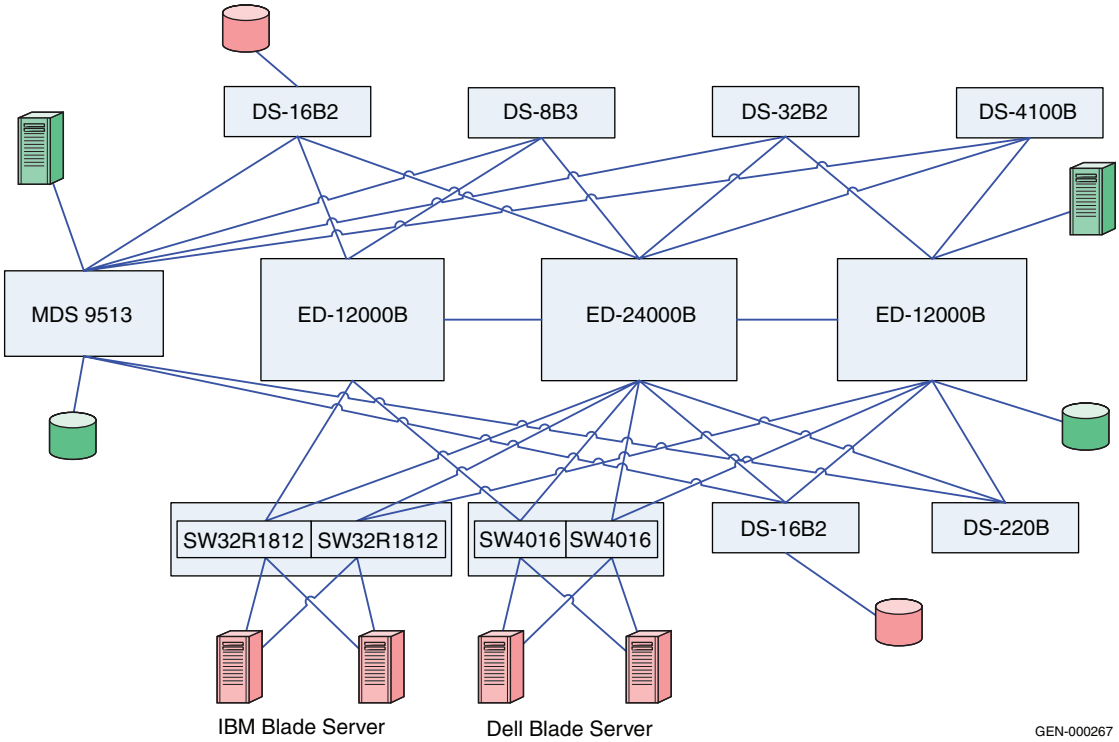


Figure 49 **Phase 3: Moving half the host and storage ports**

As shown in Figure 49, the same host-target pair is now connected to both core switches: the ED-24000B and the Connectrix MDS 9513. This is an intermediate phase when moving the host and storage ports to the Connectrix MDS 9513. It ensures that the zones are appropriately pushed from the Connectrix B to the Connectrix MDS switches and that the traffic between the host and storage zone is not disrupted. The host and storage ports can be successfully moved to another core switch without any downtime.

Next, you should physically pull cables from one switch and plug them into the other core switch.Additionally, you should review the "Checkpoints" on page 305, to validate that this transition did not affect the connectivity and functioning of the fabric.

Phase 4: Completely moving the host and storage ports from Connectrix B core to the Connectrix MDS 9513

Topology

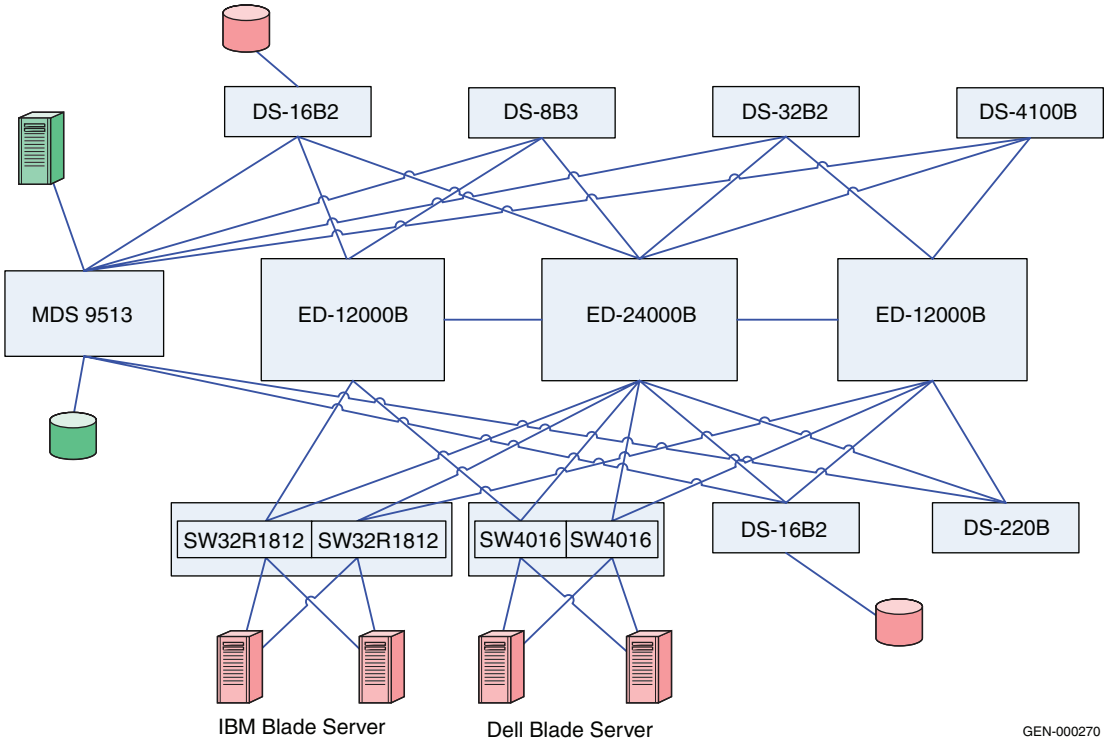


Figure 50 **Phase 4: Completely moving host and storage ports**

This phase is an extension of "Phase 3: Moving half of the host and storage ports from Connectrix B core to the Connectrix MDS 9513" on page 309. After validating that a stable fabric exists, pull the remaining host-storage pairs from the ED-24000B switch and transferring them to the Connectrix MDS 9513 core.

Review the "Checkpoints" on page 305 using both Connectrix M and Connectrix MDS fabric management tools.

Phase 5: Adding a Connectrix MDS 9216 to the edge

Topology

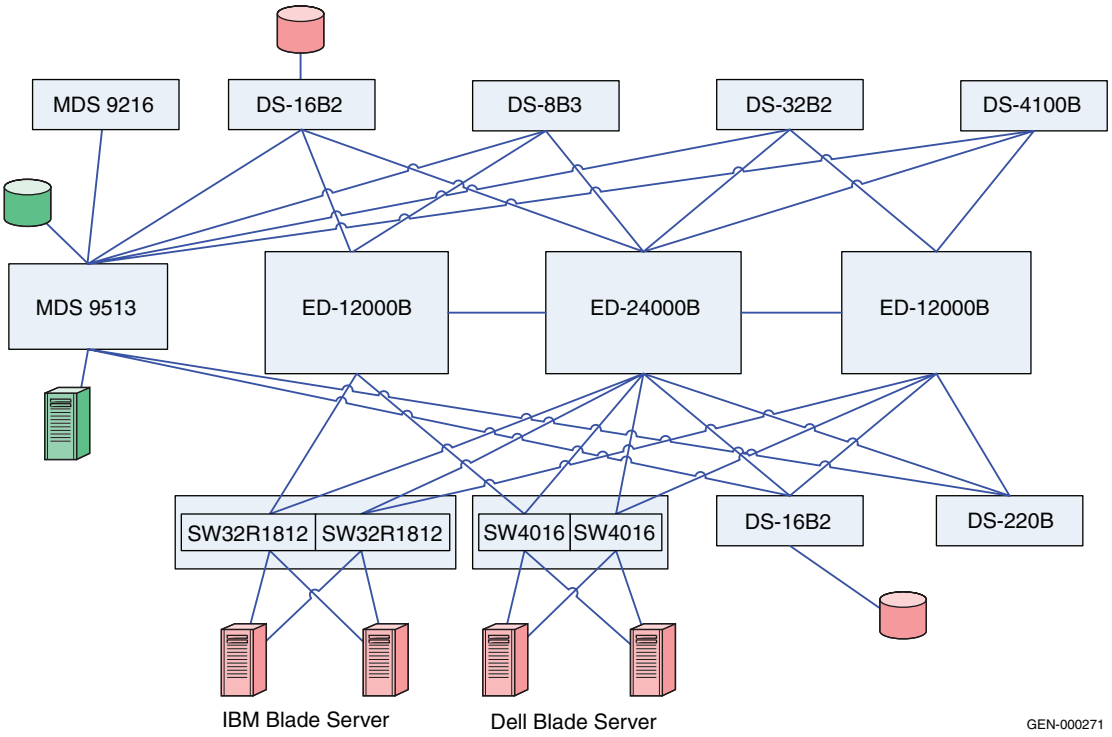


Figure 51 **Phase 5: Adding Connectrix MDS 9216**

Perform the following procedure. After connecting these switches with matching VSAN IDs, the zoning information merges. After executing these steps, this edge switch must be linked using an ISL to both the Connectrix MDS and Connectrix B core switches.

1. Create a VSAN on the Connectrix MDS switch.

   A Virtual SAN can be used to create multiple logical SANs over the same physical infrastructure.

   a. Click **Create VSAN** on the **Cisco Fabric Manager**.

   b. When the **Create VSAN** window appears, select the switches that you want to include in the VSAN. In this case study, only the Connectrix MDS 9513 is included.

c. Fill in the **VSAN ID** field with an unused ID number and the **VSAN name** field with an appropriate name. In this case study, the **VSAN ID number = 801** is assigned.

d. For VSAN attributes, leave everything other than the **InterOperValue** and the **AdminState** at default. VSAN attributes can be seen in the **Information** pane.

e. Set **InterOperValue** to **Interop-1** and the **AdminState** to **Suspended**.

f. Statically assign VSAN membership for an interface using Fabric Manager, by selecting **Interfaces > FC Physical** from the **Physical Attributes** pane.
The interface configuration appears in the **Information** pane.

g. Select the **General** tab on this window, double-click and complete the **PortVSAN** field by entering the **VSAN ID number** (**801**) for every port desired for this fabric.

For more details on VSAN settings and for information on configuring the Interop-1 specific settings for Connectrix MDS 9000 Family switches refer to the following Cisco document: http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_configuration_guide_chapter09186a00805c77de.html

2. Set port settings on the Device Manager for the Connectrix MDS 9513:

a. Select the desired switch, in this case the Connectrix MDS 9513. The **Device Manager** for this switch appears.

b. Select the ports participating in this fabric, and then set the port speed to **automax 2 G**. Leave the other settings as **default**.

c. Set the **Admin** option to **up**.

3. Unsuspend the VSAN.

After configuring all the settings as stated above, go back to the **VSAN attributes** and change the **AdminState** to **Active**.

4. Link, using an ISL, the Connectrix MDS 9513 to all the Connectrix B edge switches referring to the topology diagram.

After connecting these switches with matching VSAN IDs, the zoning information merges. After executing the steps above, this edge switch must be linked with an ISL to both the Connectrix MDS and Connectrix M core switches.

Phase 6: Moving hosts and storage to new edge

Topology

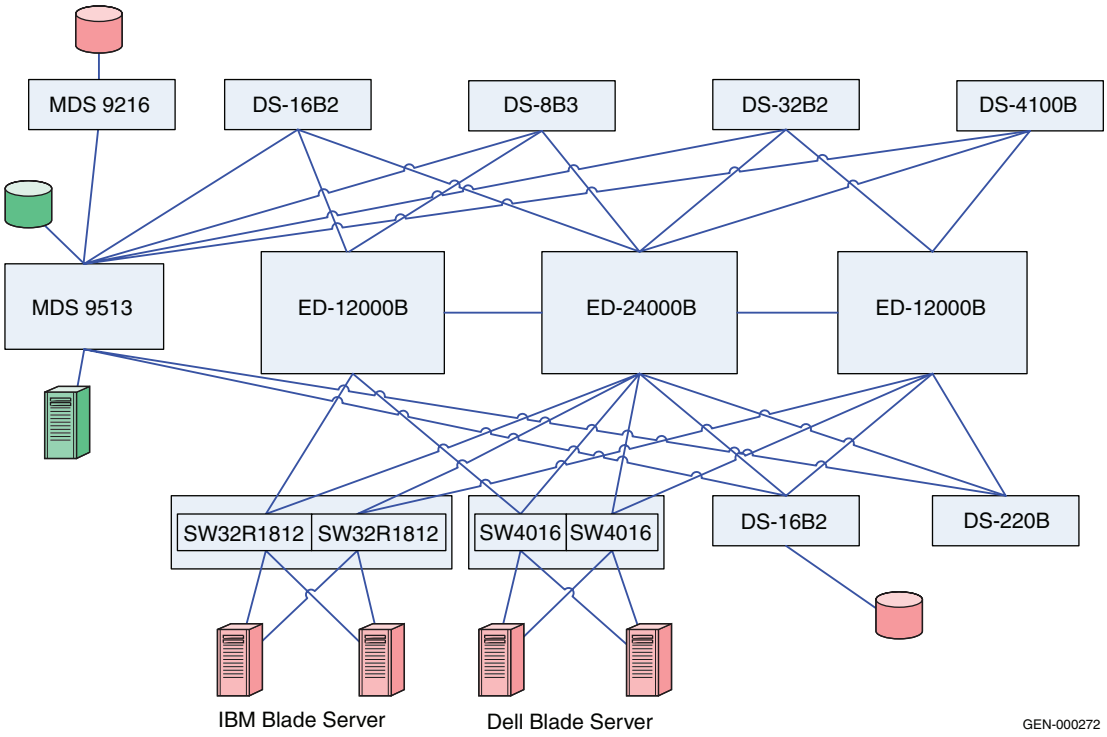


Figure 52 **Phase 6: Moving hosts and storage to new edge**

Hosts and storage ports shared by the other Connectrix B edge switches in the fabric can be completely or partially moved to the Connectrix MDS 9216 edge switch. It is evident in this transitional phase that both the switches, Connectrix B and Connectrix MDS, can co-exist in a stable fabric with the Connectrix B operating in **interopmode 1**and the Connectrix MDS operating in **Interop-1 mode.**

The settings on the switches in this phase can be used to dictate the configuration settings when setting up a Connectrix B–Connectrix MDS interop fabric. All the non-default settings previously discussed for the Connectrix MDS 9513, Connectrix MDS 9216, and the Connectrix B switches apply to setting up a heterogeneous Connectrix MDS-Connectrix B fabric from the ground up.

Phase 7: Adding a Connectrix M switch to the core

Topology

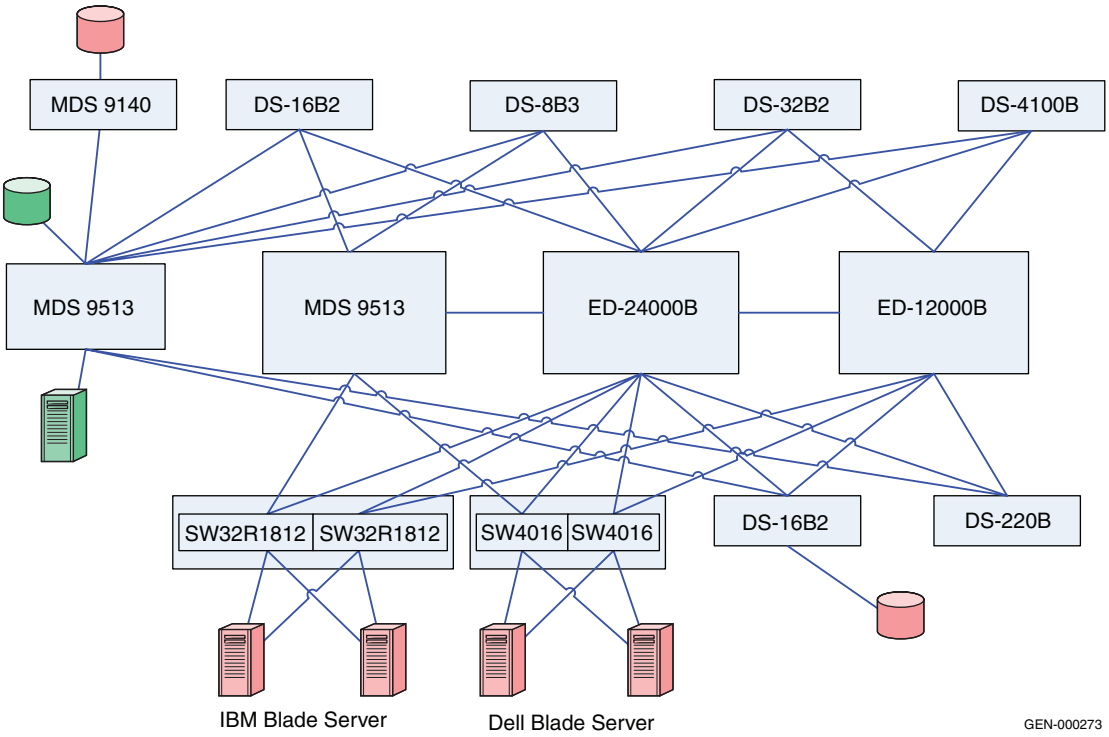


Figure 53 **Phase 7: Adding Connectrix M switch to core**

As shown in Figure 53, another Connectrix MDS switch, the Connectrix MDS 9513, can be added to the core with similar settings as the previous Connectrix MDS 9513, and can be linked with an ISL to the existing core Connectrix MDS 9513 director and the other edge switches in the fabric. Again, it is important to note that the VSAN must be configured with the same VSAN ID and attributes as the existing VSAN to ensure a clean fabric merge.

Complete migration to Connectrix MDS

At the end of case study 4, a migration from a Connectrix B-only fabric to a Connectrix B-Connectrix MDS fabric with Connectrix MDS switches at the core, connected using ISLs to every edge switch in the fabric, is complete. The host and storage ports (except for the blade

server host ports) can be completely moved to the Connectrix MDS switches as specified in Phase 3 (refer to "Phase 3: Moving half of the host and storage ports from Connectrix B core to the Connectrix MDS 9513" on page 309), Phase 4 (refer to "Phase 4: Completely moving the host and storage ports from Connectrix B core to the Connectrix MDS 9513" on page 310) and Phase 6 (refer to "Phase 6: Moving hosts and storage to new edge" on page 313). The Connectrix B edge switches (except for the blade server Brocade switch modules), and the Connectrix B core switches can then be pulled out from the fabric. This results in a fully operational Cisco-only fabric. This is a complete migration from one switch type (Connectrix B) to another type (Connectrix MDS).

IMPORTANT

In a Connectrix MDS-only fabric, all the Connectrix MDS switches should operate in native (default) mode. The Connectrix MDS switches are in interopmode at the end of the migration.

You need not reboot Connectrix MDS switches to change the interopmode. The **Interop mode** attribute on the currently active VSANs must be changed from **Interop-1** to **Default**.

However, the active zones that were pushed on to the Connectrix MDS from the Connectrix M switches cannot be modified once the interopmode on the Connectrix MDS is changed. Zoning must be reconfigured. It is highly recommended that you backup the configuration to avoid losing all zoning information that was present in the active zoneset for Connectrix MDS switches.

Warnings or caveats

Please refer to EMC Knowledgebase solution emc149735 for all interop issues.

Case study #5: **Migrating from a Connectrix M homogeneous fabric in Open Fabric 1.0 mode to a Brocade fabric.**

Assumptions specific to this case study:

- Interoperability mode settings on the switches are:
  - **Brocade M-Series Open Fabric mode 1.0** on the Brocade M-Series switches in the fabric.
  - **Connectrix B interop mode (interopmode 1)** mode on the Connectrix B switches in the fabric.

- Fabric management applications used for managing the fabric:
  - The Connectrix Manager is used to manage the fabric.

Note: The Brocade Fabric Manager and Web Tools cannot be used for management of a Connectrix B-Connectrix M heterogeneous fabric.
Web Tools can be used to change the switch settings on the Connectrix B.

Phase 1: Base configuration – Pre-existing Connectrix M core-edge fabric

Topology

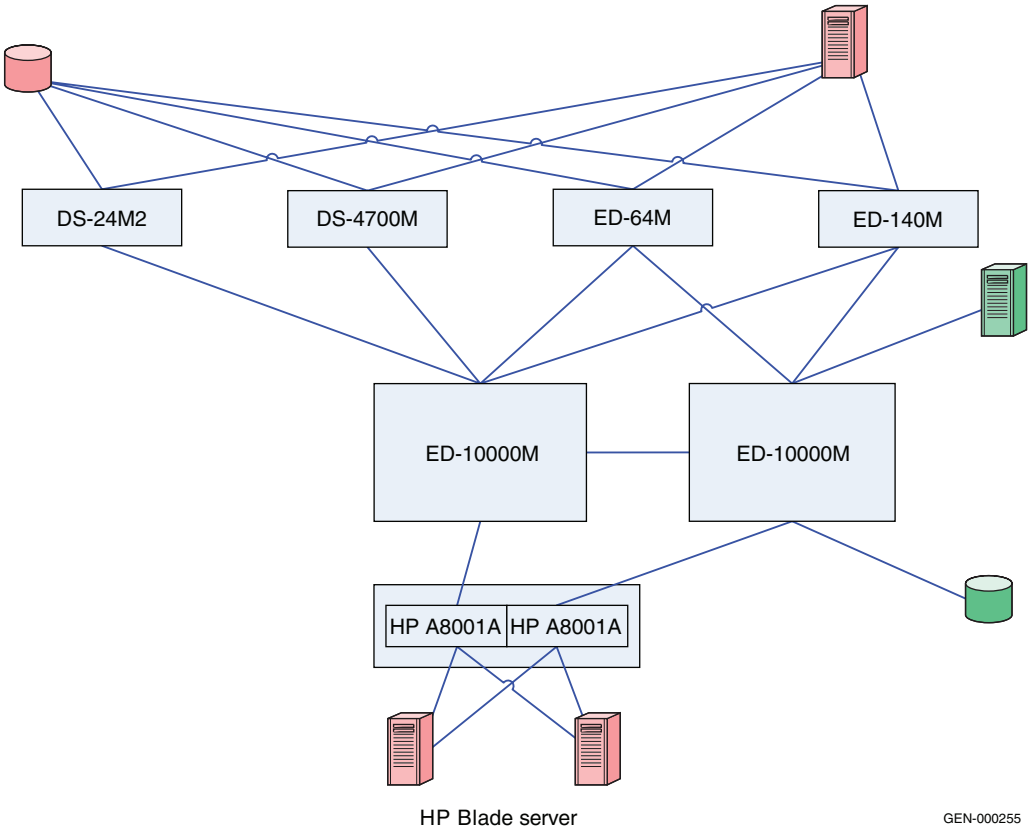


Figure 54 **Phase 1: Basic configuration**

As shown in Figure 54 on page 316, the ED-10000M is at the core of the fabric while the departmental switches: the DS-4500M, DS-4700M, and director switches: ED-64M, ED-140M, and ED-10000M are at the edge. All the switches must be set to the **Open fabric Mode 1.0**. The HP A8001A Brocade M-Series switch modules have also been included as edge switches in this fabric and must be set to the **Standard Mode**. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* to obtain a list of all Connectrix M switches that are supported in a heterogeneous setup, and for supported operating modes.

Specific configuration settings, best practices, host and storage layouts, and topology design that can withstand failures for a Brocade M-Series core-edge homogeneous fabric, were discussed in the "Connectrix M example" on page 102. These settings apply to the base topology in this case study.

Checkpoints

Before adding the ED-24000B switch to the homogeneous Brocade M-Series fabric, verify that the following fabric characteristics exist. Use either the McDATA Connectrix Manager or the McDATA CLI.

- **Proper distribution of Name Server information.**

  Verify that the all host (standalone and server blade) HBA and storage ports that are logged into the fabric appear in the name server.

- **Proper distribution of Zoning information.**

  Verify that the active zoneset consists of the zones containing the desired mapping of host and storage ports.

- **Proper display of fabric and N_Port elements on the management applications.**

  Verify that the physical topology and domain count of the fabric is correct.

- **No disruption in data transfer.**

  Verify that data traffic is running appropriately through the SAN.

Phase 2: Adding the ED-24000B to the core of the fabric

Topology

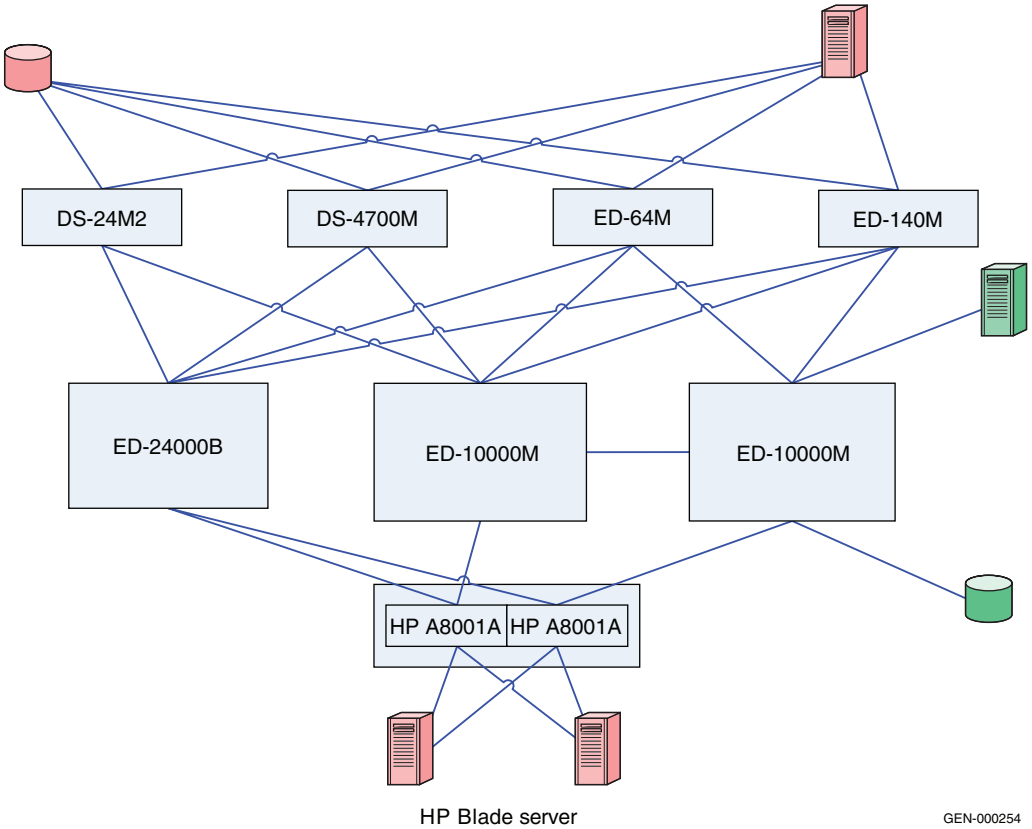


Figure 55 **Phase 2: Adding ED-24000B**

As shown in Figure 55, the ED-24000B switch is added to the core of the fabric. It is important to note that there is no ISL between the ED-10000M core and the Connectrix B core. Please refer to the "Switched Fabric Topology Parameters" section of the *EMC Support Matrix* for a list of all Connectrix B series switches that can be supported in a heterogeneous environment with Connectrix M switches, and for the operating modes (previously specified) for this case study.

Before adding the ED-24000B to the fabric, follow these steps for the Connectrix B switch:

1. Telnet into the ED-24000B.
2. Run the **interopmode** command at the switch command prompt.
3. If interopmode is **On**, the switch is ready to be a part of a fabric where a Cisco switch will be added. If it is **Off**, execute Step 4 through Step 8.
4. Disable the switch by running the **switchdisable** command at the switch prompt.
5. Issue the **interopmode 1** command.
6. Enter a confirmation yes: **y** to switch the interopmode to **On**.
   A warning appears.
   Enter **y** to continue.
7. Reboot the switch to implement configuration settings.
8. After reboot, repeat Step 1 through Step 3 to verify that interopmode was set to **On**.
9. Run the **version** command at the switch command prompt to verify that the switch is running the supported firmware version for interop with Connectrix M.
10. The **msplmgmtdeactivate** command must be run prior to connecting from a Connectrix B switch to a Connectrix M switch.

IMPORTANT

At the end of this phase of the case study, make sure the zoning information was distributed appropriately to the Connectrix B core.

Note: Unlike the Connectrix B-Connectrix MDS interop where the active zoneset or *Effective Configuration* in Connectrix B terminology gets pushed into Connectrix B *Defined Configuration*, the active zoneset from Connectrix M is not pushed into Connectrix B Defined Configuration. Therefore no zoning edits or changes can be made using Connectrix B switch management applications in a Connectrix B-Connectrix M interop environment.

"Checkpoints" on page 317 must be revisited at the end of this phase to verify that the name server information distribution and other fabric parameters based on this topology have been distributed appropriately.

Phase 3: Moving half of the host and storage ports from Connectrix M core to the ED-24000B

Topology

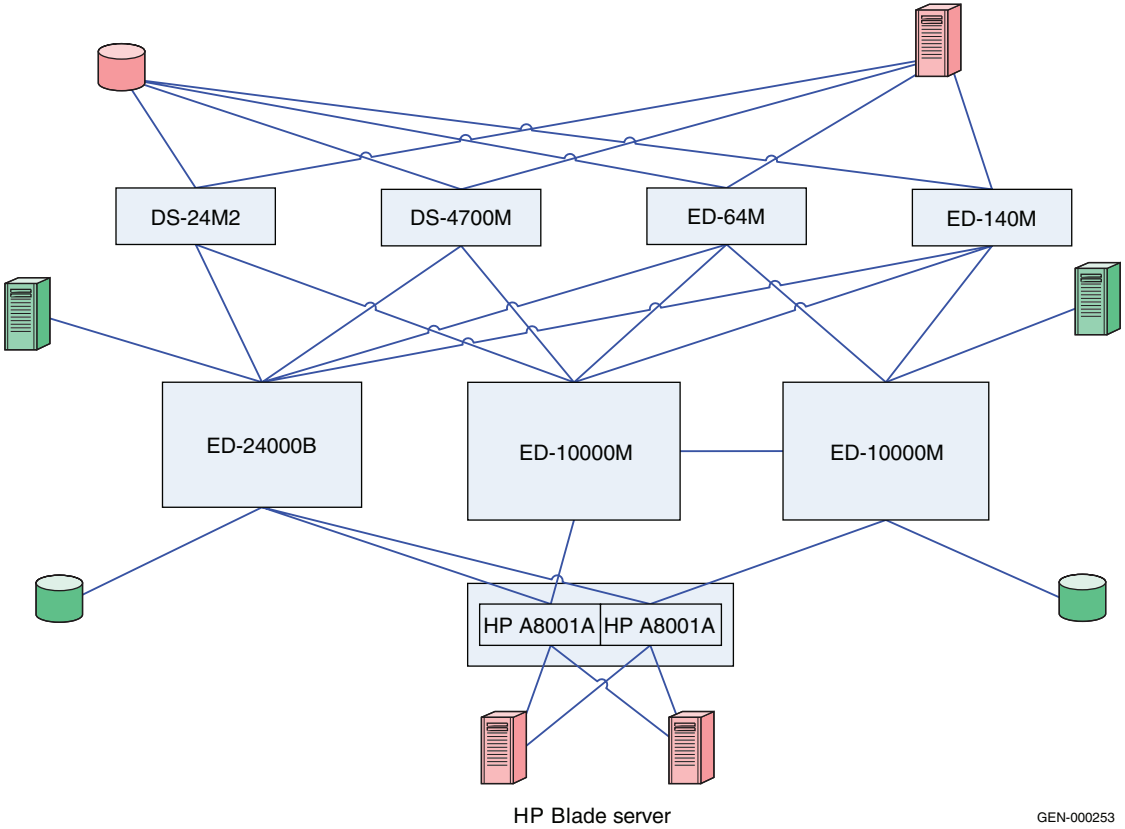


Figure 56 **Phase 3: Moving half the host and storage ports**

As shown in Figure 56, the same host-target pair is now connected to both core switches (the existing ED-10000M core and the newly-added ED-24000B).

This phase is an intermediate phase prior to completely moving the host and storage ports to the Connectrix B core. This phase ensures that the zones are appropriately pushed from the Connectrix M switches to the Connectrix B switches, and that the traffic between the host and storage (which are zoned together) is not disrupted. The host and storage ports are moved to another core switch successfully without any downtime.

It is important to note that at this time *only* Connectrix Manager can be used to make any changes to the active zoneset. The QLogic Connectrix Manager also updates accurately with the changes in the fabric topology, and can be used to manage the HP Brocade M-Series switch modules in this case study.

Next, physically pull cables from one switch and plug them into the other core switch. Review the "Checkpoints" on page 317 and verify that this action did not affect the connectivity and functioning of the fabric.

Phase 4: Moving the host and storage ports from Connectrix M core to the ED-24000B

Topology

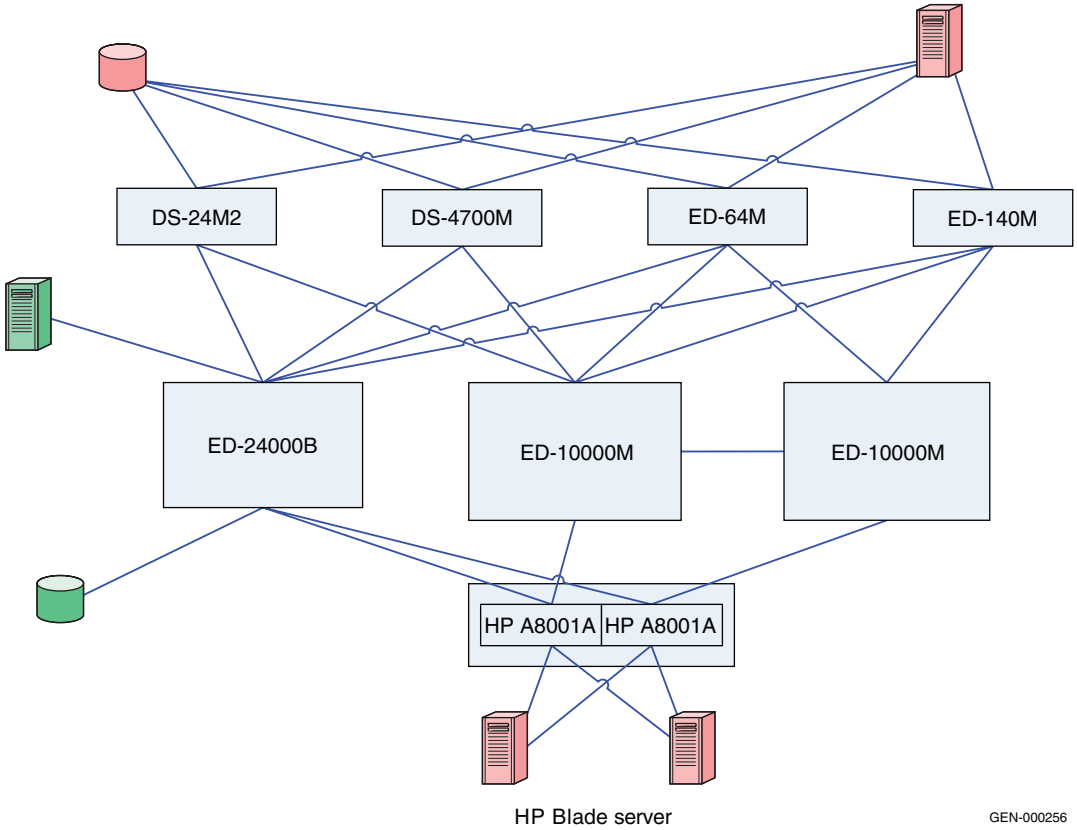


Figure 57 **Phase 4: Completely moving host and storage ports**

Phase 4 is an extension of "Phase 3: Moving half of the host and storage ports from Connectrix M core to the ED-24000B" on page 320. After validating that stable fabric exists after the completion of Phase 3, execute Phase 4 by pulling the remaining host-storage pairs from the ED-10000M switch and transferring them to the ED-24000B core.

Review "Checkpoints" on page 317. Use Web Tools for Connectrix B switch settings, parameters, and connectivity-related information, and Connectrix Manager for the Connectrix M / Connectrix B heterogeneous fabric-related information.

Phase 5: Adding a DS-16B2 and the DS-220B to the edge

Topology

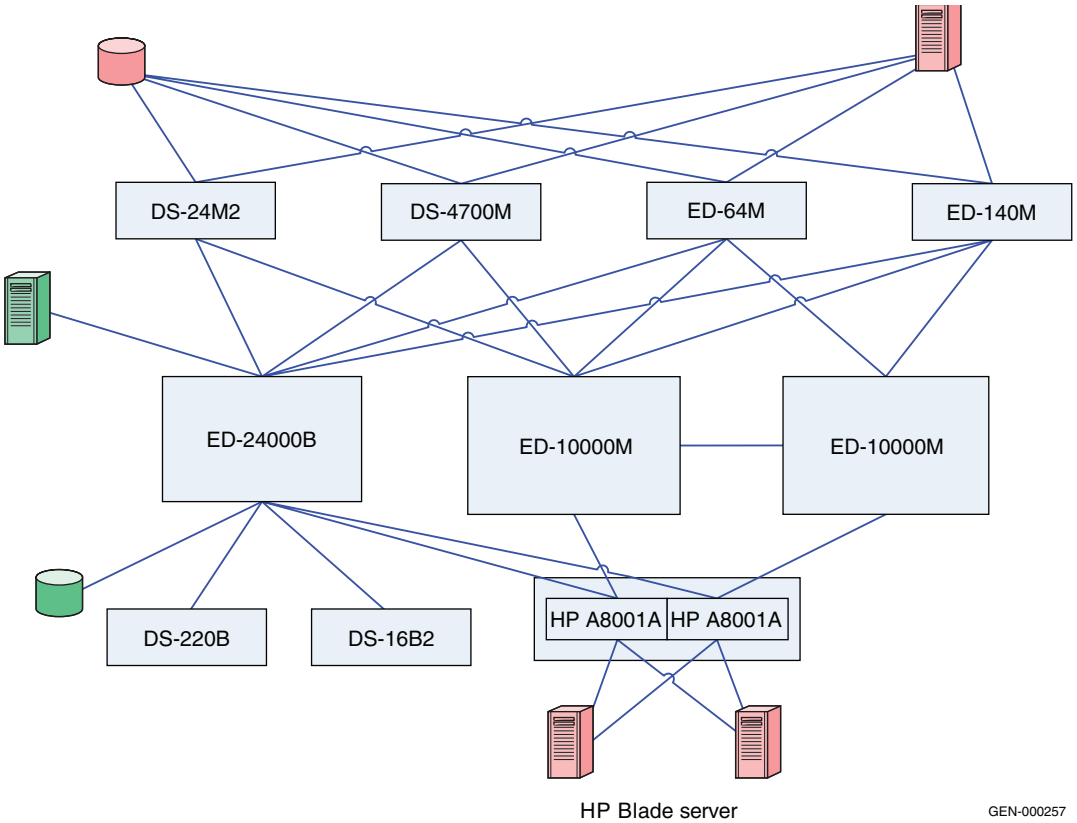


Figure 58 **Phase 5: Adding DS-16B2 and DS-220B**

Follow these steps to configure the ED-24000B before you add it to the core of the fabric. These steps must be executed on the DS-16B2 and the DS-220B.

1. Telnet into the ED-24000B.
2. Run the **interopmode** command at the switch command prompt.
3. If interopmode is **On**, the switch is ready to be a part of a fabric where h a Cisco switch will be added. If it is **Off**, execute Step 4 through Step 8.
4. Disable the switch by running the **switchdisable** command at the switch prompt.
5. Issue the **interopmode 1** command.
6. Enter **y** to set the interopmode to **on**.
   A warning appears.
   Enter **y** at this prompt.
7. Reboot the switch to implement new configuration settings.
8. Once the switch is rebooted, repeat Step 1 through Step 3 to verify that the interopmode has been set to **On**.

If required, after performing the previous procedure, you can link the edge switches with an ISL to both the ED-24000B and ED-10000M core switches.

Follow the "Checkpoints" on page 317 to validate the fabric merge.

Phase 6: Moving hosts and storage to the new edge switches

Topology

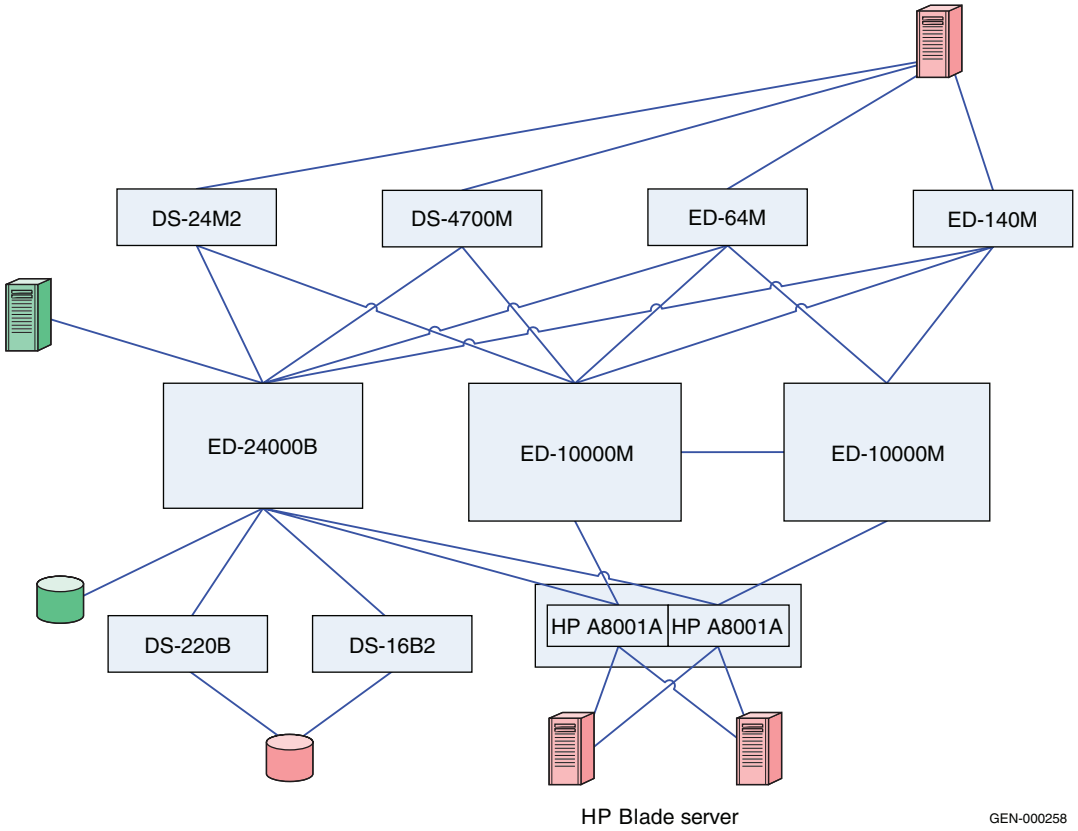


Figure 59 **Phase 6: Moving hosts and storage to the new edge switches**

Hosts and storage ports shared by the other Connectrix M edge switches in the fabric (except for the Server Blade N_Ports logged into the Brocade M-Series Fibre Channel Blade server switch modules) can be completely or partially moved to the DS-16B2 and DS-220B edge switches. In this transitional phase both switches, Connectrix B and Connectrix M, can co-exist in a stable fabric with Connectrix B operating in Connectrix B interopmode and the Connectrix MDS operating in Open Fabric mode 1.0.

The settings on the switches in this phase can be used to dictate the configuration settings to set up a Connectrix B–Connectrix M interop fabric. To set up a successful heterogeneous Connectrix M /

Connectrix B fabric from the ground up, execute all the non-default settings discussed in the previous phases for the Connectrix B and Connectrix M switches and validate the "Checkpoints" on page 317 for all switches across the fabric.

Phase 7: Adding a Connectrix B switch to the core

Topology

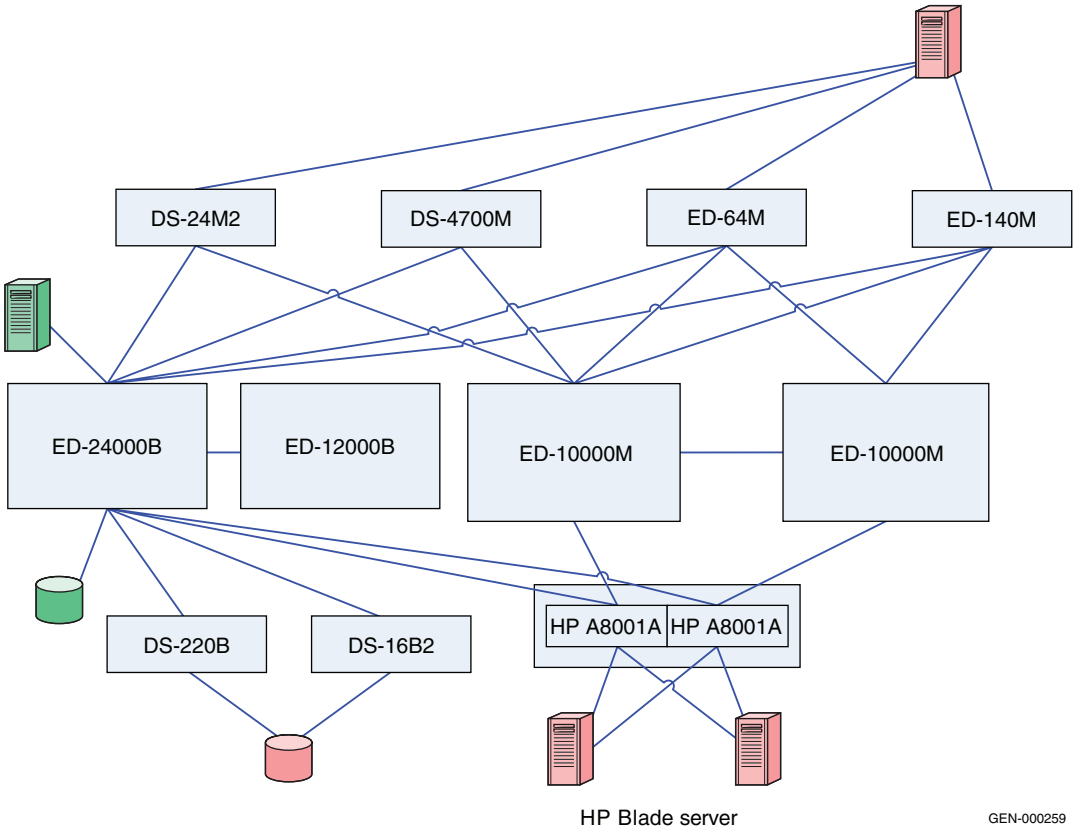


Figure 60 **Phase 7: Adding Connectrix B switch**

As shown in Figure 60, another Connectrix B director switch, the ED-12000B can be added to the core with similar settings as the ED-24000B that was previously added to the fabric, and can be connected through an ISL to the existing core ED-24000B switch and the other edge switches in the fabric. The "Checkpoints" on page 317 must be followed to ensure a clean fabric merge.

Complete migration to Connectrix B

At the end of case study 5, a full migration from a Connectrix M-only fabric to a Connectrix B-Connectrix M fabric with Connectrix Bs at the core, connected through ISLs to every edge switch in the fabric occurred. The host and storage ports (except for the blade server host ports) can be completely moved over to the Connectrix B switches as specified in Phase 3 (refer to "Phase 3: Moving half of the host and storage ports from Connectrix M core to the ED-24000B" on page 320), Phase 4 (refer to "Phase 4: Moving the host and storage ports from Connectrix M core to the ED-24000B" on page 321), and Phase 6 (refer to "Phase 6: Moving hosts and storage to the new edge switches" on page 324). The Connectrix M edge switches (except for the blade server Brocade M-Series switch modules) and the Connectrix M core switches can then be pulled out from the fabric. This results in a fully operational Connectrix B-only fabric. This is a complete migration from one switch type (Connectrix M) to another type (Connectrix B).

IMPORTANT

In a Connectrix B-only fabric, make sure Connectrix B switches operate in their native (interopmode 0) mode. Connectrix B switches are in their interopmode at the end of the migration and it is not possible to make any zoning change to the existing active zoneset on Connectrix B switches since it has been pushed to Connectrix B switches by Connectrix M switches.

In order to change the interopmode on the Connectrix B switches, it is necessary to reboot the switches. As a best practice, it is recommended to reboot the edge switches and then the core switches, depending on how many N_Ports (host and storage) are connected to the edge or core switches, and depending on how rebooting a particular switch affects the I/O running through the fabric.

After rebooting the switches, the active zoning configuration (effective and defined configuration) on Connectrix Bs is erased, and it is necessary to reconfigure the zoning on Connectrix B switches. It is highly recommended that you backup the configuration and activate it from the Connectrix B switch management application once it is back up (after the reboot) to avoid losing all zoning information previously on the Connectrix B switches.

Warnings or caveats

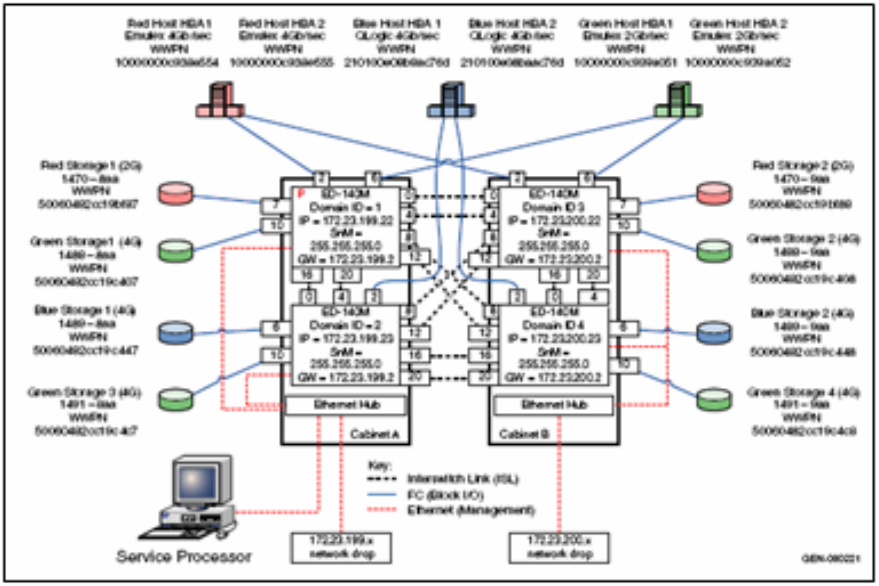
Please refer to EMC Knowledgebase solution emc149735 for all interop issues.

Case study #6: Migrating from a Connectrix M-Series, McDATA mode to Connectrix B-Series, Interopmode 0

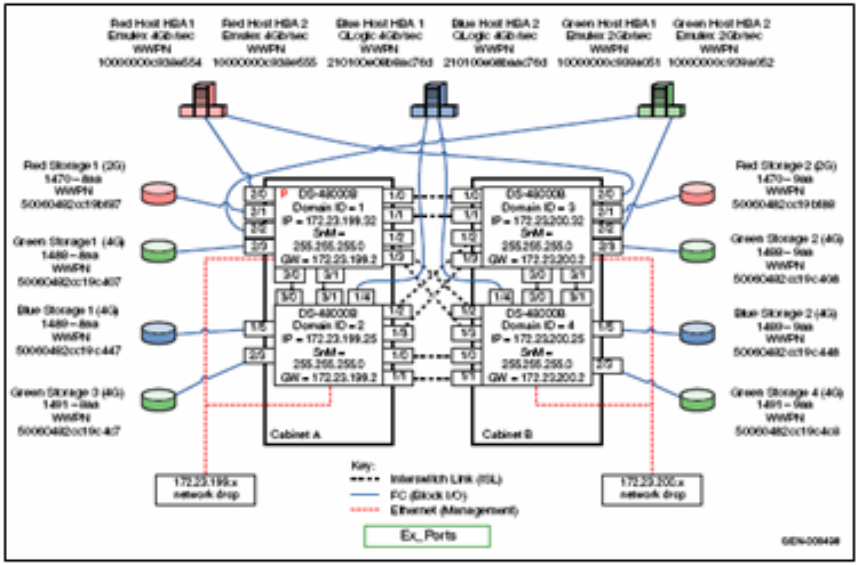
In response to customer demand for migration solutions, EMC E-Lab and Brocade solution center collaborated on the qualification of the following topology.

This example will walk the user through one possible way of migrating from a Brocade M-EOS fabric running in McDATA Fabric Mode (the donor fabric, also referred to as the source) to a Brocade FOS fabric running in its native mode (the target fabric).

Figure 61 on page 328 shows the donor (source) and target topology. For more details on the donor topology, refer to the "Connectrix M example" on page 171. For more details on the target topology, refer to the "Connectrix B example" on page 201.



Donor (Source) : Connectrix M



Target: Connectrix B

Figure 61 Donor and target topology

Note: A Brocade SAN Router is used to support the migration of devices from a Connectrix M fabric running in McDATA Fabric mode (native mode) to a Connectrix B fabric running in Interopmode 0. A Connectrix B SAN Router is available in two form factors:

– As a blade (FR4-18i) installed in a slot of an ED-48000B director,
– As a stand alone, 1U pizza box (SAN Router MP-7500B).

Either can be used for migration. The FR4-18i blade is described in this case study. Refer to "Brocade SAN Routing — FCR" on page 944 for more details about SAN routers and SAN routing.

Assumptions specific to this case study

- Interoperability mode settings on the switches:
  - **Brocade M-Series Fabric** — The Connectrix M-EOS switches in the fabric are configured to operating in McDATA Fabric mode (M-EOS native mode / McDATA Fabric mode 1.0).
  - **Brocade B-Series Fabric** — Connectrix B-FOS switches have been configured to operate in Interopmode 0 (FOS native mode).
- Fabric Manager applications used for managing the fabric:
  - Although this environment can be managed by Connectrix Manager (Brocade EFCM), Brocade Web Tools, or Brocade CLI, the configuration in this case study will be done via Connectrix Manager for the M-EOS products and Brocade CLI for the FOS- based products. :
- HA configuration
  - Mirrored fabrics for both the Connectrix M and Connectrix B fabrics are supported in configurations similar to this example. However, for simplicity, only one Connectrix M and Connectrix B fabric are shown. The same procedures would be applied to other fabrics to complete the migration..
  - Multipath software is installed on all hosts to allow a non-disruptive migration of fabric connections from Connectrix M to Connectrix B fabric.

Phase 1: Base (donor) configuration of Connectrix M-Series fabric

Topology

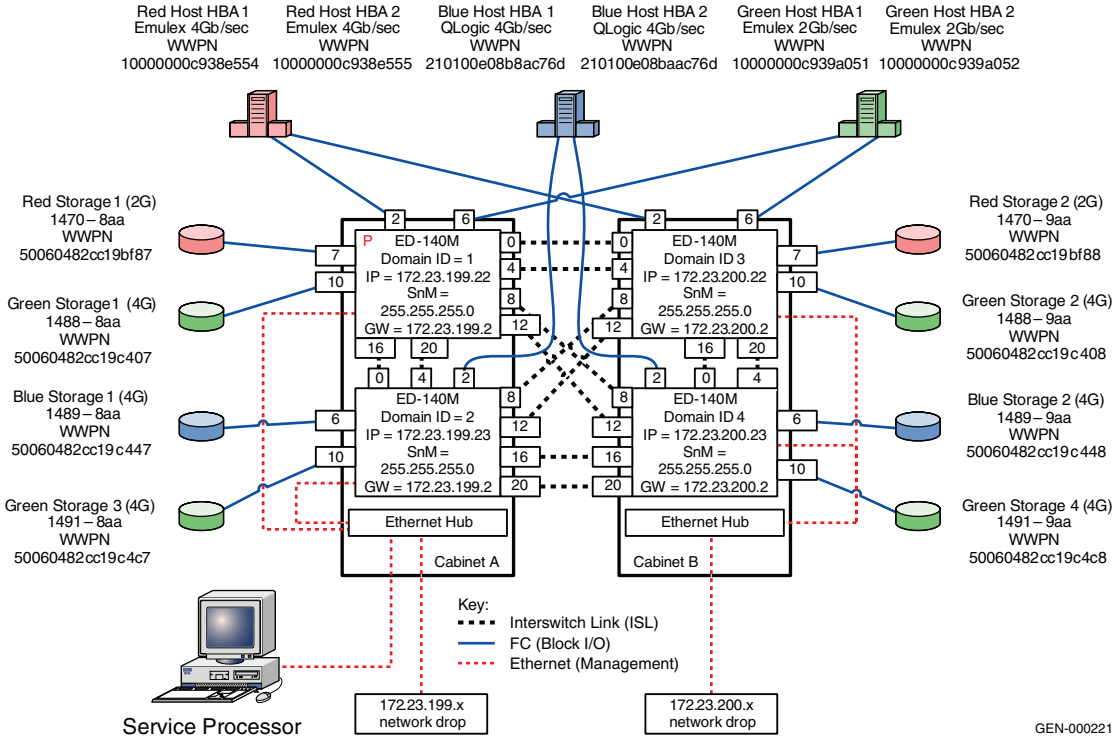


Figure 62 **Phase 1: Basic configuration of Connectrix M fabric**

As illustrated in Figure 62, the donor fabric topology is a four switch mesh fabric using ED-140M director switches. Please refer to the "Switched Fabric Topology Parameters"section of the *EMC Support Matrix* for a list of all the other Brocade M-Series / EMC Connectrix M Series switches that can be supported in a heterogeneous setup and for the operating modes specified above for this case study.

Note: At the time of this publication, the ED-10000M (Brocade M i10K) and the Blade server Brocade M-Series modules (QLogic modules in Brocade M-Series McDATA mode (native mode)) are not supported in an M-series fabric connected to a Brocade SAN router, and thus cannot participate in the switch migration procedure discussed in this case study. Please refer to the *EMC Support Matrix* for the most up-to-date support.

The specific configuration settings, best practices, host and storage layouts, and topology design for this configuration are discussed in the "Connectrix M example" on page 171.

Checkpoints

It is advisable to verify the following fabric characteristics:

- **Proper distribution of the name server information**.

  Verify that all the host HBA ports and storage ports logged into the fabric are listed in the name server. This can be accomplished by looking at the node list in the switch's Element Manager from within Connectrix Manager or by using the **show nameserver** CLI command. Depending on the switch this command was issued on, different information will be displayed.

- **Proper distribution of zoning information.**

  Use the **Show zoning** CLI command or the zoning dialog in Connectrx Manager to verify that the active zone set comprises zones that contain the desired mapping of host and storage ports.

- **Proper display of fabric and N_Port elements on the management applications**.

  Use the **show fabric topology** CLI command or the topology view in Connectrix Manager to verify that you have the desired physical topology and domain count in the fabric.

- **No disruption in data transfer**.

  Verify that the data traffic is running appropriately through the SAN. This can be done in any number of ways. One possibility is to use the EMC `inq` utility to ensure that the appropriate number of LUNs are visible to the OS. Be sure to save the`inq` output to a file for future reference. The `inq` utility is available at ftp://ftp.emc.com/pub/symm3000/inquiry/ (anonymous login).

Before adding the Connectrix B-series fabric, it is advisable to verify the following fabric characteristics of the Connectrix M-series fabric and to backup critical configuration information using either Connectrix Manager, Connectrix Manager Basic, or the Brocade M-Series CLI:

- **Save active zone set.**

  Verify that a backup copy of the active zone set has been saved for the Connectrix M fabric.

This can be accomplished via the zoning dialog in Connectrix Manager. Simply select the fabric of interest by single clicking a switch in the fabric. Next, click the zoning icon from the top of the topology view and the zoning dialog displays. Select the active zone set tab and click **save active**. Provide a name and then click **OK**.

This can also be done via CLI. Telnet into a switch in the fabric and use the **config zoning replacezoneset** command.

- **Backup active switch configurations.**

  Create a backup of the switch configurations in the Connectrix M fabric. This can be accomplished from the switches Element Manager within Connectrix Manager or by using SANPlicity.

- **Perform a "data collection" from each switch in the fabric**.

  Create a data collection for each switch the existing Connectrix M fabric. This is best accomplished via the switch Element Manager within Connectrix Manager.

Phase 2: Create the target Connectrix B fabric

Topology

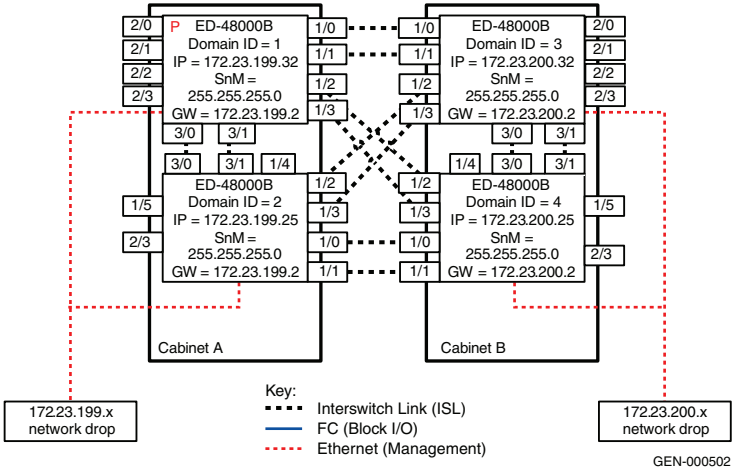


Figure 63 **Phase 2: Adding Connectrix B fabric and SAN router**

As shown in Figure 63, a Connectrix B fabric (Interopmode 0) is created added using in a four switch mesh topology. This will be the "Target" fabric to which the host and storage ports will be migrated. For simplification, the target topology shown matches the base donor

fabrics topology used for the Connectrix M fabric, but in practice, the target topology can be any suitable topology supported for Connectrix B-series fabrics.

The specific configuration settings, best practices, host and storage layouts, and topology design for this configuration are discussed in the "Connectrix B example" on page 201.

Phase 3: Configuring Connectrix B SAN routing

Topology:

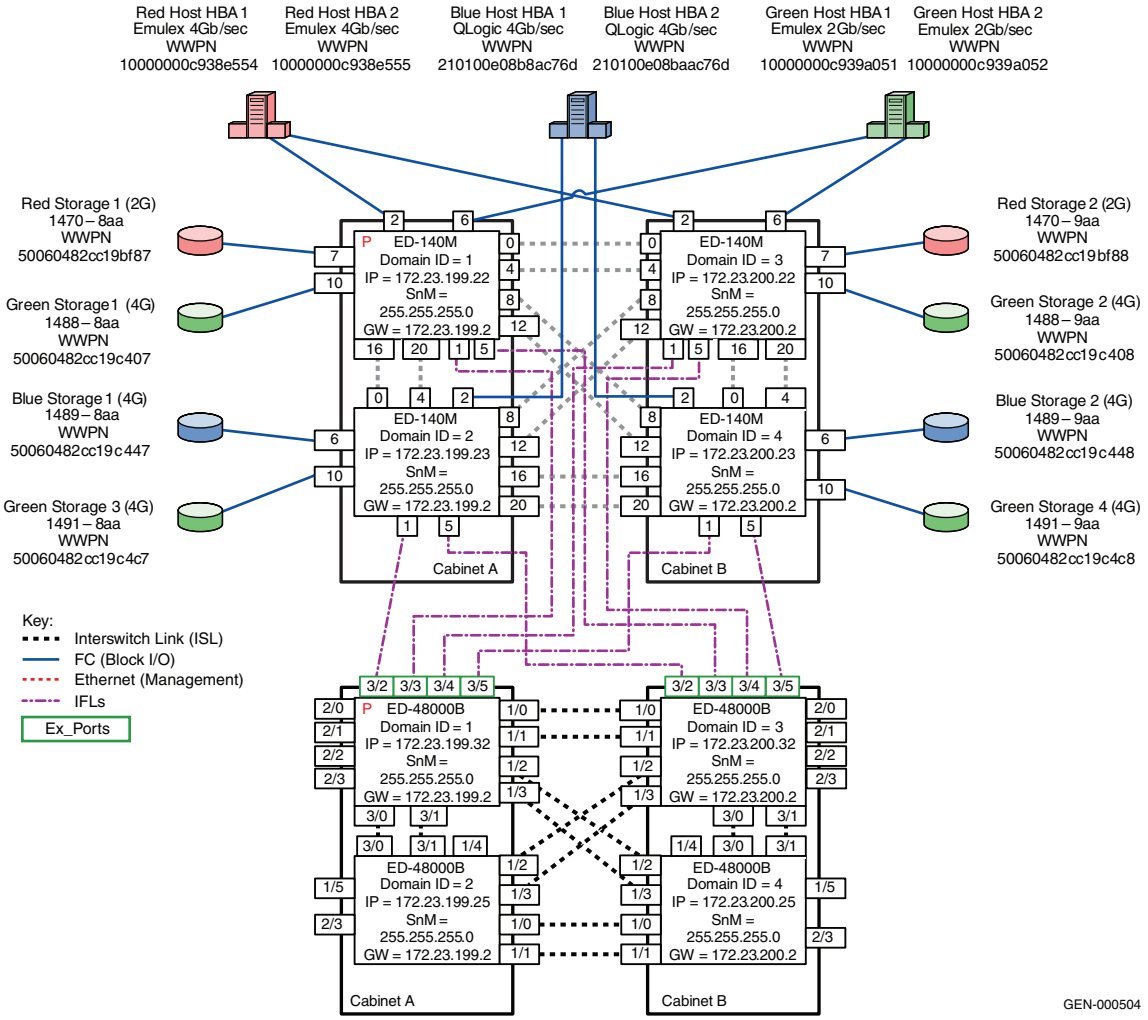


Figure 64 **Phase 3: Configuring Connectrix B SAN routing**

In this phase, the Connectrix B SAN routing services are configured. The SAN Router is the FR4-18i SAN Router blade and is installed in two of the ED-48000B directors, Domain 1 and Domain 3. These blades were previously installed in Phase 2 and two of their ports were deployed as E_Ports within the Connectrix B fabric. In this phase, four of the SAN Router ports are configured as Inter-Fabric

Links (IFL). The IFLs use EX_Ports in the Connectrix B fabric and use E_Ports to attach to the Connectrix M fabric. Refer to "Brocade SAN Routing — FCR" on page 944 for additional background information.

To set up the SAN routing service, perform the following tasks:

1. "Perform verification checks.," discussed next.
2. "Configure EX_Ports and Interfabric Links (IFLs)." on page 337
3. "Configure Logical SAN (LSAN) zoning" on page 340

Each of these tasks are further described in this section.

1. Perform verification checks.

   Before configuring a Connectrix B SAN Router to connect to another fabric, you must perform the following verification checks on the director with the SAN Router. Since there are two SAN routers installed, these checks should be done on both ED-48000B directors identified as Domain 1 and Domain 3..

   a. Log in to the ED-48000B as admin and enter the **version** command. Verify that Fabric OS v5.3.0 (or higher) is installed on the SAN router, as shown in the following example.

   ```
   mw116:admin> version
   Kernel:     2.6.14
   Fabric OS:  v5.3.0
   Made on:    Mon Apr 2 21:44:31 2007
   Flash:      Fri Apr 6 00:34:34 2007
   BootProm:   4.6.2
   ```

   b. Enter the **slotShow** command to verify that the FR4-18i blade is present The FR4-18i will be shown as a blade type of "AP BLADE" with an ID of 24.

   ```
   smw116:admin> slotshow

   Slot   Blade Type   ID    Status
   --------------------------------
     1     SW BLADE   4      ENABLED
     2     SW BLADE   39     ENABLED
     3     AP BLADE   24     ENABLED
     4     UNKNOWN           VACANT
     5     UNKNOWN           VACANT
     6     UNKNOWN           VACANT
     7     UNKNOWN           VACANT
     8     UNKNOWN           VACANT
     9     UNKNOWN           VACANT
    10     UNKNOWN           VACANT
   ```

c. Enter the **chassisConfig** command to verify that the director is using configuration option

```
switch:admin_06> chassisconfig
Current Option: 5

All Supported Options
-----------------------------------------------------
Option 1: One 128-port switch
    Blade ID's 4, 24, 39 in slots 1-4, 7-10
    Blade ID's 5, 16 in slots 5-6
Option 5: One 384-port switch
    Blade ID's 4, 24, 39 in slots 1-4, 7-10
    Blade ID 16 in slots 5-6

Please use slotshow to see Blade IDs currently in the
system.
```

d. Enter the **interopMode** command and verify that Connectrix B switch interoperability with switches from other manufacturers is disabled.

```
mw116:admin> interopmode
InteropMode: Off

Usage: InteropMode 0|1
         0: to turn it off
         1: to turn it on
```

e. Enter the **secModeShow** command to verify that security is disabled.

```
mw116:admin> secmodeshow
Secure Mode: DISABLED.
```

f. Enter the **msPlatShow** command to verify that Management Server Platform database is disabled in the backbone fabric.

```
switch:admin_06> msplatshow
*MS Platform Management Service is NOT enabled.
```

If any of the items listed in the prior steps are enabled, refer to the *EMC Connectrix B Series Fabric OS Command Reference Manual* for information on how to disable the option.

Note: When it is in strict mode, ACL cannot support Fibre Channel routing in the fabric. Before connecting an edge fabric to a FC router, and before setting up the FC router in the BB, verify that the Fabric Wide Consistency Policy is not in 'strict' mode by issuing the **fddCfg --showall** command.

If the Fabric Wide Consistency Policy has the letter 'S' in it in the edge fabric or the BB fabric, do not connect the edge fabric or the BB to the FC router. See the *EMC Connectrix B Series Fabric OS Command Reference Manual* for details.

2. Configure EX_Ports and Interfabric Links (IFLs).

   Configuring an Inter fabric Link (IFL) involves disabling ports and cabling them to the switches in the Connectrix M fabric, configuring those ports for their intended use, and then enabling the ports.

   Note: You cannot configure both IFLs (EX_Ports) and ISLs (E_Ports) from a single SAN Router to the same edge fabric.

   - To configure EX_Ports from Connection to Connectrix M fabric, perform the following steps:

     a. On the SAN Router with the IP Address of 172.23.199.32, disable the ports that you are configuring as an EX_Port for connection to the Connectrix M-series fabric by issuing the **portDisable** command.

     ```
     switch:admin_06> portdisable 3/2
     switch:admin_06> portdisable 3/3
     switch:admin_06> portdisable 3/4
     switch:admin_06> portdisable 3/5
     ```

     You can verify that ports have been disabled by issuing the **portShow** command for each port.

     ```
     mw116:admin> portshow 3/2
     portName:
     portHealth: OFFLINE
     … [rest cut off]
     ```

     Repeat this command for ports 3, 4 and 5 to verify they are of-line.

b. Use the **portCfgExPort** command to configure the EX_Ports connected to the Connectrix M fabric. Assign a Fabric ID of 10 (-f 10) and a fabric mode of McDATA Fabric Mode (-m 2) for the Connectrix M fabric.

```
switch:admin_06> portcfgexport 3/2 -a 1 -f 10 -m 2
switch:admin_06> portcfgexport 3/3 -a 1 -f 10 -m 2
switch:admin_06> portcfgexport 3/4 -a 1 -f 10 -m 2
switch:admin_06> portcfgexport 3/5 -a 1 -f 10 -m 2
```

c. Use the **portCfgExPort c**ommand to verify the configuration of each of the EX_Ports just configured.

```
mw116:admin> portcfgexport 3/2
        Port   3/2   info
Admin:                    enabled
State:                    NOT OK
Pid format:               Not Applicable
Operate mode:             McDATA Fabric
Edge Fabric ID:           10
Preferred Domain ID:      160
Front WWN:                50:06:06:98:00:fa:ae:1e
Fabric Parameters:        Auto Negotiate
R_A_TOV:                  Not Applicable
E_D_TOV:                  Not Applicable
Authentication Type: None
DH Group: N/A
Hash Algorithm: N/A
Edge fabric's primary wwn: N/A
Edge fabric's version stamp: N/A
```

d. Enter the portEnable command to enable the ports that you disabled in Step a.

```
switch:admin_06> portenable 3/2
switch:admin_06> portenable 3/3
switch:admin_06> portenable 3/4
switch:admin_06> portenable 3/5
```

e. Repeat the above steps for the switch with an IP address of 172.23.200.32.

- To connect EX_Ports and verify EX_Ports are online, perform the following steps:

  a. Physically attach the ISLs from the SAN router to the Connectrix M edge fabric switches as shown in Figure 64 on page 334.

b. On the switch with an IP Address of 172.23.199.32, enter the **switchShow** command to verify the EX_Port, edge fabric ID, and name of the edge fabric switch (containing the E_Port).

```
mw116:admin> switchshow
switchName:     mw116
switchType:     42.2
switchState:    Online
switchMode:     Native
switchRole:     Subordinate
switchDomain:   116
switchId:       fffc71
switchWwn:      10:00:00:60:69:80:0f:aa
zoning:         ON (SanMig)
switchBeacon:   OFF
blade1 Beacon:  OFF
blade2 Beacon:  OFF
blade8 Beacon:  OFF
FC Router:      ON
FC Router BB Fabric ID: 1

Index Slot Port Address Media Speed State     Proto
=====================================================
  0    1    0   740000   --    10   No_Module
  1    1    1   740100   --    10   No_Module
  2    1    2   740200   --    10   No_Module
  3    1    3   740300   --    10   No_Module
  4    1    4   740400   id    10   In_Sync          LS
  5    1    5   740500   id    10   No_Light         LS
... [removed lines for slot 1]

 16    2    0   741000   --    N4   No_Module        Disabled (Persistent)
 17    2    1   741100   --    N4   No_Module        Disabled (Persistent)
... [removed lines for slot 2]

 80    3    0   745000   id    N2   Online          E-Port  10:00:00:05:1e:35:e0
 81    3    1   745100   id    N2   Online          E-Port  10:00:00:05:1e:35:e0
 82    3    2   745200   id    N2   Online          EX-Port  10:00:08:00:88:a0:d4:8d
"6140_80" (fabric id = 10 )
 83    3    3   745300   --    N4   Online          EX-Port  10:00:08:00:88:a0:d4:ff
"6140_81" (fabric id = 10 )
 84    3    4   745400   id    N2   Online          EX-Port  10:00:08:00:88:a0:d5:01
"6140_82" (fabric id = 10 )
 85    3    5   745500   --    N4   Online          EX-Port  10:00:08:00:88:a0:ef:50
"6140_83" (fabric id = 10 )
 86    3    6   745600   --    N4   No_Module
 87    3    7   745700   --    N4   No_Module
...[removed lines for slot 3]
```

c. Enter the **fcrFabricShow** command to view the Connectrix M fabric's switch names and ensure links are working as expected:

```
mw116:admin> fcrfabricshow
FC Router WWN: 10:00:00:60:69:80:0f:aa, Dom ID: 1,
Info: 10.66.14.116, "mw116"
   EX_Port       FID    Neighbor Switch Info (enet IP, WWN, name)
   ---------------------------------------------------------------------
      82         10     172.23.199.22    10:00:08:00:88:a0:d4:8d   "6140_80"
      83         10     172.23.199.23    10:00:08:00:88:a0:d4:ff   "6140_81"
      84         10     172.23.200.22    10:00:08:00:88:a0:d5:01   "6140_82"
      85         10     172.23.200.23    10:00:08:00:88:a0:ef:50   "6140_83"
```

d. Repeat Step b and Step c on the switch with an IP Address of 172.23.200.32.

3. Configure Logical SAN (LSAN) zoning

Device connections across fabrics relies on zoning. A special type of zone, the LSAN zone, is created in each edge fabric containing the devices in any edge fabrics which need to connect.

- To create LSAN zones in Connectrix B fabric, perform the following steps:

  a. Telnet to the switch with the IP Address of 172.23.199.32 and log in as admin.

  b. Use the **zoneCreate** command to create the LSAN zone for the red host and targets.

```
switch:admin_06> zonecreate "LSAN_RedHBA1_1470_8aa", "10:00:00:00:c9:38:e5:54;
50:06:04:82:cc:19:bf:87"
```

c. Using the **zoneCreate** command to add the other LSAN zones.

```
switch:admin_06> zonecreate "LSAN_RedHBA2_1470_9aa ", "10:00:00:00:c9:38:e5:55;
50:06:04:82:cc:19:bf:88"
switch:admin_06> zonecreate "LSAN_BlueHBA1_1489_8aa ", "21:01:00:e0:8b:8a:c7:6d;
50:06:04:82:cc:19:c4:47"
switch:admin_06> zonecreate "LSAN_BlueHBA2_1489_9aa ", "21:01:00:e0:8b:aa:c7:6d;
50:06:04:82:cc:19:c4:48"
switch:admin_06> zonecreate "LSAN_GreenHBA1_AllGreenStorage",
"10:00:00:00:c9:39:a0:51; 50:06:04:82:cc:19:c4:07; 50:06:04:82:cc:19:c4:08;
50:06:04:82:cc:19:c4:c7; 50:06:04:82:cc:19:c4:c8"
switch:admin_06> zonecreate "LSAN_GreenHBA2_AllGreenStorage",
"10:00:00:00:c9:39:a0:52; 50:06:04:82:cc:19:c4:07; 50:06:04:82:cc:19:c4:08;
50:06:04:82:cc:19:c4:c7; 50:06:04:82:cc:19:c4:c8"
```

d. Enter the **cfgAdd** command to add the LSAN zones to the active zone set.

```
switch:admin_06> cfgadd "zone_cfg", "LSAN_RedHBA1_1470_8aa"
switch:admin_06> cfgadd "zone_cfg", "LSAN_RedHBA2_1470_9aa "
switch:admin_06> cfgadd "zone_cfg", "LSAN_BlueHBA1_1489_8aa "
switch:admin_06> cfgadd "zone_cfg", "LSAN_BlueHBA2_1489_9aa "
switch:admin_06> cfgadd "zone_cfg", "LSAN_GreenHBA1_AllLGreenStorage "
switch:admin_06> cfgadd "zone_cfg", "LSAN_GreenHBA2_AllGreenStorage "
```

e. Enter the **cfgShow** command to display the zones in the active zone set.

```
mw116:admin> cfgshow
Defined configuration:
 cfg:   zoneconfig  LSAN_RedHBA1_1470_8aa; LSAN_RedHBA2_1470_9aa;
LSAN_BlueHBA1_1489_8aa; LSAN_BlueHBA2; LSAN_GreenHBA1_AllGreenStorage;
LSAN_GreenHBA2_AllGreenStorage

 zone:  LSAN_RedHBA1_1470_8aa
                10:00:00:00:c9:38:e5:54; 50:06:04:82:cc:19:bf:87
 zone:  LSAN_RedHBA2_1470_9aa
                10:00:00:00:c9:38:e5:55; 50:06:04:82:cc:19:bf:88
 zone:  LSAN_BlueHBA1_1489_8aa
                21:01:00:e0:8b:8a:c7:6d; 50:06:04:82:cc:19:c4:47
 zone:  LSAN_BlueHBA2_1489_9aa
                21:01:00:e0:8b:aa:c7:6d; 50:06:04:82:cc:19:c4:48
 zone:  LSAN_GreenHBA1_AllGreenStorage
                10:00:00:00:c9:39:a0:51; 50:06:04:82:cc:19:c4:07;
                50:06:04:82:cc:19:c4:08; 50:06:04:82:cc:19:c4:c7;
        50:06:04:82:cc:19:c4:c8;
 zone:  LSAN_GreenHBA2_AllGreenStorage
                10:00:00:00:c9:39:a0:52; 50:06:04:82:cc:19:c4:07;
                50:06:04:82:cc:19:c4:08; 50:06:04:82:cc:19:c4:c7;
                  50:06:04:82:cc:19:c4:c8;

Effective configuration:
 No Effective configuration: (No Access)
```

f. Enter the **cfgEnable** command to enable the updated active zone set.

```
switch:admin_06> cfgenable "zoneconfig"
You are about to enable a new zoning configuration.
This action will replace the old zoning configuration with the
current configuration selected.
Do you want to enable 'zoneconfig' configuration (yes, y, no, n): [no] y
zone config "zoneconfig" is in effect
Updating flash ...
```

g. Display the defined, and now effectiv,e configurations using the **cfgshow** command.

```
mw116:admin> cfgshow
Defined configuration:
 cfg:   zoneconfig  LSAN_RedHBA1_1470_8aa; LSAN_RedHBA2_1470_9aa;
LSAN_BlueHBA1_1489_8aa; LSAN_BlueHBA2; LSAN_GreenHBA1_AllGreenStorage;
LSAN_GreenHBA2_AllGreenStorage

 zone:  LSAN_RedHBA1_1470_8aa
                10:00:00:00:c9:38:e5:54; 50:06:04:82:cc:19:bf:87
 zone:  LSAN_RedHBA2_1470_9aa
                10:00:00:00:c9:38:e5:55; 50:06:04:82:cc:19:bf:88
 zone:  LSAN_BlueHBA1_1489_8aa
                21:01:00:e0:8b:8a:c7:6d; 50:06:04:82:cc:19:c4:47
 zone:  LSAN_BlueHBA2_1489_9aa
                21:01:00:e0:8b:aa:c7:6d; 50:06:04:82:cc:19:c4:48
 zone:  LSAN_GreenHBA1_AllGreenStorage
                10:00:00:00:c9:39:a0:51; 50:06:04:82:cc:19:c4:07;
                50:06:04:82:cc:19:c4:08; 50:06:04:82:cc:19:c4:c7;
                  50:06:04:82:cc:19:c4:c8;
 zone:  LSAN_GreenHBA2_AllGreenStorage
                10:00:00:00:c9:39:a0:52; 50:06:04:82:cc:19:c4:07;
                50:06:04:82:cc:19:c4:08; 50:06:04:82:cc:19:c4:c7;
                  50:06:04:82:cc:19:c4:c8;

Effective configuration:
 cfg:   zoneconfig
 zone:  LSAN_RedHBA1_1470_8aa
                10:00:00:00:c9:38:e5:54;
                50:06:04:82:cc:19:bf:87
 zone:  LSAN_RedHBA2_1470_9aa
                10:00:00:00:c9:38:e5:55;
                    50:06:04:82:cc:19:bf:88
 zone:  LSAN_BlueHBA1_1489_8aa
                21:01:00:e0:8b:8a:c7:6d;
                50:06:04:82:cc:19:c4:47
 zone:  LSAN_BlueHBA2_1489_9aa
                21:01:00:e0:8b:aa:c7:6d;
                 50:06:04:82:cc:19:c4:48
 zone:  LSAN_GreenHBA1_AllGreenStorage
                10:00:00:00:c9:39:a0:51;
                50:06:04:82:cc:19:c4:07;
                50:06:04:82:cc:19:c4:08;
                50:06:04:82:cc:19:c4:c7;
               50:06:04:82:cc:19:c4:c8

 zone:  LSAN_GreenHBA2_AllGreenStorage
                10:00:00:00:c9:39:a0:52;
                50:06:04:82:cc:19:c4:07;
                50:06:04:82:cc:19:c4:08;
                50:06:04:82:cc:19:c4:c7;
                   50:06:04:82:cc:19:c4:c8
```

h. Enter the **cfgSave** command to save the updated active zone set.

- To create LSAN zones in Connectrix M fabric, perform the following steps:

  Although the only requirement to create an LSAN between the Connectrix M and Connectrix B fabric is that the LSAN zones need to have the same devices zoned together and the zone names of the LSAN zones must start with the string "LSAN_", it is still a best practice to create identical zones with identical zone names in order to maximize ease of use.

  a. Open the **Zoning** dialog in Connectrix Manager by right-clicking the appropriate fabric topology and selecting the zoning menu item.

  b. Create a zone by clicking on the **New Zone** button under the Zones "Tree".

  c. Provide a descriptive name for the zone. In this case,we are going to zone "Red host HBA 1" and "Red Storage 1". We will enter "**LSAN_RedHBA1_1470_8aa**" and press **Enter**.

  d. Locate and then single-click on "**Red Host HBA 1**" **(WWPN 10000000c938e554)** in the **Potential Zone Members** list.

  e. Click the right-pointing arrow on the divider between the **Potential Members** list and the **Zones** list to add the HBA to the zone.

  f. Locate and then single-click on "**Red Storage 1**" **(WWPN 50060482cc19bf87)** in the**Potential Zone Members** list.

  g. Click the right-pointing arrow on the divider between the **Potential Members** list and the **Zones** list to add the storage port to the zone.

  h. Repeat Step b through Step g for all host and storage pairs in the environment.

  i. Select the active zone set (for example, Oct_31_06_1140) in the **Zone Set** panel. Select one of the LSAN zones you just created in the Zone panel, and click the right-pointing arrow on the divider to add the zones to this zone set.

  j. Add all of the new LSAN zones to the zone set. When completed, the zone set should be similar to what is shown below.

```
Zone set name = "Oct_31_06_1140"
       Zone name = "LSAN_RedHBA1_1470_8aa"
          Zone Member = "10000000c938e554"
          Zone Member = "50060482cc19bf87"

       Zone name = "LSAN_RedHBA2_1470_9aa"
          Zone Member = "10000000c938e555"
          Zone Member = "50060482cc19bf88"

       Zone name = "LSAN_BlueHBA1_1489_8aa"
          Zone Member = "210100e08b8ac76d"
          Zone Member = "50060482cc19c447"

       Zone name = "LSAN_BlueHBA2_1489_9aa"
          Zone Member = "210100e08baac76d"
          Zone Member = "50060482cc19c448"

       Zone name = "LSAN_GreenHBA1_AllGreenStorage"
          Zone Member = "10000000c939a051"
          Zone Member = "50060482cc19c407"
          Zone Member = "50060482cc19c408"
          Zone Member = "50060482cc19c4c7"
          Zone Member = "50060482cc19c4c8"

       Zone name = "LSAN_GreenHBA2_GreenStorage"
          Zone Member = "10000000c939a052"
          Zone Member = "50060482cc19c407"
          Zone Member = "50060482cc19c408"
          Zone Member = "50060482cc19c4c7"
          Zone Member = "50060482cc19c4c8"

Zone set name = "Oct_31_06_1140"
       Zone name = "RedHBA1_1470_8aa"
          Zone Member = "10000000c938e554"
          Zone Member = "50060482cc19bf87"

       Zone name = "RedHBA2_1470_9aa"
          Zone Member = "10000000c938e555"
          Zone Member = "50060482cc19bf88"

       Zone name = "BlueHBA1_1489_8aa"
          Zone Member = "210100e08b8ac76d"
          Zone Member = "50060482cc19c447"

       Zone name = "BlueHBA2_1489_9aa"
          Zone Member = "210100e08baac76d"
          Zone Member = "50060482cc19c448"

       Zone name = "GreenHBA1_AllGreenStorage"
          Zone Member = "10000000c939a051"
          Zone Member = "50060482cc19c407"
          Zone Member = "50060482cc19c408"
```

```
      Zone Member = "50060482cc19c4c7"
      Zone Member = "50060482cc19c4c8"

   Zone name = "GreenHBA2_GreenStorage"
      Zone Member = "10000000c939a052"
      Zone Member = "50060482cc19c407"
      Zone Member = "50060482cc19c408"
      Zone Member = "50060482cc19c4c7"
      Zone Member = "50060482cc19c4c8"
```

k. Click the **Activate** button on the right side of the screen, and then click **OK** at the bottom of the screen to active the zone set.

- To confirm LSAN zoning configuration, perform the following steps:

  a. Log in as an admin in the Connectrix B fabric and connect to the SAN Router.

  b. Enter the following commands to display information about the LSANs:

  – **lsanZoneShow -s** shows the LSAN.

```
switch:admin_06> lsanzoneshow -s

Fabric ID: 10 Zone Name: LSAN_RedHBA1_1470_8aa
      10:00:00:00:c9:38:e5:54 EXIST
      50:06:04:82:cc:19:bf:87 EXIST

Fabric ID: 10 Zone Name: LSAN_RedHBA2_1470_9aa
      10:00:00:00:c9:38:e5:55 EXIST
      50:06:04:82:cc:19:bf:88 EXIST

Fabric ID: 10 Zone Name: LSAN_BlueHBA1_1489_8aa
      21:01:00:e0:8b:8a:c7:6d EXIST
      50:06:04:82:cc:19:c4:47 EXIST

Fabric ID: 10 Zone Name: LSAN_BlueHBA2_1489_9aa
      21:01:00:e0:8b:aa:c7:6d EXIST
      50:06:04:82:cc:19:c4:48 EXIST

Fabric ID: 10 Zone Name: LSAN_GreenHBA1_AllGreenBlueStorage
      10:00:00:00:c9:39:a0:51 EXIST
      50:06:04:82:cc:19:c4:07 EXIST
      50:06:04:82:cc:19:c4:08 EXIST
      50:06:04:82:cc:19:c4:c7 EXIST
      50:06:04:82:cc:19:c4:c8 EXIST

Fabric ID: 10 Zone Name: LSAN_GreenHBA2_AllGreenStorage
      10:00:00:00:c9:39:a0:52 EXIST
```

```
50:06:04:82:cc:19:c4:07 EXIST
50:06:04:82:cc:19:c4:08 EXIST
50:06:04:82:cc:19:c4:c7 EXIST
50:06:04:82:cc:19:c4:c8 EXIST
```

– **fcrPhyDevShow** shows the physical devices in the LSAN.

```
mw116:admin> fcrphydevshow
 Device           WWN            Physical
 Exists                            PID
in Fabric
----------------------------------------
   10    10:00:00:00:c9:38:e5:54  670913
   10    10:00:00:00:c9:38:e5:55  670813
   10    10:00:00:00:c9:39:a0:51  610a13
   10    10:00:00:00:c9:39:a0:52  610613
   10    21:01:00:e0:8b:8a:c7:6d  670613
   10    21:01:00:e0:8b:aa:c7:6d  670713

… [rest of list ]

Total devices displayed: 18
```

Phase 4: Moving half of the host and storage ports from Connectrix M fabric to Connectrix B fabric

Topology

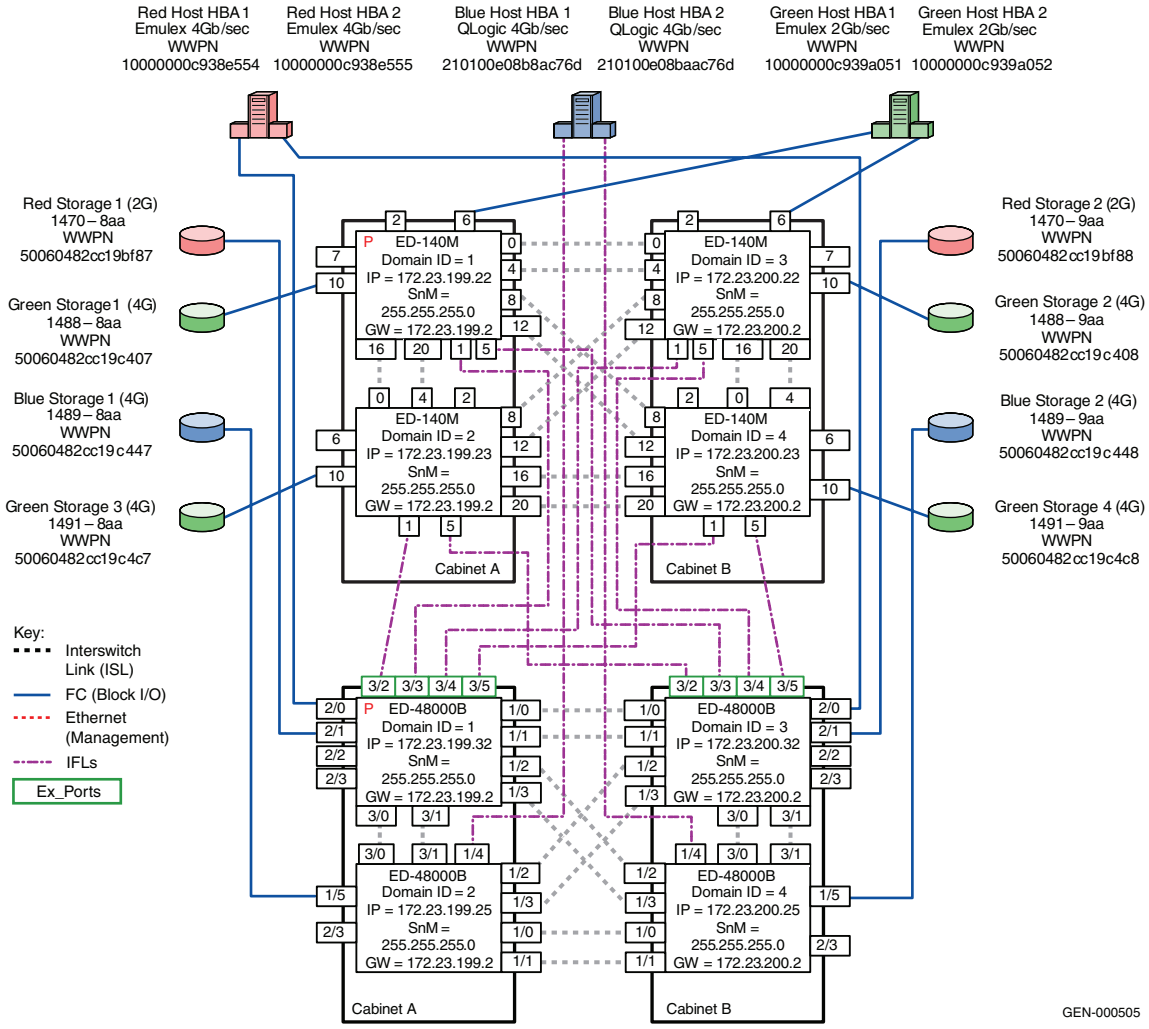


Figure 65 **Phase 4: Moving half of the host and storage ports**

Note: This phase is an intermediate phase prior to completely moving the core host and storage ports to the Connectrix B core. This phase ensures that the LSAN zones are appropriately configured, and that the traffic between the host and storage (which are zoned together via an LSAN zone) is not disrupted.

As shown in Figure 65 on page 347, some of the hosts and storage connections in the Connectrix M fabric are moved to the Connectrix B fabric. This requires connections to be physically removed from the Connectrix M director and connected to the Connectrix ED-48000B director in the the Connectrix B fabric. The LSAN zones created in phase 3 ensure connectivity is maintained even when the host connection is in one fabric and the storge connection is in the other fabric.

Note: You can move host and storage ports in any order you wish. There is no requirement to move one or the other first.

1. Move one device at a time. Start by moving the cable for "Red host HBA 1".
2. After it has been connected to the appropriate port in the Connectrix B fabric (Domain ID 1, 2/0), use the **fcrproxydevshow** command to verify the creation of a proxy device for the moved connection.

```
mw116:admin> fcrproxydevshow
  Proxy           WWN              Proxy      Device   Physical    State
 Created                            PID       Exists     PID
in Fabric                                    in Fabric
---------------------------------------------------------------------------
     1   10:00:00:00:c9:38:e5:54  01f001       10      670613   Imported

Total devices displayed: 1
```

3. Review the “Checkpoints” on page 331 and verify that this action did not affect the connectivity and functioning of the fabric.
4. Repeat this procedure as you move each connection for the red host and red storage, and the blue host and storage from the Connectrix M fabric to the Connectrix B fabric.

Below is the **fcrproxydevshow** command after moving the red storage 1 port to the Connectrix B fabric.

```
mw116:admin> fcrproxydevshow
  Proxy           WWN             Proxy      Device   Physical    State
 Created                           PID       Exists     PID
in Fabric                                   in Fabric
--------------------------------------------------------------------------
     1   10:00:00:00:c9:38:e5:54  01f001       10      670613   Imported
     1   50:06:04:82:cc:19:bf:87  01f002       10      671113   Imported

Total devices displayed: 2
```

5. When you have completed moving the red host / storage and blue host / storage connections to the Connectrix B fabric, use the **lsanshow** command to confirm the configuration. Note that the moved connections now show "IMPORTED" next to them instead of "EXIST".

```
switch:admin_06> lsanzoneshow -s

Fabric ID: 10 Zone Name: LSAN_RedHBA1_1470_8aa
      10:00:00:00:c9:38:e5:54 IMPORTED
      50:06:04:82:cc:19:bf:87 IMPORTED

Fabric ID: 10 Zone Name: LSAN_RedHBA2_1470_9aa
      10:00:00:00:c9:38:e5:55 IMPORTED
      50:06:04:82:cc:19:bf:88 IMPORTED

Fabric ID: 10 Zone Name: LSAN_BlueHBA1_1489_8aa
      21:01:00:e0:8b:8a:c7:6d IMPORTED
      50:06:04:82:cc:19:c4:47 IMPORTED

Fabric ID: 10 Zone Name: LSAN_BlueHBA2_1489_9aa
      21:01:00:e0:8b:aa:c7:6d IMPORTED
      50:06:04:82:cc:19:c4:48 IMPORTED

Fabric ID: 10 Zone Name: LSAN_GreenHBA1_AllGreenBlueStorage
      10:00:00:00:c9:39:a0:51 EXIST
      50:06:04:82:cc:19:c4:07 EXIST
      50:06:04:82:cc:19:c4:08 EXIST
      50:06:04:82:cc:19:c4:c7 EXIST
      50:06:04:82:cc:19:c4:c8 EXIST

Fabric ID: 10 Zone Name: LSAN_GreenHBA2_AllGreenStorage
      10:00:00:00:c9:39:a0:52 EXIST
      50:06:04:82:cc:19:c4:07 EXIST
      50:06:04:82:cc:19:c4:08 EXIST
      50:06:04:82:cc:19:c4:c7 EXIST
      50:06:04:82:cc:19:c4:c8 EXIST
```

Phase 5: Complete moving the host and storage ports in Connectrix M core to Connectrix B core

Topology

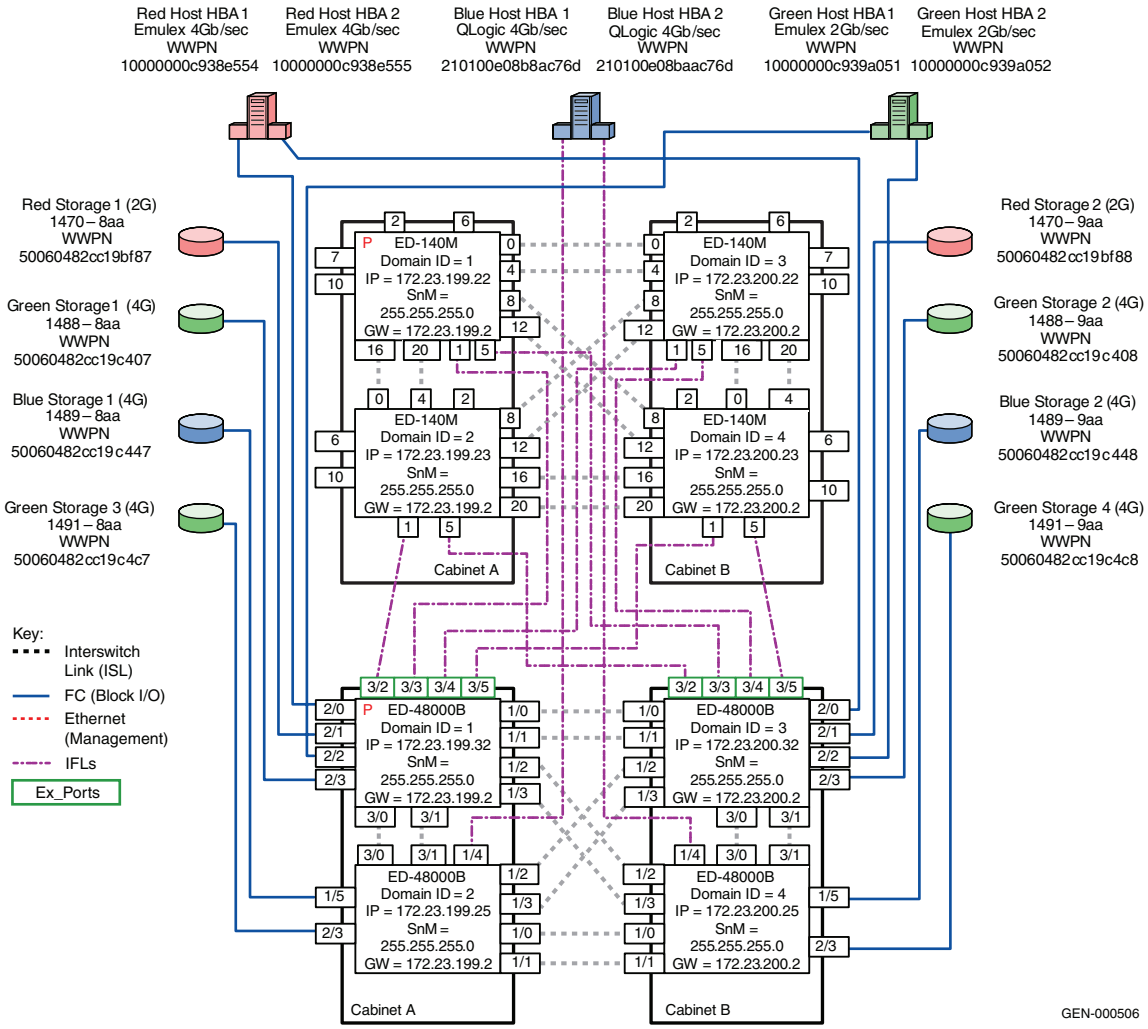


Figure 66 **Phase 5: Complete moving host and storage ports in the core**

Phase 5 is an extension of "Phase 4: Moving half of the host and storage ports from Connectrix M fabric to Connectrix B fabric" on page 347.

After validating that stable fabrics exists after the completion of Phase 4, execute Phase 5 by pulling the remaining host and storage connections in the Connectrix M fabric and transferring them to the Connectrix B fabric.

Review the "Checkpoints" on page 331.

Complete migration to Connectrix B fabric

At the end of case study 6, a full migration from a Connectrix M fabric to a Connectrix B fabric is completed. During the migration, SAN routing was deployed using interfabric links (IFL) between switches in both fabrics without disruption to application I/O.

The Connectrix M fabric can be removed after all devices have been moved to the Connectrix B fabric. This results in a fully operational Connectrix B fabric. At no time was there a fabric outage, nor any disruption to application IO, as devices were moved from the Connectrix M to the Connectrix B fabric.

The zones in the Connectrix B fabric were initially created with the special "LSAN_" prefix. When all devices are moved to the Connectrix B fabric, you may wish to rename the Connectrix B zones by removing the "LSAN_" prefix in the zone names. his prevents later confusion while managing the Connectrix B fabric since no LSAN zones are required.

If desired, the SAN router can be removed from the Connectrix B fabric as well, or it can be used to connect to any other Connectrix M or Connectrix B fabrics as desired.

Warnings or caveats

Please refer to EMC Knowledgebase solution emc149735 for all interop issues.

Cisco Inter VSAN Routing (IVR) in a heterogeneous environment

Cisco Inter VSAN Routing allows the Connectrix MDS series switches to consolidate multiple physical fabrics of different modes into a single physical fabric separated by VSANs while leveraging IVR to provide the necessary connectivity between the VSANs on a per-device basis. IVR provides additional benefits, including the ability to connect a core PID 0 Connectrix B switch to a core PID 1

Connectrix B switch. Normally, an outage would be required to change the core PID of one of the switches. IVR allows the Connectrix MDS switch to act as a conduit between two different legacy switch interop modes.

WARNING

RDI mode must be explictly enabled in IVR configurations which will use interop mode 1 to connect to non-Cisco FC switches. RDI mode is automatically enabled in configurations that use interop mode 2, 3 or 4. According to documentation from Cisco:
When third-party switches are involved in an IVR configuration, use the ivr virtual-fcdomain-add command to enable RDI mode, as third-party switches do not often query the remote name server unless it is present in the domain list. Issuing this command causes all virtual domains to appear in the domain list. This command should only be issued in the VSANs (border VSANs) in which the specified switch is present. Issuing the ivr virtual-fcdomain-add command is required only in edge or border VSANs. It is not required in MDS native VSANs.

Set up IVR with interop mode

To set up IVR with interop mode VSANs:

1. Establish ISL connectivity between the Connectrix MDS switch and the third-party switch. VSANs, ISLs, and third-party switches should be configured as specified in the appropriate sections of this document.
2. Enable IVR and configure IVR topologies, zones, and zone sets on the IVR-enabled Connectrix MDS switch. IVR zoning must be managed from the Connectrix MDS only.

   Note: In all cases, only manage IVR Zoning from the Connectrix MDS.

   To configure IVR and IVR zones using the IVR Zone Wizard in Cisco Fabric Manager:

   a. Click the **IVR Zone Wizard** icon in the **Zone** toolbar.
   b. Select the VSANs that will participate in IVR in the fabric.
   c. Select the end devices that will communicate over IVR.

Note: Fabric Manager displays an error message if all the switches participating in IVR do not have unique Domain IDs.These switches must be reconfigured before configuring IVR.

d. Enter the **VSAN ID** of the VSAN you want to use as the transit VSAN between the VSANs selected for the IVR zone, and then click **Next**.

e. Optionally, configure a unique **AFID** for switches in the fabric that have non-unique VSAN IDs in the **Select AFID** dialog box.

f. Verify the transit VSAN, or configure the transit VSAN, if Fabric Manager cannot find an appropriate transit VSAN.

g. Set the **IVR zone** and **IVR zone set** (now or later) using the **Edit Local Full Zone** database from the **Zone** menu.

   Please note:

   – A zone corresponding to each active IVR zone is automatically created in each edge VSAN specified in the active IVR zone. All pWWNs in the IVR zone are members of these zones in each VSAN.

   – Sometimes zone names beginning with prefix **IVRZ** and a zone set with name **nozoneset** appear in logical view. The zones with prefix **IVRZ** are IVR zones that get appended to regular active zones. The prefix **IVRZ** is appended to active IVR zones by the system. Similarly the zone set with name **nozonese** is an IVR active zone set created by the system if no active zone set is available for that VSAN, and if the **ivrZonesetActivateForce** flag is enabled on the switch.

h. Verify all steps that Fabric Manager uses to configure IVR in the fabric.

i. Click **Finish** if you want to enable the configured IVR topology and the associated IVR zones and IVR zone set.

j. The **Save Configuration** dialog appears. The configuration of the master switch to be copied to other IVR-enabled switches may be saved. Depending on what is required, click **Continue Activation**" or click **Cancel**.

k. Click **Finish**.

For more detailed information on setting up different types of IVR or IVR-NAT configurations please refer to the Cisco documentation at:

`http://www.cisco.com/en/US/products/ps5989/products_configuration_guide_chapter09186a0080662d37.html`

Warnings or caveats

Please refer to EMC Knowledgebase solution emc149735 for all interop issues.

- If the IBM Blade Center is involved in an IVR configuration, use the **ivr virtual-fcdomain-add** command since specified switches do not query the remote name server unless it is present in the domain list. Issuing this command causes all virtual domains to appear in the domain list. This command should only be issued in the VSANs (border VSANs) in which the specified switch is present. With Connectrix MDS SAN-OS Release 1.3(4a), issue the **ivr virtual-fcdomain-add** command only in edge or border VSANs. Do not issue the command in transit VSANs. If multiple IVR-enabled switches exist in an edge or border VSAN, the following error message may be displayed in the system messages: (`DEVICE_ON_WRONG_NATIVE_VSAN`). This message can be ignored.
- When routing between an interop mode 1 or 4 VSAN and a non-interop mode 1 or 4 VSAN, make sure that there are no domains in the native, or legacy switch interop modes 2 or 3 VSANs that are outside of the 97 to 127 range. Domains outside of this range cannot be created in the interop mode 1 VSAN. This is not a limitation of IVR-1, but a restriction imposed upon the configuration by interop mode 1. Since IVR-2 rewrites the Domain ID and FCID of the device in the non-interop mode 1 or 4 VSAN, this is not an issue.
- If an IVR zone set is active in a VSAN running in an interop mode, regular zone changes (such as zone set activation, reactivation or deactivation for zones that do not require IVR services) should always be initiated from the IVR border switch (the IVR-enabled Connectrix MDS 9000 switch) for that VSAN. This prevents disruption to ongoing traffic between IVR initiators and targets that are allowed by both the previous active zone set and the new active zone set. Otherwise, traffic disruption occurs, and a zone change can also fail because only the border switch has the IVR configuration for that VSAN.

- Domain conflicts can occur between switches in different VSANs and thus the Domain ID of a switch needs to be reset to avoid a Domain ID clash with a switch in a different VSAN (which is accessible with IVR).
- No zone names on a third-party switch participating in an interop environment with IVR can start with **IVRZ**. All IVR zoning must be managed through the Cisco switches.

Vendor-specific switch settings for interop

This section provides a more generic set of instructions or non-default settings that need to be configured on all switches belonging to a specific vendor type before they are introduced into a heterogeneous switched fabric. In addition, this section also provides the techniques, best practices, caveats, and unavailable features for a switch interoperability implementation.

Checklist

This checklist can be used to serve as a tool for completing all steps *before* actually merging fabrics.

- Verify that each switch has a unique Domain ID. When merging fabrics, ensure that there are no duplicate Domain IDs among all switches that will be part of the merged fabric.
- Verify that all switches have been set up to work in a supported interop mode.
- Verify that the E_D_TOV and R_A_TOV are set the same on all switches that will be part of the new fabric. (By default, they should all be the same; if necessary, refer to the appropriate user manual for information on how to set up operating parameters.)

  Note: Switches use different units to represent the same values; for example a value of 2000 on a Brocade switch or Cisco switch is the equivalent of 20 on a Brocade M-Series switch.

- Verify that the active zone set has been checked (with the respective switch fabric management tools) and does not contain illegal characters.
- If a switch is not operational and the zoning definition on that switch is not required, be sure to clear the zoning configuration on that switch.

- If a switch is operational and the zoning configuration on that switch is required, be sure to check that there are no duplicate active zone names. If there are duplicate zone names, rename one of the zones.
- Ensure that all switches are configured with WWN zoning.
- Ensure that all switches comply with proper zone naming.
- Back up the switch configuration by issuing the appropriate commands.

Some of the items on this checklist can be considered as recommended best practices to configure a stable interoperable environment with minimal disruption to existing data flow, if any.

Recommended best practices

- Set the Domain IDs rather than allowing the fabric to set them.
- Set the *core* switch as a *principal* switch. This reduces Class F traffic by ensuring that it goes directly from core to edge. For example, if an edge switch is the principal switch, build fabric traffic must go through the core to get from edge to edge.
- In a vendor switch migration from either Brocade or Brocade M-Series, leave Brocade or Brocade M-Series switches in their native modes and use Cisco's legacy modes, Interop-3 and Interop-4 (only if supported) for a non-disruptive and seamless migration.

  Note: At the time of this release, EMC supports Interop-4 using RPQ for migration purposes only. Please refer to the *EMC Support Matrix* for the most up-to-date support.

- Manage all IVR-based zoning in a heterogeneous environment with Cisco switches using the Cisco Fabric Manager. No zones or zone sets created on the Brocade or Brocade M-Series switch, or any other vendor switch in an IVR setup can start with an IVRZ prefix.
- Refer to the specific vendor interoperability caveats before setting up an interoperable environment. Use only the correct and supported version of switch firmware and Fabric Management software.

Configuration of non-default settings on supported vendor switches

The vendor-specific non-default settings that must be enabled on each of the supported vendor switches before introducing them into an interop environment follow:

Configure a Brocade switch

Telnet can be used to change the Domain ID or set the interop mode on a Brocade switch. These actions require disabling the switch; therefore, plan accordingly.

Principal switch setting

To configure a Brocade director switch in the core of a fabric as the *principal* switch, issue the **fabricprincipal** command on the switch CLI running FOS 4.x and higher. This setting becomes active on the next reboot, or after the build fabric (BF) event.

Configure the Domain ID

To set the preferred Domain ID:

1. Start a Telnet session.
2. Enter **switchdisable** to disable the switch.
3. Enter **configure**.
4. Answer **yes** to the **Fabric parameters** questions.
5. The first field, **Domain**, is the Domain ID, in the range 97 through 127, with these exceptions:
   - The lowest numbers are generally allocated to the Brocade M-Series switches if any, in the fabric.
   - Domain ID 104 (decimal 8) is reserved by HP.

     Starting with 97, allow one number for each Brocade M-Series switch. Then assign the next available number to the Brocade switch, and number the remaining Brocade switches in order (skipping 104).

     Accept all default values for the other parameters in this section, and answer **No** to all other sections until the switch enters the phase where it commits the new Domain ID to flash.

     The Domain ID becomes the value that the switch requests when attempting to join the fabric from an offline mode.
6. Enter **switchenable**.

Setting the interopmode

To set the interop mode:

1. Start a Telnet session

2. Enter **switchdisable**.
3. Enter **interopmode 1**to enable interopmode: **interopmode 1**.
4. Enter **reboot** or **fastboot**.

Configure a Cisco switch

To configure an MDS series switch the Domain ID and persistent FCIDs must be set, in addition to creating an interop VSAN. If you configure the switch as a core switch, you must set the switch priority.

To create an interop VSAN:

1. Log into Cisco Fabric Manager.
2. In the left pane, expand **All VSANs**.
3. Select **VSAN Attributes**.
4. In the right pane, select **Create row**.
5. Uncheck any MDS series switch that will not participate in the new VSAN.
6. In the **VSAN** field, assign a VSAN ID, 2 through 4093.
7. If desired, give the VSAN a name.
8. Leave **Loadbalancing** at default.
9. Set **InterOperValue** to **interop-1**.

Note: Interop-2 and Interop-3 are the legacy Brocade modes and Interop-4 is the legacy Brocade M-Series mode, and can be set depending on the requirement.
At the time of this release, EMC supports Interop-4 using RPQ for migration purposes only. Please refer to the *EMC Support Matrix* for the most up-to-date support.

---

10. Click **Create**. The new VSAN appears in the left pane.
11. Click **Close** in the VSAN creation window.

Configure the Domain ID and setting the switch priority

1. Log into the Cisco Fabric Manager.
2. In the left pane, expand the VSAN that has been designated for interop VSAN; for example, VSAN0002.
3. Select **Domain Manager**.

4. In the right pane, select the **Configuration** tab.
5. Set a Domain ID under the DomainID field, 97 through 127, (except 104, which is reserved for HP).
6. If configuring the switch as a core switch, the **Priority** field must be set to **1**. (If the switch will be an edge switch, the field must be left at the default setting of 128).
7. Under **Restart**, select **nondisruptive**.
8. Click **Apply Changes**.

Configure a Brocade M-Series switch

This is an offline operation and must be planned accordingly.

Configure the Domain ID and setting up the interop mode

Use the Connectrix Manager application to configure the Brocade M-Series switch.

1. Set the switch offline:
   a. Double-click the switch product icon in the **Connectrix Manager Products** or **Fabric** views.
   b. From the **Maintenance** drop-down menu, select **Set Online State**. The **Set Online State** window appears.
   c. Click **Set Offline**.
   d. To confirm, click **OK** in the **Warning** box.
2. On the **Connectrix Manager Hardware** view, select the appropriate option(s) from the **Configure** menu. For all switches except the ED-1032M, select the **Switch Parameters** option. (For the ED-1032M, you must select **Operating Parameters**.)
3. Configure the **Domain ID** on the next window.
4. To set the **Interopmode** and the **Switch Priority**, go back to the **Configure** menu, and select the **Operating Parameters**, **Fabric Parameters** option. This displays the **Configure Fabric Parameters** dialog box. The **Switch Priority** can be set as **Principal** or **Default** depending on the functionality. The director with the lower WWN becomes the principal switch, if the principal switch becomes unavailable.

You can set interopmode to **Open Fabric 1.0** if the switch is brought into an interop environment with Brocade or the Cisco in Interop-1.

When finished, click **Activate**.

5. Set the switch to **online**.

Unavailable features

Some of the features that are available prior to enabling the interoperability mode on FC switches are subsequently disabled. For example, when **interop** mode is enabled on Brocade switches, it disables domain / port-based zoning, Virtual Channel flow control, trunking, etc. This creates operational challenges for storage administrators who have to give up functionality in order to build heterogeneous fabrics.

Some features that are not available on the specific vendor switches when operating in Interop Fabric Mode include:

Brocade switches

- QuickLoop
- QuickLoop Fabric Assist
- Remote Switch
- Extended Fabrics
- Trunking
- Secure Fabric OS
- Alias Server
- Platform Service
- Virtual Channels
- FCIP

Cisco switches

- TE_ports (trunking expansion ports) and Port-Channels cannot be used to connect MDS to non-MDS switches. However, TE_ports and Port Channels can still be used to connect an MDS to other MDS switches even when in Cisco Fabric mode VSANs.

- The Quality of Service feature is intended to provide nodes with high bandwidth needs and greater access to the fabric resources. Quality of Service is applied end to end (host to storage), and can be implemented only if host and storage are attached to MDS models.

Brocade M-Series switches

- Show route (where targets or initiators are located on Brocade or Cisco switches)
- Show zone (where members are located on Brocade or Cisco switches)
- Fabric Binding
- Enterprise Fabric Mode
- SANtegrity (with the exception of Switch Binding)

Supported interoperability modes

Table 4 on page 362 shows EMC-tested modes. Read the table as follows:

The vendor switch operating in the respective mode listed in the two left columns can interoperate with the vendor switch in the top row operating in the mode specified by the intersection of that row and column.

Note: For the most current support information, refer to the *EMC Support Matrix*.

Table 4 EMC-tested modes

		Connectrix B/ Brocade	Connectrix MDS/ Cisco	Connectrix M/ Brocade M-Series	QLogic
Connectrix B/ Brocade	Interopmode0	Interopmode0	Interop-2 (if Brocade core PID=0) Interop-3 (if Brocade core PID=1)	N.A	N.A
	Interopmode1	Interopmode1	Interop-1	Open Fabric	Standard
Connectrix MDS/ Cisco	Native	**N.A**	Native	N.A	Standard
	Interop-1	Interopmode1	Interop-1	Open Fabric	N.A
	Interop-2	Interopmode0 (if Brocade core PID=0)	Interop-2	N.A	N.A
	Interop-3	Interopmode0 (if Brocade core PID=1)	Interop-3	N.A	N.A
	Interop-4	N.A	Interop-4	Brocade M-Series Fabric	N.A
Connectrix M/ Brocade M-Series	Connectrix M/Brocade M-Series Fabric	N.A	Interop-4	Connectrix M/ Brocade M-Series Fabric	Connectrix M/ Brocade M-Series Fabric
	Open Fabric	Interopmode1	Interop-1	Open Fabric	Standard
QLogic	Connectrix M/Brocade M-Series Fabric	N.A	Interop-4	Connectrix M/ Brocade M-Series Fabric	Connectrix M/ Brocade M-Series Fabric
	Standard	Interopmode1	Native	Open Fabric	Standard

Heterogeneous interoperability test information

Interoperability test plan

A list of tests that EMC E-Lab performs to validate vendor switch interoperability follows. These tests were performed for most of the applicable phases of the seven phase switch migration procedure discussed earlier in "Case studies" beginning on page 261.

Generic Switch features/performance tests

- I/O runs in an interop environment
- Switch reboots and power cycle
- Repeated Switch reboot
- Port type and Speed sensing (1/2/4 G performance based tests)
- NDCLA: Online code loads
- Firmware downgrade/upgrade backward compatible to two (2) firmware revisions
- Threshold alert-based tests
- Routing Table Management for switches
- Block and unblock ports

Tests with a cable puller

- Constant N_Port failures (Cable pull test)
- Constant E_port failures: Constant ISL breaks

Jammer and analyzer tests

- RSCN format/Domain format check
- N_Port login process into fabric
- Sending corrupted frames to switch
- Sending out of order frames to switch
- Loss of synchronization

Zoning

- Member addition to/removal from zone
- Mixed port and WWN zoning tests
- Creating/deleting zones
- Fabric merge/conflict tests for same zone set name, different zone names, same zone members

- Fabric merge/conflict tests for different zone tests, different zone names, different zone members
- Fabric merge/conflict tests for same zone set names, same zone names, different zone members

Fabric: Principal switch test

- Principal switch selection with priority settings
- Principal switch selection with no priority settings

Command line

- Performance monitoring
- Switch configuration

External Host based

- Host power cycle
- Host Reboot

Automated patch panel

Throughout this FC-SW based interoperability section, several topologies have been discussed that can be stepped through sequentially. Normally, the initial setup and incremental changes require a number of hours to configure. In the past, this setup time proved to be a significant barrier to performing frequent interoperability testing, and customer-requested migration processes at E-Lab.

To reduce setup time and help facilitate the migration testing, E-Lab recently purchased a Brocade M-Series UCS 2910 Automated Patch Panel. This enables E-Lab to approach an interop deployment in stages, similar to the way deployment is performed by an end user. E-Lab can also now take point-in-time snapshots of any configuration for future reference when confirming a fix or reproducing a problem.

More about the UCS 2910

The UCS 2910 is an Optical to Electrical to Optical (O-E-O) physical layer switch that handles all serial communication protocols up to 2 GB/s. (4 GB/s interfaces are in development.) It allows multiple users to control subsets of ports. The user interface allows for simple point-and-click connections for single connections as well as configuration backup and restore functions for multiple connections.

FICON connectivity

This section provides information on FICON topologies.

Overview

FICON (Fibre Connectivity) is an I/O channel designed to support low-latency, high-bandwidth connections between a mainframe and a storage controller. FICON builds on the technology of Fibre Channel, sharing the lower levels of Fibre Channel, including FC-0 (Physical Interface), FC-1 (8b/10b Encode/Decode), FC-2 (Framing and Signaling), and FC-3 (Common Services).

FICON is an FC-4 type. Its place in the Fibre Channel architecture is analogous to FCP, which is the FC-4 type used by all Open System host bus adapters and storage systems. FCP was designed to support SCSI. FICON was designed as a replacement for ESCON to support mainframe-attached CKD-formatted storage systems.

FICON is developed and maintained by the T11 Technical Committee of the International Committee for Information Technology Standards. Drafts and other technical documentation can be found at www.T11.org under FC-SB.

Table 5 lists some attributes of FICON and ESCON.

Table 5 FICON compared to ESCON

Attribute	FICON	ESCON
Link rate	212 MB/s	20 MB/s
Effective max data rate	200 MB/s	17 MB/s
Duplex	Full duplex	Half duplex
Type of switching	Packet switching with frame-by-frame routing using FSPF and Classes 2 and 3	Virtual Circuits (Class 1)
Multiple concurrent I/O	Yes	No
Maximum distance without droop	Depends on BB_Credit and link speed (typically 100 km at 100 MB/s).	9 km

The FICON technologies described in this Topology Guide are organized as follows:

- General FICON connectivity information — Covers the topic with switch vendor-neutral terminology and concepts. Refer to "General information" on page 366.
- Vendor-specific information — Covers the topic by switch vendor, applying the vendor's terminology, product names, support specifics, and unique considerations. Refer to:
  - "Connectrix B series" on page 371
  - "Connectrix M series" on page 374
  - "Connectrix MDS series" on page 379

Reference material

Refer to the following for more information:

- Connectrix documentation on Powerlink
- HCD/IOCP: http://www.ibm.com/servers/resourcelink
  Select **Library**, **z900**, **z800**, or **G5/G6**.
- IBM Redbooks, at http://www.ibm.com/redbooks:
  - *FICON Native Implementation and Reference Guide*, PN SG24-6266-01
  - *Getting Started with the Brocade M-Series Intrepid FICON Director*, PN SG24-6857-00
  - *Getting Started with the Inrange FC/9000 FICON Director*, PN SG24-6858-00

General information

Topology support

Topology support for FICON covers:

- Switch support for FICON-to-Symmetrix connectivity by switch vendor, model number, and recommended firmware levels
- Intermixing FICON and FCP on the same switch
- Intermixing FICON and SRDF on the same switch
- Cascading (multiswitch fabric support)
- Size of fabric
- Interoperability among switch vendors (ISLs between Connectrix B and Connectrix M switches)

- Compatibility between a FICON environment and EMC ControlCenter®

Each of these topology capabilities is determined on a per-vendor, per-model basis. Refer to the following for more information:

Switch series	Topology Guide reference	Vendor website
M series	"Connectrix M series" on page 374	http://www.Brocade M-Series.com
B series	"Connectrix B series" on page 371	http://www.Brocade.com
MDS series	"Connectrix MDS series" on page 379	http://www.Cisco.com

Zoning practices

The recommended zoning practice for FICON environments is to build a single FICON zone and include all FICON N_Ports, channels, and control units in that zone. FICON channels depend on State Change Notifications (SCNs) from the switch to perform error recovery.

A single zone is administratively simple, and insures that all FICON N_Ports receive their SCNs. Zoning based on World Wide Port Name (WWPN) is recommended, but port-based zoning is also supported. This practice varies from the Single Initiator zoning of Open Systems, because mainframe channels do not depend on the name server for device discovery. Instead, mainframes use the device address information in the IOCDS for discovery. ("IOCP considerations" on page 369 provides more information.)

If the switch or fabric includes Open Systems (intermix) or SRDF ports, these ports should be zoned per practices. For Open Systems, this is single-initiator zoning based on WWPN. For SRDF, place all SRDF ports into a single SRDF zone.

Note: Refer to "Use single initiator zoning." on page 38 for information on SRDF zoning.

Cascading

Cascading is a FICON topology where the channel and the control unit are on different switches, connected through an interswitch link (ISL). There are additional security and port addressing considerations in a cascaded switch environment.

Terminology

- **Entry switch** — A FICON switch that is connected to the processor's FICON channel(s) and to a cascaded switch. An entry switch can also be a cascaded switch.
- **Cascaded switch** — A FICON director that connects to a control unit (, for example, a storage array) and to an entry switch. A cascaded switch can also be an entry switch.
- **Switch ID** and **switch address** (1 byte) — Ways to address a FICON director. The Switch ID and Switch address describe fields used in the IOCDS. Both values should be set equal to the Domain ID of the switch.

  The IOCDS expects physical values in hex. In this context, the physical value refers to the value that would be seen in a Finisar / I-Tech trace. The logical value would be the value as seen from a management application. Since the switch vendors may use logical values (a physical value plus an offset) and display the Domain ID in decimal, care must be taken when transferring this information from the switch's management application to the IOCDS. Refer the specific switch vendor information:

  - “Connectrix B series” on page 371
  - “Connectrix M series” on page 374
  - “Connectrix MDS series” on page 379

  The switch ID is assigned by the customer, and must be unique within the scope of the IOCP (Input / Output Configuration Program) or HCD (Hardware Configuration Definition). The switch ID is technically an arbitrary unique number given to identify the switch. It is highly recommended that the switch ID be set equal to the Domain ID.

  The link statement in the IOCDS is where the actual switch Domain ID and port address are defined. For example, the link statement **LINK= (6104,6104)** defines a connection to port 4 on Domain ID 1.

  The switch address is the Domain ID of the switch. These addresses can be customized to a preplanned value. They must be unique within a fabric.
- **Port address** — Address (with a one-byte value) of the physical FICON director port.

- **Insistent Domain ID** — A switch using the Insistent Domain ID feature ensures that it will join a fabric if (and only if) its administratively assigned Domain ID is granted during the fabric initialization procedure.

Notes on cascading

Switch addresses for FICON directors must be unique across all of the processor's CHPIDs (Channel Path Identifiers).

With SYSPLEX installations, the switch address must be unique across the SYSPLEX complex.

Cascading requires the use of two-byte link addresses that consist of a high-order byte defining the switch address (Domain ID) and a low-order byte that defines the FICON director port address.

Two-byte link addresses require that the FICON director have the Fabric Binding and Insistent Domain ID features installed. The channel checks this during initialization by sending the Security Attributes Extended Link Services query to the switch. In the response from the switch, the channel checks that Fabric Binding Enforcement and Insistent Domain ID bits are set. If they are not set, the link does not complete initialization.

IOCP considerations

Figure 67 shows an example of cascaded IOCP.

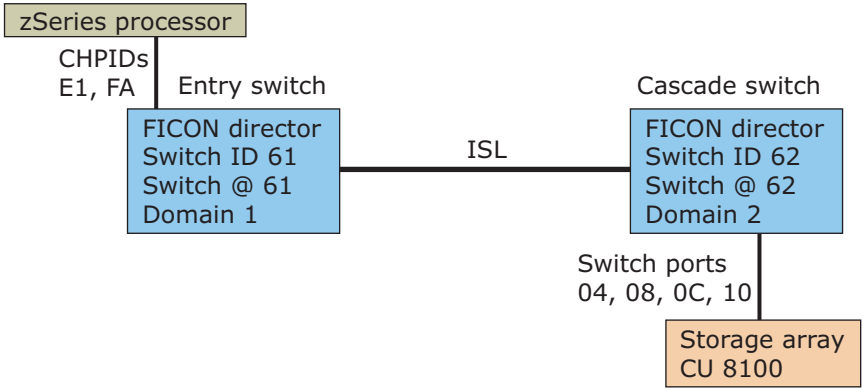


Figure 67 Cascaded IOCP

A cascaded FICON IOCP based on the figure would look like this:

Type keyword
CHPID will operate in FICON native mode, required for cascading

Switch keyword
Logical switch number
Channel entry switch.id, required for FICON (FC mode)

```
CHPID PATH=(E1),SHARED,TYPE=FC,SWITCH=61

CHPID PATH=(FA),SHARED,TYPE=FC,SWITCH=61

CNTLUNIT CUNUMBR=8100,PATH=(E1,FA),UNIT=2032,
    UNITADD=((00,1)),LINK=(6204,6208,620C,6210
```

Link keyword
Two bytes
Destination port address, one for each switched path
Switch address in high byte plus port address in low byte

`IODEVICE` (Same as for non FICON)

Once a two-byte link address has been specified for a channel path:

- It is recommended that you specify two-byte link addresses for all paths from the same CEC that have the same entry switch.
  - This allows IOCP to perform better checking of the switch configuration.
  - HCD requires additional information and performs a different checking method; therefore, the above is an IOCP recommendation only.
- All link addresses defined for access from the same channel path *must* be two-byte link addresses.

FICON and EMC ControlCenter

The EMC ControlCenter switch management functions (including zoning) are qualified on a switch vendor basis. Refer to the specific switch vendor information for the latest support.

Any time there is more than one management control point, care must be taken to insure that administrative conflicts do not arise. With ControlCenter there can be three separate and potentially conflicting control points: ControlCenter, Connectrix Manager, and System Automation OS/390. It is largely the administrator's responsibility to insure that conflicting controls are avoided.

As a result, having a single administrator in an intermix environment for all Open Systems and FICON switch management is the preferred practice.

CUP (fabric management server)

Control Unit Port (CUP) is an inband communications protocol that allows an OS/390 to manage a switch with the same level of control and security that S/390 has on an ESCON switch. Control functions include blocking and unblocking ports, as well as error reporting and monitoring.

CUP is also required for other mainframe management functions, such as:

- HCD — Port activation/deactivation
- HCM — Performance display
- SMF — Performance statistics gathering for RMF
- ZOS — Error handling, problem determination
- DCM — Dynamic CHPID management

The services provided by CUP are used by IBM management software System Automation for OS/390. In the case of the Connectrix M series product line, this management system depends on storage services provided by the Connectrix service processor for maintaining CUP-controlled configuration files.

CUP is not a requirement for mainframe/storage FICON connectivity. Typically, large mainframe data centers utilize its services.

Connectrix B series

Supported products

These B series products support FICON connectivity:

- DS-32B2
- DS-4100B
- ED-12000B
- ED-24000B
- ED-48000B
- DS-4900B
- PB-48K-18i blade (FC-IP)
- Silkworm 7500 (FC-IP)

Configuring

Topology support

- Single switch is supported.
- Cascading is supported with with FOS 5.2.1 or higher
  - Cascading is *not* supported with ED-12000B.
- Intermixing FICON and FCP on the same switch is supported.

  In a FCP / FICON intermix environment, as long as the FICON N_Ports, Channel, and Control Unit reside on the same domain, and the IOCDS uses single byte addressing, FICON I / O in a cascaded Open Systems Brocade fabric is supported.
- Intermixing FICON and SRDF on the same switch is supported.
- B series / M series interoperability is not supported in a FICON environment.

Recommended FICON environment configuration settings:

- Configure ports that are connected to 1 GB / s channels for fixed 1 GB / s speed. Otherwise, when using fixed 1 GB / s channels (both G5 and FICON Express), the FICON host might generate erroneous link incidents when the channels are coming online. These link incidents will result in a call home. Other than the generated link incident, the channel will come online and function normally.
- Enable in-order delivery (**iodset** command).

OCP considerations

Switch ID Definition — No offsets on switch ID or port address, but every value must be in hex for the mainframe.

CUP support

CUP is a separately licensed inband management service installed on each switch. Table 6 lists the EMC CUP models.

Table 6 EMC CUP models

CUP model	Applicable switch
DS32B-CU	DS-32B2
DS4100-CU	DS-4100B
SW-12000B-CU	ED-12000B
SW-24000B-CU	ED-24000B
SW-48000-CU	ED-48000B

You can manage CUP using Web Tools or Fabric Manager. Limited support for CUP is provided through the Fabric OS CLI.

CUP provides the following advantages:

- Single point of control and monitoring for channel, director (switch), and control unit.
- Automated tools on the mainframe can take advantage of the statistics to move channels where they are needed. This cannot be done from the switch alone.
- Seamless integration into existing management tools that are also used to manage ESCON directors (switches). This makes migration from ESCON to FICON smoother.

You can monitor the FICON director (switch) using CUP to obtain the following port statistics:

- Number of words transmitted
- Number of words received
- Frame Pacing Time (the number of 2.5 ms units that frame transmission is blocked due to zero credit).

Refer to the IBM document *FICON Director Programming Interface with Cascading Support*.

Switch node identifier

You can find switch node information such as the serial number and manufacturer name. This information is the same as the Switch Node ID in the RNID ELS:

- Configuration file information — Provides a list of configuration files on the switch. You can also obtain the actual file content, including port address name and port connectivity.
- History summary (director history buffer) — The history buffer logs each change in status or configuration of the ports. You can retrieve the history buffer using CUP.
- Switch configuration data — Provides switch configuration data such as time-out values and number of ports per card.

You can find information on CUP functions and commands in the IBM-proprietary document *FICON Director Programming Interface With Cascading Support*.

EMC documentation

Refer to these documents for more information on FICON connectivity:

- *EMC Connectrix B Series Version 4.4 Features Guide* (P / N 300-001-702)
- *EMC Connectrix B Series Fabric OS Version 4.4 Procedures Guide* (P / N 300-001-708)

Connectrix M series

Supported products

Supported M series products are:

- ED-10000M (described under "Connectrix ED-10000M" on page 1096)
- ED-140M (described under "Connectrix ED-140M" on page 1108)
- ED-64M (described under "Connectrix ED-64M" on page 1122)
- ED-1032 (described under "Connectrix ED-1032" on page 1124)
- DS-32M2 (described under "Connectrix DS-32M, DS-32M2" on page 1143)
- DS-4700M (described under "Connectrix DS-4700M" on page 1127)

Configuring

Topology support

Note the following:

- Intermixing FICON and FCP on a Brocade M-Series fabric is supported. It is recommended that all E / OS-based products use firmware version 6.01 or higher to take advantage of improvements in availability.
- Intermixing FICON and SRDF on the same fabric is supported.
- FICON is supported in homogeneous fabrics; all switches in the fabric must be from the same switch vendor.
- The EMC ControlCenter switch management functions (including zoning) can be used to manage a switch in a FICON or FICON / FCP intermix environment. ControlCenter 5.1.2 or higher is required.

Note: Any time there is more than one management control point, care must be taken to insure that administrative conflicts do not arise. With EMC ControlCenter, there may be three separate and potentially conflicting control points: ControlCenter, Connectrix Manager, and System Automation OS/390. It is largely the administrator's responsibility to insure that conflicting controls are avoided. As a result, having a single administrator in an intermix environment for all Open Systems and FICON switch management is the preferred practice.

IOCP configuration

Note the following:

- Brocade M-Series switch ID definition — The switch ID is the Domain ID plus x'60'.

  Example: Domain ID 1 would be switch ID x'61'.

- Brocade M-Series port address:
  - On the ED-10000M and DS-4700M, the physical port mapping maps directly to the logical port mapping.
  - On any other supported M series switch, the logical port address is the physical port address plus 4 in hex (if there has been no port swapping). For these switches, you must use the logical port address, not the physical port address.

    The physical port address refers to the address associated with the physical position of the port card in the chassis, and the specific port on the port card.

    The logical port address can be swapped using the **Swap Ports** function on the Element Manager's **Maintenance** pull-down menu. By using the FICON setting in the Management Style option in the Element Manager's **Product** pull-down menu, the translation from physical (**Port #**) to logical (**Addr**) can be determined. Use the **Addr** value found in the Port List tab in the IOCDS.

    Example: Physical port 6 would be logical port x'A', assuming there has been no port swapping.

Connectrix Manager: Fabric Manager

Note the following:

- **Zoning** — Refer to "Zoning practices" on page 367.

  Prior to the support of intermix, the State Change Notification distribution in a Connectrix M Series FICON environment was accomplished by enabling default zoning and having no active zone set. This practice is supported in FICON-only environments, but is no longer recommended.

- **Default zoning** — The recommended setting *disabled*. When new N_Ports are added to the fabric, disabling the default zone prevents these new N_Ports from becoming aware of other N_Ports. This function is controlled in the **Zoning** dialog of the Connectrix Manager Fabric Manager.

- **Enterprise Fabric Mode** — Must be enabled in order to support cascading. In order to enable this mode, every switch in the fabric must have the SANtegrity feature enabled (using the Connectrix Fabric Manager **Configure** menu).

  Enabling this mode automatically enables the following parameters. These parameters cannot be disabled unless the switch is offline. Disabling any of these parameters also disables Enterprise Fabric Mode.

  - **Rerouting Delay**
  - **Domain RSCNs** (Register for State Change Notifications)
  - **Insistent Domain ID**

Switch Element Manager-level FICON configuration considerations

Note the following:

- **Open Fabric Mode** — This is the recommended interoperability mode for both FICON-only and FICON/FCP intermix fabrics.

- **Suppress Zoning RSCNs on Zone Set Activations** — FICON Channels do not use Fabric Format RSCNs. Therefore, this setting may be enabled or disabled.

  This option has been superseded in E/OS 7.01 and above. E/OS 7.01 has an RSCN filtering function, sometimes called *Zone Flexpars*. Refer to the *EMC Connectrix E/OS 7.01 Release Notes* for details.

- **Preferred Pathing** and **PDCM on E_Ports** — These are new switch features offered in Connectrix M series firmware 6.01 that offer route control for segregating FICON and FCP traffic over ISLs on a per-N_Port basis. In order to use the PDCM controls, the Element Manager Management Style must be set for FICON.
- **Unified Management Mode** — In E/OS 06.xx.xx and higher, the separate OS/390 and Open Systems modes of operation for FICON and Open Systems networks have been consolidated.

  In Connectrix Manager, this feature changes the function of the FICON Management Style (previously OS/390 operating mode) and Open Systems Management Style (previously Open Systems operating mode) options. These options are available through the **Management Style** option on the **Product** menu.

  If E/OS 06.xx.xx or higher is installed on the switch and you are using Connectrix Manager 08.xx.xx, you do not have to set the switch to the offline state before changing the Management Style. Previously you had to take the product offline before changing management styles or operating modes.

  You can now switch to Open Systems Management style even if the optional FICON Management Server feature is installed. Previously, you could only use FICON Management Style if this feature was installed.

  You can now enable the Open Systems Management Server (OSMS) and FICON Management Server features simultaneously. Previously, you could enable only one type of management server feature at a time for a switch.

Distance options

Distance options are:

- DSED-LWKT20 — Single-mode longwave optic, 20 km distance
- DSED-LWKT35 — Single-mode longwave optic, 35 km distance

SANtegrity

SANtegrity provides a layer of security to network ports, switches, and fabrics through authorization, authentication, and secure management techniques in Open Systems or mainframe environments. SANtegrity ensures that Fibre Channel traffic is not redirected by unauthorized access.

In mainframe FICON environments, SANtegrity enables FICON cascading because it provides a layer of security to the fabric required by the channel. SANtegrity is separately licensed and is a

requirement on every switch in a fabric to enable Enterprise Fabric Mode (high integrity mode).

The following are SANtegrity features:

- **Fabric Binding** — Prevents unauthorized switches from joining a fabric.
- **Switch Binding** — Allows users to restrict connections to an individual switch, for either E_Port connections, F_Port connections, or all connections. Switch binding can be enabled without enabling Enterprise Fabric Mode.
- **Insistent Domain ID** — Prohibits the use of dynamic Domain IDs to ensure that predictable Domain IDs are being enforced within the fabric. This is important for fabric binding because changing the Domain ID on a switch can cause segmentation resulting from the exclusion of a switch from the fabric.

  Insistent Domain IDs are critical to FICON addressing, and are set automatically set when Enterprise Fabric mode or Fabric Binding is enabled.
- **Rerouting Delay** — Ensures that frames are delivered through the fabric in order to reach their destination.
- **Domain RSCNs** — Allows the switch to transmit Domain RSCNs between end devices in a fabric to provide additional connection information to host bus adapters and storage devices.

Address swapping

The logical address of a physical port is specified in the HCD on the mainframe. Typically the HCD must be changed when moving a FICON link between switch ports. However, a Port address Swapping feature in the Connectrix Manager Element Manager can be used to avoid changing the HCD. The feature swaps two switch port addresses. Refer to the switch's user guide for more details.

CUP

In the past, EMC sold Brocade M-Series-branded CUP switches as Connectrix models ED-140MCUP and ED-64MCUP. These products included the FICON Management Server license, which activates the CUP functions. This option has been replaced with a new option, in which a software license key enabling CUP support on EMC-branded Brocade M-Series switches can be ordered through CQS.

The following CQS model numbers apply to field upgrades and new orders:

Switch model	License key model number for CUP support
ED-64M	ED-CUPSW-64M
ED-140M	ED-CUPSW-140M
ED-10000M	ED-10K-CUP [a]

a. When enabled on the ED-10000M, the FICON Management Server (FMS) consumes port addresses 0xFE and 0xFF, which are ports 254 and 255, if there has been no swapping of ports. These ports always transmit OLS (offline sequence).

As a result, if 10 GB paddles are being installed in an environment with CUP enabled, placing the 10 GB paddle in the upper-left corner of the chassis is a good practice since the 10 GB paddle does not use addresses 0xFE or 0xFF.

The FMS limits the ED-10000M to a single partition. This limitation is characteristic of E/OSn 6.x and this restriction is expected to be lifted in future firmware.

You can find more information on CUP in *Getting Started with the Brocade M-Series Intrepid FICON Director (*P / N SG24-6857-00), available on the IBM website:

`http://www.ibm.com/redbooks`

Connectrix MDS series

Supported products

These MDS series products support FICON connectivity:

- MDS 9513
- MDS 9509
- MDS 9506
- MDS 9216
- MDS 9216A
- MDS 9216i

Requirements

Requires the purchase and installation of the MDS mainframe license package. Note that MDS 9500 and 9200 series have different model numbers for licenses.

Configuring

Topology support

Note the following:

- You must create a FICON VSAN.
- A multiswitch environment is supported.

- Cascading is supported with one ISL hop.
- Intermixing FICON and FCP on the same switch is supported using separate VSANs.
  Use standard or interop mode VSANs for FCP traffic, and FICON VSANs for FICON traffic.
- Intermixing FICON, SRDF, MirrorView™, Open Replicator, and SAN Copy™ on the same switch is supported.
- Connectrix M series/B series interoperability is supported if those switches exist in another VSAN set for *interop-1* mode.
- FICON over IP is not supported.
- FC Write Acceleration is not supported for FICON.
- TE_Ports/EISLs may support FICON and FCP traffic.

  If dedicated bandwidth for FICON is required, create multiple ISLs dedicated to the FICON VSAN.
- Port Channels for FICON or mixed FCP/FICON traffic is supported.

  FICON ports must be bound to a Port Channel.

Recommended FICON environment configuration settings:

- Configure ports that are connected to 1 GB/s channels for fixed 1 GB/s speed. Otherwise, when using fixed 1 GB/s channels (both G5 and FICON Express), the FICON host might generate erroneous link incidents when the channels are coming online. These link incidents will result in a call home. Other than the generated link incident, the channel will come online and function normally.
- Enable in-order delivery.

OCP considerations

Switch ID Definition — No offsets on switch ID or port address, but every value must be in hex for the mainframe.

CUP support

Control Unit Port (CUP) is supported by switches and directors in the Cisco MDS 9000 Family. The CUP function allows the mainframe to manage the Cisco MDS switches.

CUP comes with the Mainframe license package required for FICON support.

CUP provides the following advantages:

- Single point of control and monitoring for channel, director (switch), and control unit.
- Automated tools on the mainframe can take advantage of the statistics to move channels where they are needed. This cannot be done from the switch alone.
- Seamless integration into existing management tools that are also used to manage ESCON directors (switches). This makes migration from ESCON to FICON smoother.

You can monitor the FICON director (switch) using CUP to obtain the following port statistics:

- Number of words transmitted
- Number of words received
- Frame Pacing Time (the number of 2.5 ms units that frame transmission is blocked due to zero credit).

The IBM document *FICON Director Programming Interface with Cascading Support* contains more information on CUP.

FICON configuration files

You can save up to 16 FICON configuration files on each FICON-enabled VSAN (in persistent storage).

FICON configuration files contain the following configuration for each implemented port address:

- Block
- Prohibit mask
- Port address name

Switch node identifier

You can find switch node information, such as the serial number and manufacturer name. This information is the same as the switch node ID in the RNID ELS:

- Configuration file information — You can get a list of configuration files on the switch. You can also obtain the actual file content, including port address name and port connectivity.
- History summary (director history buffer) — The history buffer logs each change in status or configuration of the ports. You can retrieve the history buffer using CUP.
- Switch configuration data — Provides switch configuration data as time-out values and number of ports per card.

You can find more information on CUP functions and commands in the IBM-proprietary document *FICON Director Programming Interface With Cascading Support*.

FICON port numbering

Table 7 represents the logical port number for the MDS 9216, MDS 9506, and MDS 9509.

Table 7 FICON port numbering in the MDS 9000 family

Product	Slot number	Implemented port allocation		Unimplemented ports	Notes
		To ports	To portchannel/FCIP		
MDS 9200 Series	Slot 1	0 through 31	64 through 89	90 through 253 and port 255	Similar to a switching module
	Slot 2	32 through 63			The first 16 port numbers in a 16-port module are used and the rest remain unused.
MDS 9506 Director	Slot 1	0 through 31	128 through 153	154 through 253 and port 255	The first 16 port numbers in a 16-port module are used and the rest remain unused.
	Slot 2	32 through 63			
	Slot 3	64 through 95			
	Slot 4	96 through 127			
	Slot 5	None			Supervisor modules are not allocated port numbers.
	Slot 6	None			
MDS 9509 Director	Slot 1	0 through 31	224 through 249	250 through 253 and port 255	The first 16 port numbers in a 16-port module are used and the rest remain unused.
	Slot 2	32 through 63			
	Slot 3	64 through 95			
	Slot 4	96 through 127			
	Slot 5	None			Supervisor modules are not allocated port numbers.
	Slot 6	None			
	Slot 7	128 through 159			The first 16 port numbers in a 16-port module are used and the rest remain unused.
	Slot 8	160 through 191			
	Slot 9	192 through 223			

References

These documents and resources contain for more information on FICON connectivity:

- E-Lab Navigator, for supported MDS firmware for FICON usage
- Connectrix MDS Release Notes, for specific information related to new firmware
- *Cisco MDS 9000 Family Configuration Guide*, available at http://www.Cisco.com
- *Implementing the Cisco MDS 9000 in an Intermix FCP, FCIP, and FICON Environment* (part number SG24-6397-00), available at http://www.ibm.com

2

iSCSI SAN Topologies

This chapter provides the following information on iSCSI SAN topologies.

- TCP/IP technology overview 386
- iSCSI technology overview 405
- iSCSI solutions 410
- Summary 429

TCP/IP technology overview

The Transmission Control Protocol (TCP) is a connection-oriented transport protocol that guarantees reliable in-order delivery of a stream of bytes between the endpoints of a connection. TCP achieves this by assigning each byte of data a unique sequence number by maintaining timers, acknowledging received data through the use of acknowledgements (ACKs), and retransmitting data if necessary. Data can be transferred after a connection is established between the endpoints. The data stream that passes across the connection is considered a single sequence of eight-bit bytes, each of which is given a sequence number. This section contains the following information:

- "IPv6," next
- "TCP terminology" on page 397
- "TCP error recovery" on page 400
- "Network congestion" on page 403
- "IP Security" on page 404

IPv6

Internet Protocol version 6 (IPv6) is a network layer protocol for packet-switched internets. It is designated as the successor of IPv4.

Note: For the most up-to-date support information, please refer to the *EMC Support Matrix*>PDF and Guides> Miscellaneous> Internet Protocol.

Note: The information in this section was acquired from Wikipedia.org, August 2007, which provides further details on many of these topics.

The main improvement of IPv6 is the increase in the number of addresses available for networked devices. IPv4 supports 2^{32} (about 4.3 billion) addresses. In comparison, IPv6 supports 21^{28} (about 34×1037) addresses, or approximately 5×10^{28} addresses for each of roughly 6.5 billion people. However, that's not the intention of the designers.

The extended address length simplifies operational considerations, including dynamic address assignment and router decision-making. It also avoids many complex workarounds that were necessary in IPv4, such as Classless Inter-Domain Routing (CIDR). Its simplified

packet header format improves the efficiency of forwarding in routers. More information on this topic is provided in "Larger address space" on page 387 and "Addressing" on page 389.

Features of IPv6

To a great extent, IPv6 is a conservative extension of IPv4. Most transport- and application-layer protocols need little or no change to work over IPv6. The few exceptions are applications protocols that embed network-layer addresses (such as FTP or NTPv3). Applications, however, usually need small changes and a recompile in order to run over IPv6.

The following features of IPv6 will be further discussed in this section:

- "Larger address space," next
- "Stateless autoconfiguration of hosts" on page 388
- "Multicast" on page 388
- "Jumbograms" on page 388
- "Network-layer security" on page 388
- "Mobility" on page 389

Larger address space

The main feature of IPv6 is the larger address space: 128 bits long (versus 32 bits in IPv4). The larger address space avoids the potential exhaustion of the IPv4 address space without the need for network address translation (NAT) and other devices that break the end-to-end nature of Internet traffic.

Note: In rare cases, NAT may still be necessary, but it will be difficult in IPv6 so should be avoided whenever possible.

It also makes administration of medium and large networks simpler, by avoiding the need for complex subnetting schemes. Ideally, subnetting will revert to its original purpose of logical segmentation of an IP network for optimal routing and access.

There are a few drawbacks to larger addresses. For instance, in regions where bandwidth is limited, IPv6 carries some bandwidth overhead over IPv4. However, header compression can sometimes be used to alleviate this problem. IPv6 addresses are also harder to memorize than IPv4 addresses, which are, in turn, harder to memorize than Domain Name System (DNS) names. DNS protocols have been modified to support IPv6 as well as IPv4.

For more information, refer to "Addressing" on page 389.

Stateless autoconfiguration of hosts

IPv6 hosts can be automatically configured when connected to a routed IPv6 network. When first connected to a network, a host sends a link-local (automatic configuration of IP addresses) multicast (broadcast) request for its configuration parameters. If configured suitably, routers respond to such a request with a router advertisement packet that contains network-layer configuration parameters.

If IPv6 autoconfiguration is not suitable, a host can use stateful autoconfiguration (DHCPv6) or be configured manually.

Note: Stateless autoconfiguration is suitable only for hosts. Routers must be configured manually or by other means.

Multicast

Network infrastructures, in most environments, are not configured to route multicast. The link-scoped aspect of multicast (that is, on a single subnet) will work but the site-scope, organization-scope, and global-scope multicast will not be routed.

IPv6 does not have a link-local broadcast facility. The same effect can be achieved by multicasting to the all-hosts group (FF02::1).

The m6bone is catering for deployment of a global IPv6 multicast network.

Jumbograms

IPv6 has optional support for packets over the IPv4 limit of 64 KB when used between capable communication partners and on communication links with a maximum transmission unit larger than 65,576 octets. These are referred to as *jumbograms* and can be as large as 4 GB. The use of jumbograms may improve performance over high-MTU (Maximum Transmission Unit) networks.

An optional feature of IPv6, the *jumbo payload* option, allows the exchange of packets larger than this size between cooperating hosts.

Network-layer security

IPsec (IP security), the protocol for IP network-layer encryption and authentication, is an integral part of the base protocol suite in IPv6. In IPv4, this is optional (although usually implemented). IPsec is not widely deployed except for securing traffic between IPv6 Border Gateway Protocol (BGP) routers (the core routing protocol of the Internet).

Mobility

Mobile IPv6 (MIPv6) avoids triangular routing and is as efficient as normal IPv6. This advantage is mostly hypothetical, since neither MIP nor MIPv6 are widely deployed.

Deployment status

As of December 2005, IPv6 accounts for only a small percentage of the live addresses in the Internet, which is still dominated by IPv4. Many of the features of IPv6 have been ported to IPv4, with the exception of stateless autoconfiguration, more flexible addressing, and Secure Neighbor Discovery (SEND).

IPv6 deployment is primarily driven by IPv4 address space exhaustion, which has been slowed by the introduction of classless inter-domain routing (CIDR) and the extensive use of network address translation (NAT).

Estimates as to when the pool of available IPv4 addresses will be exhausted vary widely, ranging from around 2011 (2005 report by Cisco Systems) to Paul Wilson's (director of APNIC) prediction of 2023.

To prepare for the inevitable, a number of governments are starting to require support for IPv6 in new equipment. The U.S. Government, for example, has specified that the network backbones of all federal agencies must deploy IPv6 by 2008 and bought 247 billion IPv6 addresses to begin the deployment. The People's Republic of China has a 5-year plan for deployment of IPv6, called the "China Next Generation Internet".

Addressing

128-bit length

The primary change from IPv4 to IPv6, as discussed in "Larger address space" on page 387, is the length of network addresses. IPv6 addresses are 128 bits long (as defined by RFC 4291), compared to IPv4 addresses, which are 32 bits. IPv6 has enough room for 3.4×10^{38} unique addresses, while the IPv4 address space contains about 4 billion addresses.

IPv6 addresses are typically composed of two logical parts: a 64-bit (sub-)network prefix and a 64-bit host part, which is either automatically generated from the interface's MAC (Media Access Control) address or assigned sequentially. Globally unique MAC addresses offer an opportunity to track user equipment (and thus users) across time and IPv6 address changes. In order to restore some of the anonymity existing in the IPv4, RFC 3041 was developed to reduce the prospect of user identity being permanently tied to an

IPv6 address. RFC 3041 specifies a mechanism by which time-varying random bit strings can be used as interface circuit identifiers, replacing unchanging and traceable MAC addresses.

Notation

IPv6 addresses are normally written as eight groups of four hexadecimal digits. For example, the folllowing is a valid IPv6 address:

2001:0db8:85a3:08d3:1319:8a2e:0370:7334

If one or more four-digit group(s) is 0000, the zeros may be omitted and replaced with two colons(::). For example, 2001:0db8:0000:0000:0000:0000:1428:57ab can be shortened to 2001:0db8::1428:57ab. Following this rule, any number of consecutive 0000 groups may be reduced to two colons, as long as there is only one double colon used in an address. Leading zeros in a group can also be omitted (as in ::1 for localhost). For example, the addresses below are all valid and equivalent:

2001:0db8:0000:0000:0000:0000:1428:57ab
2001:0db8:0000:0000:0000::1428:57ab
2001:0db8:0:0:0:0:1428:57ab
2001:0db8:0:0::1428:57ab
2001:0db8::1428:57ab
2001:db8::1428:57ab

Note: Having more than one double-colon abbreviation in an address is invalid, as it would make the notation ambiguous.

A sequence of 4 bytes at the end of an IPv6 address can also be written in decimal, using dots as separators. This notation is often used with compatibility addresses. For example, the following two addresses are the same:

::ffff:1.2.3.4

::ffff:0102:0304 and 0:0:0:0:0:ffff:0102:0304.

Additional information can be found in RFC 4291 — IP Version 6 Addressing Architecture.

Literal IPv6 addresses in URLs

In a URL the IPv6-Address is enclosed in brackets. For example:

http://[2001:0db8:85a3:08d3:1319:8a2e:0370:7344]/

This notation allows parsing a URL without confusing the IPv6 address and port number:

https://[2001:0db8:85a3:08d3:1319:8a2e:0370:7344]:443/

Additional information can be found in "RFC 2732 — Format for Literal IPv6 Addresses in URLs" and "RFC 3986 — Uniform Resource Identifier (URI): Generic Syntax."

Network notation

IPv6 networks are written using CIDR (Classless Inter-Domain Routing) notation.

An IPv6 network (or subnet) is a contiguous group of IPv6 addresses, the size of which *must* be a power of two. The initial bits of addresses, identical for all hosts in the network, are called the network's *prefix*.

A network is denoted by the first address in the network and the size in bits of the prefix (in decimal), separated with a slash. For example:

2001:0db8:1234::/48

stands for the network with addresses:

2001:0db8:1234:0000:0000:0000:0000:0000 through
2001:0db8:1234:FFFF:FFFF:FFFF:FFFF:FFFF

Because a single host can be seen as a network with a 128-bit prefix, you will sometimes see host addresses written followed with:

/128.

Types of IPv6 addresses

IPv6 addresses are divided into the following three categories:

- **Unicast Addresses** — Identifies a single network interface. A packet sent to a unicast address is delivered to that specific computer.
- **Multicast Addresses** — Used to define a set of interfaces that typically belong to different nodes instead of just one. When a packet is sent to a multicast address, the protocol delivers the packet to all interfaces identified by that address. Multicast addresses begin with the prefix FF00::/8. Their second octet identifies the addresses scope, that is, the range over which the multicast address is propagated. Commonly used scopes include link-local (2), site-local (5), and global (E).
- **Anycast Addresses** — Also assigned to more than one interface, belonging to different nodes. However, a packet sent to an anycast address is delivered to just *one* of the member interfaces, typically the “nearest” according to the routing protocol’s idea of distance. Anycast addresses cannot be easily identified. They

have the structure of normal unicast addresses, and differ only by being injected into the routing protocol at multiple points in the network.

Special addresses

There are a number of addresses with special meaning in IPv6:

- ::/128 — The address with all zeros is an unspecified address, and is to be used only in software.
- ::1/128 — The loopback address is a localhost address. If an application in a host sends packets to this address, the IPv6 stack will loop these packets back to the same host (corresponding to 127.0.0.1 in IPv4).
- ::/96 — The zero prefix was used for IPv4-compatible addresses. It is now obsolete.
- ::ffff:0:0/96 — This prefix is used for IPv4 mapped addresses (see "Transition mechanisms" on page 395).
- 2001:db8::/32 — This prefix is used in documentation (RFC 3849). Addresses from this prefix should be used anywhere an example IPv6 address is given.
- 2002::/16 — This prefix is used for 6to4 addressing.
- fc00::/7 — Unique Local Addresses (ULA) are routable only within a set of cooperating sites. They were defined in RFC 4193 as a replacement for site-local addresses. The addresses include a 40-bit pseudorandom number that minimizes the risk of conflicts if sites merge or packets somehow leak out. This address space is split into two parts:
  - fc00::/8 — ULA Central, currently not used as the draft is expired.
  - fd00::/8 — ULA, as per RFC 4193, Generator and unofficial registry.
- fe80::/64 — The link-local prefix specifies that the address is valid only in the local physical link. This is analogous to the Autoconfiguration IP address 169.254.0.0/16 in IPv4.
- fec0::/10 — The site-local prefix specifies that the address is valid only inside the local organisation.

Note: Its use has been deprecated in September 2004 by RFC 3879 and systems must not support this special type of address.

- ff00::/8 — The multicast prefix is used for multicast addresses[10] as defined by in "IP Version 6 Addressing Architecture" (RFC 4291).

There are no address ranges reserved for broadcast in IPv6. Instead, applications use multicast to the all-hosts group. IANA maintains the official list of the IPv6 address space. Global unicast assignments can be found at the various RIRs or at the GRH (Ghost Route Hunter) DFP pages.

Zone indices

Link-local addresses present a particular problem for systems with multiple interfaces. Because each interface may be connected to different networks and the addresses all appear to be on the same subnet, an ambiguity arises that cannot be solved by routing tables.

For example, host A has two interfaces that automatically receive link-local addresses when activated (per RFC 2462): fe80::1/64 and fe80::2/64), only one of which is connected to the same physical network as host B which has address fe80::3/64. If host A attempts to contact fe80::3, how does it know which interface (fe80::1 or fe80::2) to use?

The solution, defined by RFC 4007, is the addition of a unique zone index for the local interface, represented textually in the form <address>%<zone_id>. For example:

http://[fe80::1122:33ff:fe11:2233%eth0]:80/

However, this may cause the following problems due to clashing with the percent-encoding used with URIs.

- Microsoft Windows IPv6 stack uses numeric zone IDs: fe80::3%1
- BSD applications typically use the interface name as a zone ID: fe80::3%pcn0
- Linux applications also typically use the interface name as a zone ID: fe80::3%eth0, although Linux ifconfig as of version 1.42 (part of net-tools 1.60) does not display zone IDs.

Relatively few IPv6-capable applications understand zone ID syntax (with the notable exception of OpenSSH), rendering link-local addresses unusable within them if multiple interfaces use link-local addresses.

IPv6 packet

A packet is a formatted block of data carried by a computer network. Figure 68 shows the structure of an IPv6 packet header.

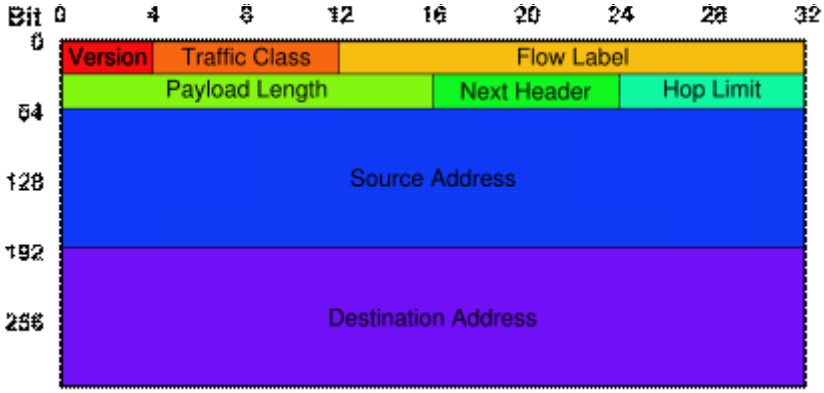


Figure 68 **IPv6 packet header structure**

The IPv6 packet is composed of two main parts:

- Header

  The **header** is in the first 40 octets (320 bits) of the packet and contains:

  - Both source and destination addresses (128 bits each)
  - Version (4-bit IP version)
  - Traffic class (8 bits, Packet Priority)
  - Flow label (20 bits, QoS management)
  - Payload length in bytes (16 bits)
  - Next header (8 bits)
  - Hop limit (8 bits, time to live)

- Payload

  The **payload** can be up to 64 KB in size in standard mode, or larger with a jumbo payload option (refer to "Jumbograms" on page 388).

Fragmentation is handled only in the sending host in IPv6. Routers *never* fragment a packet, and hosts are expected to use PMTU (Path MTU) discovery.

The **protocol** field of IPv4 is replaced with a **Next Header** field. This field usually specifies the transport layer protocol used by a packet's payload. In the presence of options, however, the **Next Header** field

specifies the presence of an **Extra Options** header, which then follows the IPv6 header. The payload's protocol itself is specified in a field of the **Options** header. This insertion of an extra header to carry options is analogous to the handling of AH and ESP (Encapsulating Security Payload) in IPsec for both IPv4 and IPv6.

Transition mechanisms

Until IPv6 completely supplants IPv4, which is not likely to happen in the near future, a number of so-called *transition mechanisms* are needed to enable IPv6-only hosts to reach IPv4 services and to allow isolated IPv6 hosts and networks to reach the IPv6 Internet over the IPv4 infrastructure. The following transition mechanisms are briefly discussed in this section.

- "Dual stack," next
- "Tunneling" on page 395
- "Automatic tunneling" on page 396
- "Configured tunneling" on page 396
- "Proxying and translation" on page 396

Dual stack

Since IPv6 is a conservative extension of IPv4, it is relatively easy to write a network stack that supports both IPv4 and IPv6 while sharing most of the code. Such an implementation is called a *dual stack*. A host implementing a dual stack is called a *dual-stack host*. This approach is described in RFC 4213.

Most current implementations of IPv6 use a dual stack. Some early experimental implementations used independent IPv4 and IPv6 stacks. There are no known implementations that implement IPv6 only.

Tunneling

In order to reach the IPv6 Internet, an isolated host or network must be able to use the existing IPv4 infrastructure to carry IPv6 packets. This is done using a technique somewhat misleadingly known as *tunnelling* that consists of encapsulating IPv6 packets within IPv4, in effect using IPv4 as a link layer for IPv6.

IPv6 packets can be directly encapsulated within IPv4 packets using protocol number 41. They can also be encapsulated within UDP packets, for example, in order to cross a router or NAT device that blocks protocol 41 traffic. They can also use generic encapsulation schemes, such as AYIYA (Anything In Anything) or GRE (Generic Routing Encapsulation).

Automatic tunneling

Automatic tunneling refers to a technique where the tunnel endpoints are automatically determined by the routing infrastructure. The recommended technique for automatic tunneling is 6to4 tunneling, which uses protocol 41 encapsulation. Tunnel endpoints are determined by using a well-known IPv4 anycast address on the remote side, and embedding IPv4 address information within IPv6 addresses on the local side. 6to4 tunneling is widely deployed today.

Another automatic tunneling mechanism is Intra-Site Automatic Tunnel Addressing Protocol (ISATAP). This protocol treats the IPv4 network as a virtual IPv6 local link, with mappings from each IPv4 address to a link-local IPv6 address.

Teredo is an automatic tunneling technique that uses UDP encapsulation and is claimed to be able to cross multiple NAT boxes. Teredo is not widely deployed today, but an experimental version of Teredo is installed with the Windows XP SP2 IPv6 stack.

Note: IPv6, 6to4, and Teredo are enabled by default in Windows Vista.

Configured tunneling

Configured tunneling is a technique where the tunnel endpoints are configured explicitly, either by a human operator or by an automatic service known as a Tunnel Broker. Configured tunneling is usually more deterministic and easier to debug than automatic tunneling, and is therefore recommended for large, well-administered networks.

Configured tunneling typically uses either protocol 41 (recommended) or raw UDP encapsulation.

Proxying and translation

When an IPv6-only host needs to access an IPv4-only service (for example, a web server), some form of translation is necessary. The one form of translation that actually works is the use of a dual-stack application-layer proxy (for example, a web proxy).

Techniques for application-agnostic translation at the lower layers have also been proposed, but they have been found to be too unreliable due to the wide range of functionality required by common application-layer protocols. As such, they are commonly considered to be obsolete.

TCP terminology

This section provides information for TCP terminology.

Acknowledgements (ACKs)

The TCP acknowledgement scheme is cumulative as it acknowledges all the data received up until the time the ACK was generated. As TCP segments are not of uniform size and a TCP sender may retransmit more data than what was in a missing segment, ACKs do not acknowledge the received segment, rather they mark the position of the acknowledged data in the stream. The policy of cumulative acknowledgement makes the generation of ACKs easy and any loss of ACKs do not force the sender to retransmit data. The disadvantage is the sender does not receive any detailed information about the data received except the position in the stream of the last byte that has been received.

Delayed ACKs

Delayed ACKs allow a TCP receiver to refrain from sending an ACK for each incoming segment. However, a receiver should send an ACK for every second full-sized segment that arrives. Furthermore, the standard mandates a receiver must not withhold an ACK for more than 500 ms. The receivers should not delay ACKs that acknowledge out-of-order segments.

Maximum segment size (MSS)

The maximum segment size (MSS) is the maximum amount of data, specified in bytes, that can transmitted in a segment between the two TCP endpoints. The MSS is decided by the endpoints, as they need to agree on the maximum segment they can handle. Deciding on a good MSS is important in a general inter-networking environment because this decision greatly affects performance. It is difficult to choose a good MSS value since a very small MSS means an underutilized network, whereas a very large MSS means large IP datagrams that may lead to IP fragmentation, greatly hampering the performance. An ideal MSS size would be when the IP datagrams are as large as possible without any fragmentation anywhere along the path from the source to the destination. When TCP sends a segment with the SYN bit set during connection establishment, it can send an optional MSS value up to the outgoing interface's MTU minus the size of the fixed TCP and IP headers. For example, if the MTU is 1500 (Ethernet standard), the sender can advertise a MSS of 1460 (1500 minus 40).

Maximum transmission unit (MTU)

Each network interface has its own MTU that defines the largest packet that it can transmit. The MTU of the media determines the maximum size of the packets that can be transmitted without IP fragmentation.

Retransmission

A TCP sender starts a timer when it sends a segment and expects an acknowledgement for the data it sent. If the sender does not receive an acknowledgement for the data before the timer expires, it assumes that the data was lost or corrupted and retransmits the segment. Since the time required for the data to reach the receiver and the acknowledgement to reach the sender is not constant (because of the varying Internet delays), an adaptive retransmission algorithm is used to monitor performance of each connection and conclude a reasonable value for timeout based on the round trip time.

Selective Acknowledgement (SACK)

TCP may experience poor performance when multiple packets are lost from one window of data. With the limited information available from cumulative acknowledgements, a TCP sender can only learn about a single lost packet per round trip time. An aggressive sender could choose to retransmit packets early, but such retransmitted segments may have already been successfully received. The Selective Acknowledgement (SACK) mechanism, combined with a selective repeat retransmission policy, helps to overcome these limitations. The receiving TCP sends back SACK packets to the sender confirming receipt of data and specifies the holes in the data that has been received. The sender can then retransmit only the missing data segments. The selective acknowledgment extension uses two TCP options. The first is an enabling option, **SACKpermitted**, which may be sent in a SYN segment to indicate that the SACK option can be used once the connection is established. The other is the **SACK** option itself, which may be sent over an established connection once permission has been given by **SACKpermitted**.

TCP segment

The TCP segments are units of transfer for TCP and used to establish a connection, transfer data, send ACKs, advertise window size and close a connection. Each segment is divided into three parts:

- Fixed header of 20 bytes
- Optional variable length header, padded out to a multiple of 4 bytes
- Data

The maximum possible header size is 60 bytes. The TCP header carries the control information. **SOURCE PORT** and **DESTINATION PORT** contain TCP port numbers that identify the application programs at the endpoints. The **SEQUENCE NUMBER** field identifies the position in the sender's byte stream of the first byte of attached data, if any, and the **ACKNOWLEDGEMENT NUMBER** field identifies the number of the byte the source expects

to receive next. The **ACKNOWLEDGEMENT NUMBER** field is valid only if the ACK bit in the **CODE BITS** field is set. The 6-bit **CODE BITS** field is used to determine the purpose and contents of the segment. The **HLEN** field specifies the total length of the fixed plus variable headers of the segment as a number of 32-bit words. TCP software advertises how much data it is willing to receive by specifying its buffer size in the **WINDOW** field. The **CHECKSUM** field contains a 16-bit integer checksum used to verify the integrity of the data as well as the TCP header and the header options. The TCP header padding is used to ensure that the TCP header ends and data begins on a 32-bit boundary. The padding is composed of zeros.

TCP window

A TCP window is the amount of data a sender can send without waiting for an ACK from the receiver. The TCP window is a flow control mechanism and ensures that no congestion occurs in the network. For example, if a pair of hosts are talking over a TCP connection that has a TCP window size of 64 KB, the sender can only send 64 KB of data and it must stop and wait for an acknowledgement from the receiver that some or all of the data has been received. If the receiver acknowledges that all the data has been received. The sender is free to send another 64 KB. If the sender gets back an acknowledgement from the receiver that it received the first 32 KB (which is likely if the second 32 KB was still in transit or it is lost), then the sender could only send another 32 KB since it cannot have more than 64 KB of unacknowledged data outstanding (the second 32 KB of data plus the third).

The primary reason for the window is congestion control. The whole network connection, which consists of the hosts at both ends, the routers in between, and the actual connections themselves, might have a bottleneck somewhere that can only handle so much data so fast. The TCP window throttles the transmission speed down to a level where congestion and data loss do not occur.

The factors affecting the window size are as follows:

Receiver's advertised window:
The time taken by the receiver to process the received data and send ACKs may be greater than the sender's processing time, so it is necessary to control the transmission rate of the sender to prevent it from sending more data than the receiver can handle, thus causing packet loss. TCP introduces flow control by declaring a receive window in each segment header.

Sender's congestion window:
The congestion window controls the number of packets a TCP flow has in the network at any time. The congestion window is set using an Additive-Increase, Multiplicative-Decrease (AIMD) mechanism that probes for available bandwidth, dynamically adapting to changing network conditions.

Usable window:
This is the minimum of the receiver's advertised window and the sender's congestion window. It is the actual amount of data the sender is able to transmit. The TCP header uses a 16 bit field to report the receive window size to the sender. Therefore, the largest window that can be used is 2**16 = 65K bytes.

Window scaling:
The ordinary TCP header allocates only 16 bits for window advertisement. This limits the maximum window that can be advertised to 64 KB, limiting the throughput. RFC 1323 provides the window scaling option, to be able to advertise windows greater than 64 KB. Both the endpoints must agree to use window scaling during connection establishment.

The window scale extension expands the definition of the TCP window to 32 bits and then uses a scale factor to carry this 32- bit value in the 16-bit Window field of the TCP header (SEG.WND in RFC-793). The scale factor is carried in a new TCP option — Window Scale. This option is sent only in a SYN segment (a segment with the SYN bit on), hence the window scale is fixed in each direction when a connection is opened.

TCP error recovery

In TCP, each source determines how much capacity is available in the network so it knows how many packets it can safely have in transit. Once a given source has this many packets in transit, it uses the arrival of an ACK as a signal that some of its packets have left the network and it is therefore safe to insert new packets into the network without adding to the level of congestion. TCP uses congestion control algorithms to determine the network capacity. From the congestion control point of view, a TCP connection is in one of the following states.

- **Slow start**: After a connection is established and after a loss is detected by a timeout or by duplicate ACKs.

- **Fast recovery**: After a loss is detected by fast retransmit.
- **Congestion avoidance**: In all other cases. Congestion avoidance and slow start work hand-in-hand. The congestion avoidance algorithm assumes that the chance of a packet being lost due to damage is very small. Therefore, the loss of a packet means there is congestion somewhere in the network between the source and destination. Occurrence of a timeout and the receipt of duplicate ACKs indicates packet loss.

When congestion is detected in the network it is necessary to slow things down, so the **slow start** algorithm is invoked. Two parameters, the congestion window (**cwnd**) and a slow start threshold (**ssthresh**), are maintained for each connection. When a connection is established, both of these parameters are initialized. The **cwnd** is initialized to one MSS. The **ssthresh** is used to determine whether the slow start or congestion avoidance algorithm is to be used to control data transmission. The initial value of **ssthresh** may be arbitrarily high (usually ssthresh is initialized to 65535 bytes), but it may be reduced in response to congestion.

The **slow start** algorithm is used when **cwnd** is less than **ssthresh**, while the **congestion avoidance** algorithm is used when **cwnd** is greater than **ssthresh**. When **cwnd** and **ssthresh** are equal, the sender may use either **slow start** or **congestion avoidance**.

TCP never transmits more than the minimum of **cwnd** and the receiver's advertised window. When a connection is established, or if congestion is detected in the network, TCP is in **slow start** and the congestion window is initialized to one MSS. Each time an ACK is received, the congestion window is increased by one MSS. The sender starts by transmitting one segment and waiting for its ACK. When that ACK is received, the congestion window is incremented from one to two, and two segments can be sent. When each of those two segments is acknowledged, the congestion window is increased to four, and so on. The window size increases exponentially during slow start as shown in Figure 69.

When a time out occurs or a duplicate ACK is received, **ssthresh** is reset to one half of the current window (that is, the minimum of **cwnd** and the receiver's advertised window). If the congestion was detected by an occurrence of a timeout the **cwnd** is set to one MSS. When an ACK is received for data transmitted the **cwnd** is increased, but the way it is increased depends on whether TCP is performing **slow start** or **congestion avoidance**. If the **cwnd** is less than or equal to the **ssthresh**, TCP is in **slow start** and **slow start** continues until TCP is

halfway to where it was when congestion occurred, then **congestion avoidance** takes over. **Congestion avoidance** increments the **cwnd** by MSS squared divided by **cwnd** (in bytes) each time an ACK is received, increasing the **cwnd** linearly as shown in Figure 69. This provides a close approximation to increasing **cwnd** by, at most, one MSS per RTT.

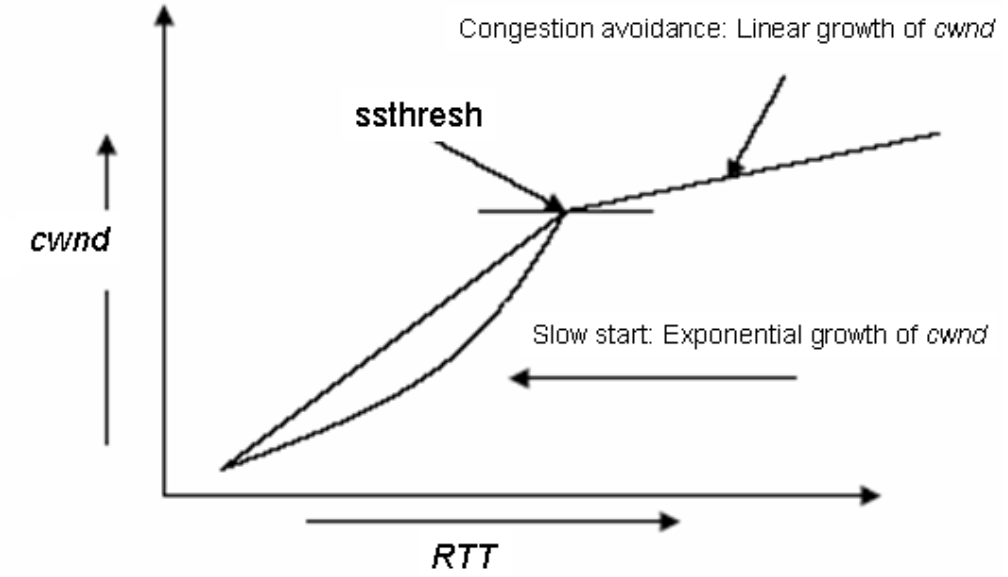


Figure 69 **Slow start and congestion avoidance**

A TCP receiver generates ACKs on receipt of data segments. The ACK contains the highest contiguous sequence number the receiver expects to receive next. This informs the sender of the in-order data that was received by the receiver. When the receiver receives a segment with a sequence number greater than the sequence number it expected to receive, it detects the out-of-order segment and generates an immediate ACK with the last sequence number it has received in-order (that is, a duplicate ACK). This duplicate ACK is not delayed. Since the sender does not know if this duplicate ACK is a result of a lost packet or an out-of-order delivery, it waits for a small number of duplicate ACKs, assuming that if the packets are only reordered there will be only one or two duplicate ACKs before the reordered segment is received and processed and a new ACK is generated. If three or more duplicate ACKs are received in a row, it implies there has been a packet loss. At that point, the TCP sender

retransmits this segment without waiting for the retransmission timer to expire. This is known as *fast retransmit* (Figure 70 on page 403).

After fast retransmit has sent the supposedly missing segment, the **congestion avoidance** algorithm is invoked instead of the **slow start**; this is called *fast recovery*. Receipt of a duplicate ACK implies that not only is a packet lost, but that there is data still flowing between the two ends of TCP, as the receiver will only generate a duplicate ACK on receipt of another segment. Hence, fast recovery allows high throughput under moderate congestion.

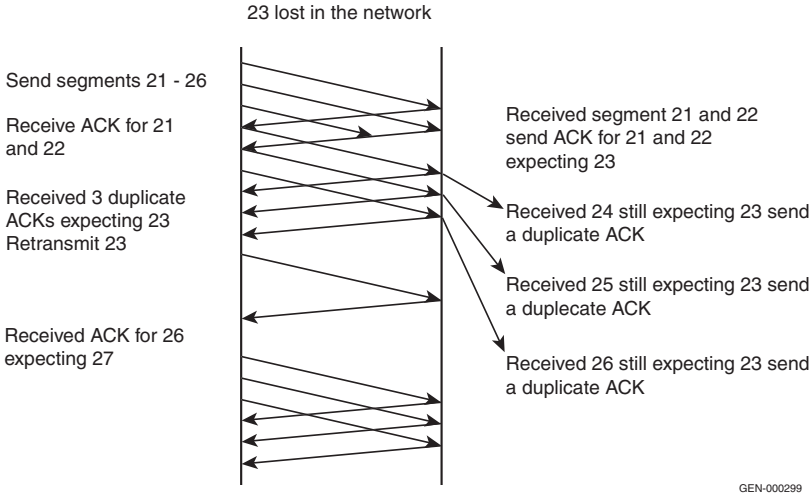


Figure 70 **Fast retransmit**

Network congestion

A network link is said to be congested if contention for it causes queues to build up and packets start getting dropped. The TCP protocol detects these dropped packets and starts retransmitting them , but using aggressive retransmissions to compensate for packet loss tends to keep systems in a state of network congestion even after the initial load has been reduced to a level which would not normally have induced network congestion. In this situation, demand for link bandwidth (and eventually queue space), outstrips what is available. When congestion occurs, all the flows that detect it must reduce their transmission rate. If they do not do so, the network will remain in an unstable state with queues continuing to build up.

IP Security

IP Security is a set of protocols developed by the IETF to support secure exchange of packets in the IP layer. IP Security has been deployed widely to implement Virtual Private Networks (VPNs).

IP security supports two encryption modes:

- Transport
- Tunnel

Transport mode encrypts only the payload of each packet, but leaves the header untouched. The more secure Tunnel mode encrypts both the header and the payload.

On the receiving side, an IP Security compliant device decrypts each packet. For IP security to work, the sending and receiving devices must share a public key. This is accomplished through a protocol known as Internet Security Association and Key Management Protocol / Oakley (ISAKMP / Oakley), which allows the receiver to obtain a public key and authenticate the sender using digital certificates.

iSCSI technology overview

iSCSI (Internet Small Computer System Interface) is an IP-based storage networking standard for linking data storage facilities developed by the Internet Engineering Task Force. By transmitting SCSI commands over IP networks, iSCSI can facilitate block-level transfers over the intranet/internet.

The iSCSI architecture is similar to a client/server architecture. In this case, the client is an initiator that issues an I/O request and the server is a target (such as a device in a storage system). This architecture can be used over IP networks to provide distance extension. This can be implemented between routers, host-to-switch, and storage array-to-storage array to provide asynchronous/synchronous data transfer.

This section contains the following information:

- "iSCSI discovery", next
- "Digests" on page 406
- "iSCSI error recovery" on page 406
- "iSCSI security" on page 407

iSCSI discovery

In order for an iSCSI initiator to establish an iSCSI session with an iSCSI target, the initiator needs the IP address, TCP port number, and iSCSI target name information. The goals of iSCSI discovery mechanisms are to provide low overhead support for small iSCSI setups and scalable discovery solutions for large enterprise setups.

Send target

An initiator may log in to an iSCSI target with session type of discovery and request a list of target WWUIs through a separate **SendTargets** command. All iSCSI targets are required to support the **SendTargets** command.

iSNS

The iSNS protocol is designed to facilitate the automated discovery, management, and configuration of iSCSI and Fibre Channel devices on a TCP/IP network. iSNS provides intelligent storage discovery and management services comparable to those found in Fibre Channel networks, allowing a commodity IP network to function in a similar capacity as a storage area network. iSNS also facilitates a seamless integration of IP and Fibre Channel networks, due to its

ability to emulate Fibre Channel fabric services, and manage both iSCSI and Fibre Channel devices. iSNS thereby provides value in any storage network comprised of iSCSI devices, Fibre Channel devices, or any other combination.

SLP

iSCSI targets are registered with SLP as a set of service URLs, one for each address on which the target may be accessed. Initiators discover these targets using SLP service requests. Targets that do not directly support SLP, or are under the control of a management service, may be registered by a proxy service agent as part of the software providing this service.

Digests

Digests enable the checking of end-to-end, non-cryptographic data integrity beyond the integrity checks provided by the link layers and the cover the entire communication path including all elements that may change the network level PDUs such as routers, switches, and proxies.

Optional header and data digests protect the integrity of the header and data, respectively. The digests, if present, are located after the header and PDU-specific data and cover the header and the PDU data, each including the padding bytes, if any. The existence and type of digests are negotiated during the Login phase. The separation of the header and data digests is useful in iSCSI routing applications, where only the header changes when a message is forwarded. In this case, only the header digest should be recalculated.

iSCSI error recovery

iSCSI supports three levels of error recovery: 0, 1,and 2:

- Error recovery level 0 implies session level recovery.
- Error recovery level 1 implies level 0 capabilities as well as digest failure recovery.
- Error recovery level 2 implies level 1 capabilities as well as connection recovery.

The most basic kind of recovery is called *session* recovery. In session recovery, whenever any kind of error is detected, the entire iSCSI session is terminated. All TCP connections connecting the initiator to the target are closed, and all pending SCSI commands are completed

with an appropriate error status. A new iSCSI session is then established between the initiator and target, with new TCP connections.

Digest failure recovery starts if the iSCSI driver detects that data arrived with an invalid data digest and that data packet must be rejected. The command corresponding to the corrupted data can then be completed with an appropriate error indication.

Connection recovery can be used when a TCP connection is broken. Upon detection of a broken TCP connection, the iSCSI driver can either immediately complete the pending command with an appropriate error indication, or can attempt to transfer the SCSI command to another TCP connection. If necessary, the iSCSI initiator driver can establish another TCP connection to the target, and the iSCSI initiator driver can inform the target the change in allegiance of the SCSI command to another TCP connection.

iSCSI security

Historically, native storage systems have not had to consider security because their environments offered minimal security risks. These environments consisted of storage devices either directly attached to hosts or connected through a Storage Area Network (SAN) distinctly separate from the communications network. The use of storage protocols, such as SCSI over IP-networks, requires that security concerns be addressed. iSCSI implementations must provide means of protection against active attacks (such as, pretending to be another identity, message insertion, deletion, modification, and replaying) and passive attacks (such as, eavesdropping, gaining advantage by analyzing the data sent over the line). Although technically possible, iSCSI should not be configured without security. iSCSI configured without security should be confined, in extreme cases, to closed environments without any security risk.

Security mechanisms

The entities involved in iSCSI security are the initiator, target, and IP communication end points. iSCSI scenarios in which multiple initiators or targets share a single communication end points are expected. To accommodate such scenarios, iSCSI uses two separate security mechanisms:

- In-band authentication between the initiator and the target at the iSCSI connection level (carried out by exchange of iSCSI Login PDUs.)

- Packet protection (integrity, authentication, and confidentiality) by IPsec at the IP level.

The two security mechanisms complement each other. The in-band authentication provides end-to-end trust (at login time) between the iSCSI initiator and the target while IPsec provides a secure channel between the IP communication end points.

Authentication methods

The authentication methods that can be used are:

CHAP (Challenge Handshake Authentication Protocol)

The Challenge-Handshake Authentication Protocol (CHAP) is used to periodically verify the identity of the peer using a three-way handshake. This is done upon establishing initial link and *may* be repeated anytime after the link has been established. CHAP provides protection against playback attack by the peer through the use of an incrementally changing identifier and a variable challenge value. The use of repeated challenges is intended to limit the time of exposure to any single attack. The authenticator is in control of the frequency and timing of the challenges. This authentication method depends upon a "secret" known only to the authenticator and that peer. The secret is not sent over the link.

SRP (Secure Remote Password)

This mechanism is suitable for negotiating secure connections using a user-supplied password, while eliminating the security problems traditionally associated with reusable passwords. This system also performs a secure key exchange in the process of authentication, allowing security layers (privacy and / or integrity protection) to be enabled during the session. Trusted key servers and certificate infrastructures are not required, and clients are not required to store or manage any long-term keys.

KRB5 (Kerberos V5)

Kerberos provides a means of verifying the identities of principals, (such as a workstation user or a network server) on an open (unprotected) network. This is accomplished without relying on authentication by the host operating system, or basing trust on host addresses, or requiring physical security of all the hosts on the network, and under the assumption that packets traveling along the network can be read, modified, and inserted at will. Kerberos performs authentication under these conditions as a trusted third-party authentication service by using conventional cryptography such as a shared secret key.

SPKM1 & 2 (Simple Public Key GSS-API Mechanisim)
This mechanism provides authentication, key establishment, data integrity, and data confidentiality in an on-line distributed application environment using a public-key infrastructure. SPKM can be used as a drop-in replacement by any application which makes use of security services through GSS-API calls (for example, any application which already uses the Kerberos GSS-API for security).

iSCSI solutions

This section contains the following information:

- "General best practices", next
- "General supported configurations" on page 411
- "EMC native iSCSI targets" on page 412
- "Configuring iSCSI targets" on page 414
- "Bridged solutions" on page 416

General best practices

This section lists general best practices.

Network design

The network should be dedicated solely to the IP technology being used and other traffic should be carried over it.

The network must be a well-engineered network with *no* packet loss or packet duplication. This would lead to retransmission, which is undesirable.

While planning the network, care must be taken to ascertain that the utilized throughput will never exceed the available bandwidth. Oversubscribing available bandwidth will lead to network congestion, which causes dropped packets and leads to TCP slow start. Network congestion must be considered between switches as well as between the switch and the end device. For more information on network congestion, refer to "Network congestion" on page 403.

The MTU must be configured based on the maximum available MTU supported by each component on the network.

Header and data digest

Header and data digest are mandatory when using a routed network (Layer 3) or when using Layer 2 network with VLAN tagging.

In a plain LAN (other than those mentioned above) digests are not mandatory.

General supported configurations

This section lists general supported configurations.

Direct connect

Direct connect between an iSCSI-enabled host and storage system is the most basic configuration. This configuration allows proof of concept for pre-rollout testing and simple secure connectivity using iSCSI.

Figure 71 **Direct connect example**

Network connect

Storage array iSCSI ports are connected to an IP network. The network can either be a local area network or a routed network. iSCSI servers are connected to the network as well and communicate with the storage array iSCSI ports. All standard network equipment is supported. However, it is recommended to use enterprise class equipment since they provide more memory to address network congestion.

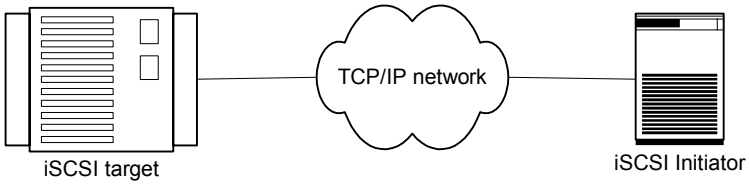


Figure 72 **Network connect example**

EMC native iSCSI targets

This section discusses the following EMC native iSCSI targets:

- "Symmetrix", next
- "CLARiiON" on page 412
- "Celerra" on page 413

Symmetrix

DMX3

The iSCSI channel director supports iSCSI channel connectivity to IP networks and to iSCSI-capable open systems server systems for block storage transfer between hosts and storage. The primary applications are storage consolidation and host extension for stranded servers and departmental workgroups. The Symmetrix iSCSI director provides up to four 1 Gb/s Ethernet ports and connects through LC connectors. The iSCSI directors support the iSNS protocol. Authentication mechanism is CHAP (Challenge Handshake Authentication Protocol).

LUNs are configured in the same manner as for Fibre Channel directors and are assigned to the iSCSI ports. LUN masking is available.

References

For configuration of Symmetrix iSCSI target please check the Symmetrix configuration guide.

For up-to-date iSCSI host support please refer to *EMC Support Matrix*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com.

For configuration of iSCSI server, please check the respective host connectivity guide.

CLARiiON

CX3 20/40 iSCSI/FC array

This can be configured as an iSCSI array or Fibre Channel array. All iSCSI ports on the array are 1 Gb/s ethernet ports. iSNS protocol is supported. Authentication mechanism is CHAP (Challenge Handshake Authentication Protocol).

LUNs are configured in the same manner as for Fibre Channel array and are assigned to a storage group.

CX 300i/500i

These are dedicated iSCSI arrays. All iSCSI ports on the array are 1 Gb/s ethernet ports. iSNS protocol is supported. Authentication

mechanism is CHAP (Challenge Handshake Authentication Protocol).

LUNs are configured in the same manner as for Fibre Channel array and are assigned to a storage group.

AX 150i/100i

These are dedicated iSCSI arrays. All iSCSI ports on the array are 1 Gb/s ethernet ports. iSNS protocol is supported. Authentication mechanism is CHAP (Challenge Handshake Authentication Protocol).

LUNs are configured in the same manner as for Fibre Channel array and are assigned to a storage group.

References

For configuration of CLARiiON® iSCSI target please check the CLARiiON configuration guide.

For up-to-date iSCSI host support please refer to *EMC Support Matrix*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com.

For configuration of iSCSI server, please check the respective host connectivity guide.

Celerra

Celerra® Network Server provides iSCSI target capabilities combined with NAS capabilities. The Celerra iSCSI system is defined by creating a file system. The file system is build on Fibre Channel LUNs accessible on EMC Symmetrix or CLARiiON arrays. The file system is then mounted on the Celerra server data movers. Out of the file system iSCSI LUNs are defined and allocated to iSCSI targets. The targets are then associated with one of the Celerra TCP/IP interfaces.

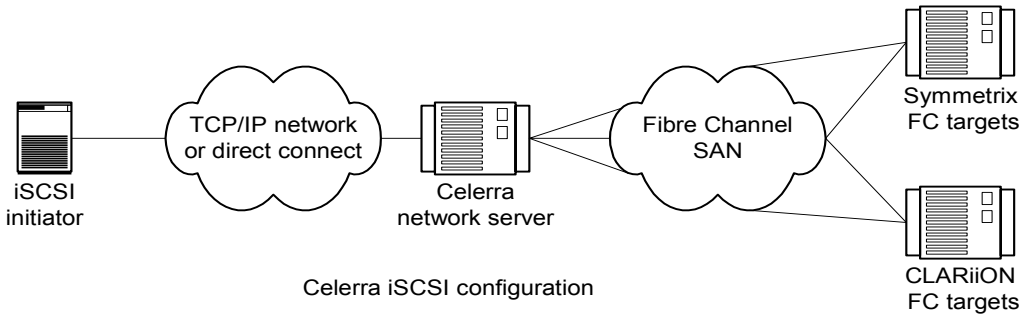


Figure 73 **Celerra Network server**

All Celerra network servers can be configured to provide iSCSI services.

The following are some of the characteristics of the Celerra Network Server:

- iSCSI error recovery level 0 (session-level recovery).
- Supports CHAP with unlimited entries for one-way authentication and one entry for reverse authentication.
- Uses iSNS protocol for discovery.
- Provides 10 Gb/s and 1 Gb/s interfaces
- Supports EMC storage Symmetrix and CLARiiON on the backend.

Implementation best practices

The following information is provided to help you estimate size requirements for iSCSI LUNs and provides guidelines for configuring iSCSI on the Celerra Network Server.

- Estimate size requirements for the file system.

  When using regular iSCSI LUNs, the file system should be large enough to hold the LUNs and the planned snapshots of those LUNs. Each iSCSI snapshot may require the same amount of space on the file system as the LUN.

- Create and mount file systems for iSCSI LUNs.

  The next step in configuring iSCSI targets on a Celerra Network Server is to create and mount one or more file systems to provide a dedicated storage resource for the iSCSI LUNs. Create and mount a file system through Celerra Manager or the CLI. The Celerra Manager Online Help and the technical module *Managing Celerra Volumes and File Systems Manually* provide instructions.

Configuring iSCSI targets

This section lists the tasks you must perform to configure iSCSI targets and LUNs on the Celerra Network Server.

The online Celerra man pages and the *Celerra Network Server Command Reference Manual* provide detailed descriptions of the commands used in these procedures.

1. Create iSCSI targets:

   You need to create one or more iSCSI targets on the Data Mover so an iSCSI initiator can establish a session and exchange data with the Celerra Network Server.

2. Create iSCSI LUNs:

   After creating an iSCSI target, you must create iSCSI LUNs on the target. The LUNs provide access to the storage space on the Celerra Network Server. From the point of view of a client system, a Celerra iSCSI LUN appears as any other disk device.

3. Create iSCSI LUN masks:

   On the Celerra Network Server, a LUN mask on a target controls incoming iSCSI access by granting or denying an iSCSI initiator access to specific iSCSI LUNs on that target. When created, an iSCSI target has no LUN masks, which means no initiator can access LUNs on that target. To enable an initiator to access LUNs on a target, you need to create a LUN mask to specify the initiator and the LUNs it can access.

4. Configure iSNS on the Data Mover (optional):

   If you want iSCSI initiators to automatically discover the iSCSI targets on a Data Mover, you can configure an iSNS client on the Data Mover. Configuring an iSNS client on the Data Mover causes the Data Mover to register all of its iSCSI targets with an external iSNS server. iSCSI initiators can then query the iSNS server to discover the available targets on the Data Movers.

5. Create CHAP entries (optional):

   If you want a Data Mover to authenticate the identity of each iSCSI initiator, configure CHAP authentication on the Data Mover. To configure CHAP, you must:

   a. Set the appropriate parameters so targets on the Data Mover require CHAP authentication.

   b. Create a CHAP entry for each initiator that contacts the Data Mover. CHAP entries are configured on each Data Mover—each initiator has a unique CHAP secret for the Data Mover.

   c. In some cases, initiators authenticate the identity of the targets as well. In this case, you must configure a CHAP entry for reverse authentication. Reverse authentication entries differ

from regular CHAP entries because each Data Mover can have only one CHAP secret. The Data Mover uses the same CHAP secret when any iSCSI initiator authenticates a target on the Data Mover.

6. Start the iSCSI service:

   Before using iSCSI targets on the Celerra Network Server, you must start the iSCSI service on the Data Mover.

References

For more information please refer to *Configuring iSCSI Targets on Celerra*, available on Powerlink.

Bridged solutions

The following switches are discussed in this section:

- "Brocade", next
- "Cisco" on page 419
- "Brocade M-Series" on page 425

Brocade

The FC4-16IP iSCSI gateway service is an intermediate device in the network, allowing iSCSI initiators in an IP SAN to access and utilize storage in a Fibre Channel (FC) SAN.

Supported configurations

The iSCSI gateway enables applications on an IP network to use an iSCSI initiator to connect to FC targets. The iSCSI gateway translates iSCSI protocol to Fibre Channel Protocol (FCP), bridging the IP network and FC SAN.

Note: The FC4-16IP iSCSI gateway service is not compatible with other iSCSI gateway platforms, including Brocade iSCSI Gateway or the SilkWorm Multiprotocol Router.

Figure 74 on page 417 shows a basic iSCSI gateway service implementation.

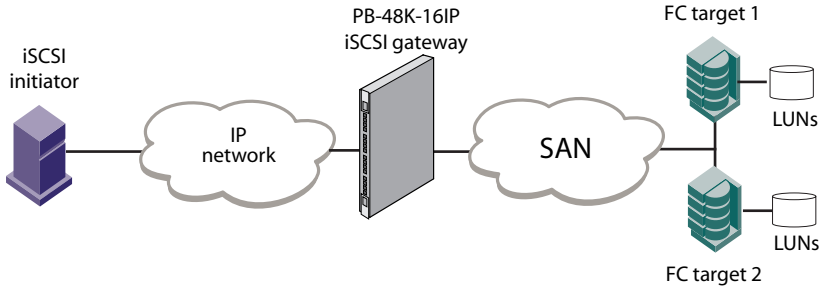


Figure 74 **iSCSI gateway service basic implementation**

The Brocade FC4-16IP blade acts as an iSCSI gateway between FC-attached targets and iSCSI initiators. On the iSCSI initiator, iSCSI is mapped between the SCSI driver and the TCP/IP stack. At the iSCSI gateway port, the incoming iSCSI data is converted to FCP (SCSI on FC) by the iSCSI virtual initiator and then forwarded to the FC target. This allows low-cost servers to leverage an existing FC infrastructure.

To represent all iSCSI initiators and sessions, each iSCSI portal has one iSCSI virtual initiator (VI) to the FC fabric that appears as an N_Port device with a special WWN format. Regardless of the number of iSCSI initiators or iSCSI sessions sharing the portal, Fabric OS uses one iSCSI VI per iSCSI portal.

Fabric OS provides a mechanism that maps LUNs to iSCSI VTs, a one-to-one mapping with unique iSCSI Qualified Names (IQNs) for each target. It presents an iSCSI VT for each native FC target to the IP network and an iSCSI VI for each iSCSI port to the FC fabric.

Fabric OS also supports more complicated configurations, allowing each iSCSI VT to be mapped to one or more physical FC targets. Each FC target can have one or more LUNs. Physical LUNs can be mapped to different virtual LUNs.

Implementation best practices

The following are scalability guidelines, restructions, and limitations:

# of iSCSI sessions per port =	64
# of iSCSI ports per FC4-16IP blade =	8
# of iSCSI blades in a switch	4
# of iSCSI sessions per FC4-16IP blade =	512
# of iSCSI sessions per switch =	1024
# of TCP sessions per switch =	1024
# of TCP connections per iSCSI session =	2
# of iSCSI sessions per fabric =	4096
# of TCP connections per fabric =	4096
# of iSCSI targets per fabric =	4096
# CHAP entries per fabric =	4096
# LUNS per iSCSI target =	256
# Members per discovery domain =	64
# Discovery domains per discovery domain set =	4096
# of Discovery domain sets =	4

The following are installation tips and recommendations:

- All iSCSI Virtual Initiators should be included in the zone with specified target.
- All iSCSI VIs must be registered on the CLARiiON array and added to the appropriate storage groups.
- All iSCSI VIs must be added to the Symmetrix VCM database, if utilizing the device masking functionality.
- If the FC targets use access control lists / database, you must add the FC NWWN / PWWN of the Ironman blade to the ACL / database (fclunquery -s to determine Ironman FC NWWN / PWWN).

- Recommend masking all LUNS for all VIs and performing the LUN masking functionality from the Ironman blade by creating individual iSCSI Virtual Targets and assigning the LUNS to the appropriate iSCSI Virtual Target.
- Firmware upgrades are not online events for the Ironman GigE ports, so plan accordingly.
- The **fcLunQuery** command only gets addresses from targets that support the **ReportLuns** command.

References

All Brocade documentation can be located at http://www.brocade.com. Click **Brocade Connect** to register, at no cost, for a user ID and password.

The following documentation is available for Fabric OS:

- *Fabric OS Administrator's Guide*
- *Fabric OS Command Reference*
- *Fabric OS MIB Reference*
- *Fabric OS Message Reference*
- *Brocade Glossary*

The following documentation is available for SilkWorm 48000 director and iSCSI blade:

- *SilkWorm 48000 Hardware Reference Manual*
- *iSCSI Gateway Service Administrator's Guide*
- *FC4-16IP Hardware Reference Manual*

Cisco

Cisco MDS 9000 storage switches are multiprotocol switches that support the Fibre Channel and Gigabit Ethernet (FCIP and iSCSI) protocols. Each switch model can be used as a Fibre Channel-iSCSI gateway to support iSCSI solutions with Fibre Channel targets (Symmetrix and CLARiiON).

Cisco MDS 9000 family IP storage (IPS) services extend the reach of Fibre Channel SANs by using open-standard, IP-based technology. The switch allows IP hosts to access Fibre Channel storage using the iSCSI protocol. The iSCSI feature is specific to the IPS module and is available in Cisco MDS 9200 Switches or Cisco MDS 9500 Directors. The Cisco MDS 9216i switch and the 14/2 Multiprotocol Services (MPS-14/2) module also allow you to use Fibre Channel, FCIP, and

iSCSI features. The MPS-14/2 module is available for use in any switch in the Cisco MDS 9200 Series or Cisco MDS 9500 Series.

Supported configurations

Initiator presentation modes (transparent and proxy)

The two modes available to present iSCSI hosts in the Fibre Channel fabric are transparent initiator mode and proxy initiator mode.

- In transparent initiator mode, each iSCSI host is presented as one virtual Fibre Channel host. The benefit of transparent mode is it allows a finer level of Fibre Channel access control configuration (similar to managing a "real" Fibre Channel host). Because of the one-to-one mapping from iSCSI to Fibre Channel, each host can have different zoning or LUN access control on the Fibre Channel storage device.
- In proxy-initiator mode, there is only one virtual Fibre Channel host per one IPS port that all iSCSI hosts use to access Fibre Channel targets. In a scenario where the Fibre Channel storage device requires explicit LUN access control for every host, the static configuration for each iSCSI initiator can be overwhelming. In such case, using the proxy-initiator mode simplifies the configuration.

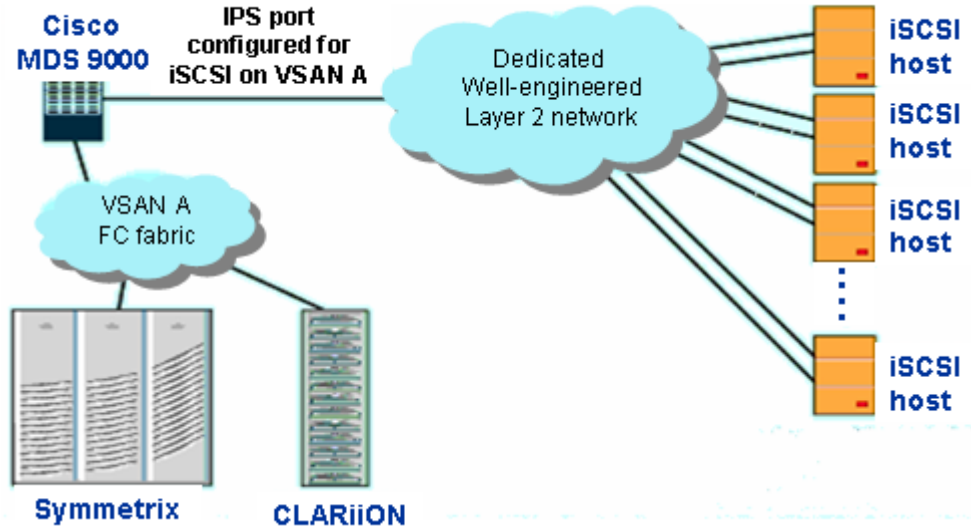


Figure 75 **Supportable configuration example**

The following iSCSI configurations are supported:

- The Cisco MDS switches can be used as Fibre Channel-iSCSI gateway to run applictions using an iSCSI initiator to Symmetrix and CLARiiON storage devices.
- Host-based redundancy is supported through the use of EMC PowerPath®.

iSCSI configuration has the following limits:

- The maximum number of iSCSI initiators supported in a fabric is 1800.
- The maximum number of iSCSI sessions supported by an IPS port in either transparent or proxy initiator mode is 300.
- The maximum number of iSCSI session support by switch is 5000.
- The maximum number of iSCSI targets supported in a fabric is 6000.

Configuration overview

To use the iSCSI feature, you must explicitly enable iSCSI on the required switches in the fabric. By default, this feature is disabled in all switches in the Cisco MDS 9000 family. Each physical Gigabit Ethernet interface on an IPS module or MPS-14/2 module can be used to translate and route iSCSI requests to Fibre Channel targets and responses in the opposite direction. To enable this capability, the corresponding iSCSI interface must be in an enabled state.

Presenting Fibre Channel Targets as iSCSI Targets

The IPS module or MPS-14/2 module presents physical Fibre Channel targets as iSCSI virtual targets, allowing them to be accessed by iSCSI hosts. It does this in one of two ways:

- Dynamic mapping — Automatically maps all the Fibre Channel target devices/ports as iSCSI devices. Use this mapping to create automatic iSCSI target names.
- Static mapping — Manually creates iSCSI target devices and maps them to the whole Fibre Channel target port or a subset of Fibre Channel LUNs. With this mapping, you must specify unique iSCSI target names.

Presenting iSCSI hosts as virtual Fibre Channel hosts

The IPS module or MPS-14/2 module connects to the Fibre Channel storage devices on behalf of the iSCSI host to send commands and

transfer data to and from the storage devices. These modules use a virtual Fibre Channel N_Port to access the Fibre Channel storage devices on behalf of the iSCSI host. iSCSI hosts are identified by either iSCSI qualified name (IQN) or IP address.

Initiator identification

iSCSI hosts can be identified by the IPS module or MPS-14/2 module using the following:

- iSCSI qualified name (IQN)

  An iSCSI initiator is identified based on the iSCSI node name it provides in the iSCSI login. This mode can be useful if an iSCSI host has multiple IP addresses and you want to provide the same service independent of the IP address used by the host. An initiator with multiple IP addresses (multiple network interface cards—NICs) has one virtual N_Port on each IPS port to which it logs into.

- IP address

  An iSCSI initiator is identified based on the IP address of the iSCSI host. This mode is useful if an iSCSI host has multiple IP addresses and you want to provide different service-based on the IP address used by the host. It is also easier to get the IP address of a host compared to getting the iSCSI node name. A virtual N_Port is created for each IP address it uses to log in to iSCSI targets. If the host using one IP address logs in to multiple IPS ports, each IPS port will create one virtual N_Port for that IP address.

You can configure the iSCSI initiator identification mode on each IPS port and all the iSCSI hosts terminating on the IPS port will be identified according to that configuration. The default mode is to identify the initiator by name.

iSCSI access control

Two mechanisms of access control are available for iSCSI devices.

- Fibre Channel zoning-based access control
- iSCSI ACL-based access control

Depending on the initiator mode used to present the iSCSI hosts in the Fibre Channel fabric, either or both access control mechanisms can be used.

Fibre Channel zoning-based access control

Cisco SAN-OS VSAN and zoning concepts have been extended to cover both Fibre Channel devices and iSCSI devices. Zoning is the standard access control mechanism for Fibre Channel devices, which is applied within the context of a VSAN. Fibre Channel zoning has been extended to support iSCSI devices, and this extension has the advantage of having a uniform, flexible access control mechanism across the whole SAN.

- Fibre Channel device pWWN.
- Interface and switch WWN. Device connecting through that interface is within the zone.

In the case of iSCSI, multiple iSCSI devices may be connected behind an iSCSI interface. Interface-based zoning may not be useful because all the iSCSI devices behind the interface will automatically be within the same zone.

In transparent initiator mode (where one Fibre Channel virtual N_Port is created for each iSCSI host), the standard Fibre Channel device pWWN-based zoning membership mechanism can be used if an iSCSI host has static WWN mapping.

Zoning membership mechanism has been enhanced to add iSCSI devices to zones based on the following:

- IPv4 address / subnet mask
- IPv6 address / prefix length (currently EMC does not support ip version 6)
- iSCSI qualified name (IQN)
- Symbolic-node-name (IQN)

For iSCSI hosts that do not have a static WWN mapping, the feature allows the IP address or iSCSI node name to be specified as zone members. Note that iSCSI hosts that have static WWN mapping can also use these features. IP address-based zone membership allows multiple devices to be specified in one command by providing the subnet mask.

iSCSI-based access control

iSCSI-based access control is applicable only if static iSCSI virtual targets are created. For a static iSCSI target, you can configure a list of iSCSI initiators that are allowed to access the targets.

By default, static iSCSI virtual targets are not accessible to any iSCSI host. You must explicitly configure accessibility to allow an iSCSI

virtual target to be accessed by all hosts. The initiator access list can contain one or more initiators. The iSCSI initiator can be identified by one of the following mechanisms:

- iSCSI node name
- IPv4 address and subnet
- IPv6 address (currently EMC does not support IP version 6)

Note: For a transparent mode iSCSI initiator, if both Fibre Channel zoning and iSCSI ACLs are used, for every static iSCSI target that is accessible to the iSCSI host, the initiator's virtual N_Port should be in the same Fibre Channel zone as the Fibre Channel target.

iSCSI session authentication

The IPS module or MPS-14/2 module supports the iSCSI authentication mechanism to authenticate the iSCSI hosts that request access to the storage devices. By default, the IPS modules or MPS-14/2 modules allow CHAP or None authentication of iSCSI initiators. If authentication is always used, you must configure the switch to allow only CHAP authentication.

For CHAP user name or secret validation, you can use any method supported and allowed by the Cisco MDS AAA infrastructure. AAA authentication supports a RADIUS, TACACS+, or local authentication device.

iSCSI immediate data and unsolicited data features

Cisco MDS switches support the iSCSI immediate data and unsolicited data features if requested by the initiator during the login negotiation phase. *Immediate data* is iSCSI write data contained in the data segment of an iSCSI command protocol data unit (PDU), such as combining the write command and write data together in one PDU. *Unsolicited data* is iSCSI write data that an initiator sends to the iSCSI target, such as an MDS switch, in an iSCSI data-out PDU without having to receive an explicit ready to transfer (R2T) PDU from the target.

These two features help reduce I/O time for small write commands because it removes one round-trip between the initiator and the target for the R2T PDU. As an iSCSI target, the MDS switch allows up to 64 KB of unsolicited data per command. This is controlled by the FirstBurstLength parameter during iSCSI login negotiation phase.

If an iSCSI initiator supports immediate data and unsolicited data features, these features are automatically enabled on the MDS switch with no configuration required.

Implementation best practices

Symmetrix setup

Symmetrix SRDF ports should be configured as standard Fibre Channel SRDF ports. In a Fibre Channel environment, the Cisco MDS switch provides all the services of a Fibre Channel switch, similar to those provided by any other Fibre Channel switch.

CLARiiON setup

CLARiiON ports should be configured as standard Fibre Channel target ports for iSCSI configurations.

References

All documentation can be found at www.cisco.com. Please search on *Cisco MDS Configuration Guide* and choose the guide relevant to the code running in your environment.

Brocade M-Series

The Brocade M-Series Eclipse switches are multiprotocol switches (see Figure 76 on page 426). They support Fibre Channel and Gigabit Ethernet. In an FC environment, the switch provides all the services similar to those provided by any other FC switch.

The Brocade M-Series multiprotocol switches are used in distance-extension configurations running SRDF or MirrorView over IP, for SAN Copy deployment over IP, or as a gateway device to perform FC-iSCSI translation.

On these switches:

- Ports 1 through 12 should be configured as FC ports (F_Ports or R_Ports).
- Ports 13 to 16 can be configured as either iFCP (TCP) ports (to cover long distances) or as iSCSI ports.

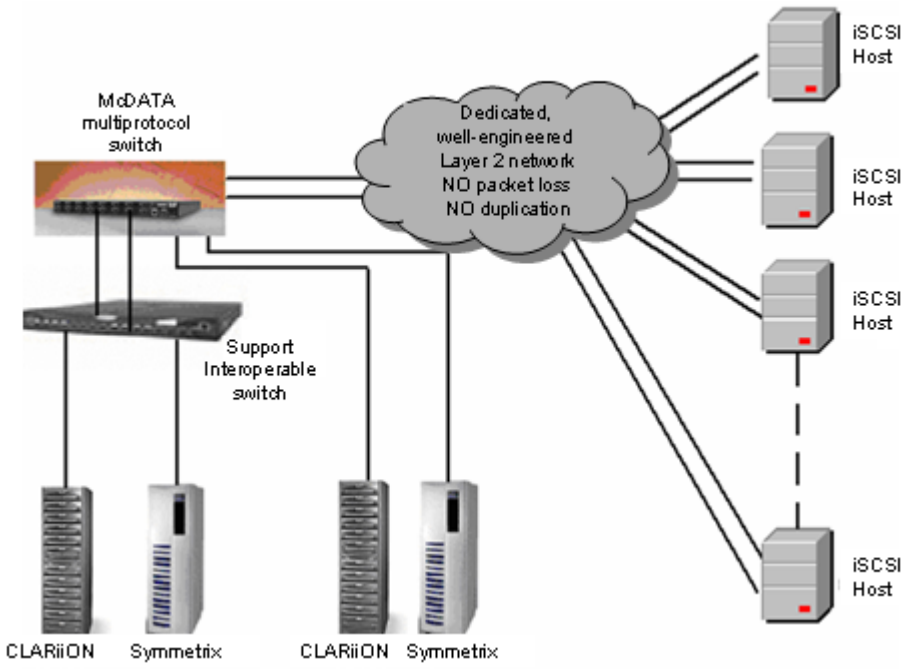


Figure 76 **Brocade M-Series multiprotocol switch**

Supported configurations

The following iSCSI configurations are supported:

- The Eclipse switches can be used as Fibre Channel-iSCSI gateway to run applications using an iSCSI initiator to Symmetrix and CLARiiON storage devices.
- The iSCSI host must run the MS iSCSI driver on a Windows 2000 or Windows 2003 system.
- Host-based redundancy is supported through the use of EMC PowerPath.
- The configuration supports a fan-out ratio of up to 50 iSCSI hosts connected to each Eclipse iSCSI port.

Implementation best practices

Certain limitations and restrictions need to be enforced for iSCSI configurations:

- The use of a single Brocade M-Series Eclipse switch port for both iSCSI and iFCP is not supported.

- The network must be a local Layer 2 network. Currently, EMC does not support deployment over Layer 3 networks or over extended distances.
- The network must be dedicated solely to the iSCSI configuration and no traffic apart from iSCSI traffic should be carried over it.
- The network must be a well-engineered network with no packet loss or packet duplication. While planning the network, care must be taken in making certain that the utilized throughput will never exceed the available bandwidth. The MTU must be configured based on the maximum available MTU supported by each component on the network.
- iSCSI sessions may need to be manually reestablished. If the iSCSI session breaks due to an FC-side cable pull, it will need to be enabled manually from the host.

Symmetrix setup

For iSCSI configurations, the Symmetrix ports should be configured as standard FA ports for access to any Windows system.

CLARiiON setup

For iSCSI configurations, the CLARiiON ports should be configured as standard FC ports. If a CX200 is used in the iSCSI configuration, connections to only one port per SP are supported.

Settings for Brocade/ Brocade M-Series/Cisco/ QLogic switches

While setting up the Brocade/ Brocade M-Series/Cisco/QLogic switches to work with the Brocade M-Series Eclipse switch, some settings are essential, as follows:

- The Brocade/Brocade M-Series/Cisco/QLogic switches must use the latest code supported by EMC for Interop Solutions. The *EMC Support Matrix*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com, provides more information.
- The E_Port on the Brocade must be in Brocade native mode, or Brocade M-Series in Brocade M-Series mode (to which the Brocade M-Series Eclipse switch is connected to) must be hard set to the actual port speed and autonegotiation must be disabled.
- The Cisco and QLogic switches must be set up to work in the Interop mode.

- Once the zoning is defined on the Brocade M-Series Eclipse switches, it appears in the Brocade / Brocade M-Series / Cisco / QLogic active configurations as zones beginning with SoIP_xxx. When changes are made through the SANvergence Manager, these zones will be updated accurately. The Brocade, Brocade M-Series, Cisco, and QLogic zoning management utilities must *not* be used to alter these zones.
- When zones are created through Brocade M-Series Eclipse switches, they are appended to the active zone configurations on the other switches. These zones are not automatically saved to the zoning library on all switches in the SAN. There may be cases where updates can be made using more than one application and this could result in errors. To avoid this, after every zoning change made using Brocade M-Series Eclipse switch applications, the updated zone sets should be saved to the zoning libraries. Refer to documentation on Connectrix™ Manager, EMC ControlCenter™, or VisualSAN®, as applicable.

Zoning and device importation

Fibre Channel targets need to be imported from their respective fabrics. This is done using the SANVergence manager. By importing the devices, only the devices that you want to participate in the Brocade M-Series Router fabric are exposed and the other devices in the attached fabric are hidden.

In order to have an iSCSI initiator communicate with a Fibre Channel target it first needs to perform a discovery login. This discovery login is rejected but the initiator information is added to the SNS.

Create a zone with the iSCSI initiator and the Fibre Channel target using the SANVergence manager. At this point the iSCSI initiator needs to perform another discovery session and then continue with a normal session.

References

All Brocade M-Series documentation can be found at http://www.Brocade.com.

Please reference the *Brocade M-Series Eclipse User Guide* (1620 / 2640)for additional information regarding:

- Command Line Interface reference
- SANVergence Manager
- Recommended settings

Summary

Table 8 compares the iSCSI solution features available.

Table 8 iSCSI solution features comparison table

	Celerra	Brocade	MDS	Brocade M-Series	Symmetrix (N)	CLARiiON (N)
Jumbo frames	yes	yes	yes	yes	yes	no
O/S Support	everything but AIX	Refer to the *EMC Support Matrix*	Refer to the *EMC Support Matrix*	Refer to the *EMC Support Matrix*	Refer to the *EMC Support Matrix*	Refer to the *EMC Support Matrix*
# of initiator per port/box	256, but check the *EMC Support Matrix*	64/512 (This is per port/blade)	300/2000 Refer to the *EMC Support Matrix*	50/200 Refer to the *EMC Support Matrix*	256 = # of boards	10/10 Refer to the *EMC Support Matrix*
Proxy initiator	yes	yes	500/2000 Refer to the *EMC Support Matrix*	no	n/a	n/a
header/data digest	yes	yes	yes	yes	yes	yes
Immediate data	yes	yes	yes	yes	no	no
Initial R2T	yes	yes	yes	yes	no	no
Authentication	yes, CHAP	yes, CHAP	yes	yes	yes	yes
Encryption	no	no	no	no	no	no
PP Support	yes, for all supported environments	yes Refer to the *EMC Support Matrix* for all supported environments	yes Refer to the *EMC Support Matrix* for all supported environments	yes Refer to the *EMC Support Matrix* for all supported environments	yes Refer to the *EMC Support Matrix* for all supported environments	yes Refer to the *EMC Support Matrix* for all supported environments

The *EMC Support Matrix* is available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com.

3

Extended Distance

This chapter provides the following information on extended distance.

- Distance extension technologies overview 432
- Distance extension consideration 450
- IP-based distance extension solutions 471
- Summary 493

Distance extension technologies overview

To comprehend the distance extension solutions for Storage Area Networks it is important to understand and recall the challenges when implementing SAN connectivity over remote distances.

Early implementations of SAN environments

To increase a single port between two Fibre Channel switches separated by a large geographical distance, every two strands (transmit, receive) of optical fiber cable was required to be physically added by the distance provider. The customer would generally incur expensive construction, service, and maintenance costs when adding a bulk of fiber cables intended to satisfy current E_Port connectivity requirements while allowing future growth potential and redundancy against accidental fiber breaks. Existing fibers that were used for Ethernet implementations could not be shared and required separate dedicated channels per protocol. The challenges involved with this process would stem anywhere from mandatory to extraneous costs associated with fiber cable maintenance.

In addition to costs, there were physical hardware limitations to achieving connectivity between (at least) two geographically separated sites. Fibre Channel optics installed on the Fibre Channel switch was at the mercy of the limited optical output transmission power. Even with repeater technology, distortion of the optical wavelength transmitted by the optics can occur over several hops.

The Fibre Channel switches provided limitations as well. Link initialization and flow control were solely controlled by the Fibre Channel switches. The Fibre Channel standard would actually dictate the thresholds in regards to supporting large distances through optical connectivity and the obtainable bandwidth between two Fibre Channel ports.

To finalize the list of challenges that SAN environments had to overcome, each Fibre Channel switch provider had their own non-standard and standard ways of implementing their native environments. This may deviate from the mass interpretation of the Fibre Channel standards.

DWDM

Dense Wavelength Division Multiplexing (DWDM) is a process in which different channels of data are carried at different wavelengths over one pair of fiber-optic links. This is in contrast with a conventional fiber-optic system in which just one channel is carried over a single wavelength traveling through a single fiber.

Using DWDM, several separate wavelengths (or channels) of data can be multiplexed into a *multicolored* light stream transmitted on a single optical fiber (*dark fiber*). This technique to transmit several independent data streams over a single fiber link is an approach to opening up the conventional optical fiber bandwidth by breaking it up into many channels, each at a different optical wavelength (a different color of light). Each wavelength can carry a signal at any bit rate less than an upper limit defined by the electronics, typically up to several gigabits per second.

Different data formats being transmitted at different data rates can be transmitted together. Specifically, IP data, ESCON SRDF, Fibre Channel SRDF, SONET data, and ATM data can all be traveling at the same time within the optical fiber.

DWDM systems are independent of protocol or format, and no performance impacts are introduced by the system itself.

Figure 77 describes the DWDM technology concept:

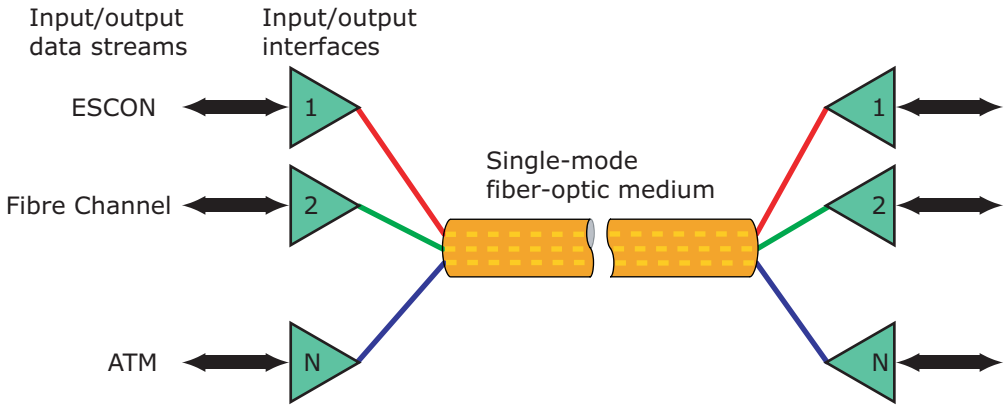


Figure 77 **DWDM example**

For EMC customers it means that multiple SRDF® channels and Fibre Channel ISL (Inter Switch Links) can be transferred over one pair of fiber links along with traditional network traffic. This is especially important where fiber links are at a premium. For example, a customer may be leasing fiber, so the more traffic they can run over a single link, the more cost effective the solution.

With today's technology, the capacity of a single pair of fiber strands is virtually unlimited. The limitation comes from the DWDM itself. Optical-to-electrical transfers for switching and channel protection are required and limit the input traffic per channel.

Available DWDM topologies include point-to-point and ring configurations with protected and unprotected schemas. DWDM technology can also be used to tie two or more metro area data centers together as one virtual data center.

DWDM systems can multiplex and de-multiplex a large amount of channel quantities. Each channel is allocated its own specific wavelength (lambda) band assignment. Each wavelength band is generally separated by 10 nm spacing(s). As optical technologies improve, separations between each channel may be further reduced enabling more channels to be packed (tighter) onto a single duplex dark fiber.

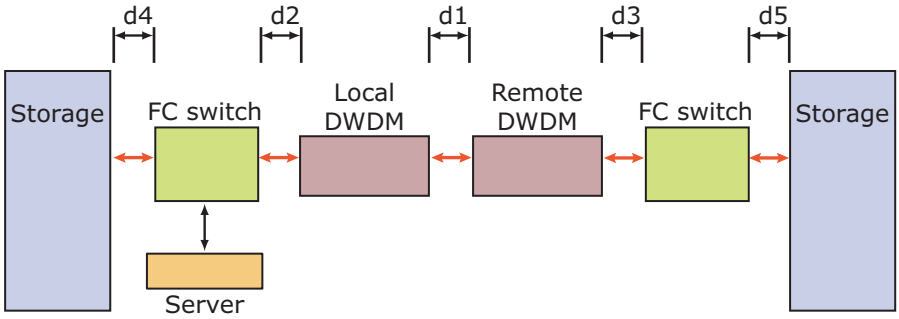
DWDM has a higher cost associated due to greater channel consolidation, flexibility, utilization of higher quality hardware precision-cooling components (to prevent low frequency signal drift) and the capabilities of regenerating, re-amplifying and reshaping (3R) wavelengths assigned to channels to ensure optical connectivity over vast distances.

Varying circuits pack capabilities are also offered in a DWDM environment. DWDM circuit packs / blades can provide the following protocol conversions:

- Fibre Channel to SONET
- Fibre Channel to Gigabit Ethernet
- Fibre Channel to IP

In addition, some circuit packs can enable features such as write acceleration and buffer-to-buffer credit spoofing. To verify the latest supported distance systems and features, refer to the *EMC Support Matrix*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com.

Figure 78 shows a general concept of Fibre Channel link extension using DWDM.



d1 = DWDM signal over dark fiber medium.
d2 and d3 = Local ISL connections between switches and DWDM input. Can be SM or MM depending on DWDM and switch interfaces or local distance requirements.
d4 and d5 = Local storage or server connections into the fabric.

Figure 78 **Fibre Channel link extension**

Note: All components are randomly selected and do not reflect a specific setup or configuration.

Note: Distance limitation may also be affected by application response time-out values and should consider signal propagation delay over site distance.

The following list provides general envelope guidelines for using DWDM systems:

- May be used for ESCON RDF distance extension, with direct connection between Symmetrix ESCON director ports and DWDM input ports.
- May be used for ISL extension of Fibre Channel switched fabrics. (E-Lab Navigator describes switch compatability.)
- Fabric topology guidelines are provided per Fibre Channel switch topology documentation.

- Direct connections between host HBA or Symmetrix Fibre Channel director to a DWDM port are not supported. E-Lab Navigator contains specific DWDM distance and topology guidelines.
- As a general approach, two distances need to be measured. The shorter of the two is the maximum distance to be supported in the site.

CWDM

Coarse Wave Division Multiplexing (CWDM), like DWDM, uses similar processes of multiplexing and de-multiplexing different channels by assigning different wavelengths to each channel. CWDM is intended to consolidate environments containing a low number of channels at a reduced cost.

CWDM contains 20 nm separations between each assigned channel wavelength. CWDM technology generally uses cost-effective hardware components that require a reduced amount of precision-cooling components usually dominant in DWDM solutions due to the wider separations. With CWDM technology the number of channel wavelengths to be packed onto a single fiber is greatly reduced.

CWDM implementations, like DWDM, utilize an optical-to-electrical-to-optical technology where all the channels are multiplexed into a single CWDM device performing the "optical-to-electrical-to-optical"conversion.

A CWDM connectivity solution can use optics generating a higher wavelength with increased output optical power. Each channel is designated its own specific wavelength by the specific hot-pluggable CWDM GBIC/SFP optic installed on the Fibre Channel Switches. With clean fibers, minimal patch panel connections, and ample optical power, CWDM optics alone can provide connectivity distances of up to 100 km per channel. To complete this solution a passive MUX/DEMUX is required to consolidate multiple channel-wavelengths into a single duplex 9-micron dark fiber.

Differences between DWDM and CWDM

The following are differences between DWDM and CWDM:

- Number of channels that are supported per solution.

  DWDM systems can support channels ranging from 16 channels or above while CWDM supports 16 channels or below.

- CWDM GBIC / SFP optics can be used to increase the wavelength output of a channel (i.e., FC-switch optics).

  The CWDM GBIC / SFP optics is usually installed in the Fibre Channel switch or client device. The wavelength and optical power enhanced links are then multiplexed and de-multiplexed to and from a single-mode 9-micron dark fiber.

- Costs.

  Hardware components included with DWDM units are higher in cost due to precision-cooling techniques required to prevent signal drift. DWDM offers greater channel flexibility and capacity.

- Configurations can be complex with CWDM.

  CWDM requires specific optics for each specific wavelength. Growth for a CWDM environment is limited and difficult to manage when supporting environments growing to larger channel support. More cabling would be required, thereby increasing complexity.

- DWDM devices offer circuit packs with numerous features (i.e., protocol conversions, buffer-to-buffer credit spoofing, write acceleration).

SONET

SONET (Synchronous Optical NETwork), is a standard for optical telecommunications transport, developed by the Exchange Carriers Standards Association for ANSI. SONET defines a technology for carrying different capacity signals through a synchronous optical network. The standard defines a byte-interleaved multiplexed transport occupying the physical layer of the OSI model.

Synchronization is provided by one principal network element with a very stable clock (Stratum 3), which is sourced on its outgoing OC-N signal. This clock is then used by other network elements for their clocks (loop timing).

SONET is useful in a SAN for consolidating multiple low-frequency channels (Client ESCON and 1, 2 Gb Fibre Channel) into a single higher-speed connection. This can reduce DWDM wavelength requirements in an existing SAN infrastructure. It can also allow a distance solution to be provided from any SONET service carrier, saving the expense of running private optical cable over long distances.

The basic SONET building block is an STS-1 (Synchronous Transport Signal), composed of the transport overhead plus a Synchronous Payload Envelope (SPE), totaling 810 bytes. The 27-byte transport overhead is used for operations, administration, maintenance, and provisioning. The remaining bytes make up the SPE, of which an additional nine bytes are path overhead. It is arranged as depicted in Figure 79. Columns 1, 2, and 3 are the transport overhead.

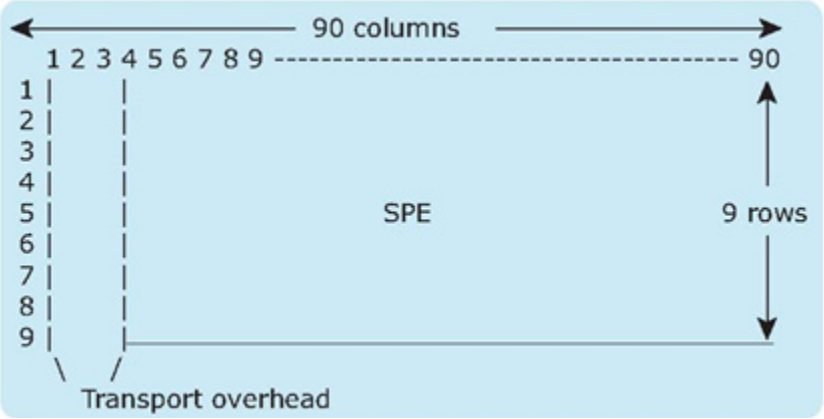


Figure 79 **STS-1 organization**

An STS-1 operates at 51.84 Mb/s, so multiple STS-1s are required to provide the necessary bandwidth for ESCON, Fibre Channel, and Ethernet, as shown in Table 9. Multiply the rate by 95 % to obtain the usable bandwidth in an STS-1 (reduction due to overhead bytes).

Table 9 **SONET/Synchronous Digital Hierarchy (SDH)**

STS	Optical Carrier	Optical Carrier Rate (Mb/s)
STS-1	OC-1	51.840
STS-3	OC-3	155.520
STS-12	OC-12	622.080
STS-48	OC-48	2488.320
STS-192	OC-192	9953.280

One OC-48 can carry approximately 2.5 channels of 1 Gb/s traffic, ss shown in Table 9. To achieve higher data rates for client connections, multiple STS-1s are byte-interleaved to create an STS-N. SONET

defines this as byte-interleaving three STS-1s into an STS-3, and subsequently interleaving STS-3s.

By definition, each STS is still visible and available for ADD/DROP multiplexing in SONET, although most SAN requirements can be met with less complex point-to-point connections. The addition of DWDM can even further consolidate multiple SONET connections (OC-48), while also providing distance extension.

GbE

Gigabit Ethernet (GbE) is a terminology describing an array of technologies involved in the transmission of Ethernet packets at the rate of 1024 megabits (Mb/s) or 1 gigabit per second. Gigabit Ethernet is specifically designed to surpass the traditional 10/100 Mb/s link speeds. GbE is defined by the IEEE publication 802.3z, which was standardized in June, 1998. This is a physical layer standard following elements of the ANSI Fibre Channel's physical layer. This standard is one of many additions to the original Ethernet standard (802.3 - Ethernet Frame) published in 1985 by the IEEE organization. The following are nomenclature and characteristics of GbE.

- 1000Base-SX is defined as a fiber optic Gigabit Ethernet standard encompassing the use of multi-mode (50 or 62.5 micron) fiber with 850 nanometer wavelengths. Distances of over 500 meters can be achieved.
- 1000Base-Lx is defined as a fiber optic Gigabit Ethernet standard encompassing the use of single-mode (9 micron) fiber with 1310 nanometer wavelengths. Distances of 10 Km or more can be achieved.
- Copper coaxial cabling, multi-mode fiber-optic cabling (50 & 62.5 micron) and single-mode (9 micron) cabling are available choices for 802.3z standard.
- GbE is mainly used in distance extension products as the transport layer for protocol such as TCP/IP. However, in some cases the product is based on a vendor-unique protocol.
- Distance products using GbE may offer features such as compression, write acceleration, and buffer credit spoofing

TCP/IP

The Transmission Control Protocol (TCP) is a connection-oriented transport protocol that guarantees reliable in-order delivery of a stream of bytes between the endpoints of a connection. TCP achieves this by assigning each byte of data a unique sequence number, maintaining timers, acknowledging received data through the use of acknowledgements (ACKs), and retransmission of data if necessary. Once a connection is established between the endpoints data can be transferred. The data stream that passes across the connection is considered a single sequence of eight-bit bytes, each of which is given a sequence number.

TCP terminology

This section provides information for TCP terminology.

Acknowledgements (ACKs)

The TCP acknowledgement scheme is cumulative as it acknowledges all the data received up until the time the ACK was generated. As TCP segments are not of uniform size and a TCP sender may retransmit more data than what was in a missing segment, ACKs do not acknowledge the received segment, rather they mark the position of the acknowledged data in the stream. The policy of cumulative acknowledgement makes the generation of ACKs easy and any loss of ACKs do not force the sender to retransmit data. The disadvantage is that the sender does not receive any detailed information about the data received, only the position in the stream of the last byte that has been received.

Delayed ACKs

Delayed ACKs allow a TCP receiver to refrain from sending an ACK for each incoming segment. However, a receiver should send an ACK for every second full-sized segment that arrives. Furthermore, the standard mandates a receiver must not withhold an ACK for more than 500 ms. The receivers should not delay ACKs that acknowledge out-of-order segments.

Maximum Segment Size (MSS)

The Maximum Segment Size (MSS) is the maximum amount of data, specified in bytes, that can transmitted in a segment between the two TCP endpoints. The MSS is decided by the endpoints, as they need to agree on the maximum segment they can handle. Deciding on a good MSS is important in a general internetworking environment because it affects the performance. It is difficult to choose a good MSS value because a very small MSS means an underutilized network, whereas a very large MSS means large IP datagrams that may lead to IP

fragmentation, greatly hampering the performance. An ideal MSS size would be when the IP datagrams are as large as possible without any fragmentation anywhere along the source to destination path. When TCP sends a segment with the SYN bit set during connection establishment, it can send an optional MSS value up to the outgoing interface's MTU minus the size of the fixed TCP and IP headers. For example, suppose the MTU is 1500 (Ethernet standard), then the sender can advertise a MSS of 1460 (1500 minus 40).

Maximum Transmission Unit (MTU)

Each network interface has its own Maximum Transmission Unit (MTU) that defines the largest packet it can transmit. The MTU of the media determines the maximum size of the packets that can be transmitted without IP fragmentation.

Retransmission

A TCP sender starts a timer when it sends a segment and expects an acknowledgement for the data it sent. If the sender does not receive an acknowledgement for the data before the timer expires, it assumes that the data was lost or corrupted and retransmits the segment. Since the time required for the data to reach the receiver and the acknowledgement to reach the sender is not constant (because of the varying Internet delays), an adaptive retransmission algorithm is used to monitor performance of each connection and conclude a reasonable value for timeout based on the round trip time.

Selective Acknowledgement (SACK)

TCP may experience poor performance when multiple packets are lost from one window of data. With the limited information available from cumulative acknowledgments, a TCP sender can only learn about a single lost packet per round trip time. An aggressive sender could choose to retransmit packets early, but such retransmitted segments may have already been successfully received. The Selective Acknowledgment (SACK) mechanism, combined with a selective repeat retransmission policy, helps to overcome these limitations. The receiving TCP sends back SACK packets to the sender informing the sender of data received and specifies the holes in the data that has been received. The sender can then retransmit only the missing data segments.

The Selective Acknowledgment extension uses two TCP options:

- An enabling option, **SACKpermitted**, which may be sent in a SYN segment to indicate that the SACK option can be used once the connection is established.
- The **SACK** option, which may be sent over an established connection once permission has been given by **SACKpermitted**.

TCP segment

The unit of transfer for TCP. The segments are used to establish a connection, transfer data, send ACKs, advertise window size and close a connection. Each segment is divided into three parts:

- Fixed header of 20 bytes
- Optional variable length header, padded out to a multiple of 4 bytes
- Data

The maximum possible header size is 60 bytes. The TCP header carries the control information. **SOURCE PORT** and **DESTINATION PORT** contain TCP port numbers that identify the application programs at the endpoints. The **SEQUENCE NUMBER** field identifies the position in the sender's byte stream of the first byte of attached data, if any, and the **ACKNOWLEDGEMENT NUMBER** field identifies the number of bytes the source expects to receive next. The **ACKNOWLEDGEMENT NUMBER** field is valid only if the ACK bit in the **CODE BITS** field is set. The 6-bit **CODE BITS** field is used to determine the purpose and contents of the segment. The **HLEN** field specifies the total length of the fixed plus variable headers of the segment as a number of 32-bit words. TCP software advertises how much data it is willing to receive by specifying its buffer size in the **WINDOW** field. The **CHECKSUM** field contains a 16-bit integer checksum used to verify the integrity of the data, the TCP header, and the header options. The TCP header padding is used to ensure that the TCP header ends and data begins on a 32-bit boundary. The padding is composed of zeros.

TCP window

A TCP window is the amount of data a sender can send without waiting for an ACK from the receiver. The TCP window is a flow control mechanism and ensures no congestion occurs in the network. For example, if a pair of hosts are talking over a TCP connection that has a TCP window size of 64 KB (kilobytes), the sender can only send 64 KB of data and must stop and wait for an acknowledgment from the receiver that some or all of the data has been received. If the receiver acknowledges that all the data has been received, the sender is free to send another 64 KB. If the sender gets back an acknowledgment from the receiver that it received the first 32 KB (which could happen if the second 32 KB was still in transit or it could happen if the second 32 KB got lost), the sender could only send another 32 KB since it cannot have more than 64 KB of unacknowledged data outstanding (the second 32 KB of data plus the third).

The primary reason for the window is congestion control. The entire network connection, which consists of the hosts at both ends, the routers in between, and the actual connections themselves, will have a bottleneck somewhere that can only handle so much data so fast. The TCP window throttles the transmission speed down to a level where congestion and data loss do not occur.

The factors affecting the window size are as follows:

Receiver's advertised window:
The time taken by the receiver to process the received data and send ACKs may be greater than the senders processing time, so it is necessary to control the transmission rate of the sender to prevent the sender from sending more data than the receiver can handle, thus causing packet loss. TCP introduces flow control by declaring a receive window in each segment header.

Sender's congestion window:
The congestion window controls the number of packets a TCP flow has in the network at any time. The congestion window is set using an Additive-Increase, Multiplicative-Decrease (AIMD) mechanism that probes for available bandwidth, dynamically adapting to changing network conditions.

Usable window:
This is the minimum of the receiver's advertised window and the sender's congestion window. It is the actual amount of data the sender is able to transmit. The TCP header uses a 16 bit field to report the receive window size to the sender. Therefore, the largest window that can be used is 2**16 = 65K bytes.

Window scaling:
The ordinary TCP header allocates only 16 bits for window advertisement. This limits the maximum window that can be advertised to 64 KB, limiting the throughput. RFC 1323 provides the window scaling option, to be able to advertise windows greater than 64 KB. Both the endpoints must agree to use window scaling during connection establishment.

The window scale extension expands the definition of the TCP window to 32 bits and then uses a scale factor to carry this 32-bit value in the 16-bit Window field of the TCP header (SEG.WND in RFC-793). The scale factor is carried in a new TCP option, Window Scale. This option is sent only in a SYN segment (a segment with the SYN bit on), hence the window scale is fixed in each direction when a connection is opened.

TCP error recovery

In TCP, each source determines how much capacity is available in the network so it knows how many packets it can safely have in transit. Once a given source has this many packets in transit, it uses the arrival of an ACK as a signal that some of its packets have left the network and it is therefore safe to insert new packets into the network without adding to the level of congestion. TCP uses congestion control algorithms to determine the network capacity. From the congestion control point of view, a TCP connection is in one of the following states.

- **Slow start:** After a connection is established and after a loss is detected by a timeout or by duplicate ACKs.
- **Fast recovery:** After a loss is detected by fast retransmit.
- **Congestion avoidance:** Congestion avoidance and slow start work hand-in-hand in all other cases. The congestion avoidance algorithm assumes that the chance of a packet being lost due to damage is very small. Therefore, the loss of a packet means that there is congestion somewhere in the network between the source and destination. Occurrence of a timeout and the receipt of duplicate ACKs indicate packet loss.

When congestion is detected in the network, it is necessary to slow things down, so the **slow start** algorithm is invoked. Two parameters, the congestion window (**cwnd**) and a slow start threshold (**ssthresh**), are maintained for each connection. When a connection is established, both of these parameters are initialized. The **cwnd** is initialized to one MSS. The **ssthresh** is used to determine whether the slow start or congestion avoidance algorithm is to be used to control data transmission. The initial value of **ssthresh** may be arbitrarily high (usually ssthresh is initialized to 65535 bytes), but it may be reduced in response to congestion.

The **slow start** algorithm is used when **cwnd** is less than **ssthresh**, while the **congestion avoidance** algorithm is used when **cwnd** is greater than **ssthresh**. When **cwnd** and **ssthresh** are equal, the sender may use either **slow start** or **congestion avoidance**.

TCP never transmits more than the minimum of **cwnd** and the receiver's advertised window. When a connection is established, or if congestion is detected in the network, TCP is in **slow start** and the congestion window is initialized to one MSS. Each time an ACK is received, the congestion window is increased by one MSS. The sender

starts by transmitting one segment and waiting for its ACK. When that ACK is received, the congestion window is incremented from one to two, and two segments can be sent. When each of those two segments is acknowledged, the congestion window is increased to four, and so on. The window size increases exponentially during slow start as shown in Figure 80.

When a timeout occurs or a duplicate ACK is received, **ssthresh** is reset to one half of the current window (i.e., the minimum of **cwnd** and the receiver's advertised window). If the congestion was detected by an occurrence of a timeout the **cwnd** is set to one MSS. When an ACK is received for data transmitted the **cwnd** is increased, but the way it is increased depends on whether TCP is performing **slow start** or **congestion avoidance**. If the **cwnd** is less than or equal to the **ssthresh**, TCP is in **slow start** and **slow start** continues until TCP is halfway to where it was when congestion occurred, then **congestion avoidance** takes over. **Congestion avoidance** increments the **cwnd** by the square of MSS divided by **cwnd** (in bytes) each time an ACK is received, increasing the **cwnd** linearly as shown in Figure 80. This provides a close approximation to increasing **cwnd** by, at most, one MSS per RTT.

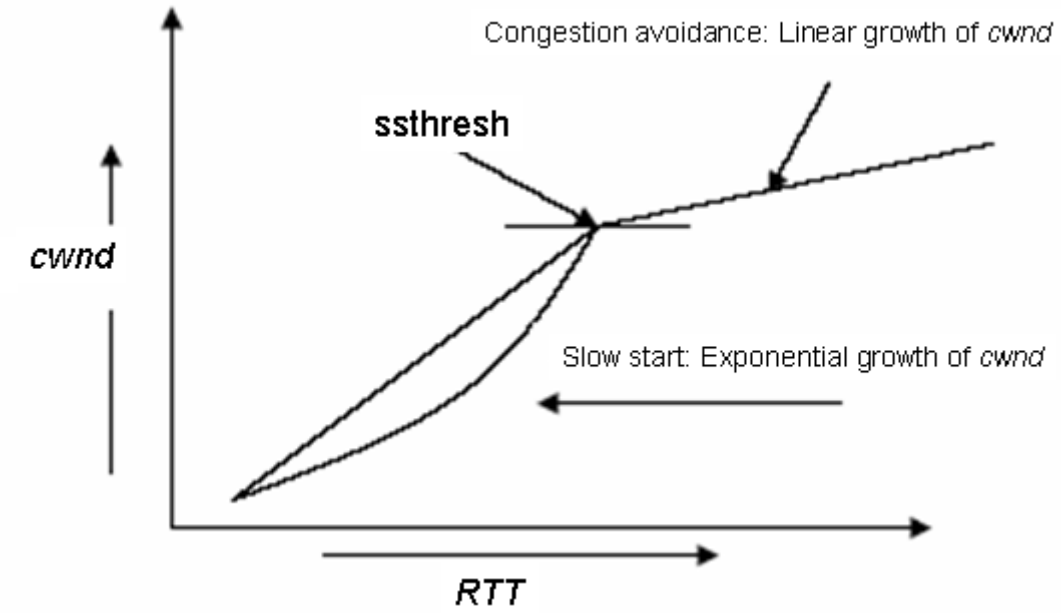


Figure 80 **Slow start and congestion avoidance**

A TCP receiver generates ACKs on receipt of data segments. The ACK contains the highest contiguous sequence number the receiver expects to receive next. This informs the sender of the in-order data that was received by the receiver. When the receiver receives a segment with a sequence number greater than the sequence number it expected to receive, it detects the out-of-order segment and generates an immediate ACK with the last sequence number it has received in-order (i.e., a duplicate ACK). This duplicate ACK is not delayed. Since the sender does not know if this duplicate ACK is a result of a lost packet or an out-of-order delivery, it waits for a small number of duplicate ACKs, under the assumption that if the packets are only reordered there will be only one or two duplicate ACKs before the reordered segment is received and processed and a new ACK is generated. If three or more duplicate ACKs are received in a row, it implies that there has been a packet loss. At that point, without waiting for the retransmission timer to expire, the TCP sender retransmits this segment. This is known as *fast retransmit (see* Figure 81 on page 447).

After fast retransmit has sent the supposedly missing segment, the **congestion avoidance** algorithm instead of the **slow start** is invoked; this is called *fast recovery*. Receipt of a duplicate ACK implies that not only is a packet lost, but that there is still data flowing between the two ends of TCP, as the receiver will only generate a duplicate ACK on receipt of another segment. Hence, fast recovery allows high throughput under moderate congestion.

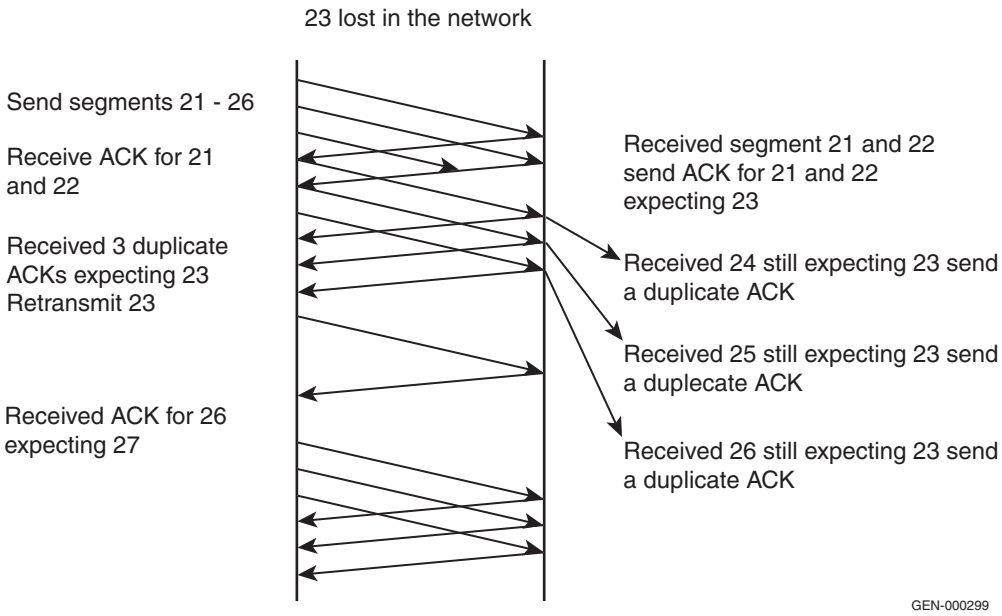


Figure 81 **Fast retransmit**

Network congestion

A network link is said to be congested if contention for it causes queues to build up and for packets to start getting dropped. The TCP protocol detects these dropped packets and starts retransmitting them, but using aggressive retransmissions to compensate for packet loss tends to keep systems in a state of network congestion even after the initial load has been reduced to a level which would not normally have induced network congestion. At such a time, demand for link bandwidth (and eventually queue space), outstrips what is available. When congestion occurs, all the flows that detect it must reduce their transmission rate. If they do not do so, the network will remain in an unstable state with queues continuing to build up.

IP security

IP security is a set of protocols developed by the IETF to support secure exchange of packets at the IP layer. IP security has been deployed widely to implement Virtual Private Networks (VPNs).

IP security supports two encryption modes:

- Transport
- Tunnel

Transport mode encrypts only the payload of each packet, but leaves the header untouched. The more secure tunnel mode encrypts both the header and the payload.

On the receiving side, an IP security compliant device decrypts each packet. For IP security to work, the sending and receiving devices must share a public key. This is accomplished through a protocol known as Internet Security Association and Key Management Protocol / Oakley (ISAKMP / Oakley), which allows the receiver to obtain a public key and authenticate the sender using digital certificates.

Tunneling and IPSec

Internet Protocol security (IPSec) uses cryptographic security to ensure private, secure communications over Internet Protocol networks. IPSec supports network-level data integrity, data confidentiality, data origin authentication and replay protection. It helps secure your SAN against network-based attacks from untrusted computers, attacks that can result in the denial-of-service of applications, services, or the network, data corruption, and data and user credential theft.

By default, when creating an FCIP tunnel, IPSec is disabled.

FCIP tunneling with IPSec enabled will support maximum throughput as follows:

- Unidirectional: approximately 104 MB / sec
- Bidirectional: approximately 90 MB / sec

Used to provide greater security in tunneling on an FR4-18i blade or a Brocade SilkWorm 7500 switch, the IPSec feature does not require you to configure separate security for each application that uses TCP / IP. When configuring for IPSec, however, you must ensure that there is an FR4-18i blade or a Brocade SilkWorm 7500 switch in each end of the FCIP tunnel. IPSec works on FCIP tunnels with or without IP compression (IPComp).

IPSec requires an IPSec license in addition to the FCIP license.

IPSec terminology

AES Advanced Encryption Standard. FIPS 197 endorses the Rijndael encryption algorithm as the approved AES for use by US government organizations and others to protect sensitive information. It replaces DES as the encryption standard.

AES-XCBC Cipher Block Chaining. A key-dependent one-way hash function (MAC) used with AES in conjunction with the Cipher-Block-Chaining mode of operation, suitable for securing messages of varying lengths, such as IP datagrams.

AH Authentication Header. Like ESP, AH provides data integrity, data source authentication, and protection against replay attacks but does *not* provide confidentiality.

DES Data Encryption Standard is the older encryption algorithm that uses a 56-bit key to encrypt blocks of 64-bit plain text. Because of the relatively shorter key length, it is not a secured algorithm and no longer approved for Federal use.

3DES Triple DES is a more secure variant of DES. It uses 3 different 56-bit keys to encrypt blocks of 64-bit plain text. The algorithm is FIPS-approved for use by Federal agencies.

ESP Encapsulating Security Payload is the IPSec protocol that provides confidentiality, data integrity, and data source authentication of IP packets, as well as protection against replay attacks.

MD5 Message Digest 5, like SHA-1, is a popular one-way hash function used for authentication and data integrity.

SHA Secure Hash Algorithm, like MD5, is a popular one-way hash function used for authentication and data integrity.

MAC Message Authentication Code is a key-dependent, one-way hash function used for generating and verifying authentication data.

HMAC A stronger MAC because it is a keyed hash inside a keyed hash. SA Security association is the collection of security parameters and authenticated keys that are negotiated between IPSec peers.

Distance extension consideration

This section provides information to consider when using distance extension.

- "Link speed", next
- "Data buffering and flow control" on page 451
- "TCP/IP" on page 461
- "Active and passive devices" on page 462
- "FC-SONET/GbE/IP" on page 469

Link speed

Link speed is an important aspect of distance extension configurations. Within the SAN networks link speeds equate to the amount of maximum bandwidth reachable on an E_Port and/or an F_Port. There are a variety of link speeds that are supported in a SAN network. Table 10 compares and contrasts the STS, optical carrier, and Fibre Channel link speed rates.

Table 10 STS-1s and optical carrier rates

STS	Optical carrier	Optical carrier rate	Fibre Channel link speeds
STS-1	OC-1	51.84 Mb/s	
STS-3	OC-3	155.52 Mb/s	
STS-12	OC-12	622.08 Mb/s	
STS-24	OC-24	1244.16 Mb/s	1.0625 Gb/s or 100 MB/s
STS-48	OC-48	2488.32 Mb/s	2.125 Gb/s or 200 MB/s
STS-96	OC-96	4976.64 Mb/s	4.250 Gb/s or 400 MB/s
STS-192	OC-192	9953.28 Mb/s	10.51875 Gb/s or 12.75 Gb/s

Data buffering and flow control

Fibre Channel

Fibre Channel uses the BB_Credit (buffer-to-buffer credit) mechanism for hardware-based flow control. This means that a port has the ability to pace the frame flow into its processing buffers. This mechanism eliminates the need of switching hardware to discard frames due to high congestion. EMC testing has shown this mechanism to be extremely effective in its speed and robustness.

BB_Credit management occurs between any two Fibre Channel ports that are connected. For example:

- One N_Port and one F_Port
- Two E_Ports
- Two N_Ports in a point-to-point topology
- In Arbitrated Loop different modes

The standard provides a frame-acknowledgement mechanism in which an R_RDY (Receiver Ready) primitive is sent from the receiving port to the transmitting port for every available buffer on the receiving side. The transmitting port maintains a count of free receiver buffers, and will continue to send frames if the count is greater than zero.

The algorithm is as follows:

1. The transmitter's count initializes to the BB_Credit value established when the ports exchange parameters at login.

   In an Arbitrated Loop environment the credits are established by the receiving port sending in advance R_RDY primitives after the login to establish the credit.

2. The transmitting port decrements the count per transmitted frame.
3. The transmitting port will stop sending frames when the credit reaches zero.
4. When a link reset occurs, the credit values are reestablished to values negotiated upon login.
5. The transmitting port increments the count per R_RDY it receives from the receiving port.

Figure 82 provides a view of the BB_Credit mechanism.

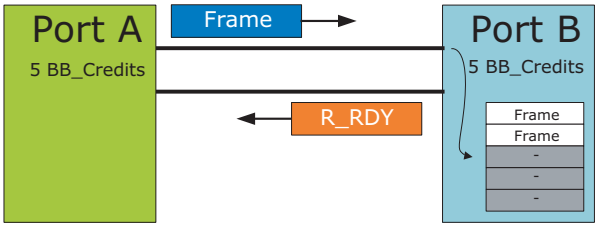


Figure 82 BB_Credit mechanism

As viewed from Port A's perspective, when a link is established with Port B, BB_Credit information is exchanged. In this case, Port B provided a BB_Credit count of 5 to Port A. For Port A, this means it can transmit upto five Fibre Channel frames without receiving an R_RDY.

Maximum supported distance per Fibre Channel BB_Credit guidelines

In order to achieve maximum utilization of the Fibre Channel link it is highly advisable that both ports, connected on either side of the long haul setup provided by the DWDM, be capable of high BB_Credit counts. Use the following formula to calculate the approximate BB_Credit(s) required for the specific long haul application.

To calculate for BB_Credits, use the following formula for calculating the required BB_Credit count:

Speed	Formula
1 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 1
2 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 2
4 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 4
10 Gb/s	BB_Credit=ROUNDUP [2 * one-way distance in km/4] * 12

The factor of **2** in the formulas accounts for the time it takes the light to travel the entire roundtrip distance: frame from transmitter to receiver and R_RDY back to transmitter.

Maximum allowable distance is based on optical power measurements of the site. These measurements should be approved by DWDM and fiber services provider(s). The distance between an

ISL ports on a Fibre Channel switch to a DWDM port should be included as part of the total distance (d1+d2+d3). (Refer to Figure 78 on page 435.)

The following BB_Credit charts will aid in providing estimates in regards to the amount of credits that should be present on the link when factoring Fibre Channel link speeds and link distances between the E_Ports.

Assuming the following is true:

- Light propagation in glass is 5 microsec / km, or 5^9 sec / m.
- Frame size is 2148 bytes / frame.
- Fibre Channel bit rate depends on the Fibre Channel speed.

Maximum distances assume 100% utilization of the ISL. If the ISL is not fully utilized, greater distances can be achieved since more BB_Credits become available. For example, for a 2 Gb / s switch port with 120 BB_Credits and with an ISL that is only 50% utilized, the maximum distance is 240 km.

Since Brocade's credit information is provided by ASIC types, please review Table 11 to correlate between switch ASIC and model numbers.

Table 11 Switch ASIC and model numbers

Vendor	ASIC/Family	EMC name	Vendor name
Brocade	Loom	Connectrix DS-16B	SilkWorm 2800
	Bloom	Connectrix DS-16B2	Silkworm 3800
	Bloom	Connectrix DS-32B2	SilkWorm 3900
	Bloom	Connectrix ED-12000B	SilkWorm 12000
	Bloom2	Connectrix ED-24000B	SilkWorm 24000
	Bloom2	Connectrix DS-16B3	Silkworm 3850
	Bloom2	Connectrix DS-8B3	SilkWorm 3250
	Condor	Connectrix DS-4100B	SilkWorm 4100
	Condor	Connectrix ED-48000B	SilkWorm 48000
	Condor	Connectrix DS-4900B	SilkWorm 4900
	Goldeneye	Connectrix DS-220B	SilkWorm 200

Table 11 Switch ASIC and model numbers (continued)

Vendor	ASIC/Family	EMC name	Vendor name
Cisco	Vegas (MDS generation 1)		9509
			9506
			9216
			9216A
			9216i
			9120
			9140
	Isolla (MDS generation 2)		9513
Brocade M-Series	Stitch	ED-1032	ED-5000
	Viper / Fuji-Shasta	DS-16M	ES-3016
		DS-16M2	ES-3216
		DS-32M	ES-3032
		DS-32M2	ES-3232
		ED-64M	ED-6064
		ED-140M	ED-6140
	Posideon/Teton	N/A	ES-4300
		DS-24M2	ES-4500
	Sanera	ED-10000M	Intrepid 10000
	Pegasus/Teton	DS-4400M	ES-4400
		DS-4700M	ES-4700

Buffer-to-Buffer credit information

Determining sufficient amount of Buffer-to-Buffer credits is crucial when provisioning Fibre Channel environments prior to utilization. Miscalculating the amount of credits may lead to less than desired performance (i.e., Buffer-to-Buffer credit, starvation, or backpressure).

Credit starvation occurs when the amount of available credits reaches a zero state preventing all forms of Fibre Channel IO-transmission from occurring. Once this condition is reached a timeout value will be triggered causing the link to reset.

Refer to the next sections for basic credit table for switches and storage arrays for Brocade, Cisco, and Brocade M-Series.

Brocade credit chart

With regards to flow control, Brocade switches support at least two forms of flow control options on the E_Port. VC_RDY and R_RDY flow control are both available options for all Brocade switch types.

For VC_RDY flow control, Brocade switches require an "Extended Fabric Mode" which will require to be activated through license code. Table 12, next, Table 13 on page 456, and Table 14 on page 456, are provided to display the supported distances for an E_Port when activating these modes in a Fibre Channel point-to-point switched fabric environment. These tables are broken down by ASIC type.

Table 12 Bloom and Bloom II ASICs

Mode	Description	Buffer allocation @ 1 Gb/s	Buffer allocation @ 2 Gb/s	Distance @ 1 Gb/s	Distance @ 2 Gb/s	Earliest Fabric OS release	Extended Fabric license required?
L0	Level 0 static mode; default	5	5	10 Km	5 Km	All	No
LE	Level E Static Mode;	13	19	n/a	10 Km	v3.x, v4.x	No
L0.5	Level 0.5 static mode	19	34	25 Km	25 Km	v3.1.0, v4.1.0, 5.x	Yes
L1	Level 1 static mode	27	54	50 Km	50 Km	All	Yes
L2	Level 2 static mode	60	65 / 108 for Bloom II	100 Km	60 Km 100 Km for Bloom II	All	Yes
LD	Dynamic mode; auto detects distance upon initialization	Auto	Auto	Auto (Max is 200 Km)	Auto (Max is 200 Km)	v3.1.0, v4.1.0, v4.4.0, 5.x – depending on model	Yes
LS	Static long distance mode (user specified)	User specified	User specified	User specified	User specified	v5.1.0	Yes

Table 13 Condor ASIC

Mode	Buffer allocation @ 1 Gb/s	Buffer allocation @ 2 Gb/s	Buffer Allocation @ 4 Gb/s	Distance @ 1 Gb/s	Distance @ 2 Gb/s	Distance @ 4 Gb/s	Earliest Fabric OS release	Extended Fabric license required?
L0	5	5	5	10 Km	5 Km	2 Km	All	No
LE	11	16	26	10 Km	10 Km	10 Km	3.x, 4.x	No
L0.5	18	31	56	25 Km	25 Km	25 Km	3.1.0, 4.1.0, 4.x, 5.x	Yes
L1	31	56	106	50 Km	25 Km	50 Km	All	Yes
L2	56	106	206	100 Km	100 Km	100 Km	All	Yes
LD	Auto	Auto	Auto	Auto (max 500 Km)	Auto (max 250 Km)	Auto (max 100 Km	3.1.0, 4.1.0, 4.x, 5.x – depending on model	Yes
LS	User specified	User specified	User specified	User specified (max 500 Km)	User specified (max 250 Km)	User specified (max 100 Km)	5.1.0	Yes

Table 14 Goldeneye ASIC

Mode	Buffer allocation @ 1 Gb/s	Buffer allocation @ 2 Gb/s	Buffer allocation @ 4Gb/s	Distance @ 1 Gb/s	Distance @ 2 Gb/s	Distance @ 4Gb/s	Earliest Fabric OS release	Extended Fabric license required?
L0	3	3	3	6 Km	3 Km	1.5 Km	All	No
LE	11	16	31	10 Km	10 Km	10 Km	3.x, 4.x	No
L0.5	18	31	56	25 Km	25 Km	25 Km	5.1.0	Yes
L1	31	56	106	50 Km	50 Km	50 Km	5.1.0	Yes
L2	56	106	n/a	100 Km	100 Km	n/a	5.1.0	Yes
LD	Auto	Auto	Auto	Auto	Auto	Auto	5.1.0	Yes
LS	User Specified	User Specified	User Specified	User Specified (max 293 Km)	User Specified (max 146 Km)	User Specified (max 73 Km)	5.1.0	Yes

Keep in mind that each Brocade switch family, ASIC, and mode type (i.e, L1, L2, LD, etc.) will have unique VC_RDY amounts and characteristics depending on specific fabric configurations. Please refer to the *EMC Support Matrix* for specific configuration information.

Brocade also supports R_RDY flow control (through Portcfgislmode). Brocade R_RDY mode can be activated when connecting to distance extension devices providing additional Buffer-to-Buffer Credits.

Brocade Extended Fabrics

Brocade's Extended Fabrics is a licensed feature that extends Storage Area Networks (SANs) across longer distances for disaster recovery and business continuance operations by enabling a modified buffering scheme in order to support long distance fibre channel extensions, such as MAN/WAN optical transport devices. This bulletin is suitable for external dissemination.

Flow control

The Fibre Channel standards specifications (for example, FC-PH and FC-SW) define a method of flow control called R_RDY to manage and control traffic as it flows across data links. Although the standards define how R_RDY flow control should be used, it does not prohibit the use of other vendor unique methods. By default, Brocade switches use Virtual Channel (VC) flow control over E_Port connections within a fabric.

VC flow control provides the following advantages over R_RDY:

- The ability to differentiate fabric internal traffic from end-to-end device traffic.

  In this case, switches generate fabric internal traffic which communicate state information to each other, such as link state information for routing, and device information for Name Service. This type of traffic is given a higher priority so that switches can distribute the most up-to-date information across the fabric even under heavy device traffic.

- The ability to differentiate data flows of end-to-end device traffic to avoid head-of-line blocking.

  In the case of (2), when there are multiple I/Os multiplexed over a single ISL, by assigning different VCs to different I/Os and giving them the same priority, each I/O can have a fair share of the bandwidth so that a large-size I/O will not consume the whole bandwidth and starve a small-size I/O, thus balance the

performance of the different devices communicating across the ISL. To identify a VC between two end-points of a link, VC_RDY is used.

Buffer allocation

When a switch port is configured for Extended Fabrics, additional credit is given to virtual channels that carry class 2 or 3 data traffic. This allows distances between switches to be extended over greater distances while maintaining maximum performance over ISLs. The Brocade Extended Fabrics license allows ISLs to be connected at up to 60 km for 2 Gbit/sec links and up to 100 km for 1 Gbit/sec links without degradation of performance.

When Extended Fabrics is enabled on Fabric OS v3.x and v4.x switches, two changes occur:

- Additional buffer credits are allocated to certain Virtual Channels on the long distance E_Port, and
- ARB(vc) is used as inter-frame gap instead of idles.

The additional buffers allow the E_Port "pipe" to be fully utilized over long distances and the ARB(vc) ordered set is used to notify the receiving switch as to which VC queue the next incoming frame should be placed on. There is a different ARB(vc) primitive for each of the eight possible virtual channels.

MAN/WAN optical transport devices

Vendors of optical transport devices may not be aware of E_Port functionality on Brocade switches, which may cause interoperability issues under certain configurations. Although there are certain workarounds, any vendor wishing to understand this functionality can contact Brocade. All devices tested in the Fabric Aware program are verified to operate under ideal switch configurations.

If the extension devices between the Brocade switches transparently propagate all traffic as is, these ARB(vc)s will not cause any problems. However, recently some transport devices have been introduced which do more than simply pass through the fibre channel frames. In some cases, and in some modes, these devices have been shown to have problems processing the ARB(vc) frames resulting in disruption of traffic over the long distance connection.

In these cases there are at least three solutions to this issue:

- If the extension device is capable of being configured in a mode which transparently passes fibre channel frames, there should be no disruption of traffic due to the ARB(vc) frames.
- If the 'fabric.ops.mode.longDistance' bit is set to '1' on all Brocade switches in the fabric, the ARB(vc) primitives will not be sent. The default setting of this parameter is '0'. In order to set this bit the switches will need to be disabled and the bit set using either the **configure** command in a telnet or serial console window or through a GUI management interface. In the Web Tools GUI this bit can be set by selecting the **Admin** button from the main screen and then clicking the enable button under **Extended Fabrics Mode** on the **Extended Fabric** tab. Despite the label of this button, it does not actually enable / disable Extended Fabrics and, in fact, the only effect this button has to set or unset the **fabric.ops.mode.longDistance** bit.

  Note: This parameter will need to be set on all switches in the fabric, not just the switch that has the long distance connection. Also note that this parameter affects all E_Ports on the switch (long distance or otherwise) by changing the amount of buffer credits allocated to the port.

- Since optical transport devices are designed to provide connectivity over long distance, many vendors provide their own method of managing flow control over long distance connections. This can allow FC performance to be maintained at up to hundreds or even thousands of kilometers without degradation. If the vendor supports this type of configuration, Brocade switches can be configured to use standards based R_RDY flow control using the **portCfgISLMode** CLI command. Extended Fabrics would not be necessary.

Note: The latest updated firmware levels and hardware levels support the combination of both Extended Fabric Modes with R_RDY mode implementation. This allows the customer to bypass the old challenges of configuring the Brocade Fabric environment to its pure native mode.

Please refer to the EMC Brocade switch documentation for further details.

Brocade M-Series credit chart

Brocade M-Series supports only R_RDY flow control. Each Brocade M-Series Family type switch will have unique credit amounts. Please refer to Table 15 for details of the Brocade M-Series credit chart.

Table 15 Brocade M-Series credit chart

Switch type (EMC/Brocade M-Series)	Module / Optic	Link speed	Number of credits	Notes
ED-1032 / ED-5000	Multi-mode, single-mode	1Gb / 2Gb	60	
DS-16M / ES -3016	Multi-mode, single-mode	1Gb / 2Gb	60	
DS-16M2 / ES-3216	Multi-mode, single-mode	1Gb / 2Gb	60	
DS-32M / ES3032	Multi-mode, single-mode	1Gb / 2Gb	60	
DS-32M2 / ES-3232	Multi-mode, single-mode	1Gb / 2Gb	60	
ED-64M / ED-6064	Multi-mode, single-mode	1Gb / 2Gb	60	
ED-140M / ED-6140	Multi-mode, single-mode	1Gb / 2Gb	60	
N/A / ES-4300	Multi-mode, single-mode	1Gb / 2Gb	12 / 7	12 on the first 4 and 7 on the rest… Credit increases applies to specified quad areas.
DS-24M2 / ES-4500	Multi-mode, single-mode	1Gb / 2Gb	12 / 7	12 on the first 4 and 7 on the rest… Credit increases applies to specified quad areas.
ED-10000M / Intepid 10000	Multi-mode, single-mode	1Gb/2Gb/10Gb	1373	
DS-4400M/ ES-4400	Multi-mode, single-mode	1Gb / 2Gb/ 4Gb	?	
DS-4700M / ES-4700	Multi-mode, single-mode	1Gb / 2Gb / 4Gb	?	

Cisco MDS credit chart

Cisco MDS switches only utilizes R_RDY flow control. Table 16 displays the number of BB-credits are available per E_Port.

Table 16 Cisco MDS credit chart

Switch Type	Blade/Optic Support	Link Speed	Number of Credits	Notes
9509	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	
9506	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	
9216	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	
9216A	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	
9216i	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	
9120	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	Based on the first quad
9140	Mult-mode, single-mode, CWDM	1Gb / 2Gb	255	

Symmetrix fibre adapter credit chart

Symmetrix boards uses R_RDY flow control. Table 17 displays the number of BB-credits available per Fibre Channel Adapter F_Port.

Table 17 Symmetrix fibre adapter credit chart

Symmetrix Family	Board Type / Optic	Link Speed	Number of Credits
Symmetrix 5.0	Fibre Adapter / multi-mode	1Gb / 2Gb	7
Symmetrix 6.0	Fibre Adapter / multi-mode	1Gb / 2Gb	7
Symmetrix 7.0	Fibre Adapter / multi-mode	1Gb / 2Gb	7

TCP/IP

TCP/IP window

A TCP window is the amount of data a sender can send without waiting for an ACK from the receiver. The TCP window is a flow control mechanism and sees to it that no congestion occurs in the network. For example, if a pair of hosts are talking over a TCP connection that has a TCP window size of 64 KB (kilobytes), the sender can only send 64 KB of data and then it must stop and wait for an acknowledgment from the receiver that some or all of the data has been received. If the receiver acknowledges that all the data has been received then the sender is free to send another 64 KB. If the sender gets back an acknowledgment from the receiver that it received the first 32 KB

(which could happen if the second 32 KB was still in transit or it could happen if the second 32 KB got lost), then the sender could only send another 32 KB since it cannot have more than 64 KB of unacknowledged data outstanding (the second 32 KB of data plus the third).

The primary reason for the window is congestion control. The whole network connection, which consists of the hosts at both ends, the routers in between, and the actual connections themselves, will have a bottleneck somewhere that can only handle so much data so fast. The TCP window throttles the transmission speed down to a level where congestion and data loss do not occur. The factors affecting the window size are as follows:

- Receiver's advertised window
- Sender's congestion window
- Usable window
- Window scaling

For more information on these factors refer to "TCP window" on page 442.

For complete overview of TCP/IP flow control please refer to "TCP/IP" on page 440.

Active and passive devices

Buffer-to-buffer local termination

In Fibre Channel, BB_Credits are a method of maintaining the flow control of transmitting Fibre Channel frames. BB_Credits help maintain a balanced flow of I/O transmissions while avoiding underutilization or oversubscription of a Fibre Channel link. Figure 83 on page 463 shows what the buffering flow control would normally follow without the local termination. This places the burden on the end nodes to maintain and track the BB-credit flow control on the Fibre Channel link.

The flow control distance will be determined by the amount of credits and the link speed that is supported by the end nodes. The end nodes can be an E_Port or F_Port.

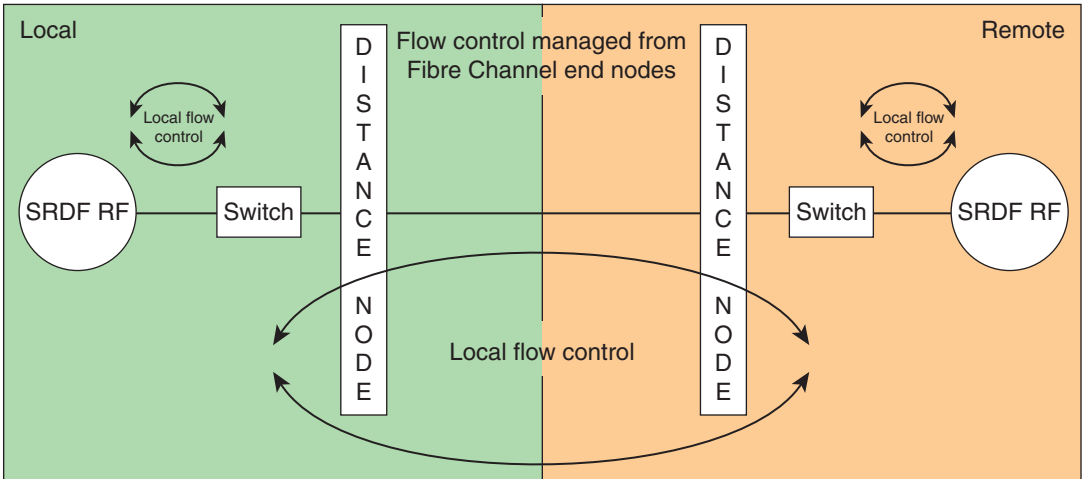


Figure 83 **Flow control managed by Fibre Channel switch (without buffering from distance extension devices)**

Distance extension devices may offer BB_Credit enhancements by maintaining the extended credits to the Fibre Channel end devices by providing an immediate R_RDY response. This is considered a form of local termination of credits where the distance device will maintain the flow control. Figure 84 illustrates the difference in flow control patterns when using the offerings from the distance device.

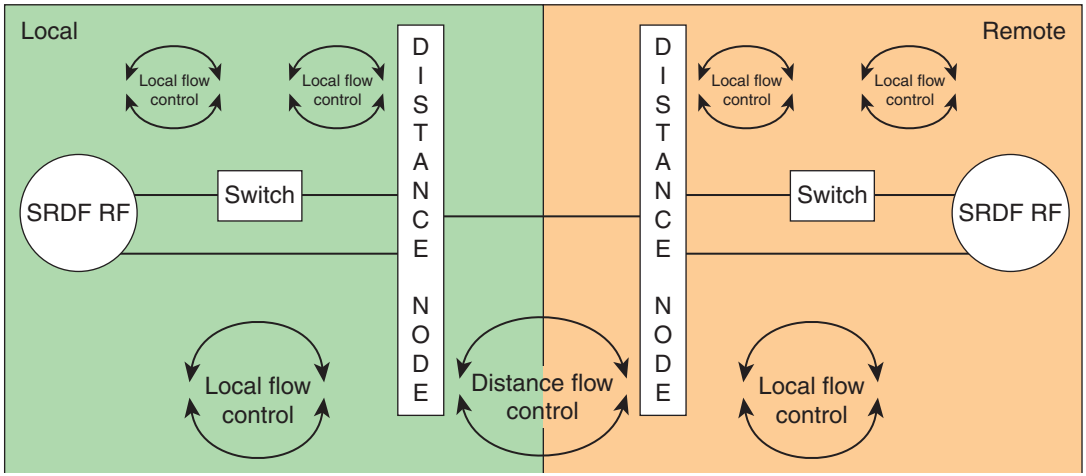


Figure 84 **Flow control (with buffering from distance extension devices)**

Write acceleration

Products with the write acceleration feature can increase SRDF throughput for direct-attach or Fibre Channel switched fabric configurations over extended distances. For SRDF synchronous and adaptive copy modes, immediate performance benefits are realized in a SCSI-write intensive SRDF-link. Figure 85 shows the normal write process as compared to Figure 86 on page 465 showing a write command with write acceleration.

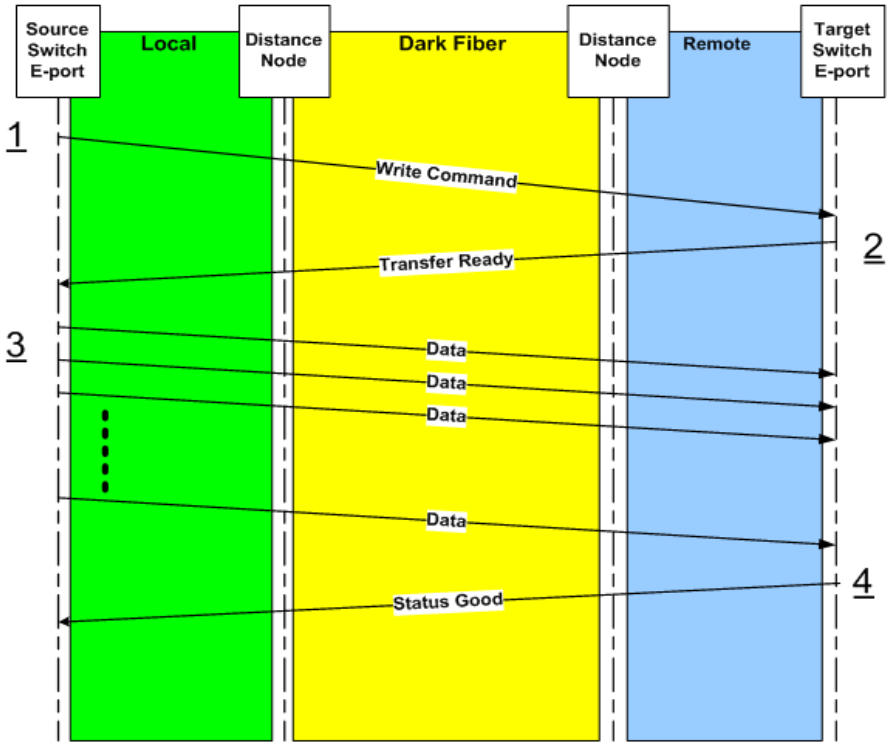


Figure 85 **Normal write command process**

A non-write accelerated write process is displayed in Figure 85. A write command request traverses across the dark fiber distance from the source switch to the target switch requesting for an acknowledgement which comes in the form of a transfer ready (X_RDY). If the write command request is received and there are no issues, the transfer ready is sent back to the source switch. Once this has been established the source switch will begin transmitting data

from the source switch to the target switch. Upon a successful error free completion of this step, the target switch will acknowledge back to the source switch with a "status good" ending the write request.

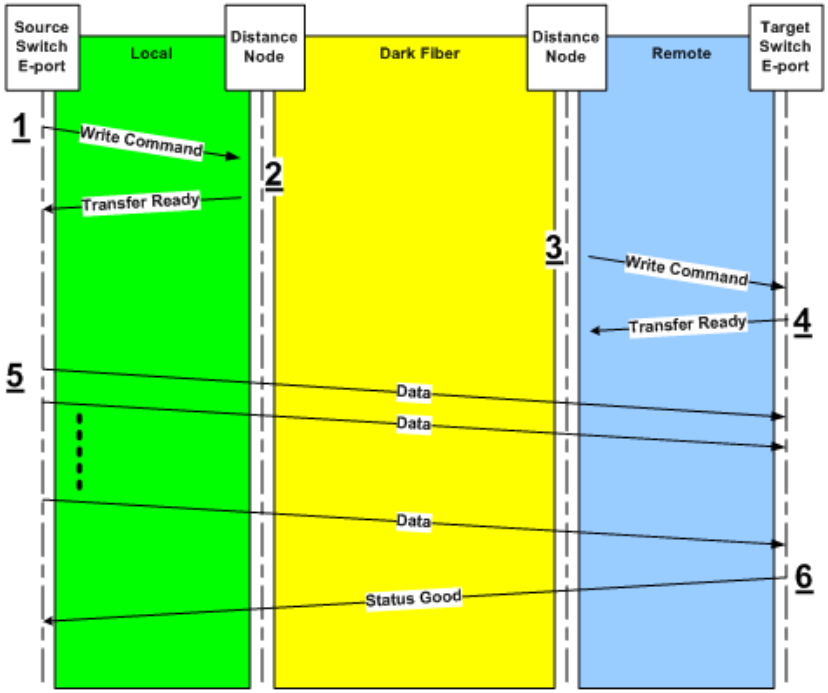


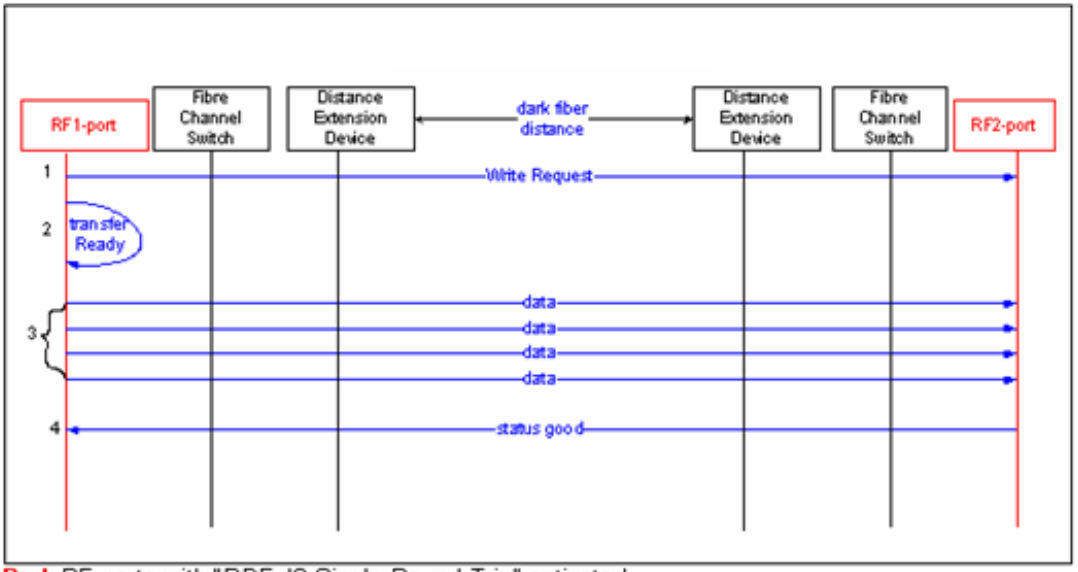
Figure 86 **Write command with write acceleration**

The write acceleration enabled diagram displays an immediate Transfer Ready (X_RDY) response to the source switch port once a recognized write command is received by the local distance node port (step 1, step 2). By having a localized response, additional latency times can be reduced by avoiding the X_RDY response originating from the remote target switch over the dark fiber distance. To further improve performance, once the localized X_RDY is received, data transmission can begin immediately (step 3a). The data transmissions should arrive closely after the X_RDY is sent from the remote switch (step 3b, step 4).

RDF JO Single Round Trip with distance vendor write acceleration

Beginning in August of 2007, EMC Rio Bravo microcode will enable an additional Fibre Channel SRDF synchronous feature known as "RDF J0 Single Round Trip".

Synchronous SRDF will achieve higher IO-throughput by issuing a (RF1-internal) local transfer-ready response for every RF1-port issued write command. This mode functions similarly to the write acceleration mode (described in "Write acceleration" on page 464) regarding distance extension devices with the exception that the emphasis is placed on the SRDF RF ports (F_Ports) only. Please refer to Figure 87.



Red: RF ports with "RDF J0 Single Round Trip" activated.
Blue: A step-by-step of a single write command with "RDF J0 Single Round Trip" enabled.

Figure 87 **RDF JO Single Round Trip**

The intended purpose of this feature is to maintain SRDF J0 synchronicity while improving performance by localizing the transfer ready response to the local RF Port, thereby reducing an unnecessary acknowledgement response (trip) over the dark fiber distance (step 2). Immediate benefits are apparent upon activation in a transparent WDM SRDF Synchronous distance extension environment. If applicable, multiple RDF Synchronous links can maximize their IO-performance over the transparent WDM environment.

In Figure 88, RF1 (R1 F_Port) and RF2 (R2 F_Port) are managing the "RDF J0 Single Round Trip" flow control.

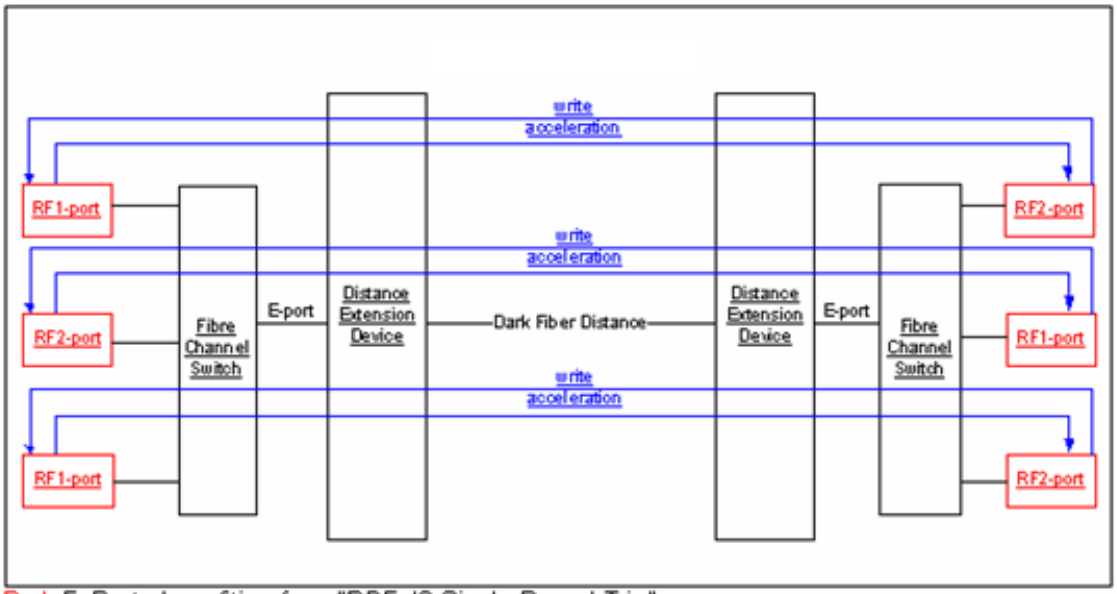


Figure 88 **RF ports benefit only**

With this in mind, RDF J0 Single Round Trip usage, in combination with the distance extension device-offered write acceleration mode, must be addressed. Essentially, for environments where the distance extension device is already servicing write commands on an E_Port level, it is recommended to disable RDF J0 Single Round Trip (refer to Figure 89 on page 468).

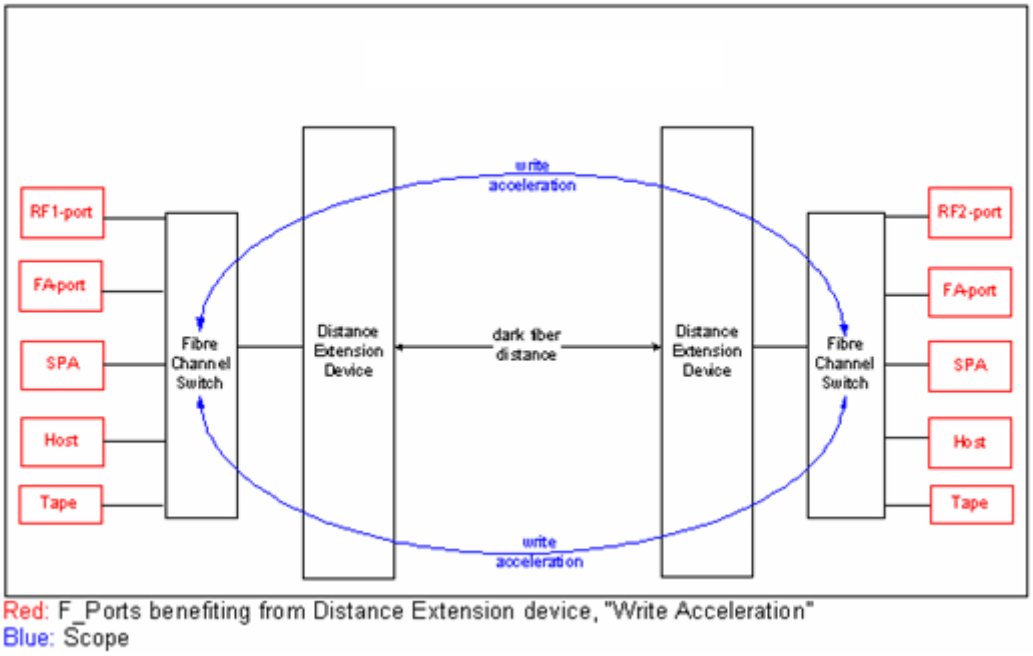


Figure 89 **All F_Ports will benefit**

As shown in Figure 89, by enabling the write acceleration feature on the distance extension device, potentially all F_Ports (RF ports, FA ports, SPA ports, Tape, etc.) issuing writes traversing across the E_Port attached to the distance extension client port can also take advantage of the throughput benefits from the activated write acceleration feature.

Link initialization

For link initialization of a Fibre Channel port, Fibre Channel specifications state that the maximum tolerable response time for a response is 100 milliseconds roundtrip time. This timeframe coincides with the limited timeframe of the Receiver-Transmitter Timeout Value (R_T_TOV), which is how long an FC-port listens for a link response to a link service before an error is detected.

FC-SONET/GbE/IP

Distance devices or circuit packs/blades performing protocol conversions from Fibre Channel to and from an alternate backbone protocol are required to maintain the lowest link initialization timeout value. In contrast to Fibre Channel's R_T_TOV, the SONET, GbE, and IP implementations can extend well beyond the 100 millisecond roundtrip time. For these environments, the distance extension devices should offer a setting enabling "local initialization" to occur between the "local" Fibre Channel port and the "local" distance extension client port rather than initializing the "local" Fibre Channel port across the actual physical distance to its "remote" Fibre Channel port (Figure 90).

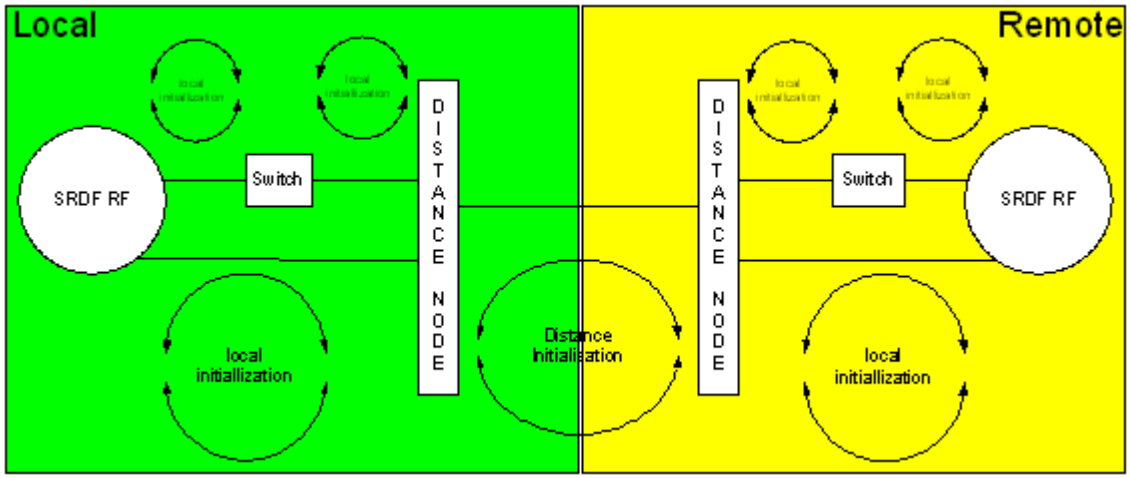


Figure 90 **Link initialization (More than 100 ms R_T_TOV)**

Network stabiliity

Error recovery

This section explains how the following handle error recovery.

CWDM:
CDWM devices do not participate in error recovery at any level. The device to handle the recovery depends on the level the error occurred. In case of link events, it will be handled by the Fibre Channel ports (switch or storage) across the CWDM link). In case of

SCSI level errors, the application (SRDF or MirrorView) will handle the error recovery. Link bit errors will cause SCSI level errors.

DWDM:
Error recovery is based on the attach client circuit pack that the Fibre Channel ports attached to. If the Fibre Channel ports attached to a Buffer-to-Buffer credit spoofing circuit then link events will be handled locally with the attached Fibre Channel port. SCSI level errors will be handled by the application. Link bit errors will cause SCSI level errors.

SONET:
Error recovery is based on the attach client circuit pack that the Fibre Channel ports attached to. If the Fibre Channel ports attached to a Buffer-to-Buffer credit spoofing circuit, link events will be handled locally with the attached Fibre Channel port. SCSI level errors will be handled by the application. Link bit errors will cause SCSI level errors.

GE:
Error recovery is based on the attach client circuit pack that the Fibre Channel ports attached to. If the Fibre Channel ports attached to a Buffer-to-Buffer credit spoofing circuit then link events will be handled locally with the attached Fibre Channel port. SCSI level errors will be handled by the application. Link bit errors will cause SCSI level errors.

TCP/IP:
Error recovery will be handled by the TCP/IP distance device (please review "TCP/IP" on page 440). If the errors persist and do not provide sufficient quality for the link to recover, the errors will be propagated to the attached Fibre Channel ports.

IP-based distance extension solutions

General best practices

Network design

The network should be dedicated solely to the IP technology being used and other traffic should *not* be carried over it.

The network must be well-engineered with *no* packet loss or duplication. This would lead to undesirable retransmission. While planning the network, care must be taken to ensure that the utilized throughput will never exceed the available bandwidth. Oversubscribing available bandwidth will lead to network congestion, which causes dropped packets and leads to TCP slow start. Network congestion must be considered between switches as well as between the switch and the end device. For more information on network congestion, refer to "Network congestion" on page 447.

The MTU must be configured based on the maximum available MTU supported by each component on the network.

Network conditions impact on effective throughput

Table 18 on page 472 demonstrates the impact of network conditions on TCP/IP effective throughput (data provided to the distance extension device by the Fibre Channel devices – the amount of data on the link will be greater due to TCP retransmission).

The distance between the sites have a significant impact on the distance system effective throughput. However, it is a fixed value. Packet loss, on the other hand is not a fixed value and can be relatively high due to TCP recovery mechanism and therefore has a greater impact. When designing the distance extension solution, network conditions must be taken into account to ensure that the effective throughput is sufficient for the solution needs. Over-utilization of the effective throughput will result in errors at the application levels.

Please review "TCP/IP" on page 461 for how to maximize effective throughput.

Table 18 Network impact on effective throughput example

Compression	Network conditions	Effective throughput
enabled	100ms RTT with 1% packet loss	1 MB/s
	100ms RTT with no packet loss	6 MB/s
	50ms RTT with 1% packet loss	3 MB/s
	50ms RTT with no packet loss	15 MB/s
	200ms RTT with 1% packet loss	800 KB/s
	200ms RTT with no packet loss	3.7 MB/s
disabled	100ms RTT with 1% packet loss	360 KB/s
	100ms RTT with no packet loss	5.4 MB/s
	50ms RTT with 1% packet loss	650 KB/s
	50ms RTT with no packet loss	12 MB/s
	200ms RTT with 1% packet loss	160 KB/s
	200ms RTT with no packet loss	2.98 MB/s

EMC-Brocade distance extension solution

The following Brocade terminology is used throughout this section.

Backbone Fabric

Routers provide a Backbone (BB) Fabric to interconnect routers for more scalable and flexible routed SANs. Each router may have many edge fabric connections, but only one BB fabric. Routers connect to the BB fabric through E_Ports, and all N_and NL_ port connections on a router are part of the BB fabric. With 4 Gb routers, a number of hosts and storage devices may be connected to the BB fabric.

Edge Fabric	Fibre Channel fabric connected to a router through an EX_Port (IFL). This is largely the same as any standard Fibre Channel fabric. This is, for the most part, where the hosts and storage are attached.
E_Port	A port on an FC switch or router, which connects to another switch or router, forming an ISL. If the devices previously formed separate fabrics, these fabrics merge, putting all fabric services into one distributed image.
EX_Port	FC Routers use EX_Ports instead of E_Ports on routed interfaces. To connect a router to a switch, you connect its EX_Port to another switch's E_Port using an appropriate cable. Routers still use E_ or VE_Ports to form a backbone fabric.
IFL	The connection between an E_Port and an EX_Port is an "Inter-Fabric Link".
ISL	The connection between two E_Ports is an Inter-Switch Link.
LSAN	LSAN Logical SANs are zones which span fabrics. They will traverse at least one EX_Port or VEX_Port. LSANs are how connectivity is configured across routers.
VE_Port	An FCIP port on an FC switch will create a "Virtual E_Port". This is physically an IP/Ethernet interface, but each FCIP tunnel "looks" like an FC E_Port to the rest of the fabric.
VEX_Port	In addition to supporting virtual E_Ports, Brocade platforms allow the FCIP and FC Router features to be combined, creating a Virtual EX_Port. FC Router features to be combined, creating a Virtual EX_Port.

Brocade SilkWorm AP7420

Brocade Silkworm Multiprotocol Router Model AP7420 can be used to route between remote fabrics across IP networks using the FCIP protocol. The AP7420 can be used to link Symmetrix and CLARiiON storage devices for disaster recovery applications (SRDF and MirrorView) and for data migration (SAN Copy). The AP7420 supports Fibre Channel and Gigabit Ethernet.

Supported configurations

Note: EMC does not support anything directly attached to the AP7420.

Supported configurations include XPath OS 7.4.1b, Extended Fabric, Trunking, Fabric Manager, Web Tools, IP over FC. The AP7420 supports host to target I/O, SRDF familiies, MirrorView, and SANCopy. The AP7420 only reaches 50 MB on the gigabit links.

Figure 91 shows an example of a Brocade AP7420 configuration.

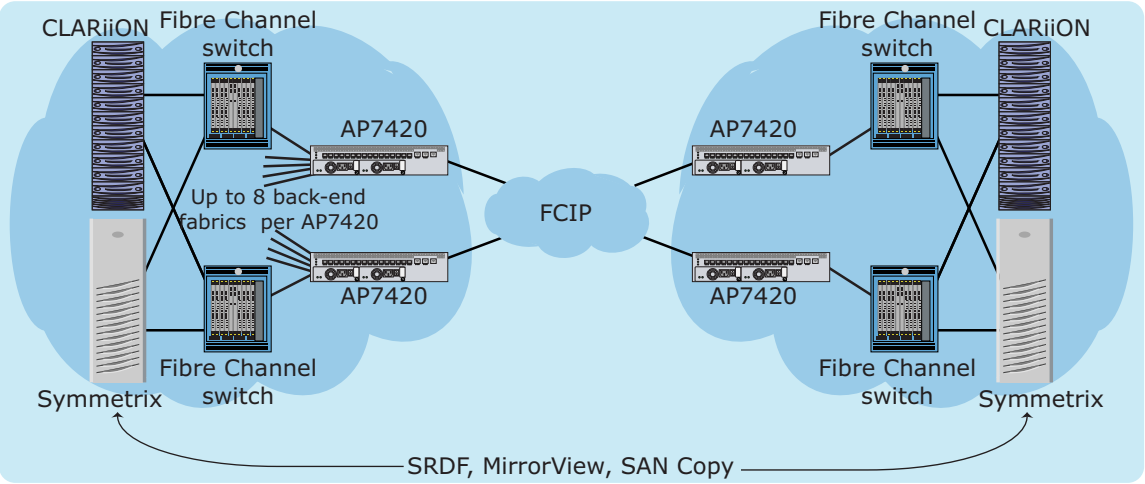


Figure 91 **Brocade Silkworm AP7420 distance extension example**

The platform operating system is based on the Brocade XPathTM Architecture, which is compatible with switches running Brocade Fabric OS. This allows users to enhance their existing SANs without costly "forklift" upgrades. The AP7420 can make switching decisions using any protocol layers up to the very top of the protocol stack. This means that the platform hardware is able to function as a standard Fibre Channel fabric switch, an FC-FC router, or a virtualization switch.

The platform has considerable flexibility, since every port has its own ASIC with multiple embedded CPUs and both Ethernet and FC protocol support. However, because it does not support 4 Gb FC, and has fewer hardware optimizations for routing, the SilkWorm 7500 is considered to be the future replacement for the AP7420 for FCIP and FC routing applications. The newer hardware on the 7500 and FR4-18i provides features which are not backward compatible. Mixing the AP7420 with the newer products in an FC-to-FC routing solution is supported, but can limit the utility of some new features such as backbone-to-edge routing.

Best practices

To make an analogy, an FC router is to an FC fabric as an IP router is to an Ethernet subnet. That is not a perfect analogy. In fact, an FC router is more like a switching firewall, or a layer 3 IP switch with hardware-enforced access control lists: it provides selective data connectivity between sites, through LSANs. (Logical SANs are zones which span fabrics.) However, FC routers prevent fabric services, faults, or administrative control regions from propagating between fabrics. A disruption in a fabric at either site, or of the WAN itself, will be contained and have less impact than would be the case without a router. Also, the feature allows each site to maintain independent fabric administration. This prevents fabric administrator errors from propagating between sites in addition to isolating failures in hardware or software. For example, if an administrator at a disaster recovery site accidentally deletes the fabric-wide zoning configuration, with FC routers in place, this will not propagate to the primary site.

Note: The AP7420 should be used only as a router in an FCIP configuration, and should not be directly connected (through Fibre Channel) to storage arrays.

Note these configuration rules:

- The AP7420 can be used as part of a DR (disaster recovery) and/or data migration SAN.
- FCIP can be used only as a backbone connection between the routers.
- SRDF, MirrorView, and SAN Copy are supported.
- Host I/O across the FCIP link can be supported if the application can tolerate the latency incurred due to the FCIP link.

- N_Port devices can be connected only to the Fibre Channel switches in the fabric connected to the router.

Note: E-Lab Navigator describes the latest supported configurations and minimum code requirements.

Symmetrix setup

Symmetrix SRDF ports should be configured as standard Fibre Channel SRDF ports. In a Fibre Channel environment, the AP7420 provides all the services of a Fibre Channel switch, similar to those provided by any other Fibre Channel switch.

CLARiiON setup

CLARiiON MirrorView ports should be configured as standard Fibre Channel MirrorView ports.

References

For additional information, refer to http://www.brocade.com/education_services/docs/reference_materials

Brocade SilkWorm 7500

FCIP tunneling enables you to connect one central office to different branch offices using different VE_Ports or VEX_Ports, thereby enabling branch offices to connect with each other without having to merge data center and branch office fabrics.

Fibre Channel frame encapsulation on one VE_Port and the reconstruction of Fibre Channel frames on the other VE_Port is transparent to the initiator and target, but the administration of VE_Ports is different from other Fibre Channel port types.

Fabric OS supports FCIP ISLs between two SilkWorm switches (SilkWorm 7500 or SilkWorm 48000 with a FR4-18i blade) or routers.

FCIP also supports:

- Configuration and management of GbE ports
- Compression and decompression of Fibre Channel frames moving through FCIP tunnels
- Statistics gathering on several layers
- Traffic shaping that adheres to a rate limit on a per tunnel basis
- FCIP tunnel/GbE port event notification
- Fibre Channel Router capabilities over VE_Ports

FCIP Tunneling introduces the following concepts:

- Tunnel

  An FCIP tunnel carries Fibre Channel traffic (frames) over IP networks such that the Fibre Channel fabric and all Fibre Channel devices in the fabric are unaware of the IP network's presence. Fibre Channel frames "tunnel" through IP networks by dividing frames, encapsulating the result in IP packets on entering the tunnel, and then reconstructing them as they leave the tunnel.

- VE_Port

  Special types of ports, called VE_Ports (virtual E_Port), function somewhat like an E_Port. The link between an VE_Port and a VE_Port is called an interswitch link (ISL). You can configure multiple ISLs from a SilkWorm 7500 or SilkWorm 48000 with an FR4-18i blade. After you configure the VE_Ports on either two SilkWorm 7500s or SilkWorm 48000s with the FR4-18i blade, an FCIP connection is established between them. VE_Ports do not prevent fabric merging. Using a VEX_Port is one way to prevent fabrics from merging.

- VEX_Port

  A VEX_Port enables routing functionality through an FCIP tunnel. VEX_Ports are virtual FC_Ports that are exposed by FCIP tunnels connecting to either the SilkWorm 7500 or SilkWorm 48000 with a FR4-18i blade; they run interfabric links (IFLs) as EX_Ports to enable Fibre Channel router capability. You can have up to 8 VEX_Ports per GbE on the SilkWorm 48000 with a FR4-18i blade.

- GbE

  Gigabit Ethernet ports are available on the SilkWorm 7500 and SilkWorm 48000 with a FR4-18i blade. These ports support FCIP with link speeds up to 1-Gbit/sec. Each GbE port (ge0, ge1) supports up to 8 FCIP tunnels.

Note: You cannot create more than one FCIP tunnel on a given pair of IP address interfaces (local and remote). However, that you can create multiple FCIP tunnels on an IP interface so that, minimally, either the local or remote IP interface will be unique and not have any other FCIP tunnel on it. When the GbE port has a valid SFP and is physically connected to any other GbE port, the status output from the **switchShow** command is online.

Supported environment

Figure 92 shows an example of a Brocade SilkWorm 7500 configuration.

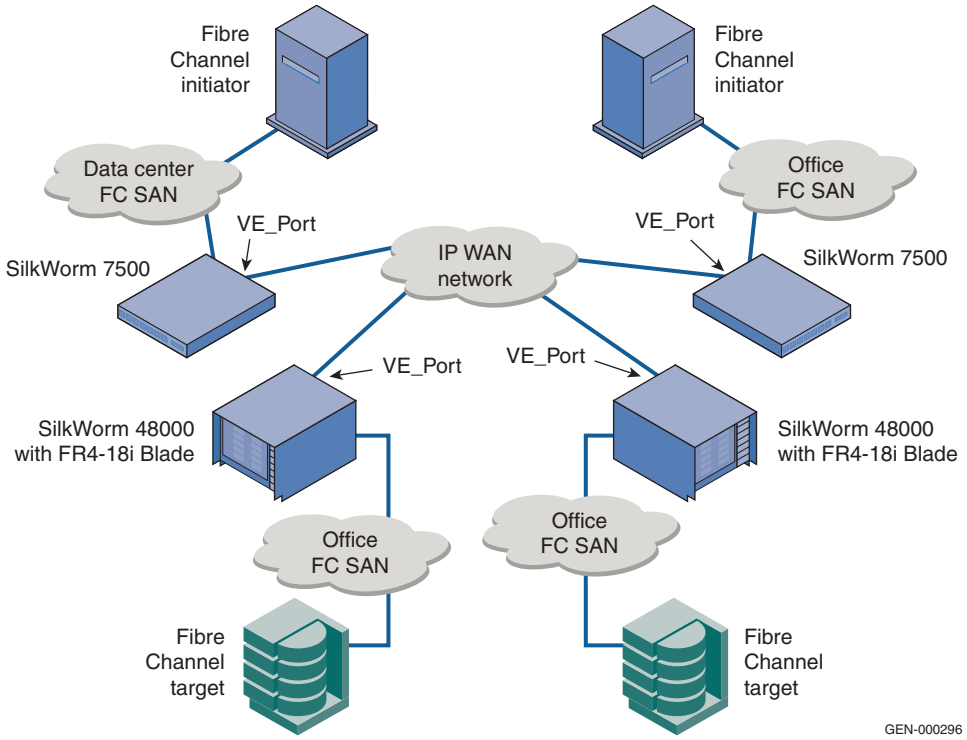


Figure 92 **Brocade SilkWorm 7500 configuration example**

References

For more information, refer to www.brocade.com. For configuration help, refer to the *FOS 5.1 Administration Guide.*

Configuring IPSec

IPSEC requires predefined configurations for IKE and IPSEC. You can enable IPSEC only when these configurations are well-defined and properly created in advance.

The following steps provide an overview of the IPSec protocol. All of these steps require that the correct policies have been created. Because policy creation is an independent procedure from FCIP tunnel creation, you must know which IPSec configurations have

been created. This ensures that you choose the correct configurations when you enable an IPSEC tunnel.

1. Some traffic from an IPSec peer with the lower local IP address initiates the IKE negotiation process.
2. IKE negotiates SAs and authenticates IPSec peers during phase 1 that sets up a secure channel for negotiation of phase 2 (IPSec) SAs.

   IKE negotiates SA parameters, setting up matching SAs in the peers. Some of the negotiated SA parameters include encryption and authentication algorithms, Diffie-Hellman group and SA lifetimes.
3. Data is transferred between IPSec peers based on the IPSec parameters and keys stored in the SA database.
4. IPSec tunnel terminates. SA lifetimes terminate through deletion or by timing out.

The first step to configuring IPSec is to create a policy for IKE and a policy for IPSec. Once the policies have been created, you assign the policies when creating the FCIP tunnel.

IKE negotiates SA parameters and authenticates the peer using the preshared key authentication method. Once the two phases of the negotiation are completed successfully, the actual encrypted data transfer can begin.

IPSec policies are managed using the **policy** command.

You can configure up to 32 IKE and 32 IPSec policies. Policies cannot be modified; they must be deleted and recreated in order to change the parameters. You can delete and recreate any policy as long as the policy is not being used by an active FCIP tunnel.

Each FCIP tunnel is configured separately and may have the same or different IKE and IPSec policies as any other tunnel. Only one IPSec tunnel can be configured for each GbE port.

Limitations

Be aware of the following limitations:

- IPv6, NAT, and AH are not supported.
- You can only create a single secure tunnel on a port; you cannot create a nonsecure tunnel on the same port as a secure tunnel.
- IPSec specific statistics are not supported.

- Fast Write and tape pipelining cannot be used in conjunction with secure tunnels.
- To change the configuration of a secure tunnel, delete the tunnel and recreate it with the desired options.
- Jumbo frames are not supported for IPSec.
- There is no RAS message support for IPSec.
- Only a single route is supported on an interface with a secure tunnel.

Fast Write and tape pipelining

In cases where the FCIP link is the slowest part of the network, and where this affects speed, consider using Fast Write and tape write acceleration (tape pipelining). Fast Write and tape pipelining are two individual features that provide accelerated speeds to FCIP tunnels in some configurations. Because of their similarities, they are both described in this section.

Supported only in Fabric OS 5.2.x and higher, Fast Write accelerates the SCSI write I/Os over FCIP.

Tape pipelining accelerates SCSI write I/Os to sequential devices (such as tape drives) over FCIP. This reduces the number of roundtrip times needed to complete the I/O over the IP network and speeds up the process. In order to use tape pipelining, you must enable Fast Write as well.

Both sides of an FCIP tunnel must have matching configuration for these features to work.

Compression, Fast Write, and tape pipelining features do not require any predefined configurations like IPSEC does. This makes it possible to enable these features when you create the FCIP tunnels by adding optional parameters such as –c, -f, or -t.

Table 19 on page 481 provides a comparison of Fast Write and tape pipelining.

Table 19 Fast Write and tape pipelining comparison

Fast Write	Tape pipelining
Does not support multiple equal-cost path configurations. (Refer to "Supported configurations" on page 481.)	Does not support multiple equal-cost path configurations or multiple non-equal-cost path configurations . (Refer to "Supported configurations" on page 481.)
Class 3 traffic is accelerated with Fast Write.	Class 3 traffic is accelerated between host and sequential device.
	With sequential devices (tape drives), there are 1024 initiator-type (IT) pairs per GbE Port, but 2048 initiator-tape-LUN (ITL) pairs per GbE Port. The ITL pairs are shared amont the IT pairs.[a] • Example 1: You can have two ITL pairs for each IT pair as long as the target has two LUNs. • Example 2: If a target has 32 LUNs, you can have 32 ITL pairs for IT pairs. In this case, only 64 IT pairs are associated with ITL pairs. The rest of the IT pairs are not associated to any ITP pairs, so no tape pipelining is performed for those pairs. By default, only Fast Write-based acceleration is performed on the unassociated pairs.
	Does not support multiple non-equal-cost path between host and sequential device.

a. Total of 2048 simultaneous exchanges combined for Fast Write and tape pipelining.

Supported configurations

To help understand the supported configurations, review the supported configurations shown in Figure 93 and Figure 94 on page 482. In Figure 93, there is a single tunnel with Fast Write and tape pipelining enabled.

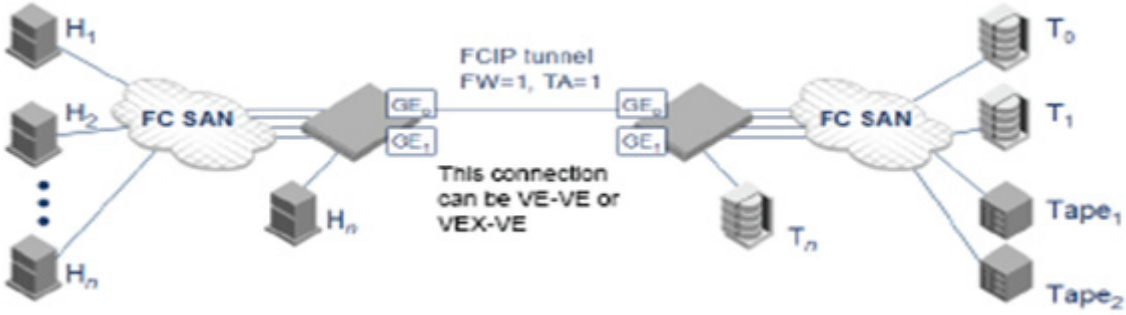


Figure 93 Single tunnel, Fast Write and tape pipelining enabled

In Figure 94, there are multiple tunnels, but none of them create a multiple equal-cost path. Fast Write and tape pipelining are enabled on a per-tunner, per-port basis.

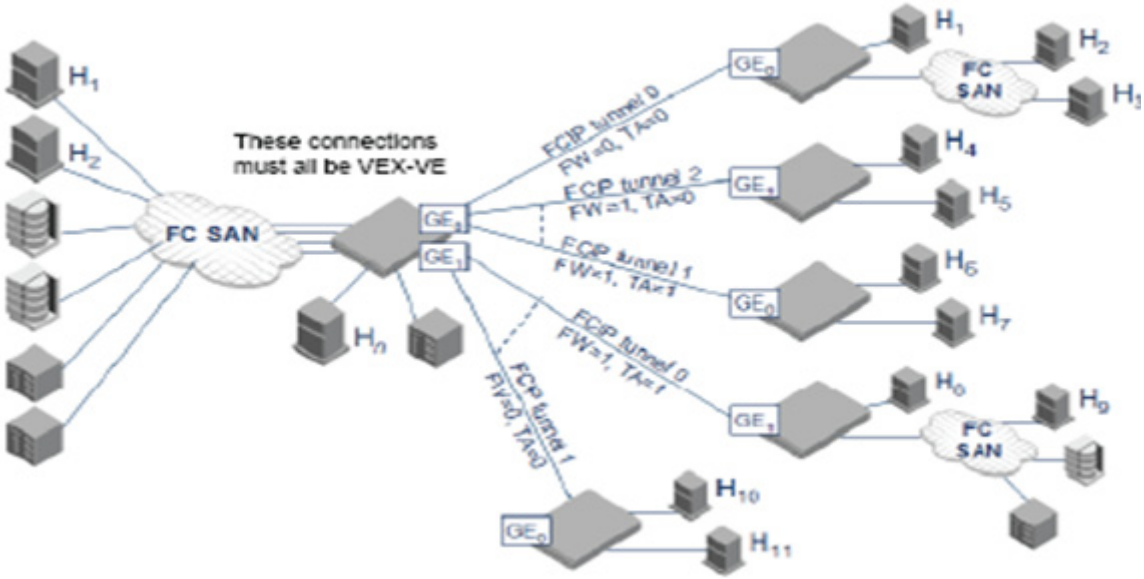


Figure 94 **Multiple tunnels to multiple ports**

EMC-Cisco MDS distance extension solution

The Cisco MDS 9000 family of switches can be used to link EMC storage devices (Symmetrix and CLARiiON) across IP networks using the FCIP protocol for disaster recovery applications (SRDF and MirrorView) and for data migration (SAN Copy). The MDS 9000 family supports the Fibre Channel and Gigabit Ethernet protocols.

Supported configurations

Figure 95 shows an example of Cisco MDS 9000 distance extension.

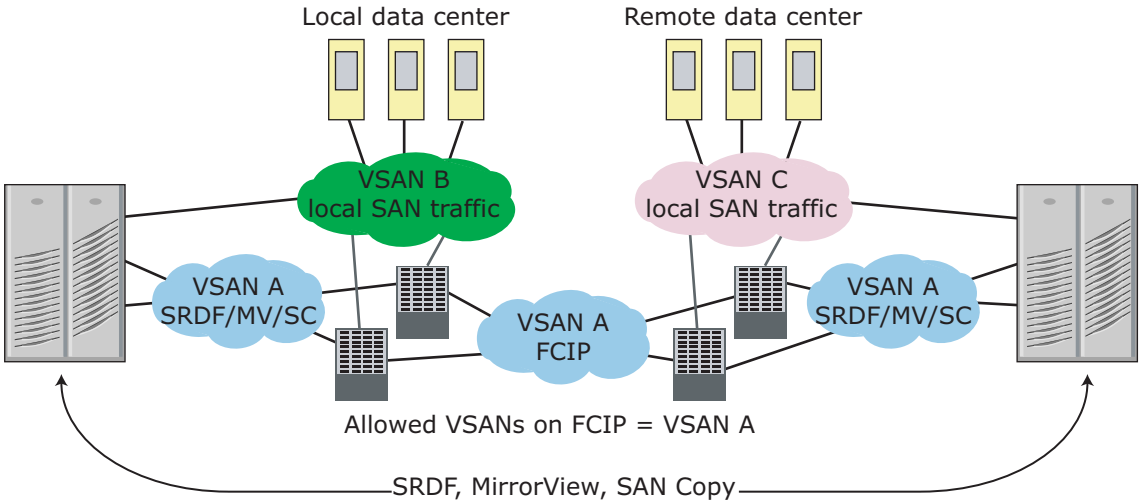


Figure 95 **Cisco MDS 9000 distance extension example**

Note these configuration rules:

- Cisco MDS switches can be used as part of a DR (disaster recovery) and / or data migration SAN only.
- SRDF, MirrorView, and SAN Copy are the only supported configurations.
- Remote host I / O configurations are supported across the FCIP link.
- Host I / O across the FCIP link can be supported if the application can tolerate the latency incurred due to the FCIP link

Note: E-Lab Navigator describes the latest supported configurations and minimum code requirements.

Symmetrix setup

Symmetrix SRDF ports should be configured as standard Fibre Channel SRDF ports. In a Fibre Channel environment, the Cisco MDS switch provides all the services of a Fibre Channel switch, similar to those provided by any other Fibre Channel switch.

CLARiiON setup

CLARiiON MirrorView ports should be configured as standard Fibre Channel MirrorView ports.

References

Additional documentation can be found at http://www.cisco.com. Search for the *Cisco MDS Configuration Guide* and select the document relevant to the code running on your box.

EMC-Brocade M-Series distance extension solution

Brocade M-Series Eclipse switches are multiprotocol switches. They support the bridging and routing for Fibre Channel and Gigabit Ethernet. In an FC environment, the switch provides all the FC services similar to those provided by any other FC switch.

The Brocade M-Series multiprotocol switches are also used in distance-extension configurations running SRDF or MirrorView over IP, for SAN Copy deployment over IP, or as a gateway device to perform FC-iSCSI translation. On these switches, ports 1 through 8 should be configured as FC ports (F_Ports or R_Ports). Ports 9 to 12 can be configured as HA ports, if want to have HA at the router. Ports 13 to 16 can be configured as either iFCP (TCP) ports (to cover long distances) or as iSCSI ports. There can be minimum of two HA links and maximum of 4 HA links between two Eclipse routers (2640). All the iFCP and iSCSI connections should be made on the primary router. HA is *not* supported on the Eclipse 1620 routers.

Supported configurations

Figure 96, Figure 97 on page 486, and Figure 98 on page 486 show examples of a Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment.

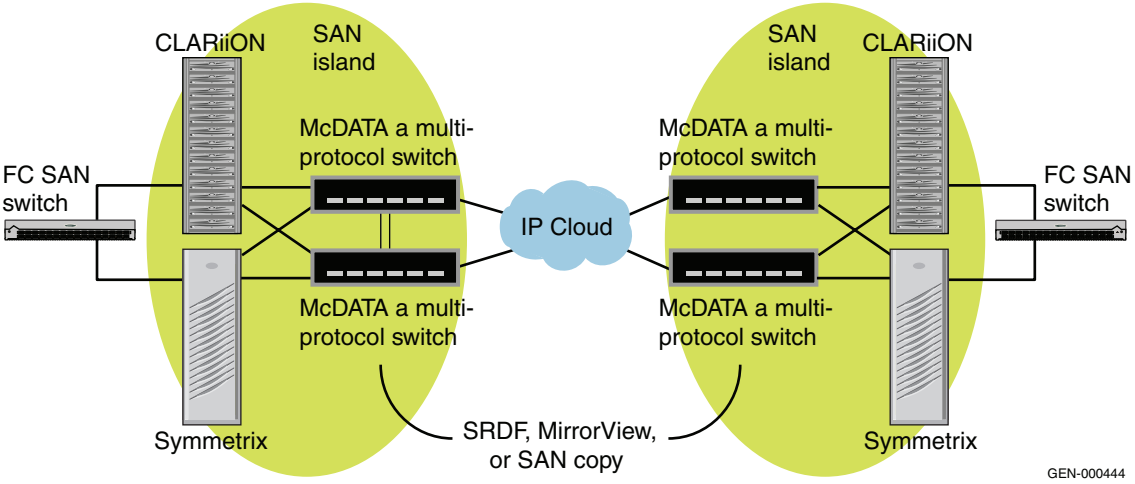


Figure 96 **Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment example 1**

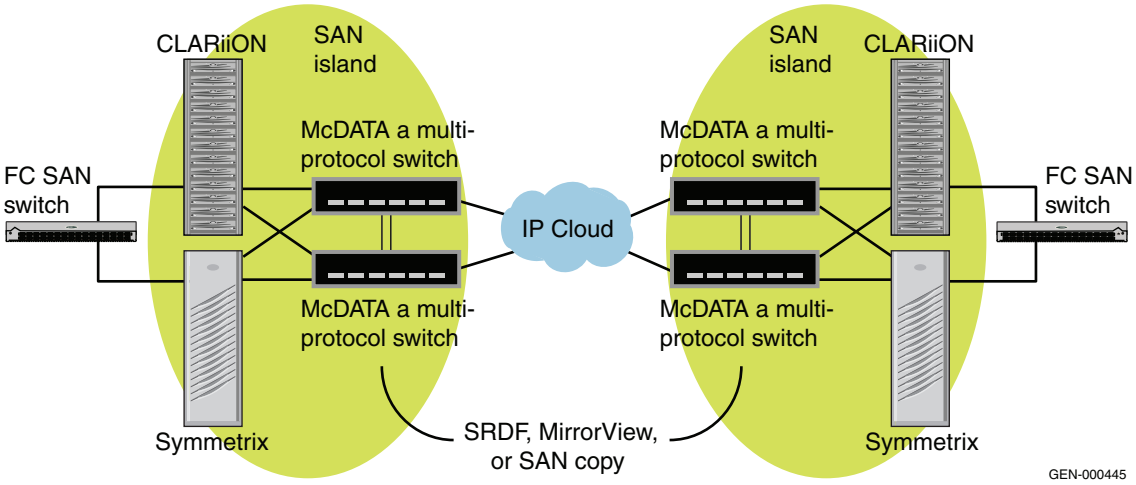


Figure 97 **Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment example 2**

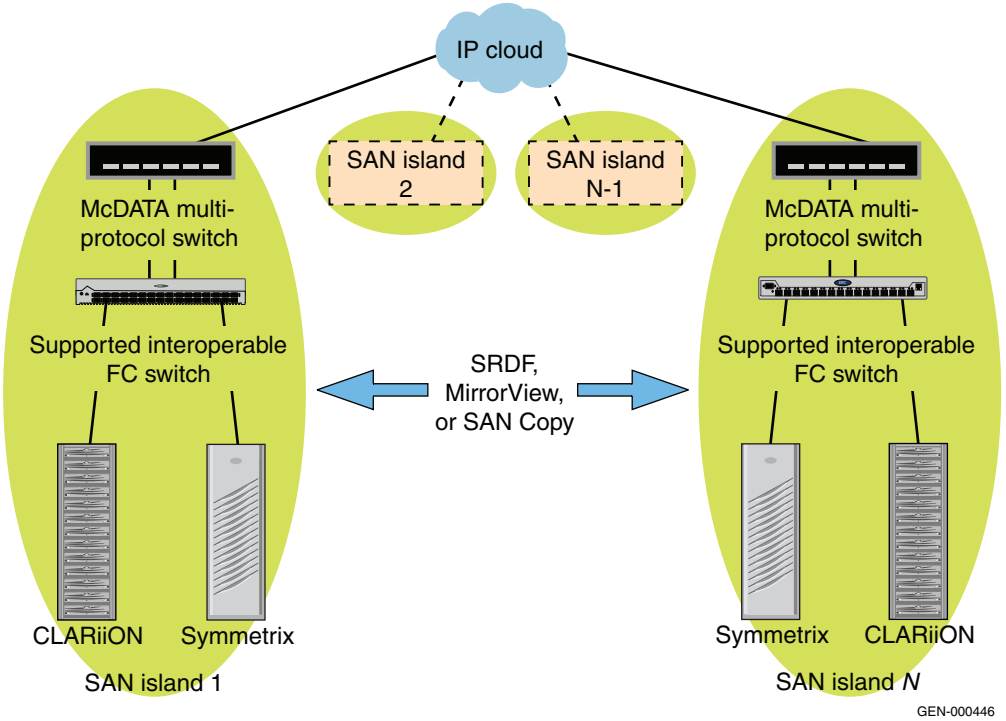


Figure 98 **Brocade M-Series in an SRDF, MirrorView, or SAN Copy environment example 3**

Implementation best practices

Consider the following best practices:

- Multivendor switches (Brocade, Brocade M-Series, Cisco, or QLogic) within a single SAN Island are not supported.
- There should be minimum 2 HA links and maximum of 4 HA links between two Brocade M-Series Eclipse switches.
- All the iFCP and iSCSI connections should be made on the primary router.
- A maximum of four interswitch links (ISLs) between the Brocade M-Series Eclipse switch and the Brocade / Brocade M-Series / Cisco / QLogic is supported.

- The Symmetrix (SRDF) or CLARiiON (MirrorView / SAN Copy) ports can be directly connected to the FC ports on the Brocade M-Series Eclipse switch, or to Brocade / Brocade M-Series / Cisco / QLogic switches in the local SAN.

Configuration of the local SAN ID and iFCP gateway

The local SAN ID is conceptually similar to the Domain ID used in FC, but is local to the switch and will not affect the 24-bit addresses assigned to any N_Ports connected to the switch. Brocade M-Series's *Eclipse 2640 Series User Manual* provides details for setting up this parameter.

The connection between SAN Islands is carried through the iFCP port (ports 13-16). Each iFCP port on the local site can be connected to iFCP ports on one or more remote sites. The association between the local site and the remote one is done through the iFCP Gateway definition. Refer to Brocade M-Series documentation for further details.

Multisite configuration: One local SAN Island to many remote SAN Islands

When designing a multisite configuration, note the following restrictions:

- The number of remote SAN Islands connected to one iFCP port on the local SAN Island will not exceed eight connections per iFCP port (eight entries per port in the Gateway definition table per iFCP port).
- The total number of FC initiator-target pairs, zoned across the iFCP link, will not exceed 16 per iFCP port.

All configuration changes must be saved to flash memory. Changes require a switch reboot to become active.

Symmetrix setup

For iFCP configurations, the Symmetrix ports should be configured as standard FC SRDF ports. The Brocade M-Series Eclipse switch, similar to any FC switch, provides FC switching services such as zoning, RSCN, name services, and so on, to the SRDF ports connected to it.

CLARiiON setup

For iFCP configurations, the CLARiiON ports should be configured as standard FC MirrorView or SAN Copy ports. The MirrorView application runs only in Synchronous mode.

Settings on Brocade/ Brocade M-Series/Cisco/ QLogic switches

While setting up the Brocade / Brocade M-Series / Cisco / QLogic switches to work with the Brocade M-Series Eclipse switch, the following settings are essential:

- The Brocade / Brocade M-Series / Cisco / QLogic switches must use the latest code supported by EMC for Interop Solutions. Refer to the *EMC Support Matrix*, "Edge Fabric Heterogeneity," for more information.
- The E_Port on the Brocade must be in Brocade native mode or Brocade M-Series mode (to which the Brocade M-Series Eclipse switch is connected) must be hard set to the actual port speed and autonegotiation must be disabled.
- The Cisco and QLogic switches must be set up to work in the Interop mode (Cisco interop mode 1 and QLogic open fabric).
- Once the zoning is defined on the Brocade M-Series Eclipse switches, it appears in the Brocade / Brocade M-Series / Cisco / QLogic active configurations as zones beginning with SoIP_xxx. When changes are made through the SANvergence Manager, these zones will be updated accurately. The Brocade, Brocade M-Series, Cisco, and QLogic zoning management utilities must not be used to alter these zones.
- When zones are created through Brocade M-Series Eclipse switches, they are appended to the active zone configurations on the other switches. These zones are not automatically saved to the zoning library on all switches in the SAN. There may be cases where updates can be made using more than one application and this may result in errors. To avoid this, after every zoning change made using Brocade M-Series Eclipse switch applications, the updated zone sets should be saved to the zoning libraries. Refer to documentation on Connectrix™ Manager, EMC ControlCenter™, or VisualSAN® as applicable.

Additional documentation

All Brocade M-Series documentation can be found at http://www.Brocade M-Series.com.

Please reference the Brocade M-Series Eclipse User Guide (1620 / 2640) for additional information regarding:

- Command Line Interface reference
- SANVergence Manager
- Recommended settings

EMC-QLogic distance extension solution

The QLogic iSR-6142 Storage Router is a low cost FC/iSCSI solution designed to enable users to replicate data between FC SANs over a LAN/WAN utilizing iSCSI/GigE as the transport over distance.

The router contains two 1/2 GB/s FC ports and dual 10/100/1000 MB/s iSCSI/GigE ports. The routers interconnect through the dual GigE/iSCSI links allowing the replication data to be transmitted between two end devices. The two routers allow up to 4 FC SANs to be connected as NL_ports (that is, 2 per router) and prevent the SANs from merging into one large SAN.

This router is intended for low to mid-range environments where distance extension and device replication, such as EMC's CLARiiON MirrorView software, are essential.

Supported configurations

The iSR-6142 Storage Router supports one distinct topology in an EMC environment:

WAN Topology – Interconnecting remote SAN Islands (also known as Remote SAN Island Connectivity).

The SANbox 6142 Intelligent Router supports inter-connecting remote SAN islands. This does not result in the merging of the two end fabrics but will allow communication to occur between two end nodes when correctly configured (Figure 99).

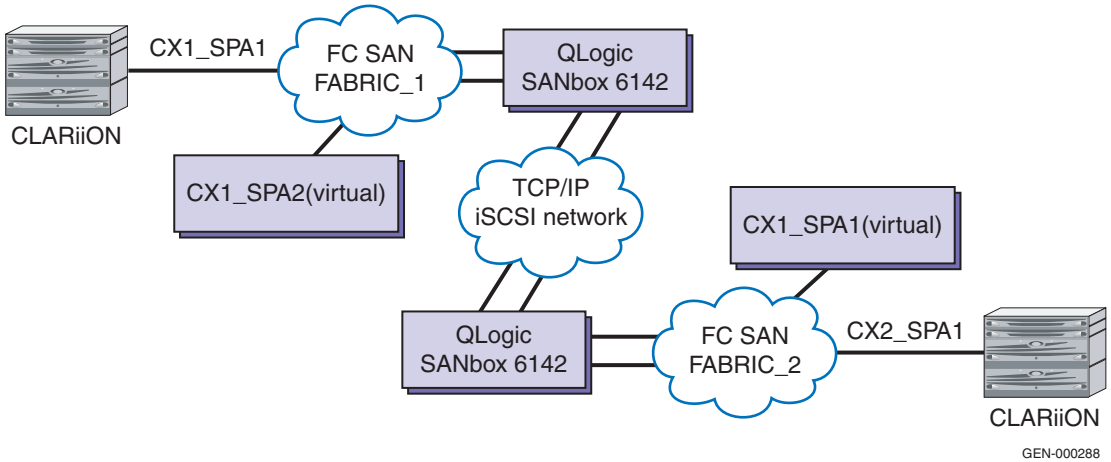


Figure 99 **SANbox 6142 Intelligent Router**

As shown in Figure 99, CX1_SPA1 (CLARiiON MirrorView port) is attached to Fabric_1. CX2_SPA1 (CLARiiON MirrorView port) is attached to Fabric_2. Using the QLogic SANbox 6142 it is possible to establish the communication between the Mirror view ports while maintaing two seperate Fabrics. the Qlogic Sanbox 6142 will create virtual entities on each fabric to represent the remote device. The mechanism to establish the connection is called remotemap. The remotemap is created using the CLI/GUI from either of the routers and is communicated to the remote router over the WAN. This remotemap presents CX1_SPA1 to Fabric_2 and CX2_SPA1 to Fabric_1 as an NL_Port. This NL_Port needs to be zoned local CX N_Port to allow communication between the two arrays over distance.

Scalability

The following are scalability guidelines, restrictions, and limitations:

- Maximum number of connections = 1024.
- Maximum number of virtual FC ports = 64 per unit (31 per FC port with 1 additional dedicated to each FC port for discovery - VP0 and VP1).
- Maximum number of concurrent I/O's = 1024 per unit (typically 32 per session).
- Maximum number of initiators/targets = 62 per unit (31 per port).

Best practices

Requirements for this configuration are as follows:

- At least one FC Port of the iSR-6142 should be connected to FC SAN.
- iSCSI/ GE Port IP addresses of remote router and iSCSI/GE port IP addresses of local routers must be accessible by each other.
- Remote iSR-6142 management port IP address and local SANbox 6142 management port IP address must be accessible by each other.

Recommendations for this configuration are as follows:

- Both GigE links are utilized with load balancing enabled.
- Compression is enabled over distance.
- Smart Writes enabled.
- Windows Scaling enabled with the recommended Windows Scaling Factor setting.
- Header and Data Digest enabled.
- Zone each N_Port that will have a remotemap to both of the router FC ports.
- Use WWPN zoning.

SmartWrite

When connecting SAN over long distances, round trip delays create significant impact to the performance. Typically, data writes involve two or more round trip latencies that result in a significant barrier to the data replication performance. SmartWrite™ technology is designed to minimize the round trip latency of any write IO to a single round-trip latency. Benefits realized with this feature key include:

- Minimizes round trip delays for any data write operation to a single round trip latency.

- Allows load balancing over multiple IP links.
- Provides failover and failback between two gigabit ethernet links.
- Allows data compression. This is very useful when data round trip latencies between two routers exceed more than 25 ms or long distance link rate is equal or less than 4500 Mb/s (DS-3 line rate).

References

For more information, refer to http://www.QLogic.com.

Please reference the *QLogic SANbox 6142 Intelligent Storage Router User Guide* for additional information regarding:

- Command Line Interface reference
- SANsurfer Router Manager GUI
- Recommended Windows Scaling Factor determined by latency between routers
- Hardware

Additional documentation regarding the QLogic SANbox 6142 Intelligent Storage Router includes:

- *QLogic SANbox 6142 Quick Start Guide*

Summary

Table 20 compares the distance extensions solutions features for TCP/IP products.

Table 20 Distance extension comparison table for TCP/IP products

Feature	Symmetrix (N)	Brocade	Cisco MDS	Brocade M-Series	QLogic
Fast Write	n/a	no	yes	yes	yes
Jumbo frames	yes	yes	yes	yes	no
Packet Loss/ roundtrip time	1%, 200 ms rtt	1%, 200 rtt	1%, 200 rtt	1%, 200 rtt	1%, 50 ms rtt
Encryption	no	no	yes	no	no
Applications	all families of srdf	srdf, srdfa, mva, mvs, sancopy	srdf, srdfa, mva, mvs, sancopy, ors	srdf, srdfa, mva, mvs, sancopy	mva,mvs,sancopy
Host I/O	n/a	yes	yes	yes	no
Protocols	tcp	fcip	fcip	ifcp	iscsi
Authentication	no	yes	yes	no	yes
Number of Sessions per port	64		1	64	32
Load Balancing	yes	yes	yes	no	yes
Compression	yes	yes	yes	yes	yes

Table 21 compares the distance extensions solution features for non TCP/IP products.

Table 21 Distance extension comparison table for non TCP/IP products

Distance extension chassis	Client/WDM/Protocol conversion						Link speed				Features					Switch vendor support			
	Client Side		WAN/ side/Line side				1 Gb	2 Gb	4 Gb	10 Gb	BBC	CLB	WA	FEC	COM	Brocade	Cisco	Brocade M-Series	QLogic
	FC-SW	FC-Direct	CWDM	DWDM	GbE	SONET													
ADVA FSP2000	X			X			X	X	X	X						X	X	X	
ADVA FSP3000	X			X			X									X	X	X	
Ciena CN2000	X	X	X		X	X	X	X			X	X			X	X	X	X	
Ciena CN4200	X			X				X	X	X						X	X	X	
Cisco ONS 15454	X			X		X	X	X	X	X	X			X		X	X	X	
Cisco ONS 15540	X			X			X	X								X	X	X	
Nortel 5200	X			X		X	X	X			X			X		X	X	X	
Nortel 3500	X	X				X	X	X			X			X		X	X	X	

Legend:

BBC:	BBC spoofing
WA:	Write Acceleration
CLB:	Channel Load Balancing
WA:	Write Acceleration
FEC:	Forward Error Correction
COM	Compression

4

Secure SANs

This chapter identifies and exemplifies some common SAN security attacks, presents some built-in and bolt-on mechanisms to enhance SAN security, and provides some insight on how to implement various product-specific security mechanisms.

- Understanding how to secure your SANs 496
- Security attacks against SANs..... 498
- Secure SAN architecture, components, and mechanisms..... 502
- Building secure SANs..... 514
- Security implementation examples 516
- EMC Select security appliances 535

Understanding how to secure your SANs

Note: This chapter discusses security in a data storage environment. Some readers may associate the acronym SANS with System Administration Networking and Security, while others are more familiar with SANs being defined as storage area networks. This chapter will use the latter definition of storage area networks whenever the acronyms SAN or SANs are used.

Valuable information such as intellectual property, personal identities, and financial transactions is routinely processed through, and stored in, a storage area network (SAN). The security of a SAN is critical because of the value of the extractable information. Designing, testing, and securing SANs is IT intensive, but necessary, to manage and protect vital information. Without proactive security measures, outsiders can easily obtain information from your SAN.

To secure your SAN you should:

- Assess configurations
- Test secure mechanism effectiveness
- Identify holes
- Quantify risks
- Implement practical security solutions for high risk exposures.

SAN configurations using block data include both Fibre Channel (FC) and Internet SCSI (iSCSI) protocols. Basic security *concepts* such as authentication, authorization, administration, and encryption are similar for these protocols, but *mechanisms* to protect Fibre Channel and iSCSI SANs may differ. Refer to "Secure SAN architecture, components, and mechanisms" on page 502 for more information on concepts and mechanisms.

The practice of interconnecting storage in the SAN can bridge several internet (TCP/IP) and security (such as DMZ, departmental, and corporate) zones that are purposefully segmented through host network access. Figure 100 on page 497 shows how a SAN can bridge external and internal security segments.

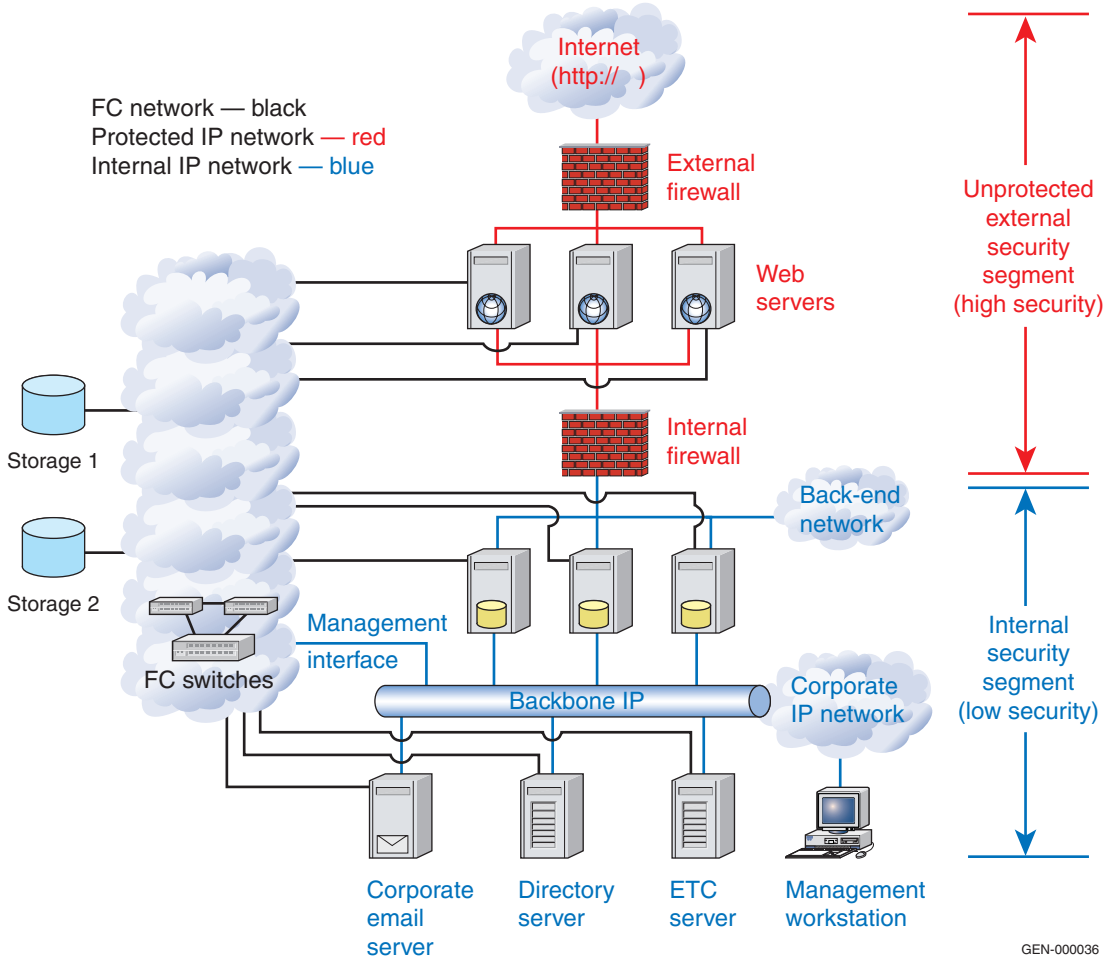


Figure 100 SAN security bridges

Security attacks against SANs

Security attacks against SANs are similar to security attacks against IP networks. Breaches of security can include breaches of authorization, authentication, data confidentiality, and/or data integrity. iSCSI SANs and Fibre Channel SANs have similar security flaws, including significant weaknesses with authentication and authorization.

Examples provided in this section describe some common security attacks in a SAN. These include:

- "Snooping (Eavesdropping)", next
- "Spoofing (Modification/ Fabrication)" on page 500
- "Denial-of-service (Disruption)" on page 500

Several other attacks such as Fibre Channel Name Server/iSNS pollution, session hijacking, zone hopping, E_Port replication, F_Port replication, LUN mask subversion, iSCSI Qualifier Name (IQN) spoofing, CHAP username sniffing, CHAP hash spoofing, and SAN management attacks can also expose a tremendous amount of information. These attacks can either provide the attacker with needed information for a subsequent attack, or directly compromise data. In many cases a breach of security involves a combination of security attacks.

For information on components and mechanisms to prevent security attacks, refer to "Secure SAN architecture, components, and mechanisms" on page 502.

Snooping (Eavesdropping)

Snooping is a deliberate act to access data without authorization. Different methods include, but are not limited to, eavesdropping, sniffing, session hacking, intercepting, copying, and monitoring.

Figure 101 on page 499 shows an example of snooping in a Fibre Channel SAN. Similar snooping methods can be accomplished in an iSCSI SAN. In this example, the name server table is modified by an unauthorized management session. The rogue Node C will have its Source ID (S_ID) and World Wide Names (WWN) strategically placed in the name server and zone tables. As a result, the switch forwards frames from the initiator (Node A) to the rogue Node C. This type of attack can be prevented with the implementation of security

mechanisms to prevent unauthenticated and unauthorized manipulation of the Fibre Channel name server or iSNS. (Refer to "Secure SAN architecture, components, and mechanisms" on page 502).

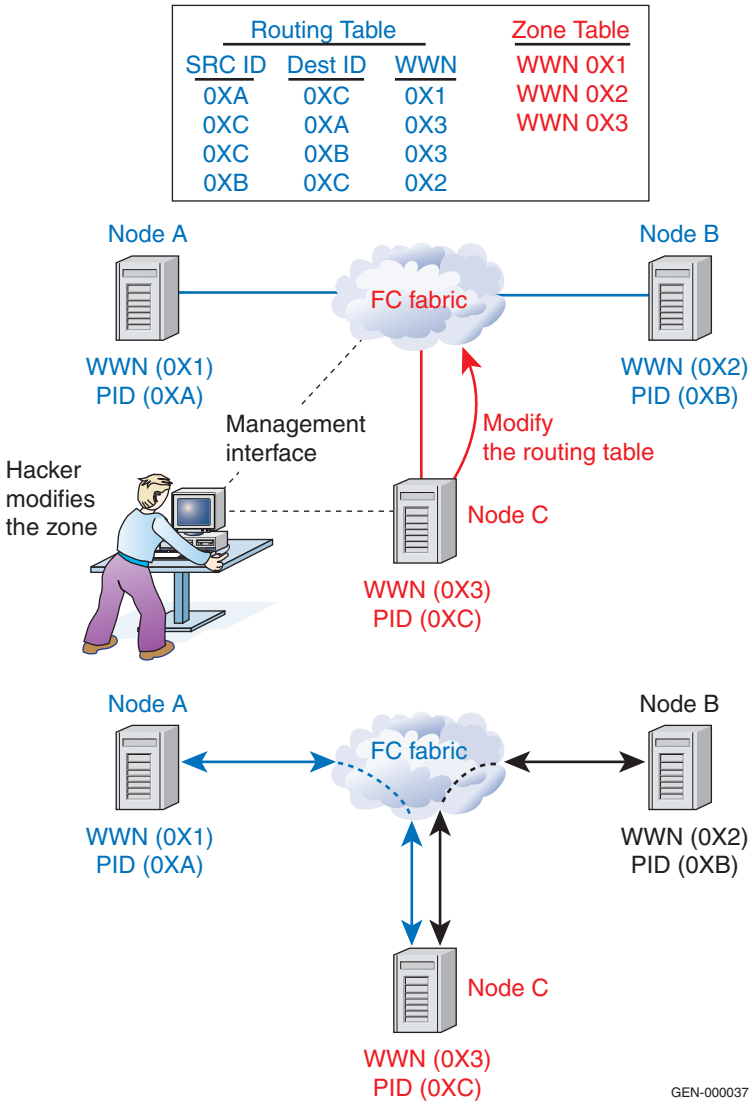


Figure 101 Snooping in a SAN

Spoofing (Modification/ Fabrication)

Spoofing is a deliberate act to assume an identity in order to gain unauthorized access to the data of the company or another user. WWNs are used to identify nodes in a Fibre Channel SAN, whereas in an iSCSI SAN a node is identified by an iSCSI Qualified Name (IQN). Without proper security mechanisms in place, both are easy to change and spoof. Some methods modify part of the information on the fly or use native host bus adapter (HBA) utilities to change the node identity.

Figure 102 shows an example of how to implement a spoofing attack in either a Fibre Channel or iSCSI SAN. In this example, the node WWN of the FC HBA is modified to match another HBA node WWN that is authenticated to log in to a storage port. Similarly, in iSCSI SANs, the IQN value is easily changed through native HBA utilities. Fortunately, this type of attack can be mitigated with the implementation of appropriate security mechanisms and configuration changes. (Refer to "Secure SAN architecture, components, and mechanisms" on page 502.)

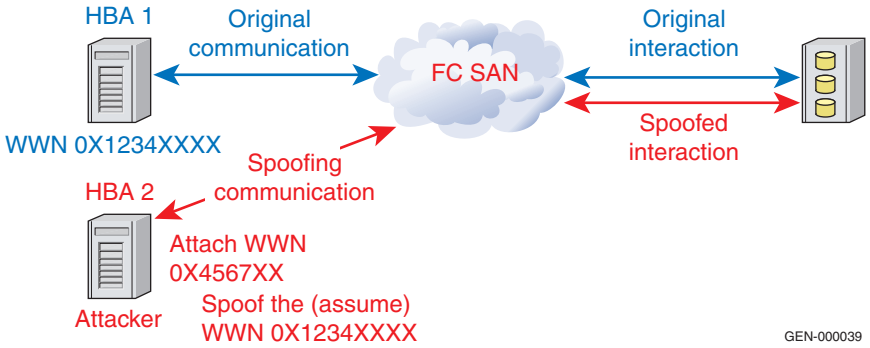


Figure 102 Spoofing in a SAN

Denial-of-service (Disruption)

A denial-of-service (DoS) attack is a deliberate act to prevent an authorized user from accessing data. Limitations of FC and iSCSI SAN protocols can be exploited in order to bring down the network. For instance, the network interface can be flooded with undesired traffic or conflicts can be created that cannot be resolved by the SAN, thereby preventing access to data.

Figure 103 shows an example of DoS in a SAN. In this example an attacker can, after snooping the transaction between the two nodes, issue a port discovery (PDISC) to change the exchange service information to a node. As a result, the service between the two nodes is disrupted.

Other DoS attacks can also precede data theft or destruction. For example, in an iSCSI SAN two iSCSI node names, one authorized and one spoofed, can try to gain access to the same target LUN. Some storage targets will drop the authorized node and grant access to the spoofed node thinking that this is a failover attempt. Other implementations do not allow two node names to connect to the same LUN at the same time, resulting in DoS. Without proper precautions, an unsuspecting administrator might elect to troubleshoot by rebooting the authorized node, thereby giving sole read and write permissions to the spoofed node for the duration of the reinitialization.

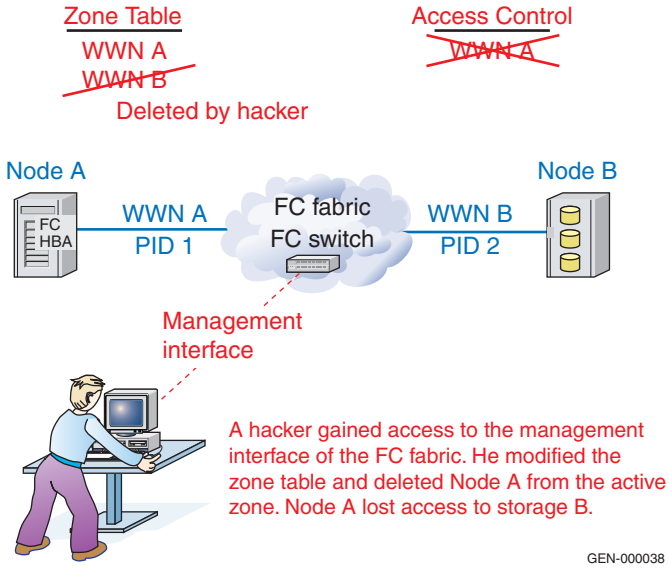


Figure 103 **Denial-of-service attack**

Secure SAN architecture, components, and mechanisms

Securing viable SANs with multiple servers, storage targets, distance extension components, and administrators is complex. Designing security into SANs requires cross-functional investigation, quantification of risks, and breach of security contingency planning before the design can be implemented and tested to satisfy business and regulatory requirements.

There is no single comprehensive security solution. Many argue that this is a result of security features not being accounted for in the early development of industry standards. Initial iSCSI SAN and FC SAN protocols have inherent weaknesses with authentication and authorization. SAN vendors have implemented partial solutions to address security issues; however, many of these solutions are not compatible with emerging standards or interoperable with other existing vendor security offerings.

Standards committees, such as the Fibre Channel Security Protocol (FC-SP) and the IETF IP Security Working Group (IPSEC), are making progress in strengthening the security aspects of these protocols. The FC-SP standard describes the protocols used to implement security in Fibre Channel fabrics. This standard includes the definition of protocols to authenticate Fibre Channel entities, protocols to set up session keys, protocols to negotiate the parameters required to ensure frame-by-frame integrity and confidentiality, and protocols to establish and distribute policies across a Fibre Channel fabric. IPSEC standards are similar, as evidenced in the standardized security elements found in IPv6.

The evolution of SANs includes both the Fibre Channel and IP storage transport protocols. Security vulnerabilities for FC and IP SANs are similar; however, their solution mechanisms and effectiveness may differ. Table 22 on page 505 presents some of the mechanisms available to mitigate security risks.

This section discusses:

- "SAN architectures" on page 503
- "Security components" on page 504
- "Security mechanisms" on page 504

SAN architectures

Secure SAN architectures usually require that multiple security domains, or zones, be implemented and that these security zones be formally documented and controlled to meet regulatory auditing requirements. Security zones can exist between servers and switches, between switches, between SAN management systems and switches, and between administrators and access control management systems.

Figure 104 depicts such security zones. The boxes indicate some of the relevant security components and mechanisms that can be deployed to control security in each of these security zones.

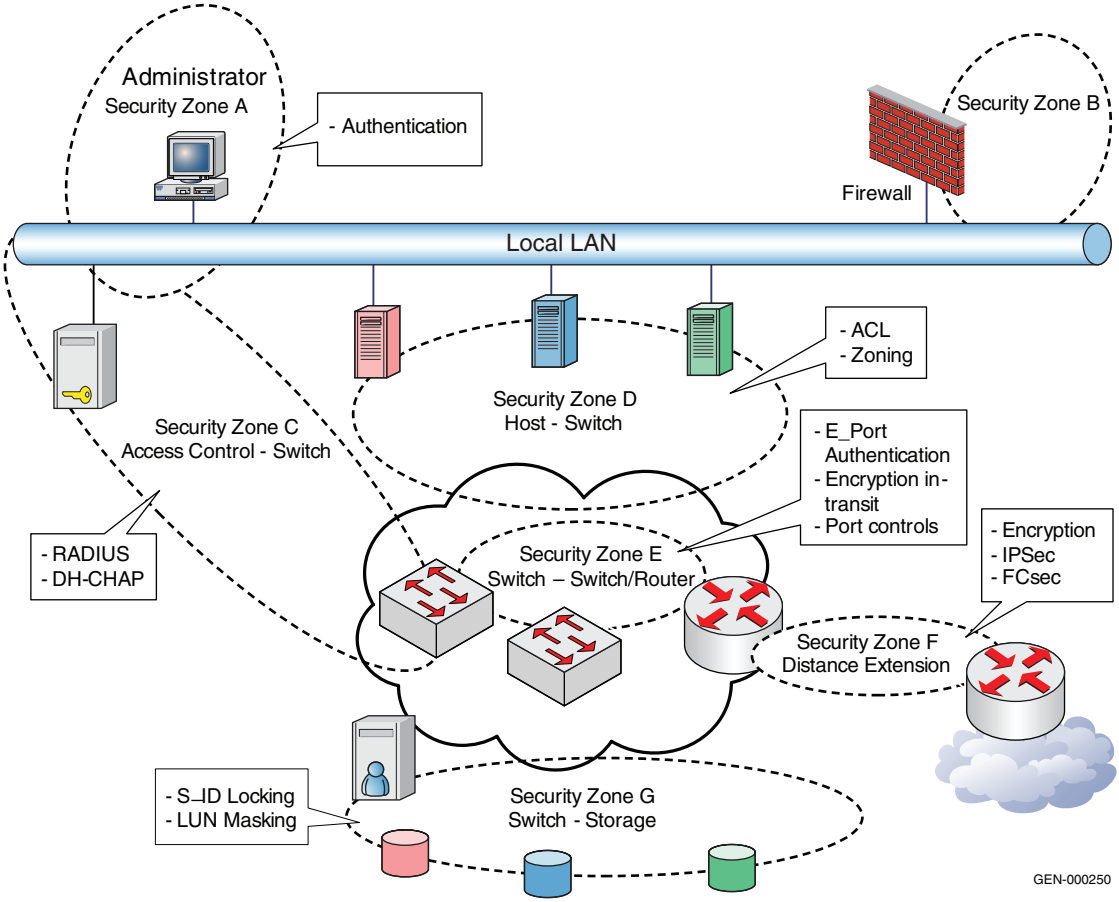


Figure 104 Security zones

Security components

Authentication, authorization, auditing (AAA), data integrity, availability, and encryption are frequently referred to as *components* of a secure implementation. Within SANs, these fundamental components of security are implemented through various *mechanisms* (described next) to provide secure data access, secure management, and data integrity.

Security mechanisms

Available mechanisms that promote a secure SAN include:

- Access control
- Zoning
- LUN masking
- Port binding
- Management keys
- Protocols
- Encryption
- Physical access control mechanisms

These mechanisms can vary by topology, vendor, and business needs. Table 22 lists some of these component mechanisms. Specific design considerations and vendor-specific implementations for some of these mechanisms are presented in "Building secure SANs" on page 514.

Table 22 Secure SANs component mechanisms

Security category	FC SAN mechanisms	IP SAN mechanisms
Authentication	DH-CHAP RADIUS TACACS FCAP FCPAP FC-SP (ESP) FC-SP (AH) FC-GS-4 CT E_Port	CHAP KBR RADIUS TACACS Kerberos SRP
Authorization	Hard port-based zoning Soft port-based zoning LUN Masking RSCN Port controls Port binding	iSCN LUN Masking VLAN Tagging Port controls
Administration / Auditing	ACL SSH SSL	ACL SSH SSL
Encryption	FC-SP (ESP) In-transit Algorithms At-rest Algorithms	IPSec In-transit Algorithms At-rest Algorithms
Integrity	FC-SP (AH) MD5 hash SHA-1 hash	IPSec (AH) MD5 hash SHA-1 hash

The following sections provide further information and suggestions to secure data access, SAN management, and data integrity:

- "Securing data access," next
- "Securing SAN management" on page 508
- "Securing data confidentiality and integrity" on page 510

Securing data access

Securing access to data within a SAN includes authentication and authorization components to control access to data as well as access to SAN management functions. Mechanisms can be deployed in

zones that exist at the ends of the data path or in zones that provide the interconnection in the network (for example, FC fabric switches, TCP/IP network switches, routers). Access control lists, zoning, LUN masking, binding, and port controls represent such mechanisms. Each of these mechanisms are further discussed in this section.

Access control

Implementing access control includes securing access to the management console, maintaining access control lists, and securing access to ports within a SAN. Access control can be complex and may fail to meet auditing requirements in the absence of documented policies. Moreover, without strong authentication and encryption in the data path, access control is not as secure as one would expect.

Secure access to a management console should include restricting the IP address for the management application. Switch management should be done only from secure networks. Remote access should be provided through a secure tunnel network, such as a Virtual Private Network (VPN). Secure access can be improved by implementing two-factor authentication that uses an additional security component, such as a smart card, so that the access is only given when a username, a password, and the key are authenticated. Other user login measures should include policies for password generation/renewal, key pass phrase requirements, and hash time limitations.

Access control lists (ACL) and policies provide authorization for access to SAN resources. Different access levels provide an additional level of accessible features and functions. Moreover, ACLs may be used to separate data access from data management. User login permissions for different access levels, such as equipment manager, security officer, recovery officer, network admin, storage admin, and so on, should be set.

Physical access to network ports should also be secured. Key card access to restricted areas, as well as port controls that are available from Fibre Channel and IP switch vendors, should be utilized.

Zoning

Fibre Channel switch zoning and IP-based switch VLANs are similar. Nodes are segmented into zones and given access to other nodes within the same zone. Zone members have any-to-any connectivity within the zone, whereas non-zone members have no connectivity. Zoning, VSAN, LSAN, and VLAN tagging are SAN technologies that use hardware or software to create segments in a single SAN fabric.

The purpose is to allow the grouping of a set of nodes. As a result, access is limited to a certain set of nodes (either initiators or targets) whereby both intentional and unintentional access to data is restricted. Other advantages of zoning include lowering the effects of State Change Notifications (SCN) and broadcasts, as well as controlling traffic of particular nodes for performance.

Zoning can be *soft* or *hard*. *Soft zoning* is based on the nodes World Wide Name (nWWN) regardless of the physical port on the switch to which it is connected. Soft zoning does not perform any filtering of frames among zones; it merely restricts routing information from being passed to unauthorized zone members. One significant advantage of soft zoning is its flexibility and ease of management. For example, with soft zoning if a node must be physically moved and plugged in to a new port, its zone membership stays intact. A significant security disadvantage is that a malicious node could spoof any given nWWN on the fabric and access data located on a LUN to which it normally would not have access.

Hard zoning, using physical switch ports for zone membership, is a more secure zoning method. Not only is routing information not passed to unauthorized zone members, but frames are filtered to ensure only authorized zone members can communicate. WWN spoofing and route-based attacks would be defended. Hard zoning security benefits come at the cost of SAN management.

LUN masking

Storage device logical unit numbers (LUNs) connected to a SAN are made visible to host devices. LUNs can be presented to one or more hosts. The ability to allow a given host to see a LUN is referred to as *LUN masking*. LUN masking is a common method of controlling host to LUN access. LUN masking can be implemented at the host node, within the switch, or at a storage node.

EMC recommends implementing LUN masking at the storage node in combination with hardware-enforced WWPN zoning. This affords the authorized storage administrator control while protecting the LUNs from spoofing attempts.

Initiator access to storage LUNs can be controlled by such LUN masking tools as Symmetrix Device Masking and CLARiiON Access Logix. Symmetrix systems implement a Volume Configuration Management Database (VCMDB) so that a HBA can only access a particular set of LUNs. EMC CLARiiON implements a storage group for the same purpose. In both implementations, hard zoning should

be used to circumvent malicious or accidental changes to the LUN masking table.

Fabric binding

Fabric binding mechanisms allow only authorized switches to join or disrupt the current fabric. ISLs are only enabled between specified switches in the fabric. Membership data is used to ensure that the list of authorized switches is identical in all switches in the fabric. Fabric binding helps to ensure that unauthorized switches are segmented from the fabric and that only an authorized switch is merging into the expected fabric.

Port binding/S_ID lock down

Port binding is a SAN security mechanism that uses switch software or hardware to associate between a physical port ID and host WWN. This association will mitigate snooping attempts. A malicious node on a bound port attempting to spoof another node's WWN would not be able to connect to another node since the spoofed WWN is not associated with the port ID. When using port binding, all nodes need to be bound to their specific port since nodes that are not bound can still spoof.

A similar security mechanism to mitigate spoofing in shared storage port configurations is to use a Source ID (S_ID) Lock Down feature, like the one available on Symmetrix systems. By mapping the S_ID with the WWN of a HBA into the VCMDB, only a HBA physically connected to a specific switch port is allowed to log in to the storage port. Once a S_ID is locked down, a spoofed HBA WWN cannot log in. Further, if a spoofed WWN is already logged in, that storage user loses all access from that HBA.

Port controls

Port security controls complement fabric binding by helping to prevent unauthorized access to a switch. Port controls, such as port locking and port-type locking, can help to protect the overall SAN infrastructure. Port locking persistently prohibits an unused port from joining a fabric. Port-type locking forces a G_Port to act as only an F_Port, N_Port, or E_Port. Implementing such controls limits the expansion of the fabric and lowers the risk of a physical security attack.

Securing SAN management

SAN management consoles are primary targets for attackers. Risks include usage of clear text management protocols, weak username and passwords, un-segmented communication networks, and shared

accounts. Administrators should deploy strong authentication and authorization mechanisms to secure SAN management. Implementation decisions are necessary to secure SAN management functions while balancing business needs for accessibility and performance. This section discusses switch authentication and management authorization.

Switch authentication

The FC-SP specification provides for host-to-switch and switch-to-switch authentication. The Diffie-Hellman Challenge Handshake Authentication Protocol (DH-CHAP) provides a secure mechanism for authentication between SAN management consoles and other SAN components.

Remote Authentication Dial In User Service (RADIUS) is a security protocol in network environments and is often embedded in switch and router components. User profiles are maintained in a database that remote components can share to authenticate and authorize access to the network. Provisions are typically made for local or centralized authentication through RADIUS or TACACS servers.

The FC-SP specification also defines authentication protocols as being *secret-based*, *certificate-based*, or *password-based*. *Secret-based* component implementations share a common secret or centralize secret administration in a RADIUS server. Components may authenticate by using the DH-CHAP protocol. *Certificate-based* implementations allow components to authenticate with a certificate-based infrastructure whereby the components are certified by a Certificate Authority. The resulting certificates bind each entity to a public-private key pair that can be used to mutually authenticate with other certified entities through the FCAP protocol. *Password-based* authentication protocols establish a security relationship by having knowledge of the password-based credential material of other entities. Entities may use this credential material to mutually authenticate with other entities using the FCPAP protocol. Similarly, iSCSI SANs authentication mechanisms include MD5 CHAP, SRP, and iSCSI Kerberos. In addition to these authentication mechanisms, management communication channels should utilize encrypted protocols such as SSH and SSL.

FC and IP switch vendors provide these various security mechanisms to authenticate switch management and merge activity. Authentication in a secure fabric verifies the identity of a new switch before it is allowed to join the fabric and also ensures the new switch is joining the expected fabric. Properly implemented, these

mechanisms only allow switches that are authenticated and authorized to join a fabric. Since many of these mechanisms have vendor-unique features, secure switch interoperability needs to be considered by the administrator in mixed-vendor environments.

Management authorization

Security management addresses authentication and authorization, as well as the logging of operations for auditing. Authentication and authorization of management access to the network needs to be implemented and enforced with policies and rules. Two-factor authentication for management access is a best practice, as is segmenting the management node from internal communications networks. Separation of management duties should be implemented by establishing ACLs and Role Based Access Control (RBAC), whereby levels of access and functions that a user can perform are controlled and monitored.

Note: Every authenticated user should *not* be granted authorization for all SAN management functions.

Securing data confidentiality and integrity

Data confidentiality includes the encryption of data to prevent reading by others, while data integrity verifies that data sent has not been altered. Securing data confidentiality and integrity can be done with physical security and with encryption, each discussed in more detail in this section.

Physical security

Controlling physical access to SAN equipment is a basic security feature. Mechanisms to implement physical security consist of access cards, locks, gates, in-circuit video surveillance, guards, and armored mobile transport. Policies to physically secure multiple data sites and the network between distributed sites are difficult to implement and costly. The industry trend is leaning towards building security into the SAN to avoid sole reliance on perimeter security deficiencies and its associated costs.

Encryption in the SAN

Zoning, ACLs, and authentication security tools address storage access, but they do not address data payload. *Encryption* and *decryption* provide a level of security beyond the access-control or perimeter-guards. These mechanisms create another security layer to block an attacker from reading the data. Data can be encrypted

"in-transit" or when "at rest." The use of encryption is dependent on how and where it is deployed.

Encryption refers to algorithms that encode plain-text data from the sender into a non-legible form or cipher-text data. The reverse process is called *decryption*. The receiver of the encrypted data uses "key (s)" to decrypt the message, returning it to its original plain-text data. A good, strong encryption will make the data confidential between the sender and the receiver; however, encryption does not guarantee data confidentiality and integrity, especially if the keys are lost or stolen. In order to accomplish secure connections, the recipient of the data must also be authenticated.

Advance encryption algorithms also implement digital signatures or certificates for authentication between the sender and the receiver. Encryption algorithms are categorized as being *symmetric* or *asymmetric*. *Symmetric* key encryption algorithms, such as Blowfish, AES, and DES, work with a single, prearranged key that is shared between sender and receiver. This key both encrypts and decrypts text. In *asymmetric* encryption schemes, such as RSA and Diffie-Hellman, the scheme creates a "key pair" for the user: a public key and a private key. The public key can be published online for senders to use to encrypt text that will be sent to the owner of the public key. Once encrypted, the cipher-text cannot be decrypted except by the one who holds the private key of that key pair. This algorithm is based on the two keys working in conjunction with each other. Asymmetric encryption is considered one step more secure than symmetric encryption because the decryption key can be kept private.

Note: Encrypted data does not protect against a malicious user's ability to destroy encrypted data or against denial of service attacks. Backup policies and additional security mechanisms need to be implemented even when data is encrypted.

Encryption of Data-in-Transit incorporates components that provide a secure communication tunnel in the middle of the session between the initiator-target or in the ISL between two switches. The iSCSI implementation in Microsoft Windows allows the host to embed IPsec encryption with the iSCSI initiator software. In FC SANs, FCsec can provide for encryption of Data-in-Transit.

A major consideration when implementing encryption is the management of the keys used for encryption and decryption. Lost keys, restores, performance, and encryption component failovers

cannot be ignored to meet availability needs for encrypted data. Key management systems need to be established to ensure that data can be recovered when encryption is used improperly and to mitigate the effect of lost or stolen keys.

Encryption of data-at-rest

If data security is a concern behind the "perimeter," then one solution is to encrypt the data on the storage media. This is known as encryption of data-at-rest. Encryption of data-at-rest can mitigate internal attacks against LUN masking, WWN spoofing, and E_Port replication, as well as protect data that needs to be transported or stored externally. As with data-in-transit encryption, the management of the data and keys can become complex and can increase the cost and performance of IT operations.

Some examples of encryption of data-at-rest include:

- File-based encryption on host operating systems, as with Microsoft EFS (Encrypted File System) in a Microsoft Windows environment.
- HBA-based encryption that provides a security gateway in the path of the host. The encryption is usually done on the block level based on the SCSI ID and LUN numbering.
- Appliance-based encryption, such as with Neoscale and Decru, where data is encrypted in the SAN prior to being written to disk or tape. These encryption appliances provide a security gateway in the path between the host and storage.

Disk encryption mechanism

The Advanced Encryption Standard (AES) or Data Encryption Standard (DES) encryption algorithm of the data-at-rest in the disk is restricted on the Open System Logical Block Address (LBA) boundary, which is 512 bytes block. Both Decru and Neoscale do not add encryption rule (metadata) into the encrypted block, hence there is always a one-to-one correspondence from the unencrypted (clear text) block of data to the encrypted (crypto text) block of data.

Each vendor implements different algorithm to ensure that each of the 512 bytes block is encrypted using different keys in order to ensure the security of the data. The algorithm involves the position of the data block in the disk that is based on the LBA. As the result, the host can randomly access any LBA location through the disk encryption appliance to randomly retrieve the unencrypted data from the encrypted disk.

Impact of the storage application to the encrypted data

The disk encryption appliance provides the rule of how a particular key applies to a set of the data. This rule affects the SCSI addressing level, such as the I-T-L (Initiator-Target-LUN Nexus). The encrypted dataset does not carry any information about the rule (metadata), therefore the key must be appropriately associated with the same dataset regardless of the I-T-L mapping.

The storage application data (such as Snapview, Mirrorview, Clone, and SRDF) provides a point-in-time duplicate of data, or a synchronous duplication of data. These applications guarantee that the data is duplicated or transferred as the whole block of the Logical Block Address. The encrypted data can be transferred into another volume with different association of the I-T-L. The user must be actively involved to ensure the correct association of the key with the dataset because of the loose correlation of the keys to the dataset.

Follow these rules for encrypted data mobility:

- The encrypted data must be transferred in the whole entirety of the dataset.
- The encrypted data must be transferred to the same location of the LBA vicinity from the original source

The impact of online encryption data preparation and online encryption data re-keying with the storage applications

Consider the following scenarios between disk encryption and storage application:

- The data is currently converted from clear text to the crypto text.
- The encrypted data is replaced with other encrypted data with different keys.

In the above scenarios, part of the dataset is in either clear text (unencrypted) or crypto text (encrypted) with different key. As a result, the integrity of the point-in-time snapshot or synchronous replication of this dataset *cannot* be guaranteed.

Additional information on storage security can be found on EMC's Powerlink. Select best practices for implementing various secure SAN mechanisms are contained throughout this chapter. Some vendor-specific best practices are included in the "Security implementation examples" on page 516 section.

Building secure SANs

Many factors need to be considered when building secure SANs. Network and security requirements are often unique to each business environment. The advice of professional security consultants is often sought to balance business security, information accessibility, and performance needs.

Design considerations

Before designing a secure fabric, you need to consider current and pending regulations, management costs, data availability, and network maintenance.

The availability of data is arguably the most important aspect of good SAN design. Prior to the application of security mechanisms, whether they are initiated from customer security policy or IT initiatives, the existing SAN needs to be in a state of stability. The complexity of adding layers of security policy in different zones, from the core to the perimeter of the data center, requires a stable SAN in order to troubleshoot any issues when the security variables are introduced.

Assuring availability can come in many forms. For example, knowing who has physical and administrative access to all components of a SAN, in a well-documented format, is a simple best practice. One main reason for costly downtime is due to physical, unintentional breaches in the environment and not knowing who owns administrative access to the affected components.

When designing the security aspects of the SAN, administrators need to be aware of business availability requirements to avoid "over-securing," which can be costly. SAN scaling and general maintenance impact on security features need to be considered to prevent loss of availability or security features. Through proper redundancy and failover practices, secure availability can almost be guaranteed.

Change control and maintenance of records and procedures have become popular security practices within IT organizations. Change control practices must assure that proposed changes to an environment are documented from start to finish with appropriate approval processes in place throughout the IT organization.

Contingency plans need to be documented as well. All personnel who have access to the environment must be made aware of changes.

Natural disasters are also to be considered when planning secure storage environments. Typically, offsite redundancy and replication is preferred to localization of resources for disaster recovery. Suggestions to maintain secure availability in disaster scenarios range from locating resources to create a sense of obscurity to frequently fire-drilling such disasters that include the loss of strategic key management resources.

Security implementation examples

The following security implementation examples are discussed in this section, with emphasis on supported features and best practices.

- "EMC Celerra iSCSI data security," next
- "Brocade" on page 518
- "Brocade M-Series" on page 524
- "Cisco" on page 529
- "EMC Select security appliances" on page 535

EMC Celerra iSCSI data security

Celerra iSCSI features can be leveraged to establish a secure environment for iSCSI sessions. Each of these measures provides a level of security when implemented separately. When combined, security is increased immensely. For instance, just implementing LUN masking provides a level of security, but an unscrupulous person could spoof an Initiator Qualified Name (IQN) in order to gain access to a LUN. Implementing CHAP authentication and firewall rules, in addition to LUN masking, severely restricts such attempts.

Supported features

The following are EMC-supported features:

- iSCSI LUN Masking
- CHAP Authentication
- iSCSI Port Protection
- VLAN Tagging
- Filtering by IP address and port

Best practices

The following is a list of best practices to consider:

- Celerra iSCSI security features can be implemented through wizards. LUNs can be masked while allowing for multiple access options for cluster node joint access through an easy-to-use DataMover graphical interface (Figure 105 on page 517).

A LUN mask creates an association between the iSCSI target (LUN) and the IQN. The Celerra system will deny access to a LUN unless a mask has been configured to associate the LUN with the IQN.

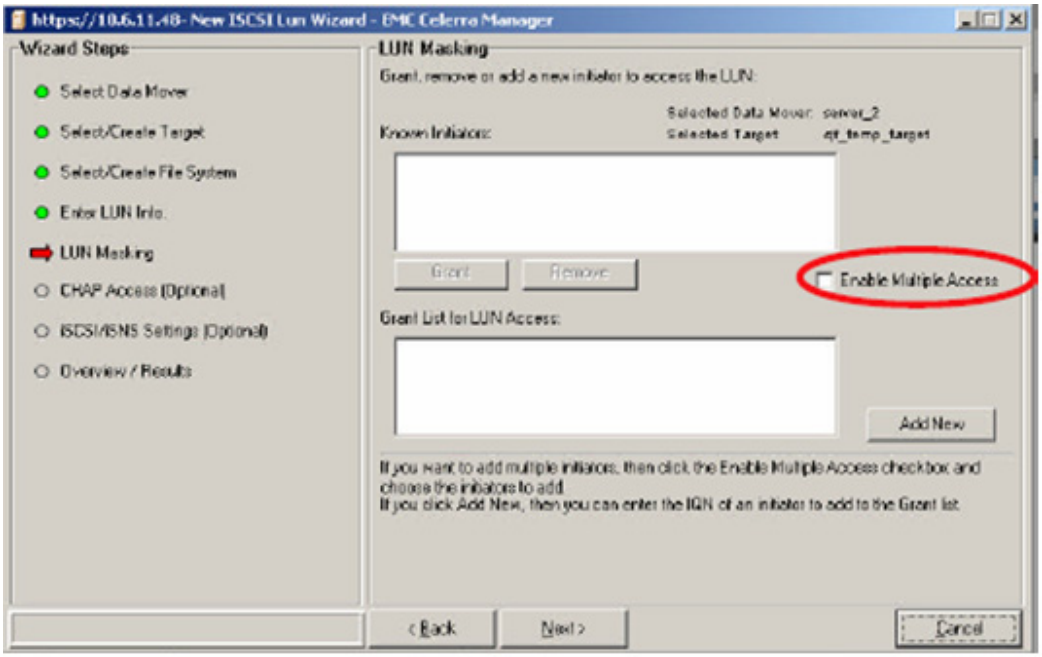


Figure 105 **LUN masking**

- VLAN technology allows for the creation of smaller network domains of trusted systems. Since it may not be desirable to allow all of the systems in the environment to have iSCSI access to the Celerra system, enabling VLAN tagging on the iSCSI interface for Celerra provides access control at the switch to police which systems are able to mount iSCSI devices.
- Implement CHAP authentication on initiators and targets, based on a shared secret, random challenge, and secure one-way hash. Celerra supports CHAP authentication (see Figure 106 on page 518). The default settings do not require CHAP be enabled; however, data access in real time (DART) parameters can be configured to force all initiators to use CHAP to authenticate to the Celerra target. The Celerra system offers several options to help manage secure environments.

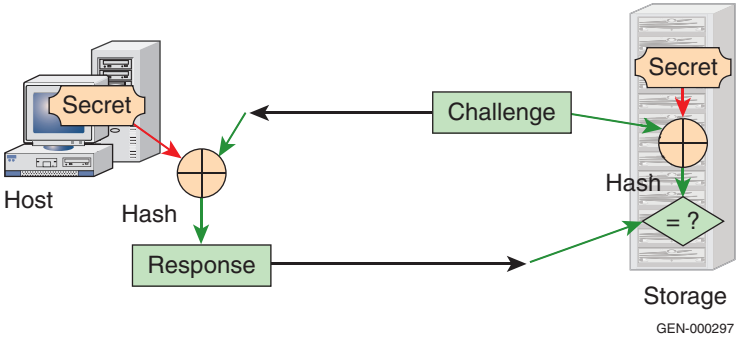


Figure 106 **CHAP authentication**

References

For information on installing iSCSI host components, refer to:

```
http://powerlink.emc.com/km/live1/en_US/Offering_Technic
   al/Technical_Documentation/300-002-695_a01_elccnt_0.pd
   f?mtcs=ZXZlbnRUeXBlPUttQ2xpY2tTZWFyY2hSZXN1bHRzRXZlbnQ
   sZG9jdW1lbnRJZD0wOTAxNDA2NjgwMTQyZmEyLGRhdGFTb3VyY2U9R
   ENUTV9lbl9VU18w
```

For information on best practices for Celerra iSCSI, refer to:

```
http://powerlink.emc.com/km/live1/en_US/Offering_Technic
   al/White_Paper/C1311_Celerra_iSCSI_ldv.pdf?mtcs=ZXZlbn
   RUeXBlPUttQ2xpY2tTZWFyY2hSZXN1bHRzRXZlbnQsZG9jdW1lbnRJ
   ZD0wOTAxNDA2NjgwMTc5MGMwLGRhdGFTb3VyY2U9RENUTV9lbl9VU1
   8w
```

Brocade

Brocade provides flexible features to assist the IT professional in safeguarding the SAN. In addition to traditional hardware security features, since Brocades Fabric OS version 5.2, more Secure Fabric OS™ security features are part of Brocades standard license. In this section, some of these security features are noted with implementation recommendations. Implementation details are provided in Chapter 1, "Fibre Channel SAN Topologies," or in the referenced Brocade documentation.

Security features

Brocade provides policy-based security protection for more predictable change management, assured configuration integrity, and reduced risk of downtime. Some of these Brocade security features are listed in Table 23 under the categories of Authentication,

Authorization, Accountability, Encryption, and Integrity. Arguably, a feature may be listed under one or more of these categories. Product-specific, conditionally and unsupported features are specifically listed in "EMC conditionally or unsupported features" on page 520.

Table 23 Brocade security features

Category	Feature	Description
Authentication	Switch-to-switch	Switch-to-switch (E_Port) authentication using Fibre Channel Certificate Authentication Protocol (FCAP). Brocade provides commands to disable switch-to-switch FCAP authentication and to select an alternate authentication protocol such as DH-CHAP. DH-CHAP is a secrete-based authentication and key management protocol that supports both switch-to-switch and host-to-switch authentication.
	Password administration and account lockout	Features and policies are included in Admin Domain.
	Switch-Host	RADIUS or TACACS+ is supported to ensure that only authorized devices access protected networks.
Authorization	Zoning	Hardware enforced WWN zoning.
	Port controls	Supports E_Port, L_Port, and G_Port lock down, port binding, and persistent port disable controls.
	Insistent Domain ID [FICON]	Allows a switch to insist on a specific Domain ID prior to joining a fabric.
	LSAN technology	Proprietary to Brocade and similar to Cisco VSAN ideology. Allows storage administrators the flexibility to isolate environments logically.
	Role-Based Access Control (RBAC)	ACL support (Switch and Port Binding) policies are identified by the following specific names and cannot co-exist. The policies are case sensitive. The ACL policies can be distributed to databases per-switch or fabric-wide. The two policy types are: • Device Connection Control (DCC) policies – used to restrict which device ports can connect to other switch ports • Switch Connection Control (SCC) policies – used to restrict switches from joining the fabric or joining another switch.

Table 23 Brocade security features (continued)

Category	Feature	Description
Accountability	SNMP v3 standard for monitoring and managing	SNMP access-control lists prevent/get/set operations to individual host or IP addresses. SNMP agent and Management Information Base (MIB) are located on each switch.
	HTTPS	Can be used with Web Tools for managing the switch.
	SNMP trap configurations	ISL and security thresholds can be set.
	Virtual Fabric with Administrative Domains	Provides for auditing of user and security related events, such as configuration changes, security, zoning events, and firmware download.
Encryption	SSH (Secure shell access) for ensuring network security encrypted sessions	Uses the SSH daemon, server-side initiated certificate. Nothing is needed on the switch.
	SSL (Secure Socket Layer) supporting SSLv3 (128-bit encryption) for support of HTTPS.	Certificates must be generated and installed on each switch to enable SSL.
	IPSec Encryption	Services for the PB-48000B-18i line card (FR4-18i) and 7500.
Integrity	MD5 & SHA-1 hashes	DH-CHAP supports MD-5 and SHA-1 algorithm-based authentication.
	SFC (Secure File Copy)	SFC of configuration uploads and downloads.

EMC conditionally or unsupported features

For the most up-to-date listing of EMC conditionally or unsupported product features, please refer to the current product release notes available on Powerlink.

- PKI (Public Key Infrastructure) — The Fabric OS 5.2.0 SFOS migration strategy of removing dependencies on a Brocade PKI is based on the assumption that there is limited demand for PKI-based solutions and that administrators will use FC-SP standards for switch to switch authentication. Brocade no longer includes a PKI Certificate as part of the installed Secure Fabric OS. If you wish to activate Secure Fabric OS on a supported director or switch, you must contact Brocade to obtain a PKI certificate.
- Secure FOS is not supported by EMC. FR4-18i (ED-48000B) blades can be used in a switch running Secure FOS. However, adding a reference to one of the ports added by PB-48K-48 (known as ports 256-383) in a DCC policy can only be performed on a FOS 5.2 or

newer switch since older versions of FOS limit the port numbers to a maximum of 255. Because of this limitation, the FCS switches must be running FOS 5.2 or later firmware. The FCS switch is the only switch that can be used to modify security policies.

- Fast Write and Tape Pipelining are not supported in conjunction with secure tunnels.
- Jumbo frames are not supported on secure tunnels.
- ICMP redirect is not supported for IPSec-enabled tunnels.
- Only a single secure tunnel is allowed on a port. Non-secure tunnels are not allowed on the same port as secure tunnels. Modifying operations are not allowed on secure tunnels. To change secure tunnel configurations, you must first delete the tunnel and then create a new one with the desires options. Only a single route is supported on an interface with a secure tunnel.
- An IPSec tunnel cannot be created using the same local IP address if ipperf is active and using the same local IP address (source IP address). Unidirectional supported throughput is ~104Mbytes / sec and bidirectional supported throughput is ~90Mbytes / sec. An IPSec tunnel takes longer to come online than a non-IPSec tunnel. The user is not informed with the IPSec mismatch RAS event when configuring a tunnel with IPSec mismatch on either ends.

Best practices

The best practice recommendations described in this section were developed based on EMC's current understanding of the general security environment and the capabilities of EMC's product(s). Security is also impacted by your requirements and your IT environment along with newly discovered vulnerabilities, security threats, and technologies identified on an almost daily basis. Changes to any of these factors may impact the applicability and effectiveness of these best practice recommendations.

Implement Role Based Access Control (RBAC)

Fabric OS 5.0.1 has three different roles: User, Admin, and Switch Admin. Another four roles (Operator, Fabric Admin, Zone Manager, and Basic Admin) have been added as of Fabric OS 5.2. Table 24 on page 522 shows the functional area and capabilities for each of these roles. Refer to the *Brocade Administrator's Guide* for implementation details.

Table 24 Implement Role Based Access Control (RBAC

Functional area	User	Switch Admin	Admin	Operator	Zone Manager	Fabric Admin	Basic Admin
Zone Configuration	V	V	M	V	M	M	V
Environmentals	V	M	M	M	V	M	P
Enable/Disable Ports	V	M	M	M	V	M	P
Logs (RAS)	V	M	M	M	V	M	P
Security Settings	V	V	M	V	V	M	V
Switch Configuration	V	M	M	V	V	M	P
Switch Management	V	M	M	M	V	M	P
Port Configuration	V	M	M	V	V	M	P
SNMP	V	M	M	V	V	M	P
Diagnostics	V	M	M	M	V	M	P
Devices	V	M	M	V	V	M	P
User Management	X	X	M	X	X	X	X
Fabric Watch	V	M	M	M	V	M	P
APM	V	M	M	V	V	M	V
Admin Domain Mgmt.	X	X	M	X	X	X	X

Legend:

V = View Only

M = View and Modify

P = Partial

X = N/A

Implement auditing

Use capabilities and procedures to capture, record, and track significant events. Captured event information should account for:

- Which users are making changes from an external source (username).
- Where they are logged in from (IP address).
- Type of management interface (Telnet, Web Tools, etc.).

Since the types of events being audited may be somewhat frequent, the handling of these events can be streamed off the switch to a remote host running a SYSLOG server. Scheduled at a logical and regular interval (such as every day at midnight) an auditor could review events to look for activity such as configuration changes, firmware downloads, etc., that that may be useful for network troubleshooting or security breaches. Refer to Brocade documentation for implementation details.

Implement hardware-enforced WWPN zoning

Securing ports can be a hindrance to flexibility of managing switches but is typical of securing an environment. Trade-offs should be expected. Recommendation is to implement hardware-enforced WWPN zoning. No license is required and this feature is enabled by default. Refer to Brocade documentation for implementation details.

Implement port controls

Use port controls such as E_Port lockout, L_Port lockdown, G_Port lockdown, and persistent port disable. These features help to prevent physically connected nodes from logging into an unintended or restricted environment. By manually setting the port configuration, you bypass the ports default auto-configuration. Refer to the *Brocade Fabric Manager's Administrator's Guide* for switch dependent implementation details.

Secure CLI sessions

Minimally, password protect service port and IP connections. Additional uses of switch-to-switch and switch-to-host authentication mechanisms are suggested. Refer to Brocade documentation for implementation details.

Related Brocade documentation

The following documentation is available on Brocade's website at http://www.Brocade.com:

- *Brocade Fabric OS Administrator's Guide*; Publication Number:53-1000043-02
- *Brocade Fabric Manager's Administrator's Guide*; Publication Number 53-1000042-01
- *Brocade Fabric Manager's Administrator's Guide*; Publication Number 53-1000043-02
- *Secure Fabric OS Administrator's Guide*, Chapter 2, "Adding Secure Fabric OS to the Fabric," for a description of how to obtain certificates from the Brocade Certificate Authority.

Brocade M-Series

Brocade-McDATA (hereafter referred to as Brocade-M) switches, directors, and software management applications offer components of Brocade's SANtegrity Security Suite.

SANtegrity ensures that Fibre Channel traffic is not redirected by unauthorized access. SANtegrity supports hardware-enforced zone parameters to protect from DoS attacks. Both zone member specification by port or WWN and software-enforced Fabric Name Server zoning is supported. In mainframe FICON environments, FICON cascading is enabled when SANtegrity is installed.

The SANtegrity Security Suite includes both standard and enhanced licensed features.

Standard features include:

- Advanced Zoning

  Advanced zoning enforces zone parameters at the ingress port to protect from DoS attacks on the Fabric.

- Role Based Accounting Control (RBAC)

  RBAC provides management zoning.

- Secure Access

  Provides locked down interface management and improved performance.

- Secure Platform

  Provices secure techniques for communication with storage network devices and for secure client server communication.

Enhanced licensed features include:

- Binding

  Binding is a set of features that locks a fabric into an intended configuration.

- Authentication for both FC and IP block-based protocols

  SANtegrity Authentication is based upon the interoperable authentication protocol, DH Challenge Handshake Authentication Protocol (CHAP), recommended as the Fibre Channel standard for security (FC-SP) in storage networks. Each

device participating in a storage network proves its identity through the supported FC-SP, iSCSI, FC-GS, FC-SB, and iFCP protocols.

- Security Center management for Brocade's EFCM and SANavigator products.

  SANtegrity Security Center manages the administration of device settings, generates logs and reports based on security policies, and promotes both monitoring and planning.

A license key specifies the maximum number of switch ports, the number of clients you can run, the expiration date of a temporary license, and the licensed optional features.

Security features

Brocade-M features are noted in Table 25 under the categories of Authentication, Authorization, Accountability, Encryption and Integrity. Arguably, a feature may be listed under one or more of these categorie

Table 25 Brocade M security features

Category	Features	Description
Authentication	Users	Adding, deleting, editing, defining roles, and password maintenance for authorized users.
	Software	Configure authentication settings for management access of both in-band and out-of-band software access.
	Devices	Define CHAP Secrets, Port Authentication Sequences (i.e., RADIUS, then Local: Radius Only; Local Only), and adding/configuring/removing authenticated devices.
	Ports	Port Authentication allows user to override authentication settings of specific ports.

Table 25 Brocade M security features (continued)

Category	Features	Description
Authorization	Zoning	Hardware-enforced WWN zoning is enabled by default and does not require configuration. Zoning is enforced in the ASIC route tables at the ingress port. When Class 3 frames are sent to none zone members, they are dropped.
	Port Binding	Port Binding has access control policy by port. You must configure F_Ports or E_Ports WWNs on an individual port basis.
	Switch Binding	Permits specific F_Ports or E_Ports to connect to a particular switch. The feature may be enabled by activating the “Enterprise Fabric” mode or by activating it directly.
	Fabric Binding	SANtegrity licensed option that permits specific switches to join a fabric. Binding has an access control policy by switch. The Insistent Domain ID feature allows the switch Domain ID to be statically set. **Note:** This feature MUST be enabled with Fabric Binding.
	Port Controls	Standard port type controls as well as persistent port blocking..
Accountability	Accounting Services	Log information for each management session in a switch. Can be implemented locally or remotely (RADIUS). Syslog available for centralized repository of audit logging; different logs can be configured through E/OS.
	Logs	Email notifications are sent and switch logs are updated when devices attempt to log in to a port with Port Binding.
	Role-Based Access Control	Access via CLI or SNMPv3. Secure management zone and reporting.
	Password Management	Password expiration through CLI and Connectrix Manager. To eliminate the possibility of password “guessing”, as of E/OS 9.1 switches have the feature set of configuring and retaining a history of up to three passwords. Administrators who require users to change passwords periodically can do so with this feature.
Encryption	SSH/SSL	SSH and SSL for ensuring network secure encrypted sessions. SSH through CLI since E/OS 7.1. SSH for Connectrix Manager (EFCM) between GUI and switches since E/OS 8.1. SNMPv3 for MIB management.
Integrity		Supports SSH v2, SNMPv3, SFTP, AES, MD5 and SHA 1

Best practices

The best practice recommendations described in this section were developed based on EMC's current understanding of the general security environment and the capabilities of EMC's product(s). Security is also impacted by your requirements, your IT environment, and by newly discovered vulnerabilities and technologies. Changes to any of these factors may impact the applicability and effectiveness of these best practice recommendations.

Common recommendations include changing default passwords, using port controls to disable unused ports or management interfaces, using SSL or SSH, and limiting physical access to FC switches. The following are additional best practices when configuring Brocade-M secure fabrics.

- "Hardware enforced WWPN zoning," next
- "Fabric binding" on page 528
- "Switch binding" on page 528
- "Port binding" on page 528
- "EFCM reporting and auditing" on page 528

Hardware enforced WWPN zoning

Zoning in Brocade-M devices is similar to Brocade and Cisco devices in that zoning minimizes fabric disruptions. The security administrator should take advantage of zoning's ability to increase security measures by preventing data loss, corruption, or attacks by specifying which devices can access one another. Brocade-M switches allow for this by implementing hardware-enforced WWPN zoning. No license is required for this feature and it is enabled by default.

Brocade-M recommends creating logical subsets of zones / user groups. Authorize access rights to those specific zones for specific user groups so that unintentional usage will not occur and the protection of confidential data from unauthorized access is consistent. Creating groups of devices that are separate from devices in the rest of the fabric, and zoning them, will allow processes to be performed on devices in one group without interrupting the other. For example, use single initiator zoning and keep your backup traffic separate from primary traffic.

Zoning practices and the Safe zoning mode (introduced in EOS 9.01.00) prevents switch merges from inadvertently creating unintended zone sets and prevents default zones from being enabled. It also prevents merges of fabrics with zone sets that do not match.

After the environment has been configured and all hosts see the appropriate volumes, configure the following Enterprise Fabric mode security features (each discussed in more detail in this section):

- Fabric binding
- Switch binding
- Port binding

Fabric binding

The Brocade-M fabric binding feature provides security from intentional and accidental fabric merges. Fabric binding permits only specified switches to join a fabric. It is set on a fabric-wide basis, not on an individual switch basis. World Wide Names (WWNs) of the switches are used to determine which switches may participate in a fabric. In McDATA Enterprise Fabric Mode, fabric binding cannot be disabled. Configuration details and information on how to modify the Fabric Binding Membership List (FBML) can be found in the "Configuring Security" chapter in the *EFCM Basic User Manual*, located at http://www.Brocade.com.

Switch binding

Switch binding permits specific F_Port or E_Port connections to a particular *switch*. The feature may be enabled by activating the Enterprise Fabric mode or by activating it directly. WWNs of a node or switch are used to determine which nodes or switch may connect to a switch. Switch binding is configured on an individual switch basis. Nodes already participating in the fabric when switch binding is activated will automatically be included in the Switch Binding Membership List (SBML). In switch binding, a node in the SBML may connect to any port on the switch.

Port binding

Port binding permits specific F_Ports or E_Ports to connect only to a particular *port* of a switch.

Port binding is more granular then switch binding. WWNs of F_Ports or E_Ports must be configured on an individual port basis. In port binding, a WWN or an Alias of a node is assigned to a port and it may only connect to that port. Configuration details can be found in the "Configuring Security" chapter in the *EFCM Basic User Manual*, located at http://www.Brocade.com.

EFCM reporting and auditing

Use the reporting and auditing tools in Connectrix Manager to track logins, operating parameters, and zoning changes. The following events are logged and may be used to satisfy auditing requirements.

Implementation details can be found in the referenced *EMC Connectrix Manager User Guide* located on Powerlink.

- Defining a new product
- Modifying product attributes
- Deleting product definitions
- Defining a new user
- Modifying user administration
- Deleting user administration
- Creating a zone set
- Modifying SNMP options

Related Brocade-M documentation

The following documentation is available for reference:

- "Configuring Security" chapter in the *EFCM Basic User Manual*, located at http://www.Brocade.com.
- *Brocade M-EOS Command Line Interface User Manual*, located at http://www.Brocade.com.
- *EMC Connectrix Manager User Guide*, located on Powerlink.

Cisco

Cisco provides a comprehensive security framework within SAN-OS. Licensing is required for some enhanced security features including FC-SP authentication, port security, LUN zoning, IPSec, and VSAN-based access control. Licensing and implementation details can be found in the referenced *Cisco MDS 9000 Family Configuration Guide* and *Cisco MDS 9000 Family Fabric Manager User Guide*, both available at http://www.cisco.com.

Security features

Cisco features are listed in Table 26 on page 530 under the categories of Authentication, Authorization, Accountability, Encryption, and Integrity. Arguably, a feature may be listed under one or more of these categories. Product-specific conditionally and unsupported features are listed following Table 26.

Table 26 Cisco security features

Category	Feature	Description
Authentication	Switch-to-Switch	FC-SP for Switch-to-Switch and Host-to-Switch Authentication using DH-CHAP. Supports MD5 and SHA-1 algorithm based authentication. Configuring the DH-CHAP feature requires the ENTERPRISE_PKG license.
	Host-to-Switch	RADIUS or TACACS+ is supported to ensure that only authorized devices access protected networks.
Authorization	Zoning	Supports N_Port, Fx_Port, iSCSI, LUN, Read-Only, and Broadcast Zones, as well as, hardware enforced WWN zoning.
	VSAN technology	Proprietary to Cisco and similar to Brocade LSAN ideology that allows storage administrators the flexibility to isolate environments logically. Port Security is activated on a per VSAN basis.
	Switch Access Control	Supports SSH v2, SNMP v3, and Role Based ACLs.
	Persistent Port Disable	A disabled port remains disabled through offline/online changes and reboots.
	Fabric Binding	FICON Pre-3.x; VSAN Post 3.x extends port security to enable ISLs only between specified switches.
Accountability	Accounting Services	Log information for each management session in a switch can be implemented locally or remotely (RADIUS).
	Port Tracking	Monitors and detects failures that cause topology changes.
	Role-Based Access Control	RBAC via RADIUS/TACACS+ features pre-defined roles and customization ability to limit administrators and users on a VSAN basis. For management access via CLI or SNMPv3, the same usernames and passwords are used to gain access to the switch.

Table 26 Cisco security features (continued)

Category	Feature	Description
Encryption	Digital Certificates	Management tools utilize SNMPv3 to communicate between switch and GUI.
	SSH	SSHv2 encrypts and authenticates traffic between switches and management stations instead of Telnet. Host keys RSA, RSA1, DSA, and AES are available.
	IPSec	Encryption services for the MPS 14/2 or MDS 9216i.
Integrity	Protocol Methods	Supports SSH v2, SNMPv3, SFTP, AES, MD5, and SHA 1
	IPSec	FCIP and iSCSI connections use IKEv1 and IKEv2 protocols. Protects and authenticates IP packets between participating devices. Enterprise Package required.

Conditionally or unsupported features

For a listing of EMC conditionally or unsupported product features, please refer to the most current version of the *Connectrix MDS 9000 SAN-OS Release Notes* residing on Powerlink.

- Cisco implementations of LUN zoning are unsupported
- Read-only zones are unsupported
- Cisco implementations of IPSec and AES Encryption for iSCSI are unsupported

Best practices

The following best practices are recommended.

Implement Zoning and VSANs

Zoning and VSANs are Cisco's natural security features that can be applied to isolate traffic within the fabric. By segregating groups of hosts and storage you are not protecting the environment via authentication or data encryption at this layer, but you are providing availability and fault tolerance. See Cisco documentation for implementation details.

VSANs should be created where applicable and where management of the VSANs will not over-complicate the environment. Ports available for VSANs are 1-4094. By default, all ports fall under VSAN port 1. To avoid possible WWN spoofing, place all unused ports in backup VSAN 4094 and then build separate VSANs in 2-4093.

It is recommended to use static domains to avoid any possible merge conflicts and manageability conflicts. Inter-VSAN routing should be used sparingly for the resources that need to be shared.

Implement hardware-enforced WWPN zoning

Securing ports can hinder the flexibility of managing switches, but are typical of securing an environment. Tradeoffs should be expected. Recommendation is to implement hardware-enforced WWPN zoning. Hardware-enforced zoning has a level of fabric enforcement not provided by software-enforced zoning. Hardware-enforced zoning is applied to every FC frame at the switch port.

Implement port controls and security

Use port controls to eliminate the dangers of having users, intentionally or not, misuse a port that has the default 'auto' mode port settings. To avoid this danger, configure 'port' mode on all switch ports, shut down all used ports, and only allow connections from expected device types by specifying N_Port, E_Port, F_Port, and FL_Port settings.

Port security prevents unauthorized access to a switch port by binding specific WWN access to one ore more given switch ports. Complimentary to port security is WWN-based zoning which zones the switching logic to frames based on the WWN, and not the physical port, on a device. With this logical security, spoofing can be a problem.

Use port security features to:

- Bind devices to switch
- Bind devices to a port
- Bind to a group of ports in case of individual port failure
- Bind switches at ISL ports; not just individual switch

When enabling port binding, consider the impact of choosing whether or not device-to-switch or switch-to-switch port security is enabled and assure that it does not impact something that should be accessing a specific port. When these features are enabled, it will reject login requests from unauthorized FC devices as well as report attempts to the SAN administrator. Refer to Cisco documentation for implementation details.

Implement Fabric Binding

Fabric Binding allows access to fabric devices based on identity attributes. All Cisco MDS 9000 switches enable fabric-wide authentication from one switch to another or from a switch to a device.

You can perform switch and host authentication locally or remotely through the DH-CHAP protocol. To configure the Enterprise Package licensed DH-CHAP authentication feature using the local password database, please refer to the *Cisco MDS 9000 Family Configuration Guide*. DH-CHAP based authentication should be employed to prevent spoofing or hijacked WWNs where port security is vulnerable.

RADIUS / TACAC+ provides for centralized management of access control when multiple switches are used. Please note that for switch-to-switch authentication, relying only on remote authentication creates a single point of failure.

Enable port tracking

The port tracking feature monitors and detects failures that cause topology changes or bring down links connecting devices. When enabled and explicitly configured, the Cisco SAN-OS software monitors the tracked ports and alters the operational state of the linked ports upon detecting a link state change.

Secure Management Access

Securing Management Access can be done at all points of the data path. Cisco supports SNMPv3, SSH, and SSL. It is recommended that Telnet, SNMPv2, and HTTP be disabled. Use of VPNs (Virtual Private Networks) are recommended for remote management of an environment.

Another Cisco features known as a VLAN (IP Networks) is recommended to isolate management traffic that commonly connects to the SAN management stations, like Cisco Fabric Manager.

Implement Role Based Access Control accounting services

SAN-OS provides Role Based Access Control (RBAC) for management access of the Command Line Interface (CLI) and Simple Network Management Protocol (SNMP). In addition to predefined roles, up to sixty four (64) roles can be defined. The roles describe the access control policies for various feature-specific commands on one or more VSANs. CLI and SNMP share the same usernames and passwords (i.e., a single administrative account for each user) to gain access to the switch. Predefined roles include:

- Network Operator (read only access)
  - Show Commands
  - Exec mode file system commands
  - Basic network diagnostics

- Network-admin (read / write)
  - Access to all CLI commands

Best practices are to use roles to give access to groups of commands and to deploy them on a per-VSAN basis and by administrative function to avoid over-managing the environment. Roles can have access permissions assigned to them either as CLI or SNMP users. See Cisco documentation for details on how to implement RBAC.

Related Cisco documentation

For Cisco documentation, please refer to http://www.cisco.com.

- *Cisco MDS CLI Configuration Guide; Release 3.x*
- *Cisco 9000 MDS Family Configuration Guide, chapter on Configuring IPSec Network Security*
- *Cisco MDS Storage Networking Fundamentals; Version 2.1 Cisco / Firefly Communications*
- *MDS 3.0 Hardware Software Overview; March 2006 Customer Assurance Engineering*
- *EMC Connectrix MDS 9000 SAN-OS Release Notes*, available on Powerlink

EMC Select security appliances

This section discusses the following EMC Select security appliances:

- "EMC Select Decru DataFort FC-Series security appliance" on page 537
- "EMC Select Neoscale CryptoStore security appliance" on page 550

This section contains the recommended settings and topology for the Decru DataFort FC-Series security appliance and the Neoscale Crypto security appliance that sit in the storage data path. Refer to the vendor's user guide and application guide for exact implementation steps. The following recommendations are valid only for local (proximate) Fibre Channel based fabrics with block transfer storage (for example, CLARiiON and Symmetrix). These appliances implement encryption / decryption based on SCSI FCP addressing (WWN, PID, LUN, LBA). Therefore, any changes of the initiator WWN, remapping of the LUN ID, or data movement from one partition to another requires user intervention for keys management. Key Management in each Decru appliance is implemented using the Lifetime Key Management Server. For more information about key management (duplication, restore, transport, etc.), refer to the vendor's user or system administration guide.

Data mobility in the encrypted form requires data consistency. Applications must commit data transactions through any of these appliances prior to data movement. For example, a volume copy should be prepared for backup activity by coalescing the database, un-mounting or synchronizing the file system, and shutting down or quiescent relevant applications.

Disk encryption mechanism

The Advanced Encryption Standard (AES) or Data Encryption Standard (DES) encryption algorithm of the data-at-rest in the disk is restricted on the Open System Logical Block Address (LBA) boundary, which is 512 bytes block. Both Decru and Neoscale do not add encryption rule (metadata) into the encrypted block, hence there is always a one-to-one correspondence from the unencrypted (clear text) block of data to the encrypted (crypto text) block of data.

Each vendor implements different algorithm to ensure that each of the 512 bytes block is encrypted using different keys in order to ensure the security of the data. The algorithm involves the position of

the data block in the disk that is based on the LBA. As the result, the host can randomly access any LBA location through the disk encryption appliance to randomly retrieve the unencrypted data from the encrypted disk.

Impact of the storage application to the encrypted data

The disk encryption appliance provides the rule of how a particular key applies to a set of the data. This rule affects the SCSI addressing level, such as the I-T-L (Initiator-Target-LUN Nexus). The encrypted dataset does not carry any information about the rule (metadata), therefore the key must be appropriately associated with the same dataset regardless of the I-T-L mapping.

The storage application data (such as Snapview, Mirrorview, Clone, and SRDF) provides a point-in-time duplicate of data, or a synchronous duplication of data. These applications guarantee that the data is duplicated or transferred as the whole block of the Logical Block Address. The encrypted data can be transferred into another volume with different association of the I-T-L. The user must be actively involved to ensure the correct association of the key with the dataset because of the loose correlation of the keys to the dataset.

Follow these rules for encrypted data mobility:

- The encrypted data must be transferred in the whole entirety of the dataset.
- The encrypted data must be transferred to the same location of the LBA vicinity from the original source

The impact of online encryption data preparation and online encryption data re-keying with the storage applications

Consider the following scenarios between disk encryption and storage application:

- The data is currently converted from clear text to the crypto text.
- The encrypted data is replaced with other encrypted data with different keys.

In the above scenarios, part of the dataset is in either clear text (unencrypted) or crypto text (encrypted) with different key. As a result, the integrity of the point-in-time snapshot or synchronous replication of this dataset *cannot* be guaranteed.

Additional information on storage security can be found on EMC's Powerlink. Select best practices for implementing various secure SAN mechanisms are contained throughout this chapter. Some

vendor-specific best practices are included in the "Security implementation examples" on page 516 section.

EMC Select Decru DataFort FC-Series security appliance

The information in this section contains the recommended settings and topology for the implementation of the security appliance that sits in the path of the storage. Refer to the vendor's user guide and application guide for the exact implementation steps. Recommendations are only valid for local (proximate) Fibre Channel-based fabric and disk-based (block transfer) storage (CLARiiON and Symmetrix).

Both Decru appliances implement encryption / decryption based on FCP (SCSI over Fibre Channel) addressing (WWN, PID, LUN, LBA). Therefore, any changes of the initiator WWN, remapping of the LUN ID, or data movement from one partition to another, requires user intervention for keys management. For more information about keys management (duplication, restore, transport, etc.), refer to each vendor's user or system administration guide.

Data mobility in the encrypted form requires data consistency. Therefore, any application must commit the data transaction (write to disk) through any of these appliances prior to data movement. In other words, the volume copied should be prepared for backup activity (coalesce database, file-system unmounted or synchronized, application shutdown or quiescent).

Key management in each Decru appliance is implemented using the Lifetime Key Management Server.

This section provides the following information on the EMC Decru DataFort FC-series security appliance:

- "Decru virtualization" on page 538
- "Decru mapping for the encrypted storage" on page 540
- "Decru Cryptainers vault" on page 541

Following this discussion, further information is provided on supported configurations and best practices for:

- "EMC Select Decru DataFort FC-52x – Disk" on page 542
- "EMC Select Decru DataFort FC-52x – Tape" on page 546

Decru virtualization

The Decru DataFort appliance operates as a FC proxy for either the tape or disk encryption. As a FC proxy, all the ports in the DataFort FC-520, FC-525, and FC-1020 (both the host and storage ports) will assume either an N_Port or an NL_Port. Both the storage ports can automatically negotiate between 1 Gb/s and 2 Gb/s. If the Decru appliance is configured as a FC proxy that provides a single FC node, then the ports are configured as N_Port (Decru's terminology is *Single-ID mode*). In this mode, the Decru can only emulate one node from both host and storage ports.

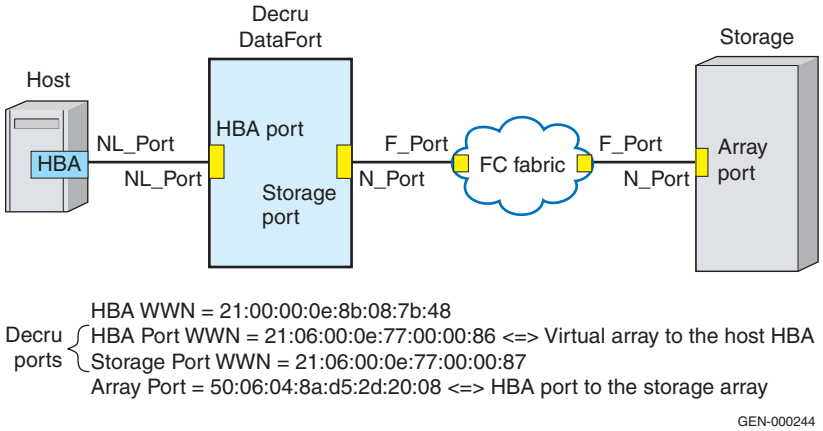
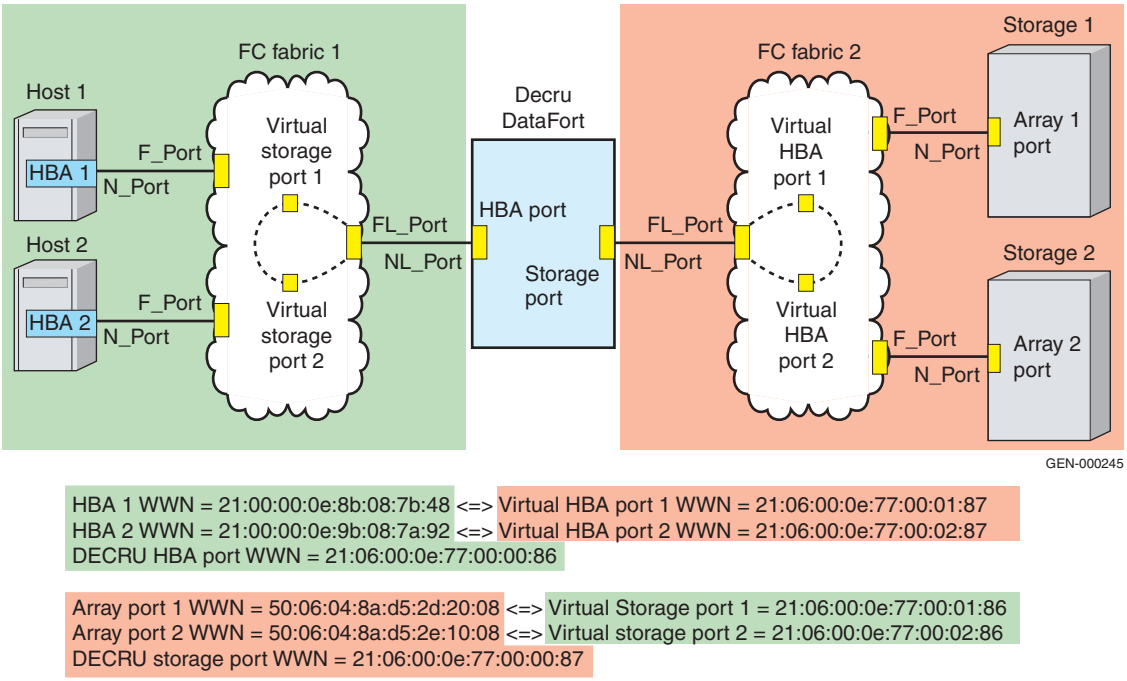


Figure 107 **Single-ID mode (virtualization disabled)**

If the Decru DataFort is configured as an FC proxy that provides emulation for multiple FC nodes, then the port will be configured as a NL_Port (Decru's terminology is *Multi-ID mode*). As the result of multi-ID mode, Decru DataFort will present several loop nodes that will emulate virtual nodes. If a storage port is virtualized, DataFort will then present a **Virtual Device** to the host through the host port. If a host port or an HBA port is virtualized, DataFort will then present a **Virtual Host** to the storage through the storage ports.



HBA 1 WWN = 21:00:00:0e:8b:08:7b:48 <=> Virtual HBA port 1 WWN = 21:06:00:0e:77:00:01:87
HBA 2 WWN = 21:00:00:0e:9b:08:7a:92 <=> Virtual HBA port 2 WWN = 21:06:00:0e:77:00:02:87
DECRU HBA port WWN = 21:06:00:0e:77:00:00:86

Array port 1 WWN = 50:06:04:8a:d5:2d:20:08 <=> Virtual Storage port 1 = 21:06:00:0e:77:00:01:86
Array port 2 WWN = 50:06:04:8a:d5:2e:10:08 <=> Virtual storage port 2 = 21:06:00:0e:77:00:02:86
DECRU storage port WWN = 21:06:00:0e:77:00:00:87

Figure 108 Multi-ID mode (virtualization enabled)

Decru virtualization will always utilize a NL_Port to represent or virtualize multiple devices connected to its host or storage ports. To accomplish virtualization, DataFort uses the emulation of private loops in each of its ports. As a result, the FC switch must be able to translate the addresses of the emulated devices in the Arbitrated Loop to the address of the FC fabric. The typical switch port that translate between the arbitrated loop address and the fabric address is identified as an FL_Port. Hence, the Decru virtualization configuration requires the support of the FL_Port on the FC switch. For direct connection to either the host, storage, or tape ports, the Decru DataFort ports must be configured as a NL_Port (Decru's terminology is *Multi-ID mode*). The HBA port, storage, or tape ports must be configured as L_Port.

For complete information about configuring the virtualization settings, please see the "DataFort Setting and Status" section in the *Decru Fibre Channel Series Administration Guide.*

Decru mapping for the encrypted storage

The concept of virtualization or emulation of a storage node in the storage port allows the mapping of a WWN storage node to a WWN of a virtual node in the DataFort host port (Decru's terminology is *Port-Mapping*). Port-mapping preserves the same LUN numbering from the storage node to the virtualized storage node in the DataFort. This feature is very useful in the disk environment where multipathing software or management software is being used.

For tapes, it is useful to be able to sync up backup apps, etc, to consistent LUN numbering. However, port-mapping is less important in the tape environment unless there are multiple tape libraries that each contain both the robot LUN and multiple tape LUNs.

Please note that the number of ports that can be mapped depends on the DataFort capacity for mapping. For more information on the capacity of the Decru DataFort please see the "Planning the Network Configuration"section in the *Decru Fibre Channel Series Administration Guide*.

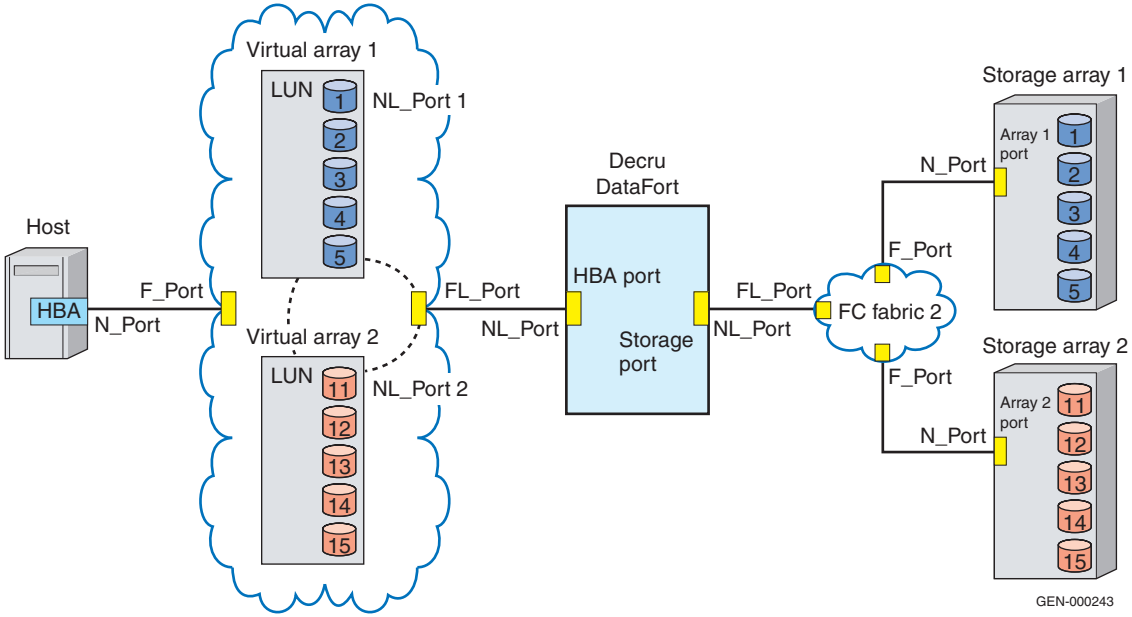


Figure 109 DataFort port mapping (storage side virtualization enabled)

Another option of DataFort mapping is known as *LUN-mapping*. LUN-mapping allows DataFort to assign a virtual LUN number to each device port (such as, storage port, robotic arm port, or tape drive port). As a result of LUN-mapping, DataFort can present a virtual storage node with multiple LUNs. For example, LUN0 can represent a robot arm while LUN1, LUN2, and LUN3 can represent tape drives. In reality, each of these LUNs are actually FC ports that are part of the tape library.

Another use of LUN-mapping is to provide the capability to map more storage ports than the maximum number of virtual devices that DataFort can virtualize.

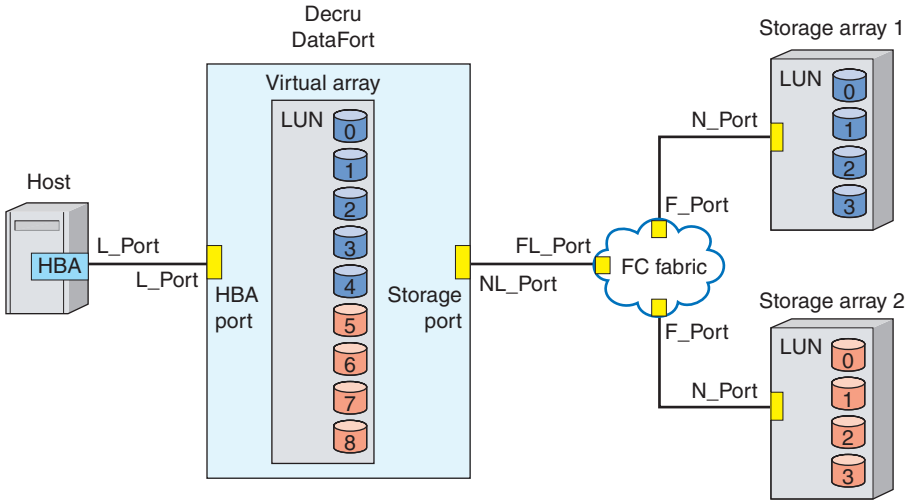


Figure 110 **DataFort LUN mapping (storage side virtualization enabled)**

Decru Cryptainers vault

Decru DataFort uses a concept called *Cryptainers* in order to manage key assignments, access control, and other security-related features. The Cryptainers are associated with storage WWNs and LUNs. Information that is related to Cryptainers is stored in a specific region of encrypted storage that has default properties in a tape environment and configurable properties in disk storage environments. The Decru DataFort term for the specific region is *Metadata*. DataFort appliances automatically associates the Cryptainers to the storage LUN when the Metadata exist. Please refer to the "Introduction of Decru" section in the *Decru Fibre Channel Series Administration Guide* for further description of the Cryptainers.

EMC Select Decru DataFort FC-52x – Disk

Supported configurations

The supported DataFort configuration for EMC disk arrays implements is *storage virtualization* and *host virtualization* mode. DataFort architecture allows up to seven hosts and seven storage devices to be virtualized by one DataFort.

Storage virtualization is the act of virtualizing the storage in the DataFort host-side (storage) port. This is the defaut setting on the DataFort. The storage virtualization mode means that hosts actually see multiple Arbitrated Loop Nodes. Each of these Arbitrated Loop Nodes corresponds to a virtualized EMC storage port. In order to achieve this configuration, the virtualization for the storage and host port must be set to ON. The use of the storage virtualization mode enables DataFort to encrypt EMC storage using the port-mapped Cryptainer (one Cryptainer vault per LUN). With a Symmetrix array, LUN 0, which does not need to be encrypted, is used as the gateway for the Symmetrix masking application.

Host virtualization is the act of virtualizing a host on the DataFort storage-side [initiator] port. Host virtualization is *not* enabled by default. Enabling host virtualization is recommended especially for clustered hosts or shared volume environments having <= 7 hosts. Note that virtualization is only supported on FC switches that support FL connections or on storage ports that support AL connections. Please refer to Decru DataFort Disk best practices and release notes for more information.

Host virtualization is beneficial, or required, for the following common scenarios:

- Multiple hosts sharing same LUN.
- Preserving existing storage array masking by combining host virtualization with the option to forward host's WWN.

  Note: Using the WWN forwarding feature is supported only if hosts and storage ports are on independent fabrics or VSANs, or if hard zoning is used.

- LUN masking in storage arrays based on host OS type.

  Host virtualization allows you to specify different host types for each virtualized host presented on the DataFort initiator loop, which is useful for this scenario.

Please refer to the *Decru Fibre Channel Series Administration Guide* for details regarding configuration of virtualization and mapping.

Note: Migration to an encrypted storage environment utilizing Decru DataFort is disruptive in nature. The disruption occurs due to zoning requirements for host-to-virtual storage nodes, and virtual host nodes-to-storage ports. In addition, the migration also requires changing the access control setting of the storage, to the Decru DataFort storage port. After zoning changes, cryptainer assignment, and granting of access controls, Decru DataFort for Disk does not affect any host applications.

For complete information about configuring Decru DataFort with EMC storage, refer to the "Troubleshooting" section in the *Decru Fibre Channel Series Administration Guide*.

In order to provide a high availability disk encryption configuration, a cluster of minimally two DataForts must be implemented in a high availability SAN configuration. Refer to the "DataFort Initialization" section in the *Decru Fibre Channel Series Administration Guide* for more information about setting up the cluster.

Based on a clustered DataFort configuration, the following list contains the supported EMC storage configurations:

- EMC Symmetrix host bit settings must be set to default Symmetrix Fibre Channel open system host bit settings (VCM, PtoP, auto speed, and SC3 enabled).
- CLARiiON host settings must be similar to the default CLARiiON Fibre Channel open system host settings (host type 1, failover mode 1, communication port enabled).
- EMC Symmetrix with multipathing (at least one EMC Symmetrix port per path).
- EMC CLARiiON with multipathing (at least one corresponding SP port per path).
- EMC PowerPath is required for high availability configurations for both Symmetrix and CLARiiON. Host multi-pathing software is required to route the IO between paths because clustered DataForts do not route IO between multipaths.
- CLARiiON requires PowerPath to support nondisruptive upgrades (NDU).

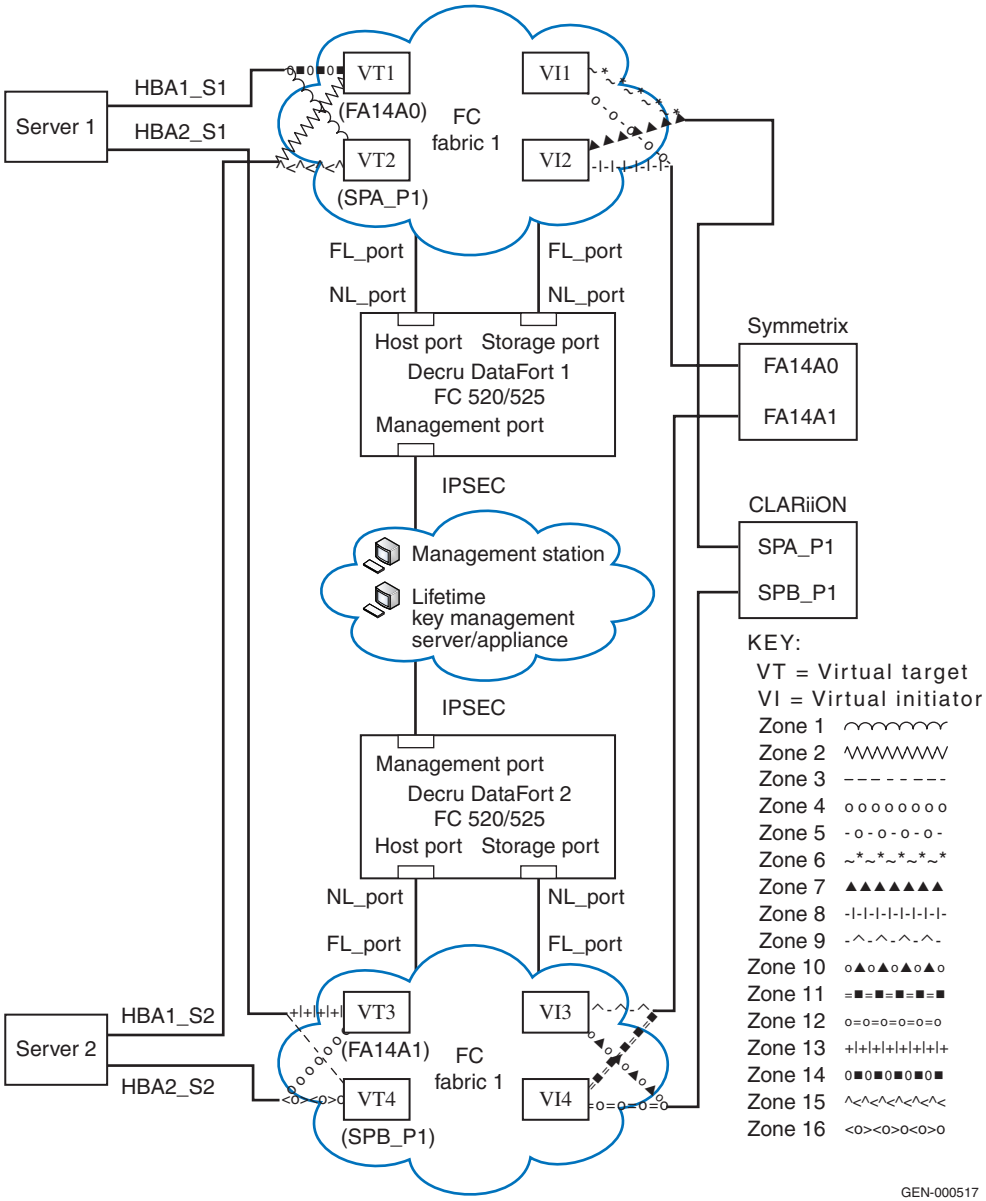


Figure 111 Recommended configuration of Decru DataFort with EMC storage

The following are recommended SAN topologies:

- Fibre Channel switches that support NL_Port nodes
- Brocade Silkworm
- Cisco MDS
- Brocade M-Series non-director switches, such as Sphereon 4500

  When using DataFort with a Brocade M-Series director, implement an edge switch to provide the FL_Port connectivity.

Note: All switch ports require auto-topology or FL_Port settings.

Best practices

This section lists best practices to consider.

- Although zone configuration can use both WWN zoning and port zoning, best practice suggests port zoning that allows multiple loop IDs in a single port.
- DataFort must be in storage virtualization mode, where the host is actually seeing multiple loop IDs, where each loop ID corresponds to a virtualized EMC storage port.
- Decru recommends adding all possible AL_PAs that can exist on the host port, such as WWN zoning (six different WWNs) for the port connected to the host. For example:

```
NL          a90725;      3;21:06:00:0e:77:00:00:86;20:06:00:0e:77:00:00:86;na
      PC4s: PCP [Decru                    _Secure_      0001]
      Fabric Port Name: 20:07:00:60:69:90:02:bd
NL          a90726;      3;21:04:00:0e:77:00:00:86;20:04:00:0e:77:00:00:86;na
      PC4s: PCP [Decru                    _Secure_      0001]
      Fabric Port Name: 20:07:00:60:69:90:02:bd
NL          a90727;      3;21:02:00:0e:77:00:00:86;20:02:00:0e:77:00:00:86;na
      PC4s: PCP [Decru                    _Secure_      0001]
      Fabric Port Name: 20:07:00:60:69:90:02:bd
NL          a90729;      3;21:00:00:0e:77:00:00:86;20:00:00:0e:77:00:00:86;na
      PC4s: PCP [Decru                    _Secure_      0001]
      Fabric Port Name: 20:07:00:60:69:90:02:bd
```

- Cryptainer mapping for EMC storage must use port mapping. As a result, each EMC storage port will be mapped to a virtual storage port. For high availability, map the LUNs from the secondary storage port as the alternate paths to the associated Cryptainers of the same LUN from the primary storage.

The *Decru DataFort Fibre Channel Series Administration Guide* contains more information about storage virtualization mode and cryptainer mapping.

- Cryptainers that provide security in the path between host and storage must be clustered together.
- DataFort provides a method to combine the references for any of the host and storage ports as a *Collection*. For complete information about configuring the Decru DataFort with EMC storage, please review the "Troubleshooting" sectopm in the *Decru Fibre Channel Series Administration Guide*.
- Access control implementation with EMC storage:
  - Requires modification of the zone and the access control list of the Symmetrix and CLARiiON arrays.
  - Requires all EMC storage LUNs to grant the access from DataFort Storage ports.
  - Uses the DataFort GUI/CLI for individual host access to the secured LUNs controlled by the DataFort Cryptainers.
- Each virtualized storage port can support up to 255 LUNs (one Cryptainer per LUN).
- Be aware of the following configuration limits:
  - Seven storage ports per Decru box, because DataFort can virtualize only seven storage ports to the host.
  - Maximum fan-out for non-virtualized host ports to the DataFort port is 32 initiator ports (Fibre Channel HBA ports) per Decru box.

Note: The introduction of a high fan-out number will increase the latency of frame transmission. The user must be aware of the scalability impact to the host application, path recovery, and storage application.

References

- The *Decru Fibre Channel Series Administration Guide* contains more information about storage virtualization mode and cryptainer mapping.
- The *Decru DataFort Fibre Channel Series CLI Guide*.

EMC Select Decru DataFort FC-52x – Tape

Decru DataFort FC-1020 provides five independent Fibre Channel host ports and five independent storage ports. However, the DataFort FC-1020 host and storage ports are engineered only for tape

encryption (no disk encryption). The FC-520, FC525, and FC-1020 can be clustered together to provide high availability for tape encryption.

Supported configuration

The supported DataFort configuration for EMC tape storage is the *Storage Virtualization Mode*. However, since tape is removable media, Decru DataFort provides a different scope of key assignments such as *Tape Pool Collection*, *Tape Pool*, or *Tape Device*. As a result, the Cryptainer configuration is different from the disk configuration. Refer to the "Tape Management" section of the *Decru Fibre Channel Series Administration Guide* for more information on tape key management. For supported configuration information, refer to the *Decru DataFort FC520, FC525 for Disk*.

Note: Migration to the encrypted storage environment utilizing the Decru DataFort is disruptive in nature. The disruption happens due to zoning requirements for host-to-virtual storage nodes and virtual host nodes-to-storage ports. In addition, the migration also requires the storage to change the access control setting to the Decru DataFort storage port (some virtual tape library environments). After the zoning changes, cryptainer assignment, and granting of access controls, Decru DataFort for tape will not affect any of the host applications. For the complete information about configuring Decru DataFort with EMC storage, refer to the "Troubleshooting" section in the *Decru Fibre Channel Series Administration Guide*.

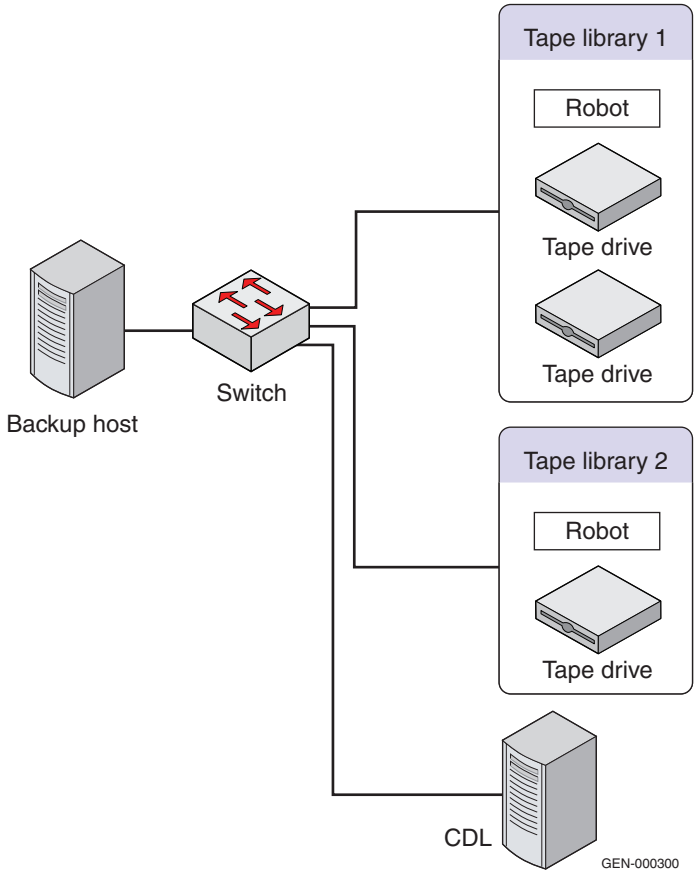


Figure 112 Topology without DataFort example

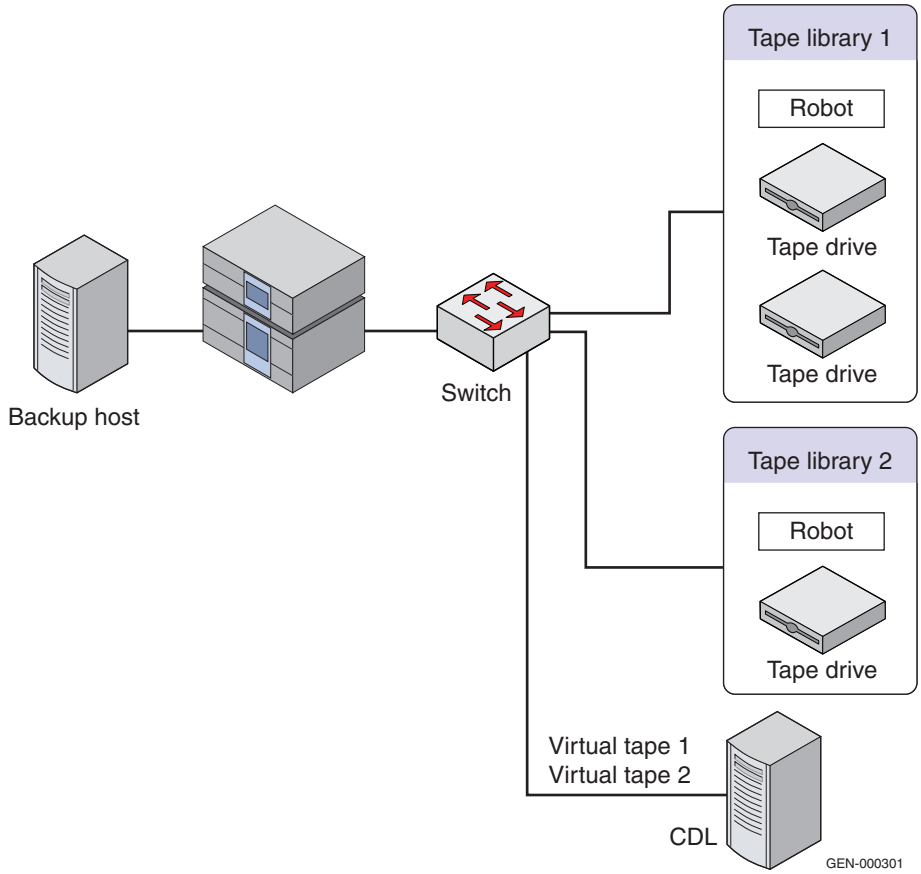


Figure 113 **Virtualization enabled with port mapping**

Best practices

This section lists best practices to consider.

- Although the DataFort is capable of providing separate paths for high availability configurations with tape encryption, DataFort depends on either user intervention or on an application mechanism to provide failover from one storage path to another. Therefore the high availability configuration for DataFort tape encryption should be recognized as hot-spare.
- Decru DataFort provides compression for the tape media in addition to encryption. Due to the nature of encryption, compression of cipher-text data can actually expand the data set.

Because of the proxy nature of the DataFort appliance, DataFort enables or disables the compression based on the host application setting.

- If the host HBA accommodates FC-Tape, Decru DataFort tape encryption appliances can accommodate the FC-Tape recovery method. It is recommended that you enable FC-Tape negotiation because the recovery for the disruption using FC-Tape is more efficient.

For further tips, limitations, and workarounds refer to the *Decru DataFort Fibre Channel Series Release Notes* and the "Troubleshooting" section of the *DecruFibre Channel Series Administrator Guide*.

References

- The *Decru Fibre Channel Series Administration Guide* contains more information about storage virtualization mode and cryptainer mapping.
- The *Decru Fibre Channel Series CLI Guide*.

EMC Select Neoscale CryptoStore security appliance

This section contains the recommended settings and topology for the implementation of the security appliance that sits in the path of the storage. Refer to the vendor's user guide and application guide for the exact implementation steps. Recommendations are valid only for local (proximate) Fibre Channel-based fabric and disk-based (block transfer) storage (CLARiiON and Symmetrix).

Both of the Neoscale CryptoStore appliances implement encryption / decryption based on FCP (SCSI over Fibre Channel) addressing (WWN, PID, LUN, LBA). Therefore, any changes of the initiator WWN, remapping of the LUN ID, or data movement from one partition to another, requires user intervention for keys management. For more information about keys management (duplication, restore, transport, etc.), refer to each vendor's user or system administration guide.

Data mobility in the encrypted form requires data consistency. Therefore, any application must commit the data transaction (write to disk) through any of these appliances prior to data movement. In other words, the volume copied should be prepared for backup activity (coalesce database, file-system un-mounted or synchronized, application shutdown or quiescent).

This section contains security implementation information for the following:

- "EMC Select Neoscale CryptoStor FC-2002 for Disk," next
- "EMC Select Neoscale CryptoStor FC702/704 for Tape" on page 553

EMC Select Neoscale CryptoStor FC-2002 for Disk

CryptoStor FC2002 is a security appliance that provides *FC pass-through* behavior. *FC pass-through* has the following characteristics:

- The FC2002 appliance does not provide ports that model any of the traditional Fibre Channel topologies such as: N_Port, L_Port, F_Port, E_Port, B_Port, and so on.
- The FC2002 ports provide speed auto-negotiation with the connected port. The negotiated speed is the common highest speed for both of the ports. The FC2002 is not capable of providing different speeds between the two ports.
- FC 2002 ports carry the mode of the port-pairs without introducing any Fibre Channel topology termination. Hence, connecting both of the ports to the same fabric may introduce unknown behavior to the related switch.

Note: Migration to the encrypted disk environment for Neoscale CryptoStor FC-2002 *does not* cause lengthy disruption. Because of the transparent nature of the FC2002, the disruption is limited to the cabling of the CryptoStor FC2002 appliance between the switch port and the storage port. For complete information about configuring the Neoscale CryptoStor FC2002W, refer to the "Overview" section in the *CryptoStor Fibre Channel Series Administration Guide.*

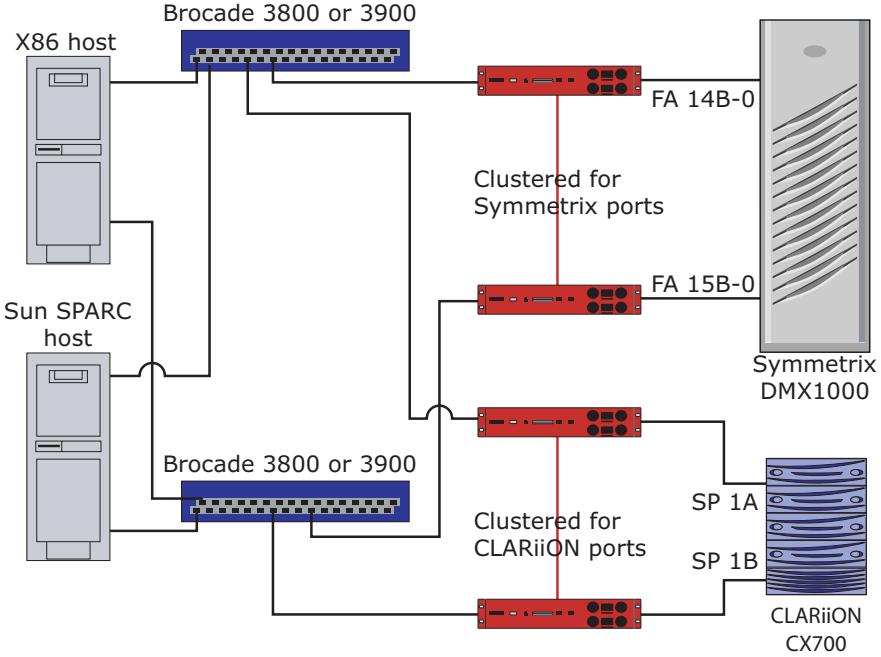
In order to provide a high-availability disk encryption configuration, a cluster of a minimum two CryptoStor FC2002 appliances must be implemented in the High Availability SAN configuration. Refer to the "System Administrator Handbook" section of the *CryptoStor Fibre Channel Series Administration Guide* for more information about setting up the cluster. The following list contains EMC-supported storage configurations based on clustered CryptoStor configurations:

- EMC Symmetrix configuration bit according to recommendation for the platform operating system of the connected host. Please refer to "Symmetrix Fibre Bit Settings" in the *EMC Support Matrix*.

- EMC CLARiiON configuration setting according to the recommendation for the operating system platform of the connected host. Refer to host connectivity guides for each operating system platform.
- EMC Symmetrix with multipathing (at least one EMC Symmetrix port per path).
- EMC CLARiiON with multipathing (at least one corresponding SP port per path).
- EMC PowerPath required for high availability configurations for both Symmetrix and CLARiiON. The host multipathing software is required to route the IO between the paths.
- CLARiiON requires PowerPath to support non-disruptive upgrades (NDU)..

Same encryption keys must be applied to the redundant Symmetrix and CLARiiON ports.

Encrypted Symmetrix LUNs must be mapped to both FA 14B-0 and FA 15B-0.



Paths between arrays and FC-2002s are clear text. All other paths are encrypted text.

Figure 114 Recommended CryptoStor 2002 configuration for EMC storage products

Supported configurations

The following are supported SAN configurations:

- Neoscale FC2002 connected inline between the hosts and the fabric (N_Port or NL_Port)
- Neoscale FC2002 connected inline between the fabric and the storage (N_Port or NL_Port)

Note: CryptoStor FC-2002 is a true transparent appliance. Therefore, there is no way to zone the CryptoStor FC-2002 ports.

Note: Zone changes and volume access control changes are not necessary between the existing host and the existing storage port.

Limitations

Note the following configuration limits:

- The limit of the host fan-out depends on the EMC storage capability.

Note: The introduction of a high fan-out number will increase the latency of frame transmission. The user must be aware of the scalability impact to the host application, path recovery, and storage application.

- All possible paths where the host port is connected to the storage port must be secured by a clustered CryptoStor FC-2002.
- Each LUN in the redundant paths must use the same Data Encryption Policy (therefore, the same key).

Best practices

Clustered Cryptostor FC 2002 must be connected on every single path where the host IO can be routed.

References

- The *CryptoStor Fibre Channel Series Administration Guide* contains detailed information on setting the date encryption poligy.
- *CryptoStor Fibre Channel Series Technical Reference Manual.*

EMC Select Neoscale CryptoStor FC702/704 for Tape

Neoscale CryptoStor for Tape functions as the proxy device between the backup host and the tape library. As the proxy device, both the CryptoStor FC702/FC704 host, and storage ports can be configured as N_Port (Point-to-Point), NL_Port (Arbitrated Loop), and auto (try Loop first before Point-to-Point). In addition to the topology, the host and storage ports are capable of supporting 1 GB/s and 2 GB/s Fibre Channel speeds. By default all of the ports are auto topology and auto

speed negotiation. The tape port virtualization requires that the host port must be configured as the NL_Port in order to emulate the tape library.

Supported configuration

The following are supported SAN configurations:

- The CryptoStor FC702 / FC704 supported configuration is multi-ID on the host port (NL_Port), and single-ID on the target port (N_Port). Using this configuration, multiple tape libraries can be emulated as a single virtual device.

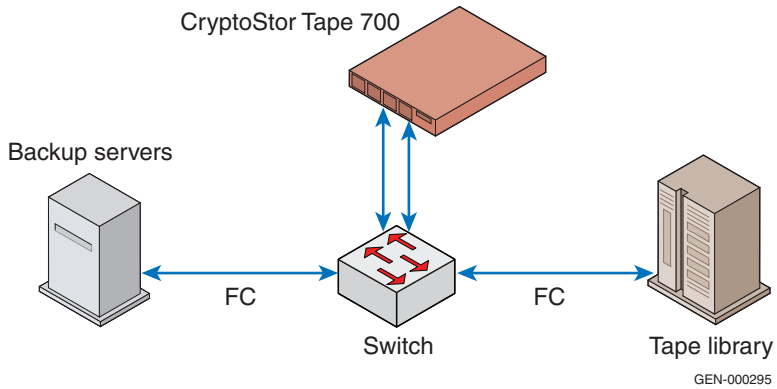


Figure 115 **CryptoStor Tape 700 deployment example**

Note: Migration to the encrypted tape environment for Neoscale CryptoStor FC Tape 700 series is disruptive in nature. The disruption happens due to the requirement of the change for the zoning both for the hosts to the virtual storage nodes and the virtual host node to the tape library ports. Furthermore, the migration also requires the tape library to change the access control setting to the Neoscale CryptoStor FC 700 series storage port (some virtual tape library environments). After the zoning changes, volume pools assignment, rules assignment, and access controls had been granted, Neoscale CryptoStor for Tapes will not affect any of the host applications. For complete information about configuring the Neoscale CryptoStor Tape 7000 series, please review the "Overview" section in the *DecruFibre Channel Series Administrator Guide*.

Note: CryptoStor FC-2002 is a true transparent appliance. Therefore, there is no way to zone the CryptoStor FC-2002 ports.

Note: Zone changes and volume access control changes are not necessary between the existing host and the existing storage port.

CryptoStor FC tape encryption appliance use three level key hierarchies:

- System key — The highest level. The key resides within the appliance and can be exported only via Smartcard. It encrypts any keys and is stored on an internal disk, Secure NVRAM, or Active Memory.
- Encryption and Pool keys — Second level. Pool / Encryption keys are used to encrypt the tape keys. The encryption / pool keys are encrypted by the System key and are backed up in a configuration file as well as the key catalog. Pool keys can be exported as part of the configuration onto a smartcard or via a file in the catalog form.
- Tape keys — Third level. The tape keys are encrypted by the encryption/pool keys. The tape keys are used to encrypt data before writing it to tape. They can reside in encrypted form on either the media (cassette) or in the catalog.

Limitations

Note the following configuration limits:

- The limit of the host fan-out depends on the EMC storage capability.

Note: The introduction of a high fan-out number will increase the latency of frame transmission. The user must be aware of the scalability impact to the host application, path recovery, and storage application.

- All possible paths where the host port is connected to the storage port must be secured by a clustered CryptoStor FC-2002.
- Each LUN in the redundant paths must use the same data encryption policy (therefore, the same key).

Best practices

This section lists best practices to consider.

- The storage port can only support FC-tape if, and only if, the port is configured as point-to-point (Single ID). Therefore, host port must be configured as point-to-point when FC-tape is expected to be implemented against the tape library port.

- It is always considered a best practice to store the tape keys in the catalog away from the data they protect. However, since tape is a removable medium and due to the limited internal disk capacity of CryptoStor FC tape appliance, the customer might opt to store the tape keys on the tape itself. If the user opts to save the tape keys on the internal catalog of the CryptoStor, then a daily backup of the catalog is highly recommended. This, however ,might very soon lead to lack of free space on the internal disk especially if the customer is using dynamic keys.

References

- *CryptoStor Tape FC 700 Administration Guide*
- *CryptoStor Tape FC 700 Installation Guide.*

5

Backup and Recovery

This chapter provides the following information on backup and recovery.

These case studies are distributed by EMC for information purposes only. EMC does not warrant that this information is free from errors. No contract is implied or allowed.

- Overview of backups 558
- Direct-attached backups 559
- LAN-based backups 561
- LAN-free backups 566
- Serverless backups 571
- Backup over long distances using FCIP and routers 574
- FCIP tape acceleration with EMC Connectrix MDS switches ... 576
- Brocade 7500 and CDL over IP with Fastwrite acceleration and tape pipelining 579
- NAS backups 582
- Backup to disk 583
- Backup media: Tape libraries 584
- Mirrored fabric backup solution 586
- Tapes and fabrics 591

Overview of backups

Terminology

The following terminology is used throughout this section.

- Back up — Create a copy of data.
- Archive — Move data with low usage patterns to a slower media for historical or other purpose.
- Restore — Copy a file or group of files from a backup to the primary storage location.
- Recover — Rebuild a system or data center from backups.

Why perform backups?

Backups are needed as an insurance policy against loss of data, which can occur because of:

- Hardware failures
- Human error
- Application failures
- Security breaches, such as hackers or viruses

High-availability storage arrays have reduced the need to recover data because of hardware failures. Hardware availability features can protect data from loss due to hardware failures; however, these availability features cannot protect against the other factors that can result in loss of data.

Backups are sometimes used as an archive; for instance, government regulations require that certain financial data must be kept for a specific number of years. In this context, a backup also becomes an archive.

Backup architectures

Refer to the appropriate sections for more information on these backup architectures:

Traditional backup architectures:

- "Direct-attached backups" on page 559
- "LAN-based backups" on page 561

SAN backup topologies:

- "LAN-free backups" on page 566
- "Serverless backups" on page 571
- "NAS backups" on page 582

Also refer to "Backup media: Tape libraries" on page 584.

Direct-attached backups

Many organizations started with a simple backup infrastructure called direct-attached. This topology is also sometimes referred to as *host-based* or *server-tethered* backup. Each backup client has dedicated tape devices. Backups are performed directly from a backup client's disk to a backup client's tape devices.

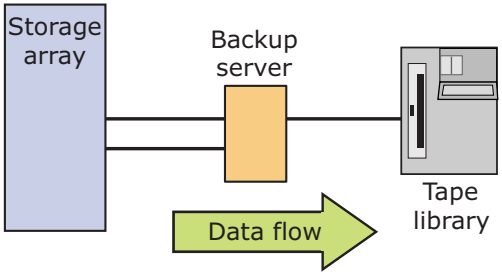


Figure 116 **Data flow for direct-attached backups**

Advantages

The key advantage of direct-attached backups is speed. The tape devices can operate at the speed of the channels. Direct-attached backups optimize backup and restore speed, since the tape devices are close to the data source and dedicated to the host.

Disadvantages

Direct-attached backups impact the host and the application, since backups consume host I/O bandwidth, memory, and CPU resources.

Direct-attached backups are generally better suited for smaller environments. Growth rates in servers and data can cause direct-attached backups to become costly and difficult to manage and operate. Organizations with large growth may experience some or all of the following issues with direct-attached backup infrastructures:

- Large numbers of tape devices might be underutilized.
- A wide variety of backup media might be in use.
- Operators could find it difficult to manage tape. Tape devices might be scattered between floors, buildings, or entire metropolitan areas.
- Each server might have unique (and possibly locally created) backup processes and tools, which can complicate backup management and operation.

- It might be difficult to determine if everything is being backed up properly.
- Dispersed backups, multiple media types, diverse tools, and operational complexity can challenge the task of business continuance recovery.

Migration paths

Organizations that have outgrown direct-attached backup infrastructures have several migration paths, described in these sections:

- "LAN-based backups" on page 561
- "LAN-free backups" on page 566

Improvement options

The following can also improve backup speed and efficiency:

- Utilize TimeFinder™ BCVs to reduce production impact
- Implement faster devices, including disk and disk libraries
- Improve disk and tape pathing

LAN-based backups

LAN backup infrastructures can be configured similar to the schematic illustrated in Figure 117.

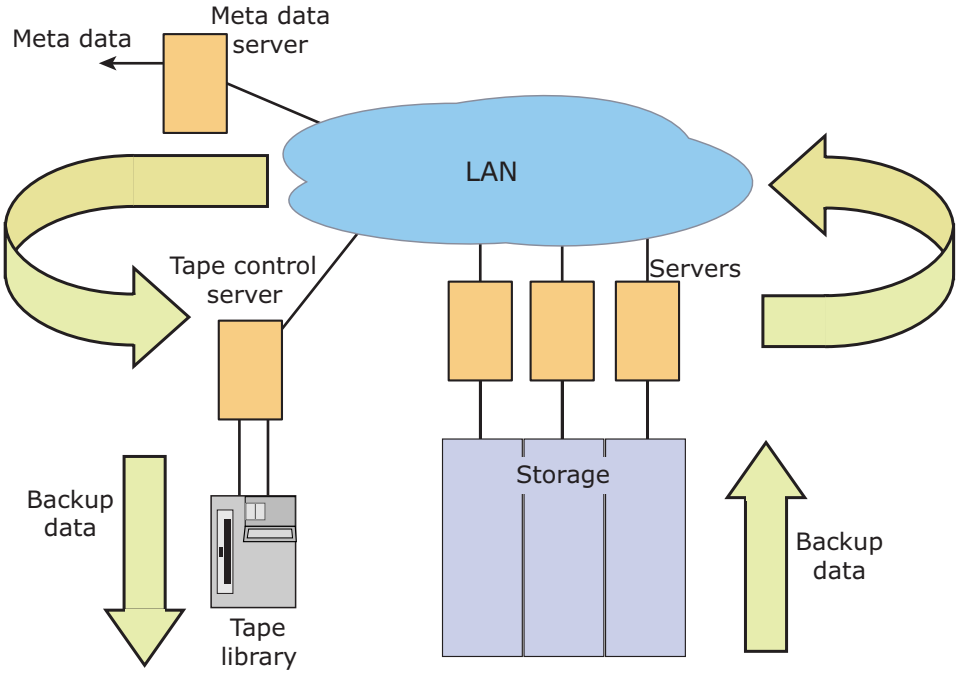


Figure 117 LAN backup infrastructures

The *metadata server* is the central control point for all backups. This is where the metadata (also known as *tape catalog* or *backup index*) and backup policies reside. *Tape control servers* manage and control backup devices, and are controlled by the meta data server. (A metadata server can also be a tape control server.) The primary use of LAN backup topologies is to centralize and pool tape resources.

Backup process overview

The backup process is as follows:

1. The metadata server invokes backup client processes on the backup client.
2. The tape control server places tapes into the tape drives.
3. The backup client determines which files require backup.

4. The backup client reads the backup data from disk and writes the backup data to the LAN.
5. The tape control server reads the backup data from the LAN and writes the backup data to the tape.
6. The backup client and the tape control servers sends meta data information to the meta data server, including what was backed up and which tapes the backups used.
7. The metadata server stores the meta data on disk.

Advantages

The key advantages of LAN-based backups compared to direct-attached backups are:

- Reduced costs — Pooling tape resources improves tape device utilization and reduces the number of tape drives required, which also results in fewer host bus adapters.

  Some small servers may require backups; because of tape drive cost or limited card slot availability, however, it might not be practical to dedicate a tape drive to one of these systems. LAN backups can address these issues.

- Improved management and operability — Centralized backups reduce management complexity; there are fewer resources to manage, and they are all in one place.

  Centralizing tape resources into tape control servers improves the productivity of the operations staff, especially when backup clients are scattered across floors of a building, campuses, or cities. Operability can be improved further by utilizing automated, robotic tape libraries.

Disadvantages

A LAN-based infrastructure introduces some disadvantages to the backup process:

- Backups impact the host and the application.
- A LAN-based backup adds two additional data movement steps.
- Backups consume host I/O bandwidth, memory, LAN, and CPU resources.
- There could be network issues.
- A LAN-based backup might require dedicated media servers.
- There could be restore and cloning issues.

Additional data movement steps

LAN backups require two additional data movement steps to put the backup data on tape, as illustrated in Figure 118.

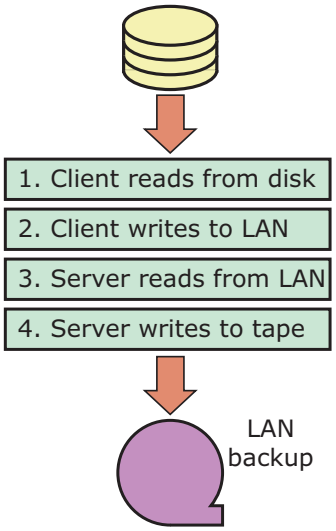


Figure 118 **LAN backup: Additional steps**

Additional CPU and memory resources are required on the backup client (compared to directly connected tape devices) to comply with network protocols, format the data, and transmit the data over the network. Note that restore processing in a LAN environment is identical except that the data flows in the opposite direction.

Resource consumption

Like direct-attached backups, LAN backups consume CPU, I/O bandwidth, and memory. Since the final destination of the backup data resides elsewhere on the LAN, additional CPU is required on a tape control server. LAN bandwidth is also required.

Network issues

LAN backups will generally not perform as well as direct-attached backups. Additional data movement steps, network protocol overhead, and network bandwidth limits reduce the speed of backups. If the network segment is not dedicated to backups, the backup performance can be erratic, since it is vulnerable to such other network activity as large FTPs, video, audio, and email.

Even the fastest available network connections can be overwhelmed by a few disk connections. Backup disk I/O consists of intense read

activity. Modern cached disk arrays like Symmetrix process I/O as fast as the channels will allow. Cache arrays with two Ultra SCSI or Fibre Channel connections are capable of exceeding the theoretical and practical rates of even the faster networking technologies. A single logical disk per path can impact the network for lengthy bursts, and multiple logical disks can saturate the network for long periods.

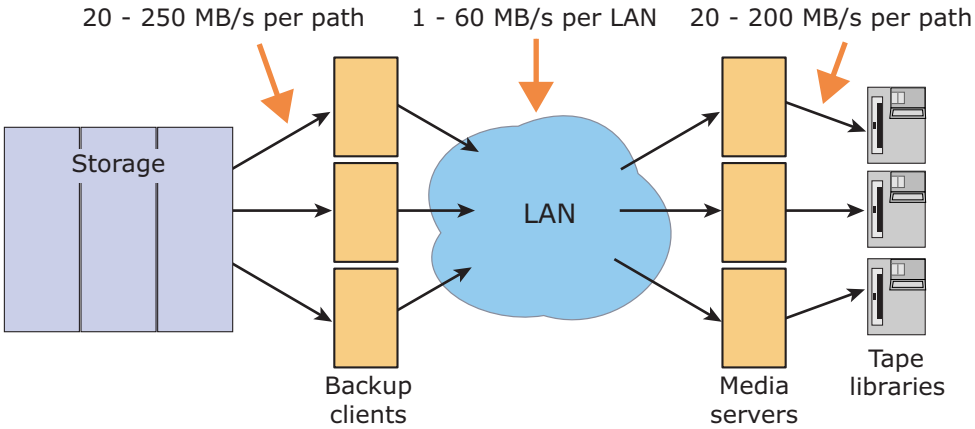


Figure 119 Bandwidth used in LAN backups

Environments that back up many logical disks to many tape libraries will constrain even the fastest network technologies. Adding additional LAN bandwidth may not always be technically feasible, since there are often limits on how many high-speed NICs (network interface cards) a server can support.

LAN backups can increase management and troubleshooting complexity. Performing backups through firewalls can be a challenge. Troubleshooting may require engagement with operations personnel, system administrators, storage administrators, and network administrators to resolve a problem.

Possible requirement for dedicated tape control servers

LAN-based backups can require dedicated tape control servers to drive the tape devices and to act as a tape central control point. Many larger organizations implement meta data servers with no tape devices, along with dedicated tape control servers for the tape libraries. Some tape device maintenance (especially SCSI) can require the server to be taken out of service. If there are multiple tape robots connected to dedicated tape control servers and the meta data server

is kept separate, restores and other backups can continue while maintenance is performed on the tape device. Dedicated tape servers are not always a technical requirement, but they are quite often an operational requirement.

Migration paths

Organizations that have outgrown direct-attached backup infrastructures have several migration paths, including the following, described elsewhere:

- "LAN-free backups" on page 566

Improvement options

The following can also improve backup speed and efficiency:

- Utilize TimeFinder BCVs to reduce production impact
- Implement faster devices, including disk and disk libraries
- Improve disk and tape pathing

LAN-free backups

LAN-free backups (Figure 120) utilize SAN (Storage Area Network) technology in conjunction with backup software that supports tape pooling. The high-speed and extended distance capabilities of Fibre Channel are used for the backup data movement path. Meta data is still moved over the LAN to the backup meta data server. This traffic is typically light and insignificant in relation to the large amounts of data moved during a backup. In a LAN-free architecture, the meta data server is also the control point for the robotic mechanism of the tape library.

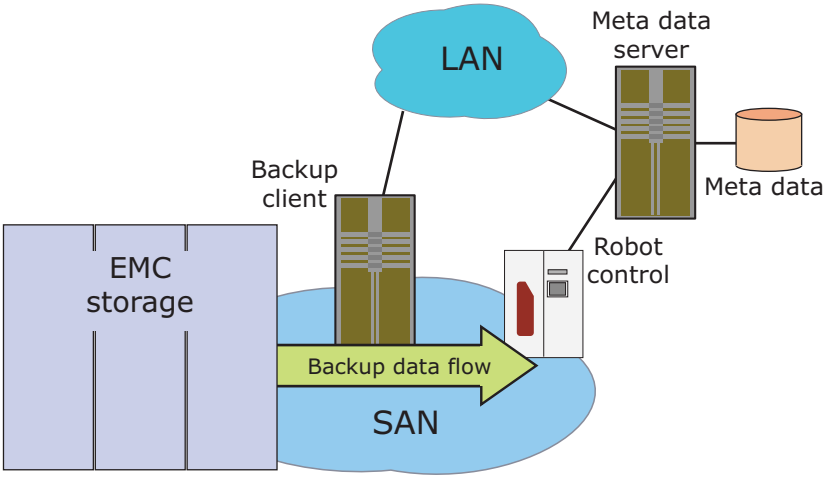


Figure 120 **LAN-free backup**

Backup process overview

The backup process is as follows:

Note: The order of steps 2, 3, and 4 may vary depending on the backup tool that is utilized.

1. The meta data server invokes backup client processes on the backup client.
2. If tape pooling is used, the meta data server assigns tape devices to the backup client.
3. The meta data server instructs the tape robot to load tapes into the assigned drives.
4. The backup client determines which files require backup.

5. The backup client reads the backup data from disk through the SAN and writes backup data to the tape device through the SAN.
6. The backup client sends meta data information to the meta data server, including what was backed up and what tapes the backups used.
7. The meta data server stores the meta data on disk.

In theory, LAN-free backups can be implemented without tape pooling capabilities. This would provide the benefits of Fibre Channel performance and distance. This approach would essentially be a direct-attached backup at extended distance and would not address the tape utilization and management issues associated with direct-attached backups.

Organizations that are evolving from a direct-attached backup topology to LAN-free backup topology can gain additional benefits from tape pooling. Organizations that are evolving to LAN-free backups from a LAN backup topology are already conceptually performing tape pooling functions. To implement LAN-free backups without tape pooling would potentially require an increase in the number of tape devices. The remainder of this section assumes that backup software with tape pooling capabilities is part of the solution.

Advantages

An SAN-enabled backup infrastructure introduces these advantages to the backup process:

- Fibre Channel performance, reliability, and distance
- Fewer processes and less overhead
- No need to use the LAN to move backup data
- Elimination or reduction of dedicated media servers
- Improved backup and restore performance

Fibre Channel performance, reliability, and distance

Fiber is designed for data movement and storage functions. This makes Fibre Channel an ideal channel for moving backup data.

The performance capabilities of Fibre Channel allow the backup application to move the backup data at the speeds required for modern backup windows. Switched fabric provides the capability to connect multiple backup clients to the tape libraries, and is key to providing tape pooling. The distance capabilities of Fibre Channel allow the solution to maintain a centralized architecture.

SAN-enabled backups provide centralization and tape pooling while operating at direct-attached backup speeds.

Fewer processes and less overhead

Figure 121 shows the process flow for SAN backups. Two steps are required to copy the data to the backup media. (LAN backups require four steps.) The reduction in data movement steps reduces CPU and memory resource requirements for backup. Since restore processing follows the same steps (except that the data flows in the opposite direction), restores will perform faster with fewer resources as well.

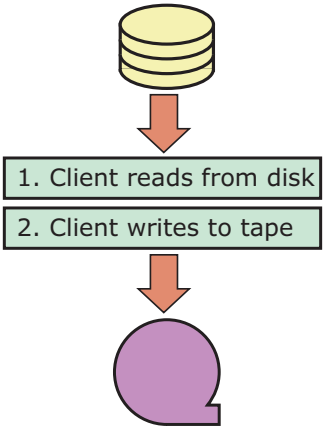


Figure 121 **Process flow for SAN backups**

No need to use the LAN

By removing the network bottleneck, the SAN allows the tape libraries to operate to full performance. Elimination of network traffic will also free CPU and memory, since the data does not need formatting for network transfer and networking protocols do not have to be managed by the system for backup traffic.

Additionally, overall system performance improves because of the reduction of backup traffic on the LAN.

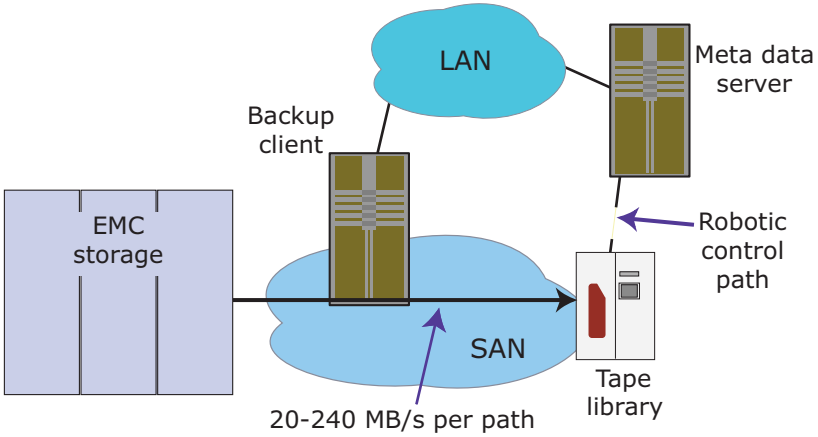


Figure 122 **Bandwidth used in LAN-free backups**

Elimination or reduction of dedicated media servers

SAN reduces or eliminates dedicated media servers. CPU consumption for backup processes is directly related to the number of data movement steps. Since a LAN-free backup requires two fewer steps to move backup data, there are less CPU resources required. This CPU savings for LAN-free backups is on the media server. Organizations evolving from LAN backups to LAN-free might have the option to downsize or eliminate dedicated media servers because of the CPU reductions.

Improved backup and restore performance

Backup and restore performance is improved because of the following factors:

- Number of processing steps is reduced.
- Backup traffic is moved to higher bandwidth and higher reliability fiber connections.
- Network protocol overhead is eliminated.

These factors also improve overall system performance. PowerPath will also help balance and manage disk path contention caused by backup load.

Disadvantages

LAN-free backups impact the host and the application. LAN-free backups also consume host I/O bandwidth, memory, and CPU resources.

Migration paths

Organizations that have exceeded the capabilities of LAN-free backups have several migration paths, including the following, described elsewhere:

- "Serverless backups" on page 571 — An emerging enhancement to LAN-free backups is the serverless backup concept. Serverless backup utilizes the third-party copy capability of the Fibre Channel standard to move the backup data over the SAN straight from disk to tape under the control of the backup application.

Improvement options

The following can also improve backup speed and efficiency:

- Faster devices
- Improved pathing
- Utilization of TimeFinder BCVs

Serverless backups

Serverless backups (Figure 123) use SAN resources to move backup data from disk to tape. These backups are called serverless because the application server does not have to utilize host resources to perform the movement of backup data. Serverless backups are the next evolutionary step in LAN-free backup.

Note: Serverless backups are sometime referred to as server-free backups.

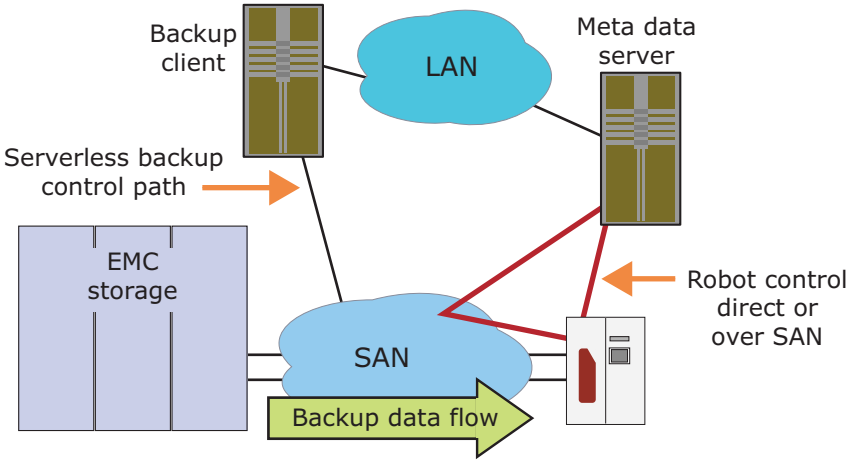


Figure 123 Serverless backup

Theory of operation

Serverless backups use the *third-party copy* function (also called *extended copy*) of the Fibre Channel standard. Third-party copy allows a device to copy data between points on the SAN. A device performing third-party copy can be thought of as a copy engine.

Currently, Cisco is supporting serverless backup in its switches with a SSM module. Cisco refers to this feature as NASB. The copy task is initiated and controlled by a host application that instructs the copy engine to copy a specific number of blocks from the source device to the target device.

NetWorker offers support for Windows and VERITAS offers support for Solaris. The copy engine then copies the data, block by block, until complete. The copy engine performs all of these copies inside the SAN and outside of the host.

With serverless backups, the backup application determines which blocks on disk require a backup. The backup application then instructs the copy engine to copy these disk blocks to the tape device.

Backup process overview

A typical backup process follows these steps:

Note: The order of steps 2, 3, and 5 may vary depending on the backup tool that is utilized.

1. The metadata server invokes backup client processes on the backup client.
2. If tape pooling is used, the metadata server assigns tape devices to the backup client.
3. The metadata server instructs the tape robot to load tapes into the assigned drives.
4. The metadata server notifies the backup client which tape drives it may use.
5. The backup client determines which files require backup.
6. The backup client instructs the copy engine to copy data, block by block, from the storage directly through the SAN to the tape drive.
7. When the backup is complete, the backup client sends metadata information to the metadata server.
8. The metadata server stores the meta data on disk.

Advantages

Serverless backups offer the following advantages over server-based backups:

- Fibre Channel performance, reliability, and distance — Fiber is designed for data movement and storage functions. This makes Fibre Channel an ideal channel for moving backup data.
- No use of host resources to move backup data — Since the backup data is moved outboard from the host CPU, memory and I/O bandwidth resources are not consumed. This substantially reduces the impact on the application by removing resource contention points. Large organizations can realize a significant CPU savings during the backup window.
- Reduced application impact.

- No use of the LAN to move backup data — By removing the network bottleneck, the SAN allows tape libraries to operate to full performance. The elimination of network traffic will also free CPU and memory, since the data does not have to format for network transfer, and networking protocols do not require management by the system for backup traffic.
- Improved backup and restored performance — Backup performance is improved primarily because of reduced host-contention points. This serverless architecture is also designed to keep the tape devices streaming so that they operate at maximum performance.

  Performance is improved because:

  - Processing steps are reduced.
  - Backup traffic is moved to higher bandwidth and higher-reliability fiber connections.
  - Network protocol overhead is eliminated.

  These factors also improve overall system performance. PowerPath also helps balance and manage disk path contention caused by backup load.

Disadvantages

Since serverless backups move data over the production SAN, there is a potential to cause an indirect impact on applications if the SAN has heavy I/O activity. Some implementations of serverless backup operate at the logical disk level, as opposed to the file level.

Backup over long distances using FCIP and routers

Using FCIP and routers allows backups over long distances within a SAN environment. It also allows longer distance vaulting for security, remote copy, real estate, etc. Tape acceleration techniques help ensure acceptable performance.

Figure 124 shows a typical SAN on both the right side and on the left. The routers link the two SANs together using FCIP. The SANs behave as if they are a SAN ISL, making the two SANs appear as one large SAN, thus extending the distance over which backups can occur.

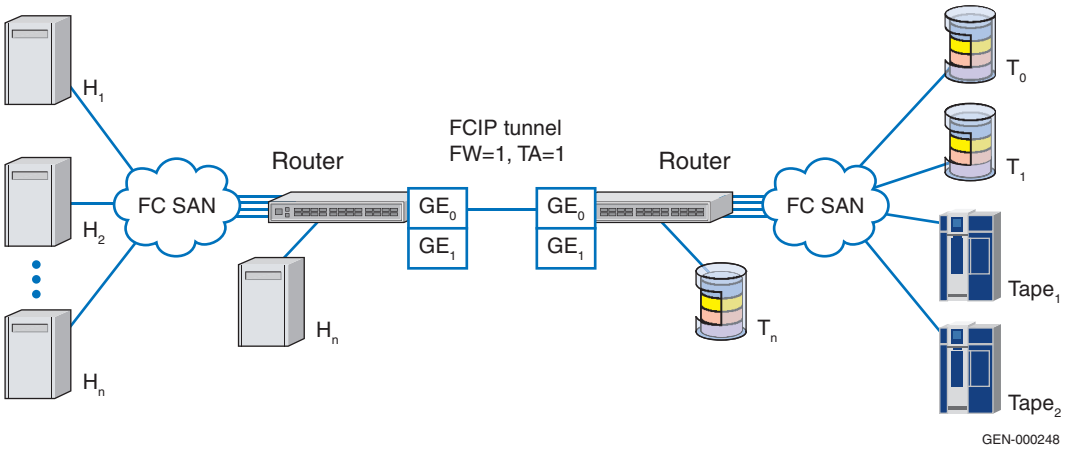


Figure 124 **Backup over long distances**

Theory of operation

Long distance backup utilizes SAN (storage area network) technology in conjunction FCIP using traditional backup software. The high-speed and extended distance capabilities of Fibre Channel and FCIP are used for the backup data movement path. Metadata is still moved over the LAN to the backup metadata server. This traffic is typically light and insignificant in relation to the large amounts of data moved during a backup. The metadata server is also the control point for the robotic mechanism of the tape library.

Advantages

An SAN-enabled backup infrastructure introduces these advantages to the backup process:

- Fibre Channel performance, reliability, and distance.
- Fewer processes and less overhead.
- Use of inexpensive dedicated LAN to move backup data to more remote locations.
- Elimination or reduction of dedicated media servers.
- Improved backup and restore performance.

Disadvantages

This is a new technology.

FCIP tape acceleration with EMC Connectrix MDS switches

Tape devices store and retrieve data sequentially. Normally, accesses to tape drives have only one outstanding SCSI write operation at any given time. This single command nature of tape writes impacts backup and archive performance because each SCSI write operation does not complete until the host receives a good status response from the tape drive.

The FCIP tape acceleration, feature introduced in MDS SAN-OS Release 2.0(1b), solves this problem. It improves tape backup and archive operations by allowing faster data streaming from the host to the tape over the WAN link.

With tape acceleration, the backup server issues write operations to a remote tape drive.

Figure 125 illustrates FCIP link tape acceleration.

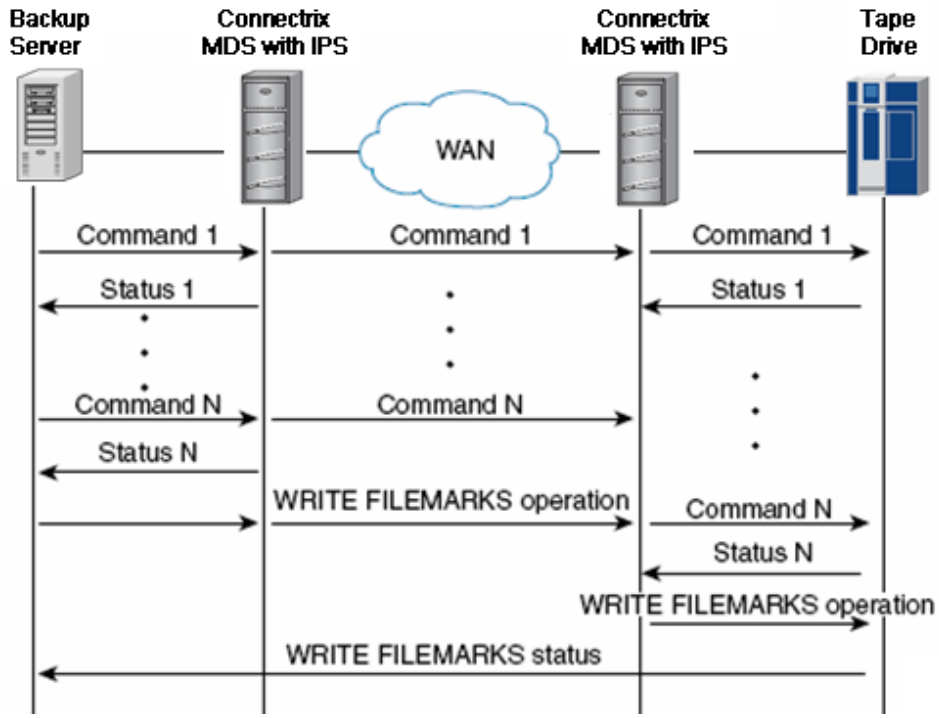


Figure 125 FCIP tape acceleration

The local Connectrix MDS switch acts as a proxy for the remote tape drive by quickly returning a Transfer Ready (Write Accelerator) signal to the host. This enables the host to more quickly begin sending the data.

After receiving all the data, the local Connectrix MDS switch responds to signal the successful completion of the SCSI write operation (Tape Accelerator). This response allows the host to start the next SCSI write operation.

This proxy method of operation results in more data being sent over the FCIP tunnel in the same time period compared to same operation with no proxying. This proxying method also improves the utilization of WAN links.

At the other end of the FCIP tunnel, another Connectrix MDS switch buffers the command and data it has received. The remote Connectrix MDS switch then acts as a backup server to the tape drive by listening to a Transfer Ready from the tape drive before forwarding the data.

The Connectrix MDS SAN-OS provides reliable data delivery to the remote tape drives using TCP/IP over the WAN. Write data integrity is maintained by allowing the Write Filemarks operation to complete end-to-end without proxying. The Write Filemarks operation signals the synchronization of the buffer data with the tape library data. While tape media errors are returned to backup servers for error handling, tape busy errors are retried automatically by the Connectrix MDS SAN-OS software.

For more information, refer to the *MDS 9000 Family Fabric Manager Configuration Guide*, available on http://www.Powerlink.EMC.com.

Notes

Note the following:

- The tape acceleration feature is disabled by default and must be enabled on both sides of the FCIP link. If it is only enabled on one side of the FCIP tunnel, the tunnel is not initialized.
- FCIP tape acceleration does not work if the FCIP port is part of a Port Channel or if there are multiple paths with equal weight between the initiator and the target port. Such a configuration might cause either SCSI discovery failure or broken write or read operations.
- When tape acceleration is enabled in an FCIP interface, a FICON VSAN cannot be enabled in that interface. Likewise, if a FCIP interface is up in a FICON VSAN, write acceleration cannot be enabled on that interface.

- Enabling the tape acceleration feature automatically enables the write acceleration feature.
- Enabling tape acceleration for an FCIP tunnel reinitializes the tunnel.
- The flow control buffer size specifies the maximum amount of write data that a Connectrix MDS switch buffers for an FCIP tunnel before it stops the tape acceleration proxying process. The default buffer size is 256 KB and the maximum buffer size is 32 MB.

Enabling FCIP tape acceleration on Connectrix MDS switches

To enable FCIP tape acceleration using Fabric Manager:

1. From Fabric Manager, choose **ISLs > FCIP** from the **Physical Attributes** pane.

   The FCIP profiles and links display in the **Information** pane.

2. From **Device Manager**, choose **IP > FCIP**.

   The **FCIP** dialog box displays.

3. Click the **Tunnels** tab.

   The FCIP link information displays.

4. Click the **Create Row** icon in Fabric Manager or the **Create** button in Device Manager.

   The **FCIP Tunnels** dialog box displays.

5. Set the profile ID in the **ProfileID** field and the tunnel ID in the **TunnelID** field.
6. Set the **RemoteIPAddress** and **RemoteTCPPort** fields for the peer IP address you are configuring.
7. Check the **Write Accelerator and TapeAccelerator** checkbox.
8. Optionally, set the other fields in this dialog box and click Create to create this FCIP link.

Brocade 7500 and CDL over IP with Fastwrite acceleration and tape pipelining

To optimize performance of backups over FCIP links, consider using Fastwrite and tape pipelining.

Note: These features are supported only in Fabric OS 5.2.x and higher.

Fastwrite and tape pipelining provide accelerated speeds to FCIP tunnels in some configurations:

- Fastwrite accelerates the SCSI write I/Os over FCIP.
- Tape pipelining accelerates SCSI write I/Os to sequential devices (such as tape drives) over FCIP, reducing the number of roundtrip times needed to complete the I/O over the IP network and accelerating the process.

  Note: You *must* enable Fastwrite in order to use tape pipelining,.

- Both sides of an FCIP tunnel *must* have matching configurations for these features to work.

Fastwrite, and tape pipelining features do not require any predefined configurations. This makes it possible to enable these features by adding optional parameters such as **–c**, **-f**, or **-t** when you create FCIP tunnels.

Refer to the *Fabric OS Administrators Guide* for further information.

Case study notes

In Figure 126 on page 580, a configuration was built with Brocade 7500 and IP Network emulating up to 40,000 kilometers distance (EMC CDL DL710 local devices and remote and Legato Networker). Fastwrite and tape pipelining were turned on for the remote configuration.

In Step 1, the backups are written to the local EMC CDL DL710 device.

In Step 2, the backup is cloned to the remote location which consists of a remote CDL DL710 connected to remote Brocade 7500. Fastwrite

and tape pipelining are turned on in both the local and remote Brocade 7500.

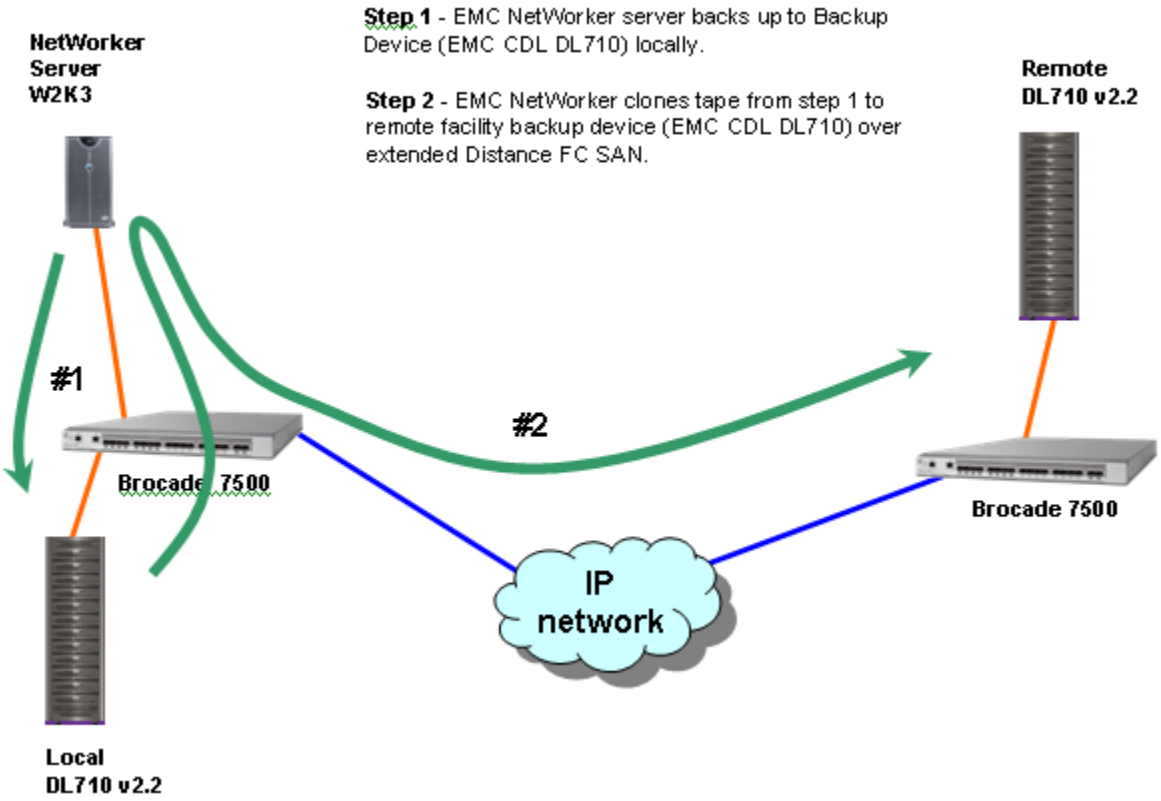


Figure 126 Brocade tape acceleration — Host remote backup

In Figure 127 on page 581, a configuration was built using EMC Celerra Brocade 7500, and IP Network emulating up to 40,000 kilometers distance, EMC CDL DL710 devices locally and remote and Legato Networker. Fastwrite and tape pipelining were turned on for the remote configuration.

In Step 1, the backups are written to the local EMC CDL DL710 device by Celerra under Networker NDMP control.

In Step 2, the backup is cloned by Celerra under NDMP control to the remote location which consists of a remote CDL DL710 connected to remote Brocade 7500. Fastwrite and tape pipelining are turned on in both the local and remote Brocade 7500s.

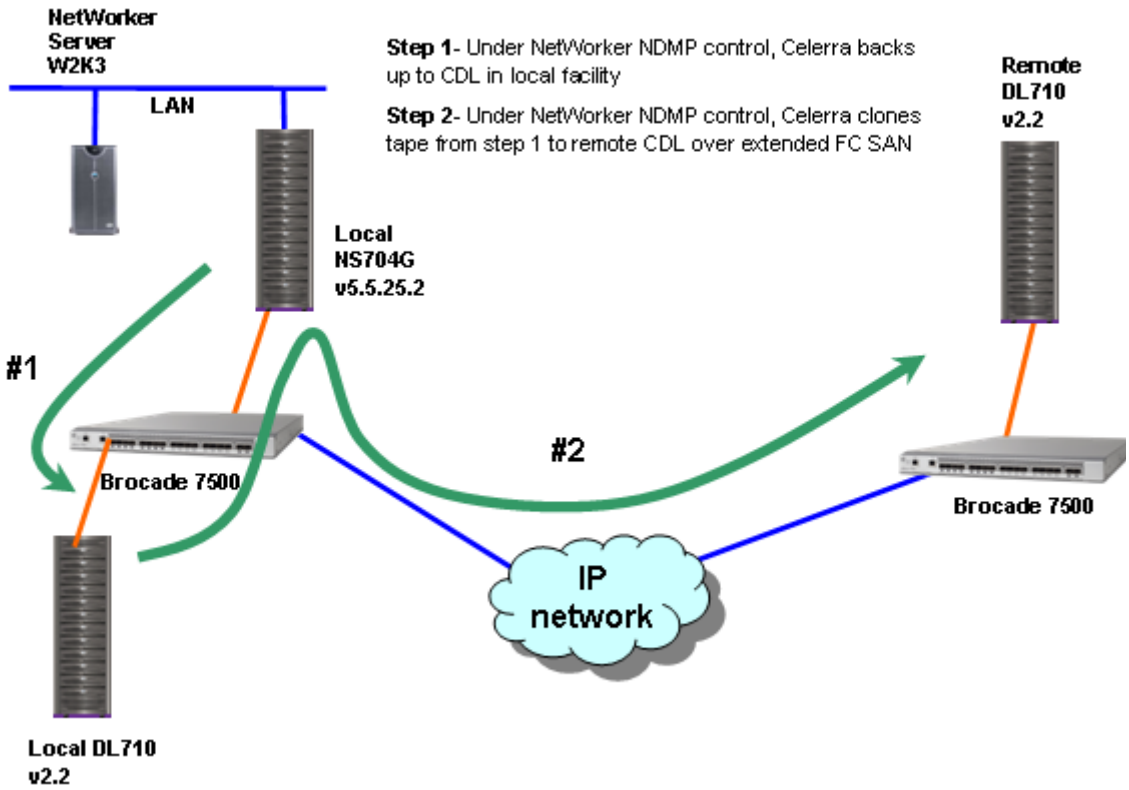


Figure 127 **Brocade tape a cceleration — NDMP remote backup**

Results Without tape pipelining performance could drop as much as 75%. When tape pipelining is enabled, performance returns to about 80% of a direct FC connection.

NAS backups

The Celerra® Network Attached Storage (NAS) device provides supports multiple backup options:

- Local and remote backup
- Network backups
- Network Data Management Protocol (NDMP) backups

Local and remote backup

Local backup uses tape devices that are directly connected to the Celerra.

Remote backup transfers the backup data to another server that contains a tape device. The backup data may be transmitted over the network or it may be transferred through direct server-to-server connection.

NDMP backup

The Network Data Management Protocol (NDMP) is an open standard communication protocol that is specifically designed for backup of network attached storage devices. NDMP enables centralized backup management and minimizes network traffic.

A key feature of NDMP is that it separates the flow of data from the management of the data. This allows third-party backup tools to interface with Celerra and maintain centralized control. NDMP backups can be performed locally (NDMP V1) with tape devices connected directly to the Celerra or remotely (NDMP V2) to another location. Both of these options are managed by the third-party backup tool.

For more information on Celerra backups can be found in various Celerra documents on Powerlink.

Backup to disk

Traditional tape-based backup systems can be slowed by the mechanical nature of tape.

Consider the following:

Tape Speed = Tape Transfer Rate + Mechanical Motion

where mechanical motion consists of the following factors:

- Robot mount / dismount time
- Load-to-ready times, find the start of the tape
- Rewind times

Disk arrays can be faster than tape for both backups and restores. Advantages include:

- Potentially faster transfer speeds
- No mount, rewind, load-to-ready issues
- Not vulnerable to streaming issues

A potential disadvantage of backup to disk is that it may require changes to existing backup processes and operations.

For more information concerning backup to disk, please review the following white papers available on EMC.com:

- *Backup-to-Disk: Building an Effective Long-Term Strategy Technology Concepts and Business Considerations*
- *Backup-to-Disk Planning Options Applied Technology*
- *Symmetrix DMX-3 Best Practices Planning*

Backup media: Tape libraries

A tape library consists of tape storage facilities, a robotic tape selection mechanism (sometimes referred to as a picker), and tape devices.

Tape libraries connect to the SAN with two types of connections: the data path and the robotic control path. The data path is the mechanism that moves data between the server and the tape device(s). The robotic control path controls the robotic mechanisms.

The robotic control (picker) path connects to the SAN with a embedded bridge or blade located in the tape library. Depending on the tape library and customer preferences, the drive data path can connect through the blade, or directly to the SAN. Some newer libraries do not use a blade for the picker. In this case, one drive is configured to be the control LUN-based drive, and the picked control path is a LUN under the drive. There are no significant technical advantages or disadvantages to either connection.

The key benefit from a SAN backup comes from the ability to connect the tape devices to the SAN.

Editing the Solaris configuration file: st.conf

If tape devices are addressed (through a blade, bridge or control drive) they can use target IDs and LUNs greater than zero. The driver configuration file **`/kernel/drv/st.conf`** must be modified to include the correct definitions for each target ID and LUN. The Solaris driver treats all entries in all target drivers (such as **`sd`** and **`st`**) as one continuous assignment per instance of the HBA. Therefore, be sure that **`st.conf`** does not contain target numbers already specified in **`sd.conf`**. Different HBAs behave in slightly different ways, so consult your HBA vendor to see if these changes are needed.

Solaris provides host application access to the tape through:

`/dev/rmt/0, /dev/rmt/1,...x`

Therefore, `/drv/rmt/<x>` addresses must be used for the tape access.

Note: The numbering in `/dev/rmt` should be sequential, starting with 0.

Note: Changes to st.conf do not take effect until the host is rebooted with the reconfigure option (`reboot -- -r`).

CAUTION

Deleting internal drive entries from `/kernel/drv/st.conf` makes your host unbootable.

HBAs

Persistent binding is recommended. Persistent binding keeps the LUNs and tape numbers consistent since there is a slight chance that they could change during a system event. Consult your HBA vendor for details.

Windows

Requires no additional settings. You must install a tape driver when working with a tape drive.

Sun Solaris (Fibre Channel tape support)

You may need to change two files:

- **`sd.conf`** — Configuration file used for disks; outlines the specific devices that can be used, and enables binding.
- **`st.conf`** — Configuration file used for tapes; all tape drives must be listed in this file, and binding can be enabled.

Tape drives

Fibre Channel

All new generation tape drives support FC-SW and FC-AL. Some drives may default to loop. EMC recommends setting your switch or drive to FC-SW. Some devices have two ports; EMC supports the use of only one port at a time.

Each Fibre Channel tape drive is configured with a static, unique World Wide Name (WWN) assigned by the manufacturer. You can use this WWN during the fabric zoning process to allow servers access to tape devices. The most popular drive technology is LTO, which is currently at generation three. Generation four is being developed. Drives with 4 GB capability are also available.

Mirrored fabric backup solution

EMC supports many Fibre Channel tape backup solutions. This section describes the solution that best fulfills the following objectives:

- Ease of management:
  - No new zone sets are required; only new tape zones.
  - Traffic patterns / routes are direct and deterministic.
- Supportability:
  - No new hardware or configuration is required.
  - Affected resources are easily identified because of consistent placement and routing.
- Flexibility:
  - PowerPath allows non-simultaneous use of the same HBAs for both disk and tape.
  - Core / edge fabric topologies allow traffic segregation and localization, so bandwidth requirements can be managed separately.
- Scalability — Provides the scalability benefits of the mirrored, core / edge fabric design.
- Availability:
  - Core / edge fabrics provide multiple paths from servers to tapes.
  - Mirrored fabrics protect against HBA failures, switch failures, tape failures, fabric failures, and maintenance activities.
- Maintainability — Since all servers have access to tapes on both fabrics, either fabric can be placed in a maintenance state without catastrophically interrupting the backup cycle.
- Maximum return on capital investments:
  - PowerPath allows hosts to provide additional disk access when backups are not being performed. (Backup customization may be required.)
  - Better utilization of bandwidth, ISLs, and switch ports.

"Solution description" on page 587 describes the solution in detail, and illustrates how each aspect of the backup solution adheres to the project requirements.

Solution description

Since tape drives are supported as single attached devices only, our solution must provide a means to protect our backup resources in the event of a failure in the environment. To accomplish this, EMC recommends that tape drive resources be distributed evenly across both sides of the proposed mirrored fabric environment.

Understanding the importance of backup resources in your business continuance plan, EMC strives to provide the highest level of protection possible, while maximizing the duty cycle of all resources. Each media server or application server (LAN-free backup) in this facility would be configured so that it could access disk and tape storage on both fabrics. This can ensure that any service interruptions or failures on one fabric do not affect the storage (disk / tape) resources attached to the other fabric.

Providing resources that are isolated from the impact of maintenance cycles or failures helps ensure that backup resources are always available when you need them.

The backup software solution should be evaluated for the best way to create media pools of devices that can include tape drives from both fabrics. PowerPath scripting should also be employed to set specific HBAs that share tape and disk into standby mode during backup cycles. Setting a path to standby allows PowerPath to use the HBA in the case of path failure, but would not use it for data traffic during normal conditions.

This procedure can ensure that the HBA is dedicated to tape access during the time of the backup, alleviating contention for bandwidth resources across this link (which could cause backup performance degradation). This procedure is especially useful for hosts that do not have enough available slots for a dedicated tape (backup) HBA. It is also useful for hosts that are already very write-intensive.

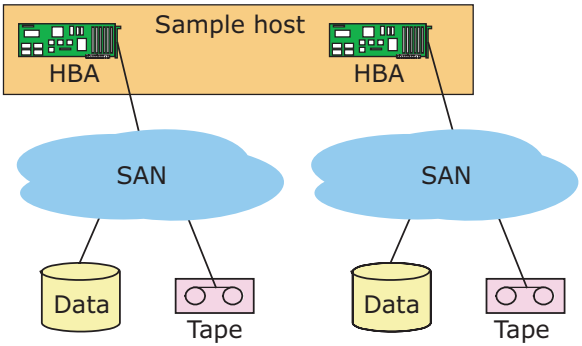


Figure 128 Example of backup software solution

Although tape drives can be placed anywhere in the fabric, EMC recommends placing the tape drives at the core of a core/edge fabric. This provides equal (single-hop) access from each of the LAN-free backup servers located at the edge switches. Centralizing the tape storage also helps to prevent any ISL bandwidth congestion that might occur by providing direct routes to the tape storage. This also aids in balancing the backup load across the entire fabric without complex backup scheduling. Backup, restore, and recovery requirements can be maintained simply by adding the required core/edge ISLs needed to handle the peak application and backup bandwidth load.

Dedicated media servers and third-party copy engines (server-free backup) should also be placed at the core of the fabric. Since these servers are dedicated to tape access, placing them at the core ensures that their data will traverse only the internal switch backplane or the core-to-core ISLs. This provides better use of the switch's internal resources, helping to minimize the required core/edge ISLs.

Utilizing the core-to-core ISLs also provides a means to reduce the contention for ISL resources by segregating the backup traffic from the application traffic. Maintaining the backup objectives for these dedicated media servers and third-party copy engines is managed by adding the proper number of core-to-core ISLs. Because the number of dedicated media servers and third-party copy engines is usually much lower than the number of application servers (backup clients), the fan-out benefits of the core/edge fabric design is not adversely affected. Providing a level of traffic segregation also provides better

control over our application server SLAs by minimizing the impact of backup traffic on our application servers and their data routes.

Figure 129 is a sample of a core/edge fabric with the recommended application server, tape drive, disk storage, and media server placement.

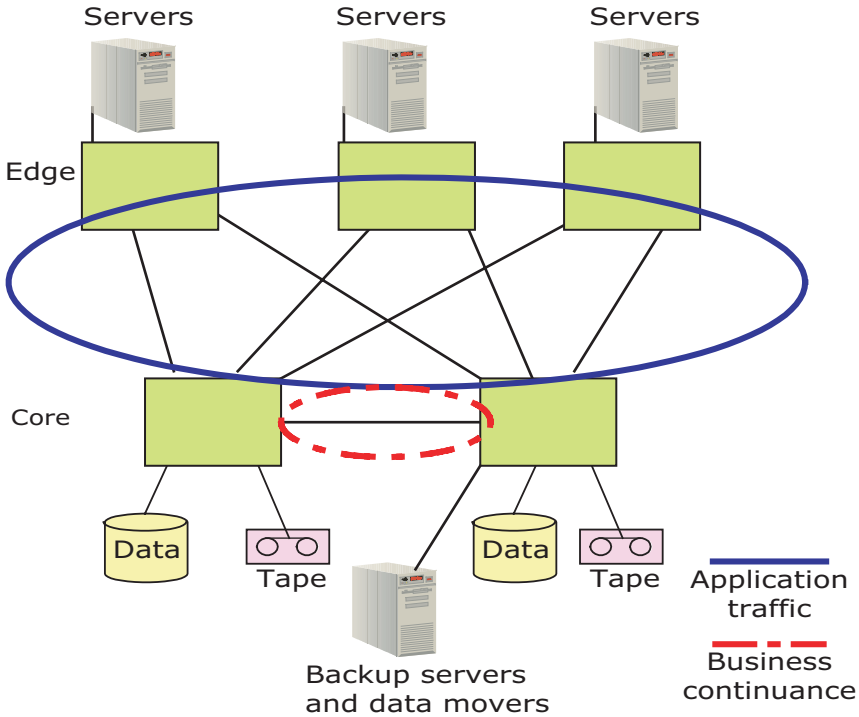


Figure 129 **Core/edge fabric example with recommended component placement**

Physical backup device centralization

Creating a centralized backup location that contains both the trained backup administrators and backup media resources can increase the duty cycles of your backup media, and increase the efficiency of your personnel resources.

Centrally locating the personnel provides the following benefits:

- Fosters communication between group members.
- Provides better access and transfer of skills among the members.
- Eliminates the necessity for operators to move between locations for tape processing and troubleshooting.

Summary

While EMC supports many different fabric environments, core/edge mirrored fabric environments offer an efficient, robust design that can fulfill both your application and business continuance requirements.

Tapes and fabrics

This section discusses the following:

- "SCSI tape", next
- "Fibre Channel tape" on page 591
- "Sharing tape and disk on the same HBA" on page 593
- "Tape library" on page 595

SCSI tape

SCSI-attached tape storage devices have been, and continue to be, a viable solution in the data storage environment. Whether these devices were legacy hardware, new purchases or advanced SCSI technology, they will also have to communicate with the Fibre Channel portion of the SAN environment. In order to achieve this, they have to be attached directly to SCSI-to-Fibre Channel bridges and through them to the SAN.

Fibre Channel tape

Most native Fibre Channel tape drives available today are FC-SW capable, but there are some early first-generation native FC tape drives that only support Fibre Channel Arbitrated Loop (FC-AL). Consult your tape drive vendor to obtain this information. Public loop devices are arbitrated loop devices that support Fibre Channel fabric login and services. Each Fibre Channel tape drive has one or more NL_Ports (Node Loop Ports), which can be used to connect into either a loop device or fabric device capable of communication with NL_Ports or a loop-to-fabric bridge.

Each Fibre Channel tape drive is configured with a static, unique, World Wide Name (WWN) assigned by the manufacturer. This WWN can be used during the fabric zoning process to allow servers access to the tape devices.

Supported Fibre Channel tape drives

E-Lab Navigator lists the currently supported FC-AL and FC-SW tape drives.

Connecting Fibre Channel tape drives into a fabric

Some fabric switches, routers, and bridges support both the FC-AL protocol and the Fibre Channel Point-to-Point protocol. Point-to-Point is required for SAN appliances that support switched fabric communication. A physical port on such an appliance is referred to as an FL_Port (Fabric Loop Port). FL_Ports may either automatically negotiate the method of communication with a connected N_Port or NL_Port device on initialization, or the user may be required to manually set the port type in the device's configuration file prior to connecting the devices together. In order to connect with SAN devices that do not support FC-AL (Connectrix ED-1032, for example) a loop-to-switch bridge can be used.

Configuration details

The supported switches all provide the same functionality when configuring Fibre Channel tape drives. Prior to starting the configuration, make sure that the Fibre Channel ports on the tape drives are enabled and that the tape drives are on line:

- Tape drive connection to switch — Each port on these switches is capable of auto-negotiating and auto-initializing the port for loop or switch communication. Auto-configuration occurs immediately after the tape devices are connected to the switch port. At the completion of the auto-configuration phase (almost instantaneously), the port should appear as an FL_Port in the name server list on the switch. No additional software or licensing is required for this functionality.

  Note: Do not configure these ports as QuickLoop or Fabric Assist ports. The QuickLoop or Fabric Assist mechanisms are not required for public loop device support and are not supported by EMC.

  When the negotiation and initialization is complete, you will also be able to view the tape drive's WWPN in the switch's name server list.

- Server connection to switch — Each server connected to the switch that will communicate with the FC-AL tape drive should be configured as a Fibre Channel switch fabric device. When the server is connected to the switch it will automatically run through the port negotiation and initialization procedures, but the resultant port configuration will appear as a fabric-capable N_Port in the switch's name server list. When the negotiation and initialization is complete, you should be able to view the server's HBA WWPN in the switch's name server list.

Note: E-Lab Navigator identifies the latest drivers and firmware associated with these HBAs.

- Zoning — The switch's Web browser configuration tool will allow you to zone the unique WWPN of the tape drive with the WWPN of the server's HBA.

  "Sharing tape and disk on the same HBA" describes the considerations on sharing the same HBA for both disk and tape.

- Server-to-tape communication — Translation of FC-AL protocols associated with the tape drive from / to FC-SW associated with the server's HBA are all handled automatically, internal to the switch. No special settings on the switches are necessary to allow translation and communication.

Sharing tape and disk on the same HBA

While EMC supports simultaneous tape and disk access over the same HBA, some access scenarios impact the feasibility of such a solution. E-Lab Navigator contains specific details about your configuration.

When disk and tape share the same HBA, the HBA must support four I/O activities:

- Production disk reads to support your applications
- Production disk writes to support your applications
- Disk reads to feed the backup operation
- Tape writes to support the backup operation

In heavy I/O environments, these four activities combined can result in I/O path and device contention.

I/O path contention occurs when more than one application is trying to access the same I/O path at the same time. Both operations must arbitrate for control of the path, and one must wait while the other is using the path. All technologies also have their own path bandwidth limits, which must be shared between the operations.

Device contention occurs when multiple operations are trying to access the same information at the same time. Again, while one operation is accessing a resource, other resources must wait their turns.

Path and device contention can result in both reduced application performance and reduced tape backup performance.

Tape devices are sequential media that expect data to be sent in large continuous blocks. This continuous flow of data (called *streaming*) keeps the tape media progressing at its fastest speed. If streaming is interrupted due to congestion, the drive must stop and back up to the end of the last block transferred before it can accept new data. Stopping and repositioning takes time because of the effort required to gradually slow down the tape drive in an orderly fashion. (Rapid deceleration could damage the media; by stretching the tape, for example.)

Backing up to the last block endpoint and stopping again also involves a change of direction, acceleration and deceleration, a stop, a re-tension, and a position check. Some tape devices (depending on the drive) may also require that the tape be kept at a specific tension level. In this situation, if a data transfer stops, the drive must stop, and back up to the end, after which the tape will move forward and backwards slowly until the data transfer starts again. This method (often called *shoe shining*) is used to maintain the proper tension on the tape.

Another way that congestion can interfere with your backup is its effect on data compression. Most modern tape devices also have built-in hardware compression. The data is written to a buffer in the tape device, compressed, and then written to the physical media. Part of the compression mechanism involves using standard algorithms; however, part of the compression mechanism also combines the contents of the tape buffer into mega blocks to reduce inter-record gaps on the tape device.

Heavy congestion can also cause tape compression rates to drop, as the drives are unable to use completely full buffers when they create the mega blocks. Since the data is not optimally compressed, more write operations are required to get the data to the tape media. More write operations results in degradation of backup performance. This could also result in more tape cartridges being required and subsequent elongated recovery times because there are now more tape cartridges to recover.

If your backup environment is in a congestion situation, the congestion can be partially addressed by placing disk and tape on separate HBAs. The potential benefits of separate tape and disk HBAs (in congestion situations) include reduced production disk path contention and improved tape performance. If the number of

HBAs on the server is limited, you can also employ PowerPath during the backups to manage the load across the adapters. PowerPath provides the ability to set an HBA into *standby* mode, which allows the HBA to be used if there is a failure, but not for disk traffic (while in standby mode). When the backups were complete, you could set the HBA back into an *active* state. Once the HBA was in an active state, PowerPath would rebalance the disk traffic, using this HBA in its load-balancing calculations.

Limitations

Moving the Fibre Channel cables associated with the tape or with the server communicating with the tape to a different switch port while I/O is running will result in I/O failures and device lockout. If this occurs, you may be required to either power cycle the tape device or return the Fibre Channel cables to their original ports, and manually release the tape device to return the system to working order. It will then be necessary to restart the backup job.

EMC supports only single-port usage of the STK 9840. Simultaneous use of both ports can result in contention for devices on both server boot and setting of device reservation.

Tape library

A tape library consists of tape storage facilities, a robotic tape selection mechanism (sometimes referred to as a picker), and tape devices. The robotic control (picker) path may or may not be connected into the SAN depending on the tape library and customer preferences. SCSI-based control paths may be connected through a bridge.

6

Clusters

The case studies presented here are informational only, and should not be presumed to be firm recommendations for specific environments other than the ones described.

These studies are provided as technology demonstrations; specific implementations should be tailored to customer requirements.

These case studies are distributed by EMC for information purposes only. EMC does not warrant that this information is free from errors. No contract is implied or allowed.

- Requirements for Oracle 10g R1 RAC with ASM on AIX 5L 598
- Oracle 10g release 2 (10.2) RAC installation notes 604
- SLES9 Oracle 10g with EMC arrays and EMC PowerPath 607

Requirements for Oracle 10*g* R1 RAC with ASM on AIX 5L

This section provides information on requirements for Oracle 10*g* R1 RAC on EMC Symmetrix or CLARiiON storage systems. The section describes a reference configuration using EMC storage systems with Oracle 10*g* R1 RAC and Oracle ASM.

CAUTION

Oracle10*g* RAC is a complex operating environment. Even with the information provided in this case study, only EMC-qualified personnel should attempt to configure this environment.

Related documentation

Refer to the following for more information:

- Oracle 10g RAC Release Notes, available at:

  `www.Oracle.com/support/metalink/index.html`

- Latest AIX 5L OS patches, available at IBM Fix Central.
- The following documentation available at Oracle.com:
  - *Oracle Real Application Clusters Installation and Configuration 10g Release 1 (10.1) Guide*, P / N B10766-08
  - *Oracle Real Application Clusters Administrator's Guide 10g Release 1 (10.1)*, P / N B10765-01
  - *Oracle Database Release Notes 10g Release 1 (10.1.0.2.0) for AIX-Based Systems*, P / N B13611-05
  - For information on Oracle ASM: *Oracle Database10g Automatic Storage Management Technical Best Practices*, Nitin Vengurlekar, Oracle Corporation, January 2004
- *PowerPath for AIX Version 4.3 Installation and Administration Guide*, P / N 300-001-683, available on Powerlink.
- E-Lab Navigator, for supported host server and HBA combinations

Supported hardware and software

Hardware requirements and considerations for RAC configurations: EMC Symmetrix DMX 800 / 1000 / 2000 or CLARiiON CX series

Note: E-Lab Navigator lists the supported services and switches. Use the HBA driver and firmware (or fcode) revisions listed on E-Lab Navigator at install time.

Software requirements and considerations:

- EMC PowerPath 4.3.0
- AIX 5L operating system:
  - Version 5.2 with maintenance Level 4 or later, *or*
  - Version 5.3 with maintenance Level 1 or later

  Note: Refer to note 282036.1 in the Oracle metalink.

- Oracle10*g* RAC version 10.1.0.2 plus patch 3761843, which contains Oracle 10*g* RAC 10.1.0.3

  Note: Oracle 10*g* RAC version 10.1.0.3 is the minimum version supported.

Setup and configuration summary

Perform the following:

1. Install the EMC CLARiiON / Symmetrix system and IBM AIX 5L hosts into the SAN fabric.
2. Install the Host AIX 5L operating system and all patches per Oracle 10*g* documentation for AIX. Refer to the Oracle metalink notes for manadatory patches.
3. Install and configure EMC PowerPath 4.3.0.
4. Install and Configure Oracle 10*g* CRS.
5. Install Oracle 10*g* Database 10.1.0.2.
6. Apply the Oracle patch for 10.1.0.3 to Oracle 10*g* CRS and Database.
7. Apply the Oracle patch for ASM.
8. Apply the Oracle patch for private heartbeat cable failure.
9. Create and configure the Oracle 10*g* database with ASM disks.

General installation guidelines

Array guidelines

Install and configure the EMC storage array for operation in the IBM AIX operating system environment.

For a CLARiiON array, FLARE version 16 or later is recommended.

E-Lab Navigator and the *EMC Host Connectivity Guide for IBM AIX* (available on Powerlink) contain more information.

Network guidelines

Refer to the *Oracle 10g RAC Installation and Configuration Guide*, Chapter 2, "Pre-Installation Tasks for Installing RAC on AIX-Based Systems."

Host guidelines

The following general guidelines describe an installation process from the beginning. The *Oracle 10g RAC Installation and Configuration Guide* contains additional details:

1. On hosts:
   - AIX 5.2 (with maintenance Level 4 or higher) — Install patch *IY62117: I/O Hang after 1 second cable pull and recovery.*
   - AIX 5.3 (with maintenance Level 1 or higher) — Refer to the Oracle metalink note on OS patches pre-requistes.
2. Install the following odm filesets:
   - CLARiiON — EMC.CLARiiON.fcp.rte 5.2.0.1
   - Symmetrix — EMC.Symmetrix.aix.rte 5.2.0.1
3. Install EMC PowerPath 4.3.0:
   - With PowerPath installed, hdiskpowerX devices must be used for Oracle 10*g* RAC CRS and Database disk devices.
   - For all 10*g* disks (Oracle 10*g* CRS and Database disks), refer to the *Oracle Database Release Notes 10g Release 1 (10.1.0.2.0)* for AIX-Based Systems.
   - Run **lsattr -E -l hdiskpowerX** to verify that **reserve_lock=no**.
   - Run **chdev** with the following syntax for every Oracle 10*g* CRS and Database disk on every node in the cluster to change it:

     ```
     chdev -l hdiskpowerX -a reserve_lock=no
     ```
4. Install the 10g 10.1.0.2 CRS.
5. Do CRS setup configuration. Make sure CRS is running. (Refer to the *Oracle 10g RAC Installation and Configuration Guide*, Chapter 8, "Installing Cluster Ready Services on UNIX-Based Systems.")
6. Install 10.1.0.2 Database Server. (Refer to the *Oracle 10g RAC Installation and Configuration Guide*, Chapter 10, "Installing Oracle Database 10g with Real Application Clusters.")
7. Run **vipca** and **netca**.

Note: Do not create a Starter Database.

8. Install Oracle 10*g* 10.1.0.3 Patchset patch 3761843.

   Note: Be sure to upgrade both the CRS and DB packages per the patch README.

9. Install Oracle 10*g* patch 2617419: opatch 1.0.0.0.50. (Refer to note 224346.1 in the Oracle metalink.)
10. Install Oracle 10*g* patch 3785451: *asm patch to get emc disk reported quickly and with proper size* (dated 10/1/04).
11. Install Oracle 10*g* patch 4051111: *pvt cable pull causes thread deadlock (BUNDLE FOR BUGS3992673 AND 3942568)*.
12. Install Oracle 10*g* patch 3626461: *CRS HUNG AFTER REBOOTING ON MASTER NODE*.
13. Create ASM Diskgroups and Database Instance. (Refer to the *Oracle 10g RAC Installation and Configuration Guide*, Chapter 11, "Creating RAC Databases with the Database Configuration Assistant.")

    For HACMP/GPFS installations, refer to the Oracle documentation and metalink website.

14. For Oracle 10*g* RAC installations using a CLARiiON array, run the following command to increase the misscount value for the CRS:

    ```
    $Oracle_CRS_HOME/bin/crsctl set css misscount 180
    ```

    Then reboot all the cluster nodes for this new value to take effect.

    Note: Refer to note 294430.1 in the Oracle metalink for misscount setting changes.

Oracle 10g RAC installation and configuration flowchart

Review the *Oracle 10g RAC Installation and Configuration Guide* and the *PowerPath 4.3.0 Installation Guide* before attempting to install.

Figure 130 on page 602 shows the Oracle 10g RAC installation and configuration flowchart.

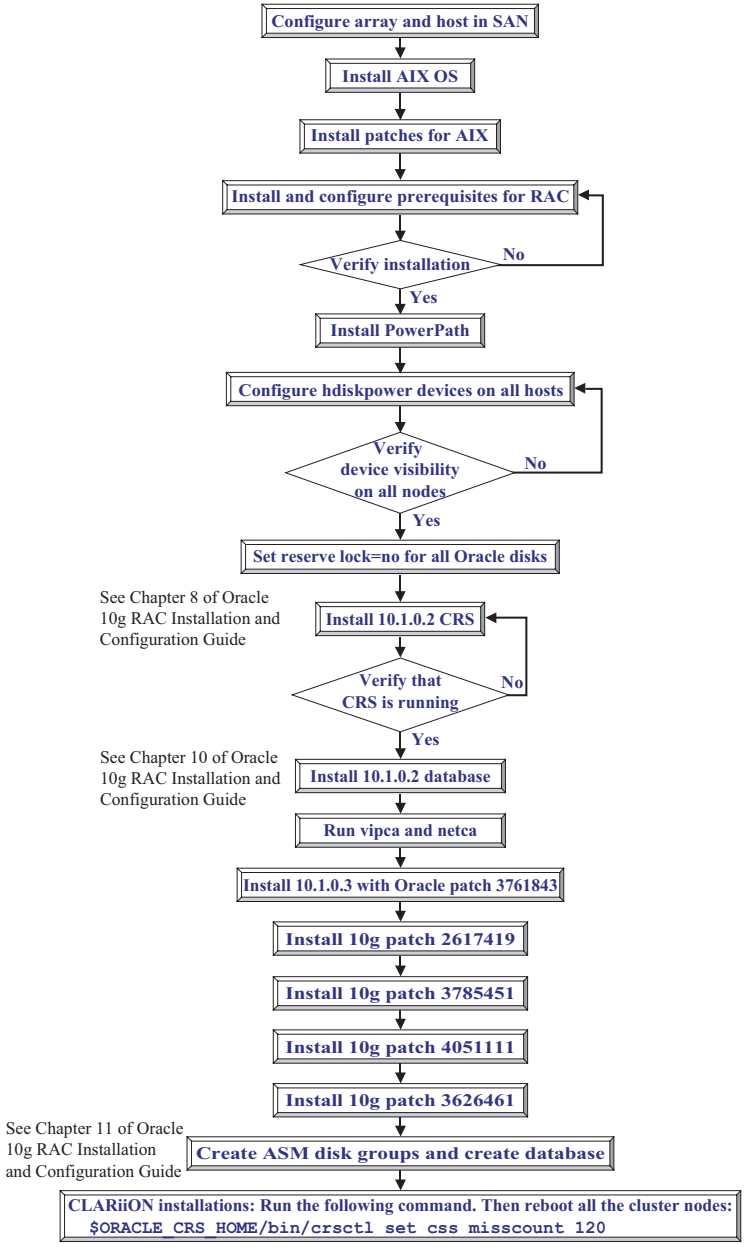


Figure 130 **Oracle 10g RAC installation and configuration flowchart**

Oracle 10*g* RAC supported configuration

Figure 131 shows a possible configuration for an EMC CLARiiON or Symmetrix system in an Oracle 10*g* RAC with ASM environment.

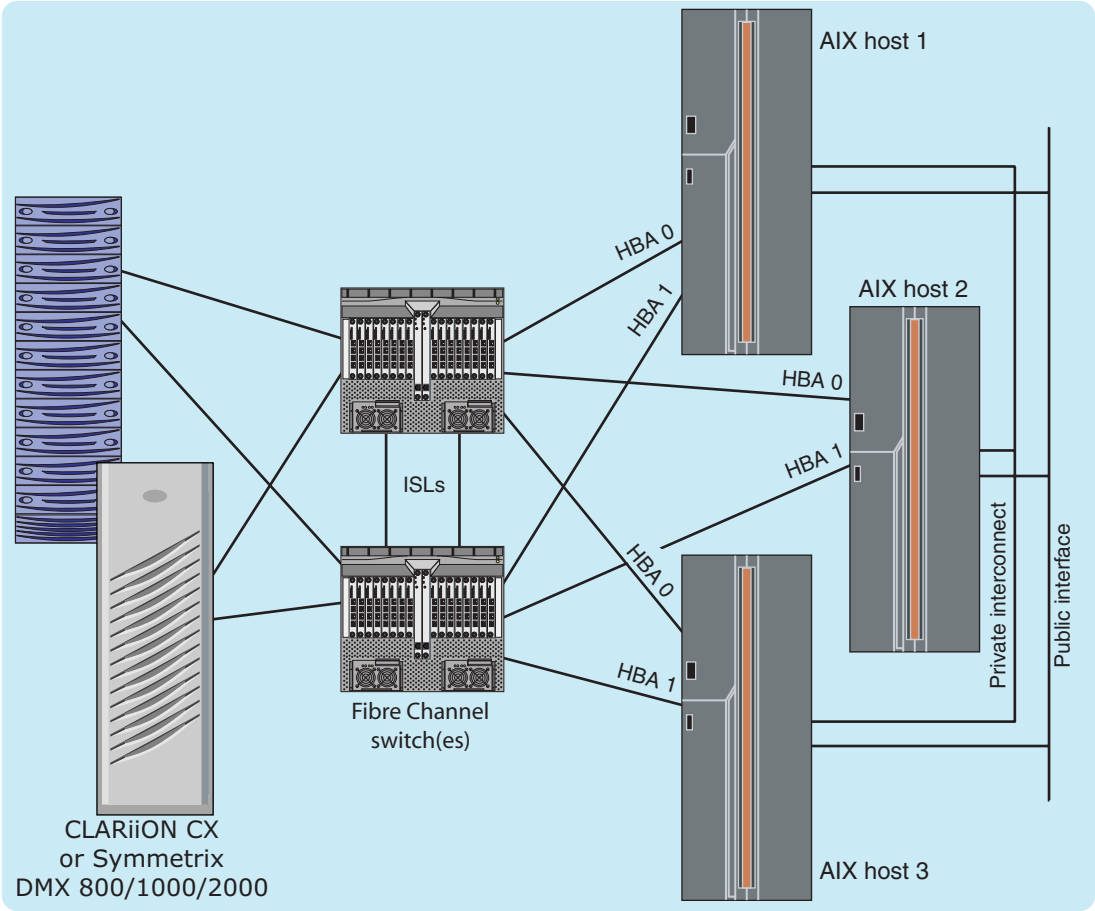


Figure 131 Typical three-node fabric topology with IBM AIX hosts

Oracle 10*g* release 2 (10.2) RAC installation notes

Oracle 10g release 2 (10.2) RAC installation notes for Linux platforms with PowerPath 4.4 connected to EMC arrays

This installation note is to be used in selecting device names for use with Oracle Clusterware and database files.

References

For more information, refer to:

- *PowerPath for Linux Version 4.4 Installation Guide* (EMC P / N 300-002-580), available on Powerlink.
- *Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for Linux* (Oracle P / N B14203-01), June 2005

Installation notes

Use native device names instead of PowerPath pseudo-names to avoid a problem with PowerPath 4.4 that prevents Oracle installation of Clusterware.

To ensure successful installation of Clusterware, stop PowerPath and choose a raw device name bound to a native device name prior to installation.

To ensure successful installation of the database, restart PowerPath and use the raw device name bound to the PowerPath pseudo-name.

Using native names or PowerPath pseudo names will invoke the PowerPath driver and provide all the features of PowerPath, such as path failover and load balancing, after Clusterware and Database are installed.

Oracle RAC installation procedure with EMC arrays and PowerPath 4.4

The installation procedure is as follows:

1. On all nodes in the cluster, install PowerPath 4.4 as described in the *PowerPath for Linux Version 4.4 Installation Guide*.
2. Proceed with Oracle RAC installation as described in the *Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for Linux*.
3. Refer to *Binding the Partitions to Raw Devices* on page 3-10 of the Oracle RAC Installation Guide.

   For the Oracle Cluster Registry and the Oracle Clusterware voting disks, choose native device names (such as **/dev/sdr1**) instead of PowerPath pseudo names (such as **/dev/emcpowera1**).

4. Type **powermt display dev=all** and refer to this output to see the mapping between native device and PowerPath pseudo device.
5. Type **/etc/init.d/Navisphere stop**.
6. Type **/etc/init.d/PowerPath stop**.
7. Refer to *Identify shared storage for Oracle Clusterware files and prepare disk partitions if necessary* on page 4-6 of the Oracle RAC Installation Guide.

   During installation, you are asked to provide paths for files that must be shared across all nodes of the cluster (OCR and Voting devices). Provide the raw device name bound to the native device name chosen for OCR and Voting disks from step 3.

   Example:

   ```
   /dev/raw/raw5 -> OCR device
   /dev/raw/raw7 -> Voting device 1
   /dev/raw/raw8 -> Voting device 2
   ```

   To make these raw devices **owner:group Oracle:dba**, refer to EMC Knowledgebase solution emc113999 to create persistent ownership of raw devices to **Oracle:dba**.
8. After Clusterware installation is complete, type the following command as **root** to confirm that clusterware is running properly:

   ```
   #<Oracle_CRS_HOME>/bin/crsctl check crs
   CSS appears healthy
   CRS appears healthy
   EVM appears healthy
   ```
9. Type **/etc/init.d/PowerPath start**.
10. Type **powermt display dev=all**, and confirm that the output lists the PowerPath pseudo device for use with Oracle ASM in the Database installation.
11. Refer to *Binding Partitions to Raw Devices for Database Files* on page 3-23 of the Oracle RAC Installation Guide.

    For the Database installation using Oracle ASM, use PowerPath pseudo names for ASM disks instead of the native device name for partioning and binding.

    Example: **/dev/emcpowerk1**

12. Install the Database and configure devices using ASM. Refer to *Installation of Oracle Database 10g with RAC Using the Oracle Universal Installer* on page 5-5 of the Oracle RAC Installation Guide.

Device file examples

SuSE device file **/etc/raw**:

```
# Clusterware Disks

raw5:sdw1 -> OCR device partition
raw7:sdp1 -> Voting device partition
raw8:sdq1 -> Voting device partition

# ASM Disks

raw1:emcpowerai1
raw2:emcpoweraj1
raw3:emcpowerak1
raw4:emcpoweral1
raw6:emcpowerae1
raw9:emcpowerah1
raw10:emcpowerm1
```

Red Hat RHEL 4.0 **/etc/sysconfig/rawdevices**:

```
# Clusterware Disks

/dev/raw/raw5 /dev/sdt1 -> OCR device partition
/dev/raw/raw7 /dev/stu1 -> Voting device partition
/dev/raw/raw8 /dev/sdv1 -> Voting device partition

# ASM Disks

/dev/raw/raw1 /dev/emcpowera1
/dev/raw/raw2 /dev/emcpowerb1
/dev/raw/raw3 /dev/emcpowerc1
/dev/raw/raw4 /dev/emcpowerd1
/dev/raw/raw6 /dev/emcpowerf1
/dev/raw/raw9 /dev/emcpoweri1
/dev/raw/raw10 /dev/emcpowerj1
```

SLES9 Oracle 10g with EMC arrays and EMC PowerPath

This section provides information for configuring persistent names for OCR/Vote devices required by Oracle Clusterware.

Note: Not applicable for file system based OCR/Vote files, e.g., OCFS2 .

SLES9 provides udev for persistent device names and device mapper with udev for multipath devices. The configuration procedure is detailed in Oracle metalink note 357472.1.

For hosts connected to EMC DMX or CLARiiON arrays and using EMC PowerPath mutlipathing software, the mechanism to persist device names is built into PowerPath software. The above-mentioned Oracle metalink Note: 357472.1 should therefore *not* be followed.

Configuration procedure

Follow these guidelines to achieve persitent names and prepare devices for use by Oracle Clusterware.

1. Install and configure PowerPath software 4.5.x on all the nodes in the cluster.
2. Prior to partitioning or using devices for Oracle Clusterware:
   a. Run **powermt display dev=all** on all nodes.
   b. Save these outputs in a safe place to ease maintenance.
3. Review the names of the devices chosen for Oracle ocr/vote device. Ensure that all the emcpower pseudo names match up on all the nodes. For example, for a two-node cluster:
   - emcpowera on hostA is for Lun 6 on a CLARiiON
   - emcpowera on hostB is for Lun 6 on the same CLARiiON

   Note: This requirement is not necessary for ASM devices but it is recommended for easier maintenance.
4. If the emcpowerX name does not match for all nodes in the cluster, use emcpadm cmdline utility on the nodes to rename the devices so that all nodes see the same LUN with the same emcpower name.

Note: Refer to *the* "emcpadm: List or Rename PowerPath Pseudo Devices" section in the *PowerPath 4.5 Product Guide*.

5. EMC recommends emcpower pseudo device names be used to prepare the /etc/raw file for SLES9 on each host. Do not use the native /dev/sdX device name. For example:

```
entry in /etc/raw
# OCR devices
raw1:emcpowerx1
raw2:emcpowery1

# VOTE devices

raw3:emcpowert1
raw4:emcpowers1
raw5:emcpowerr1
```

   Reboot all nodes.

   Note: Once names are matched on all hosts, keep a list of the powermt display dev=all output in a safe place. This is useful to maintain the Oracle RAC environment.

6. Use udev permissions file to set the permissions and ownership for the raw devices as required by Oracle RAC Clusterware Installation Guide.

   Note: Refer to the *Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for Linux*, B14203-01.

7. Follow the 10G Oracle RAC Clusterware installation procedure to install and configure Clusterware on all nodes.

   Note: Refer to the *Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for Linux*, B14203-01.

   In case a node is added, or an existing node is later replaced, repeat the above steps on the new node before following the Oracle clusterware configuration procedure.

Note: If PowerPath software is upgraded, make sure rpm –U is used to Upgrade. Following the upgrade, repeat Step 2 through Step 5.

7

SAN Management

This chapter contains the following information:

- Overview of SAN Management..... 610
- Enterprise storage management 613
- Fabric management 619

Overview of SAN Management

There are many white papers and technical documents written on SAN management. Some try and define cutting-edge tactics, or discuss the nuances of the latest trends. This chapter discusses some options and try to provide an understanding and scope of what you may need to manage your environment. The purpose of this chapter is not to become a user guide, as those have already been written, but rather to help you gain some insight and understanding on some options available for managing your SAN environment as a whole.

This chapter will attempt to define needs and then explain some options available to fill these needs. This chapter will also explore tools and products which can help perform the job.

Define your needs

When tasked with the responsibility of selecting which tools or products your organization will need in order to manage the overall storage environment, there are many variables to weigh and consider. The most basic questions are: What am I trying to manage? Am I trying to manage a large enterprise environment or just a simple fabric?

This chapter will begin to examine these questions in an effort to provide enough insight to help you make a decision based on your needs. It is important to understand what needs to be managed before deciding what will be used to manage it.

Small data centers

Perhaps you have a small data center or are part of a large company, but due to the specific nature of your environment your data center has been isolated from the rest of the company. In other words, you have a handful of servers and a few switches to manage. In addition to the small number of servers, you have a couple of small arrays, or one single mid-sized type storage array. It may be that your environment does not change too often and although it may grow, the rate at which it grows is fairly slow. You may be in a situation where the purchase of an additional product would be hard to justify. Since your environment does not change very often, it may be worth investing the extra initial setup hours to set up the environment, and on the occasion when a new server is added or a zone needs to be changed, you can use native tools. In general, using these tools can take more time, but if you do not make constant updates it may then

meet your needs. However, it does mean that you may need a wider set of skills since you will need to know how to run each switch's management interface, the different operating systems, and the storage interface. It can become cumbersome to gather information about the different components in your environment because each of these components has different interfaces and reporting mechanisms.

Larger data centers

Keeping in mind what was previously discussed about the different interfaces and different tools you would have to access just to see storage and then imagine having to perform these tasks for as many as twenty servers, or more. It becomes extremely time-consuming and very cumbersome. In some instances you may be able to script some steps to save time, but regardless, there are still many different pieces to the task of assigning storage to a newly connected server.

Complex fabrics

What about complex fabrics? Imagine having to diagnose problems if you can't manage the whole environment. Your switch tool may tell you something is not connected, but what if there is a problem with the storage array itself, or perhaps there is a problem with your server? Maybe you have been tasked to gather performance reports on your environment. If you had to manage all these components separately, it would become very difficult and time-consuming.

Options

So how do you solve this issue? Do you just hire more people? In this day of do more with less, it is not likely an option. It may be more cost-effective, and certainly more effective from a time-management standpoint, to purchase a product which has an interface to all the different components in your data center.

An enterprise management tool, like EMC ControlCenter, could be the answer. The question then becomes, "why not buy this type of product all the time?" For some, the answer is simple: the investment of the product. First, there could be initial cost and then the time needed to learn how to use the tool, the maintenance costs, licensing, updates, and (depending on the size of the environment being managed) the infrastructure costs. Most enterprise tools require at least one dedicated server and many times these servers are not your typical desktop.

Often it all becomes a trade-off and you have to list out the pros and cons of each scenario to gain a greater understanding of your objectives and overall needs.

Some other things to consider when debating whether you need an enterprise product or a switch tool is the level of granularity you may

desire. Sometimes enterprise tools do provide ease-of-use at the cost of granularity, meaning that if you want to get down to the ones and zeros of a switch, you may still need to use the native interface. There are some purists who would argue that if you are truly going to manage a switch then you have to use the native tools, but in most cases an enterprise tool will allow a user to perform the majority of their tasks from a single interface.

This is where enterprise tools are valuable. Think in terms of your training cost. If you need to become proficient with a single tool versus multiple tools, the time savings and learning curve costs could be significant. This discussion has only focused on switch management, but if you add in the overall capabilities gained when using and distributing an enterprise tool such as EMC ControlCenter, suddenly you are able to perform many more duties in half the time, saving both time and money.

This chapter aims to provide you with enough information to make an informed decision, based on your needs, to manage your SAN environment.

Enterprise storage management

This section does not compare diffrent enterprise tools, but rather provides information on EMC ControlCenter. The benefits of this tool will be briefly discussed. For more in-depth information on EMC ControlCenter, refer to Powerlink.

This section will focus on fabric management and overall or high-level storage management options and commands. For purposes of discussion, *enterprise storage* refers to the ability to not just manage a switch, but also the ability to assign and monitor storage from a single interface.

Although this section will not specifically discuss storage or server-based management, some of the capabilities of enterprise management tools can, and do, perform those types of functions. If it is the intention to have a single tool or product which can perform this type of activity, then you would need to focus your efforts on understanding the kinds of tools available, and what functions they provide.

EMC ControlCenter

When you think in terms of enterprise, most people think about one tool to manage their whole data center environment. At the very least they would like to reduce the number of tools needed in order to keep costs down. One product which can address this very need is EMCs ControlCenter. With this one product a person can perform many functions within their environment. They can do things like assign storage, manage and monitor servers, generate performance reports, zone and monitor switches, and keep a running inventory on their hardware.

Storage Area Networks are typically composed of many components. Problems associated with setup, maintenance, and management of the storage network can multiply. IT managers face the need to use a different storage management tool from each provider. The EMC solution is a comprehensive set of storage management software tools that support the entire switched fabric infrastructure. The EMC ControlCenter suite of tools provides proactive management not only to the storage devices, but also to the storage network. SAN Manager, a component of EMC ControlCenter, provides access control of storage volumes, as well as management of heterogeneous Fibre Channel switches.

EMC ControlCenter can be broken down into areas, or applications. Each application has its intended purpose and scope. For further information, refer to documentation available on Powerlink.

The tasks for managing your data center are provided by a number of closely integrated applications. These applications fall into three Storge Resource Management (SRM) categories:

- Planning and provisioning
  - SAN Manager
  - SAN Advisor
  - Automated Resource Manager
- Monitoring and reporting
  - *Performance Manager*
  - *StorageScope*
  - *StorageScope File Level Reporter*
- Storage device management
  - *Symmetrix Manager*
  - *Symmetrix Optimizer*

For purposes of this section, applications more appropriate for this discussion will be discussed, but if you would like details on the above categories, or more in-depth information, including how to manage an array, please refer to the EMC ControlCenter documentation available on Powerlink. This section will focus specifically on SAN manager and some monitoring and reporting capabilities as it relates to switch or fabric management. This section will briefly discuss the abilities built into EMC ControlCenter.

SRM

The SRM Storage Device Management solution allows you to manage your storage environment including Symmetrix, CLARiiON, non-EMC storage arrays, NAS, and content-addressed storage (CAS).

Storage Device Management includes the following products:

- Symmetrix Manager
- Symmetrix Optimizer

Storage agents contain the functionality to perform the daily storage management tasks and problem analysis required to ensure the best utilization of resources.

Using array agents you can discover, explore, report, and view configuration attributes and status. Some agents also enable you to:

- View the topology, relationships, properties, and alerts for a storage array
- Display the configuration of the storage array through the Visual Storage view (CLARiiON, HDS, and HP StorageWorks).
- Automatically allocate storage on a CLARiiON or HP StorageWorks array through Storage Provisioning Services.

The above information just begins to define the capabilities you have with Symmetrix Manager. The tool gives a user the ability to essentially configure arrays from a GUI environment. Many administrators think this is an essential component for an enterprise tool. Others may argue this is a very dangerous element as the wrong command can have negative consequences. EMC ControlCenter has ways of controlling who has privileges to execute configuration commands based on assigned access rights. For more information about assigning privileges please refer to *ControlCenter User Guide*, available on Powerlink.

Enterprise reporting

Reporting and monitoring are key components to most tools today, especially when talking about an Enterprise product. The product has to have the capability to monitor and report against as many components in an environment as possible. Monitoring and reporting solutions allow you to effectively utilize your storage assets, manage your inventory, and assess how your storage is performing. Many fabric- or switch-specific products focus on the Fabric as a whole and may not take into consideration storage array performance or server performance. EMC ControlCenter can provide insight into your environment as a whole.

If you want to understand your fabric from the server, through the switch, and to the storage, using EMC ControlCenter will monitor the environment from one end to the other. You can gather performance metrics along the way, perform analysis as to the best path utilization, or get a quick snapshot of your overall freespace. The component within ControlCenter which provides much of the report generation is called StorageScope.

StorageScope is an EMC ControlCenter product that provides a variety of SRM reports to help customers assess their current storage environment, determine future storage requirements, and provides capacity utilization reporting.

Note: For detailed information about StorageScope reports, refer to the *EMC ControlCenter StorageScope Reference Guide*,available on Powerlink.

Because StorageScope is fully integrated with ControlCenter, it uses many common ControlCenter components and services. For example, StorageScope:

- Populates its reports with data that has been gathered by ControlCenter agents and stored in the Repository.
- Collects and correlates metrics across:
  - host databases, file systems, and volume managers
  - switches
  - arrays
  - backup applications
- Uses ControlCenter security and access management services to grant users access to StorageScope.
- Provides file-level summaries by age and size distribution.
- Employs ControlCenter object management services for grouping storage assets by line of business, location, application, billing center, or other concept. This allows users to customize StorageScope's reporting capabilities to meet specific business needs.

For increased flexibility, users can create custom report layouts and can also export reports into CSV, PDF, or XML format. Integration with open-source applications (for example, billing or customer care) is enabled through standard SQL views into the ControlCenter Repository and direct access to StorageScope XML files.

Note: For more information about using the StorageScope API to access Repository data, refer to the *EMC ControlCenter StorageScope API* document, available on Powerlink.

Enterprise SAN Management

The Storage Resource Management (SRM) Planning and Provisioning solution allows you to design, plan, and provision a multivendor storage infrastructure for the benefit of improving the utilization, performance, and the cost effectiveness of your storage assets. This solution is achieved using ControlCenter planning and provisioning applications, which allow you to quickly recognize, isolate, and respond to storage issues.

The applications included in the planning and provisioning solution include:

- SAN Manager
- SAN Advisor
- Automated Resource Manager

This section focuses on the first of the three components — SAN Manager. This is where an Administrator will spend a good deal of their time within this tool. For more information and more details about all three of these components, refer to EMC ControlCenter documentation available on Powerlink.

A storage area network (SAN) is the collection of end ports and connectivity elements in a network of interconnected hosts and storage devices. A SAN typically consists of many components including hosts, HBAs, switches, hubs, bridges, and storage devices, all potentially from different vendors. SAN Manager supports the entire switched fabric infrastructure, providing integrated network discovery, topology, and alert capabilities, and actively controlling SAN management functions such as zoning and LUN masking.

SAN Manager enables you to:

- Perform discovery of the logical and physical view of the SAN through agents.
- Monitor the health of connectivity devices and their ports to detect potential error conditions and to remedy or respond to any errors that occur.
- Monitor the connectivity device ports in your SAN by collecting statistical information. Alerts are triggered when thresholds are met or exceeded.
- Perform zoning operations for switches and mixed fabrics, including the ability to import active/inactive zoning, to create zones and zonesets, and to configure changes.

- Perform masking operations on various arrays including Symmetrix, CLARiiON, HP StorageWorks HSG-80 and HP StorageWorks, HDS, and SUN arrays. A command line interface is available.
- Monitor and troubleshoot paths from hosts to storage.

Note: For more details about SAN Manager and ControlCenter capabilities, please refer to documentation available on Powerlink.

Fabric management

Storage area administrators sometimes need more granular detail into the SAN fabric components. ControlCenter may not have the appropriate programmable interfaces or toolsets to integrate with vendor-specific native management capabilities. Viewing certain aspects within the constructs of vendor hardwar,e such as departmental or director class switches, ISLs (inter-switch links), and multiprotocol routers, is sometimes a necessity that not all proprietary applications can handle.

Native tools from switch vendors offer three tools in common. They all allow:

- Connectivity to devices through telnet or terminal through a CLI (Command Line Interface).
- Embedded 'product/element' manager.
- Interactive GUI to view the fabric.

One example where native tools may become necessary is during an episode where the management GUI cannot connect to a device, therefore troubleshooting through the CLI (command line interface) to see if a particular device is available may become necessary. This is also pertinent in understanding why, when in high-security environments, telnet through SSL may only be allowed. Also, if a network is not unavailable, an administrator can access any switch through internet browser and download their target fabric switches' management applications for quick download in a situation where alternative means are not ready. Native tools can be useful to the traveling systems administrator when robust tools cannot be connected to or moved easily.

The use of these tools is at the discretion of the customer, but understanding the tools' differences and what a customer is looking for can help narrow the choices for the administrator. Being able to discriminate between styles of reporting, auditing, and fundamental behaviors an administrator can also determine which tools best suit an environment.

Interface options

Typically when talking about interface options, the questions that come to mind are: What does it look like?" and "How easy is it to use?" Other considerations could be whether there is a CLI available, if everything can be done from a graphical user interface (GUI), or is a combination of things needed to accomplish the tasks.

Most vendors offer a graphical user interface (GUI) and also provide a command line interface (CLI). Each major vendor has developed an interface to manage their hardware components, whether these components are departmental switches or enterprise switches, and even routers. Some like the ability to use the CLI as they feel they can move around faster and do not have to wait to be prompted for the next item to click. Others like the idea of a GUI since the learning curve is not as drastic as it can be with the CLI and people can sort their way through a process or task easier. With a CLI you can often write a script to perform tasks unattended or for repetitive actions. With a GUI, you are more restricted to performing tasks since you must select the options by clicking on your choice. Today some GUIs are building in task lists, where you can define a series of actions to execute at a later time. This can be beneficial if there are restrictions to when changes can be made in the environment. Building tasks would allow the senior staff to build a series of tasks, almost like batch jobs, to be executed or initiated after hours.

The next several sections will explain some of the different GUI interfaces that may be encountered when using a Fabric Management tool. These include:

- "Discovery" on page 621
- "Zoning" on page 629
- "Alerting" on page 638
- "Monitoring" on page 645

Attention will be focused on four specific areas of Fabric Management and will illustrate what each interface may look like so you can see the similarities and differences.

Note: For EMC ControlCenter specific documentation , please refer to the documentation and white paper libraries on Powerlink.

If you would like more information on a particular vendor CLI or EMC's CLI, refer to the appropriate CLI user's guide.

Discovery

This component of SAN management is a critical path to the overall management of your environment. Obviously, without some form of discovery, there is nothing in your environment to manage. Discovery usually goes out and interfaces with the switches in a fabric, or multiple fabrics, and loads information about those switches into some form of resident database. The information would include things like hardware type, firmware versions, and port information to name a few. Once a discovery is completed, a user will have the ability to display a topology view which provides a layout of the overall fabric as it has been discovered. Again, for more detailed information or step by step procedures on how to discover a switch or fabric, please refer to the appropriate user guide. In the case of ControlCenter, please refer to the documentation area on Powerlink. In the case of a vendor application, please refer to their user guides as well.

The following examples are used in this section:

- "EMC ControlCenter example", next
- "Brocade Fabric Manager example" on page 623
- "Cisco Fabric Manager example" on page 625
- "Connectrix Manager example" on page 627

EMC ControlCenter example

ControlCenter discovers a wide range of Fibre Alliance-compliant connectivity devices in the topology and renders them accessible in both the Console tree and the topology map. When ControlCenter encounters a fully supported switch during discovery, it attempts to discover topology information for that switch, optionally prompting for connection settings, and optionally fabric information, as needed.

Figure 132 illustrates the **Search for Connectivity** dialog box, used for discovering a Connectivity device.

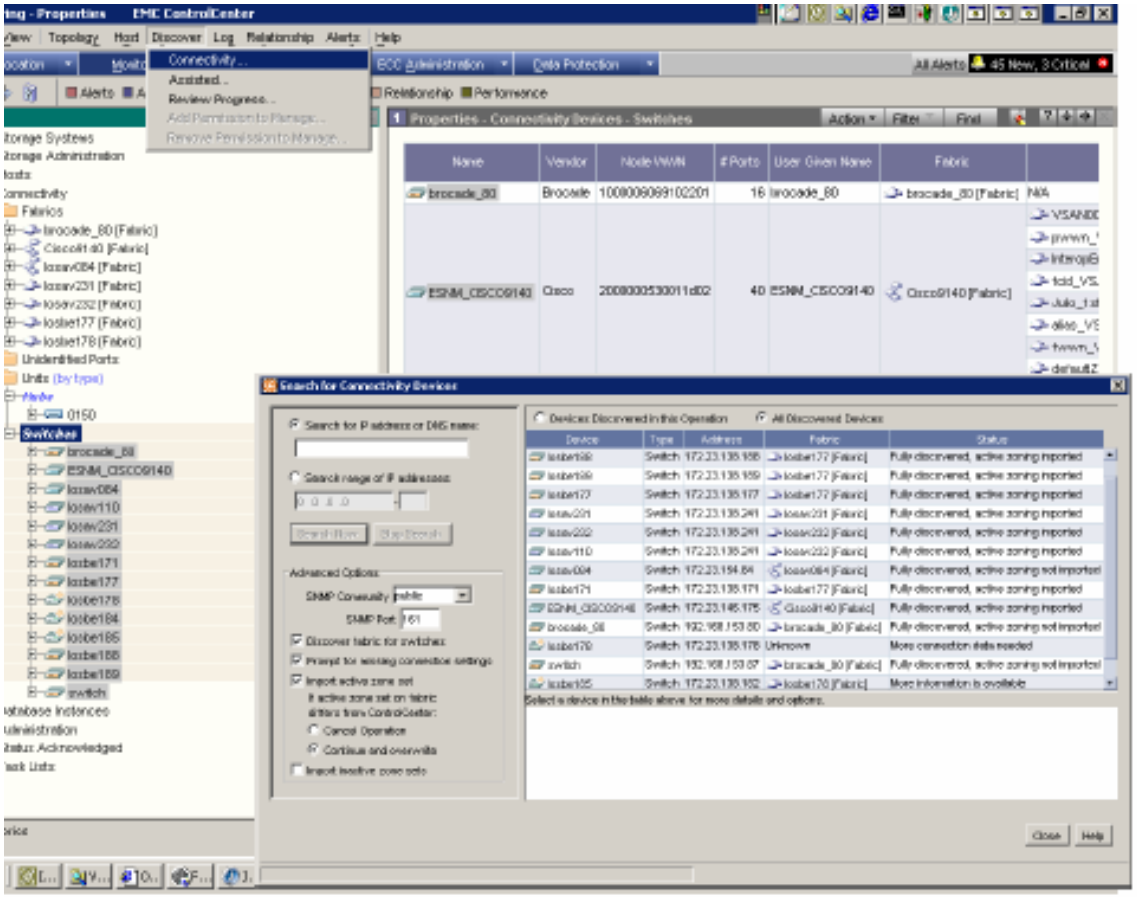


Figure 132 Search for Connectivity dialog box

Figure 133 on page 623 shows the topology view after a discovery is performed.

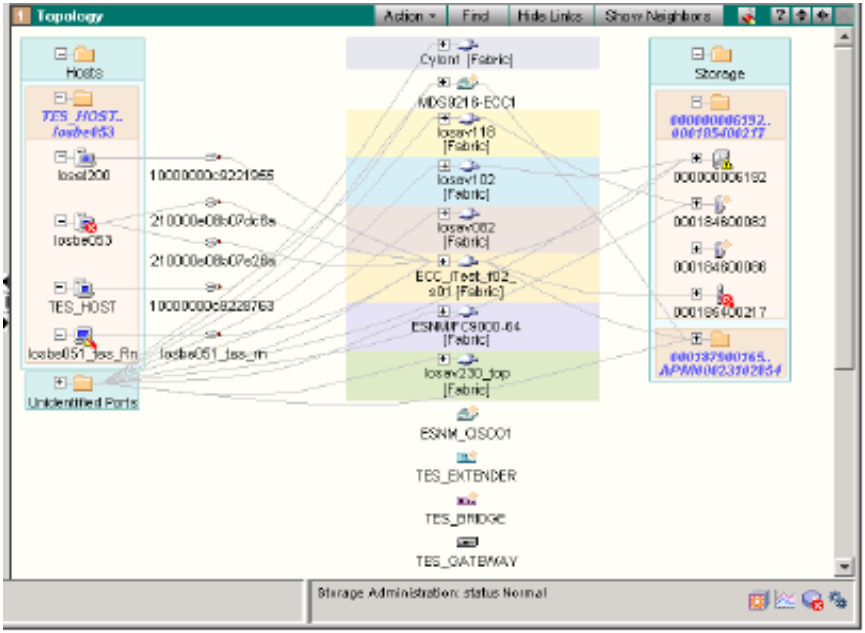


Figure 133 Topology view

Brocade Fabric Manager example

Brocade's Fabric Manager provides a few ways to perform discoveries. You can add a single switch at a time, or run a subnet scan to discover multiple switches. It provides flexibility when it comes to topology views, as you can name / rename fabrics and switches, as well as many other options. For detailed information about discovery and topology using Brocade's Fabric Manager, please refer to the Brocade Fabric Manager Administrator Guide available on Powerlink.

Figure 134 on page 624 is the main view in Fabric Manager. In order to discover a switch, simply type in the IP-address or the switch name and you will be prompted for a username and password. Upon entering the appropriate information the switch will be added to the view in Fabric Manager.

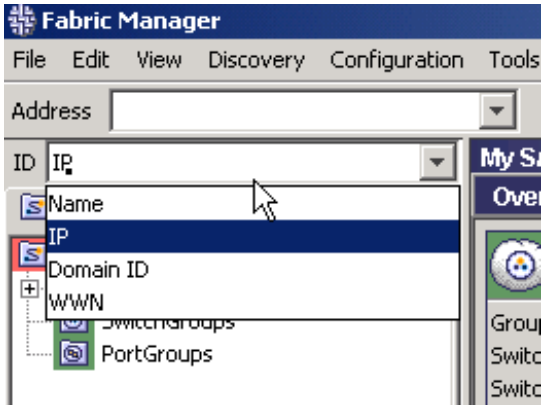


Figure 134 Fabric Manager main view

Figure 135 show different topology views available within Fabric Manager.

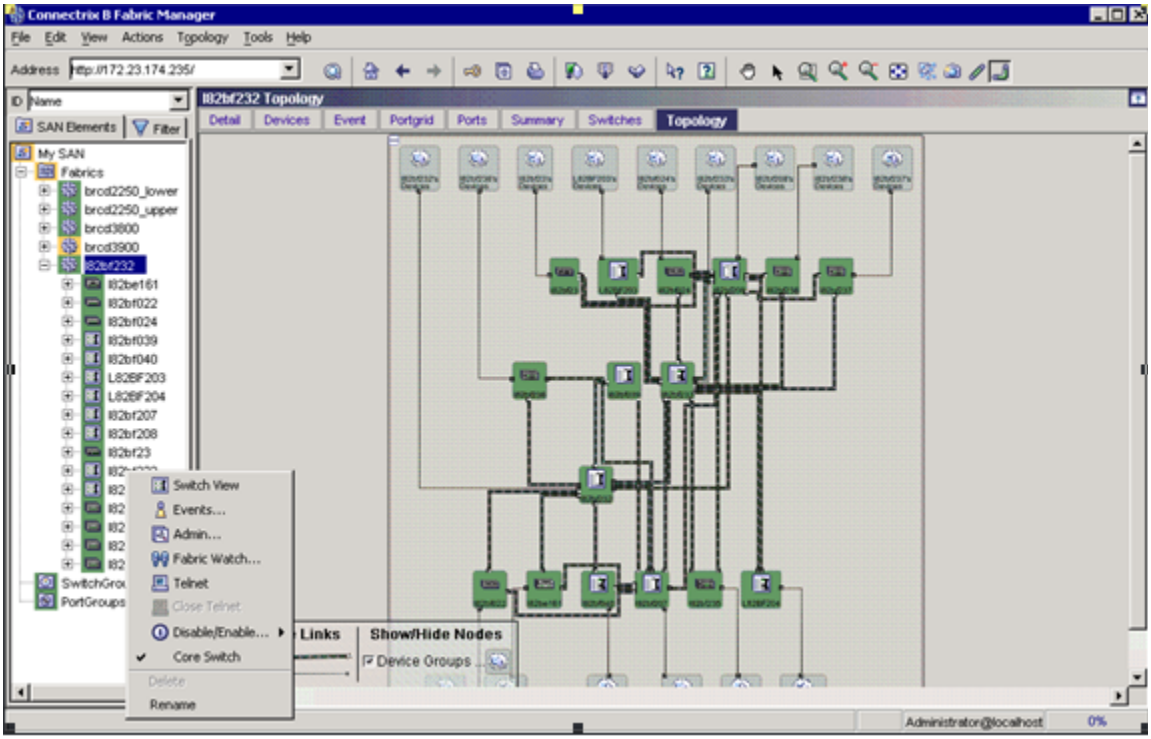


Figure 135 Core/edge topology with core switch designation

Figure 136 shows the same fabric displayed in circular and tree designs, respectively.

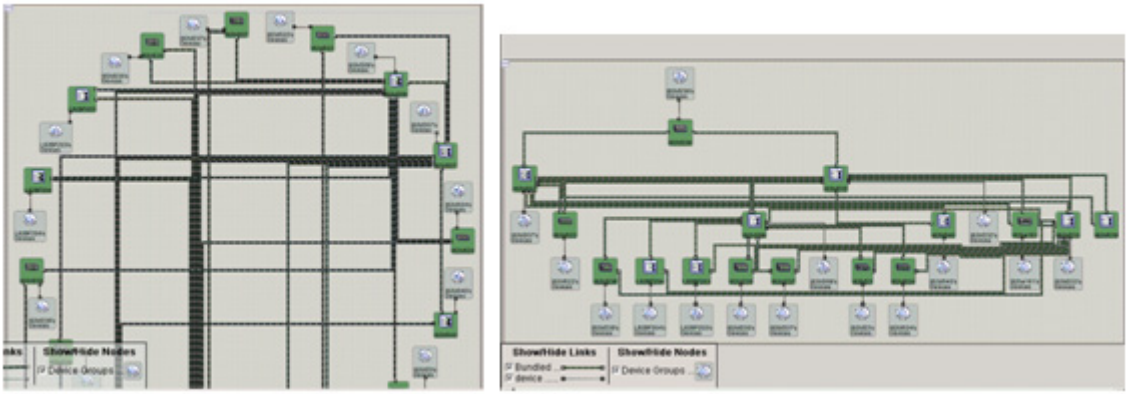


Figure 136 Circular and tree designs

Cisco Fabric Manager example

Cisco Fabric Manager is a Cisco-developed switch hardware and fabric management application. It is a Java-based application that is downloaded from the switch through opening a browser direct. There is no Service Processor style management system, like Connectrix Manager. A user may download this to any Windows NT/2000 or Solaris host on their network. By downloading the application you are, by default, discovering a fabric. You do have the ability to scan subnets for more switches. Once you have a discovery, the Fabric Manager view will display the fabric topology.

Fabric Manager has a second component, called Device Manager, that can be downloaded in the same way. It is intended to be an individual switch management tool. Nearly all switch monitoring functions can be accomplished through Fabric Manager; however, Device Manager makes single switch management easier. For more information about Fabric Manager or Device Manager refer to Cisco's website.

Figure 137 shows the main view from Cisco Fabric Manager.

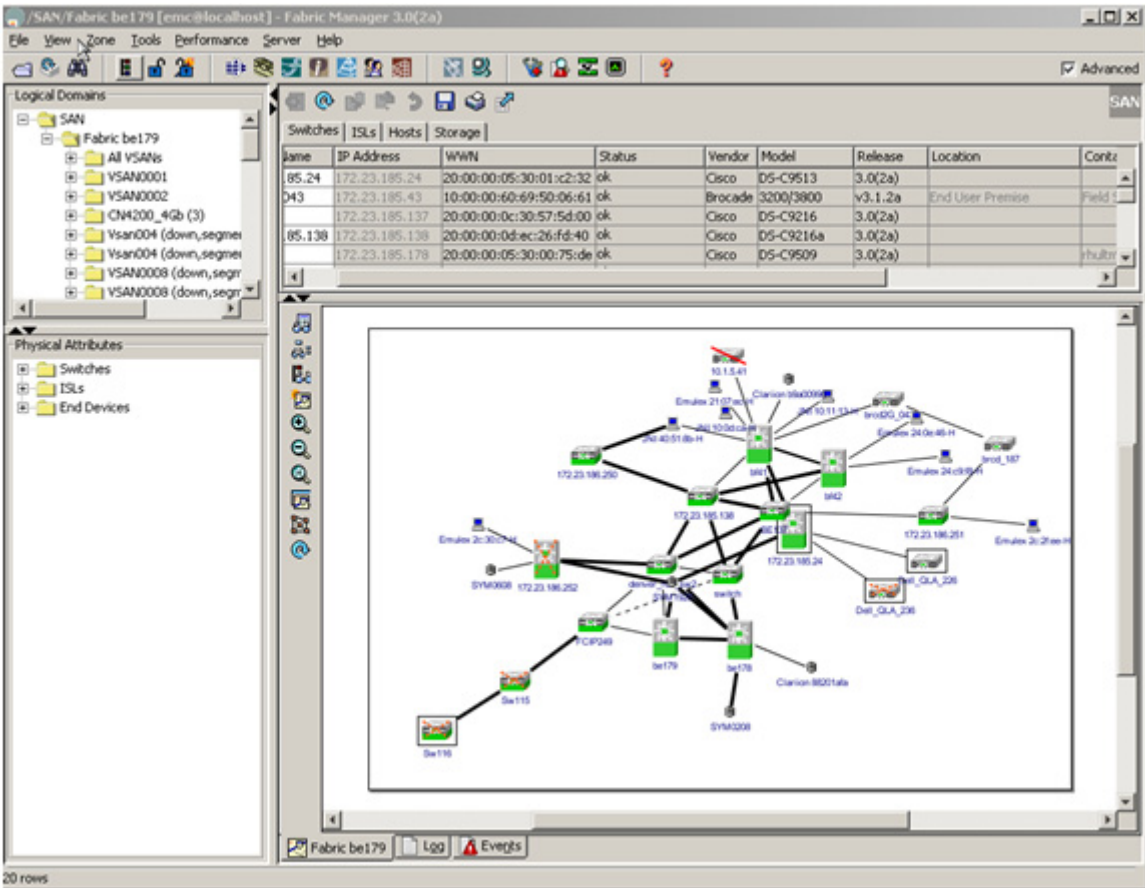


Figure 137 Cisco Fabric Manager main view

Figure 138 shows Cisco's Device Manager default view.

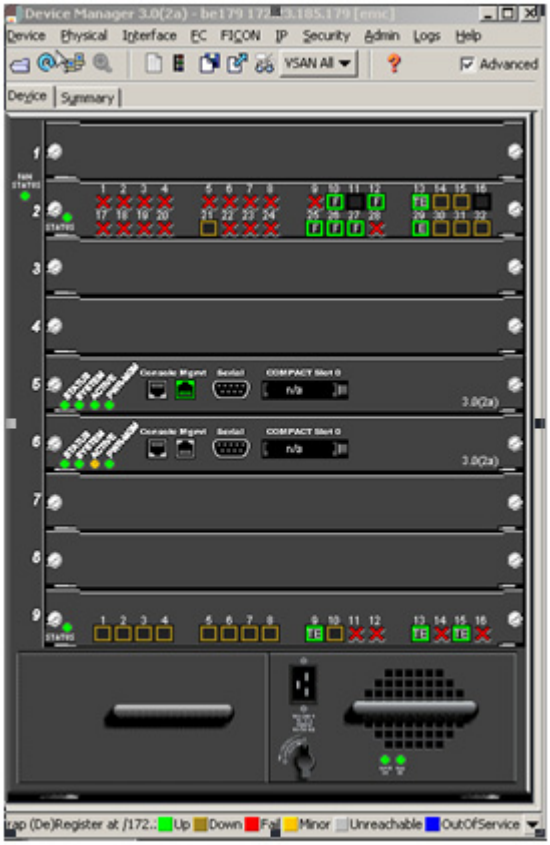


Figure 138 Device Manager default view

Connectrix Manager example

EMC Connectrix Manager is similar to Brocade M-Series's product, EFCM (Enterprise Fabric Connectivity Manager).

Using the Connectrix Manager product, there are a few screens associated with discovery. Depending on how you set things up, you can have in-band or out-of-band discovery take place.

Figure 139 allows for this basic setup.

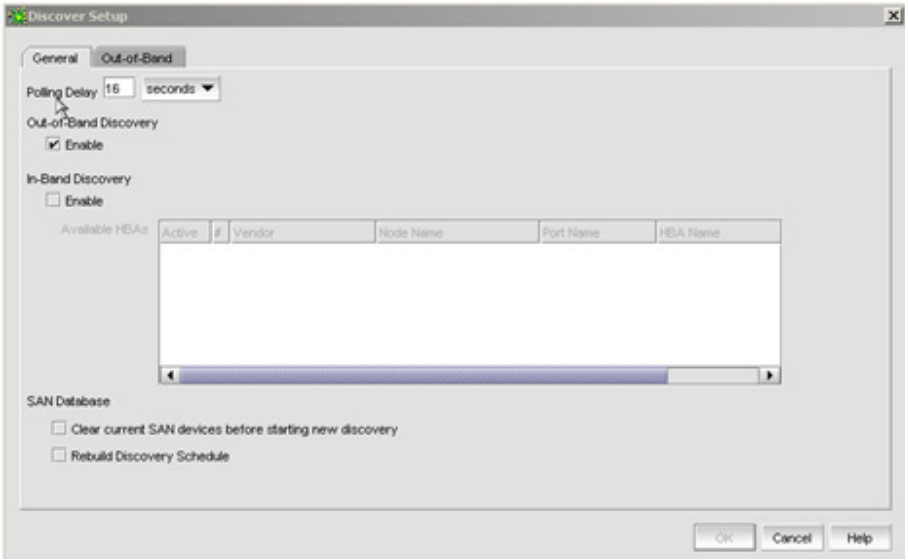


Figure 139 Connectrix Manager basic setup

Figure 140 shows how you can select different switches to be managed.

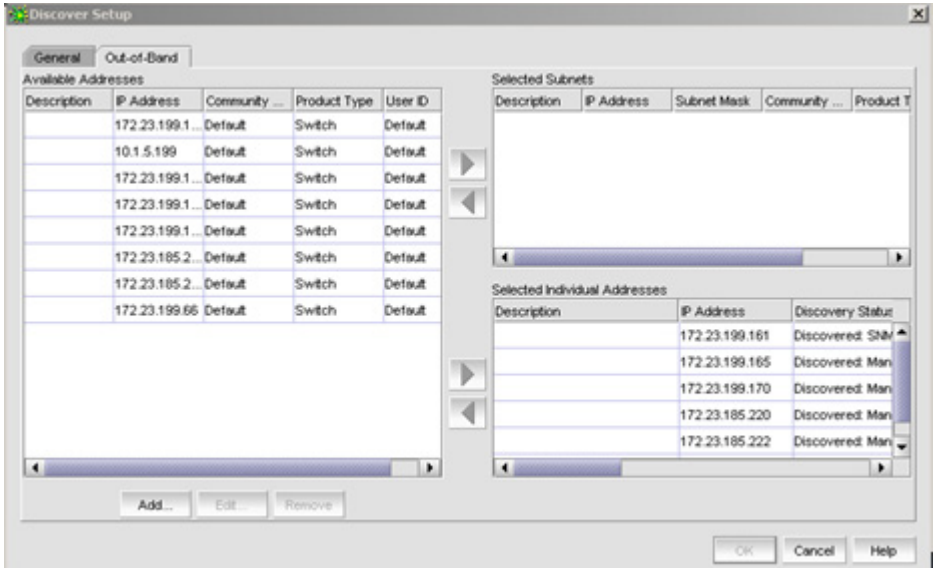


Figure 140 Selecting switches to be managed

Figure 141 shows the main view from Connectrix Manager. From this main view you have full control over the fabric or fabrics you are trying to manage as well as the ability to manage the physical switches as well.

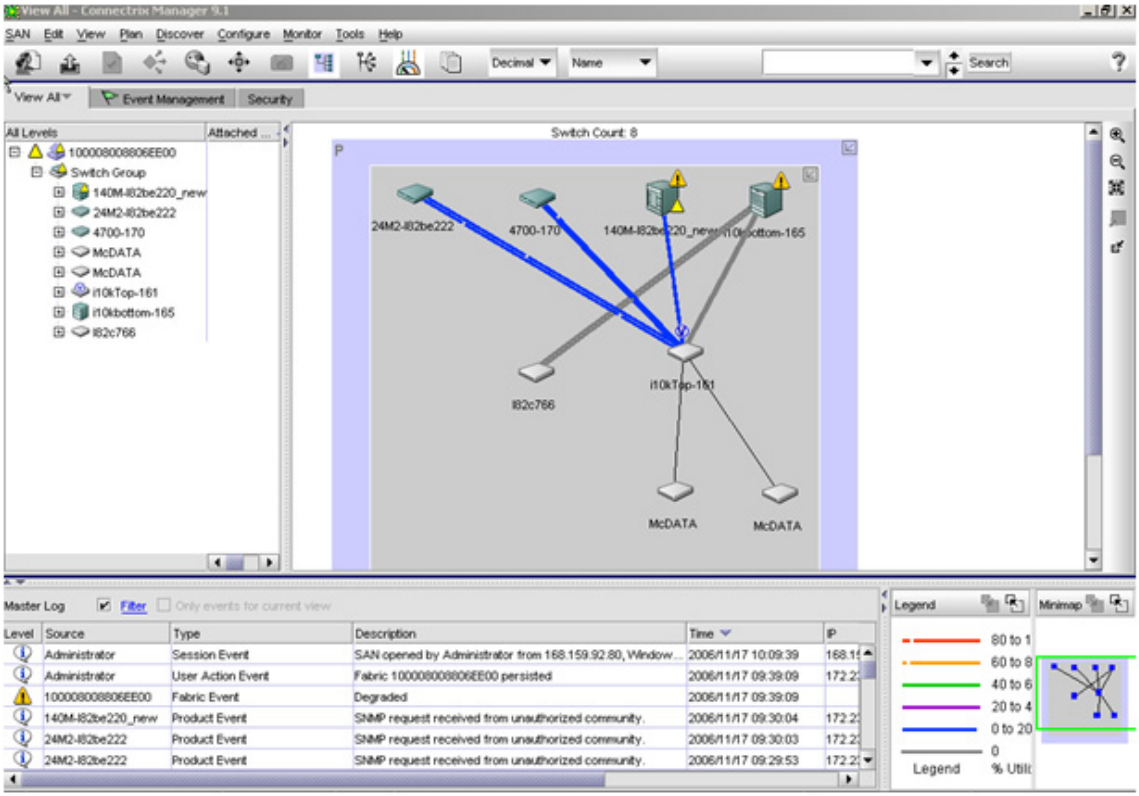


Figure 141 **Connectrix Manager main view**

For more information about Connectrix Manager see the *Connectrix Manager User Guide* available on Powerlink.

Zoning

Zoning enables a set of devices connected to a switched Fibre Channel fabric to communicate with each other; for example, a host and a storage array. Each zone groups the end ports of the devices involved, or the switch ports physically connected to those end ports. Using multiple zones, a single host can communicate with multiple

storage devices, and vice versa. A zone set is a collection of zones that can be activated together, partitioning a fabric into zones. Only one of the zone sets associated with a fabric can be active at any time. It is this active zone set that determines which of the devices connected to the fabric can communicate with each other.

Many administrators consider having an easy to use interface to create and maintain zones very important. All of this zoning information is again retained in a zoning library, which can be maintained at a switch level or in a database within the Fabric tool being used.

The following examples are used in this section:

- "EMC ControlCenter example", next
- "Brocade Fabric Manager example" on page 633
- "Cisco example" on page 634
- "Connectrix Manager example" on page 636

EMC ControlCenter example

Zones and zone sets can be created, modified, renamed, and deleted. A fabric's active zone set cannot be modified, but the active zone set can be changed by activating a different zone set. Each fabric's zoning configurations are stored in the following folders in the Console tree panel, under the fabric's folder within the parent Connectivity folder:

- Active Zone Set folder — Contains the current active zone set.
- Planned Zone Sets folder — Stores inactive zone sets that have been created or modified by users, including a copy of the active zone set and a copy of the previous active zone set. Inactive zone sets imported from the fabric are also copied to this folder. Zone sets in this folder can be modified and activated.
- Planned Zones folder — Stores inactive zones that have been created or modified by users. Inactive zones imported from the fabric are also copied to this folder. Zones in this folder can be modified.
- Zone Sets and Zones folders — Located under each switch in the Switches folder. Inactive zoning configurations imported from the fabric are placed in these folders, then copied to the Planned Zone Sets and Planned Zones folder where they can be modified.

Note: If zoning has not yet been imported from the fabric, all zoning folders are empty.

Figure 142, next, and Figure 143 on page 632, and Figure 144 on page 633, show zoning options from EMC ControlCenter.

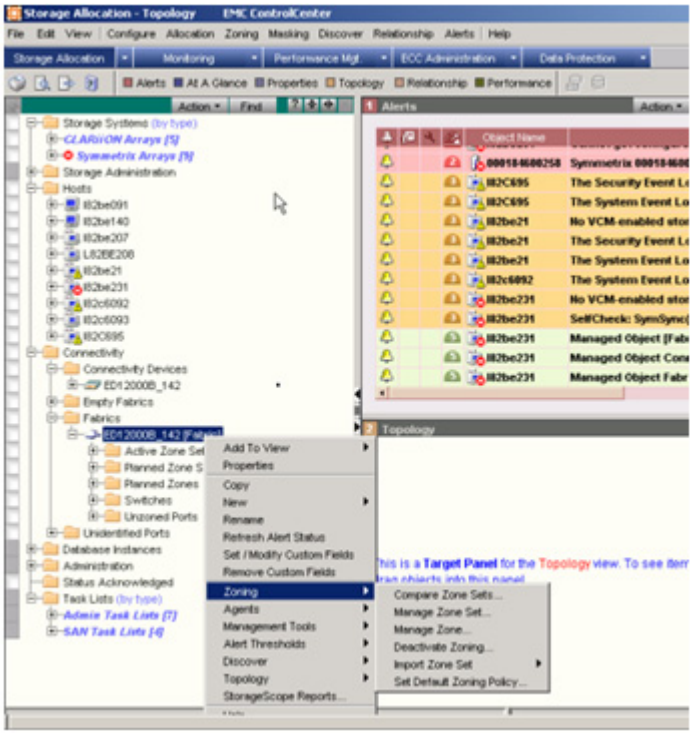


Figure 142 **Zoning option**

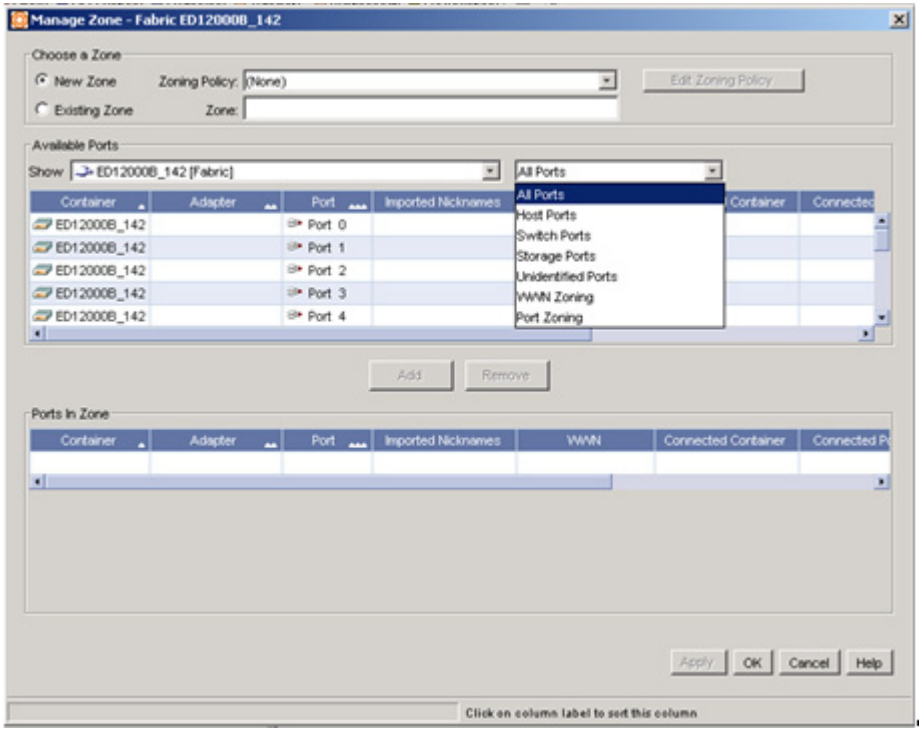


Figure 143 **Manage Zone option**

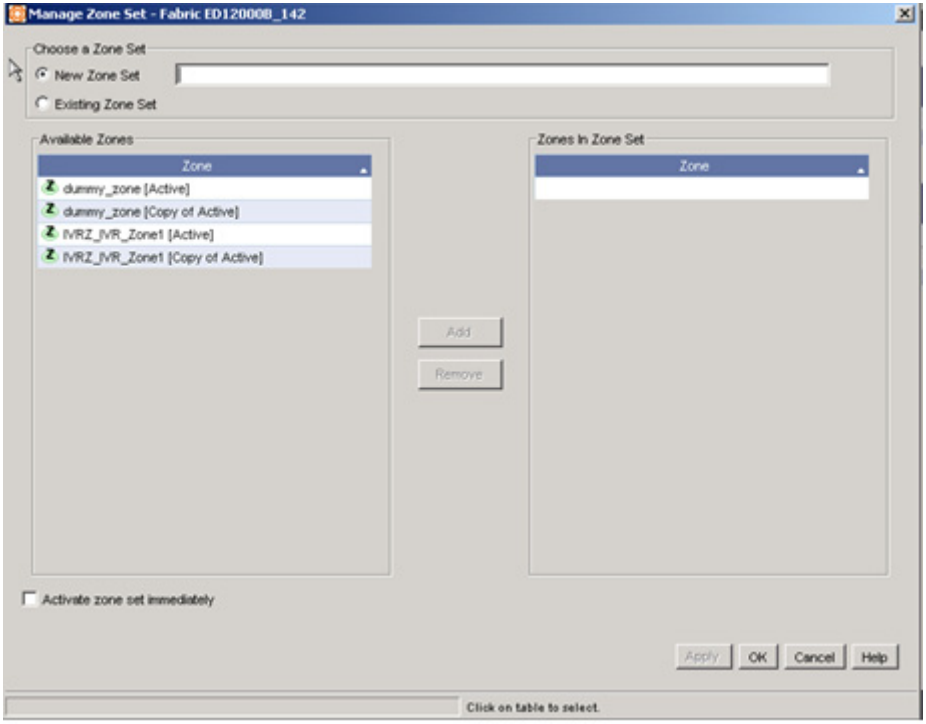


Figure 144 **Manage Zone Set option**

Brocade Fabric Manager example

Zoning is entirely based on Web Tools and the CLI. Zoning may be launched from Fabric Manager (see Figure 145 on page 634.) Review the Connectrix B-series Zoning and Command Reference Guides, available on Powerlink, for more information.

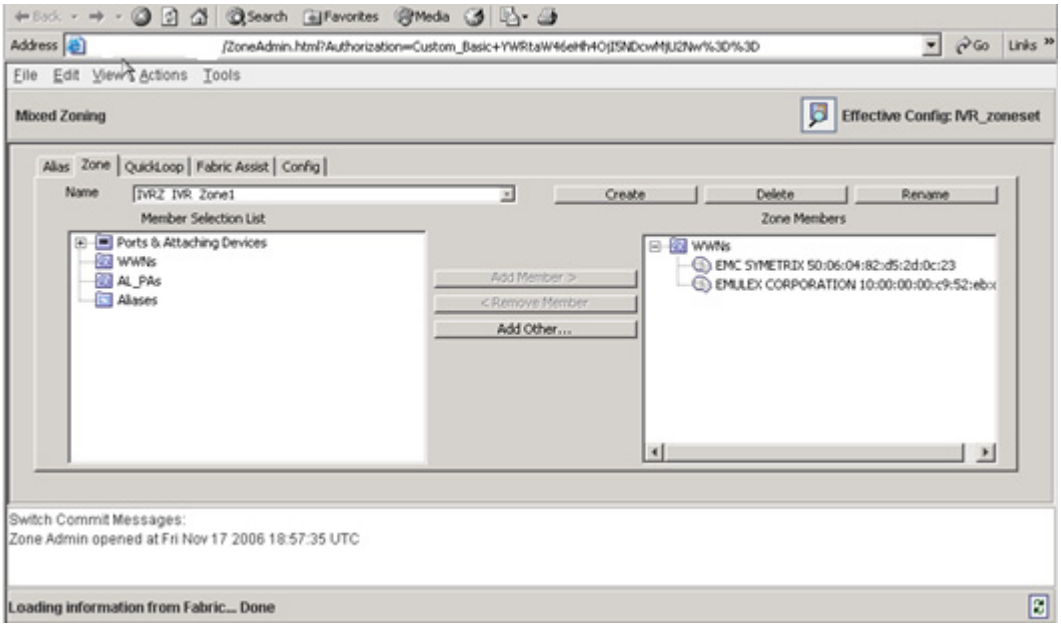


Figure 145 **Broade Fabric Manager, Zone tab**

Cisco example

Zones and zone sets are based on VSANs. Each VSAN has its own zoning database. A zone or a zone set from one VSAN cannot be applied to another VSAN. Multiple zones and zone sets can be created per VSAN; however, only one zone set is active at a time (see Figure 146 on page 635). For more information refer to Cisco's website.

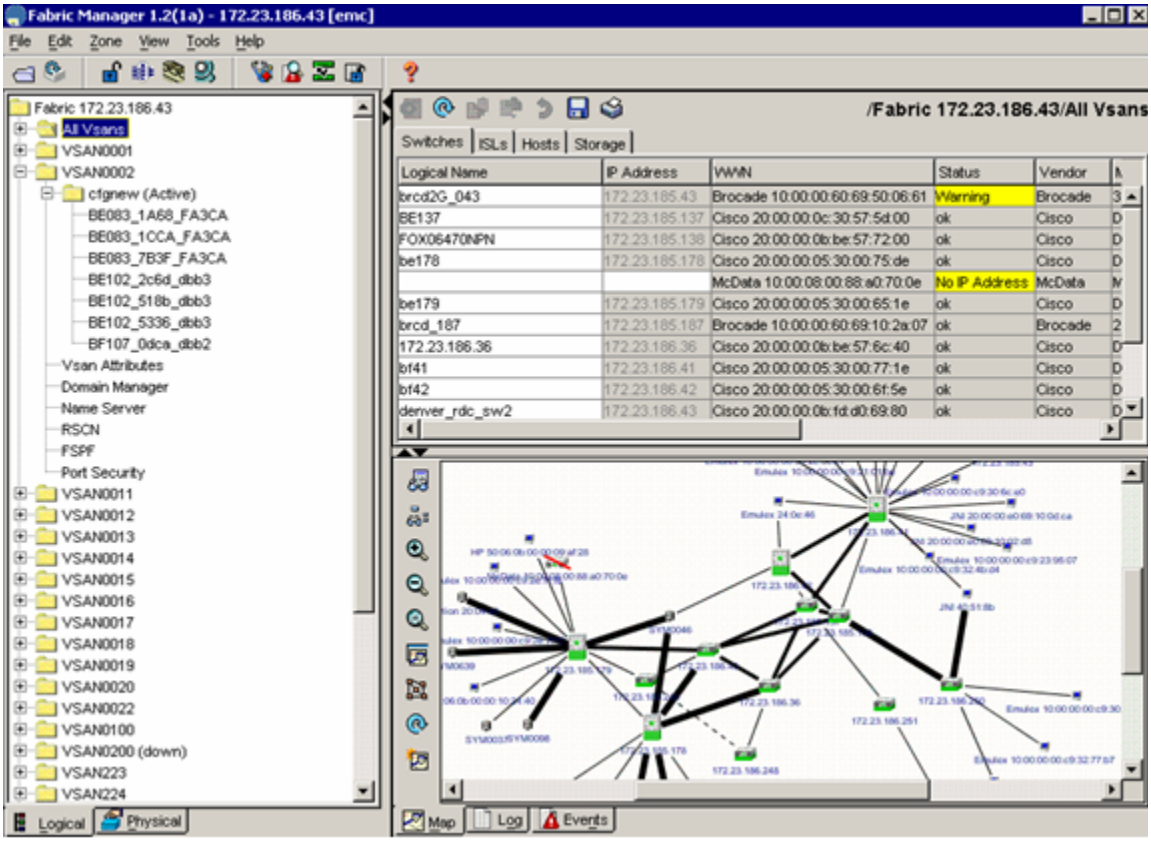


Figure 146 **Cisco Fabric Manager, Switches tab**

By highlighting a particular VSAN in the left pane, the nodes and ISL that play an active role in the zoneset are highlighted in bold in the topology map. See Figure 147 on page 636.

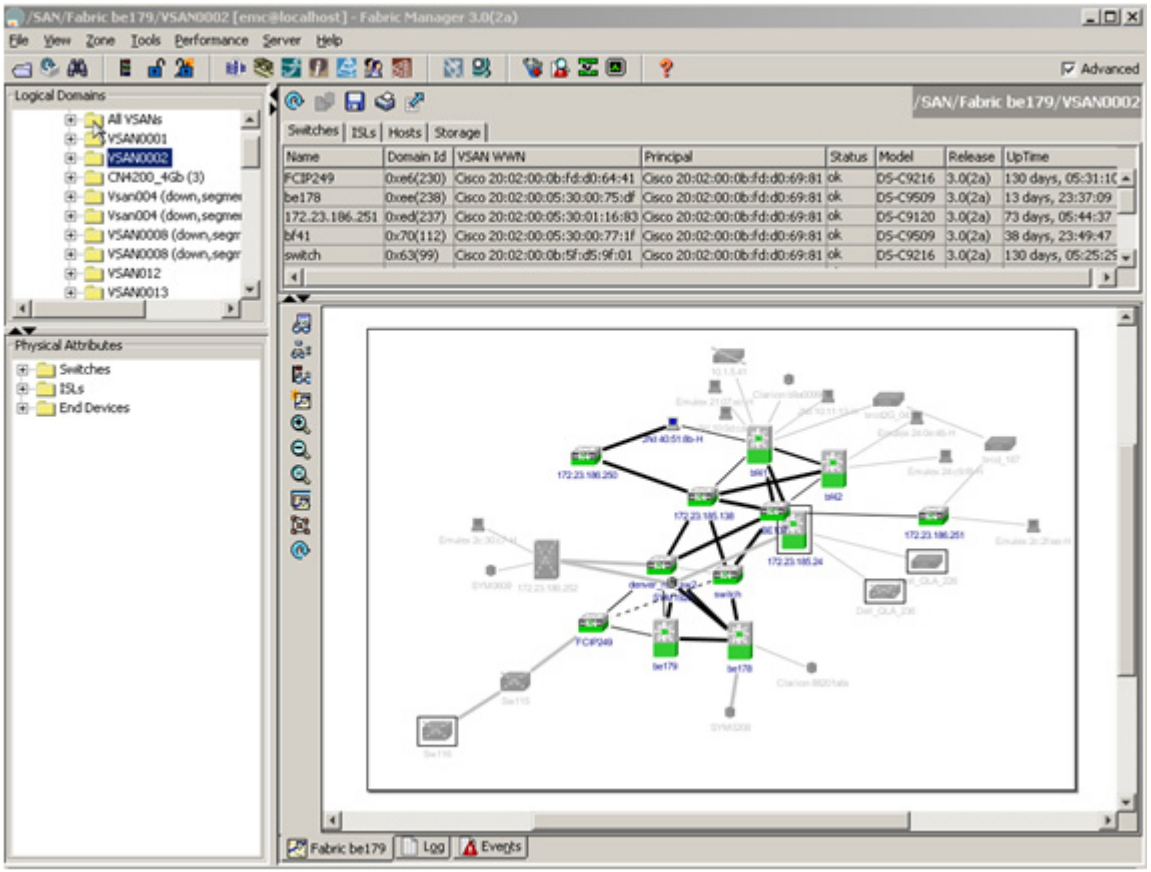


Figure 147 Fabric Manager VSANs

Connectrix Manager example

Connectrix Manager allows the user to create hundreds of zones and multiple zonesets with hundreds of zones within those zonesets. All reside in the zoning library on the Connectrix Service Processor. See the most recent Connectrix Manager release notes available on Powerlink for zoning limits.

The zoning system in Connectrix Manager provides a Nickname function where a meaningful ASCII string can be associated to a WWPN.

Remove, Replace-All, Rename, Duplicate, and Find functions are available for zone library management. Zoning by WWN or

Domain / Port is supported. A Properties function at the zone set level provides statistics for number of zones and number of unique members in a zone set.

Zoning is done from one GUI screen. Go to **Configure> Zoning**. Create zone sets and library entries here, as shown in Figure 148.

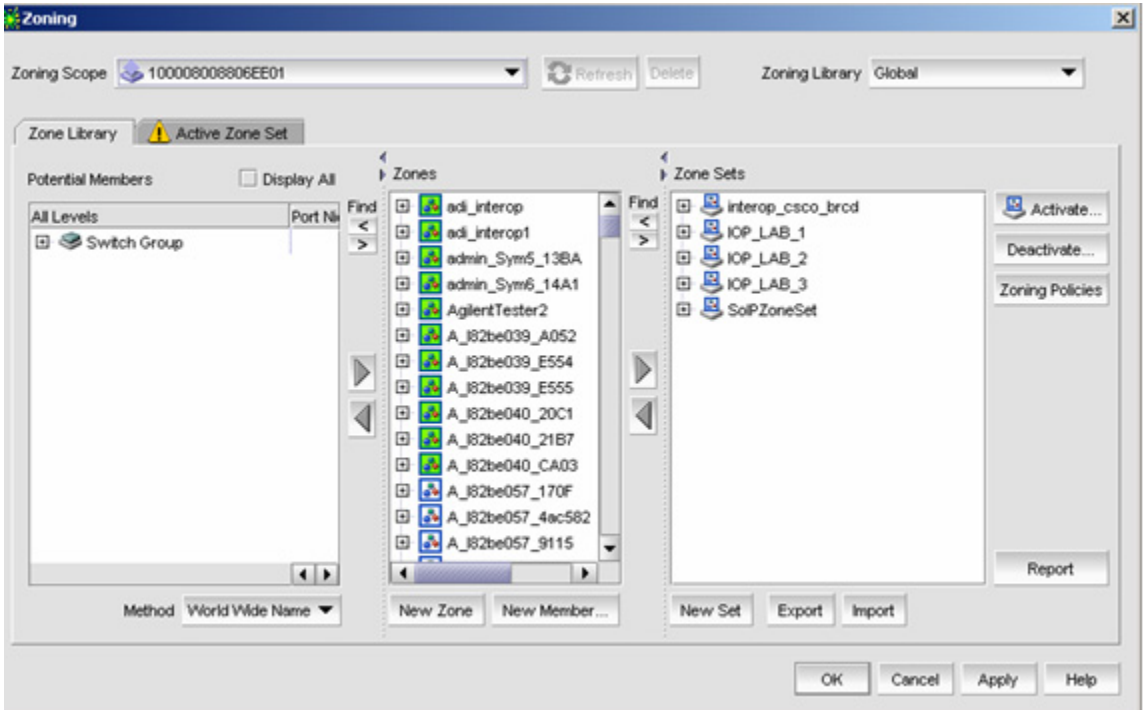


Figure 148 **Zone library**

Multiple zone sets may be present on a Service Processor. The **Zone Set** tab will display the active zone set for a selected fabric (see Figure 149 on page 638). Within the active zone set, the user can select to view by all members, members that are logged in, or members that are not logged in. View of the Active Zone Set and alerts are seen in the second tab integrated in the one-stop screen, as well.

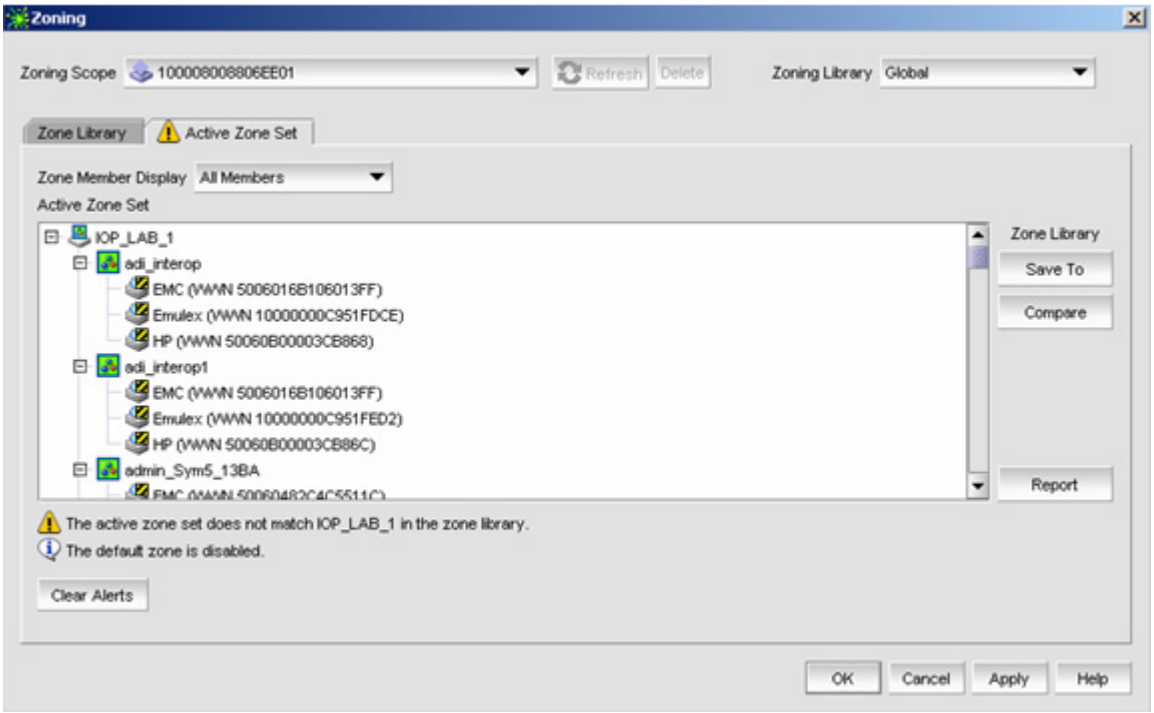


Figure 149 **Active Zone Set tab**

For more information please refer to *Connectrix Manager User Guide* found on Powerlink.

Alerting

Problem notification is an integral part of any fabric tool. Administrators need to know immediately when there are problems or issues within their environments. Notification is one component of alerting, but the ability to set thresholds for performance issues is also important.

The following examples are used:

- "EMC ControlCenter example" on page 639
- "Brocade Fabric Manager example" on page 641
- "Cisco Fabric Manager example" on page 642
- "Connectrix Manager example" on page 644

EMC ControlCenter example

ControlCenter provides numerous metrics to help monitor various aspects of your distributed storage environment, such as:

- Space availability of volumes or disks
- Size of files
- Performance of hosts and storage systems
- Status of backup operations
- Availability and status of ControlCenter components

You configure these metrics into alert definitions that tell ControlCenter how you want to be alerted when a resource meets a value that you define. Many alert definitions are enabled by default when you install ControlCenter agents.

To view alerts in the ControlCenter Console an object (such as a storage system or host) has an active alert, an icon indicating the severity of the condition appears next to the object:

In the tree panel, a small downward arrow on a folder indicates that an object within the folder has an active alert:

ControlCenter provides several methods for monitoring these metrics:

- All Alerts button

  Shows you the total number of new alerts, the level of the highest severity alert, and the total number of alerts at that severity. Click the All Alerts button to display the Active Alerts view.

- Active Alerts view
  - In Table Mode, provides a color-coded table showing triggered alerts for the selected object(s) or all alerts, and tools for tracking alerts to resolution.
  - In Chart Mode, displays charts of triggered alerts for specific storage systems, hosts, or other objects or groups of objects.

- At A Glance view

  Allows you to categorize metrics into charts, such as Host Capacity and Storage System Performance, and view metrics based on groups of systems, such as an application or business-unit group. The At A Glance view is the view displayed by default when you start the Console.

- Alert History view

  Shows a history of all triggered alerts that have since been resolved or cleared. This enables you to trace the state and severity transitions.

In any of these views, you can filter the view to reduce the number of alerts shown, you can switch between a chart or table format, and you can sort or rearrange columns. Figure 150, next, and Figure 151 on page 641 show images of some of the alert views within EMC ControlCenter.

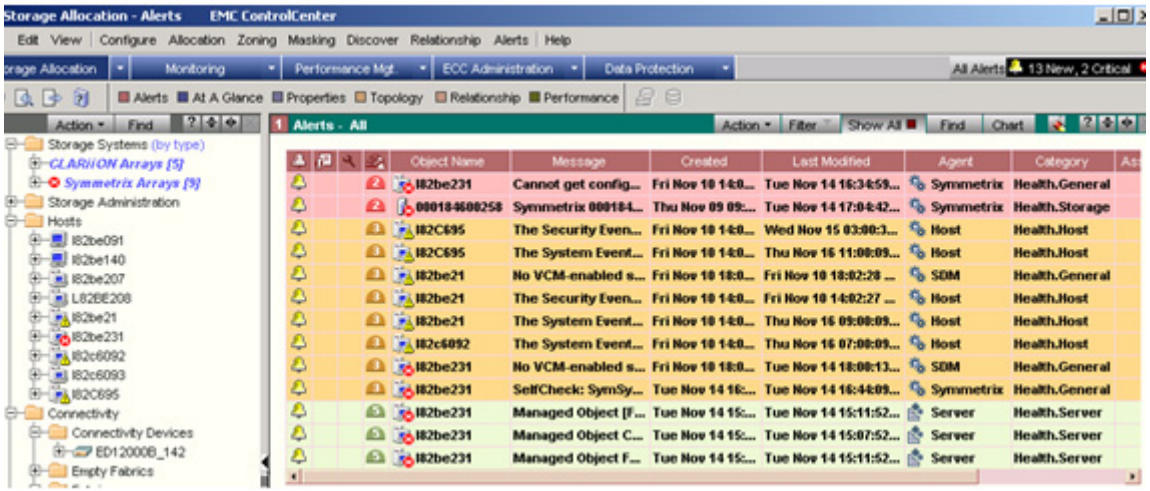


Figure 150 Storage allocation alerts

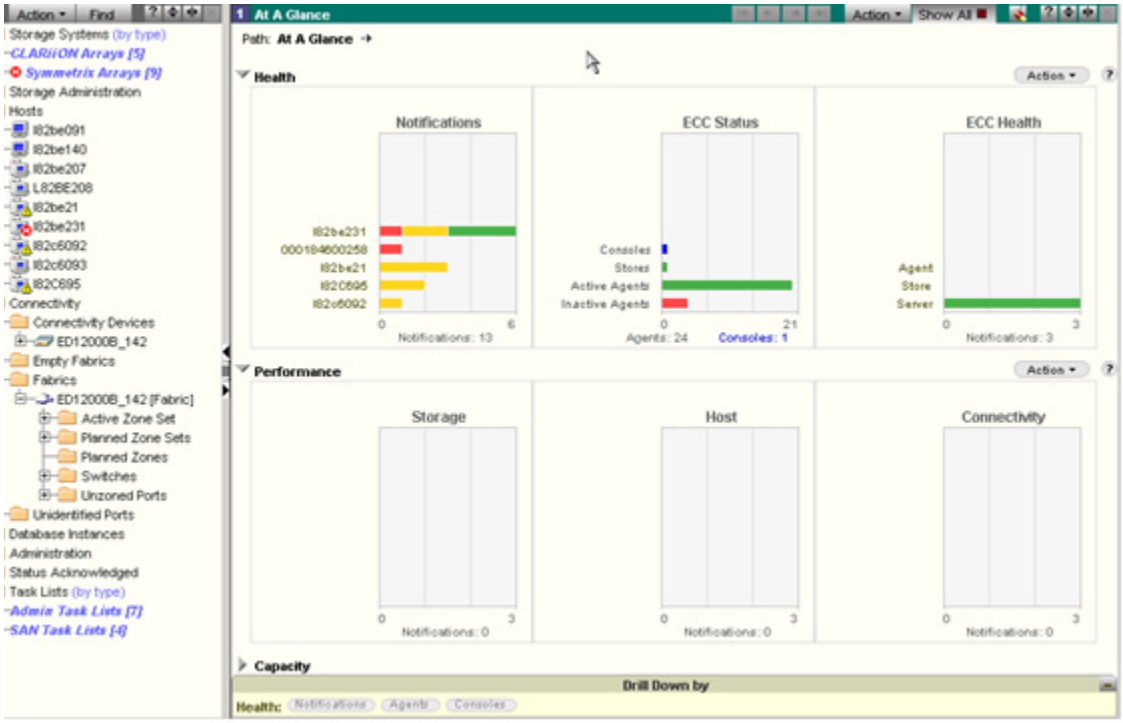


Figure 151 **At A Glance view**

Brocade Fabric Manager example

Alerts are Fabric Manager-specific events that are stored in the Fabric Manager database (see Figure 152 on page 642). For example, if a switch is rebooting while data collection is being performed, an alert appears in the Fabric Manager **Alerts** view showing the switch as *unreachable*. When that same switch has finished booting and data collection occurs again, the **Alerts** table is queried and the switch status changes to *resolved*.

All alerts, regardless of status, remain in the **Alerts** view until deleted. Unacknowledged and unresolved alert information is displayed in bold type by default in the **Alerts** view. When a fabric is deleted from Fabric Manager, any alerts generated by the fabric or its member switches are also deleted from the Fabric Manager database. Fabrics and switches with associated alerts have an alerts icon

displayed next to the fabric or switch icon in the **SAN Elements** tab and the **At A Glance** window for that element.

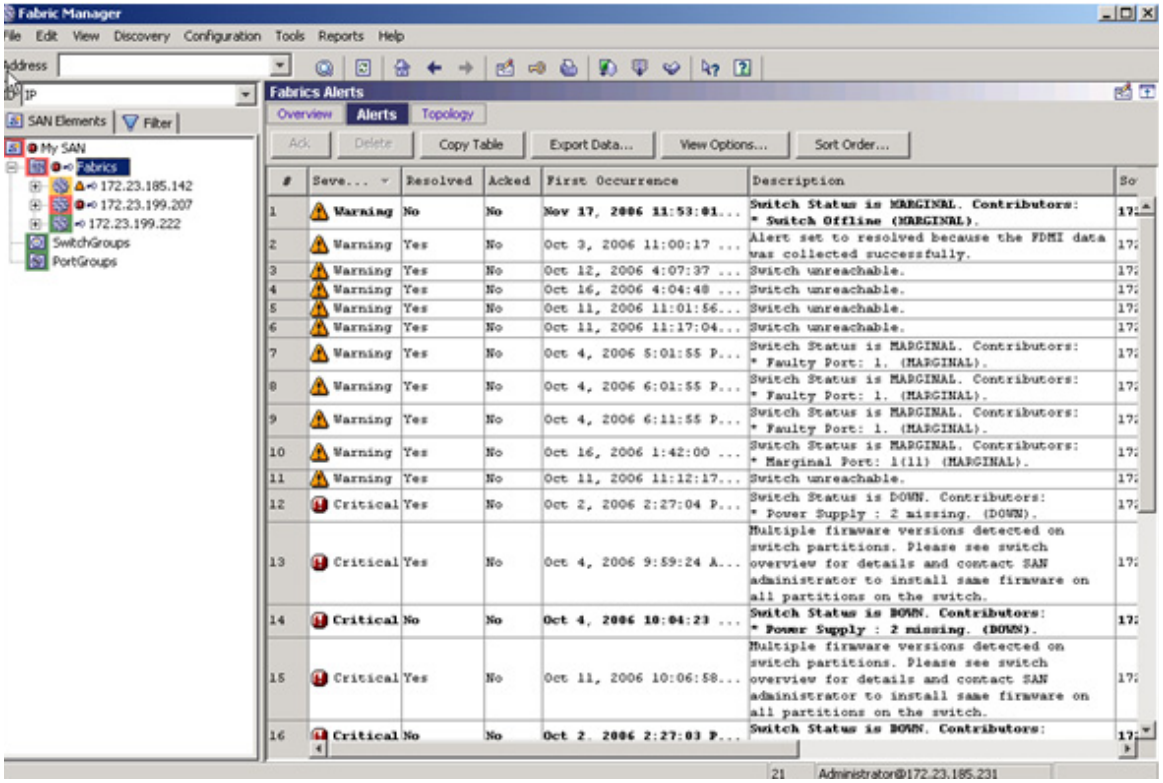


Figure 152 **Fabric Manager Alerts tab**

For more information about alerts and alert types, please refer to the *Brocade Fabric Manager Administration Guide* available on Powerlink.

Cisco Fabric Manager example

Similar to other fabric tools, Cisco's Fabric manager provides real-time alert capabilities and logs the information into log files for tracking. If a switch has an issue, the state of the switch changes in the topology view. When you select the switch you launch the Device Manager which displays a picture of the switch in question, allowing you to qiuckly check the switch for physical issues. For more information about alerting, refer to Cisco's website.

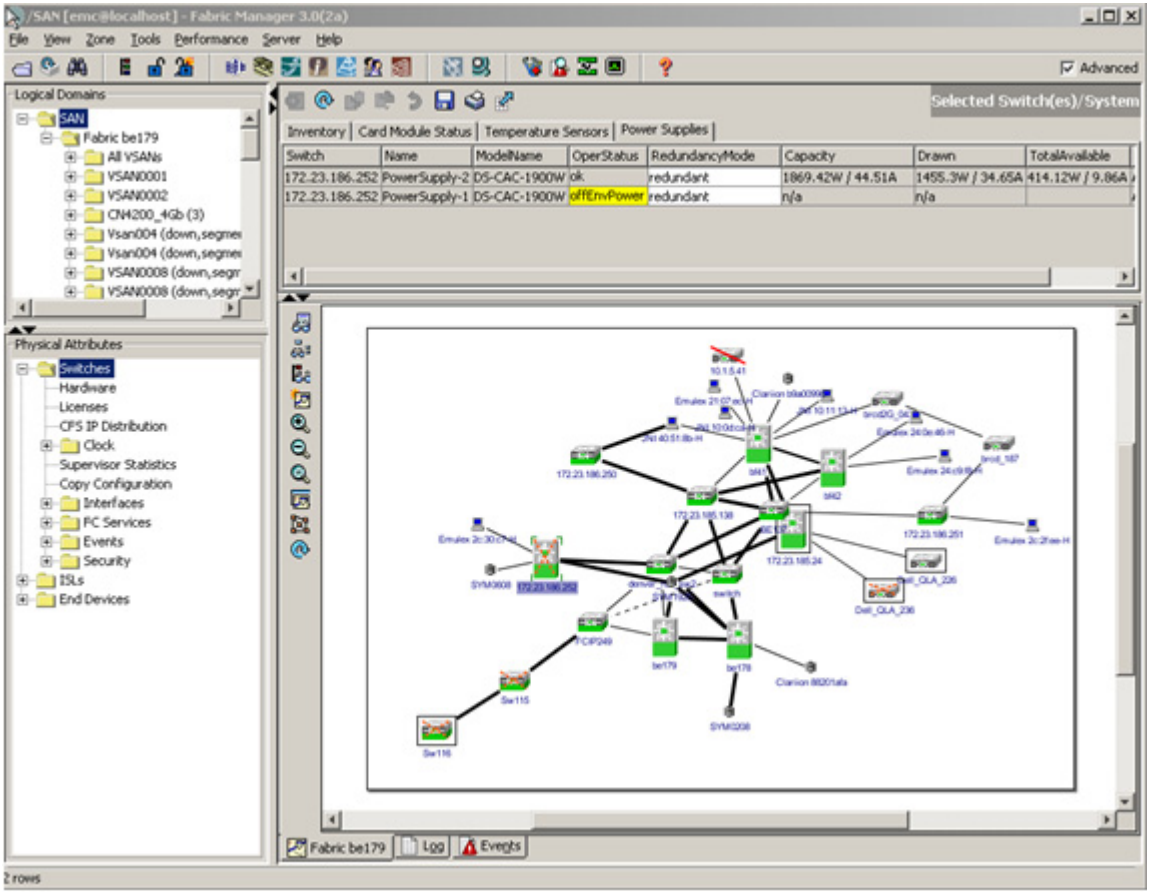


Figure 153 **Fabric Manager, Select switch**

In Figure 153, notice the highlighted switch has an issue. By selecting the switch, you can quickly check the status and then drill further down by launching the Device Manager.

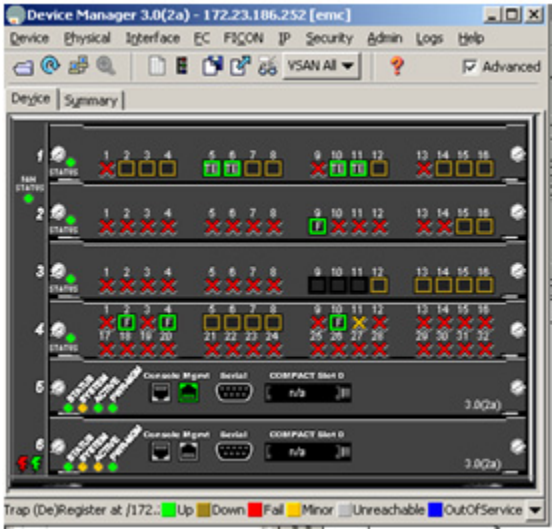


Figure 154 **Device Manager**

Figure 154 shows that the switch has a faulty power supply.

Connectrix Manager example

The main view from Connectrix Manager shows current alerts, and updates or refreshes any new alerts that occur. You can choose to generate emails or notifications when alerts occur. When an alert does occur, you can drill down to the offending component to get more details as well as examine log files to determine root cause. As seen in Figure 155 on page 645, the main view shows an overall status of the switches being monitored. By selecting a switch which has an alert, and then starting the element manager for the switch, you can drill further down to find the component which is responsible for generating the alert. Under the pull-down menu, notice you have the ability to setup SNMP, so traps generated by an alert can be sent to an Enterprise tool, or monitoring type tool , which can translate the traps.

For more information, please review the *Connectrix Manager User Guide* found on Powerlink.

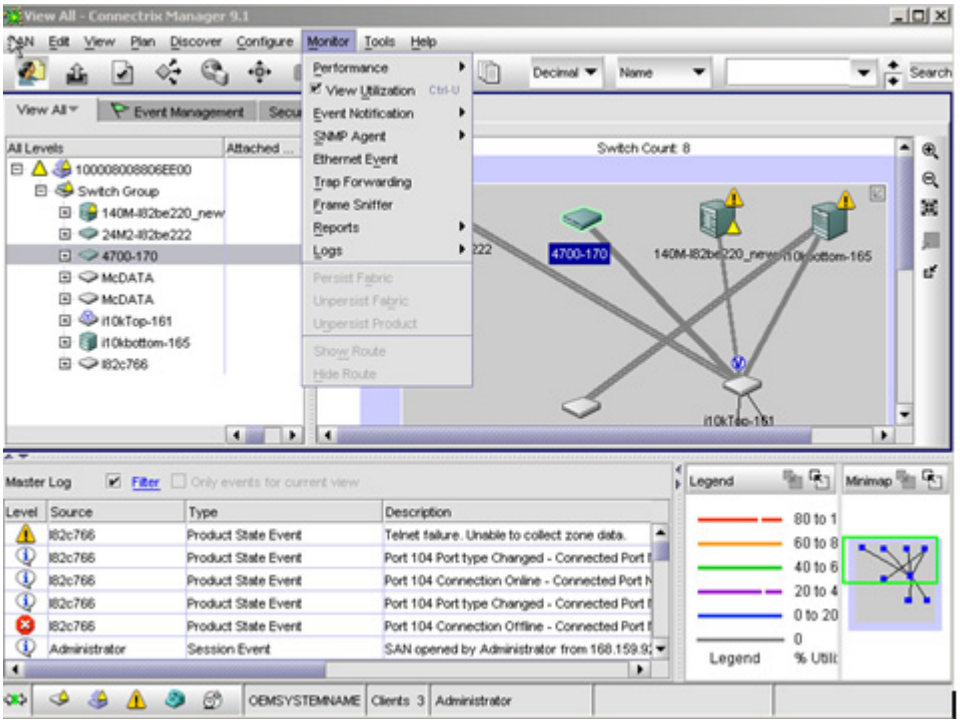


Figure 155 **View All**

Monitoring

It is essential to be able to monitor your environment. The ability to take a quick glance at your environment and see potential problems, or be aware of breakdowns as they happen, is a key element in a fabric tool. Some tools allow various modifications to tailor your environment. Almost all tools today have the ability to display a main view which allows for a quick check of your fabrics. Monitoring is not limited to just alerts or status, but should also provide an ability to follow the performance of your fabric. Alerting capabilities of the different fabric tools have been previously discussed. This section will briefly outline some performance monitoring tools.

The following examples are used:

- "EMC ControlCenter example" on page 646

- "Brocade Fabric Manager example" on page 646
- "Cisco Fabric Manager example" on page 648
- "Connectrix Manager example" on page 649

EMC ControlCenter example

ControlCenter has the ability to provide many different kinds of performance charts and graphs. Many of these are generated from StorageScope, which provides great detail based on Workload Analyzer information gathered during the course of the day. It also provides real-time performance statistics, provided the appropriate data collection policy (DCP) is enabled to gather the information for display.

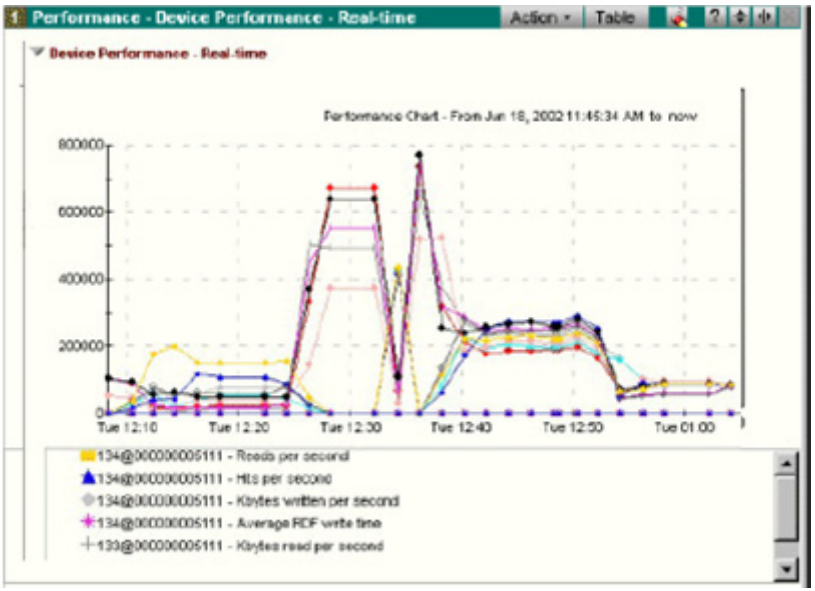


Figure 156 Real-time Device Performance

Although the performance display in Figure 156 is using a device view, similar output is generated by switches.

Brocade Fabric Manager example

As previously discussed, Brocade Fabric Manager cannot only monitor and display alerts, but it can also display fabric events and performance of the switches (see Figure 157 on page 647).

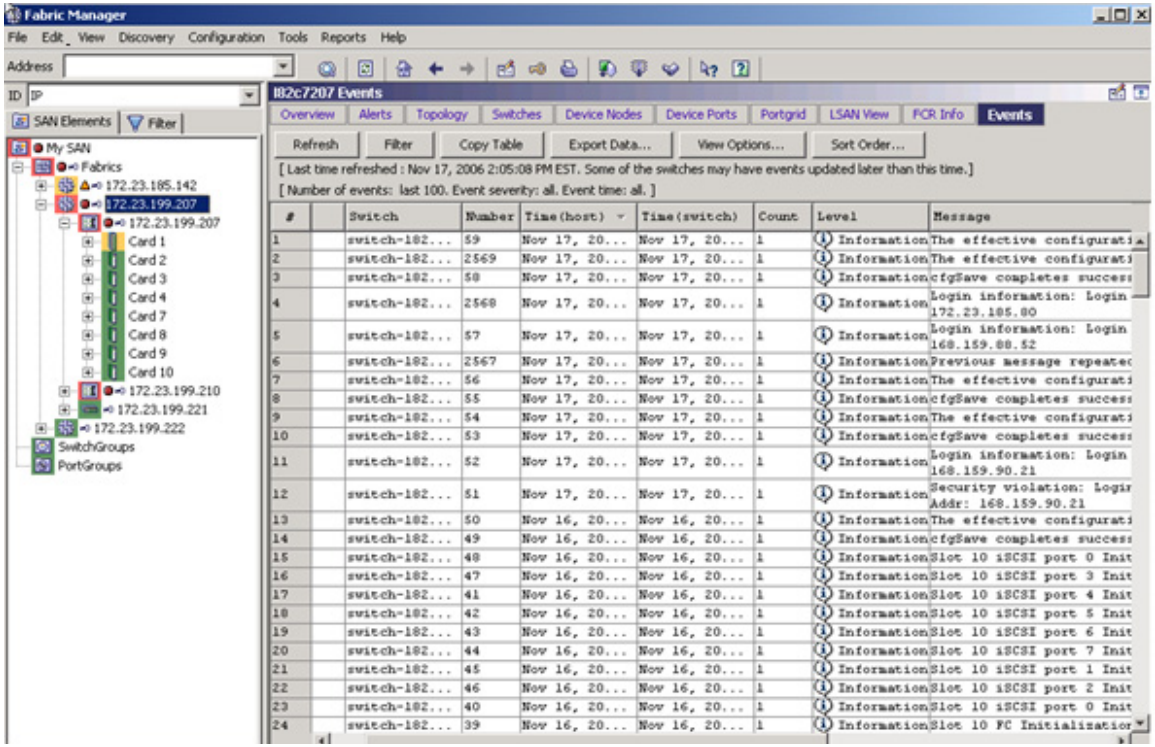


Figure 157 **Fabric Manager Event Monitoring view**

Figure 158 on page 648 shows one type of performance graph which can be generated using Brocade Fabric Manager. For more information on monitoring, please refer to the *Brocade Fabric Manager Administrator's Guide* available on Powerlink.

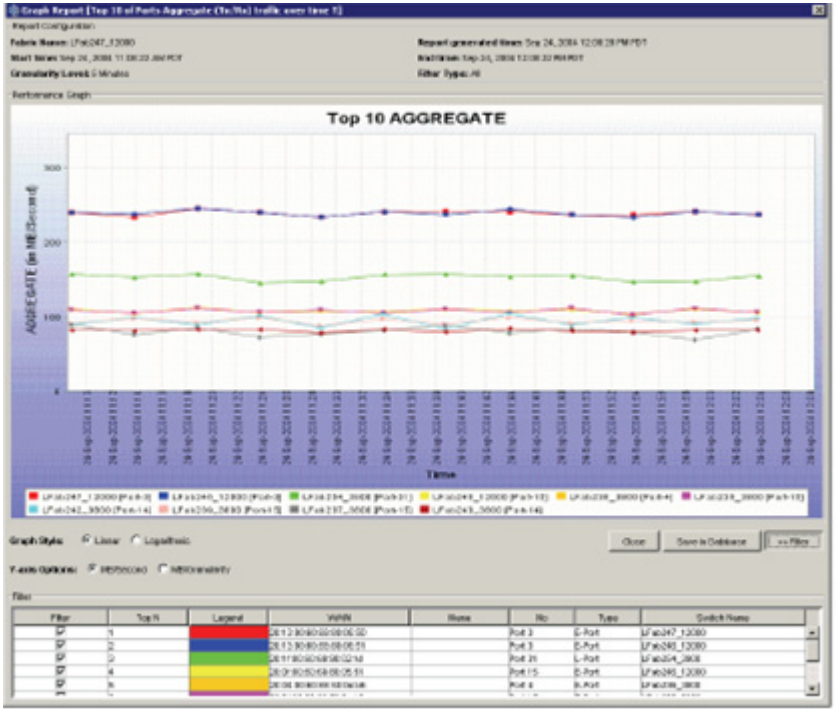


Figure 158 Performance graph example

Cisco Fabric Manager example

As previously mentioned, there is an ability to monitor overall health of your fabric using Cisco's Fabric Manager. There is also an ability to monitor performance real-time. One way to accomplish this is to use Device Manager (see Figure 159 on page 649). You can look at a summary view which looks at all of the modules in the switch and displays the overall performance of each. For a more thorough understanding of performance monitoring, refer to Cisco's website.

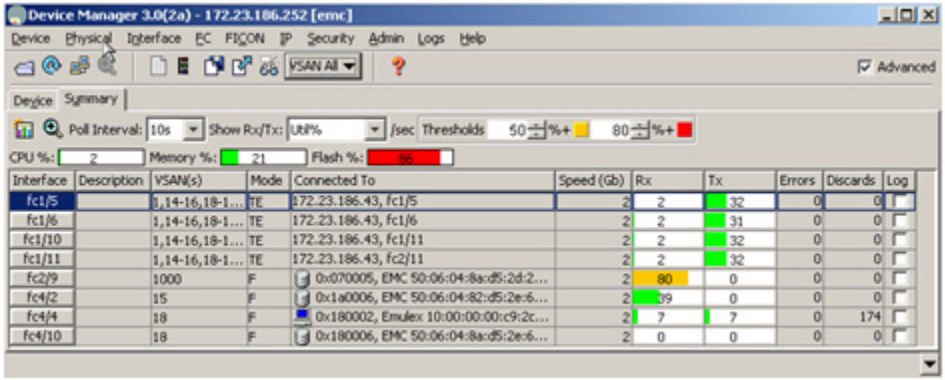


Figure 159 Device Manager

Connectrix Manager example

In addition to the alert monitoring, you also have the ability to monitor performance within or across the fabric (see Figure 160 on page 650). Again, Connectrix Manager's main view can give you an idea of link performance and load. You can also right-click on a switch and select performance graphs to see performance results of the given switch.

In addition, you can go into the element manager for a particular switch and select the **Performance** tab (refer to Figure 161 on page 651, at which point you can monitor all of the ports within the switch. For more information about monitoring, please refer to *Connextrix Manager User's Guide* available on Powerlink.

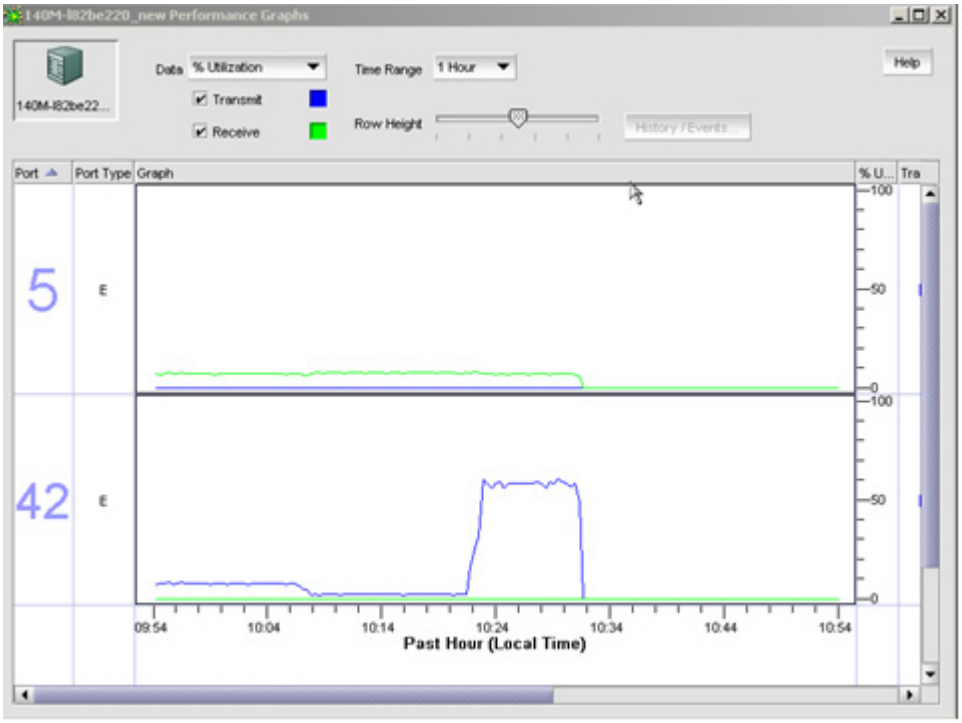


Figure 160 **Performance graph**

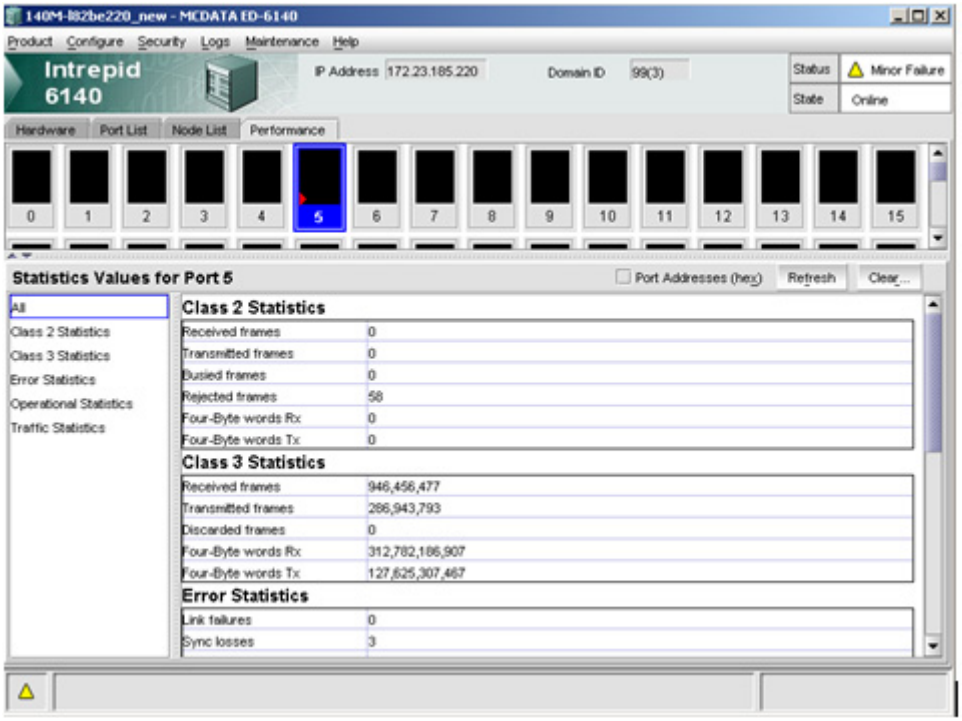


Figure 161 **Performance tab: Statistics values**

8

Storage Virtualization and Replication

This chapter contains the following information on storage virtualzation and replication:

- Invista 654
- RecoverPoint 664
- SAN connectivity 671
- Interswitch link hops 685

Invista

EMC Invista is a high-performance, block storage, SAN-based storage virtualization solution that runs on intelligent SAN switches. Invista's split-path architecture leverages intelligent SAN-switch hardware from EMC's Connectrix partners.

Invista can be deployed in small and enterprise-class organizations. It is flexible enough to virtualize a small storage environment with a few arrays, yet it is resilent and robust enough to virtualize a large datacenter.

This section provides only a brief overview of Invista. To locate more detailed information, refer to "Related Invista documentation" on page 658.

This section provides information on:

- "Benefits of virtualization" on page 654
- "Invista architecture" on page 655
- "Invista components" on page 657
- "Invista disaster recovery" on page 657
- "Intelligent switches" on page 658
- "Logical topology" on page 659

Benefits of virtualization

Block-based virtualization products share a common set of features and benefits.There are numerous benefits to virtualization. Those benefits include (but are *not* limited to) the following:

- Create an open storage environment capable of easily supporting the introduction of new technologies
- Significantly decrease the amount of planned and unplanned downtime
- Increase storage utilization and decrease the amount of "stranded" storage
- Reduce management complexity through single pane of glass management with simplified policies and procedures for multiple storage arrays

- Improve application, storage, and network performance (in optimized configurations) through load balancing and multi-pathing

Invista benefits and virtualization services

Invista supports all of the above. However, there are three key services of Invista when discussing virtualization:

- **Data mobility** — Allows administrators to move primary volumes between heterogeneous storage arrays while the application remains online. This enabler of information lifecycle management allows one to move applications non-disruptively to the appropriate storage tier based on application requirements and service levels.
- **Networked-based volume management** — The basis for what many people think of as *virtualization*. Invista enables one to create and configure virtual volumes from a heterogeneous storage pool and present them out to hosts. It makes sense for the network to be the control point for this—abstracting and aggregating the backend storage, configuring it, and making it available to all of the connected hosts.
- **Heterogeneous cloning** — Allows one to extend the use of clones to areas where their use may have previously been impossible, due to compatibility issues—for example, customers who have implemented an ILM (Infrastucture Lifecycle Management) strategy can create a clone Tier 1 storage array and extend it to a Tier 3, lower-cost storage array.

  Tiered storage is the assignment of different categories of data to different types of storage arrays in order to reduce total storage cost. Categories may be based on levels of protection, performance, usage, and more.

Invista architecture

Invista uses *split-path* architecture to supply virtualization services to a host. Split-path combines intelligent Fibre Channel switches (using purpose-built ASICs) and EMC Invista software to separate the *control* path from the *data* path. Key storage services and virtualization services are performed at the network level. (The *network* in this case refers to the Fibre Channel SAN and not the IP LAN / MAN / WAN.) With IO processing occurring in the intelligent FC switch at the network level, latency associated to “in-band”

solutions is non-existent. By using purpose-built switches with port-level processing (ASICs), the Invista architecture delivers wire-speed performance with negligible latency to meet the demands of the most intensive, random I/O applications delivering consistently scalable performance. Latency introduced by the ASIC does not impact IO on host or storage.

The intelligent switch handles the vast majority of IO traffic passing from the host to the physical storage. All virtualized IO traffic is routed through the intelligent switch. The routing of traffic does not introduce protocol overhead. That is, there is no internal ISL (interswitch link). There are no additional ULPs (upper layer protocols) associated with virtualization. All virtualized IO is routed through ASICs on the intelligent switch or blade. Only the FC frame headers are cracked and re-mapped. The payload (SCSI data frame) is not touched.

The intelligent switch running the Invista software redirects IO to the appropriate physical storage device based on the mapping operation (association of virtual volume to physical volumes) which is created by the administrator when virtual volumes are created. Well over 90% of the IO is handled by the intelligent switch. The CPC handles control path functions. See Figure 162.

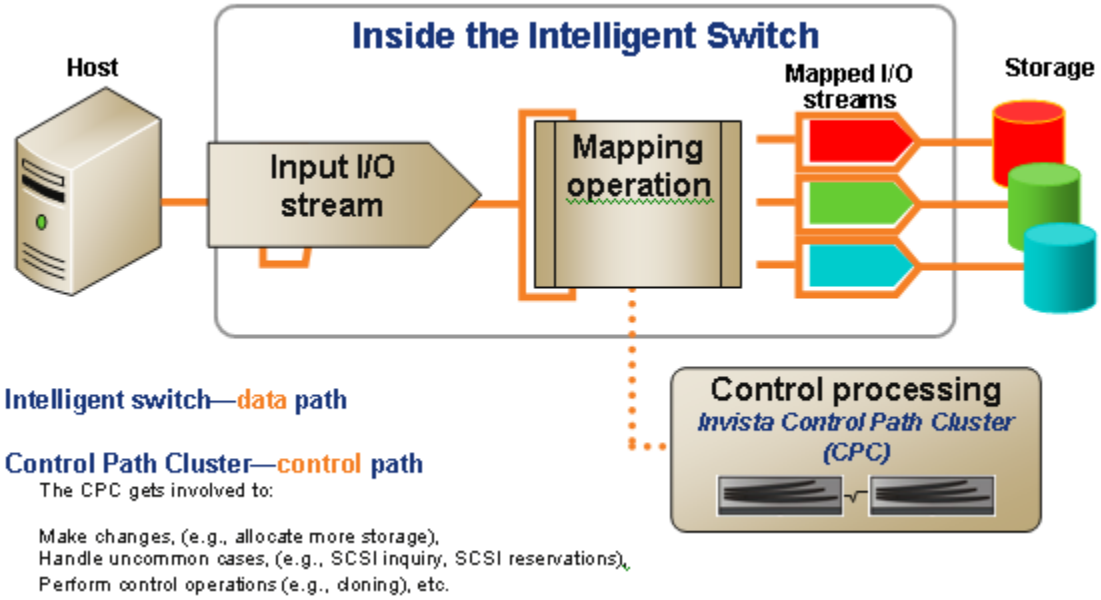


Figure 162 Control path functions

Invista components

There are four major components to an Invista *instance*:

- **Control Path Cluster (CPC) —** The CPC is a dual-node Intel architecture cluster that runs EMC's storage application and manages the virtualization metadata (mapping tables). The CPC supports active-active failover between nodes. Because Invista is a true active/active design, performance is not impacted on a failover. The metadata is stored on a non-virtualized Symmetrix or CLARiiON LUN on the SAN. The CPCs can span two racks spaced 300m (1000 ft) apart.
- **Data Path Controller (DPC) —** This is the intelligent-switch hardware that contains specialized port-level processors (ASICs) to perform virtualization operations on IO at line speed. The DPC is available from two vendors: Brocade and Cisco. Brocade's intelligent switch is the Connectrix B AP-7420B which can be connected through an ISL to a new or existing SAN.

  Cisco's intelligent blade is the Connectrix MDS Storage Services Module (SSM) which can be installed in a Connectrix MDS 9513, 9509, 9506, 9216i, or 9216A. Invista can be configured using two or four DPCs that span at least two fabrics. This creates an additional level of redundancy.
- **Invista software —** This runs on the CPC, communicates with the DPC, and provides application functionality.
- **IP routers —** These IP devices support the communication between the CPC and the DPC. The vendor is Allied Telesis. They provide a firewall to isolate the control path communication from a customer's public network while at the same time providing management access to the Invista instance's hardware components. The routers are sold in pairs to provide redundancy on the LAN. Using VRRP (virtual redundancy router protocol), only one router is active on the network at any one time to prevent IP loops from forming.

Invista disaster recovery

EMC RecoverPoint provides disaster recovery for EMC Invista.

For more information, review the Invista whitepaper located at http://powerlink.emc.com.

Related Invista documentation

There are a number of release notes and documents for more specific Invista-related information that can be found at http://powerlink.emc.com, Home > Support > Documentation/White Paper Library > Software> E-I > Invista, including the following:

- Release notes:

  Information in the release notes include:

  - Support parameters
  - Supported configurations
  - New features and functions
  - SAN compatibility
  - Bug fixes
  - Expected behaviors
  - Technical notes
  - Documentation issues
  - Upgrade information
- *Element Manager Administrator's Guide*
- *Element Manager Command Line Interface*
- Invista Product Security Default
- Various Invista-related whitepapers
- *EMC Support Matrix*

  http://elabnavigator.emc.com, PDFs and Guides>Invista

Intelligent switches

There are two intelligent switches that have purpose-built ASICs and provide port-level virtulation IO processing for the Invista instance:

- Connectrix AP-7420B
- Connectrix MDS Storage Services Module (SSM)

For more information, refer to the product date information in Appendix B, "EMC Connectrix Products Data":

- "Connectrix AP-7420B" on page 989
- "Connectrix MDS Storage Service Module (SSM)" on page 1068

For more information on specifications and performance on the Cisco SSM *with Invista*, go to:

http://www.emc.com/products/software/invista/pdf/H1615_EMC_Invista_SS_ldv.pdf

Note: The SSM module will not support multiple advanced services. That is, it will not support Fibre Channel Write Acceleration functions, SANtap, and other features while supporting Invista.

Logical topology

Invista requires the intelligent switch to physically connect to a SAN. However, it also logically incorporates itself into the SAN.

There are *Virtual Initiators* (VI) and *Virtual Targets* (VT). A host (initiator) or storage (target) is physically connected to the fabric. The VI and VT are logical entities that are accessible from any port on the chassis or in the fabric. See Figure 163 on page 659.

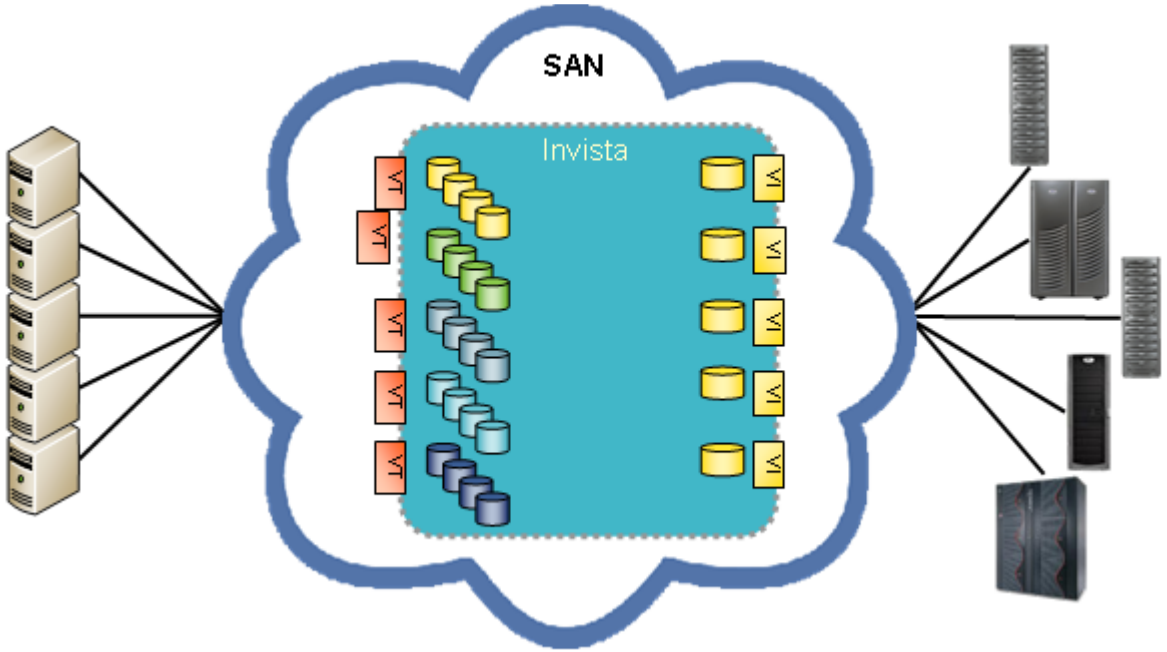


Figure 163 Logical topology

On examination of the Fibre Channel nameserver, a list of Invista VIs and VTs are visible. Figure 164 on page 661 has a list of WWNs of the 32 VTs from a single SSM blade. Note the EMC-specific WWN. These are unique to Invista.

In this case, the customer's host would be added to VSAN 250 and then zoned to an appropriate VT. This is known as the frontend VSAN.

```
LabSwitchTop# show fcns database vsan 250
VSAN 250:
--------------------------------------------------------------------------------
FCID      TYPE  PWWN                    (VENDOR)      FC4-TYPE:FEATURE
--------------------------------------------------------------------------------
0x9a0020  N     50:00:14:41:00:10:16:52 (EMC)         scsi-fcp:target vi..
0x9a0021  N     50:00:14:41:00:10:16:53 (EMC)         scsi-fcp:target vi..
0x9a0022  N     50:00:14:41:00:10:16:54 (EMC)         scsi-fcp:target vi..
0x9a0023  N     50:00:14:41:00:10:16:55 (EMC)         scsi-fcp:target vi..
0x9a0024  N     50:00:14:41:00:10:16:56 (EMC)         scsi-fcp:target vi..
0x9a0025  N     50:00:14:41:00:10:16:57 (EMC)         scsi-fcp:target vi..
0x9a0026  N     50:00:14:41:00:10:16:58 (EMC)         scsi-fcp:target vi..
0x9a0027  N     50:00:14:41:00:10:16:59 (EMC)         scsi-fcp:target vi..
0x9a0028  N     50:00:14:41:00:10:16:5a (EMC)         scsi-fcp:target vi..
0x9a0029  N     50:00:14:41:00:10:16:5b (EMC)         scsi-fcp:target vi..
0x9a002a  N     50:00:14:41:00:10:16:5c (EMC)         scsi-fcp:target vi..
0x9a002b  N     50:00:14:41:00:10:16:5d (EMC)         scsi-fcp:target vi..
0x9a002c  N     50:00:14:41:00:10:16:5e (EMC)         scsi-fcp:target vi..
0x9a002d  N     50:00:14:41:00:10:16:5f (EMC)         scsi-fcp:target vi..
0x9a002e  N     50:00:14:41:00:10:16:60 (EMC)         scsi-fcp:target vi..
0x9a002f  N     50:00:14:41:00:10:16:61 (EMC)         scsi-fcp:target vi..
0x9a0030  N     50:00:14:41:00:10:16:62 (EMC)         scsi-fcp:target vi..
0x9a0031  N     50:00:14:41:00:10:16:63 (EMC)         scsi-fcp:target vi..
0x9a0032  N     50:00:14:41:00:10:16:64 (EMC)         scsi-fcp:target vi..
0x9a0033  N     50:00:14:41:00:10:16:65 (EMC)         scsi-fcp:target vi..
0x9a0034  N     50:00:14:41:00:10:16:66 (EMC)         scsi-fcp:target vi..
0x9a0035  N     50:00:14:41:00:10:16:67 (EMC)         scsi-fcp:target vi..
0x9a0036  N     50:00:14:41:00:10:16:68 (EMC)         scsi-fcp:target vi..
0x9a0037  N     50:00:14:41:00:10:16:69 (EMC)         scsi-fcp:target vi..
0x9a0038  N     50:00:14:41:00:10:16:6a (EMC)         scsi-fcp:target vi..
0x9a0039  N     50:00:14:41:00:10:16:6b (EMC)         scsi-fcp:target vi..
0x9a003a  N     50:00:14:41:00:10:16:6c (EMC)         scsi-fcp:target vi..
0x9a003b  N     50:00:14:41:00:10:16:6d (EMC)         scsi-fcp:target vi..
0x9a003c  N     50:00:14:41:00:10:16:6e (EMC)         scsi-fcp:target vi..
0x9a003d  N     50:00:14:41:00:10:16:6f (EMC)         scsi-fcp:target vi..
0x9a003e  N     50:00:14:41:00:10:16:70 (EMC)         scsi-fcp:target vi..
0x9a003f  N     50:00:14:41:00:10:16:71 (EMC)         scsi-fcp:target vi..
```

Figure 164 **List of WWNs**

Figure 165 has a list of WWNs of the 9 VIs from a single SSM blade. Note the EMC-specific WWN. These are unique to Invista.

In this case, the customer's physical storage device would be added to VSAN 101. All VIs are zoned to all storage ports that will present LUNs to the Invista instance. This is known as the backend VSAN.

```
LabSwitchTop# show fcns database vsan 101

VSAN 101:
--------------------------------------------------------------
FCID      TYPE  PWWN                (VENDOR)     FC4-TYPE:FEATURE
--------------------------------------------------------------
0x800011  N     50:00:14:41:00:10:16:bb (EMC)     scsi-fcp:init vir..t
0x800012  N     50:00:14:41:00:10:16:bc (EMC)     scsi-fcp:init vir..t
0x800013  N     50:00:14:41:00:10:16:bd (EMC)     scsi-fcp:init vir..t
0x800014  N     50:00:14:41:00:10:16:be (EMC)     scsi-fcp:init vir..t
0x800015  N     50:00:14:41:00:10:16:bf (EMC)     scsi-fcp:init vir..t
0x800016  N     50:00:14:41:00:10:16:c0 (EMC)     scsi-fcp:init vir..t
0x800017  N     50:00:14:41:00:10:16:c1 (EMC)     scsi-fcp:init vir..t
0x800018  N     50:00:14:41:00:10:16:c2 (EMC)     scsi-fcp:init vir..t
0x800019  N     50:00:14:41:00:10:16:c3 (EMC)     scsi-fcp:init vir..t
```

Figure 165 WWNs of 9 VIs from a single SSM blade

A virtual target has its own unique World Wide Name (WWN). Groups of *virtual volumes* are associated with a virtual target for presentation to the host. The virtual volumes' hosts see standard SCSI-3 compliant volumes with their own unique Invista signature. In all other respects, virtual volumes look just like volumes from an array. They support persistent reservation for clustering and they can be either managed with a host-based volume manager or LUN-masked.

Physical storage devices are zoned to a logical construct called a *virtual initiator* (VI). A virtual initiator has its own unique World Wide Name (WWN). To the physical storage device, the Invista VI appears as a host.

From the host side, or frontend, Invista looks just like an array. After qualification, any host that can work with SCSI disks over Fibre Channel can work with Invista. Additional host drivers or other

specialized software are *not* required. Hosts connect to the intelligent switch (also called the DPC, or Virtualizer) via Fibre Channel through standard F_Ports. Optionally, as shown in Figure 165 on page 662, the connection can be through a “Layer 2” or non-intelligent SAN for greater host connectivity.

From the array side, or backend, Invista looks like a host initiator. As is the case for the front end, arrays are connected via Fibre Channel and communicate with Invista by SCSI commands. Optionally, Figure 165 on page 662, connection can be through a “Layer 2” SAN for fan-out of connectivity.

Traffic flow

Invista can be integrated into a new or existing SAN. Depending on the needs of the application, the administrator decides whether a host will access virtualized storage, non-virtualized (physical) storage, or both kinds of storage. As described in later sections, the host/storage placement and the location of the intelligent blade or switch need consideration in the overall design of a Fibre Channel switched fabric. All virtualized traffic must pass through the virtualization ASICs on the Cisco Storage Services Module or the Connectrix B AP-7420B. Non-virtualized storage IO will not pass through the intelligent blade or switch.

RecoverPoint

EMC RecoverPoint is a SAN-based replication solution for data that host applications are writing over Fibre Channel to locally attached SAN volumes. RecoverPoint incorporates an out of band hardware appliance as well a data splitter. The data splitter can run either on intelligent SAN switches or reside on the host. The RecoverPoint system is flexible and scalable enough to protect data in both small storage environments with a few arrays and in large scale datacenters.

RecoverPoint's fabric splitter leverages intelligent SAN-switch hardware from EMC's Connectrix partners. This document focuses on topology considerations where the data splitter resides on an intelligent SAN switch.

This section provides only a brief overview of RecoverPoint. To locate more detailed information, refer to "Related RecoverPoint documentation" on page 667.

Benefits of virtualization

Block-based virtualization products share a common set of features and benefits.There are numerous benefits to virtualization. Those benefits include (but *not* limited to) the following:

- Create an open storage environment capable of easily supporting the introduction of new technologies
- Significantly decrease the amount of planned and unplanned downtime
- Increase storage utilization and decrease the amount of "stranded" storage
- Reduce management complexity through single pane of glass management with simplified policies and procedures for multiple storage arrays
- Improve application, storage, and network performance (in optimized configurations) through load balancing and multi-pathing

RecoverPoint benefits and replication services

RecoverPoint provides the following key services in a data center:

- **Intelligent fabric splitting —** Allows data to be split and replicated in a manner that is transparent to the host. By having the switch intercept host writes and duplicate, or split, those writes to the RecoverPoint appliance allows for data replication with minimal impact to the host.
- **Heterogeneous replication of data volumes —** RecoverPoint is a truly heterogeneous replication solution, enabling the replication of data between different storage array volumes. This enables the ability to migratefrom a virtualized environment to a non-virtualized environment and/or replication between different tiers of storage,thus allowing greater flexibility in the data center.
- **Continous data protection —** Enables a fine granularity of recovery to a block level and refers to local replication of SAN volumes.
- **Continuous remote replication via IP —** Allows data replication across completely separate fabrics over distances limited only by IP technology. The RecoverPoint appliances transfer data from one appliance to another via TCP/IP, greatly extending the distance between source and target volumes.
- **Compression and bandwidth reduction —** By implementing block-level compression algorithms before sending traffic over the IP network, RecoverPoint significantly reduces the network cost required for matching the RPO/RTO expectations.

RecoverPoint architecture

RecoverPoint uses a combination of hardware appliances configured out of the direct data path “out of band” and a splitter driver to replicate data from the designated source volume to the configured target volume. RecoverPoint uses intelligent Fibre Channel switches to intercept and “split” incoming writes. The intelligent Fibre Channel switch then sends a copy of the incoming write to the RecoverPoint Appliance as well as passing the original write to the original destination. The RecoverPoint appliance then handles the sending of the “split” write to the proper destination volume, either via TCP/IP in the case of Continuous Remote Replication or locally via Fibre Channel in the case of Continuous Data Protection.

Traffic routed by the intelligent switch doesn't introduce protocol overhead. All virtualized IO is routed through ASICs on the intelligent switch or blade. The payload (SCSI data frame) is not affected by this operation, as the ASIC does not read or modify the content but rather reads only the protocol header information to make routing decisions.

The splitter residing in the intelligent switch is responsible for duplicating the write traffic and sending the additional write to the RecoverPoint appliance.

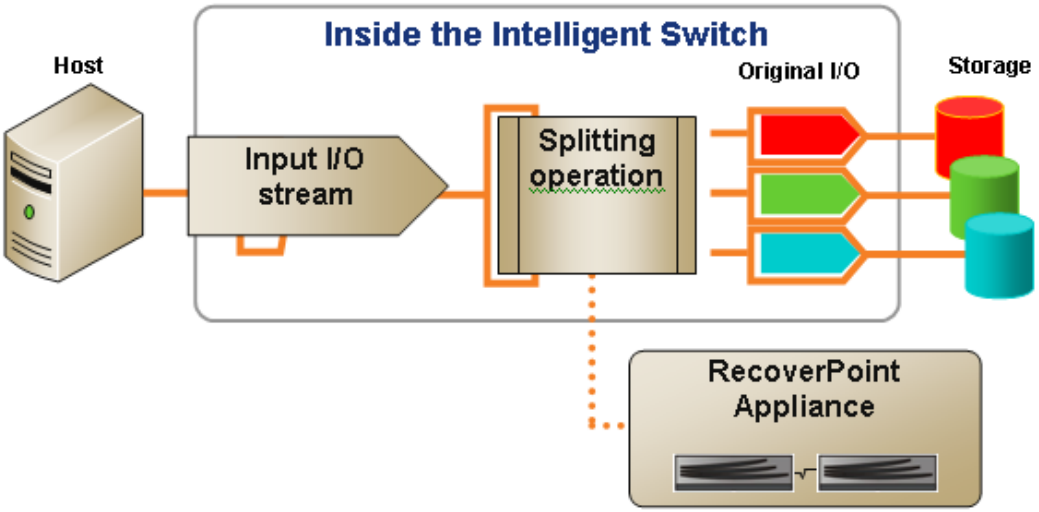


Figure 166 **Intelligent switch running RecoverPoint**

RecoverPoint components

There are four major components to a RecoverPoint installation.

- **Virtual Target (VT)** — This is the virtual representation of the physical storage that is used as either source or destination for replicated data.
- **Virtual Initiator (VI)** — This is the virtual representation of the physical host that is issuing I/O to the source volume.
- **Intelligent Fabric Splitter** — This is the intelligent-switch hardware that contains specialized port-level processors (ASICs) to perform virtualization operations on IO at line speed. This

functionality is available from two vendors: Brocade and Cisco. Brocade's intelligent switch is the AP-7420B which can be linked though ISLs to a new or existing SAN.

Cisco's intelligent blade is the Storage Services Module (SSM) which can be installed in an MDS 9513, 9509, 9506, 9216i, or 9216A.

- **RecoverPoint Appliances (RPA)** — These appliances are linux based boxes that accept the "split" data and route the data to the appropriate destination volume, either via IP or Fibre Channel. The RPA also acts as the sole management interface to the RecoverPoint installation.

Related RecoverPoint documentation

There are a number of documents for RecoverPoint related information, all which can be found at http://powerlink.emc.com, Home > Support > Documentation/White Paper Library > Software P-R> Documentation > RecoverPoint.

- Release notes:

  Information in the release notes include:

  - Support parameters
  - Supported configurations
  - New features and functions
  - SAN compatibility
  - Bug fixes
  - Expected behaviors
  - Technical notes
  - Documentation issues
  - Upgrade information
- RecoverPoint maintenance and administration documents
- RecoverPoint installation and cnofiguration documents
- RecoverPoint-related whitepapers
- *EMC Support Matrix*

  http://elabnavigator.emc.com, Home > Support > Interoperability/ Compatibility/ Configuration > E-Lab Interoperability Navigator>PDFs and Guides> RecoverPoint

Intelligent switches

There are two intelligent switches that provide fabric-level IO splitting for the RecoverPoint instance:

- Connectrix AP-7420B
- Connectrix MDS Storage Services Module (SSM)

For more information on the two intelligent switches, refer to the product date information in Appendix B, "EMC Connectrix Products Data":

- "Connectrix AP-7420B" on page 989
- "Connectrix MDS Storage Service Module (SSM)" on page 1068

Logical topology

There are *Virtual Initiators* (VI) and *Virtual Targets* (VT). A host (initiator) or storage (target) is physically connected to the fabric. The VI and VT are logical entities that are accessible from any port on the chassis or in the fabric. See Figure 167.

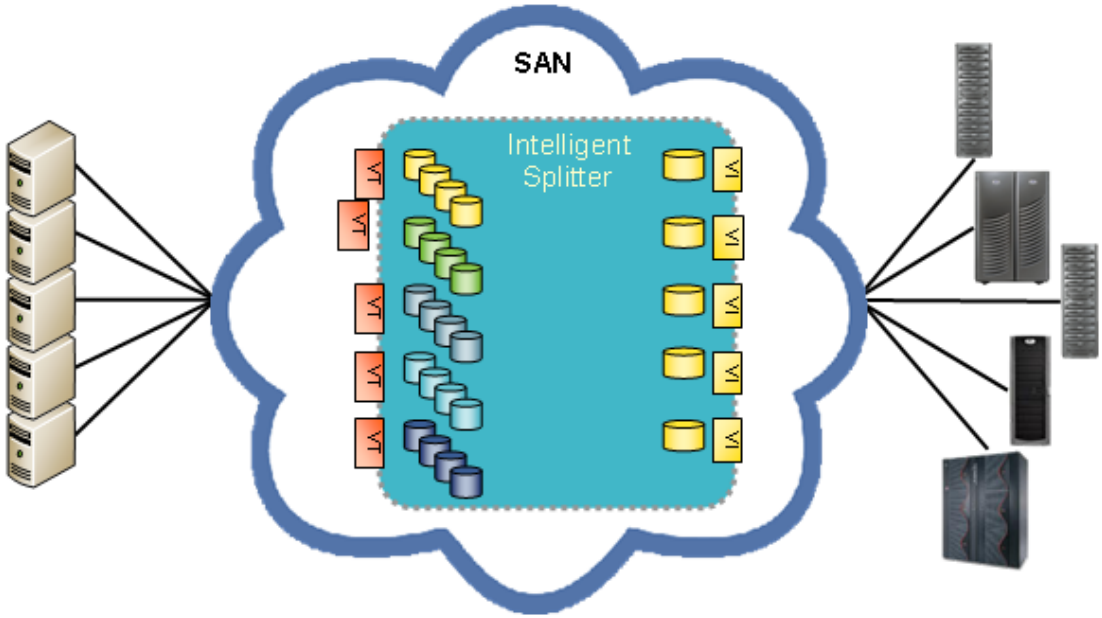


Figure 167 Logical topology

The exact method of this virtualization differs between the Cisco SANTap solution and the Brocade FAP solution.

The SANTap architecture uses VSANs to physically separate the physical host from the physical storage and bridges the VSANs via virtualization internal to the intelligent switch. When a virtual target (DVT) is created, a duplicate WWN is created and placed in the “Front End” VSAN, that is the VSAN that encompasses the physical hosts. The physical hosts then communicate with the virtual target. SANTap then also creates virtual initiators (VI) and duplicates the host WWNs and places them in the “Back End” VSAN, which is the VSAN that encompasses the physical storage. The VIs then communicate to the physical storage. SANTap also creates virtual entries representing the splitter itself and communication between the RPAs and the intelligent switch happens via these virtual entries.

Figure 168 shows example “Back End” and “Front End” VSANs and shows the duplicate WWNs in both VSANs.

```
Example "Back End" VSAN
VSAN 150:
-------------------------------------------------------------------------
FCID        TYPE  PWWN                      (VENDOR)        FC4-TYPE:FEATURE
-------------------------------------------------------------------------
0x930009    N     50:06:04:82:d5:2e:87:52 (EMC)           scsi-fcp:both 253 ← Physical Storage
0x93000b    N     22:35:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x93000c    N     22:36:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x93000d    N     22:37:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x93000e    N     22:38:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x93000f    N     22:39:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x930010    N     22:3a:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x930011    N     22:3b:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x930012    N     22:3c:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x930013    N     22:3e:00:05:30:00:3d:e0 (Cisco)         scsi-fcp:init vir..t
0x930200    N     21:01:00:e0:8b:a6:fd:cc (Qlogic)        scsi-fcp:init vir..t ← Virtual Initiator

Corresponding "Front End" VSAN
VSAN 151:
-------------------------------------------------------------------------
FCID        TYPE  PWWN                      (VENDOR)        FC4-TYPE:FEATURE
-------------------------------------------------------------------------
0x3a0100    N     50:06:04:82:d5:2e:87:52 (EMC)           scsi-fcp:target vi.. ← Virtual Storage (DVT)
0x3a0200    N     21:01:00:e0:8b:a6:fd:cc (Qlogic)        scsi-fcp:init ← Physical Initiator
```

Figure 168 Example Back End and Front End VSAN

The Brocade FAP architecture permanently binds hosts to virtual initiators and channels all host IO through the corresponding bound VI. When a host is bound to a particular VI, a virtual target is created with a new WWN. The separation in this case is done via zoning. The WWN of the bound VI is zoned with the physical storage and the

physical host is zoned with the FAP created VT. Data would flow from the physical host to the virtual target, then from the bound VI to the physical storage.

Figure 169 shows an example of the zoning with the VIs and VTs. The host WWNs are bound to one of the 16 VIs in the other zone.

```
Zone:   Host2Storage
          70:01:24:82:7f:d6:04:d9 ← Virtual Target
          21:00:00:e0:8b:03:2b:b0
          21:00:00:e0:8b:03:e3:b0
Zone:   RPA2Storage
          50:01:24:82:00:99:50:f3
          50:06:01:61:41:e0:04:d9 ← Physical Target
          50:01:24:82:01:b9:50:f3
          50:01:24:82:01:b9:38:f3
          50:01:24:82:00:99:b1:f3
          50:01:24:82:00:99:4c:f3
          50:01:24:82:00:99:38:f3
          22:01:00:05:1e:16:1f:a6 ← 16 Virtual Initiators
          22:0a:00:05:1e:16:1f:a6
          22:09:00:05:1e:16:1f:a6
          22:08:00:05:1e:16:1f:a6
          22:0e:00:05:1e:16:1f:a6
          22:0d:00:05:1e:16:1f:a6
          22:0b:00:05:1e:16:1f:a6
          22:06:00:05:1e:16:1f:a6
          22:02:00:05:1e:16:1f:a6
          22:0c:00:05:1e:16:1f:a6
          22:04:00:05:1e:16:1f:a6
          22:05:00:05:1e:16:1f:a6
          22:0f:00:05:1e:16:1f:a6
          22:03:00:05:1e:16:1f:a6
          22:00:00:05:1e:16:1f:a6
          22:07:00:05:1e:16:1f:a6
```

Figure 169 Zoning with VTs and VIs

SAN connectivity

This section includes SAN connectivity information for both Invista and RecoverPoint.

Invista's and RecoverPoint's use of intelligent switches allows for seamless connectivity into a Fibre Channel switched fabric. Using intelligent switches, a new fabric can be dedicated to Invista or Invista can be integrated into an existing fabric.

There are a number of different SAN topologies as described in this topology guide. They are classified in two categories: simple and complex topologies.

Invista and RecoverPoint can be integrated into the design of either topology type. The next sections will cover each design type and specif ic best practices for Invista solutions:

- "Simple SAN topology" on page 671

  Note: For more detailed information on simple SAN topologies, refer to "Simple Fibre Channel SAN topologies" on page 50.

- "Complex SAN topologies" on page 675

  Note: For more detailed information on complex SAN topologies, refer to "Complex Fibre Channel SAN topologies" on page 146.

Simple SAN topology

Invista is dependent on the intelligent switches to handle the port level processing of the virtualized traffic.

RecoverPoint is dependent on the intelligent switches to handle data splitting to the RecoverPoint Appliance.

The intelligent switches can act as standalone FC switches as well as Layer 2 FC switches. In small SANs, or with small virtualization implementations, this is a one possible topology. There are a few configurations based on the choice in FC switch vendors, each discussed further in this section:

- "Connectrix B AP-7420B simple SAN topology" on page 672
- "Storage Services Module (SSM) blade in a simple SAN topology" on page 672

- "Node placement in a simple SAN topology – Connectrix B" on page 674
- "Node placement in a simple SAN topology – Connectrix MDS" on page 675

Connectrix B AP-7420B simple SAN topology

The Connectrix B AP-7420B, when operating as a platform for the Invista and RecoverPoint instance, supports F_Port connections. That is, FC-SW nodes, such as hosts and storage, can be connected directly to any one of the 16 ports on the AP-7420B (see Figure 170). There is no need to connect through a Layer 2 switch such as an ED-48000B or DS-4900B, for example.

Note: Figure 170 does not show the CPCs or the AT switches. It is only displays the FC connectivity..

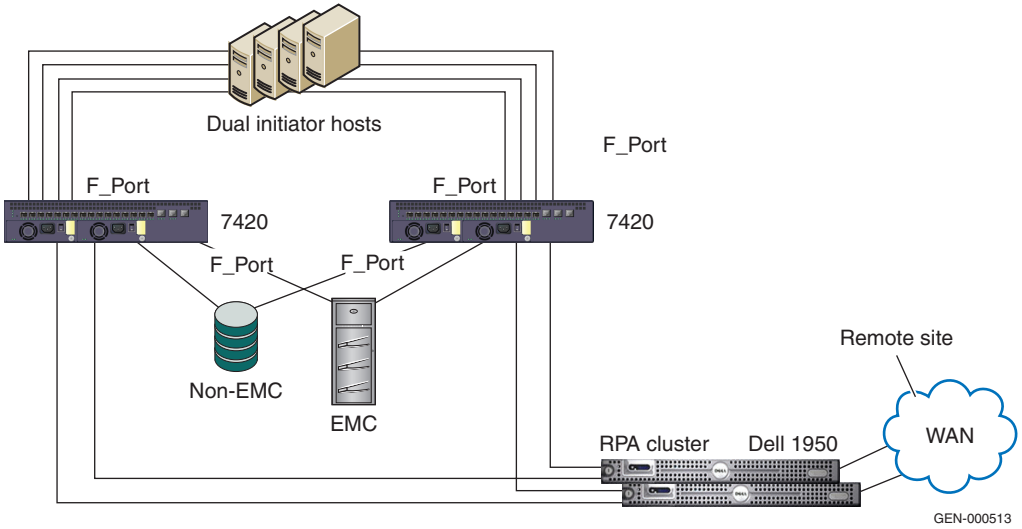


Figure 170 **Connectrix B AP-7420B simple SAN topology**

Storage Services Module (SSM) blade in a simple SAN topology

The Connectrix MDS Storage Services Module (SSM) can be installed in any MDS switch or director with open slots. This includes the MDS 9513, 9509, 9506, and 9216. In Figure 171 on page 673, hosts are connected to the SSM ports and storage is connected to the fixed ports on the MDS 9216i. Hosts may be also be connected to the fixed ports; however, since the SSM ports are oversubscribed, they are best suited for host connectivity.

Since the fixed ports are line rate mode (2 Gb on the MDS 9216), storage ports should be connected there to ensure maximum bandwidth.

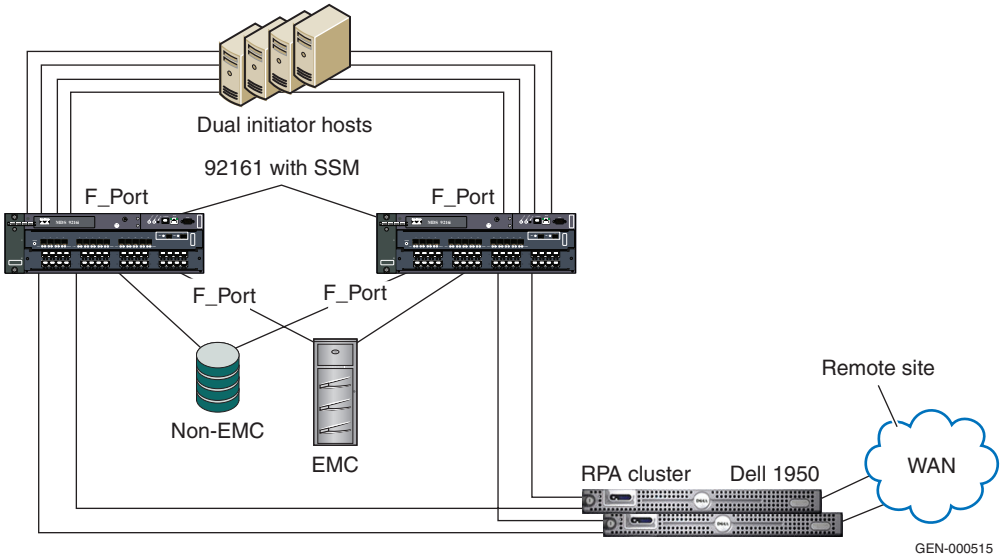


Figure 171 **Connectrix MDS SSM blade in simple SAN topology**

Figure 172 on page 674 shows an MDS 9513 with a 24-port line card and an SSM blade. Hosts are connected to the SSM ports and storage is connected to the 24-port line card. The SSM ports are oversubscribed so they are best suited for host connectivity. The 24-port line card ports are line rate mode (4 Gb), so storage ports should be connected there to ensure maximum bandwidth.

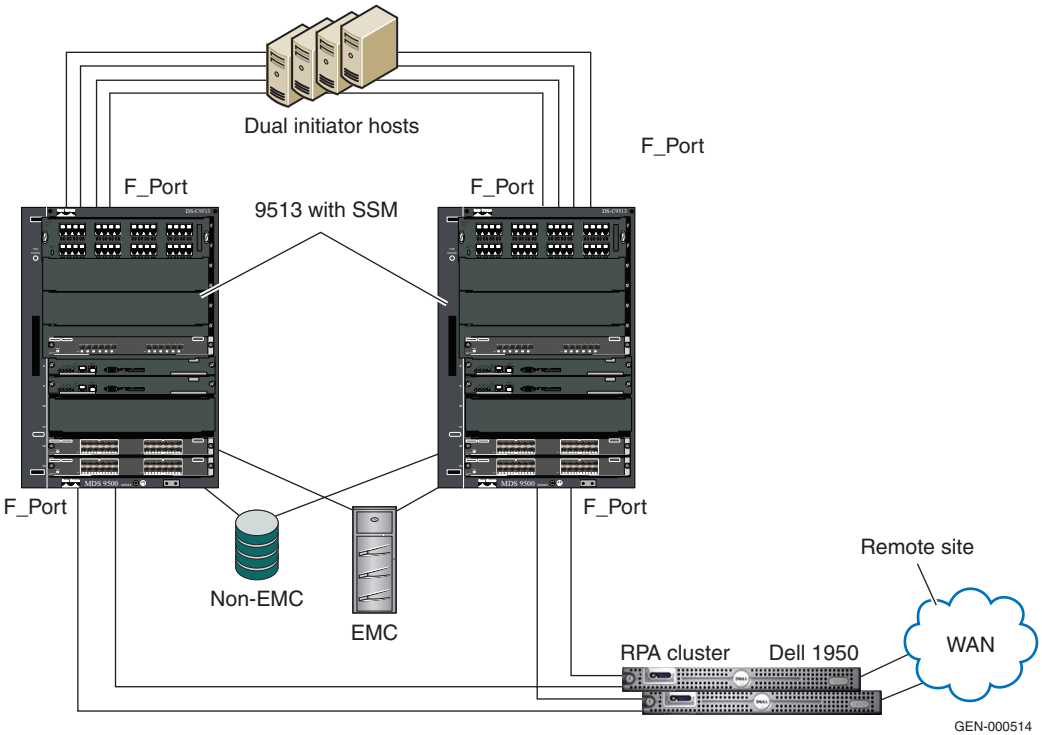


Figure 172 **MDS 9513 with a 24-port line card and an SSM blade**

Note: Figure 171 on page 673 and Figure 172 do not show the CPCs or the AT switches. It is only displays the FC connectivity.

Node placement in a simple SAN topology – Connectrix B

The AP-7420B, when used as a single switch, may have both hosts and storage nodes connected directly to any of the ports. It is unlikely that the AP-7420B will be used in this manner because it does not make effective use of the available ports. Since there are only 16 ports, only a mixture of 16 hosts and storage ports are possible. It is more likely that this intelligent switch will be used with more complex designs.

Node placement in a simple SAN topology – Connectrix MDS

The Connectrix MDS single switch / director model is sometimes known as a collapsed core / edge design. There is a mixture of line rate mode (LRM) and oversubscribed mode ports (OSM). The external ports on an SSM blade are oversubscribed (same ratio as the standard 32-port line card). The Cisco switch / director can support virtualized and non-virtualized storage simultaneously. Hosts may be connected to any port in the chassis. Storage may be connected to any port in the chassis.

Note: Hosts accessing virtualized storage, and physical storage supplying Invista or RecoverPoint, do *not* have to be connected to the SSM.

The same best practice holds true for host and storage connectivity with or without Invista or RecoverPoint. Hosts should be connected to OSM ports whenever possible unless they have high IO throughput requirements. Storage should be connected to LRM ports whenever possible because a storage port is a shared resource with multiple hosts requiring access.

If there are hosts using solely virtualized storage, then they require the SSM. Therefore, because a dependency exists, connect the host (whenever possible) directly to the SSM. As previously stated, host and storage nodes may be connected to any port on the switch / director.

Complex SAN topologies

Complex SAN topologies derive their complexity from several possible conditions:

- Domain count
- Heterogeneous switch vendors in a single SAN infrastructure
- Distance extension
- Multi-protocol configurations

For the purposes of Invista and RecoverPoint, only the first two conditions will be discussed in this section.

Domain count

SANs come in various sizes and configurations. Customers with hundreds of host and storage ports require several FC switches be linked together through ISLs. Although the size of director class FC switches has significantly increased, core / edge topologies (as

discussed in "Compound core edge switches" on page 199) are common because legacy switches are used in combination with the new high-density directors. These switches can be all departmental switches, all directors, or a combination of both.

A topology with numerous switches will most likely be in one of two designs, each discussed in the next sections:

- Core / edge (hosts on an edge switch and storage on a core switch). (See Figure 173 on page 677)
- Edge / core / edge (hosts and storage on an edge switch but on different tiers). (See Figure 174 on page 678)

Core/edge

Figure 173 on page 677 shows a core / edge design. In this model, hosts are connected to the MDS 9506s (edge tier) and storage is connected to MDS 9513s (core tier). The core tier is the centralized location and connection point for all physical storage in this model. All IO between the host and storage must flow over the ISLs between the edge and core. It is a one hop logical infrastructure. (ISL hop is the link between two switches).

The SSM blade is located in the core. To minimize latency and increase fabric efficiency, the SSM should be co-located with the storage .that is it virtualizing, just as one would do in a Layer 2 SAN.

Note: Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

Note: Direct connections to the SSM are *not* required.

- ISLs between the switches do *not* have to be connected to the SSM blade
- Hosts do *not* have to be connected to the SSM blade
- Storage should *not* be connected to the SSM blade

Since the SSM is located inside the MDS 9513 core, internal routing between blades in the chassis is *not* considered an ISL hop. There is additional latency with the virtualization ASICs; however, there is no protocol overhead associated with routing between blades in a chassis.

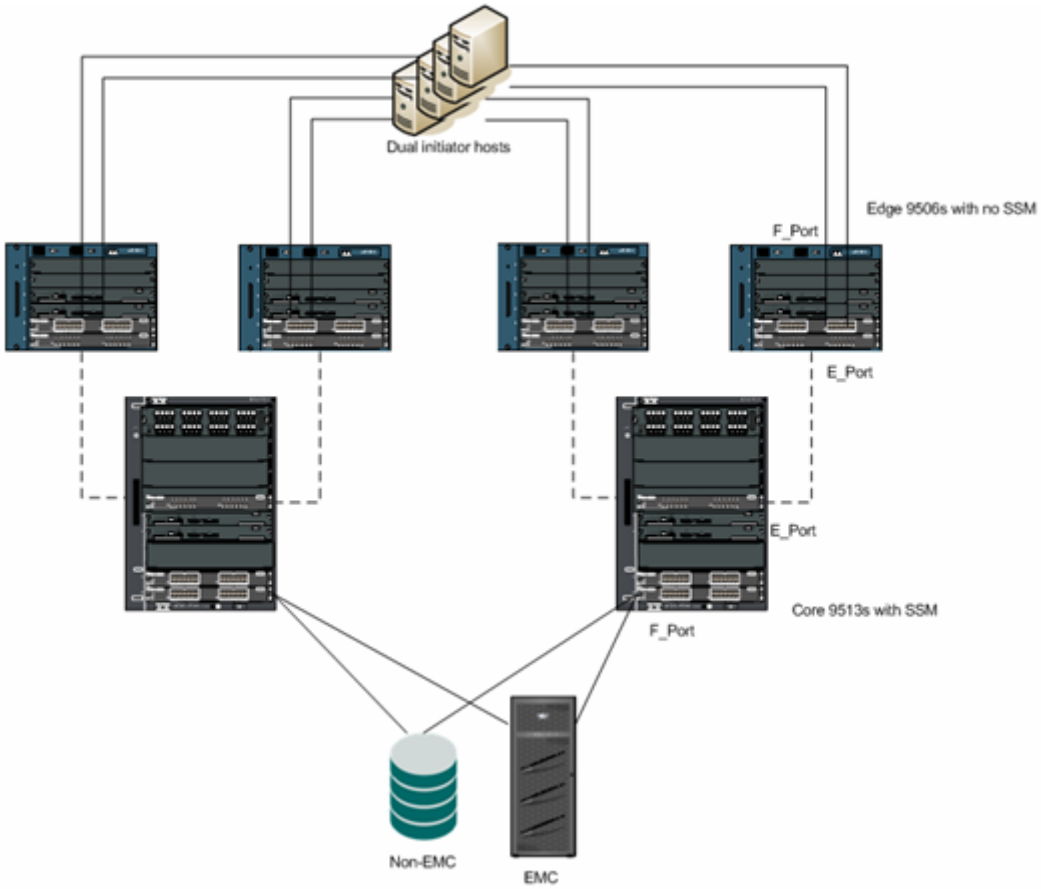


Figure 173 **Core/edge design**

Figure 174 on page 678 shows a core/edge design with a Connectrix B Invista or RecoverPoint implementation. The AP-7420B is an external intelligent switch; therefore, it must be linked through ISLs to the core switch.

Note: Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

In this model, hosts are connected to the DS-32B2 (edge tier) and storage is connected to ED-48000B (core tier). The core tier is the centralized location and connection point for all physical storage in this model. All IO between the host and storage must flow over the

ISLs between the edge and core. It is a one-hop infrastructure for non-virtualized storage. However, all virtualized storage traffic must pass through at least one of the ports on the AP-7420B. Therefore, the IO from the host must traverse an ISL between the DS-32B2 and the ED-48000B. Then, it must traverse an ISL between the ED-48000B and the 7420. Finally, it must traverse an ISL back from the AP-7420B to the ED-48000B where the IO is terminated. This is a three-hop logical topology.

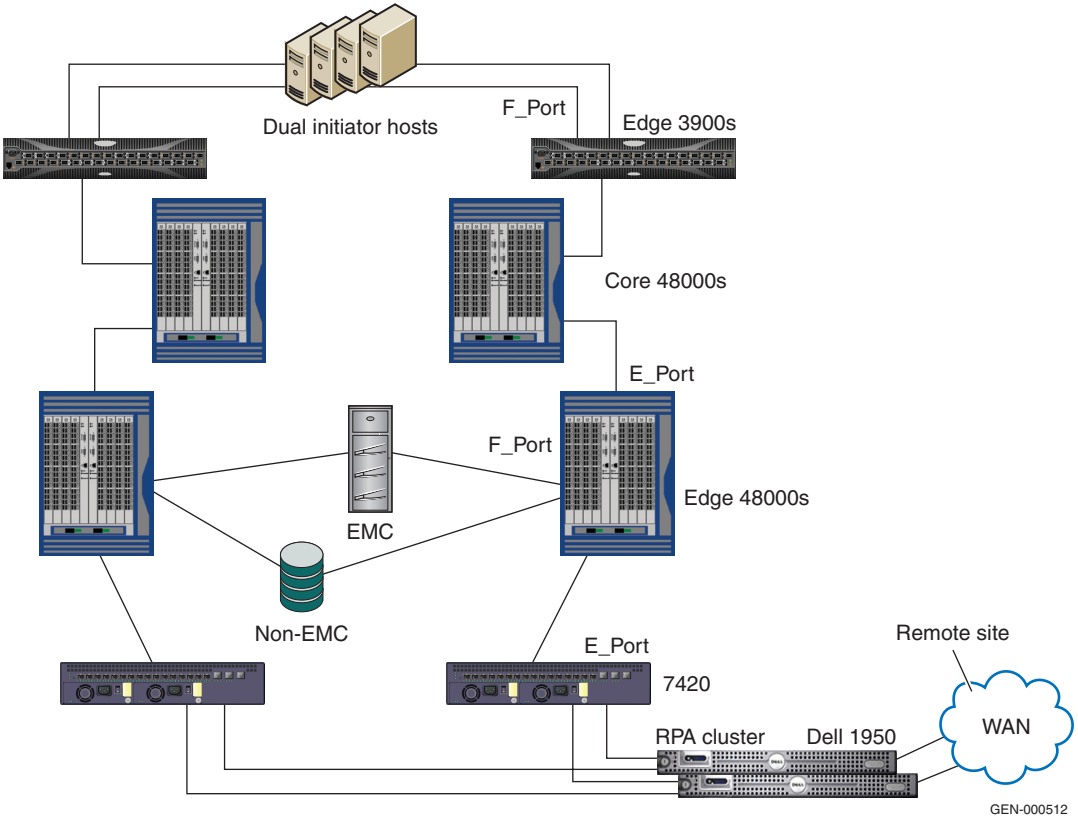


Figure 174 **Core/edge design with a Connectrix B**

Edge/core/edge

Figure 175 on page 680 shows an edge/ core/edge design. In this model, hosts are connected to the Connectrix MDS 9506s (edge tier) and storage is connected to MDS 9506 (storage tier). The MDS 9513s act as a connectivity layer. The core tier is the centralized location and connection point for edge and storage tiers. All IO between the host

and storage must flow over the ISLs between. This is a two-hop logical topology.

Within the fabric, hosts can access virtualized and non-virtualized storage. The location of the physical storage to be virtualized will determine the location of the SSM blade. There are two possibilities:

- Storage to be virtualized is located on a single switch
- Storage to be virtualized is location on multiple switches on the storage tier

Note: Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

Storage on a single switch

If the physical storage is located on one edge switch, place the SSM on the same switch. See the red box in Figure 175 on page 680.

To minimize latency and increase fabric efficiency, the SSM is co-located with the storage that it is virtualizing.

Note: Connections to the SSM are *not* required:

- ISLs between the switches do not have to be connected to the SSM blade
- Hosts do not have to be connected to the SSM blade
- Storage should *not* be connected to the SSM blade

Since the SSM is located inside the edge MDS 9506, internal routing between blades in the chassis is *not* considered an ISL hop. There is additional latency with the virtualization ASICs; however, there is no protocol overhead associated with routing between blades in a chassis.

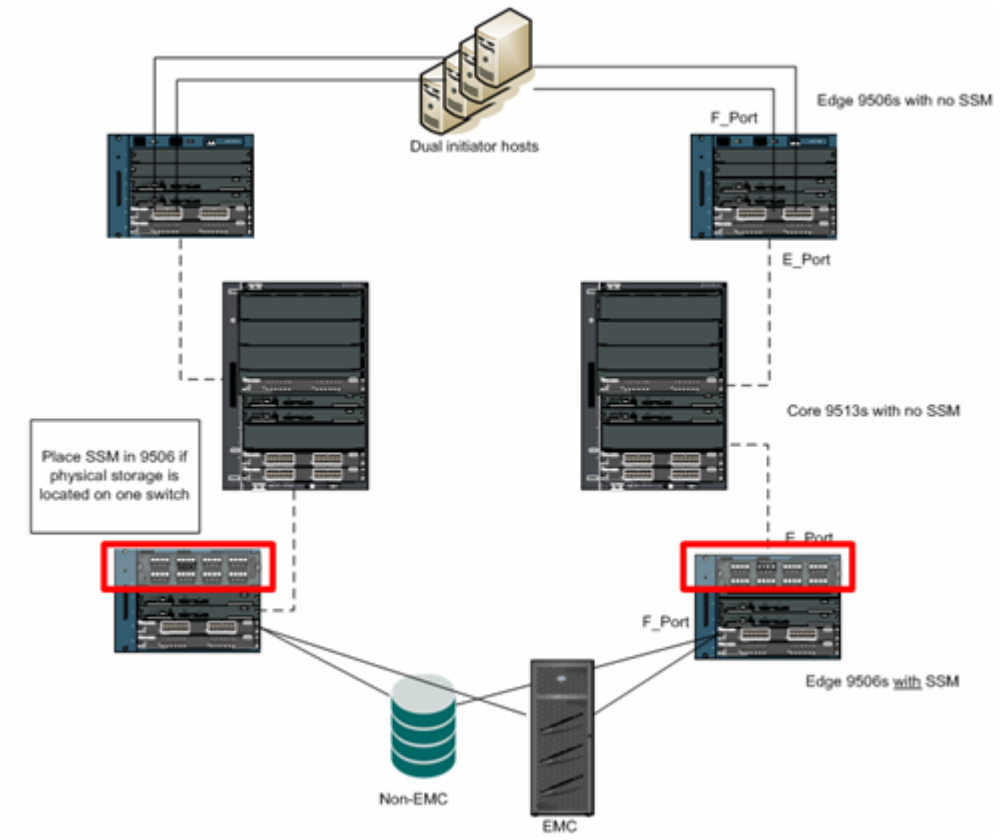


Figure 175 **Storage on a single switch**

Figure 176 on page 681 shows a core/edge/core design with Connectrix B . The AP-7420Bs are linked via ISLs to the edge ED-48000B.

Note: Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

In this model, hosts are connected to the DS-32B2 (edge tier) and storage is connected to a ED-48000B which is the other edge tier. The core ED-48000Bs are for connectivity only. All IO between the host and storage must flow over the ISLs in the core. It is a two-hop infrastructure for non-virtualized storage. However, all virtualized storage traffic must pass through at least one of the ports on the

AP-7420B. Therefore, the IO from the host must traverse an ISL between the DS-32B2 and the ED-48000B. It must then traverse an ISL between the ED-48000B and the next ED-48000B. Then, it must traverse the ISL between the ED-48000B and the AP-7420B. Finally, it must traverse an ISL back from the AP-7420B to the ED-48000B where the IO is terminated. This is a four-hop design that would require an RPQ. (Refer to "Interswitch link hops" on page 685 for more infomation on how storage and intelligent placement impacts fabric design.)

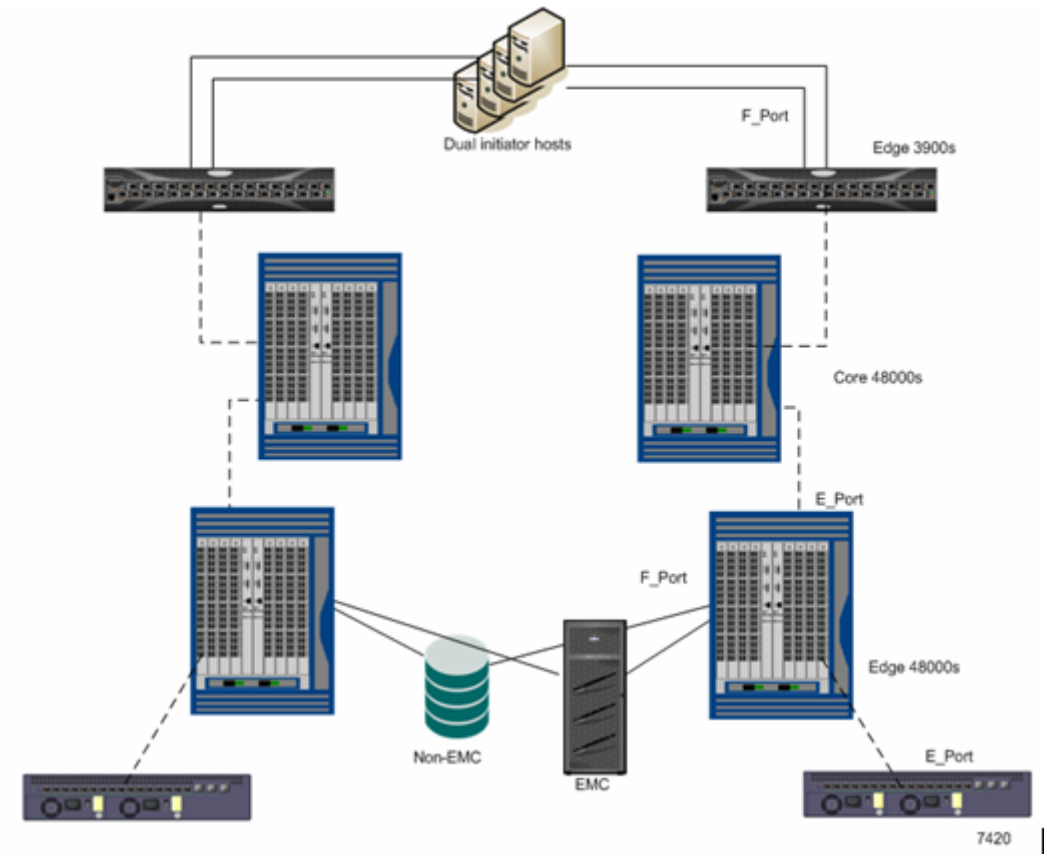


Figure 176 **Core/edge/core design with Brocade**

Storage on multiple switches

If the physical storage is spread amongst several edge switches, a single SSM must be located in a centralized location in the fabric. By doing this, one will achieve the highest possible efficiencies. In Figure 177 on page 683, only one side of the dual fabric configuration is shown. Because the physical storage ports are divided among multiple edge switches in the storage tier, the SSM has been placed in the connectivity layer in the core.

Note: Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

For Invista:

By locating the SSM in the connectivity layer, Invista's VIs will only need to traverse a single ISL in order to access physical storage. If the SSM was placed in one of the storage tier switches, most or some of the traffic between the SSM and storage would traverse two or more ISLs.

For RecoverPoint:

Just as with the Core/Edge design (or any design), all virtualized traffic must flow through the virtualization ASICs. By locating the SSM in the connectivity layer, RecoverPoint's VIs will only need to traverse a single ISL in order to access physical storage. If the SSM was placed in one of the storage tier switches, most or some of the traffic between the SSM and storage would traverse two or more ISLs.

For both Invista and RecoverPoint:

Note: Connections to the SSM are *not* required:

- ISLs between the switches do not have to be connected to the SSM blade
- Hosts do not have to be connected to the SSM blade
- Storage should *not* be connected to the SSM blade

Since the SSM is located inside the MDS 9513, internal routing between blades in the chassis is *not* considered an ISL hop. There is additional latency with the virtualization ASICs; however, there is no protocol overhead associated with routing between blades in a chassis.

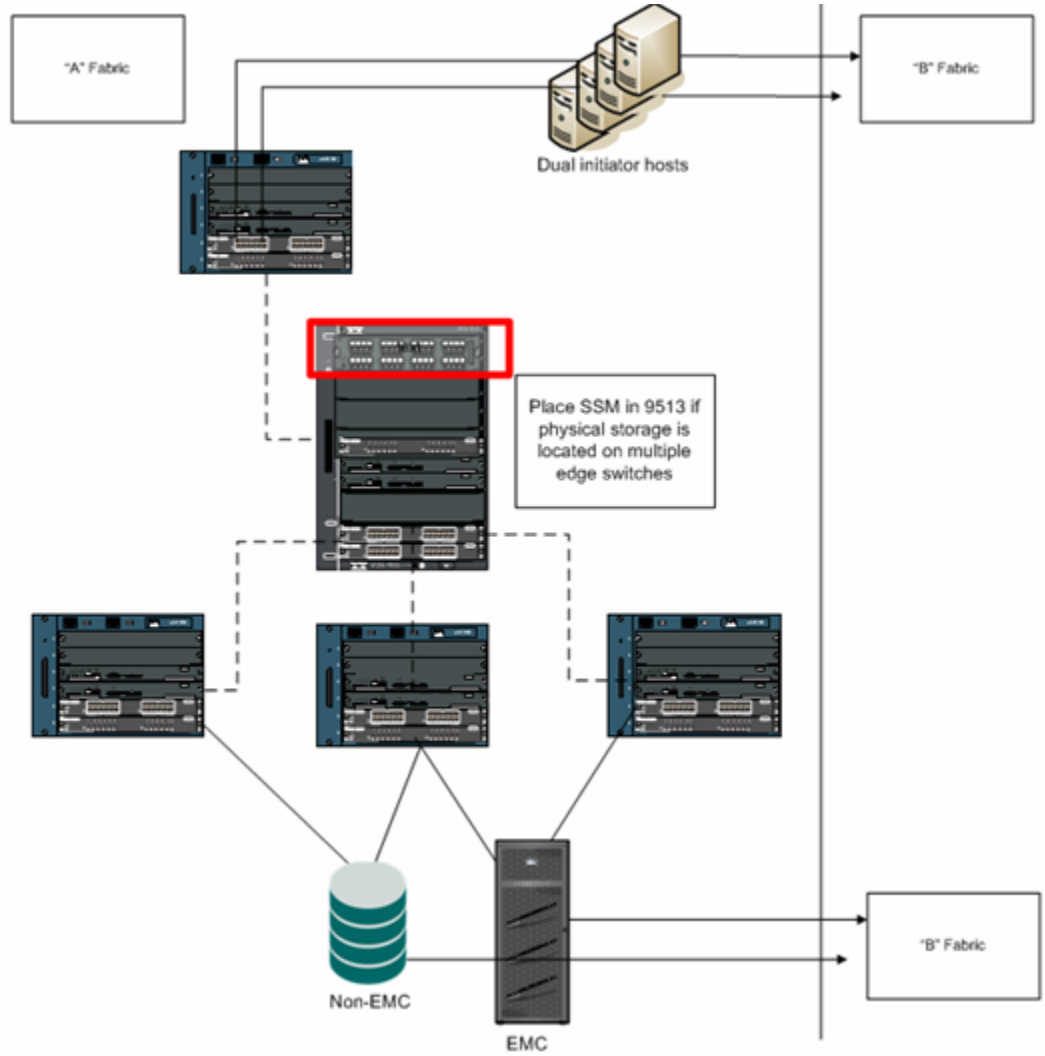


Figure 177 Storage on multiple switches

Similarly in the Connectrix B Invista and RecoverPoint solutions, the physical storage may be spread amongst several edge switches. An AP-7420B must be located in a centralized location in the fabric. This will achieve the highest possible efficiencies. In Figure 178 on page 684, only one side of the dual fabric configuration is shown. Because the physical storage ports are divided among multiple edge

switches in the storage tier, the AP-7420B has been placed in the connectivity layer in the core.

Note: When Invista and / or RecoverPoint is added to a SAN, not all physical storage has to be virtualized. Multiple storage ports may have LUNs to be virtualized. However, if more than 75% of the LUNs to be virtualized reside on a single switch (e.g., Figure 175 on page 680, then locate the AP-7420B on that switch as in Figure 175).

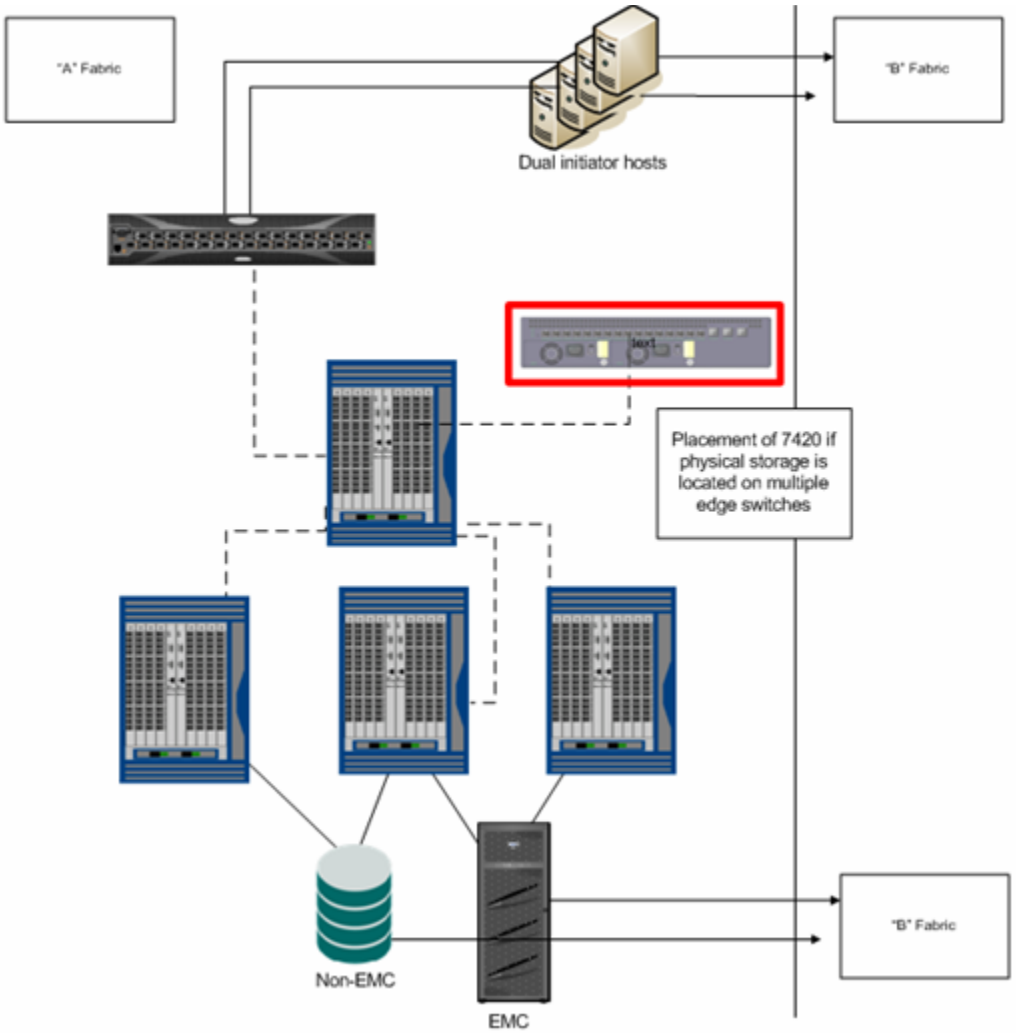


Figure 178 Storage on multiple swithces with Connectrix B

Interswitch link hops

An interswitch link hop is the physical link between two switches. The Fibre Channel switch itself is not considered a hop. The number of hops specified in the Switched Fabric Topology table of *EMC Support Matrix* states the maximum supported number of hops in a fabric under no fault conditions. For Class 3 traffic from initiator port to target port, and for Class F traffic among switches, there could be between three or five ISL hops, depending on the vendor. The physical link between host and switch as well as storage and switch is not counted as a hop. If there is a fault in the fabric, more than the *EMC Support Matrix* specified number of hops will be supported temporarily until the fault is resolved. In excess of maximum specified hops are not supported under normal operating conditions.

Note: In a fabric topology with **RecoverPoint**, all virtualized traffic must flow through an ASIC on the AP-7420B or the SSM blade.

In the previous sections, recommendations were provided for placement in simple and complex fabrics. When including an intelligent switch into a fabric, traffic patterns across ISLs may change and hop counts may increase depending on the vendor and switch type used. The following diagrams display hop counts in various fabric configurations.

In a Cisco MDS design, the SSM blade can be inserted in any MDS chassis with a slot. This includes the MDS 9513, 9509, 9506, and 9216. If the SSM is added to an existing fabric, the addition of the blade does not add any hops within the chassis itself. However, depending on the placement of the SSM in the environment, there may be a direct impact on the number of ISL hops.

Note: All virtualized traffic must pass through the SSM. All non-virtualized traffic does not have to pass through the SSM.

Figure 179 on page 686 and Figure 180 on page 687, show an example of a four-switch mesh design using MDS 9506s. (Fabric ‘B’ is not shown for simplicity of the diagram. Fabric ‘B’ would be a mirror of Fabric ‘A’.) Figure 179 shows a host attached to *sw1* and physical storage attached to *sw4*. The SSM module has been placed in *sw3*. The host IO must follow a path from *sw1* to *sw3* (first hop) because of the FSPF (fabric shortest path first) algorithm inherent to FC switches.

The SSM then re-maps the IO and sends it to its physical target on *sw4* (second hop). This is a two-hop solution.

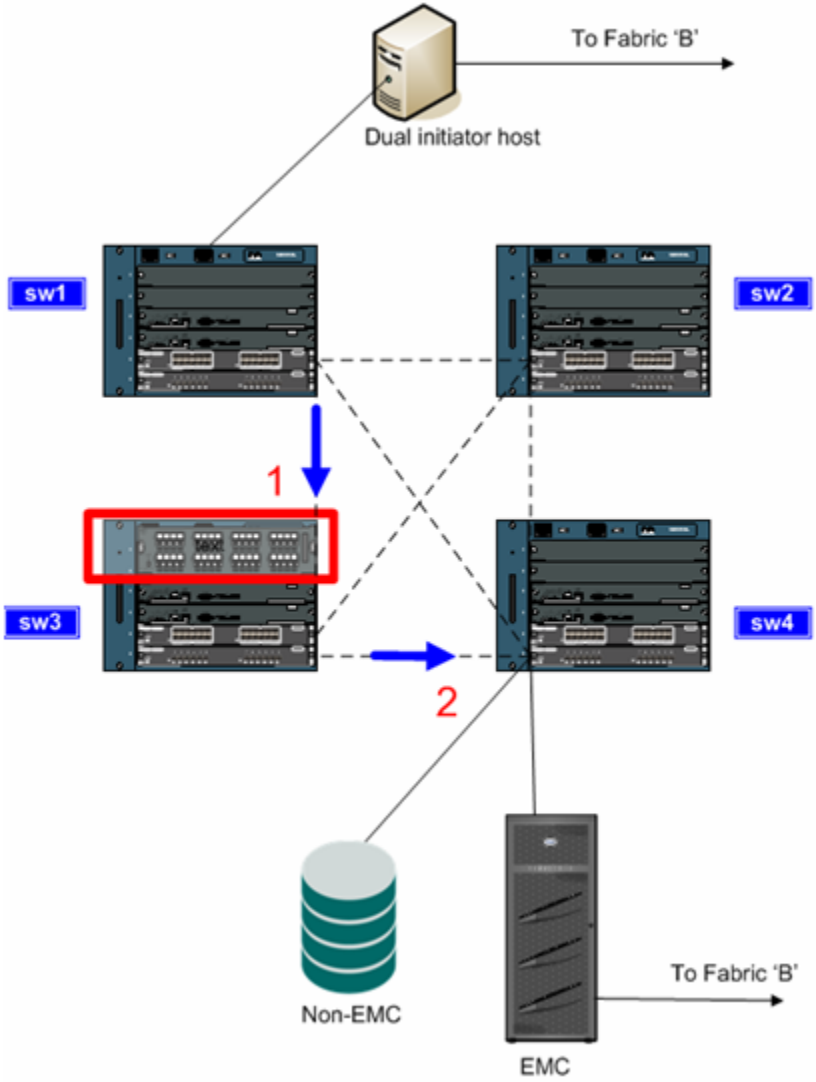


Figure 179 Four-switch mesh design using MDS 9506s (Example 1)

Figure 180 shows the same host attached to *sw1* and the same physical storage attached to *sw4*. The SSM module has been moved to *sw4*. The host IO must follow a path from *sw1* to *sw4*. The SSM then re-maps the IO and sends it to its physical target on the same switch. There is no hop increase when IO passes between blades in the MDS chassis. This is a preferred one-hop solution.

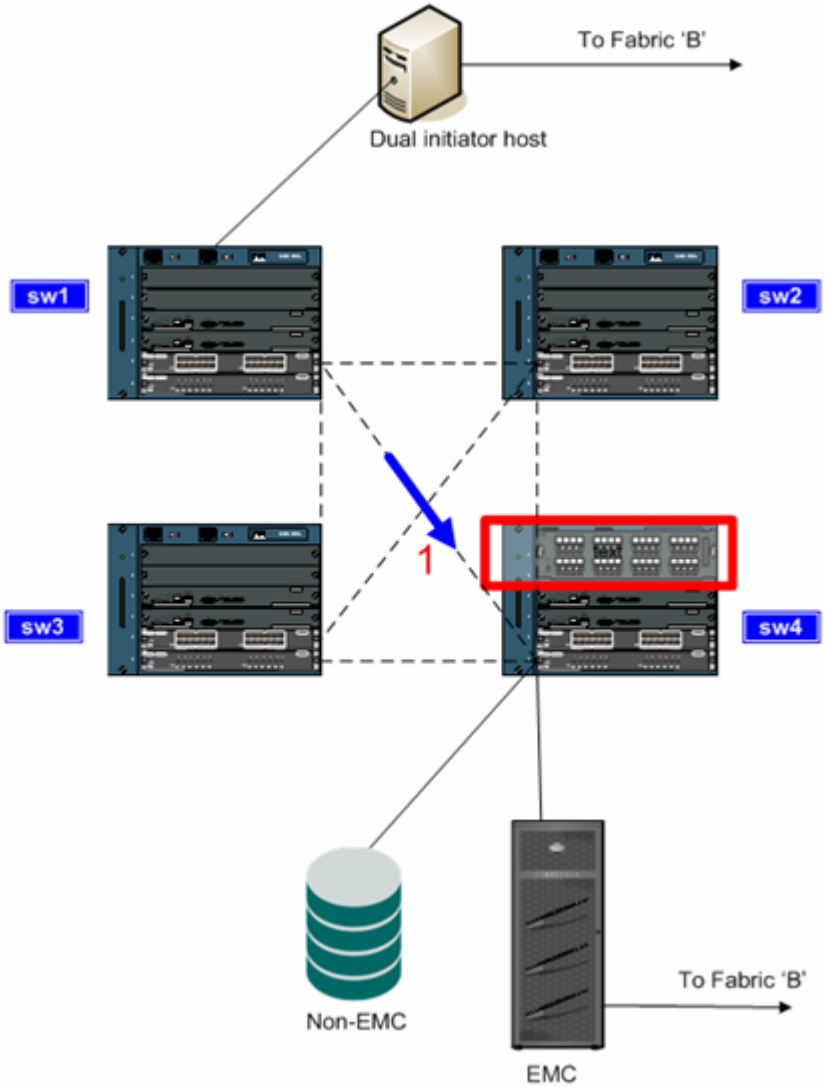


Figure 180 Four-switch mesh design using MDS 9506s (Example 2)

Figure 181 on page 689 shows a four-switch mesh of ED-48000Bs with an AP-7420B as the intelligent switch. Because the AP-7420B is an external switch, it must be linked via an ISL to the existing fabric. Again, the host is attached to *sw1* and physical storage attached to *sw4*. The AP-7420B has been linked via an ISL to *sw4*. The host IO must follow a path from *sw1* to *sw4* (first hop) because of the FSPF (fabric shortest path first) algorithm inherent to FC switches. The IO is then passed from *sw4* to the 7420 (second hop). The virtualization ASIC re-maps the IO and passes it from theAP-7420B to *sw4* where the physical storage resides (third hop). This is a three-hop solution.

Note: The number of hops and traffic patterns would be identical in an MDS configuration if an MDS 9216 with an SSM blade which is hosting the Invista instance is attached to an existing fabric of MDS directors.

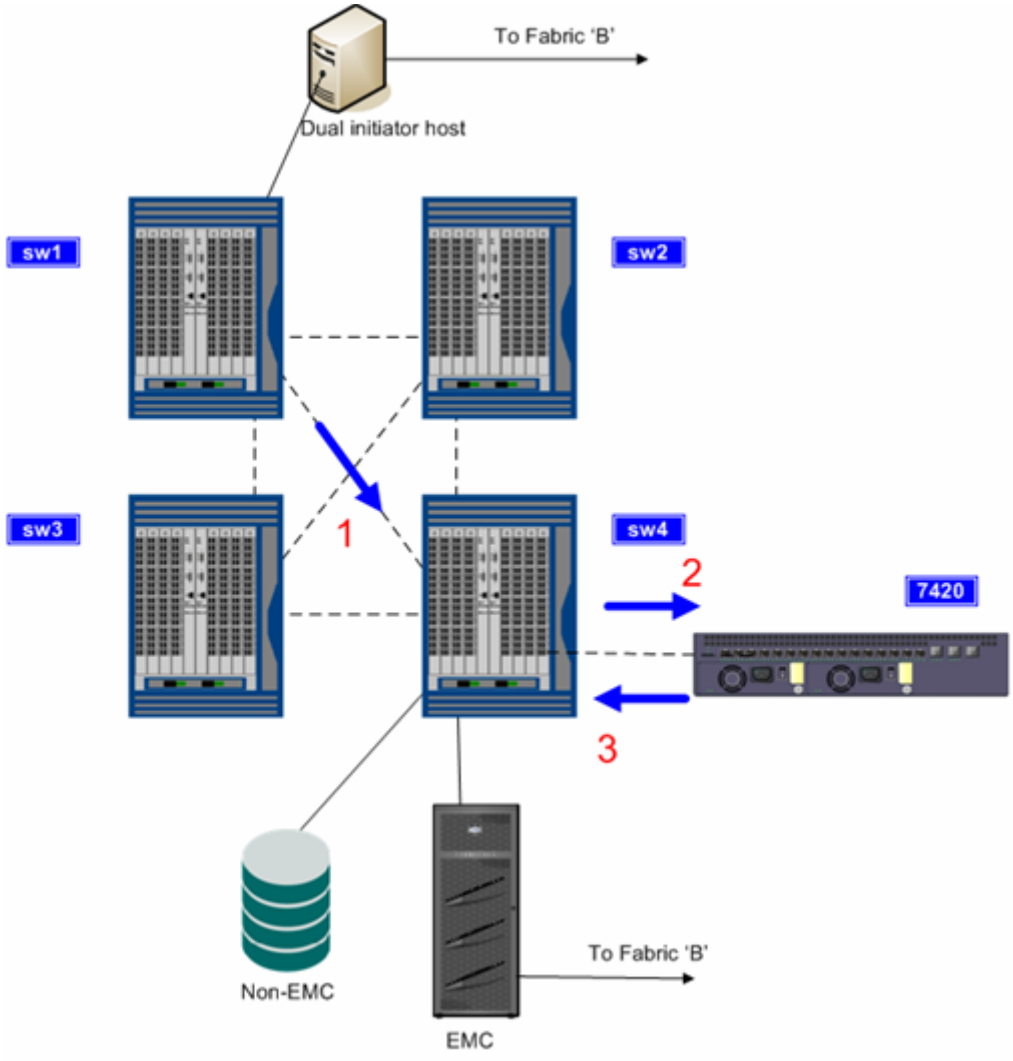


Figure 181 **Four-switch mesh of ED-48000Bs with an AP-7420B as the intelligent switch**

With both mesh and complex fabric designs, it is not always possible to maintain a one-hop solution. With higher domain-count fabrics, there is a likelihood of increased hops. This is true of a virtualized and non-virtualized environment. It is also likely with the addition of intelligent switches or blades that some hosts will have more hops than others due to the placement of the physical storage in relation to

the location of the intelligent switch. Because Invista creates a pool of storage from which it can read from and write to, the physical target may reside on a switch that does not host the intelligent switch.

Figure 182 on page 691 shows a four-switch mesh of MDS 9506s. Storage has been spread across *sw2*, *sw3*, and *sw4* switches. The SSM blade is located in *sw4*. (Again, Fabric 'B' is not shown for simplicity). Both *Host A* and *Host B* are accessing virtualized storage. All IO must pass through the SSM. IO initially passes from *sw1* to *sw4* (first hop) because of the FSPF (fabric shortest path first) algorithm inherent to FC switches. The SSM re-maps the IO. The final destination of the IO *may* be the EMC storage on *sw4*. If the IO terminates there, it is considered a one-hop solution. However, the final destination *may* be the EMC storage on *sw2* or the non-EMC storage on *sw3*. In this case, it is a two-hop solution. *Host C* is *not* accessing virtualized storage. Its physical target is the EMC storage array on *sw2*. The IO passes from *sw1* to *sw2* because of the FSPF (fabric shortest path first) algorithm inherent to FC switches. This is always one-hop solution.

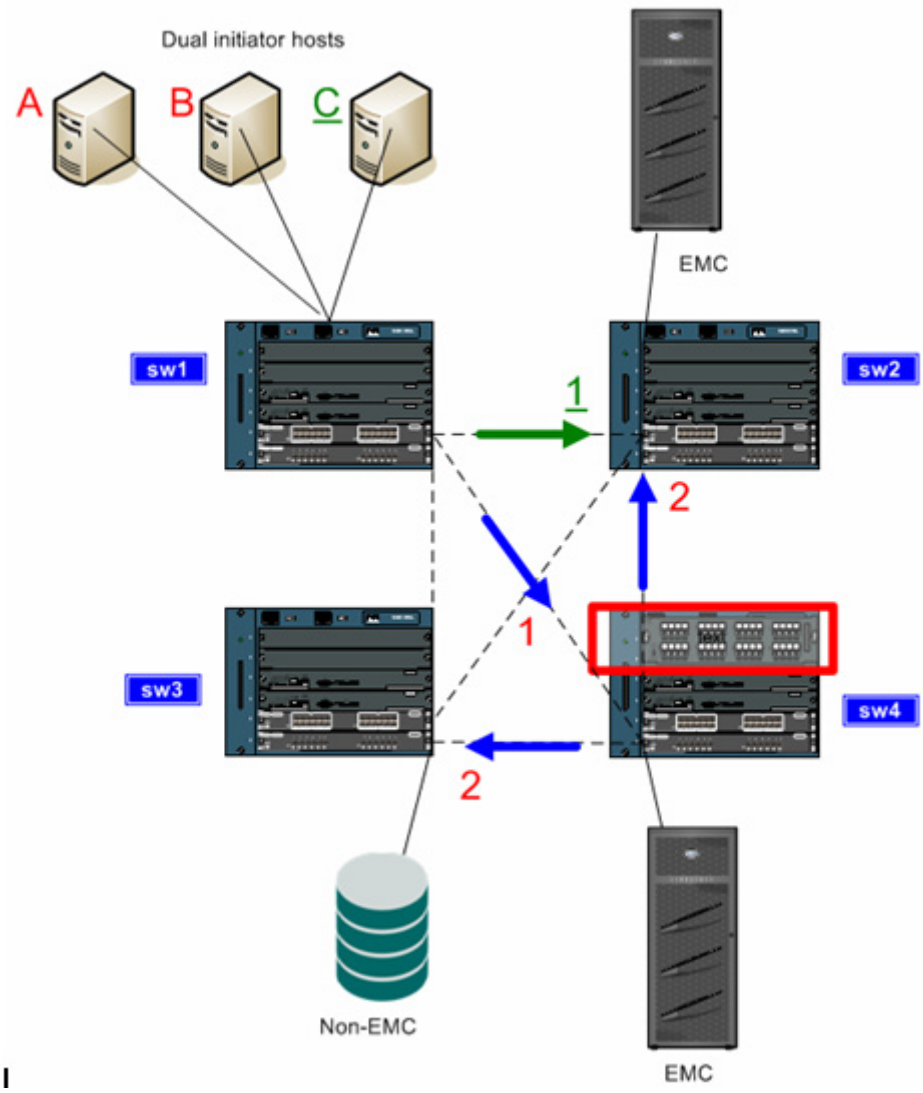


Figure 182 Four-switch mesh of MDS 9506s

A

Storage Networking Concepts and Protocols

This appendix provides the following information needed prior to building a SAN.

- Overview 694
- Channels versus networks versus SANs 695
- Physical versus logical topologies 702
- Fabric design considerations and recommendations 720
- Determining customer requirements 759
- FC SAN concepts 765
- IP SAN concepts 962
- IP over Symmetrix directors 968

Overview

This chapter provides detailed information on the fundamentals of Fibre Channel. It begins with the "Channels versus networks versus SANs" on page 695, which provides a starting point and boundaries for the discussion to follow.

"Physical versus logical topologies" on page 702 discusses the differences of physical and logical topologies.

"Fabric design considerations and recommendations" on page 720 and "Host design considerations" on page 778 discuss what to consider when designing your fabric and host and provides recommendations.

"Determining customer requirements" on page 759 provides guidelines for implementing a new SAN.

"FC SAN concepts" on page 765 contains a high-level overview of Fibre Channel.

"IP SAN concepts" on page 962 describes IP in two elements, internetworks and routing, and IP addresssing.

"IP over Symmetrix directors" on page 968 describes IP connectivity using channel directors in Symmetrix 8000 and DMX series systems.

Channels versus networks versus SANs

This section describes the differences between channels, networks, Fibre Channel with network characteristics, and SANs.

Channels

Traditionally, host computer operating systems have communicated with storage devices over channel connections, such as parallel bus and tag, ESCON, and SCSI. These channel technologies provide fixed connections between host systems and their peripheral devices.

The static connections are defined to the operating system in advance. Tight integration between the transmission protocol and the physical interface minimizes the overhead required to establish communication and transport large amounts of data to the statically defined devices.

Some characteristics of channel technologies are:

- High performance
- Low protocol overhead
- Static configuration
- Short distance (although ESCON is somewhat of an exception)
- Connectivity within a single system

Networks

Network technologies are more flexible than channel technologies, and provide greater distance capabilities. Most networks provide connectivity between client or host systems, and carry a variety of data between the devices. A simple example is a network of desktop PCs within a company. This type of setup can provide each PC with connectivity to file and print services, server-based applications, and corporate intranets.

The networks these PCs are connected to provide shared bandwidth and the ability to communicate with many different systems. This flexibility results in greater protocol overhead and reduced performance.

Some characteristics of network technologies are:

- Low performance
- High protocol overhead
- Dynamic configuration
- Long distance
- Connectivity among different systems.

Fibre Channel: Channels with network characteristics

Fibre Channel captures some of the benefits of both channels and networks. A Fibre Channel fabric is a switched network, providing a set of generic, low-level services onto which host channel architectures and network architectures can be mapped. It provides a serial data transfer interface that operates over copper wire and/or optical fiber at data rates currently up to 1200 MB/s. (This could increase since standards specifying speeds up to 1600 MB/s have recently been ratified.) Networking and I/O protocols (such as SCSI commands) are mapped to Fibre Channel constructs and then encapsulated and transported within Fibre Channel frames. This process allows high-speed transfer of multiple protocols over the same physical interface.

SCSI and Fibre Channel

Before designing your fabric, it is essential to understand why Fibre Channel is so widely used and accepted in the data center and why it is the natural technological progression from direct-attach SCSI devices. The phrase *Fibre Channel* is often used as an abbreviation of *SCSI over Fibre Channel*. Fibre Channel is a transport protocol that allows mapping other service-oriented or device-oriented protocols within its transport frames. SCSI over Fibre Channel allows us to overcome the distance, dynamic flexibility, and accessibility limitations associated with traditional direct-attach SCSI.

As with direct-attach SCSI, Fibre Channel provides the block level access to the devices that allows the host system to identify the device as a native device. The true power of native device identification is seen in our ability to use all of our current applications (for example: backup software, volume management, and raw disk management) without modification. With this understanding, the considerations and methodologies needed to build Fibre Channel SANs will be discussed.

Storage Area Networks (SANs)

Figure 183 shows an example of a SAN.

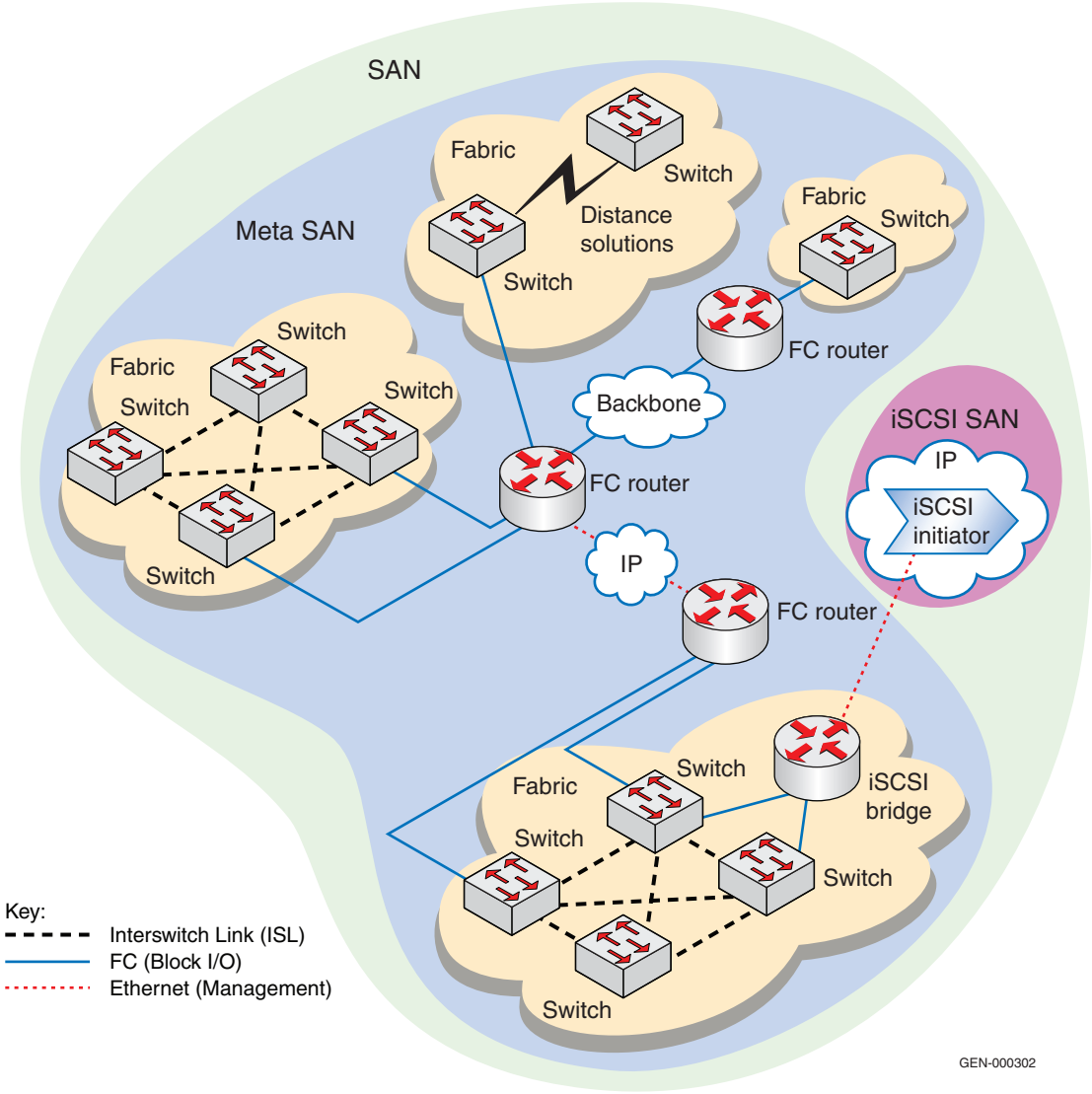


Figure 183 SAN example

Note: Although the FC initiators and FC targets have been omitted from the diagram, it should be understood that the graphic is intended to highlight the majority of supported end-to-end connectivity configurations.

A Storage Area Network is a network that has been created to facilitate block, file, or object I/O operations. SANs can be further broken down into three subcategories, each further discussed in this section:

- "FC SANs," next
- "iSCSI SANs" on page 700
- "Multiprotocol SANs" on page 700

FC SANs

Fibre Channel Storage Area Networks exclusively use Fibre Channel for the transport of command, data, or status information. Fibre Channel SANs consist of either an individual fabric or multiple fabrics interconnected using a routing function (Figure 183 on page 697). As can be seen in Figure 183, FC SANs can be further broken down into fabrics and Meta SANs. A Meta SAN could also be considered to be an ISANs, (refer to "Meta SAN" on page 700), depending on how the routing function is accomplished.

Fabric

A *fabric* is a collection of Fibre Channel ports that all share the same 24-bit address space. Information about ports that are logged into a particular fabric is distributed to all switching elements in the fabric. When fabrics were first introduced, every physical chassis connected by an ISL could be considered to be in the same fabric. Today, with the introduction of VSANs, virtual switches, and fabric routing, this physical relationship breaks down and the VSAN, virtual switch ID, or Fabric ID needs to be considered when trying to determine which ports are in the same fabric. For more information on VSANs, virtual switches, or Fabric routing, see "VSANs" on page 958, "Virtual switches" on page 958 and "Fibre Channel Routing" on page 937.

Although it is not a best practice, a fabric can span long distances by using specialized distance extension equipment or by using a protocol such as FCIP (refer to "FCIP" on page 699). Some may argue that the use of FCIP in a fabric makes it an IP SAN, but this is not the case. A fabric that spans a distance by using an FC-over-IP link still shares the same address space. An argument can be made that such a fabric is a multiprotocol SAN, as described in "Multiprotocol SANs" on page 700).

FCIP

FCIP (Fibre Channel over IP), defined in RFC (Request for Comments[1]) 3821, is an IP-based storage networking technology developed by the Internet Engineering Task Force. RFC 3821 describes mechanisms that allow the interconnection of islands of Fibre Channel SANs over IP Networks to form a unified SAN in a single Fibre Channel fabric. The motivation behind defining these interconnection mechanisms is a desire to connect physically remote FC sites allowing remote disk access, tape backup, and live mirroring.

FCIP enables the transmission of information by tunneling data between storage area network (SAN) facilities over IP networks. The tunneling functionality is achieved by encapsulating Fibre Channel frames inside of TCP/IP datagrams as defined in RFC 3643 and FC-BB-2.

As shown in Figure 184, depending on the MTU size, a FC frame can be a part of one or more TCP segments.

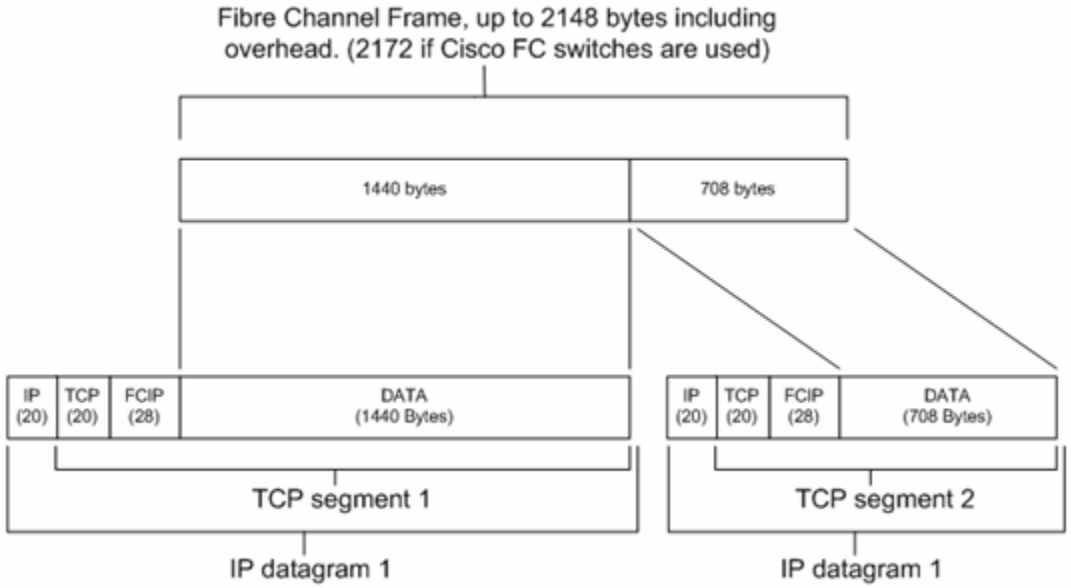


Figure 184 **FCIP-MTU 1500**

1. Request for Comments (RFC) are official documents from the Internet Engineering Task Force (iETF) with unlimited distribution. These are numbered in a series and are referred to by numbers.

FCIP is designed to be transparent to Fibre Channel. In most cases, EMC customers would utilize FCIP technology in one of two ways:

- To allow SRDF or MirrorView operations to be conducted across an IP network.
- To connect Fibre Channel SAN *islands*, essentially extending the reach of the fabric.

E-Lab Navigator describes the qualified devices and configurations when examining these options.

Note: For more information on specific use cases, refer to Chapter 3, "Extended Distance."

Meta SAN

A *Meta SAN* is a collection of fabrics that are connected together through some routing mechanism. The routing mechanism can either be a single router or multiple routers, which are connected either by the FC or IP protocol. Examples of Meta SANs are: an edge switch connected through IFL to a EX_Port on a Brocade switch, fabrics connected by Cisco's IVR-NAT, a single Brocade M-Series 2640, or two Brocade M-Series 2640s connected by an MFCP link., fabrics or Meta SANs connected by an iFCP link, two Brocade fabrics connected by VEx_Ports, or a transit Cisco VSAN which spans an IP link.

iSCSI SANs

iSCSI Storage Area Networks exclusively use the iSCSI protocol over TCP/IP for the transport of command, data, and status information.

IP is a network layer protocol that provides datagram routing services for transport layer protocols such as Transmission Control Protocol (refer to "Transmission Control Protocol (TCP)" on page 963) and UDP (User Datagram Protocol). (Refer to Figure 349 on page 964). Datagrams are the packets that carry user and application data end-to-end across router-connected networks. A TCP/IP network can actually be viewed as a communications medium designed to transport IP packets.

For more information on IP SANs, refer to "IP SAN concepts" on page 962.

Multiprotocol SANs

Multiprotocol Storage Area Networks have some elements that use Fibre Channel or iSCSI for the transport of command, data and status information. This class of SAN also includes ISANs formed by

spanning FCIP and iFCP links as well as FC SANs that have iSCSI initiators accessing FC targets through a bridge.

iFCP

iFCP, defined in RFC 4172, specifies an architecture and a gateway-to-gateway protocol for the implementation of Fibre Channel fabric functionality over an IP network. This functionality is provided through TCP protocols for Fibre Channel frame transport and the distributed fabric services specified by the Fibre Channel standards. The architecture enables internetworking of Fibre Channel devices through gateway-accessed regions with the fault isolation properties of autonomous systems and the scalability of the IP network.

iFCP supports FCP, the ANSI SCSI serialization standard to transmit SCSI commands, data, and status information between a SCSI initiator and SCSI target on a serial link. iFCP replaces the transport layer with an IP network (Ethernet), but retains the upper layer information such as FCP.

In an EMC environment, iFCP would allow Fibre Channel switches to use IP as the interswitch *fabric* protocol. This would allow customers to utilize existing IP infrastructure (cabling and switches, for example) to support storage-to-storage communications.

Physical versus logical topologies

The Fibre Channel environment consists of a *physical* topology and a *logical* topology. The *physical* topology describes the physical interconnects among devices (servers, storage, and switch in the EMC-specific environment). The *logical* topology describes the logical paths established between the operating system device names and their associated storage ports and volumes. Physical and logical topologies are further described in this section.

In Figure 185, the solid lines and their connections represent the physical topology and the dotted lines represent the logical topology.

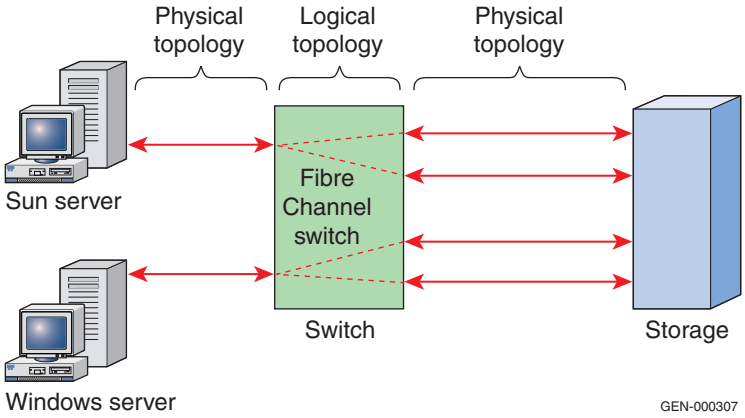


Figure 185 Physical and logical topologies

Physical topology

This section describes the details of the individual topology representations and shows how a logical topology is overlaid onto a physical topology. The physical topology can be described as actual hardware components in a fabric and the Fibre Channel cabling that interconnects them. A physical topology also includes the geographical locations of the switches and distances between them.

Some examples of the components and concepts used to describe the physical topology of a fabric are:

- Number of switches in the fabric
- Number of hops between any two switches

- Number of ports per switch
- Number of ISLs between switches
- Physical distance between any two switches

When describing a particular physical topology, it can be discussed in terms of its number of *tiers*. The number of tiers in the fabric is based on the number of switches that are traversed between the farthest two points in the fabric. It should be noted that this number is based on the infrastructure constructed by the fabric topology and does not concern itself with how the storage and server are connected across the switches.

Increasing the number of tiers in a fabric also increases the distance that a fabric management message must travel to reach every switch in the fabric. Increasing that distance can affect the time it takes to propagate and complete a fabric reconfiguration event (for example, adding a new switch), or zone set propagation event. Figure 186 displays one-tier, two-tier, and three-tier physical fabrics.

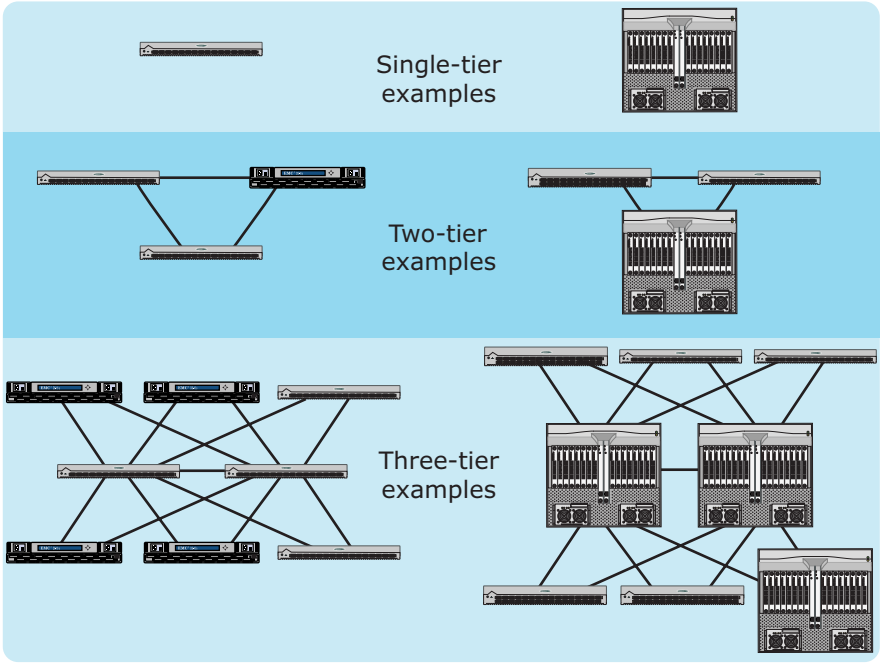


Figure 186 Tiered fabrics

As Figure 186 shows, a single-tier physical topology has a single switch. A two-tier topology has up to two switches between any two end-points in the fabric. A three-tier topology has up to three switches between any two end-points in the fabric. Currently, EMC recommends that the size of the fabric not exceed three hops, which equates to a four-tier physical fabric topology.

Final identification of your physical topology and later expansion of that topology relies on your ability to not only understand the individual impacts of the issues mentioned in this section, but also your selection of data protection schemes, logical topology, and management paradigm.

Core/edge fabric

Another common fabric design is the *core/edge* fabric. A core/edge design is based on the assumption that there will always be more host ports than storage ports and that providing equal and deterministic access to the storage from anywhere in the fabric is a marketable benefit to fabric management and storage administrators.

A core/edge design is built by consolidating storage access into a centrally accessible pool at the logical center of the fabric. From this *core* you can attach as many edge switches as necessary to service the hosts that require access to this storage. Each edge switch will be connected to each core switch, maintaining the fabric's accessibility as well its robustness. The number of core-to-edge ISLs can be based on the individual bandwidth requirements from each edge switch to the respective core switches or on overall fabric balance and flexibility. As described under "Methodology 1: Balanced fabrics" on page 713, balancing a fabric allows you easier management of componenent placement as the fabric requirements increase.

Since hosts do not have to communicate with other hosts over the Fibre Channel storage infrastructure, you can eliminate the need to connect edge switches to other edge switches over ISLs. Eliminating ISLs can greatly increase the number of host and storage ports that can be attached to the fabric. A simple core/edge fabric is designed to provide all hosts with single-hop access to all storage in the fabric.

While you strive to keep all hosts out at the edges, the attachment of excessively active hosts at the core to reduce some of the ISL bandwidth requirements is not allowed for. Core-to-core ISLs are added to facilitate RDF traffic, fabric management traffic, and backup traffic.

The core of the fabric can be extended to accept new storage access. EMC recommends that the storage core be built out as a full mesh to

perpetuate multiple paths to the storage, multiple paths for fabric management, and shortest-path access to all switches in the fabric.

No matter what fabric topology you are using in your design, you can still use the methodologies defined by mirroring fabrics or balancing fabrics, or use the steps outlined in this section on how to lay out hosts, storage, and ISLs to further enhance the accessibility, availability and manageability of your fabric.

Figure 187 demonstrate the topologies of single-core design.

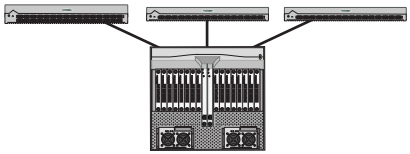


Figure 187 Single-core fabric

EMC recommends that single-core fabrics should always employ the fabric mirroring methodology to enhance the protection and robustness of the entire environment. While Figure 187 shows a director-class switch at the core and departmental switches at the edges, any switch type may be used, based on your connectivity needs.

Figure 188 demonstrate the topologies of a dual-core design.

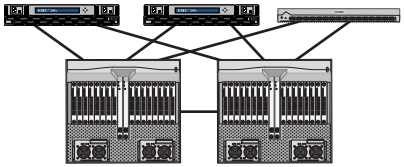


Figure 188 Dual-core fabric

Dual-core fabrics provide the ability to distribute mirrored storage ports across the core switches without going to a mirrored fabric protection scheme. (However, mirrored schemes provide even greater protection and isolation against events in the fabric.) Dual cores can also be expanded to contain more storage core switches, but for each core switches that is added, each edge switch will need ISLs to that core switch to maintain the topology.

Benefits

While the core / edge does not normally provide any zero-hop storage access, it does provide one-hop storage access to all storage in the system. Because traffic also travels in a deterministic pattern (from the edge to the core), a core / edge provides easier calculation of ISL loading and traffic patterns.

Since each tier's switch is used for either storage or hosts, you can easily identify which resources are approaching their capacity and arrive at an easier set of rules for scaling and apportioning. For example, if you are running out of available ports on our host switch, you know that if you are planning to expand your host numbers you must order the appropriate number of switches to handle the new hosts with their HBAs, as well as the ISLs from the edge to the cores to support the number of new switches.

A well-defined, easily reproducible building block approach makes rolling out new fabrics easier. Core / edge fabrics can be scaled to larger environments by linking core switches, adding more core switches, or adding more edge switches. This method can be used to extend the existing simple core / edge model or to expand the fabric into a compound or complex core / edge model. (Refer to "Compound core / edge fabric" on page 731 for more information.)

Limitations

As the number of cores increases, it may be prohibitive to continue to maintain ISLs from each core to each edge switch. When this happens, EMC recommends shifting your fabric design to a compound or complex core / edge design, discussed next.

Logical topology

In contrast to a fabric's physical topology, a logical topology is concerned with where the Fibre Channel elements are attached around the fabric and the relationships (zoning) that define how these elements will be used together.

When describing a fabric's logical topology, it can also be discussed in terms of its number of logical tiers. The number of logical tiers in the topology is based on the number of switches traversed between a server and the storage zoned to it.

Since servers and storage can be located anywhere in the fabric, there could be several different logical tier relationships in the same fabric.

Figure 189 demonstrates an example of multiple logical tier relationships in the same fabric.

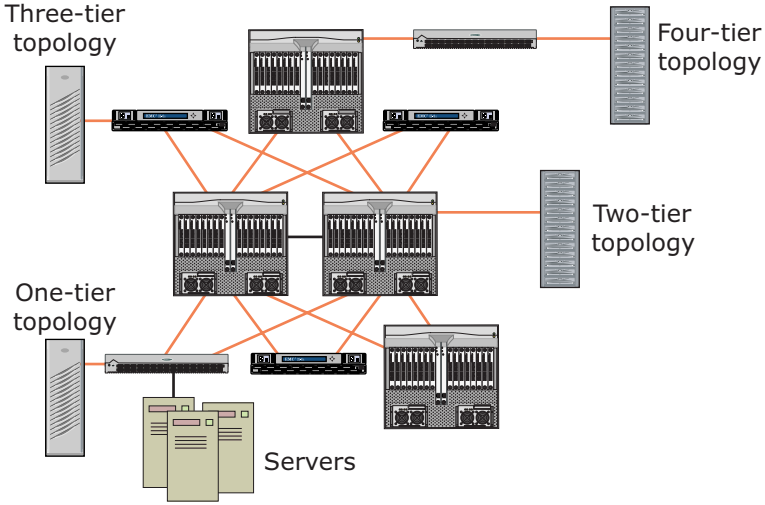


Figure 189 **Multiple tiers in the same fabric**

Note: As Figure 189 on page 707 shows, the physical fabric size might not directly correlate to the logical fabric size.

Increasing the number of logical tiers in a fabric also increases the distance the data must travel from storage to the server. Each switch that the data has to traverse, and the length of the links between them, adds to the latency involved in sending or retrieving the data. Increasing the size of the logical fabric also increases the probability that bandwidth will be aggregated across the tiers. Excessive aggregation of traffic across the physical tiers can lead to fabric congestion and increased data retrieval latencies.

EMC recommends that you limit the path between the storage and the servers that are zoned to them to three hops. As Figure 189 shows, you can construct a four-tier logical fabric with three hops.

Logical topologies in the EMC / Fibre Channel switch environment can generally be described in terms of *fan-in* (*into* the EMC storage array) and *fan-out* (*out of* the EMC storage array). The example in Figure 185 shows a fan-in rate of 1:2 for each server. Refer to "Fan-in and fan-out considerations" on page 782 for more information.

Note: You can find some recommended fan-in and fan-out ratios on E-Lab Navigator.

The logical connectivity topologies, identified in Table 27, have been developed to solve three distinct customer problems. Each of these will be discussed in this section.

Table 27 Fibre Channel topology solutions

Problem	Solution
Proximity extension	Distance topology — Uses shortwave-to-longwave conversion to extend server-to-storage distance beyond shortwave's 500-meter limitation. Also includes DWDM solutions. (Refer to "Distance topology" on page 708.)
Capacity expansion	Capacity topology — Expands the storage capacity supported per host port by allowing a host port to connect to two or more Symmetrix nodes. (fan-in). (Refer to "Capacity topology" on page 709.)
Storage consolidation	Consolidation topology — Expands the number of servers supported per Symmetrix port (fan-out). (Refer to "Consolidation topology" on page 710.)
More than one of the above problems	Combined topologies — For example, if you combine all three topologies you get high-capacity, highly available, multi-server, geographically dispersed clusters with a minimum of host I/O slots . (Refer to "Combined topologies" on page 711.)

Distance topology

Most Fibre Channel equipment is designed for shortwave lasers with multimode optical fibers, which is effective for up to 500 meters. Greater distances can be achieved using longwave lasers with single-mode fiber.

Symmetrix Fibre Channel directors support distances up to 500 meters with 50/125 fiber cable. By using the appropriate longwave adapter and 9/125 single mode cable, the distance can be extended to 10 Km from one Symmetrix port to a second longwave Symmetrix port or switch port.

Note: Although the Symmetrix uses optics which are capable of 10 Km, links over 7 Km will experience a slight droop in throughput when running at 2 GB/s due to the number of BB_Credit each FA provides. See "Determining customer requirements" on page 759 for more information.

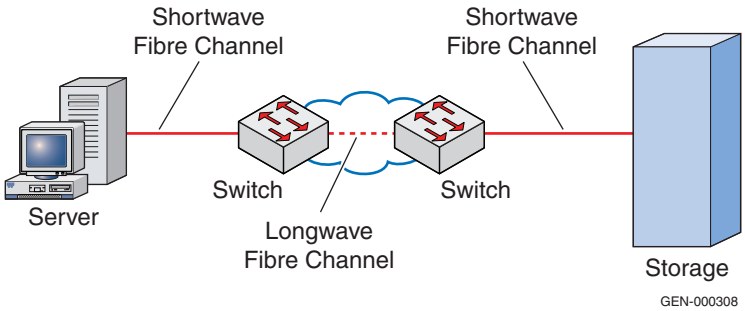


Figure 190 **FC-Switched fabric distance topology example**

Greater distances require two fabric switches connected with longwave lasers over 9/125 single-mode fiber cable. The switches interconnect through a port called an *expansion* port (E_Port) over a connection that is referred to as an Interswitch Link (ISL).

Another option for distance extension requires a DWDM (Dense Wavelength Division Multiplexer) system.

Another distance extension technique is to use either iFCP or a similar IP-based solution. The advantages for using an IP based solution over distance are error detection and the ability for the IP gateway to request the re-transmission of lost packets.

Capacity topology

This topology expands capacity in the Symmetrix environment, allowing a single host bus adapter in a file server to access a large capacity that might be stored on multiple Symmetrix devices.

The logical concept of the capacity topology in a switched fabric is described by the *fan-in* rate (*in* as in *into* the storage system).

The capacity topology uses a logical topology of fan-in, as shown in Figure 191 on page 710, which illustrates a fan-in rate of four. Each storage port in the figure connects to a shared server port.

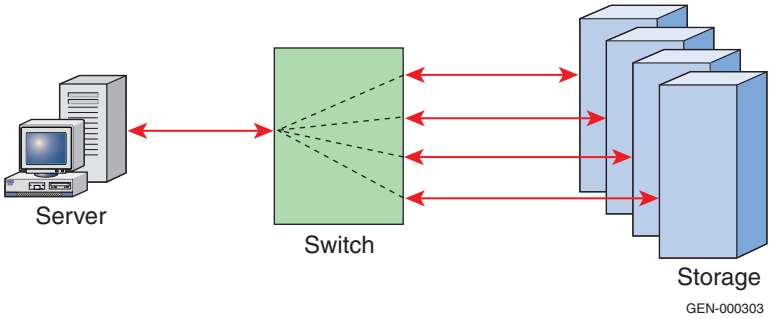


Figure 191 FC-Switched fabric capacity topology example

In Figure 191, assume that several host I/O slots are needed for tape, clustering, and networking connections, leaving only one slot in the host available for attachment to four storage ports. By using a Fibre Channel, Host Bus Adapter and a Fibre Channel Switch, the host can connect to a large pool of storage, minimizing the server I/O slot requirement.

Consolidation topology

Some environments contain low-capacity servers that must be connected to a high-capacity storage, expanding the required number of server connections. The consolidation topology provides a solution for this situation.

Both clustered and non-clustered applications are possible. Through host-based file-locking facilities, clustered hosts share information assets. Non-clustered application environments can share physical storage assets, capacity, bandwidth, and connectivity.

Channel failover and channel load-balancing software products (like PowerPath) can play a central role in providing manageability, reliability, and performance benefits.

Consolidation topology in the fabric environment

The logical concept of a consolidation topology in a switched fabric is described by the *fan-out* rate (*out* as in *out of* the storage system).

The consolidation topology uses a logical topology of fan-out. Figure 192 on page 711 illustrates a fan-out rate of four, where four server connections share a single storage port.

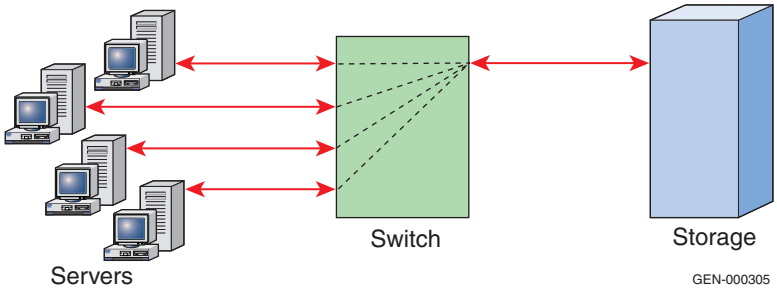


Figure 192 **FC-Switched fabric consolidation topology example**

Combined topologies

Topologies can be combined for maximum efficiency, achieving large storage capacity for many servers while minimizing the necessary number of host bus adapters and storage ports. Figure 193 shows how the three basic topologies can be combined to take advantage of the benefits of each.

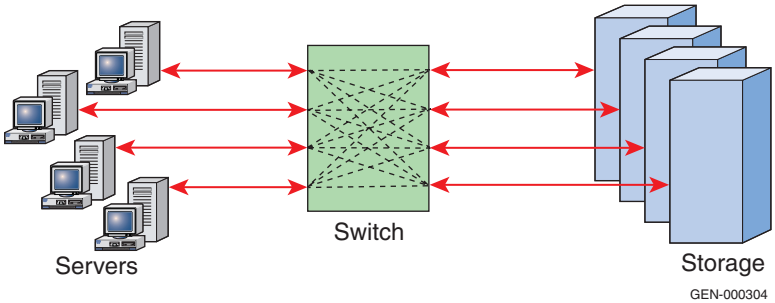


Figure 193 **FC-Switched fabric combined topologies example**

This type of configuration can take advantage of PowerPath and the redundancy of dual switches to provide alternative routes if necessary.

Combing physical and logical topologies

In the previous sections, the various physical and logical toplogy concepts were introduced. This section combines different pieces from of each of these topology concepts to higlight different design methodologies. As will be discussed, each methodology has its strengths and weaknesses.

The section starts with an introduction to the basic types of traffic that is present in a fabric and then hightlighst four different methodologies for managing these traffic patterns.

- "Traffic types" on page 712
- "Methodology 1: Balanced fabrics" on page 713
- "Methodology 2: Mirrored fabrics" on page 714
- "Methodology 3: Logical fabric segregation" on page 717
- "Methodology 4: Business unit fabrics" on page 718

An example of a combined physical and logical topology is shown in Figure 194, which displays a three-tier physical fabric that is also configured as a three-tier logical fabric. There are two hops between the server and the storage that is zoned to it. This fabric also has two separate same-cost, same-distance (hop) routes, indicated by the heavy blue lines. The thin black lines indicate ISLs that are in the fabric but are not currently being used for traffic. These unused ISLs are also referred to as *cold* secondary ISLs.

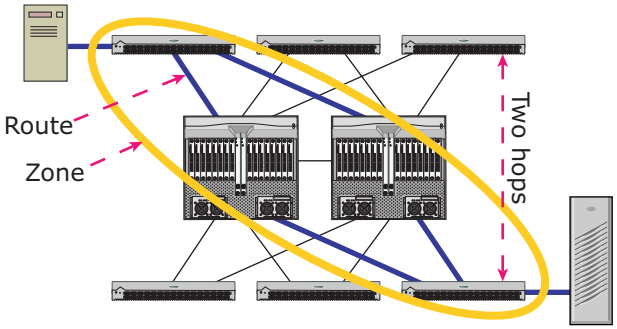


Figure 194 **Three-tier physical and logical fabric**

Traffic types

Note the following types of traffic regions that can be constructed to handle specific traffic and tasks.

Subscriber traffic

Subscriber traffic refers to that traffic associated with data being transferred between the application hosts and their storage.

Management traffic

Management traffic refers to that traffic associated with switch-to-switch traffic necessary to manage the fabric topology, user-generated traffic, or traffic from the management workstations

to the storage arrays or fabric controllers for the purpose of configuring the environment. This includes:

- Zoning changes
- Volume access / masking changes
- Fabric build messages

Business continuance traffic

Business continuance traffic refers to the data traffic generated through the use of a data replication application (SRDF or MirrorView) or the by the action of backing up, restoring, recovering, or retrieving files from a removable (tape) or near-line (digital vault / virtual tape) storage system.

Methodology 1: Balanced fabrics

Balanced fabrics, or balanced environments, are constructed by distributing the fabric resources (switches and ISLs), storage, and servers evenly across the environment so that the bandwidth and usage is also evenly distributed across each switch and ISL. No matter which specific fabric topology is chosen, you can create a balanced environment.

The first step in creating a balanced environment would be to create the fabric topology by laying out the switches in the desired pattern and attaching the ISLs evenly across the tiers of the fabrics. The aim is to create an infrastructure that allows equal bandwidth utilization for all points in the fabric. This design allows for better flexibility and more consistent performance as you increase the amount of shared storage access from your distributed hosts.

A further level of balancing can be accomplished by strategically using similar switch types (Connectrix M-series or B-series) and / or similar technologies (2 Gb / s or 1 Gb / s) in the fabric. Using similar technologies across the fabric can standardize and simplify later troubleshooting efforts. If the fabric is constructed of mixed switch types, balance can still be achieved by strategic placement of the switch technologies across the fabric for best utilization of the unique features and balanced bandwidth utilization.

The number of ISLs used should be based on both the desired level of redundancy and the estimated level of bandwidth. To create the balanced environment, the bandwidth requirement estimates should be evenly distributed across the entire fabric.

Lastly, the storage, and then servers, would be attached according to their access requirements and zoning relationships so that the

bandwidth requirements across each link would be managed evenly across all links.

In Figure 195 on page 714, each heavy line represents two ISLs. As shown, each switch has the same number of servers and, for the sake of the example, each server has the same amount of bandwidth requirements to each storage port. Storage ports would also be connected evenly across the storage level switches. Zoning and volume accessing would then be configured so that each link would carry an even portion of the total load. Fabric traffic would be constantly monitored and migrations of servers and storage could be made to perpetuate the level of bandwidth balance desired in the fabric.

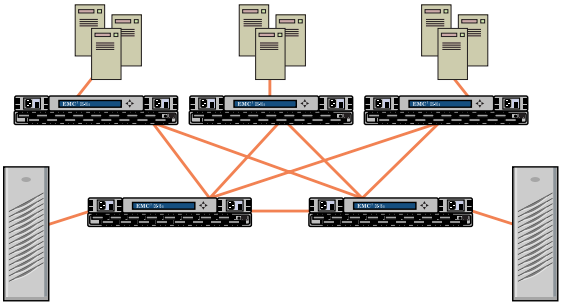


Figure 195 **Balanced fabric**

By creating the initial balance and maintaining the balance over time, you can produce an environment with predictable and deterministic traffic and fabric behaviors. Ensuring predictable fabric behavior facilitates identification of issues in the fabric components and may simplify troubleshooting. Because the data load is balanced across the ISL and switches, a more efficient use of fabric resources is achieved. Underutilized ISLs can be removed and free ports can be used for attaching more servers and storage.

Methodology 2: Mirrored fabrics

A mirrored fabric is an environment created by two or more distinct fabrics that, while they may share a common out-of-band management connection, are not directly connected to each other by fiber.

When mirroring an environment, start by configuring each physical device in the storage array so that it is presented across multiple storage ports. These storage ports would then be distributed evenly

across both fabrics. Each host accessing this storage would also have multiple ports distributed across both fabrics.

Volume access and device pathing to these storage arrays through the severs would be handled by load-balancing and path-failover software, such as PowerPath. PowerPath manages the representation of the multiple instances of the same devices to the host and manages the load balancing and path failover necessary to maintain the high availability of the environment.

Figure 196 shows how a server and storage array would be connected to both fabrics in a mirrored environment.

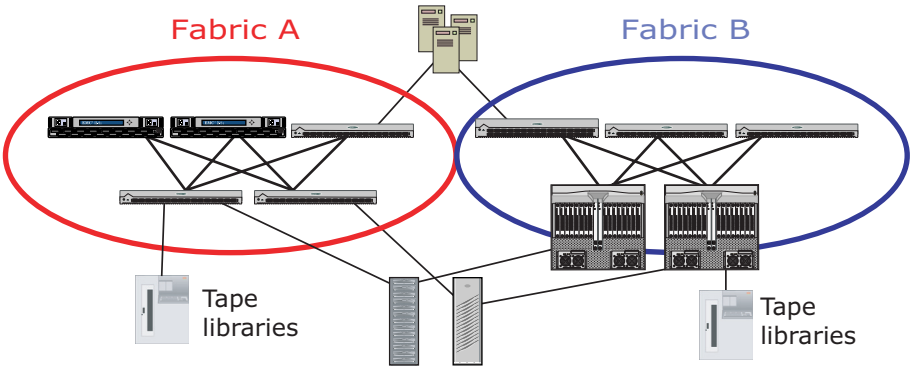


Figure 196 Mirrored fabrics

The decision to mirror a fabric can also be made independent of the individual topology used to create each mirror. Also, while EMC recommends that each side of the mirror be identical for easier management, monitoring, and performance characterizations, you can change the topology or switch vendor on one side of the mirror and phase in new technologies without any impact on the other side of the mirrored fabric.

Having different versions of firmware across a balanced fabric configuration does not pose any HBA, switch, or storage interoperability concerns, and is fully supported.

Balanced (redundant) fabric configurations can be created from any supported combination of switches. As long as the host and storage ports are supported with the switch firmware revision they will be connected to, there is no restriction on connecting host and storage ports to fabrics consisting of different vendor, model, or firmware revisions.

It is important to note that in homogeneous switch vendor environments, all switch firmware versions inside each fabric should be equivalent *except* during the firmware upgrade process.

For devices that have single attachment (such as tape or JBOD), EMC recommends that for the highest level of availability and distribution of workload, these devices be distributed evenly across both fabrics. Since each host in the environment has access to both sides of the fabric, connections to all devices are maintained. Backup environments can be configured so that device pools contain tape or disk devices from both sides of the mirrored environments.

To simplify maintenance and management, all additions to the environment should be made in pairs. For example, if you add storage access to Fabric A, ensure there is an identical addition is made to Fabric B.

Benefits of mirrored fabrics

Many individual benefits can be directly associated with creating a mirrored fabric environment. The following are some major benefits, but this should not be thought of as an exhaustive list.

- Increased availability through insulation and isolation — One of the most important benefits to implementing mirrored fabrics is the isolation your mission critical applications will receive from the increased buffering from fabric events. Issues that arise from component failures will be totally isolated to the single fabric experiencing the event. Mirrored fabrics can also be separated by distance so that a local room or floor outage or event will not affect the entire environment.
- Simplification to increase availability — By separating the environment into equal parts, you are able to manage much smaller units more confidently and effectively. Problem isolation becomes easier as there are fewer components in the environment under investigation at one time. Also, problems that occur on one side of the mirror are isolated from the other side of the mirror. Ssince each fabric is effectively a separate environment, you can maintain separate, smaller zone sets specific to each fabric.
- Increased effective domain counts in the environment — In general, EMC recommends that you limit your individual fabric designs to 31 domain or less switch topologies. By separating the required environment into two equal pieces, you increase the effective accessibility by having twice the number of switches and switch ports to work with.

- Increased N_Port availability — Since the number of switches is evenly distributed across two mirrored fabrics, the effective ISLs that would have joined the switches from both fabrics together can be freed up to attach more storage and servers. Reducing the infrastructure requirements associated with additional ISLs reduces the cost associated with the operation and management of the fabric.
- Flexibility — Since each fabric is independent of the other, you can perform maintenance and upgrade procedures on one side of the mirrored environment without affecting the other. The risk associated with updating firmware, recabling, implementing zoning, security, or other management changes, as well as physical movement of the devices, is greatly reduced because of the insulation benefits provided by mirroring fabrics.

  You may also use the isolation of the fabrics to phase in or leverage new hardware from different vendors, or hardware that uses more advanced technologies, with reduced risk to the entire environment. You may also choose to use the isolation to validate topology or management changes without affecting the entire environment. After changes are made to one side of the mirrored fabric, you can use the original design as a benchmark to contrast the performance of the changes.
- Centralized or distributed management — Management of both fabrics in the mirrored environment can be performed from either a centralized location or separate locations to further limit the dependencies. Management applications, such as EMC ControlCenter, can be used across single or multiple Connectrix service processors across multiple Ethernet segments.

Methodology 3: Logical fabric segregation

Logical fabric segregation is a set of design concepts that you can use during the planning stages to optimize a single fabric environment. While you normally think of a fabric in its entirety, the fabric can be logically segregated into separate traffic regions that are constructed to handle specific traffic and tasks. For example, a single fabric can be constructed so that subscriber traffic and business continuance traffic travel over the same fabric, yet over separate parts of the fabric.

Logically, segregating the fabric into different traffic centers limits the dependence of one storage task on another. Limiting the dependence makes management easier because the fabric resources can be designed, scaled, and maintained somewhat independently of one another. Troubleshooting is also simplified because the data traffic

related to any one task may be relegated to only a specific portion of the fabric.

As Figure 197 on page 718 shows, the subscriber traffic occurs in the area indicated inside the larger oval. SRDF traffic, MirrorView traffic, and SAN backup traffic all occur within the area indicated by the smaller oval.

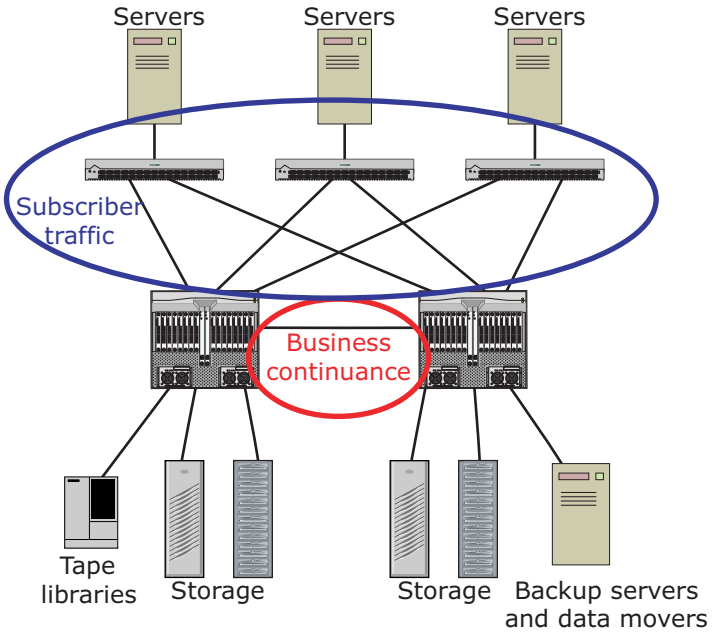


Figure 197 Logical fabric segregation

ISLs can be added or removed based on the specific bandwidth requirements of the given task.

Methodology 4: Business unit fabrics

Business unit fabrics are fabrics that are created to perform a particular function or process, or a set of business rules, bandwidth or availability requirements, support resources, or maintenance / duty cycles. Business unit fabrics can employ the benefits of both balancing and mirroring to further increase their robustness and availability.

Business unit fabrics are related to mirrored fabrics in the sense that they offer similar benefits in the isolation, limiting of dependencies of applications and processes on unrelated components, and the simplification of an environment. For example, one might implement

a unique fabric for a line of business that uses highly available Fibre Channel directors.

That line of business might have special uptime requirements. The added expense to use the new directors might be justifiable for this particular line of business, but might not be for the entire environment. The port-count requirements for this single line of business will be smaller, and thus require fewer directors. The converse is also true; there may be lines of business that require less guaranteed storage access and their requirements may allow them to work on a lower bandwidth set of switches or different topology.

Separating fabrics based on business units also allows each business unit to manage their maintenance, duty cycles, and upgrade schedules based on their availability. It will be much easier for the storage manager to coordinate the changes with the groups separately, rather than as a collective.

Business continuance traffic (backup, recover and restore, and remote data replication facilities such as SRDF and MirrorView) has been a prime candidate for business unit fabrics because of the differences in duty and maintenance cycles and the independence from application data (subscriber traffic). Customers have also chosen to build smaller test fabrics that allow them to investigate and validate components and applications before they are implemented into their production environment.

If you are considering the use of business unit fabrics in your environment, consider the following questions:

- Are all projects created equally based on their requirements for bandwidth, availability, and accessibility?
- Do the individual duty cycles and maintenance cycles of my individual business units differ to the point that they hinder one's ability to do business as efficiently as possible?
- Are there interdependencies between how the storage or servers are shared between business units that cannot be decoupled?

Fabric design considerations and recommendations

Note: Until this point, the material presented has been at the conceptual level and has been sparing in the actual design and setup detail. This section takes the first steps towards understanding how to design a solution. Throughout the next several sections, the design concepts will also have pointers to actual case studies containing cookbook setup steps.

This section describes the considerations and methodologies associated with the design and implementation of a Fibre Channel storage area network (SAN). The items detailed here represent a set of best practices that have been collected though SAN engagements with EMC customers and verified through testing in the EMC E-Lab.

While there is no substitute for tracking actual data traffic, application and storage utilization, and performance characterizations, there is still a need for determining a set of rules to approach the initial designing of a fabric. It is the objective of this section to outline some of the considerations involved in the designing of reliable fabrics and to provide some recommendations to managing a fabric's expansion over time.

It is imperative that after the initial fabric has been implemented, it be continually monitored to evaluate the success of the design and actual bandwidth utilization and to identify future needs for expansion.

To help design a fabric, the following topics are discussed in this section:

- "Fabric design considerations" on page 721
- "Common fabric topologies" on page 726
- "Nonstandard topologies" on page 739
- "Layout management" on page 739
- "Switch interoperability" on page 740
- "Multisite fabrics" on page 740
- "Tape connectivity" on page 743
- "General fabric design recommendations" on page 746

Fabric design considerations

Designing a fabric involves many variables that require consideration. With each variable consideration comes a separate design decision that must be made. Each design decision will help you create a fabric design that is appropriate for your business information model. This section discusses the following major considerations to aid you in evaluating how they are prioritized in your business:

- "Accessibility", next
- "Availability" on page 722
- "Consolidation" on page 724
- "Flexibility" on page 724
- "Scalability" on page 724
- "Security" on page 725
- "Supportability" on page 725

Note: Each of the examples presented in the "Common fabric topologies" section have been assigned a subjective rating. The subjective rating grade each topology on how well it meets each of the major considerations listed above.

Accessibility

Accessibility refers to the ability of your hosts to access the storage that is required to service their applications. Accessibility can be measured on your ability to physically connect and communicate with the individual storage arrays, as well as your ability to provide enough bandwidth resources to meet your full-access performance requirements. A storage array that is physically accessible, but cannot be accessed within accepted performance limits because of over-saturated paths to the device, may be just as useless as an array that cannot be reached physically.

Accessibility's link to available bandwidth leads us to consider the differences in building a statistical bandwidth infrastructure and a guaranteed bandwidth infrastructure. Guaranteed bandwidth infrastructures provide enough bandwidth resources for the full potential of the devices on the fabric to be used simultaneously. Statistical bandwidth fabrics are developed to handle only a fraction of the potential bandwidth.

An example of a statistical bandwidth network is the telephone system. The telephone system is not constructed with enough bandwidth resources to allow every subscriber to communicate simultaneously. You may have heard "all lines are currently busy; please try your call again later." This message indicates that the number of subscribers have saturated the bandwidth currently available, so no new connections are possible until resources are freed. Similar issues can arise in the design and implementation of a fabric.

You should also consider the internal design of the switching devices used in your fabric when considering accessibility. While switches may be designed for high levels of connectivity and allow many physical attachments, their internal designs may cause internal bandwidth congestion. (Refer to Appendix B, "EMC Connectrix Products Data," for additional information.) Even though you may develop a fabric topology that eliminates external congestion, your fabric may still be limited by the individual switches if they are designed with internal bandwidth congestion points.

Full *any-to-any* accessibility comes at a price. Increasing the accessibility of the fabric also increases the size of that fabric and the complexity of the topology. As more switches are needed to connect the new hosts and storage, more ports will be required for switch-to-switch communication. (Refer to "Determining customer requirements" on page 759 for scalability information.) You must be careful to ensure that our switch-to-switch bandwidth designs do not cause external congestion points in the fabric.

As the complexity of the fabric increases, more emphasis and reliance must be placed with the management application and adherence to well-defined management policies. Complexity can also impact how an environment is secured.

Availability

Availability is a measurement of the amount of time that your data can be accessed, compared to the amount of time the data is not accessible because of issues in the environment.

Lack of availability might be a result of failures in the environment that cause a total loss of paths to the device, or it might be an event that caused so much bandwidth congestion that the access performance renders the device virtually unavailable.

Availability is impacted not only by your choice of components used to build the fabric, but also by your ability to build redundancy into the environment.

The correct amount of redundancy will allow processes to gracefully failover to secondary paths and continue to operate effectively. Too little redundancy built into the fabric can cause bandwidth congestion, performance degradation, or (in some cases) a loss of availability.

When considering building redundancy into the environment, you must always weigh the *opportunity cost* of how much redundancy you need / want to build into the environment. Opportunity cost is the cost associated with the possible impact of not using the resources for other activities. For example, each extra port that is used for redundant ISLs cannot be used to attach more servers and storage to the environment.

Another concept that adds to the availability of an environment is *sparing*, which is the process of dedicating resources to remain unused until they are needed to take the place of a failed resource. You might also develop a sparing plan that indicates that even though a resource is currently in use, it may be swapped for a failed resource that has a higher priority.

The following must be considered in your redundancy and sparing plan:

- How much bandwidth do I need to preserve after a single event occurs?
- What other applications might be affected when the original storage resources move to a new path or down to a single path?
- Do I need to plan for scenarios that include successive failures?
- Do I want redundancy built into my connectivity components (as seen with director-class switching devices)?
- Do I want to build site redundancy and copy data to another site using the Symmetrix Remote Data Facility (SRDF) or CLARiiON MirrorView?
- Do I want to build redundancy at the host level with a load balancing and path failover application (like PowerPath)?
- How do I rank my business applications so that I can identify lower priority tasks, so these resources can be used as spares during a failure event? An example of this would be if task one had failed due to all of its fiber links being damaged and fiber links from task two were used to bring up the resources associated with task one. When the resources were back on line, both tasks would be working at 50 percent efficiency.

Consolidation

Resource consolidation includes the concepts of both physical and logical consolidation. Physical consolidation involves the physical movement of resources to a centralized location. Now that these resources are located together, you may be able to more efficiently use facility resources, such as HVAC (heating, ventilation and air conditioning), power protection, personnel, and physical security. The trade-off that comes with physical consolidation is the loss of resilience against a site failure.

Logical consolidation is associated with bringing components under a unified management infrastructure and creating a shared resource pool, such as a SAN. Logical consolidation does not allow you to take full advantage of the site consolidation benefits, but it does maintain site failure resilience.

Flexibility

Flexibility is a measure of how rapidly you are able to deploy, shift, and redeploy new storage and host assets in a dynamic fashion without interrupting your currently running environment. An example of flexibility is your ability to simply connect new storage into the fabric and then zone it to any host in the fabric. You can do this without any interruption in the I/O to any of the other hosts in your environment. Flexibility can also be seen in the ability of the Fibre Channel directors to perform code loads, component replacement, and insertion while the system is running without any noticeable impact to the hosts.

Scalability

Scalability is a measure of how easily a fabric can be extended so that it can accept more storage, more hosts, or more connectivity (switches). More storage and servers not only cover the physical connections required to attach these components, but also the internal and external bandwidth required to handle the actual throughput of these devices during usage.

External bandwidth is associated with the number of ISLs a switch can support, as well as the individual bandwidth of each ISL. Some switches now support both 2 Gb/s and 1 Gb/s ISLs. Other switches support a feature called *trunking*, which allows sharing the bandwidth resources of multiple ISLs simultaneously.

Fabric topologies that aggregate traffic into unbalanced scenarios should be avoided. One such scenario is when an intermediate switch has two ISLs coming in from the server-tier switches and only one ISL leaving towards the storage-tier switches. This scenario can lead to performance degradation, sometimes referred to as an *externally congested* architecture.

Each vendor's switch design has an associated internal bandwidth model which has a specific set of resources available to transfer data inside the switch. These internal resources may or may not aggregate bandwidth between the switching chips. Those architectures that do aggregate traffic inside the switch design may also represent an *internal congestion* architecture. Internal or external congestion architectures limit the scalability of the environment.

Scalability is enhanced by a switching component's ability to allow the online insertion of port expansion cards or additional optics. Allowing *hot* insertion of devices promotes the purchase and usage of partially-populated chassis that can be upgraded as the need arises. Partially-populated chassis can also play a role in the creation of redundancy in the fabric.

As the fabric grows, so does the ability of the switches to accept large numbers of zones and zone members. A fabric topology also plays a role in how quickly zoning or fabric changes take to propagate to all other switches in the fabric. Better fabric designs mean that your fabrics can recover from fabric events and propagate changes to the environment faster and with more predictability.

Security

Security refers to the ability to protect your operations from external and internal malicious intrusions, as well as the ability to protect accidental or unintentional data access by unauthorized parties.

Security can range from the restriction of physical access to the servers, storage, and switches by placing them in a locked room, or the logical security associated with zoning, volume accessing/masking, S_ID lockdown, or port binding.

Increasing the level of security directly impacts the flexibility of the environment. As security is increased, changes become more complicated. Whenever a new security policy is desired, it should be documented and reviewed for its impact on the accessibility, flexibility, and supportability of the environment.

For more information on security, refer to Chapter 4, ”Secure SANs.”

Supportability

Supportability is the measure of how easy it is to effectively identify and troubleshoot issues, as well as to identify and implement a viable repair solution in the environment. The ability to troubleshoot may be enhanced through good fabric designs, purposeful placement of servers and storage on the fabric, and a switch's ability to identify and report issues on the switch itself or in the fabric.

Fabric topologies can be designed so that data traffic patterns are deterministic, traffic bandwidth requirements can be easily associated with individual components, and placement policies can be documented so that troublesome components can be identified quickly.

The supportability measurement takes into effect the usefulness of internal error reporting, logging, and any diagnostic utilities that are shipped with the component. A product that is easily supported is an asset to an organization because of its ability to be brought back online without the time delays associated with shipping replacements or from scheduling visits for on-site service representatives. Many switches have the ability to identify issues and, through policy management, initiate automatic failover and recovery procedures. Some switches, as well as the Symmetrix system, also have the ability to identify issues and initiate call-home procedures to alert support personnel of the issue.

Common fabric topologies

This section provides information and recommendations on the following types of common fabric topologies.

For simple Fibre Channel SAN topologies:

- "Single switch fabric" on page 727
- "Two switch fabric" on page 727

For complex Fibre Channel SAN toplogies:

- "Full-mesh fabric" on page 729
- "Compound core/edge fabric" on page 731
- "Partial-mesh fabric" on page 734
- "Connectivity tier fabric" on page 736

Chapter 1, "Fibre Channel SAN Topologies" provides detailed implementation information for simple and complex topologies.

Simple Fibre Channel SAN topologies

A *simple* Fibre Channel SAN consists of less than four Directors and switches connected by ISLs and has no more than two hops.

Single switch fabric

A single switch fabric is the simplest of the simple Fibre channel SAN topologies and consists of only a single switch (Figure 198). The switch is also connected to a single management LAN through IP. For more information on single switch topology, refer to "Single switch fabrics" on page 50.

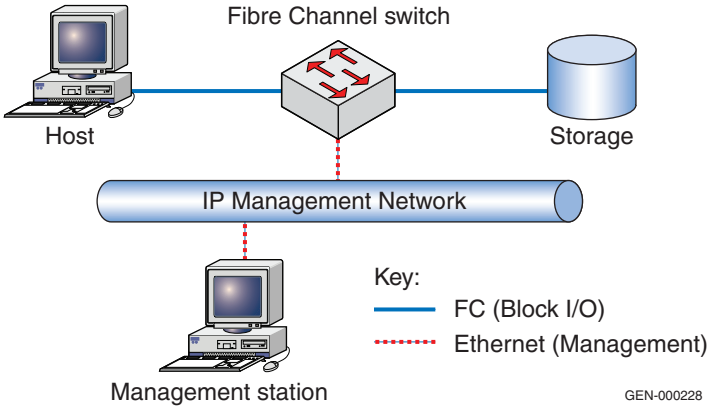


Figure 198 Single switch topology example

Table 28 shows the subjective rating for a single switch.

Table 28 Single switch subjective rating

	Most <----------> Least				
Attribute	5	4	3	2	1
Accessibility					
Availability					
Consolidation					
Flexibility					
Scalability					
Security					
Supportability					

Each of these attributes are discussed in more detail in the "Fabric design considerations" section beginning on page 721.

Two switch fabric

Figure 199 on page 728 shows an example of a two switch fabric, which is slightly more complicated than the single switch fabric. The switches are connected using two ISLs.

Both switches are connected to Management Network A. For more information on two switch topology, refer to "Two switch fabrics" on page 79.

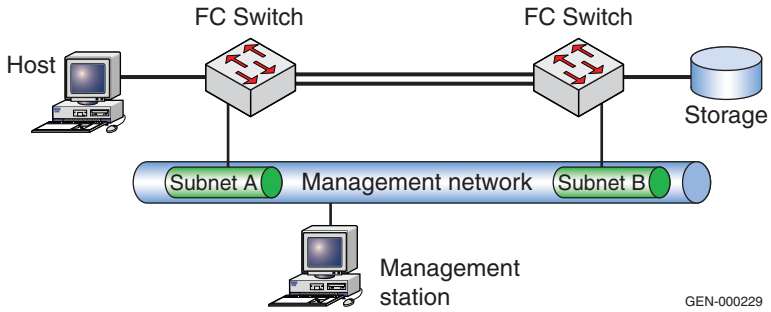


Figure 199 Two switch fabric example

Table 29 shows the subjective rating for a single switch.

Table 29 Two switch subjective rating

	Most <----------> Least				
Attribute	5	4	3	2	1
Accessibility					
Availability					
Consolidation					
Flexibility					
Scalability					
Security					
Supportability					

Each of these attributes are discussed in more detail in the "Fabric design considerations" section beginning on page 721.

Complex Fibre Channel SAN topologies

As the fabric size continues to grow, not all hosts will need one-hop access to all available storage. You may have a need to increase the storage and host connectivity while maintaining an efficient limit on the number of ISLs used in the fabric. You may also be searching for a way to add more storage accessibility to your hosts without major changes to your fabric design. These factors are all considered in the design of both a compound and a complex core/edge fabric.

A complex Fibre Channel SAN consists of four or more Directors and switches connected by ISLs and has any number of hops.

Full-mesh fabric

Table 30 shows the full mesh subject rating.

Table 30 Full-mesh subjective rating

	Most <----------> Least				
Attribute	5	4	3	2	1
Accessibility					
Availability					
Consolidation					
Flexibility					
Scalability					
Security					
Supportability					

Each of these attributes are discussed in more detail in the "Fabric design considerations" section beginning on page 721.

A *full-mesh* fabric is any collection of Fibre Channel switches in which each switch is connected to every other switch in the fabric by one or more ISLs. For best host and storage accessibility, EMC recommends that a full-mesh fabric contain no more than four switches. If you need to consider exceeding the four-switch recommendation, you should first investigate the other switch topologies described in this section.

A mesh may contain departmental switches, directors, or both, depending on your connectivity needs.

EMC recommends that when designing and implementing a full-mesh fabric you lay out the storage and servers in a single-tier logical topology design and plan your ISL requirements based on the assumption that 50% of the traffic on any one switch will remain local and the other 50% will originate from the remaining remote switches. This allows you a higher level of initial growth and resources sharing and gives you the following rule for optimizing your fabric performance:

> If you realize that more than 50% of your traffic is originating from remote switches, it may be time to migrate those components exceeding the remote access requirements to other switches so that they are either on the same switch or using a switch that is not fully utilizing its allocated switch-to-switch bandwidth.

If the initial remote storage access requirements are known, care should also be taken to distribute this load evenly across the remaining switches to equalize bandwidth utilization in the initial

design. Understanding your initial storage-to-host fan-out and relationships can greatly increase your success in implementing a design that will be subject to minimal flux.

Identification of the highest bandwidth hosts and ensuring that they remain local to their storage needs will also ensure efficient use of the available switch backplane bandwidth before utilizing your ISL resources.

Figure 200 shows examples of full-mesh fabrics.

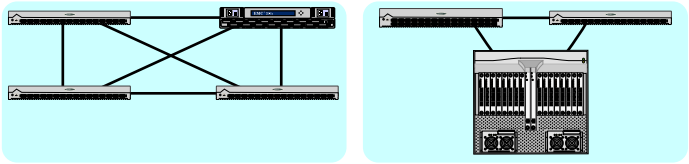


Figure 200 Examples of a full mesh

Benefits

Full-mesh configurations give you, at most, one-hop access from any server to any storage device on the fabric. This means that when you are adding or migrating storage or server attachments you have the greatest possibility of placing the server attachment and matching storage attachments anywhere in the fabric and achieving the same response time. Meshes also ensure that you always have multiple local and remote paths to the data even after fabric events have occurred.

No matter which building block you choose for your fabric, you must be mindful in planning the traffic patterns and loading to ensure that you are not exceeding the technology's ability to handle the traffic.

For more details on designing a fabric that can handle your data traffic, review these topics:

- "ISLs in the fabric" on page 751
- "Trunking" on page 868
- "Layout management" on page 739

Limitations

Scaling a full-mesh solution becomes complicated and costly with the increase in the number of switches and required ISLs to guarantee traffic performance. Without rules for sharing access to storage on remote switches or an ability to migrate components once traffic increases, ISL bandwidth can become over-utilized.

Compound core/edge fabric

Table 31 shows the shows the compound core/edge subject rating.

Table 31 Compound core/edge subject rating

	Most <----------> Least				
Attribute	5	4	3	2	1
Accessibility					
Availability					
Consolidation					
Flexibility					
Scalability					
Security					
Supportability					

Each of these attributes are discussed in more detail in the "Fabric design considerations" section beginning on page 721.

A compound core/edge fabric can be formed by merging two or more simple core/edge fabrics into a single fabric environment. Initially, hosts would be assigned volume access rights to storage that was available one hop away. Core switches that are one hop away from an edge switch in a compound or complex core/edge fabric are known as *primary core* switches. A core switch that is a primary core switch for one edge switch may also be a secondary core switch for an edge switch that is two physical hops away. Figure 201 shows an example of a compound core/edge fabric.

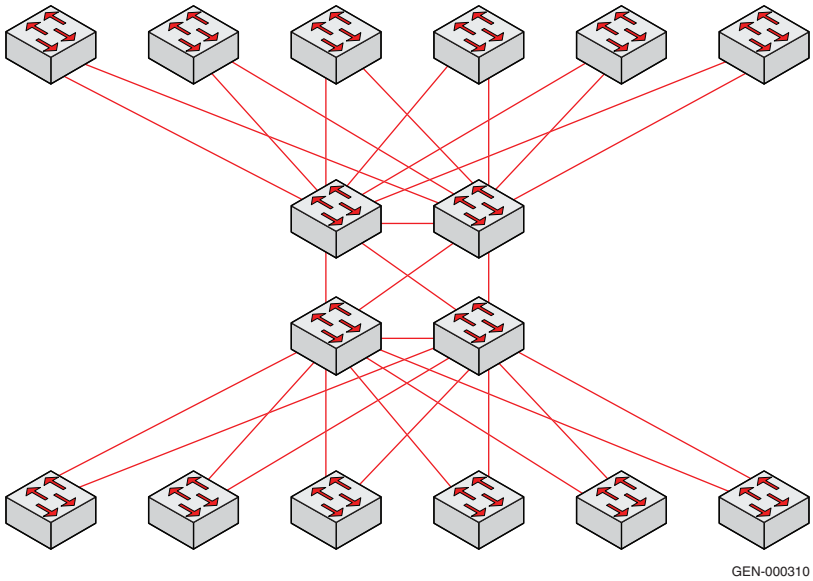


Figure 201 **Compound core/edge fabric example**

Once storage is exhausted for an edge switch's primary core switches, these hosts would then be assigned to storage that was two hops away on the secondary core switch. Access to the shared storage traffic on a secondary core switch would traverse the ISLs at the *back end* of the fabric. ISL planning for the compound core would be based on the both the subscriber traffic requirements at the *front end* of the fabric and the core-to-core bandwidth requirements at the back end of the fabric. Note that the front-end ISL requirements are independent of the back end ISL requirements. This independence ensures a more efficient approach to ISL planning and utilization.

To facilitate the performance of fabric management traffic, the principal switch would be assigned (and remain assigned) to those switches at the core of the fabric.

Benefits

The compound core/edge model maintains a robust, highly efficient traffic model while reducing the required ISLs and thus increasing the available ports for both storage and host attachments. It also offers a simple method for the expansion of two or more simple core/edge fabrics into a single environment.

By connecting the core switches from simple core / edge fabrics into a full mesh, you can easily create a compound core topology.

The compound core / edge topology creates a robust back-end fabric that can extend the opportunities for sharing of both backup and RDF resources.

The complex core / edge fabric inherits the benefits from both the simple core / edge and the compound core / edge designs. The complex core / edge increases overall fabric availability by limiting the effects on the edge switches from multiple failures on the core switches.

Since the edge switches are more evenly distributed, a failure of any two core switches would result in fewer accessibility impacts to edge switches and attached hosts.

Further availability can be added by spreading hosts across edge switches that are not connected to the same set of core switches. Note that switch failures are rare and unexpected. Multiple simultaneous failures of Fibre Channel components are rarer still.

Limitations

Both the compound and the complex core / edge design models produce a physically larger, tiered fabric which could result in slightly longer fabric management propagation times over smaller, more compact designs. Neither compound or complex core / edge fabrics provide for single-hop access to all storage.

While the potential availability of the complex core fabric is increased over other designs, the designs are more complex and may add to management and troubleshooting time if care is not taken to document the environment. Management concerns would be eased by the implementation of a comprehensive fabric management infrastructure.

Partial-mesh fabric

Table 32 on page 734 shows the partial-mesh subjective rating.

Table 32 Partial-mesh subjective rating

	Most <----------> Least				
Attribute	5	4	3	2	1
Accessibility					
Availability					
Consolidation					
Flexibility					
Scalability					
Security					
Supportability					

Each of these attributes are discussed in more detail in the "Fabric design considerations" section beginning on page 721.

A *partial-mesh* fabric is different from a full mesh in that each switch does not have to be connected to all other switches. However, to be considered a partial mesh, the fabric must be of a configuration where splitting it results in each new sub-fabric being a full mesh. Figure 202 provides an example of partial-mesh fabric.

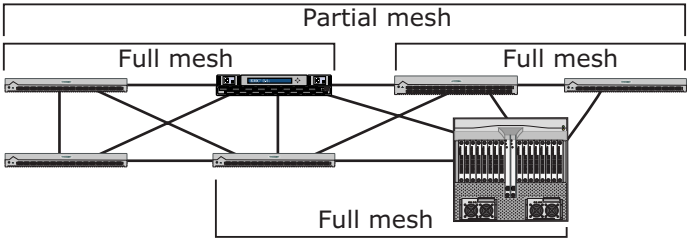


Figure 202 Partial-mesh example

EMC recommends that you use the same rules for developing a partial mesh as for a full mesh. Storage and hosts would be placed in the fabric starting with a single-logical tier approach, and additional storage access requirements would be placed one hop away.

Identifying the best selection from the pool of available storage should be evaluated based on your company's sharing, maintenance, and support procedures.

Servers with a majority of remote storage and storage that is heavily shared throughout the entire partial mesh should be moved toward the middle of the mesh. This should allevieate congestion and simplify pathing requirements evenly across the fabric.

EMC recommends limiting the size of the partial mesh to remain within the current boundaries for switch count, ISL hop counts, and ISL loading. This design has proved to be very useful and robust for configurations of five to eight switches and can be extended further by adding switches at the center of the partial-mesh fabric.

Both the managed switch (where zoning is activated) and the principal switch should be at the logical center of the fabric for best fabric response times.

Benefits

Partial-mesh designs offer extensive access to both local switch storage and single-hop storage. A partial mesh also extends the accessibility and provides many unique paths to the storage. Increasing accessibility while maintaining the same level of robustness is a design goal for every topology. Moving from many small, isolated full-mesh fabrics into fewer partial-mesh fabrics also simplifies the management paradigm by moving to fewer individual zone sets for the managed environment.

Moving to a partial mesh from two or more separate full-mesh fabrics is a simple logical progression. Many things will remain the same and only those items that are heavily shared across the new ISL boundaries will have to be migrated toward the center of the new fabric.

Partial meshes also offer a simple progression into a core/edge design. If you look at the center of the partial mesh as the core, you can create the new infrastructure by simply removing some of the ISLs at the outer edges of the fabric. Hosts would then be physically relocated to the edges and storage would be relocated at the core switches.

Figure 203 shows a sample six-switch partial mesh. Removing the two ISLs, as illustrated, could start the migration from a partial mesh to a core/edge fabric.

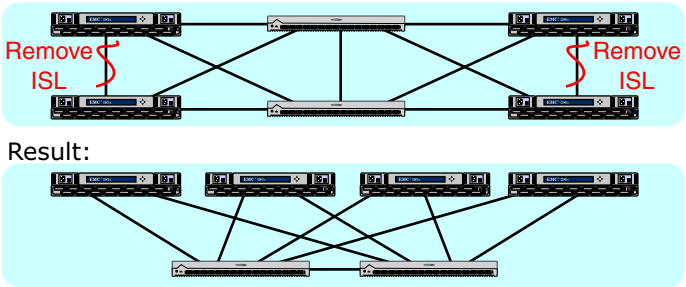


Figure 203 **Partial-mesh migration to core/edge fabric**

Limitations

Increasing the size of the fabric always increases the dependencies within the fabric. This does not cause a problem, but it does increase the complexity of troubleshooting and the impact on unrelated processes during a fabric event.

Without rules on placement of host and storage, traffic management in a partial mesh fabric can soon become cumbersome and ISLs can become overloaded due to excessive traffic aggregation across the individual physical switch tiers. Technology advances, like the introduction of 2 Gb/s switch functionality, may be unrealized because of poor placement of these new resources or ISL placement between these new resources.

Connectivity tier fabric

Table 33 shows the subjective rating for the connectivity tier.

Table 33 **Subjective rating for connectivity tier**

	Most <----------> Least				
Attribute	5	4	3	2	1
Accessibility					
Availability					
Consolidation					
Flexibility					
Scalability					
Security					
Supportability					

Each of these attributes are discussed in more detail in the "Fabric design considerations" section beginning on page 721.

Connectivity tier fabrics are based on the need to increase the accessibility between hosts and storage, and the inability of the switches at the core to support the core/edge model because of a limited number of ports per switch. The connectivity tier model creates a core of switches that are completely populated with ISLs.

The connectivity tier fabric is a three-tier physical and a three-tier logical fabric. This means that there are three switches from one side of the fabric to the other and that host to storage access traverses three switches. As for all designs, to increase the performance of the fabric management and event handling, the principal switch should be placed at the connectivity tier.

Figure 204 provides an example of a connectivity tier fabric.

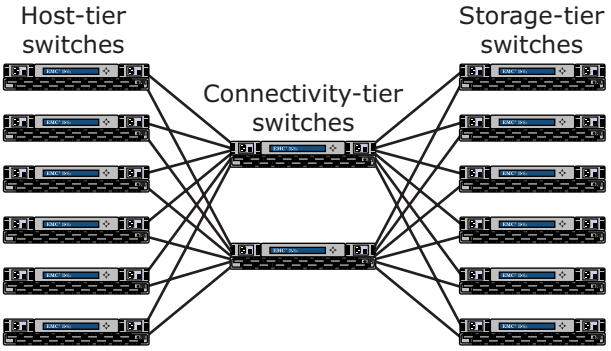


Figure 204 Connectivity tier fabric example

The tiers are:

- Host tier — The switches that are used to attach host HBAs into the fabric. Each host-tier switch should be configured using the layout recommendations for both ISLs and host attachments that are described under "Layout management" on page 739.
- Storage tier — The switches that are used to attach storage ports (disks, tapes, and so on) into the fabric. Each storage-tier switch should be configured using the layout recommendations for both ISLs and host attachments that are described under "Layout management" on page 739.

- Connectivity tier — The switches that are dedicated to aggregating and distributing the ISL traffic between storage-tier and host-tier switches. The fabric size is based on the number of ports available at this layer, as well as the individual switches' bandwidth capabilities.

  For best performance and availability, the number of ISLs coming from the host tier should equal the number of ISLs going to the storage tier. This will help to limit the chances of future traffic bottlenecks that could occur in the connectivity tier. This topology also lends itself to the placement of the principal switch in the connectivity tier, which promotes the propagation of the Class F traffic across the entire fabric.

  Zoning changes should also be initiated from the connectivity tier to facilitate the propagation of new zone sets across the fabric.

Benefits

This model allows you to extend the usage of your switches with fewer ports and increase the capacity of your fabric by aggregating bandwidth from the storage and fanning it out to the hosts through the connectivity layer.

Limitations

Without great attention to detail during the placement of component and assignment of storage access, the connectivity tier model can create severe traffic bottlenecks and eventually result in a fabric with a few overloaded ISLs, while others remain underutilized. Since these storage switches have a limited amount of resources to handle the traffic aggregated from the host tier, you may see longer latencies and lower performance with an unbalanced layout.

The connectivity tier model also offers only two-hop access to all storage. Straying from the original layout model would impact the accessibility benefits gained by moving to this model. Because of the dependencies this model has on the connectivity tier, losing a switch in the connectivity tier would have a large impact on the fabric performance.

Since connectivity tier switches are used only for ISL connectivity, your user port-to-ISL port ratio is lowered. Lowering the user port ratio increases the average cost associated with attaching storage and hosts into the fabric.

Nonstandard topologies

EMC does not recommend ring, string, or tree topologies because of their inherent flaws in availability, accessibility, and scalability. Single failures of components in these fabrics can result in switch isolation, resources inaccessibility, and poor performance. While these topologies can be configured, EMC highly recommends moving away from them and into a more robust and resilient design.

Layout management

Manual and purposeful managing of the layout of the individual fabric devices and ISLs can produce fabrics that are more stable, more available, less prone to latency, and less susceptible to traffic congestion. Managing traffic patterns also provides a set of rules for the systematic addition of components and promotes better ISL usage and planning, highest N_Port (storage and server) connectivity, better asset management, and apportioning analysis. You may not be able to implement everything that is being proposed in this section but the recommendations are listed below in the priority that they should be implemented.

General recommendations

1. Map volumes across multiple storage ports for availability.
2. When mapping storage ports that share paths to the same devices, or HBAs in the same host that share paths to the same devices:
   a. If mirrored fabrics are being used, attach the ports to different fabrics.
   b. If mirrored fabrics are not being used, attach the ports to different switches in the same fabric.
   c. If the ports must be attached to the same Fibre Channel Director, attach the ports to different port cards.
   d. If the ports must be attached to the same Departmental switch, attach the ports different ASICs on the same switch.

When mapping ISLs in your fabric:

1. Connect each switch to more than one other switch in the fabric.

2. Between any two switches there should be at least two physical connections for redundancy. Additional ISL requirements should be based on the actual/estimated performance data.
3. Choose one of the following:
   - If hardware trunking (such as with Brocade) is *not* being used:
     - For directors, connect ISLs to the same switch across different port cards.
     - For departmental switches, connect ISLs to different switch ASICs.
   - If hardware trunking *is* being used (refer to "Trunking" on page 868), you should also consider the redundancy benefits of creating two separate trunks as opposed to one larger one when planning trunking implementation.

Switch interoperability

Switch *interoperability* refers to the formation of a fabric using switches from multiple vendors. E-Lab Navigator contains fabric guidelines, the vendor switch models, and code versions that are supported in a heterogeneous fabric. For more information and implementation examples, refer to "Interoperability overview" on page 256.

Multisite fabrics

This section contains the following information related to multisite fabrics:

- "Catalysts for multisite fabrics", next
- "Creating multisite fabrics" on page 741
- "Multisite example: Application/backup and RDF" on page 742

Catalysts for multisite fabrics

Resource sharing

Connecting two or more sites into a single fabric allows you to create a sharable resources pool that includes all of the free resources available at each one of the connected sites. This is especially useful when looking for free disk storage or building a disaster recovery location. Sharing resources like hosts and storage among the different sites can increase your resources usefulness and overall company productivity.

Disaster recovery

Once connectivity is established between the sites, you might choose to create remote data facilities using the Symmetrix RDF or CLARiiON mirroring applications. More and more customers are moving to a centralized SAN backup architecture that allows them to locate their tape storage resources in one physical location.

Backup centralization

Creating a centralized backup location where both the trained backup administrators and the backup media resources are located can increase the duty cycles of your backup media, as well as increase the efficiency of your personnel resources. Centrally locating the personnel fosters communication, facilitating access to, and transfer of, skills among the group.

Digital vaulting

Digital vaulting is the procedure of performing daily backups over an extended-distance link to a remote data facility. The extent of your digital vaulting can included both the original backup and a clone, or it might include only the tape clone. Administrators might even choose to perform a backup to a disk that is part of an RDF group. This disk information would then be vaulted to a remote site's R2 device.

Digital vaulting can employ any combination of RDF, BCVs, or other backup methodology that best fits your needs for backup and recovery. No matter which procedure is used, the time and expense savings associated with switching to a digital vaulting paradigm should be evaluated in your business model.

Creating multisite fabrics

Fabrics can be created that can expand beyond the physical constraints of a room, a floor, a building or a site. Multimode Fibre Channel links support devices over 500 meters apart at a 1 Gb/s speed. When the speed is increased to 2 Gb/s, the distance is reduced to 250–300 meters. If your needs go beyond this distance, you may decide to use either longwave optics or fabric-extending DWDM technology. E-Lab Navigator lists the support distances for longwave optics and the list of currently supported DWDM vendors.

EMC recommends using longwave and DWDM connectivity to extend the length of an ISL. Extending the distance of ISLs gives you the greatest efficiency on your long-distance investment by allowing you to share the ISL link among multiple storage devices and multiple hosts devices on either side of the distance-extension device.

Sharing the ISL between multiple devices maximizes the usage of the link, and increases the usefulness of the distance extension components.

Whenever extending the distance of a device or an ISL, you should be aware of the buffer-to-buffer credit requirements to maintain data transmission performance over the link.

Multisite example: Application/backup and RDF

Mirrored fabric with DWDM across the mirrors

Figure 205 shows how two sites may be connected by DWDM equipment using a portion of their mirrored environment at both sites.

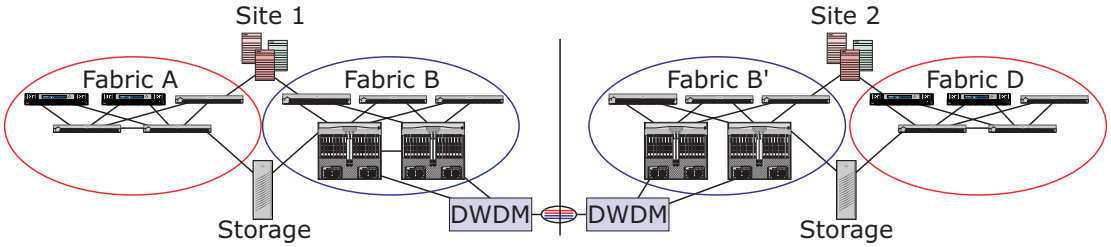


Figure 205 DWDM across mirrored fabric

While linking the mirrors on both sides creates a single larger fabric from fabrics B and B', it does not require additional switch resources at either site. This topology would allow direct sharing of storage resources between both sites, as well as RDF traffic between the two sites. Data traversing from site 1 to site 2 travels across the back end of fabric B to the back end of fabric B'. This will eliminate any impact on the subscriber traffic at one site during the sharing of resources from the other site.

ISLs could also be added from fabric A to fabric D to create another multisite fabric (A+D). Leaving them separated limits dependencies on the long-distance links between sites. An event on the long-distance links that causes a flux in the system would affect only fabric B+B'. Fabrics A and D would be isolated from the event.

RDF extended fabrics

Figure 206 on page 743 shows how to configure a mirrored fabric environment and reduce the dependencies so that only the RDF traffic traverses the link. For accessibility and performance, EMC recommends that Fibre Channel switches be placed on both sides of the DWDM equipment.

This ensures that there are enough BB_Credits to handle the bandwidth over the required distance and that there is enough bandwidth aggregation to efficiently use the long distance Fibre Channel links.

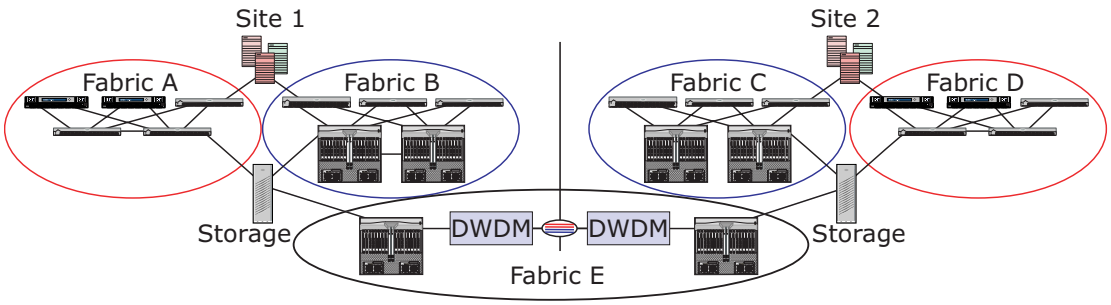


Figure 206 RDF extended fabric

Tape connectivity

EMC supports attaching tape drives into the fabric using Fibre Channel/SCSI bridges, FC-AL/FC-SW bridges, FL_Port-capable switches, or native FC-SW-capable tape drives.

Though tape drives may be placed anywhere in the fabric, EMC has developed some best practices resulting from qualification efforts and customer testimonials. These recommendations are based on optimizing the ability to share the tape drives with multiple hosts, as well as ways to minimize the impact of tape backup traffic on the application traffic. This section will further discuss:

- "Tape drive-sharing model", next
- "Traffic reduction model" on page 745

With any topology that you choose, EMC recommends that the design be mirrored and balanced whenever possible. Since tape drives are supported as single attached devices, you should consider spreading your tape drives across both sides of the mirrored fabric for highest availability of the backup solution. Your backup software solution should be evaluated for the best way to create pools of devices that can include tape drives from both fabrics.

PowerPath scripting should also be evaluated to investigate how setting an HBA which is shared to both tape and disk into standby mode may boost the performance of your backup.

Setting a path to standby still allows PowerPath to use it in case of a path failure, but would not use it during normal conditions for disk traffic. This procedure is especially useful for hosts that do not have enough available slots for a dedicated tape (backup) HBA.

Tape drive-sharing model

To promote sharing of any devices, you would place those devices at the center (core) of the fabric. This would allow all hosts on the fabric to have equal, deterministic access, equal bandwidth performance, and equal latency to the shared device. The farther a shared device is from a host, the more traffic aggregation that host must contend with to reach and access the device. If only a few media servers are solely responsible for writing the data to disk, these media servers would be placed on the same switch as the tape drives, or as close to the tape drives' switches as possible.

Figure 207 shows an example of a tape drive-sharing model of a fabric where a core-to-edge fabric topology has been deployed.

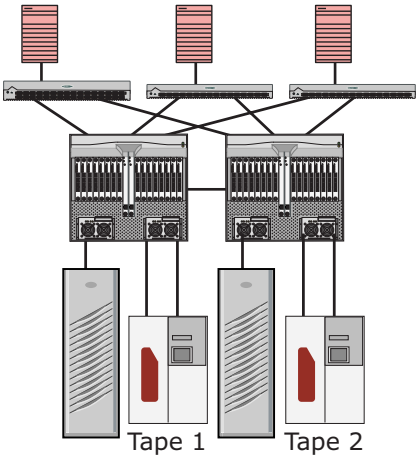


Figure 207 Tape drive-sharing model

With the advancement of server-free backups, EMC recommends that the copy engine, storage ports, and tape drives be located on the same switch or at the back end of the fabric.

You can construct a front-end / back-end design, as done in Figure 207, so that backup data will traverse only the core-to-core ISLs in a server-free backup environment.

Traffic reduction model

If tape devices are highly dedicated and seldom shared across servers, placing them on the same switch as a tape-capable application server or on the same switch as the media servers can help reduce the traffic associated with backups.

Data must be:

1. Read from the disk storage at the core.
2. Pulled across the core to edge ISL.
3. Brought into the backup (media) server.
4. Written out one of the tape accessible HBAs.
5. Sent back over the ISL from the edge to the core.
6. Written to tape.

If you were concerned only with the amount of ISL traffic, you can see that the data has traversed the ISLs twice. Though this is a powerful fabric design, you can see how the number of traversals could grow as you move to higher-tier fabrics like the partial-mesh fabrics or connectivity-tier fabric topologies.

In Figure 208, you have moved the tape devices out toward the edges on the same switches as the tape-accessible hosts (application servers or media servers). In this model you have to traverse the ISLs only once during the operation where you are initially reading from the disk array.

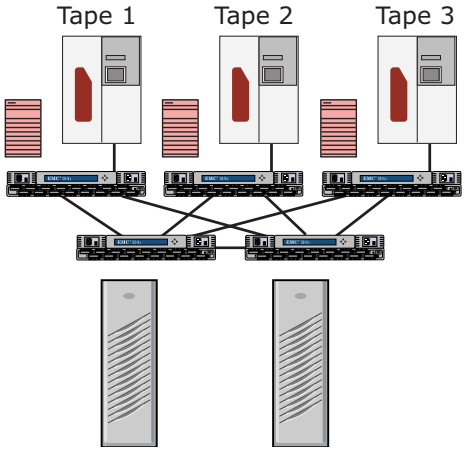


Figure 208 **Traffic-reduction model**

Sharing tape storage in this model between hosts becomes more costly because you would have to make three ISL transitions to use a tape that is not connected to the local switch. This design would also not reap the full benefit of an environment that was employing a sever-free backup solution.

General fabric design recommendations

Through testing and customer experience, EMC has compiled a list of general fabric design recommendations. These recommendations can be used no matter which switch vendor you are using or which topology and protection schemes you have employed at your site. Recommendations are discussed for the following areas in this section:

- "Physical topology hop count" on page 746
- "Maximum switch count per fabric" on page 747
- "Maximum number of Nx_Ports in the fabric" on page 748
- "Unique Domain IDs" on page 750
- "Fabric mode" on page 750
- "ISLs in the fabric" on page 751
- "Trunking" on page 868
- "Principal switch placement" on page 755
- "Zoning" on page 755
- "Port/cable sparing" on page 756

Consult the E-Lab Navigator for the most up-to-date information.

Physical topology hop count

For all EMC-qualified switches, EMC recommends that you build your initial fabrics so that all components are within three hops. If you were to build a larger fabric, you may find your fabric unable to maintain the fabric topology or that it drops frames after a failure of one or more components.

The failure of one or more components may change the topology of the fabric in such a way that you are exceeding the time a frame is allowed to stay active within the fabric. Changing the time-out values can cause undue congestion in the fabric if frames were to be allowed to remain active for long periods of time without reaching their destination. Fabric recovery and change response times are also based on these values and altering these values without a full

understanding of the consequence could result in unexpected fabric behavior.

One such example of a fabric change would be a zone set merge due to the attachment of two operational fabrics. Once the fabric parameters and link parameters were verified and the link came up, the zone set merge would propagate from the initial point of connection through the adjacent switches. Once an adjacent switch successfully merged the zone set, it would then propagate the merge request to its adjacent switches. The time that it would take to do this increases as the distance from the initial connection point increases.

If the placement of components around the fabric becomes unmanaged, the traffic patterns will become nondeterministic and as the size of the topology increases, the configuration may become more susceptive to bandwidth congestion problems. Maintaining a densely-configured fabric of three hops helps prevent nondeterministic traffic and traffic congestion issues.

As technologies and software advances are made, EMC will continue to extend the solution environment envelope. E-Lab Navigator describes changes to these recommendations.

Maximum switch count per fabric

As with the number of hops in a fabric, the number of switches in the fabric also affects the operation and timing of fabric responses. Fibre Channel switches communicate among themselves to provide distributed services for zone propagation, name resolution, and switch identification. As the number of switches increases, so does the amount of data that is propagated and the amount of components involved in the negotiation and synchronization of this data.

Since the effects of fabric size are very important to the continued reliability of the fabric, EMC continually spends a great deal of effort identifying and progressing the current fabric limits.

Fabric size increases are qualified in the same detailed fashion as would be the introduction of a new switch model or new firmware revision. Increases in the number of switches in the fabric are tested within the specific topologies listed in this document to continually validate and improve on our design practices.

Currently, EMC recommends that a maximum of 31 switches be used in a single fabric. E-Lab Navigator describes changes to these recommendations as well as the switch types capable of this fabric size.

Maximum number of Nx_Ports in the fabric

The single largest limiting factor in building large fabrics is the number of Nx_Ports that will be attached. This is due to the amount of information that must be passed between all distributed name servers in the fabric each time a new Nx_Port is added or removed.

One or two ports entering or leaving the fabric usually has a very small impact, but when you start adding and removing an almost fully-populated director to the fabric, the impact can be enormous.

Figure 209 on page 748 and Figure 210 on page 749 illustrate the impact to name server response time during a Zone Set activation and a Domain Add. The testing was being performed as a part of the validation work done to verify that a newer version of firmware from one of EMC's switch vendors could handle up to 2048 ports in the same fabric. The zone set activation chart (Figure 209) shows that on the older version of firmware, the name server would take up to 6 seconds to respond to NS Queries when a zone set was activated and 1530 Nx_Ports were in the fabric. This same chart also shows that the response time dropped to less than a second with almost 2000 Nx_Ports.

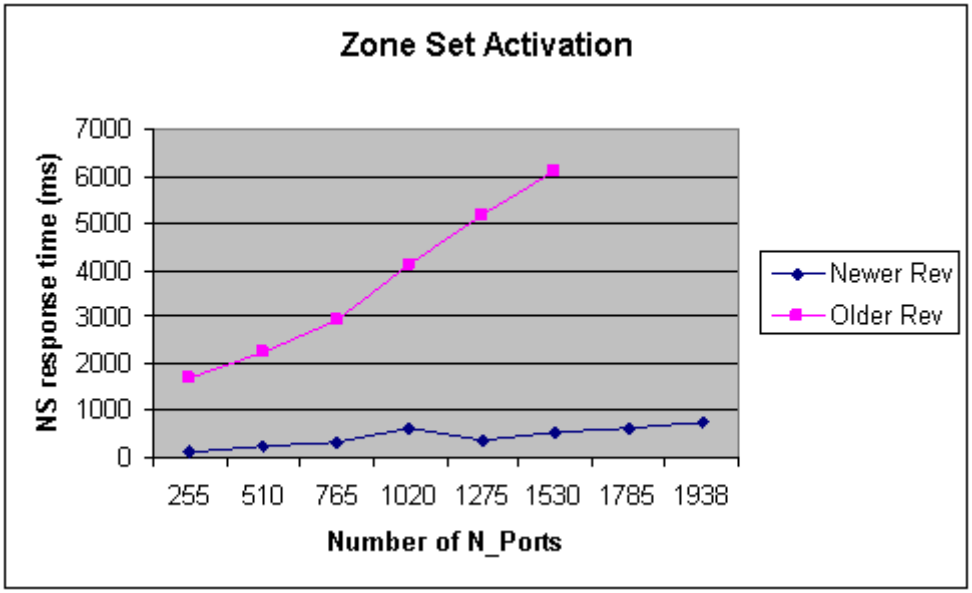


Figure 209 Zone set activation

The Domain add chart (Figure 210) shows that on the older version of firmware, the name server response time was 10 seconds when adding the 512th through 765 devices. The same testing on the newer version of firmware shows that up to 2000 ports can be added and response time remains less than 10 seconds.

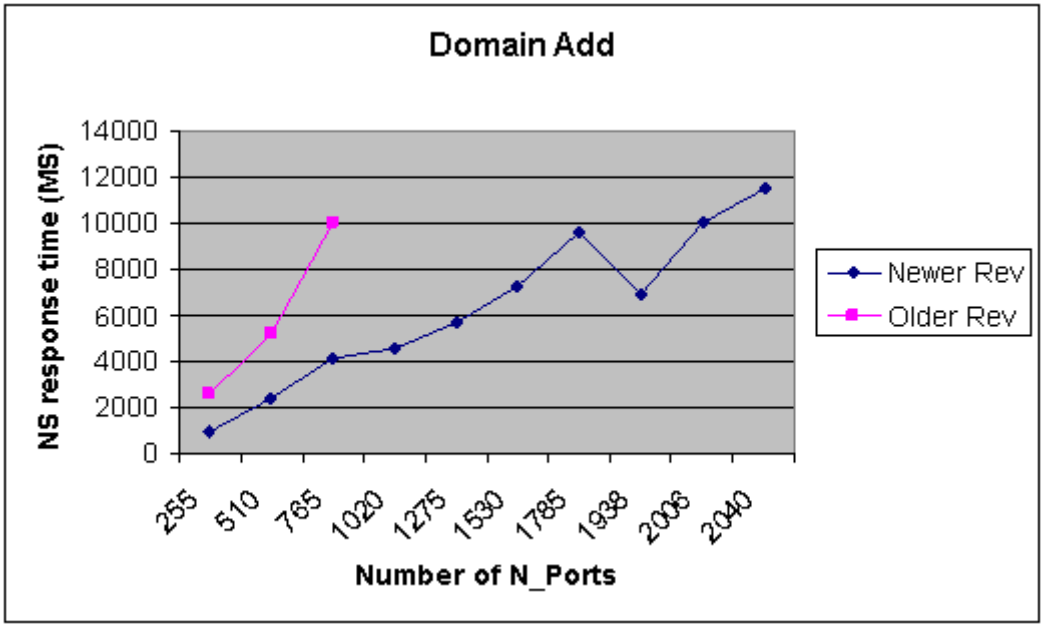


Figure 210 Domain add

Later versions of firmware from this vendor also showed improvements that allowed us to support up to 2048 Nx_Ports in the same fabric.

You may be asking why E_Lab cares about Name Server response time. Two answers spring immediately to mind.

- Name server response time impacts the amount of time it takes for two Nx_Ports to establish or re-establish connectivity through the fabric. This establishment or re-establishment of connectivity is also known as *discovery*. While it is not normal for host and storage pairs to be utilizing the name server to perform discovery during normal operation, it is possible that discovery will be invoked as a part of some error recovery sequence during a

domain add or zone set activation. If this were to happen and the name server was unresponsive for long periods of time, the outcome would be unpredictable.

- Name server response time is one way to measure how busy a fabric is. The more busy a fabric is during a simple zone set activation or domain add, the higher the likelihood that the fabric will have a more difficult time recovering from error sceanarios that are not as straight forward. The amount of time it would take to perform the recovery would also be impacted.

Unique Domain IDs

EMC recommends that you always attempt to give each switch a unique preferred Domain ID, whether it is targeted for an existing fabric or a new fabric. (The exception to this rule would be if you were swapping a new switch for a failed switch.) This provides the best opportunity to merge a switch or fabric into another fabric with the least amount of manual intervention.

Since Domain IDs become part of an HBA or storage array port's fabric address, and data is transferred to and from devices based on their address, Domain IDs cannot be changed without first bringing the switch to an offline state. Every switch that is currently online will have been assigned a Domain ID from the principal switch. If the fabric contains only one switch, that switch is the principal switch for that fabric. Domain ID negotiation done during a fabric merge will not change the current Domain ID when the switch is connected to the fabric while online.

During the exchange of parameters following a merger event, if the switches identify a domain conflict, they will not merge and the ISLs connecting them with remain segmented. Taking a switch that has segmented from the fabric due to a Domain ID conflict offline and then bringing it back online allows the switch to renegotiate for a new unique Domain ID. It should be noted that taking a switch offline terminates communication between the devices that are on that switch. EMC recommends that if a conflict does exist that you change the preferred Domain ID of the switch yourself so that you can ensure that it is unique in your entire environment and that it will always attempt to use the same ID, even after future merges and fabric changes.

Fabric mode

Fabric mode refers to the switch code that is utilized to operate the switch-to-switch communication and distributed processes involved with building and maintaining a running fabric.

Originally, each vendor created their firmware with a slightly different interpretation of the Fibre Channel standard. This meant that they would not be able to operate natively in a heterogeneous switch vendor environment.

Changes were made and a new *Open Fabric* mode was created that results in a common interpretation and implementation of these required switch-to-switch communications. To operate in a heterogeneous switch environment, all switches in the environment must be set to Open Fabric mode.

Changes to the fabric mode of a switch that allow it to operate in a single vendor fabric or a heterogeneous vendor fabric changes not only the fabric mode, but also the switch address. Both of these changes mandate that the switch first be set offline before the change is made. Once the change is made the switch may be brought back to an online state.

Since the change of the mode is an offline change, you should consider how to set the mode during the initial installation of a new switch or new fabric.

E-Lab Navigator contains more information on the specific supported code combinations between vendor switches.

ISLs in the fabric

Interswitch links (ISLs) are used to transfer host-to-storage data, as well as fabric management traffic from one switch to another. When planning the quantity and locations of ISLs around the fabric, one should consider the effects of routing, redundancy requirements, ISL load balancing, and trunking. Following a brief discussion of each of these, ISL design guidelines are provided.

Intra-fabric routing

Not to be confused with inter-fabric routing, intra-fabric routing describes the process used to route frames from an ingress port (entry point) to egress port (exit point). Frames are routed across the fabric using an algorithm called Fibre-Shortest-Path-First (FSPF) routing. Each ISL in the fabric is assigned a cost based on the link speed it is currently operating at. FSPF calculates the shortest distance between an ingress and egress port by finding the path with the lowest cost. The path with the lowest cost will typically be the one(s) with least number of hops, but this may not always be the case when mixing different line rates (4 Gb/s and 1 Gb/sec especially) in the same fabric.

EMC strongly recommends that you construct your fabric to have multiple equal, lowest-cost, shortest-path routes between any combination of host and storage. This means that you may have two ISLs between every switch in the fabric or you may have single links between switches, but would still have multiple equal-cost/length paths that travel through different switch combinations in the fabric.

Routes between storage and hosts that are not the shortest, lowest-cost path will not be used unless there is an event in the fabric that causes them to become the shortest, lowest-cost path.

Routing tables on each switch are updated and recalculated during events that change the status of links in the system. The calculation of routes and its ability to perform this function in a timely fashion is important for the stability of the fabric. For this reason, as well as for the fact that for every ISL configured, two ports are lost to attaching storage and hosts. EMC recommends using reasonable limits on the number of ISLs in a fabric. Since only the lowest cost/shortest-path routes will be used in the routing scheme, many ISLs may remain unused while others may be approaching peak utilization as long as there are no events in the fabric. For a true number of required ISLs, you should continually examine your ISL utilization and identify your level of actual protection from fabric events. You may be able to identify how changes in your current topology could enhance your performance as well as your path redundancy.

Routes are assigned to devices for each direction of the communication and the route one way may differ from the return route. The routes are assigned based on a round-robin approach that is initiated as the device is logged into the fabric. These routes are static for as long as the device is logged in or routes do not have to be recalculated due do a fabric event.

Redundancy requirements

Redundancy is added to the fabric to protect your environment from possible events and failures that may occur. The amount of redundancy that you choose to add to the fabric is a decision based on your business model and the amount of resources you can spare for increased availability.

EMC recommends a minimum of two ISLs between any two switches for redundancy. This also allows a certain level of traffic balancing across the ISLs and maintains a same-cost route in the fabric if one of the links fails.

EMC also recommends connecting each switch to more than one switch for redundancy. Having each switch connected to more than one switch in large configurations ensures multiple paths to the edge switches if one of the intermediate switches or paths to those switches fails.

If you are limited on the number of ISLs you can place in the fabric, EMC recommends configuring a single ISL from the originating switch to multiple switches if both paths from the hosts to its storage could be counted as equal, shortest-path, lowest-cost paths.

These recommendations can be used separately or combined for even higher availability. The ultimate goal is to provide each host multiple primary paths to its storage and then a number of secondary paths if failures were to occur. ISL utilization should always be monitored to identify unused, underused, or over utilized ISLs. Unused ISLs could become candidates for removal if they do not represent the only secondary path a host would have to its storage in the event of a switch or ISL failure.

E-Lab Navigator lists the ISL limits for individual switch vendors.

ISL load planning

There is no substitute for real traffic analysis and identification of burst patterns in your information model; however, there is a set of initial design rules that can be used to implement your initial fabric design.

When designing a new fabric from the ground up, use the 6:1 rule to get a general feeling for the number of ISLs that will be required between two domains. The rule states that for every 6 Nx_Ports that will be routed between two domains, there shoud be 1 ISL.

For another approach, consider the following.

Performance on a storage port is highly dependent on the number of I/O requests per second and the size of the I/O in the request. This will vary from customer to customer and from process to process. It has been found that a reasonable design guideline can be estimated when using the EMC fan-out recommendations for storage ports. EMC currently recommends two Symmetrix Fibre Channel director groups per ISL for initial fabric planning. (A Fibre Channel director group is the set of host HBAs that are communicating with a particular storage port.) This estimate uses as its basis the 1 Gb/s dual port/dual processor boards. Using the connectivity-enhanced four port/dual processor boards, the recommendation changes to four Fibre Channel director groups per ISL for initial fabric planning.

As an example, if the fan-out of a storage port is 12:1 (12 HBAs to one storage port), you would plan to allow the traffic from two storage ports to traverse from one switch to another over a single ISL. For redundancy reasons, you should add one ISL for increased availability. When considering SRDF ports in your design, the initial design estimates are similar. EMC recommends that for initial design purposed you plan to use two local SRDF ports per ISL.

After the design has been implemented, all ISLs should be monitored and readjusted based on the actual results of the fabric over time. Data bandwidth requirements continually grow, so they should be monitored and tracked with care. Regular monitoring will assist in planning future needs with greater confidence and accuracy. It should also allow the lead time required to better plan for the acquisition and implementation requirements of the new equipment.

Planning ISLs for tape backup is slightly more complex. Because individual tape drivers do not have the load balancing, path failover, or random simultaneous host access potential that disks have, you need to consider only their individual access potential and your redundancy concerns. The redundancy concerns can be satisfied by spreading out our total complement of tape drives across multiple switches. This will ensure that you do not have all of your tape drives traversing over the same single link.

To gauge the additional bandwidth requirements for tape backup, you can use the bandwidth specifications for the tape using the vendor's maximum compression numbers. This will give you the most conservative effort for allowing for the greatest possible bandwidth requirements. Once you have this number, you can simply take 80 percent of the ISL bandwidth and divide it by the bandwidth of the tape with maximum compression. This will give you the number of tapes that can be streaming across the ISL at one time.

Combining the outcomes of the SRDF, storage port, and tape drive bandwidth numbers will provide a high-confidence estimate for the quantity of ISLs required between the switch tiers.

E-Lab Navigator contains the latest information on the number of host HBAs that may be zoned to a single storage port by OS platform.

See “Determining customer requirements” on page 759 for more information.

For ideas on how to gauge your current traffic patterns and traffic requirement, you can examine the bandwidth reporting available in EMC ControlCenter and Connectrix Manager.

ISLs would then be added based on your plans for a balanced fabric and your actual bandwidth requirements. Loading plans should be such that the no ISL in the planning stage is scheduled to exceed 80 percent of its actual bandwidth potential.

Designing for greater bandwidth could result in an inability to compensate for bursty traffic or extrapolation errors that may have occurred during the planning stages.

Principal switch placement

Domain ID negotiation is governed by the entity in the fabric called the *principal switch*. Principal switch selection is made based on the switch with a combination of the highest principal switch priority and the lowest World Wide Name.

Proper placement of the principal switch in the fabric can lead to short negotiation times, resulting in the fabric returning to a normal state in an shorter amount of time. For these reasons, EMC recommends that the switch at the logical center of the fabric be made the principal switch. This would normally mean the switch with both the least amount of hops to the farthest extent of the fabric and / or the switch that has connections to the most other switches in the fabric.

These two placement strategies will help to ensure that your principal switch access times are as quick as they can be.

Zoning

To an end user, zoning is one of the most critical concepts to fully understand, yet it seems to be the most widely misunderstood. Zoning is a generic term that is commonly used to describe either the process of grouping ports together so that they may access each other over the fabric *or* the effective zoning configuration on the fabric.

The process of zoning consists of identifying host and storage pairs, adding them to a zone, adding this new zone to a zone set, and then activating the zone set onto the fabric where the host and storage pair resides. Once the zone set is activated, the host and storage pair can communicate with each other

General guidelines for zoning

- Use a consistent and detailed zone and zone set naming convention.
- Use WWPN-based zoning and not port zoning.
- Configure each zone so that it only contains one initiator.

- Practice good SAN hygiene and frequently check for "dead" zones.

Port/cable sparing

Sparing needs are associated with the resources you are leaving unused as well as the resources that can be removed and reused based on your business's processed prioritization model. Only you can decide what can be reused for another process based on your business prioritization model. Therefore this section will focus on what issues need to be considered when calculating for true spares, as well as which resources should be spared in a fabric environment.

Cable sparing

The data center is sometimes a very busy and dynamic environment. New equipment implementation, troubleshooting techniques, and everyday data center activities can place the Fibre Channel cabling at some risk of damage. Whether you are using patch panels for easier cable change management, or raised floors and cable trays for cable protection and concealment, you still may be at risk. Cable shielding and connectors can be compromised if they are under some continued weight-related stress associated with how they are attached and supported on the patch panel or Fibre Channel component. Over time, the light loss associated with the connections could increase due to these stresses. Some cables have been damaged from a heavy floor tile being dropped. For these reasons, EMC recommends that you always have spare cables on hand at your site.

When running trunks from a Fibre Channel device, EMC recommends that the links to this device be split among several different, smaller trunks, and that those trunks are run under a separate set of floor tiles some distance from each other. This is the same analogy used when DWDM rings are established; both sides of the ring should never be run through the same location. This analogy can be brought into the data center and increase the resilience here as well.

This process is completed by maintaining a cable plant diagram that lists the cable trunks, their role in the data center, their start and end points, the paths that they travel, and their position at the end devices. Maintenance of this plan can greatly reduce the time needed to troubleshoot suspected cable problems. It will also help you redeploy cable trunks as equipment assets are shifted. Cable trunk spares may be used not only for the devices that they are currently servicing, but also for neighboring devices when their trunk spares have been exhausted. This cross-utilization allows for a more efficient sparing model at the site.

EMC recommends that you run extra lines with each trunk. These lines can be used for *fast starts* associated with getting a new system up and running as soon as possible, as well as replacements for suspected troublesome cables. The number of cables you spare in trunks or lose idepends on how dynamic your environment is and how subject it is to stresses.

ISL port sparing

Data and performance requirements will continue to grow as time goes by. To manage that growth you may need to add ISLs between adjacent switches. To do this, you would first have had to reserve spare ports on both switches to connect the new ISL. Without sparing the ports in the initial design, you would first have to migrate components off of the selected switches onto other switches. This migration would cause at least a momentary dip in performance and, depending on the redundancy in the fabric design, could cause a temporarily delay of services. If care is not taken, the resultant new data path may be longer and more congested than the original, thus causing another congestion issue that will need to resolved.

While this section talked about growth of an existing fabric, there is also the growth associated with the addition of new switches and the merging of existing fabrics for the purpose of increasing our storage accessibility. When you are attempting to identify the quantity of sparing ports for both growth as well as for expansion, you need to consider both the location and direction of these new ISLs. Each topology design will have its own methods for adding ISLs for added bandwidth and merging fabrics.

Sparing for failures

Failures of switch components, while rare, can occur in your fabric. The failure of a port card in a director would force you to find free ports for most, if not all, of the attached ports on the failed card. Since failures are a temporary state and you are always expected to have the failed component replaced, you may decide to move only those items that are necessary in the short term. To provide for a card, port, link, or optic failure, you must first understand how any one failure of any one component could effect your business continuance. Depending on your current protection schemes (for example: mirroring fabrics, PowerPath, or sparing models), you can identify which resources need protection and preservation during an outage. Any event in the fabric will do its best to work around the problem automatically. For this reason, you need to consider sparing as a means to protect the relative bandwidth performance of particular

applications and lines of business. The planning of such an event, as well as its recovery procedure, is just as important as the reservation of resources to handle these unexpected events.

EMC recommends that if you are unprotected, or weakly protected, you should spare extra ports in your fabric for failure events. You may find the following useful when considering where and how many ports to spare in your environment:

- Sparing per switch — This allows you to avoid any logistics associated with moving things to a new Domain ID or recabling to reach the other switch. It may also allow you to avoid any issues associated with the rebalancing of the bandwidth load because the traffic patterns have changed.
- Sparing per fabric tier — This may allow you to avoid any logistics associated with rebalancing the bandwidth load because the traffic patterns have changed. This model may allow you to spare less if you are able to redeploy your cabling to make these changes.
- Sparing per fabric — This allows you more flexibility and could reduce the overall spare ports required, but would cause you to possibly rebalance the fabric traffic. If this is extended to include the sparing for both sides of a mirrored fabric, you would have to include both rezoning and migration of components currently attached to the other side of the fabric. Since the interdependence of the storage components may be very complex, EMC recommends that, whenever possible, you do not create a sparing model that includes the migration of components to other fabrics.
- Sparing per site — EMC strongly recommends that you always reserve some level of sparing for each individual site within your business. If you are severely limited by the number of spares that you can reserve, you may want to explore a recovery plan that calls for a merge of smaller environments that shares the spares among the entire environment. The detailed planning of such an event is the most important step in its successful execution.

Determining customer requirements

Note: Although this section assumes that the reader will be deploying a completely new SAN, the user adding capacity may find some useful information in this section, as well.

Once you have an idea of what a SAN will be used for and the physical location of each piece of equipment settled, you need to consider how many host and storage ports will be deployed initially, and how the environment is expected to grow, before you can decide on the right topology.

The following information is provided in this section:

- "Scalability", next
- "Choosing a switch type" on page 763

These are just guidelines. Your actual implementation will vary as more or less ISLs are needed between switches.

Scalability

Scalability information for full mesh with two and four ISLs and full mesh core with edge switches has been included below to help you find the right topology.

Full mesh core with two and four ISLs

In Table 34 and Table 35, combinations that result in negative numbers or result in greater than 2048 available ports are grayed out and should *not* be considered for use.

Table 34 Full mesh with two ISLs

Number of ISLs between switches	2																	
	Number of ports on switch																	
Number of switches	16			24			32			64			140			256		
	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail
1	16	**16**	100%	24	**24**	100%	32	**32**	100%	64	**64**	100%	140	**140**	100%	256	**256**	100%
2	32	**28**	88%	48	**44**	92%	64	**60**	94%	128	**124**	97%	280	**276**	99%	512	**508**	99%
4	64	**40**	63%	96	**72**	75%	128	**104**	81%	256	**232**	91%	560	**536**	96%	1024	**1000**	98%
8	128	**16**	13%	192	**80**	42%	256	**144**	56%	512	**400**	78%	1120	**1008**	90%	2048	**1936**	95%
16	256	-224	-88%	384	-96	-25%	512	**32**	6%	1024	**544**	53%	2240	1760	79%	4096	3616	88%

Table 35 Full mesh with four ISLs

Number of ISLs between switches	4																	
	Number of ports on switch																	
Number of switches	16			24			32			64			140			256		
	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail	#total	# avail	%avail
1	16	**16**	100%	24	**24**	100%	32	**32**	100%	64	**64**	100%	140	**140**	100%	256	**256**	100%
2	32	**24**	75%	48	**40**	83%	64	**56**	88%	128	**120**	94%	280	**272**	97%	512	**504**	98%
4	64	**16**	25%	96	**48**	50%	128	**80**	63%	256	**208**	81%	560	**512**	91%	1024	**976**	95%
8	128	**-96**	-75%	192	**-32**	-17%	256	**32**	13%	512	**288**	56%	1120	**896**	80%	2048	**1824**	89%
16	256	-704	-275%	384	-576	-150%	512	**-448**	-88%	1024	**64**	6%	2240	1280	57%	4096	3136	77%

The following definitions apply to these tables:

Number of switches	The number of switches in the fabric.
Number of ISLs between switches	Indicates the number of ISLs that will connect every switch to every other switch. Since this is a full mesh, all switches will connect to each other.
Number of ports on switch (i.e., 16, 24, 32, etc)	Indicates the number of ports on each switch in the fabric. It assumes all switches have the same port count.
#total	The total raw port count. As of publication, the raw port count cannot exceed 2048 in any single fabric. This number can be determined by summing the port count on each switch.
# avail	The number of ports available for Nx_Ports to attach to.
% avail	The amount of ports that were not consumed by E_Ports expressed as a percentage of the total ports. Generally, you should not use a topology that has 50% or less of the ports available.

Full mesh core with edge switches

Table 36, Table 37, and Table 38 on page 762 provide information for 64, 140, and 256 port core switches, respectively.

Table 36 **Using 64 port core switches**

Number of ports on core switches		64					
Number of ISLs between core and edge		2	Number of core switches			4	
Number of cores each edge will connect to		2	Number of ISLs between cores			2	
Number of Edge switches	Core ports	Number of ports on edge switch					
		16	24	32	64	140	256
1	**228**	12	20	28	60	136	252
2	**224**	24	40	56	120	272	504
4	**216**	48	80	112	240	544	1008
8	**200**	96	160	224	480	1088	2016
16	**168**	192	320	448	960	2176	4032
24	**136**	288	480	672	1440	3264	6048
32	**104**	384	640	896	1920	4352	8064

Table 37 **Using 140 port core switches**

Number of ports on core switches		140					
Number of ISLs between core and edge		2	Number of core switches			4	
Number of cores each edge will connect to		2	Number of ISLs between cores			2	
Number of Edge switches	Core ports	Number of ports on edge switch					
		16	24	32	64	140	256
1	**532**	12	20	28	60	136	252
2	**528**	24	40	56	120	272	504
4	**520**	48	80	112	240	544	1008
8	**504**	96	160	224	480	1088	2016
16	**472**	192	320	448	960	2176	4032
24	**440**	288	480	672	1440	3264	6048
32	**408**	384	640	896	1920	4352	8064

Table 38 **Using 256 port core switches**

Number of ports on core switches		256					
Number of ISLs between core and edge		2	Number of core switches			4	
Number of cores each edge will connect to		2	Number of ISLs between cores			2	
Number of Edge switches	Core ports	Number of ports on edge switch					
		16	24	32	64	140	256
1	**996**	12	20	28	60	136	252
2	**992**	24	40	56	120	272	504
4	**984**	48	80	112	240	544	1008
8	**968**	96	160	224	480	1088	2016
16	**936**	192	320	448	960	2176	4032
24	**904**	288	480	672	1440	3264	6048
32	**872**	384	640	896	1920	4352	8064

Table 36 on page 761, Table 37, and Table 38 are mostly self-explanatory with the exception of the field of numbers. The core ports and the appropriate value under **Number of ports on edge switch** need to be added in order to determine the total port count. When this is done, a few configurations fall outside of the support envelope of 2048 Nx_Ports. If the number of ISLs are increased between cores and edge switches, then some of the fabrics fall back into the supportable range.

Table 39 on page 763 shows the increase of ISLs between cores and edge switches.

Table 39 **Number of ISLs increased between cores and edge switches**

Number of ports on core switches		256					
Number of ISLs between core and edge		4	Number of core switches			4	
Number of cores each edge will connect to		2	Number of ISLs between cores			8	
Number of Edge switches	Core ports	Number of ports on edge switch					
		16	24	32	64	140	256
1	**912**	10	18	26	58	134	250
2	**896**	20	36	52	116	268	500
4	**864**	40	72	104	232	536	1000
8	**800**	80	144	208	464	1072	2000
16	**672**	160	288	416	928	2144	4000
24	**544**	240	432	624	1392	3216	6000
32	**416**	320	576	832	1856	4288	8000

Note: Remember, if you are considering deploying a mirrored fabric (and you should because this is a best practice), the number of ports needed on each fabric will be roughly half of the total ports needed.

Once you have decided on a topology, the number and type of switches that will be needed can be determined.

Choosing a switch type

This section provides considerations for choosing a vendor and selecting a model.

Choosing a vendor

Consider the following when choosing a vendor:

- If an environment has standardized on a switch vendor such as Brocade, Cisco, or Brocade M-Series, you should use a switch from their product line. Although improvements to test coverage of interop environments have been made, interop fabrics remain the least tested configurations as switch vendors spend much more time verifying interop with their own products rather than investing time in testing interop with another switch vendors products.

The subject of interoperability is raised because even if the fabrics are not connected when installed, there is a chance that connecting them together will be desired at some point in the future.

- An equally-important reason for using the same vendor is training. A user who has standardized on a particular vendor is less likely to need training on the product. Typically, their expectation of the products performance is more realistic and any infrastructure challenges (power, monitoring) have already dealt with.
- Sometimes it is not possible to keep with the same vendor as the decision has made to migrate to another vendor. If this is the case, read on, but refer to the "Heterogeneous switch interoperability" on page 256 for more specific ideas on how to migrate when the need eventually arises.
- If a particular vendor has not been standardized, then determine which features will work best for you.

Selecting a model

Once a vendor has been chosen, it is time to select a model. There are many different aspects to consider, but this section is only in regards to port count.

Switches provide between 8 and 64 ports of connectivity. Directors provide between 8 and 528 ports of connectivity. Keep in mind when ordering a Director that they all have minimum shipping configurations. For example, assuming 4 GB/s FC will be used:

- For Cisco: 9513, 9509, and 9506, there is a minimum of one blade per chassis (16 ports)
- For Brocade Silkworm 48000, there is a minimum of two blades per chassis (64 ports)
- For Brocade M-Series Intrepid 10000, there is a minimum of two LIMs per chassis (64 ports)
- For Brocade M-Series Intrepid 6140, there is a minimum of two UPMs (8 ports)

Factor these minimums in when considering which switch and how many of each to purchase.

FC SAN concepts

The purpose of this section is meant to familiarize the reader with how SAN technology works, rather than how to use the technology. With this as our goal, this information is organized to follow the flow of data from the initiator across the SAN to the target.

- "Fibre Channel standards" on page 766
- "Hosts" on page 778
- "HBAs" on page 783
- "ASICs (ASIC or Application Specific Integrated Circuits)" on page 783
- "8b / 10b encoding and decoding" on page 785
- "SERDES" on page 791
- "Optics" on page 791
- "Fiber" on page 798
- "Data transfer rates" on page 800
- "Fibre Channel port types" on page 804
- "FC-AL (Fibre Channel Arbitrated Loop)" on page 806
- "Hubs" on page 806
- "FC-SW (Fibre Channel switched fabric)" on page 806
- "In order delivery" on page 904
- "Inter-Switch Link (ISL)" on page 908
- "Frame services in Fibre Channel" on page 908
- "Frame structure in Fibre Channel" on page 913
- "Class of Service (C.O.S.)" on page 925
- "Buffer-to-buffer credit (BB_Credit)" on page 928
- "Zoning" on page 929
- "Storage" on page 929
- "NPIV" on page 930
- "Fibre Channel Routing" on page 937
- "DWDM" on page 953
- "CWDM" on page 953

- "Fast Write" on page 954
- "Vendor-specific features" on page 957
- "Port fencing" on page 960
- "Threshold alerts" on page 960
- "Management" on page 961

Fibre Channel standards

Fibre Channel standards define layered communications architecture similar to other networking environments. Each level of the Fibre Channel protocol stack provides a specific set of functions, summarized in Figure 211.

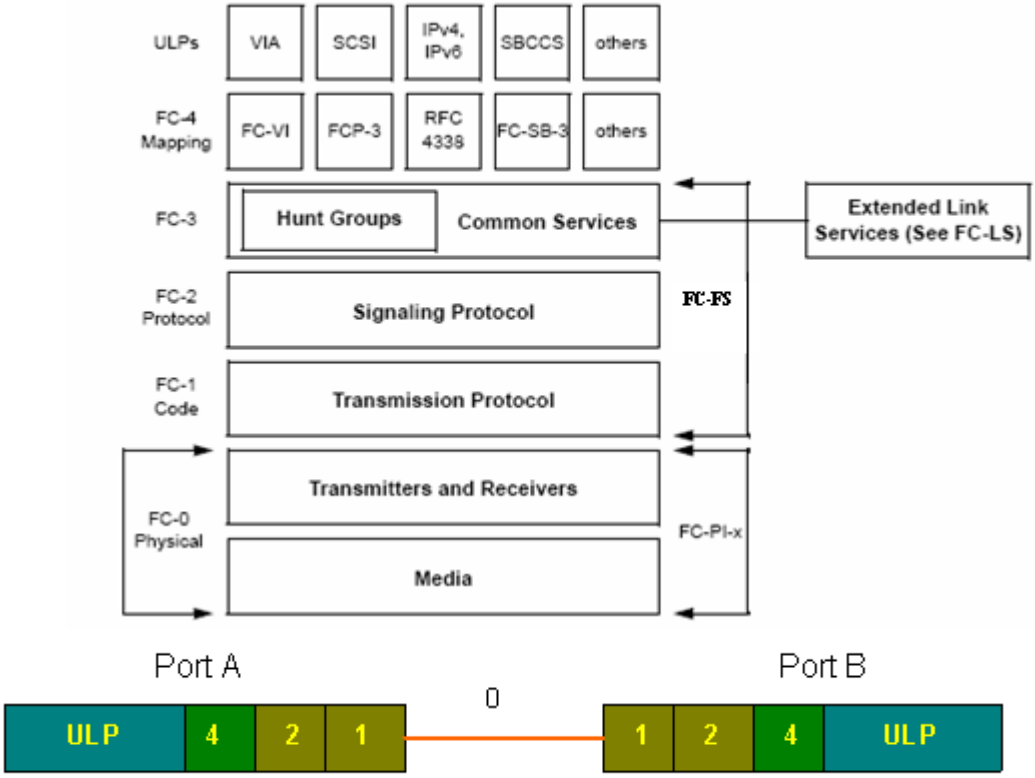


Figure 211 **Fibre Channel levels**

The highest level in Fibre Channel, FC-4, defines the interface between upper-level protocols and the underlying transport layer. This section primarily focuses on SCSI-3 Fibre Channel Protocol (SCSI-FCP) implementations, which provides a connectivity technology with the channel benefits of low protocol overhead and high performance at the upper level, and the network benefits of extended distance, dynamic configuration, and open systems connectivity at the transport layer.

For more information on Fibre Channel levels, refer to "Architectural layers," next.

Architectural layers

The Fibre Channel architecture covers all aspects of Fibre Channel from the physical interface to the transport of multiple upper layer protocols. What must be understood here is that Fibre Channel is a layered approach and each level has multiple functions each of which describes a specific aspect of Fibre Channel.

The following levels are discussed in this section:

- "FC-0 layer" on page 768
- "FC-1 layer" on page 769
- "FC-2 layer" on page 772
- "FC-3 layer" on page 777
- "FC-4 layer" on page 777
- "Upper Layer Protocol" on page 778

Figure 212 shows the architectural levels in Fibre Channel.

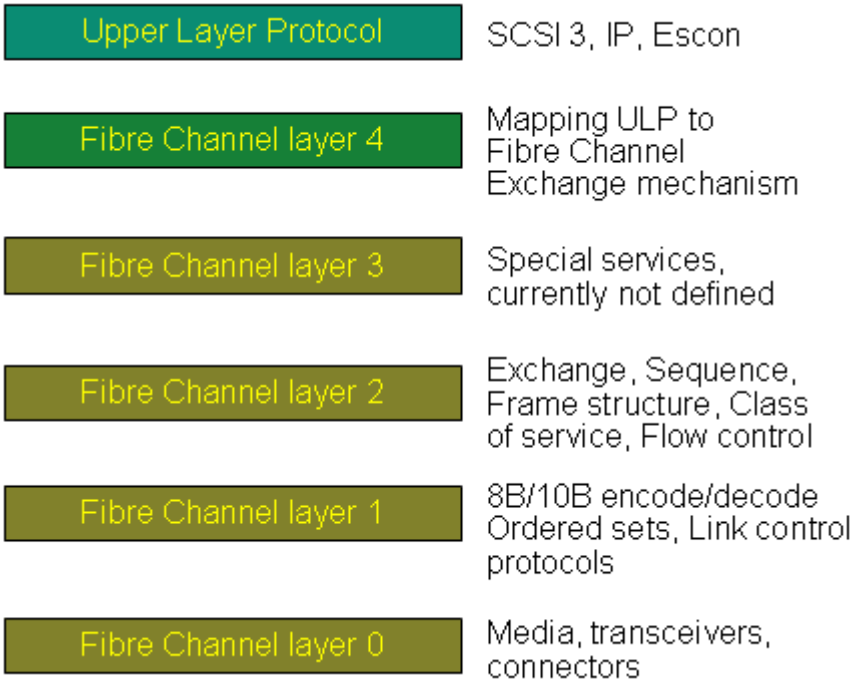


Figure 212 Fibre Channel architectural layers

The standards document Fibre Channel Physical Interface (FC-PI) defines the FC-0 and FC-1 levels of Fibre Channel. FC-FS (Fibre Channel Framing and Signaling) defines the FC-2 level. FC-0 through FC-2are available on the port while the FC3, FC4, and ULP (Upper Layer Protocol) layers are provided by the system / node being used. Each of these layers will be discussed next in detail to give a clear understanding as to how they all combine to provide the Fibre Channel architecture.

FC-0 layer

The FC-0 layer describes the physical interface including transmission media, transmitters and receivers, and their interfaces. This layer defines data rates provided by the Fibre Channel standard, optical and electrical media that can be used at each rate, connectors associated with each media type, maximum distance capabilities and other characteristic such as wavelength of light and light levels.

FC-1 layer

The FC-1 layer in Fibre Channel defines how data is encoded prior to transmission and decoded upon receipt. There are three primary functions of the FC-1 layer:

- "Encoding / decoding," next
- "Ordered sets" on page 769
- "Link initialization" on page 771

Encoding/decoding

In order to transmit data over a high-speed serial interface the data is encoded prior to transmission and decoded upon receipt. All information transmitted over Fibre Channel is encoded into 10-bit transmission characters prior to transmission and decoded back into 8-bit bytes at receipt. Encoding the data improves the transmission characteristics of the serial bit stream and facilitates successful recovery of the data at the receiver. This encoding also maintains a balance between the number of ones and zeros transmitted which allows for the recovery of a clock from the data stream and allows the receiver to be properly biased. Refer to "8b / 10b encoding and decoding" on page 785 for more information.

Ordered sets

Fibre Channel uses a number of transmission words to perform control functions. These transmission words consist of four transmission characters the first of which is the K28.5 special character. The remaining three characters are data characters (D. xx.y) that are used to define the meaning of the ordered set. Three classifications of ordered sets are defined:

- Frame delimiters — Identify the start and end of a frame. There are two types:
  - SOF delimiters — Identify the start of a frame and the class of service associated with the frame. Used to establish a class 1 connection and to signify the beginning and continuation of a sequence.

- EOF delimiters — Identify the end of a frame. Used to end a connection, to signify whether the frame is the last frame in a sequence, and to indicate certain frame errors.

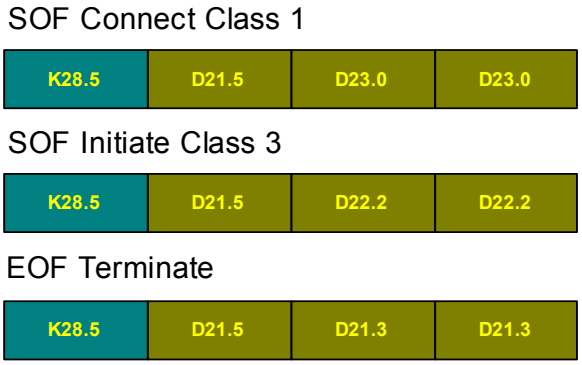


Figure 213 Examples of SOF and EOF delimiters

- Primitive signals are defined for use over a single link for indicating events at a transmitting port. There are three basic types:
  - Arbitrate (ARBx) ordered set is used to indicate that a loop port requires access to the loop.
  - An IDLE indicates the port is ready for Frame transmission and reception. They are transmitted when the port has no other specific information to send.
  - The R_RDY ordered set is used to control the transmission of frames on a link and indicates the receiver has emptied a receive buffer and is ready to receive another frame.

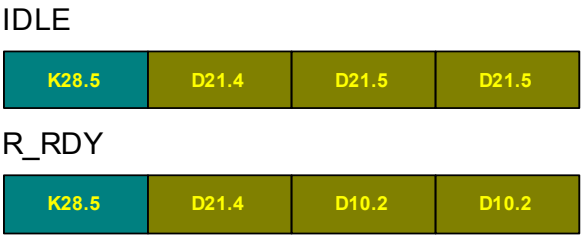


Figure 214 Example of primitive signals

- **Primitive sequences** ordered sets are used for link initialization and error recovery in Fibre Channel environments. They are used in arbitrated loop and also in switched fabric as for link initialization. (Refer to "Link initialization," next.

Link initialization

The four primitive sequences used in link initialization in switched fabric are:

- Non Operational (NOS) — Transmitted by a port to indicate that the transmitting port has detected a link failure.
- Offline (OLS) — Transmitted by a port to indicate the port is beginning link initialization, has received and recognized the NOS sequence, or the port is entering the offline state.
- Link Reset (LR) — Used to indicate a link reset.
- Link Reset response (LRR) — Transmitted by a port to indicate that it has recognized a LR sequence and performed the appropriate actions.

The flow of the primitive sequences in initializing a link is shown in Figure 215 on page 772.

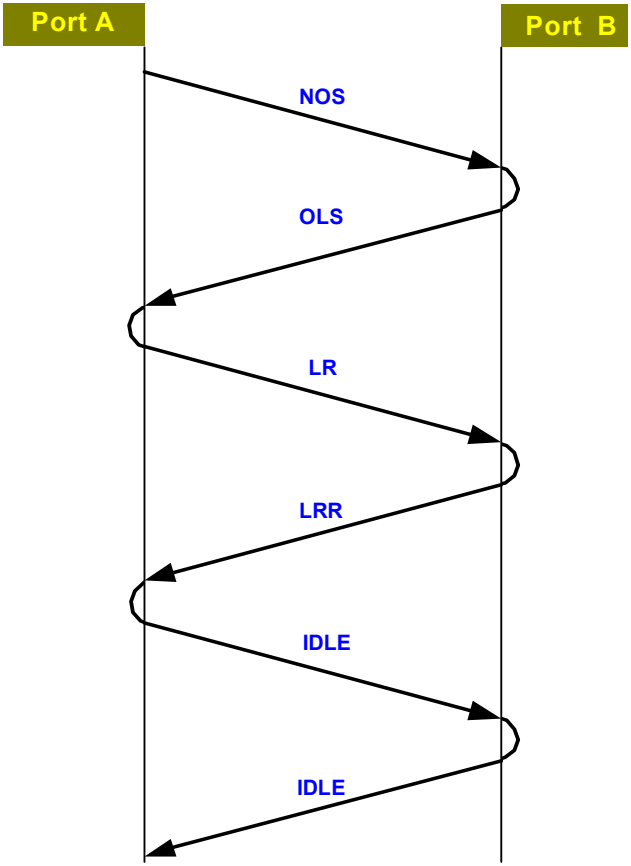


Figure 215 **Link initialization handshake**

FC-2 layer

The FC-2 layer in Fibre Channel defines the structure and organization of information being delivered and how it is controlled and managed. The FC-2 layer covers exchange and sequence management, frame structure, class of service, and flow control.

There are four tiers to the organizational structure in FC-2 to control and manage delivery of information, as shown in Figure 216 on page 773.

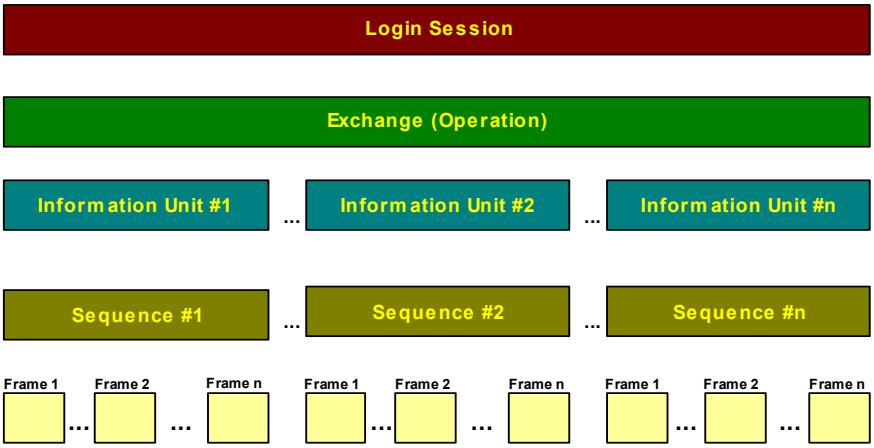


Figure 216 **Exchange sequence and frame relationship**

Login session tier

Before any I/O type operations can take place in Fibre Channel, two ports must establish a login session. During the login process information is exchanged between both ports, which is used in any further communication between the ports. As long as the login session is active between both ports normal I/O operations may take place. If the login session is lost for any reason then any current I/O is terminated and no subsequent I/O can take place until a new login session is established.

Exchange tier

The exchange is the mechanism that allows two Fibre Channel ports to identify and manage a set of information units. These information units may represent an entire operation (Command, Data, Status) or just part of an operation. Each protocol has its own pieces of information that must be sent to another port in order to perform a certain operation. These protocol specific pieces of information are called *information units*. The structure of these information units is defined in the FC-4 protocol mapping for the protocol. If an information unit needs to be sent from one port to another as part of an I/O, the request is passed to the layer beneath. Since there are no services defined in the FC-3 layer the information unit is converted into a sequence which is handled at the FC-2 layer. In effect the information unit corresponds to an FC-2 Sequence.

An exchange is composed of one or more sequences (information units) and within the exchange information units can be sent in either of two ways:

- Unidirectional exchange — Information units are sent in one direction only from exchange originator to exchange responder.
- Bidirectional exchange — Information units are sent in both directions during the course of the exchange.

When the exchange is created the originator has the initiative to send the first sequence but once that sequence has completed the originator may transfer control (initiative) to the responder so the responder can send a sequence in return. This is known as *transferring sequence initiative*.

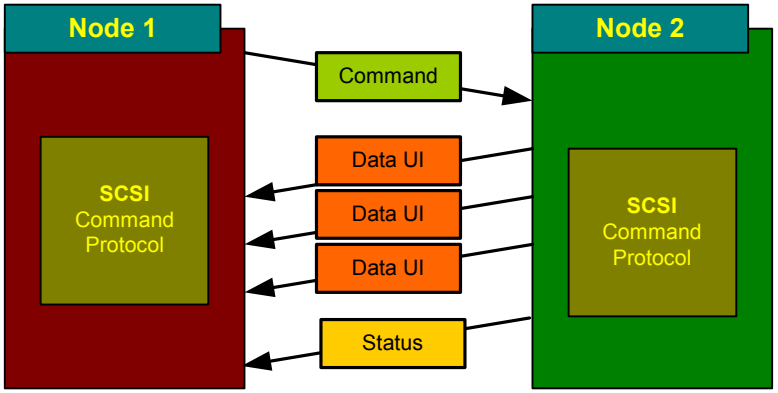


Figure 217 The Fibre Channel exchange

Exchange Identification — Since a port may have multiple exchanges open at the same time there must be a way to identify each exchange individually. Two exchange identifiers exist to provide this:

- Originator Exchange Identification (OX_ID)

  When a port wants to begin an I/O with another port it has to originate an Exchange. The exchange originator assigns an OX_ID to this exchange which has to be unique to the originator-responder pair. Once the exchange is created and the OX_ID is assigned the originator can begin transferring the first sequence to the responder. To identify frames in the exchange this OX_ID is included in every frame in the exchange whether it is sent by the originator or the responder.

One factor to consider is that if multiple originators communicate with a responder as would be the case in a SAN environment with multiple HBAs logging in to a Symmetrix FA it is possible that multiple originators may assign the same OX_ID to separate exchanges. In this case the responder uses the OX_ID in conjunction with the source address identifier (S_ID) of the originator to uniquely identify the exchange.

- Responder Exchange Identification (RX_ID)

  When the responder receives the first frame of an exchange, in order to manage the operation it assigns its own identifier to the exchange which is the RX_ID. This is not required by the standard and if the field in the frame header contains 'FFFF' it is not being implemented. If it is implemented the responder will include the RX_ID value in every frame sent for that exchange.

Sequence

When information needs to be sent from one port to another a request is made to the FC-2 layer to deliver this information. This is accomplished by putting this information into a sequence of frames (refer to "Frame" on page 777). The maximum amount of data that can be transmitted by a single frame is 2112 bytes, thus some sequences will contain more than one frame.

There are number of terms that need to be understood when dealing with sequences:

Sequence Initiator	The sequence initiator is the N_Port that initiates a sequence and transmits the data frames to the destination port.
Sequence Recipient	The sequence recipient is the N_Port that is receiving the Data frames from the sequence initiator.
Sequence ID (SEQ_ID)	Because a port can have multiple open exchanges there then exists the possibility of having multiple open sequences. Each sequence needs its own identifier which is the SEQ_ID which is a value set in the frame header of every frame in that sequence.

Sequence Count (SEQ_CNT) Each frame in a sequence is sequentially numbered by the sequence initiator. This sequential number is in the frame header in the SEQ_CNT field. The SEQ_CNT indicates the position of the frame in the sequence and is also used in verifying every frame of a sequence has been received and is used in handling responses such as acknowledgements, rejects or busy being returned for a certain frame.

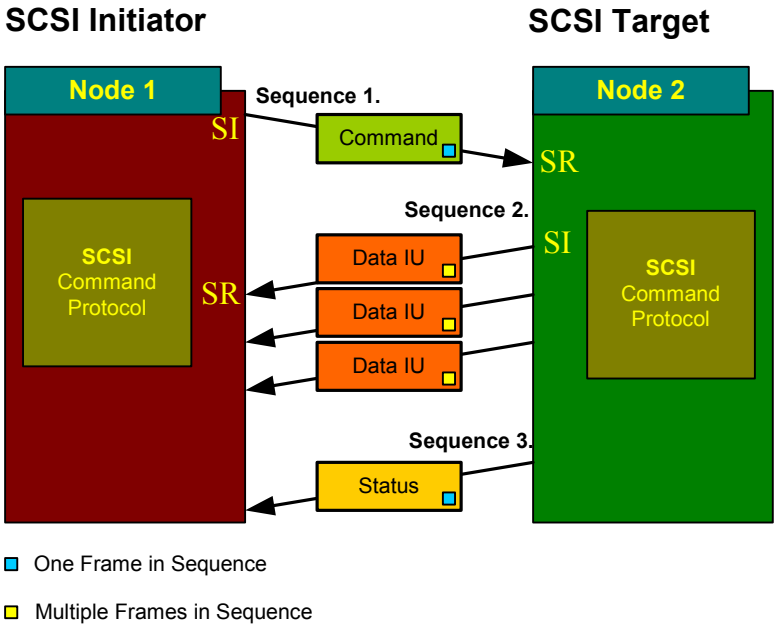


Figure 218 Sequences in an Exchange

Figure 218 shows multiple sequences in an exchange showing sequence initiative (SI) and sequence recipient (SR) changing between Initiator and Target during the exchange. Frames in the same sequence will have the same SEQ_ID. Frames in the same sequence will have their own unique SEQ_CNT value.

Frame

Information transferred in Fibre Channel is packaged into frames. Frames are structured according to a defined format and consist of a fixed length header, a variable length data field (also known as PAYLOAD) and a fixed length CRC. The beginning of a frame consists of a Start of Frame delimiter and at the end there is an End of Frame delimiter (refer to page 769). The header contains address and control information. For more information on frame structure, refer to "Frame structure in Fibre Channel" on page 913.

FC-3 layer

The FC-3 layer provides for future functions that may be implemented in Fibre Channel. These functions are known as *Common Services*. The FC-3 layer could also be used to provide data compression and encryption prior to delivery to the FC-2 layer.

FC-4 layer

The FC-4 layer defines the mapping between the protocols that can be transported by Fibre Channel and the lower transport layers of Fibre Channel (FC0, FC1 and FC2). Each Upper Layer Protocol has its own specific command, data, status or packet information that needs to be communicated with other nodes in order for the protocol to operate. The FC-4 layer defines the format and structure of the protocol specific information being delivered by Fibre Channel.

Examples of protocols that have been mapped to Fibre Channel are:

- Fibre Channel Protocol for SCSI-3 (SCSI-FCP)

  The SCSI FCP protocol mapping defines how SCSI-3 protocol operates using the Fibre Channel interface. It includes for example how SCSI CDB's, Logical Unit Number (LUN), SCSI data, SCSI status and sense information are structured and transported over the Fibre Channel interface.

- Fibre Channel Single-Byte Command Code Set-2 (FC-SB-2)

  This is a new mapping of the ESCON protocol for transport through Fibre Channel.

- Fibre Channel Link Encapsulation (FC-LE)

  This mapping can be used for different network protocols although it is mainly focussed on transport of TCP / IP.

Upper Layer Protocol

The Upper Layer Protocol (ULP) is not really a part of the Fibre Channel protocol as such, but is the command set or packet architecture that will be transported by Fibre Channel.

Examples of Upper Layer Protocols are:

- SCSI
- ESCON
- Intelligent Peripheral Interface (IPI)
- High Performance Parallel Interface (HiPPI)
- IEEE 802. 2 Logical Link Control (LLC) which defines logical content of information transported over networks.

Hosts

In the context of an FC SAN, a host services requests from applications. As a part of servicing these requests, a host may need to store or retrieve data to or from a storage array, located somewhere on the SAN. There are different kinds of hosts, but they can be broken down into two main categories, *mainframe* and *open systems*.

- Mainframe

  A mainframe host connects to the SAN via Channels, accesses storage behind Control Units and usually uses the SBCCS (Single Byte Command Contol Set) protocol.

- Open systems

  An open system host, also known as a server, connects to the SAN through HBAs (Host Bus Adapters), accesses storage behind targets, and usually uses the SCSI-FCP protocol.

This section will mainly focus on open system hosts.

Host design considerations

The following host design considerations are discussed in this section:

- "Persistent binding", next
- "Fan-in and fan-out considerations" on page 782

Persistent binding

The upper levels of all device drivers used in a fabric environment were designed for parallel SCSI. The parallel SCSI I/O model assumed a small number of targets on a bus and a small number of LUNs on each target. The operating system data structures for targets were often a three-bit or four-bit field that was directly associated

with the address signaling method used by the SCSI hardware. In a fabric, the target level of the addressing hierarchy is the 64-bit WWPN. This addressing level is indirectly associated with the fabric hardware. The association between the unique WWPN and the source / destination IDs used to route frames through the fabric is maintained by the name server in the fabric switch.

Using a 64-bit address with three-bit or four-bit device driver software architecture has proved to be a challenge for the Fibre Channel fabric device drivers.

The association of the operating system's controller-target value (the **cxty** in the **cxtydz** or the **bxty** in the **bxtylz**) and the WWPN of the Symmetrix system is called *binding*. Binding takes place in the operating system. The unique WWPN value is bound to an operating system controller-target value. When this bind is persistent from boot-to-boot and across hardware changes and failures in the fabric, it is called a *persistent bind*. To establish a persistent bind, the driver must save this association in non-volatile storage.

Binding is closely associated with the device discovery process. Before any binding associations have been established, the driver queries the name server for all ports that have logged in with an FC-4 type of SCSI. The name server delivers a list of D_IDs (Desgination ID) and the driver logins into each D_ID, then retrieves the WWPN associated with the D_ID and processes this information against the Control Target Device / Bus Target LUN (CTD / BTL) value. (The exact mechanism and storage of this information varies among HBA drivers).

If this association is not saved, the list from the name server might be different the next time the driver makes the association, resulting in a bind between the WWPN and the HBA-target value that is inconsistent with the original bind. In an operating system environment, this results in devices shifting in the device address space, depending on consistent BTL or CTD associations from boot-to-boot.

Figure 219, next, and Figure 220 on page 781 illustrate the binding association and the consequences of not having a persistent bind when a hardware failure occurs and the server goes through a boot.

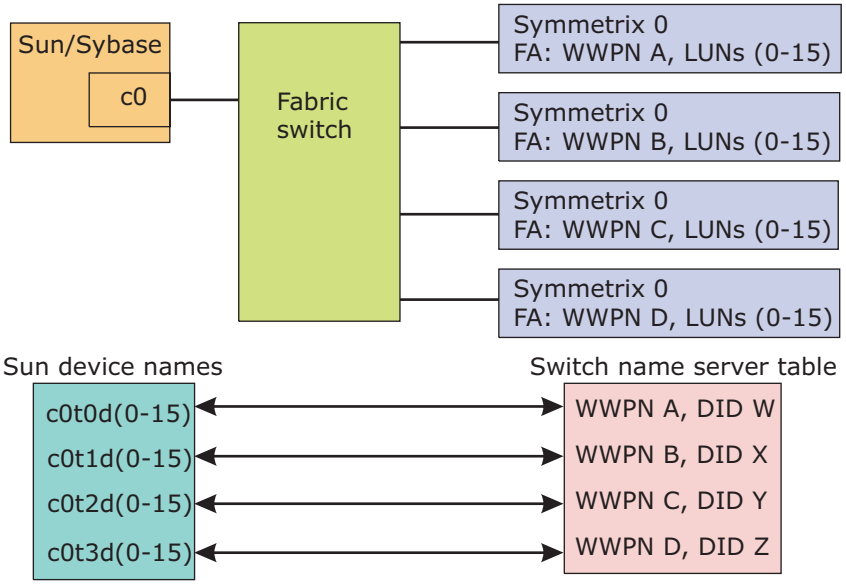


Figure 219 Before hardware failure, outage

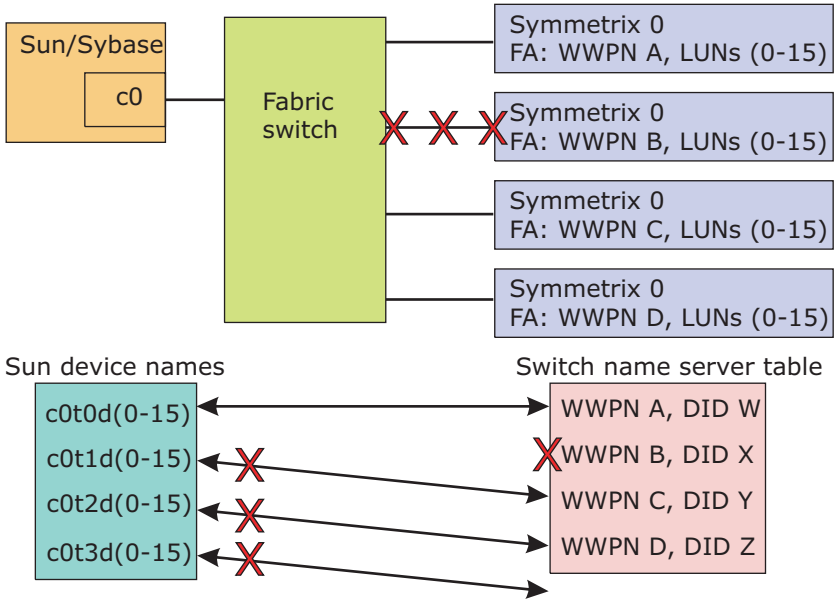


Figure 220 **After any one hardware failure, outage**

When a higher-level operating system function (such as a Windows NT disk signature) or a volume serial number is used to maintain the association between the operating system device name and the device itself, controller-target-to-WWPN binding is not essential. However, these higher-level addressing functions are not found in every environment, so persistent binding is necessary to ensure a single, simple, safe, consistent practice. This is done automatically with the Emulex/Windows NT driver, but requires editing a configuration file in the JNI/Solaris environment. The JNI tech notes and the EMC fabric driver release notes provide more information.

CAUTION

Failure to follow fabric installation instructions can result in data loss.

CAUTION

Do not rely on such functions as the Windows NT signature or a Volume ID when the application or hardware depends on raw device names (for example: controller-target-device in Solaris, bus-target-lun in Windows). For example, Oracle parallel server, Sybase, and tape devices use raw device names. If the customer plans to eventually use the HBA to access both disk and tape, the HBA should implement a persistent binding strategy. In Windows NT, the Emulex port driver is the recommended solution.

Fan-in and fan-out considerations

In the Symmetrix environment, *fan-in* and *fan-out* mean *into* and *out of* the Symmetrix system. E-Lab Navigator contains recommended fan-in and fan-out ratios. Figure 221 shows a fan-in rate of 1:4.

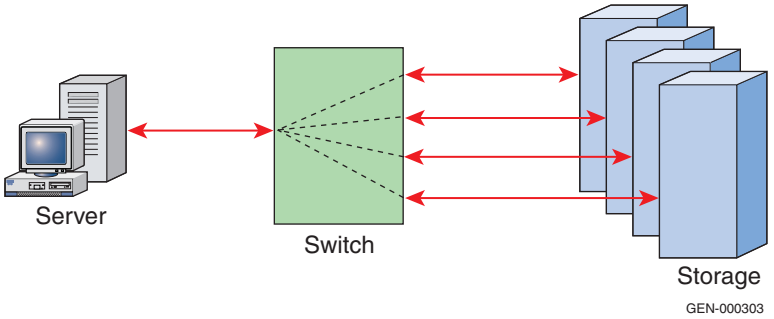


Figure 221 Example of fan-in

Figure 222 shows a fan-out rate of 4:1.

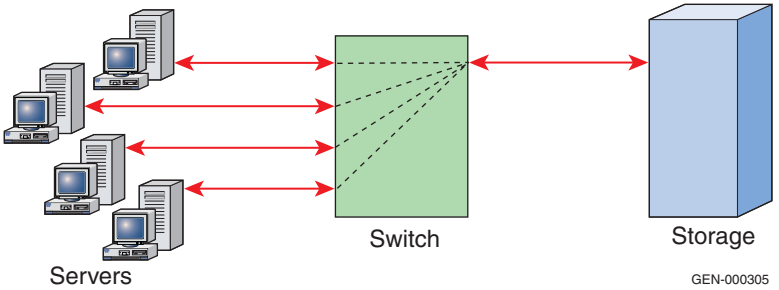


Figure 222 Example of fan-out

HBAs

In the context of an FC SAN, HBAs (Host Bus Adapters) are used to efficiently connect the host to the SAN.

Procedures for configuring most EMC-qualified HBAs for connection to EMC storage arrays can be found at the HBA vendors' websites.

Emulex

To download the Emulex documentation:

1. Access http://www.emulex.com.
2. Click **drivers, software, and manuals** at the left side of the screen.
3. Click **EMC** the upper center of the next screen.
4. Click your HBA on the list at the left side of the screen.
5. Under **Drivers for (your OS)**, click **Installation and Configuration** in the **Online Manuals** column.

QLogic

To download the QLogic documentation:

1. Access http://www.qlogic.com.
2. Click **Downloads** at the left of the screen.
3. Click **EMC** to the right of **OEM approved/recommended drivers and firmware**.
4. Find the description of your HBA driver in the **Name** column of the table for your HBA model. Then click the **Readme** link in the associated **Description** column.

ASICs (ASIC or Application Specific Integrated Circuits)

An ASIC is a highly specialized integrated circuit designed to perform a single task extremely well. Although ASICs can be used for many different purposes in FC HBAs, switches, or storage arrays, the ASICs referred to most frequently are the Port ASICs. A Port ASIC is an ASIC that at a minimum controls basic port behavior and is arguably the largest factor in determining an FC Port's basic set of behaviors or *personality*. It is commonly accepted that switches using the same Port ASIC will react almost identically under the same set of conditions. Because of this, switches are usually classified by the Port ASIC used in its design. Switches that utilize the same Port ASICs are

said to be members of the same ASICfamily. Some current examples of Port ASIC's are Brocade's Condor and Cisco's Vegas ASICs.

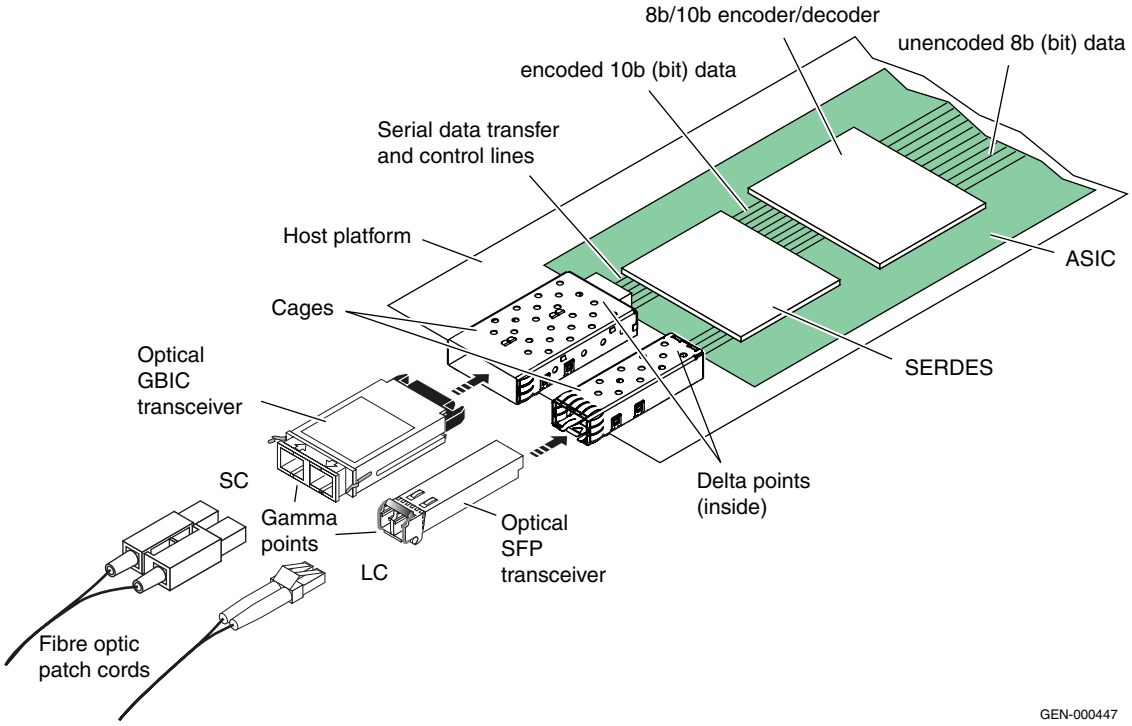


Figure 223 ASIC

8b/10b encoding and decoding

In the context of an FC SAN, 8b/10b encoding is the process of converting 8 bits of data into a 10-bit representation.

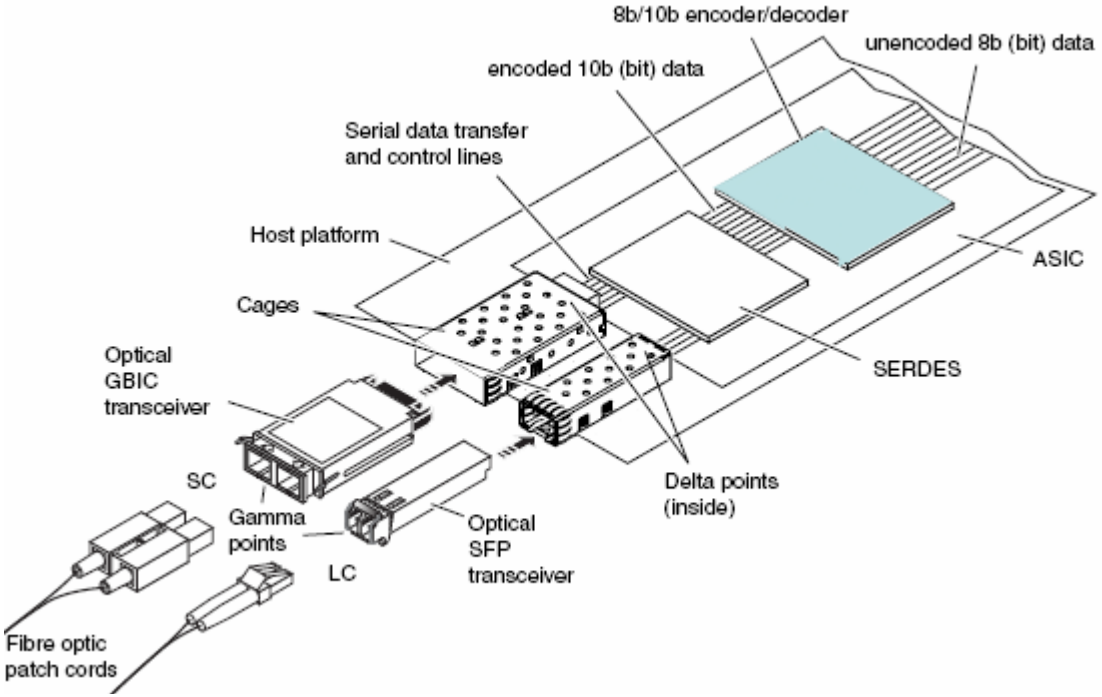


Figure 224 **8b/10b Encoder/decorder**

There are several reasons for this conversion and one of the most important is that it enables the receiver to recover a clock signal from the data stream. Other reasons include data integrity and ensuring that there are no excessive DC components a problem specific to implementations that use copper wire instead of instead of Optical Fiber. The 8b/10b encoding and decoding logic are usually built into the port ASIC.

A 10-bit code facilitates up to 1024 unique bit patterns. From these patterns 256 data characters are encoded in one or two 10-bit patterns. (For more information, refer to "Disparity" on page 788). A

number of data patterns are allocated to special characters, which will be discuss shortly. To explain this, the following is an example of a byte (0xF4) which we would like to send to the other side:

0xF4 = 1111 0100

The bit order gets reversed and the bits are divided into a 5-bit and a 3-bit block:

abcde fgh

00101 111

This pattern is now represents the data character D20. 7 which must be encoded depending on the current running disparity. Figure 225 shows how this example arrives at the D20. 7 data character.

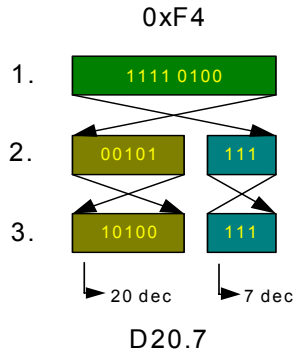


Figure 225 Example

1. Shows the binary bit pattern of our 0xF4.
2. First the bit pattern is reversed and then divided into the 5 and 3 bit sub blocks.
3. Shows how this example arrives at the data character 20. 7.

Note: In reality the 'D' (data character) is achieved by a high or low control line input to the encoding circuitry. The same method would be used for the 'K' special characters.

Figure 226 shows an example of a number of data characters being transmitted. The data byte of 0xF4 is transmitted as an encoded byte.

How this bit pattern was arrived at is discussed in "Disparity" on page 788.

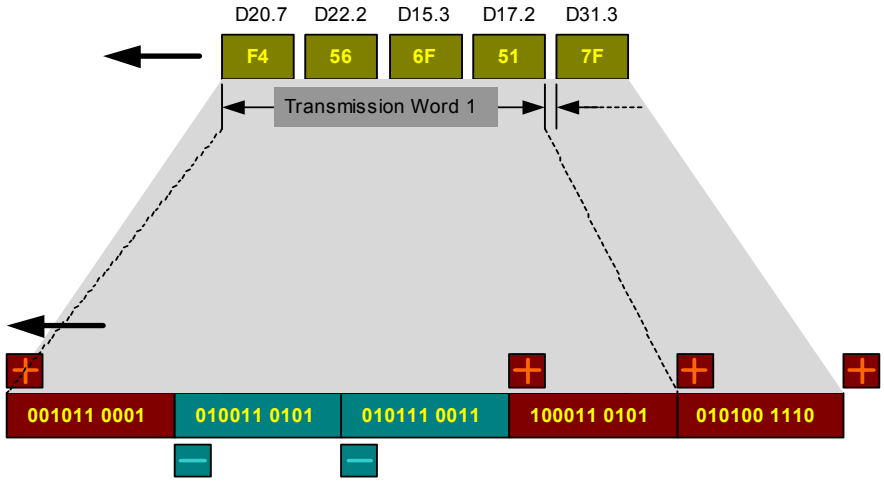


Figure 226 Transmitting 10 bit encoded bytes

All bytes are transmitted as transmission characters, four of which form a transmission word. These transmission words are either *ordered sets* or data *transmission words*. A transmission word is the smallest unit of transfer in Fibre Channel. The '+' and '-' above are an indicator of the balance of ones / zeros within the most recent encoded character. "Ordered sets" is discussed on page 769.

If one had to transmit a SCSI block of zeros, it should be obvious from Figure 226 why the encoding is needed. In a serial transmission stream the receiver would have a hard time detecting the individual bytes and knowing for sure whether it received 511 0x00s or 512.

Figure 227 shows what the difference is between an ordered set and a data transmission word.

Ordered Sets

Data

Figure 227 **Transmission words**

Disparity — The encoding process converts 8 bit input characters to 10 bit transmission characters. To prevent excessive DC components in the bit stream (ensure the same average amount of 1 bits and 0 bits) only certain characters are used. Only characters containing six ones and four zeros, five ones and five zeros, or six zeros and four ones are allowed.

Disparity is the ratio of 1 bits to 0 bits in the bit stream. The possibilities are:

- Positive disparity — Transmission character has more 1 bits than 0 bits
- Negative disparity — Transmission character has more 0 bits than 1 bits
- Neutral disparity — Transmission character has equal number of 0 and 1 bits

Note that each character that has an unequal number of 1 bits and zero bits has two different encoding patterns. One pattern contains six ones and four zeros and the other contains six zeros and four ones.

When the encoder creates a character with an unequal number of zeros than ones, it keeps track of a variable called *Current Running Disparity*. This variable is set to *positive* if the character generated has more ones than zeros and is set to *negative* if the character generated has more zeros than ones. This variable is used to select the appropriate encoding of the next character (positive or negative) to balance the number of 1 bits and 0 bits being transmitted.

For example, the hex value 0xF4 and its two encoding values depending on the value of the Current Running Disparity are shown in Figure 228.

		Current RD -	Current RD+
Data Byte Name	Bits HGF EDCBA	abcdei fghj	abcdei fghj
D20.7	111 10100	001011 0111	001011 0001

Figure 228 **Data character "F4"**

This code would be automatically generated by the encoding hardware taking the following input variables:

- 5 and 3 bit sub blocks
- D or K indicator
- Current Running Disparity

Figure 229 on page 789 shows the transmission of encoded data and the Current Running Disparity changing depending on the transmission character being sent. The positive and negative transmission characters are also shown in this diagram.

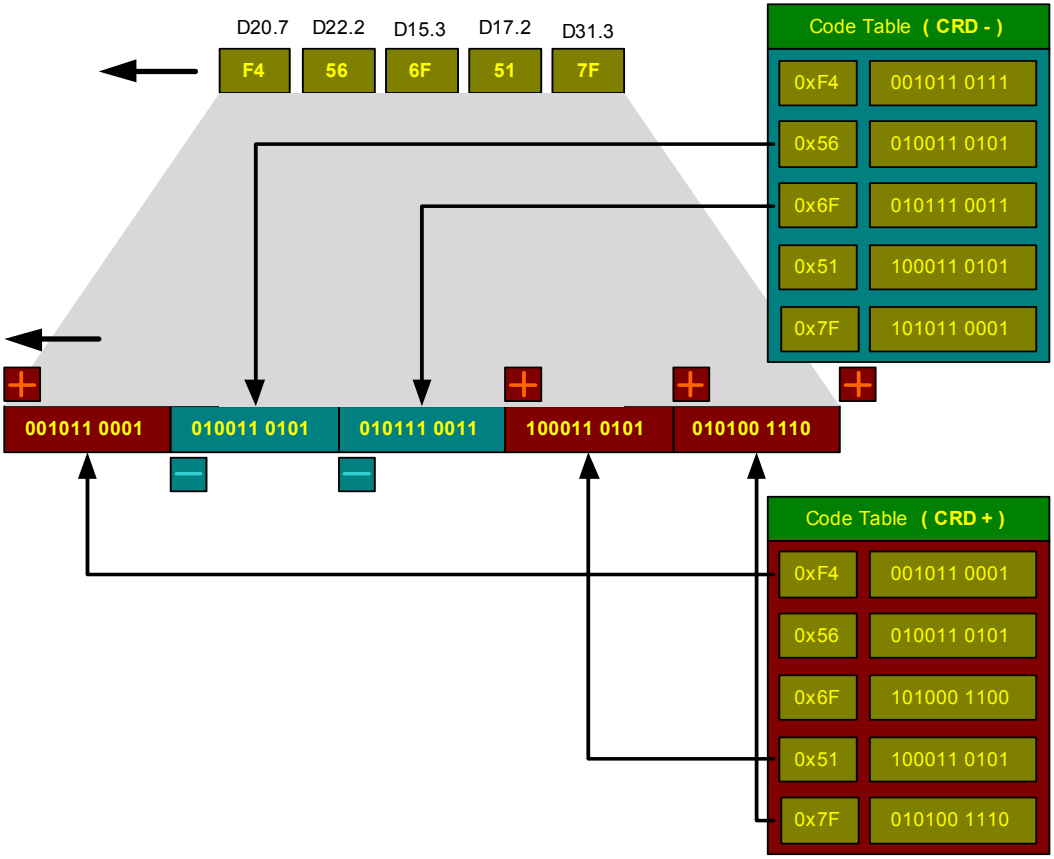


Figure 229 Sending encoded data

Special characters are derived in a similar fashion but always start with a “K”. There are 12 special character defined in the 8b / 10b encoding scheme but only the **K28. 5** special character is used by

Fibre Channel. Its use is limited to the first character in a transmission word known as an Ordered Set (see "Ordered sets" on page 769).

SERDES

The purpose of the SERDES is to convert 10 bit parallel data into 10 bit serial data. The SERDES is usually built into the port ASIC (Figure 230).

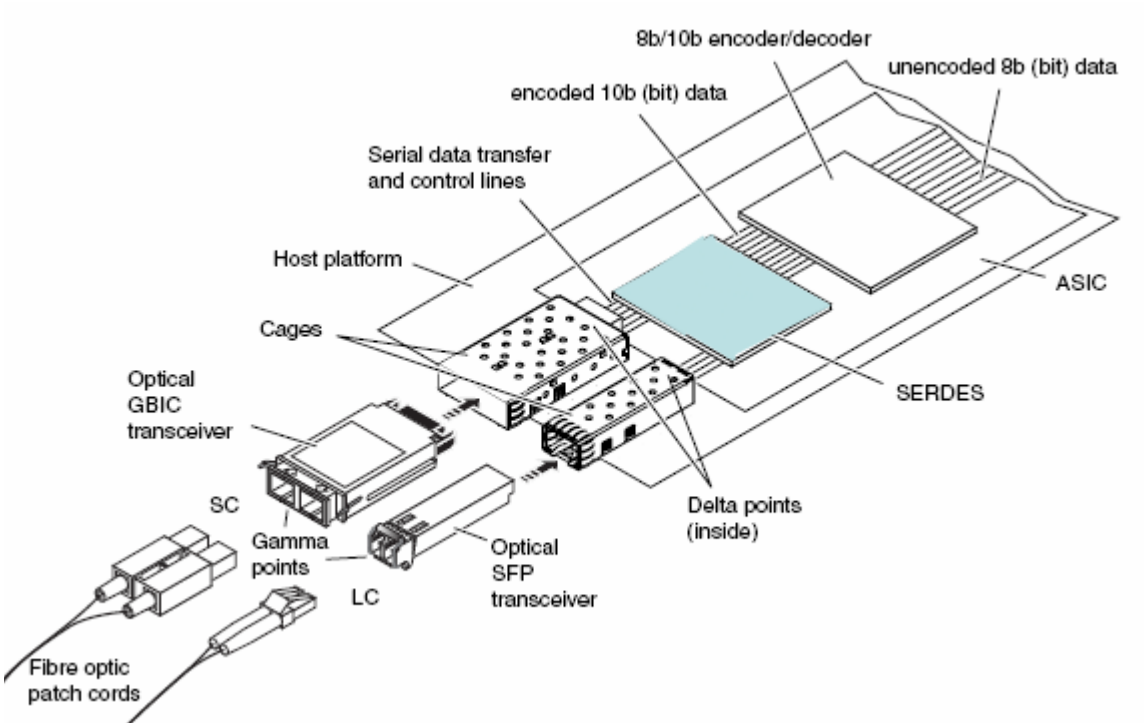


Figure 230 SERDES example

Optics

Optics (formerly known as optical transceivers) convert the high and low electrical pulses coming from the SERDES (serializer / deserializer) on the host platform, into high and low light levels which propagate down the optical medium (fiber). Optics also convert incoming high and low light intensity levels into high and low voltage outputs, which will be interpreted by the SERDES on the receiver before being passed to the 8b / 10b decoder on the receiver.

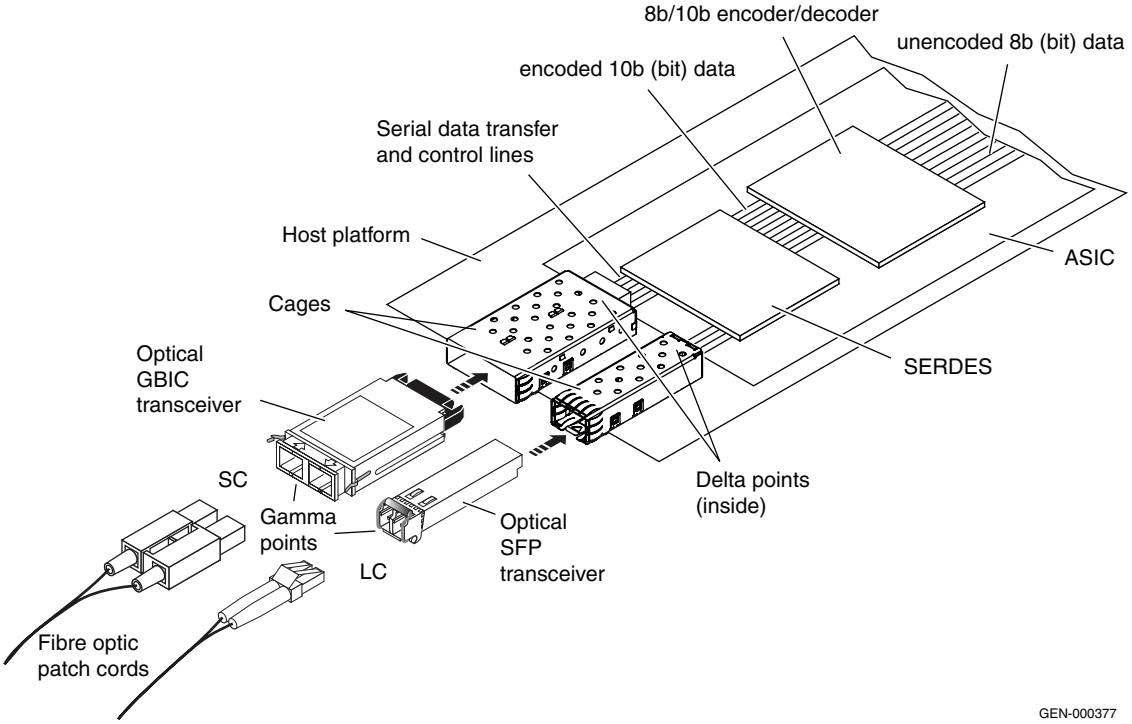


Figure 231 **Optics**

Optics come in a variety of packages (SC, LC, XFP, etc). Each package typically has its own cable type. Packages of the same type are usually interchangeable because they are built using the same set of standards that define the connector layout and other necessary information. SFPs, for example, are expected to adhere to the SFP multisource agreement.

Some important points about optics in general are:

- Optics are usually interchangeable with each other, but not always.

  Even though optics are built using the same standards, issues may be encountered due to the physical characteristics of the optic that the developer of the platform, into which the optic will be inserted, may not have considered. These problems may manifest as an inability to bring a link up (gain bit sync) or as periodic CRC errors.

Note: Because of the potential for interoperability issues in this area, it is recommended that should it be necessary to replace a failed optic, you replace it with one of the same type, or one that has been tested and is approved for the host platform.

- Optics convert an electrical signal to an optical signal and conversely.

  The electrical interface of the optical transceiver has many different connection points (pads), (labeled as delta points in Figure 231 on page 792), which are used to connect to the host platform (i.e., FC switch, HBA, front end storage port). The host platform will generate a signal (a high voltage or low voltage) at some of these connection points to cause the optic to do something (i.e., make light brighter, turn off light). The optic can also provide a high or low signal at some of these connections to indicate to the host platform that an external condition, such as loss of light, has been detected.

  The data transferred between the SERDES and the optic is in a serial 10 bit format and it not interpreted by the optics in any way. This means that bit and CRC errors are not detected by the optics, and in fact are not detected until the after SERDES when the 10 bit data is being decoded.

  Note: Although optics are not capable of detecting CRC or bit errors, they are frequently the source of them.

- Some SFPs provide extended digital data.

  Most of the optics shipping today provide extended digital diagnostics information which can be accessed if the host platform supports reading this data and provides an interface to display it. Some of the information that is provided by the SFP, and what it may indicate, is provided below.

  - Temperature: The operating temperature of the optic.
  - 3.3 Voltage: The internal voltage of the optic.
  - Current: The current being drawn by the optic.

  These measurements are specific to the optic and informational in nature. If you were to monitor each optic and were to observe a steady change in one of these values, it may be possible to predict an optic failure before it were to occur. By comparing all of the

values for the same type of optic installed in the same host platform, a single optic that is operating with a set of different values could be detected and may be in the process of failing.

Be aware, though, that these changes might not mean anything; for example, the values might be fluctuating due to changing environmental conditions. Do not worry about these values unless they are approaching a warning or alarm level. Once an optic reaches one of these thresholds, you should consider replacing it.

The next values are provided in microwatts (mW) or dBm, depending on the platform, and can be extremely useful when debugging why a link will not come up.

- TX Power: A measurement of the strength of the signal being transmitted by the optic.
- RX Power: A measurement of the strength of the signal being received by the optic.

In order to interpret the TX and RX power values, it will be useful to quickly review microwatts and dBm and determine how they are related.

- Microwatts (mW): As used in this example, mW refers to a measurement of light intensity.
- dBm: dBM is an abbreviation for the power ratio in decibels (dB) of the measured power referenced to one milliwatt.

Converting between microwatt and dBm is fairly straightforward if you have a calculator that supports the log function:

– Bm to mW:

  mW = 10(dBm / 10)

– mW to dBm:

  Note the optics measurements are given in microwatts before you can use the following formula, convert the value in microwatts to milliwatts by multiplying by 10-3.

  For example 1000 microwatts becomes 1 milliwatt.

  dBM = 10 * LOG mW

Instead of converting to get an exact value, use Figure 232 to get an idea of where the measured value falls.

microwatt	milliwatt	dBm	Description
1	0.001	-30.00	Loss of Signal
10	0.01	-20.00	
25.1	0.0251	-16.00	Minimum RX value allowed at 2 Gb/sec
31.6	0.0316	-15.00	Minimum RX value allowed at 4 Gb/sec
50	0.05	-13.01	
100	0.1	-10.00	Minimum TX value allowed at 2 Gb/sec
125.9	0.1259	-9.00	Minimum TX value allowed at 4 Gb/sec
150	0.15	-8.24	
200	0.2	-6.99	Typical operating range
250	0.25	-6.02	
300	0.3	-5.23	
350	0.35	-4.56	
400	0.4	-3.98	
450	0.45	-3.47	
500	0.5	-3.01	
550	0.55	-2.60	
600	0.6	-2.22	
650	0.65	-1.87	
700	0.7	-1.55	
750	0.75	-1.25	
800	0.8	-0.97	
850	0.85	-0.71	
900	0.9	-0.46	
950	0.95	-0.22	
1000	1	0.00	
1050	1.05	0.21	
1100	1.1	0.41	
1150	1.15	0.61	
1200	1.2	0.79	
1250	1.25	0.97	

Figure 232 Measured values

The value of -30 dBm shown in the **Loss of Signal** row is the approximate value that a multimode optic will detect a loss of signal. It was determined by reviewing a number of SFP data sheets and should not, therefore, be treated as a universal value.

On all SFPs, pin 8 (LOS) is LOW when a signal is detected and HIGH when a signal is not detected. The point at which this will change state from a LOW to a HIGH depends upon the optical transceiver, but will be in the neighborhood of -30 dBm or .001

mW. However, well before the link could get all the way down to -30 dBm, the bit error rate would become so high that the link would lose sync.

Please note the following misconceptions:

- A high electrical signal is converted into a bright light pulse and a low electrical signal is converted into dark (transmitter off).
- 8b / 10b encoding was created so that a bunch of zeros would not be confused with a loss of signal.

Note: Refer to "8b / 10b encoding and decoding" on page 785 for more information.

Both of these are misconceptions. High and low light levels are high and low with respect to themselves and *not* static values, as is the case when working with a digital circuit (high = 5V, Low = 0V).

Note: All of the light level measurements are provided by the optic to the host platform as a value between 0 and 65535, with 0 being equal to -40 dBm and 65535 being to 8.2 dBm.

The values of -9 dBm (shown in the **Minimum TX value allowed at 4 Gb/sec row**) and -10 dBm (shown in the **Minimum TX value allowed at 2 Gb/sec** row) were taken from FC-PI-2. Both represent the minimum transmit signal allow.

The values of -15 dBm shown in the **Minimum RX value allowed at 4 Gb/sec** row) and -16 dBm (shown in the **Minimum RX value allowed at 2 Gb/sec** row) were calculated by subtracting the maximum link power budget of 6 dB from the minimum TX value allowed. Both represent the minimum receive signal that needs to be decoded without exceeding the BER of 10^{-12}.

Three examples of host platforms (FC switches in this case) that provide a user interface to display enhanced digital diagnostics information are provided next.

Brocade M-EOS (McDATA) platforms:

Using the command **show port opticdata [port number]**

```
Root> show port opticdata 1
Port Number:        1
Overall Health:     Normal
Transciever:        SFP
Type          Value      Low Warning  High Warning  Low Alarm  High Alarm
-----------   --------   -----------  ------------  ---------  ----------
Temperature   36.769     -20.000      90.000        -25.000    95.000
3.3 Voltage   3.265      2.900        3.700         2.700      3.900
Current       8.374      2.000        14.000        1.000      17.000
TX Power      377.200    79.000       631.000       67.000     631.000
RX Power      294.100    15.800       794.000       10.000     1259.000
```

Brocade FOS platforms:

Using the command **sfpshow [portnumber]**

```
l82c7207:admin> sfpshow 2/20
Identifier:  3    SFP
Connector:   7    LC
Transceiver: 150c402000000000 100,200,400_MB/s M5,M6
   sw Inter_dist
Encoding:    1    8B10B
Baud Rate:   43   (units 100 megabaud)
Length 9u:   0    (units km)
Length 9u:   0    (units 100 meters)
Length 50u:  15   (units 10 meters)
Length 62.5u:7    (units 10 meters)
Length Cu:   0    (units 1 meter)
Vendor Name: AGILENT
Vendor OUI:  00:30:d3
Vendor PN:   AFBR-57R5AEZ
Vendor Rev:
Wavelength:  850  (units nm)
Options:     001a Loss_of_Sig,Tx_Fault,Tx_Disable
BR Max:      0
BR Min:      0
Serial No:   A206080273
Date Code:   060223
Temperature: 38 Centigrade
Current:     5.350 mAmps
Voltage:     3269.2 mVolts
RX Power:    0.0 uWatts
TX Power:    335.5 uWatts
l82c7207:admin>
```

Cisco SAN-OS platforms:

Using the command **show interface fc [slot #/port #] transceiver details**

```
switch# show interface fc 2/7 transceiver details
fc2/7 sfp is present
    name is CISCO-FINISAR
    part number is FTLF8524P2BNL-C2
    revision is 0000
    serial number is FNS0942B0JA
    fc-transmitter type is short wave laser w/o OFC (SN)
    fc-transmitter supports intermediate distance link length
    media type is multi-mode, 62.5m (M6)
    Supported speed is 400 MBytes/sec
    Nominal bit rate is 4300 MBits/sec
    Link length supported for 50/125mm fiber is 150 m(s)
    Link length supported for 62.5/125mm fiber is 70 m(s)
    cisco extended id is unknown (0x0)

    Failed to get runtime SFP state info.
    SFP Detail Diagnostics Information
  ---------------------------------------------------------------------------
                                   Alarms                  Warnings
                               High        Low         High          Low
  ---------------------------------------------------------------------------
  Temperature  32.92 C         0.00 C      0.00 C      0.00 C        0.00 C
  Voltage       3.30 V         0.00 V      0.00 V      0.00 V        0.00 V
  Current       8.05 mA        0.00 mA     0.00 mA     0.00 mA       0.00 mA
  Tx Power     -4.21 dBm       N/A         N/A         N/A
  Rx Power     -7.08 dBm       N/A         N/A         N/A

  Transmit Fault Count = 0
  ---------------------------------------------------------------------------
  Note: ++  high-alarm; +  high-warning; --  low-alarm; -  low-warning
```

Applicable standards /organizations

- FC-PI-2
- TIA FO workgroup
- SFP multi-source agreement

Fiber

The ports connect to the fibre topology through a link. The link is a cable, or other connection, that carries the data. The ports transmit and receive information through two separate fibers. The link may consist of optical fibers or electrical cable. The transmitters may be either longwave laser, shortwave laser, LED, or electrical.

The fibre construction, as shown in Figure 233, consists of a central core through which the light travels. The core is surrounded by cladding, the function of which is to reflect and contain the light within the core. The core and cladding are made from a glass material and can be easily damaged. To protect the fibre from physical and environmental damage to the core and cladding, it is covered in protective layers which also give the fibre considerable strength.

There is a minimum bend radius specified around which the fiber may be bent which, if exceeded, may result in a degradation in transmission over the fibre or, worst case, may result in possible damage to the fibre itself. In normal use these cables are quite robust and do not require any special handling apart from observing the minimum bend radius requirement.

The core and cladding diameter (in µm) is typically the means by which a fibre is specified. For example, 62.5 / 125µm fibre has a core diameter of 62.5µm and a cladding diameter of 125µm. Two of these fibres are normally combined in one *duplex* cable that is terminated at both ends with a duplex connector. The two fibres are used to send data in opposite directions for transmit and receive. Simultaneous transmission / reception is possible over the duplex cable.

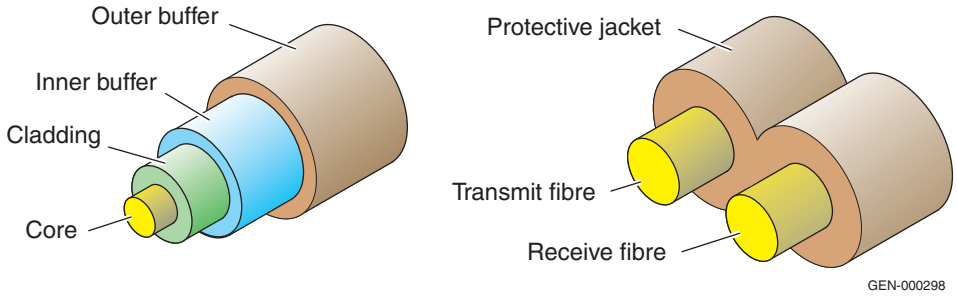


Figure 233 Fibre construction

Single-mode and multimode

There are two modes of transmission in Fibre Channel: single-mode and multimode.

Single-mode links have a fiber core of 9-10 um and use a long wavelength laser operating in the infrared portion of the spectrum at 1300 nanometers (nm) as the light source. This light is not visible to the human eye. The lower core diameter enables single-mode links to support a maximum distance of 10 km between ports as all the light

propagates along the same path in the fiber as is shown in Figure 234. Single-mode links are mainly used over long-distance transmissions and are implemented on certain versions of the Symmetrix Fibre Channel adapter. For example, the 201-333-901 is a four port 2 Gb/s capable adapter that has three multimode and one single-mode port

Multimode is less costly than single-mode and is used where the distance capabilities of single-mode are not required. The Fibre Channel links are usually based on either a 50 or a 62.5 um core diameter and support light wavelengths of approximately 800nm. This increased core diameter size, as opposed to single-mode, means the light has multiple propagation modes (paths) through the fiber. Therefore, some of the light of a pulse may take one path through the fiber and the rest of the light take another path. The result is called *Modal Dispersion*. This results in the spreading of the pulse which limits the distance that can be achieved with multimode cabling.

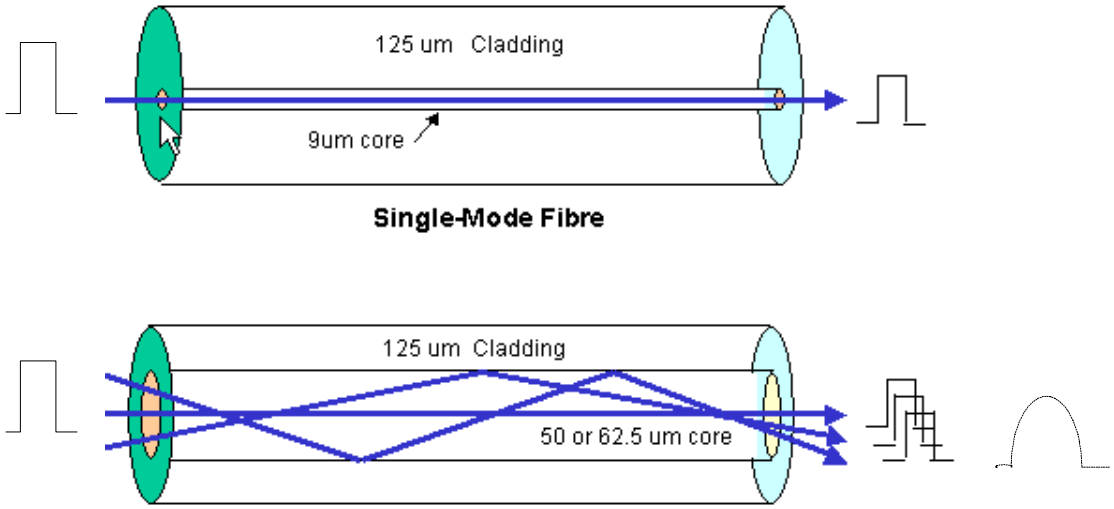


Figure 234 Single and multimode optical fibre

Data transfer rates

The rate at which data will be transferred across a given medium is dependent on a number of factors. This section will consider the medium type first. In the case studies that follow, almost all of the implementations will use a serial communication protocol over an optical medium (i.e., Fibre Channel over multimode fiber).

The next factor to be considered is the signaling rate. In these examples, the signaling rate will equal the bit rate, but this is not always the case in serial communication protocols. Table 40 on page 801 compares Fibre Channel to SONET and SDH in terms of their bit rates.

Table 40 Bit rates

FC abbreviation	MB/s definition	MB/s max theoretical	Signal rate MBd	bits per second (bps)	Signal rate Mbps	Mb/s max theoretical	Optical Carrier Level	SONET Frame format	SDH level and Frame format
				51840000	51.84	48.69	OC-1	STS-1	STM-0
				155520000	155.52	150.336	OC-3	STS-3	STM-1
				622080000	622.08	601.344	OC-12	STS-12	STM-4
1 Gb/sec	100	103.06	1,062 5	1062500000					
				1244160000	1244.16	1,202 668	OC-24	STS-24	STM-8
2 Gb/sec	200	206.13	2,125	2125000000					
				2488320000	2488.32	2,405 376	OC-48	STS-48	STM-16
4 Gb/sec	400	412.25	4,250	4250000000					
				9953280000	9953.28	9,621 504	OC-192	STS-192	STM-64
10 Gb/sec	1200	1236.63	10,518 75	10518750000					
				39813120000	39813.12	38,486 016	OC-768	STS-768	STM-256

As shown in Table 40, Fibre Channel environments can currently operate at 1 of 4 signaling rates (shown in megabaud MBd in the table) or data rates (also known as link speed, shown in MB/s in the table). In order to convert the link speed to a useable number in MB/s, you need to consider that the maximum data rate in MB/s will be decreased by protocol overhead, such as 8b/10b encoding, Start of Frame delimiter (SOF), Frame Header, CRC, End of Frame delimiter (EOF), and at least six primitives between the EOF of one frame and the SOF of another. The following formula can be used to determine maximum throughput.

Maximum throughput in MB/s = (((bit rate*.8)*.97)/8)/1000000

To work out this formula:

1. Take the bit rate and multiply it by .8 to remove the 8b/10b overhead.
2. Take the product from Step 1 and remove the minimum overhead caused by the SOF (4 bytes), Frame Header (24 bytes), CRC (4 bytes), EOF (4 bytes) and the six primitive signals (24 bytes) that are required between each EOF and SOF. This overhead works out to be 3% (actually 3.4%) when frames are 2048 bytes long (using 2048/(2048+4+24+4+4+24) it works out to .966).

To remove this minimum overhead, multiply the product from Step 1 by .97.

3. Take the product from Step 2 and divide it by 8 to get the number of bytes per second.

4. Take the quotient from Step 3 and divide it by 1000000 to get MB/s. This is the number that is displayed in the "MB/s max theoretical" column in Table 40 on page 801.

Exception

The exception to the above rule is the 10 Gb/s link. 10 Gb/s uses 64b/66b encoding and this slightly changes the math:

Maximum throughput in MB/s = (((bit rate* .97)*.97)/8)/1000000

Note that the only difference is that the .8 was removed and replaced with .97. The .97 was obtained by division of (64/66).

The maximum data rate is also affected by the amount of buffers that a receiver has granted to a transmitter. In environments where there are no long distance links, having a large number of BB_Credits is not as beneficial as one might think.

For environments that do have long distance links, the number of BB_Credits available can have a dramatic impact on the maximum data rate. See Table 41 and Table 42 on page 803.

Table 41 **200 MB/s link with max 60 BB_Credits**

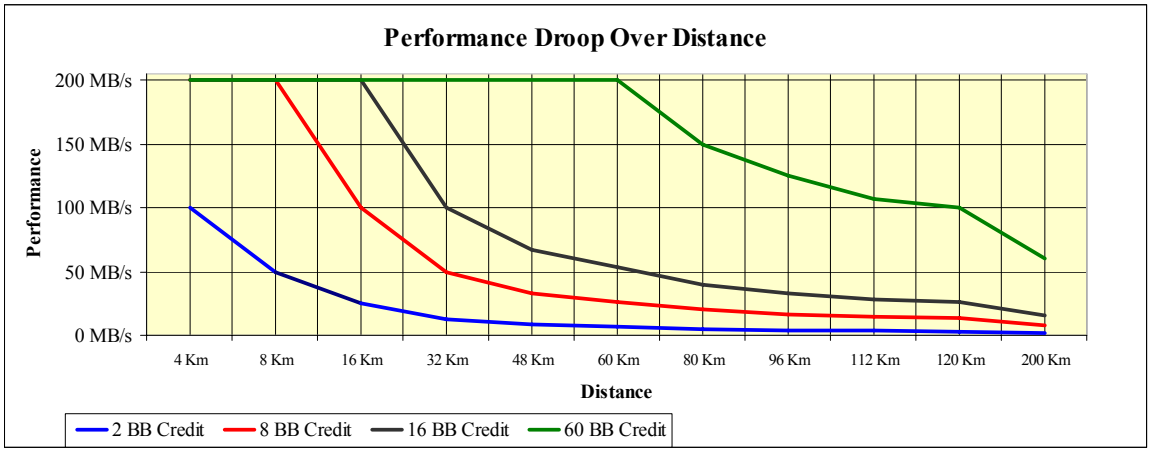
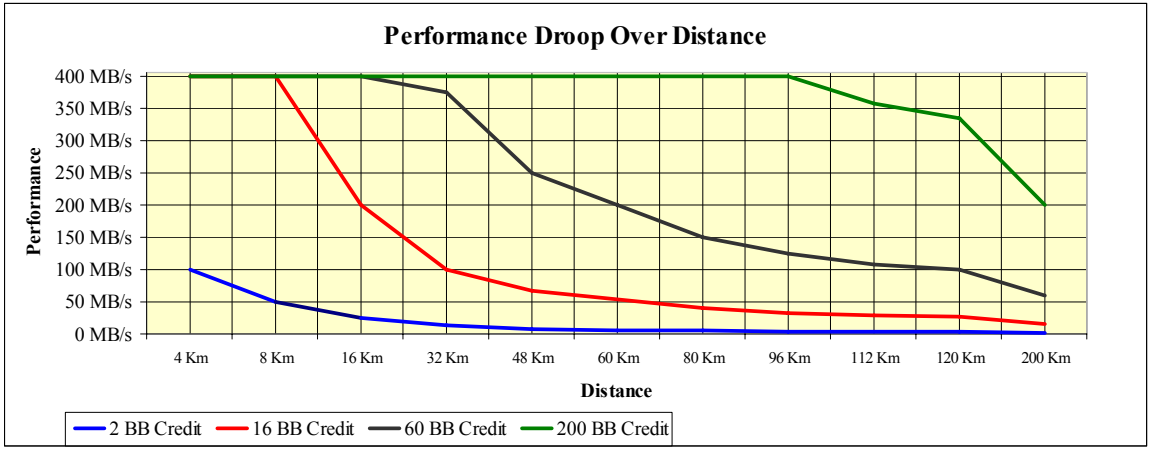


Table 42 **400 MB/s link with maximum of 200 BB_Credit**



If the environment under consideration will experience a significant performance drop due to the amount of distance involved, distance extension solutions are available to help boost the performance. Refer to Chapter 3, "Extended Distance," for more information.

Another factor that can significantly impact performance is the number of bit errors on a link. For more information, refer to "Congestion and backpressure" on page 877.

Fibre Channel port types

This section discusses types of Fibre Channel ports common to all vendors and then vendor-specific port types.

Standard Fibre Channel port types

There are several types of Fibre Channel ports, some of which are shown in Figure 235 and listed in Table 43.

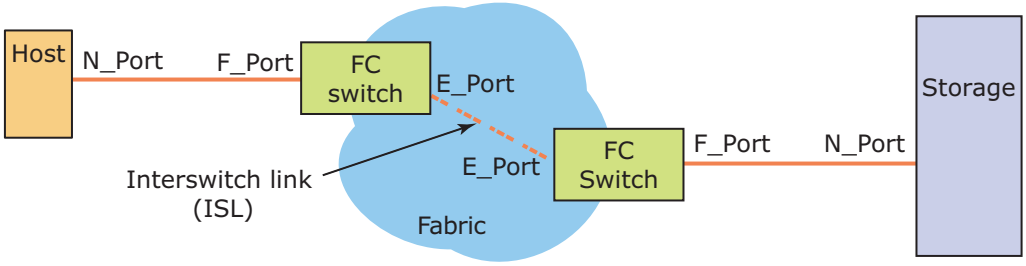


Figure 235 **Standard Fibre Channel port types**

Table 43 **Standard Fibre Channel port types**

Port Type	Description
N_Port	A port on a node outside the fabric
NL_Port	An N_Port that can also connect to a Fibre Channel arbitrated loop
F_Port	A port, located on a switching device, that connects to an N_Port and brings that connection (internally within the switch) to the fabric
FL_Port	An F_Port that can also connect to a Fibre Channel arbitrated loop
E_Port	A port, located on a switching device, that connects to an E_Port on a different switching device over an interswitch link (ISL)
G_Port	A port that can act as either an E_Port or an F_Port
GL_Port	A port that can act as E_Port, an F_Port, or an FL_Port

Vendor-specific Fibre Channel port types

In addition to the standard Fibre Channel port types each platform supports, they may also support some vendor-specific port types which provide enhanced functionality beyond what is defined in the standard. To determine which port types are supported by a specific platform, see the hardware section for that platform in Chapter 1, ”Fibre Channel SAN Topologies”. For a list of all port types supported by each platform vendor, see the following vendor sections.

Brocade

The following port types are supported by Brocade:

VE_Port:	An FCIP port on an FC switch will create a Virtual E_Port. This is physically an IP/Ethernet interface, but each FCIP tunnel “looks” like an FC E_Port to the rest of the fabric.
EX_Port:	FC Routers use EX_Ports instead of E_Ports on routed interfaces. To connect a router to a switch, you connect its EX_Port to another switch’s E_Port. The link formed is referred to as an IFL.
VEX_Port:	A port that combines both FCIP and FC Routing features to allow the creation of a Virtual EX_Port.
IFL:	The connection between an E_Port and an EX_Port is an Inter-Fabric Link.

Cisco

The following port types are supported by Cisco:

TE_Port:	Trunking port (VSAN Trunking port). This port allows a single ISL or Port Channel to transport I/O for multiple VSANs.
SD_Port:	Spanning destination port. When troubleshooting a connectivity problem, it is frequently necessary to capture a trace between the FC switch and the device being attached. This can be accomplished by inserting an analyzer inline between the device and the switch or by mirroring (spanning) the frames between the two ports to another port called the SD_Port or Spanning destination port.

FC-AL (Fibre Channel Arbitrated Loop)

When Fibre Channel was first introduced, it was a new technology and everything was expensive. Switches, hubs, and node transceivers proved to be costly. Arbitrated Loop topology lies between point-to-point and switched fabric in that it provides more connectivity than point-to-point with up to 126 NL_Ports in a loop, but less than switched fabric which has the ability in theory to support up to 16 million ports. It was a cost-effective way of connecting a limited number of ports in a loop single network.

This is considered an older technology, although it is still used. For detailed information on FC-AL, refer to "Fibre Channel arbitrated loop (FC-AL)" on page 1352.

Hubs

Fibre Channel hubs are used with Fibre Channel Arbitrated Loop (FC-AL) to increase server and storage connectivity. Using a hub, multiple servers can access multiple storage devices.

Note: EMC Symmetrix storage systems are qualified with hubs in limited configurations with HP-UX, Sun Solaris, Windows NT, Windows 2000, Windows 2003, and Siemens servers.

FC-SW (Fibre Channel switched fabric)

Fibre Channel switched fabric (FC-SW) is one or more dynamic Fibre Channel switches connecting multiple devices (see Figure 236 on page 807). FC-SW involves a switching device (the fabric) interconnecting two or more nodes. Frames are routed between source and destination by the fabric.

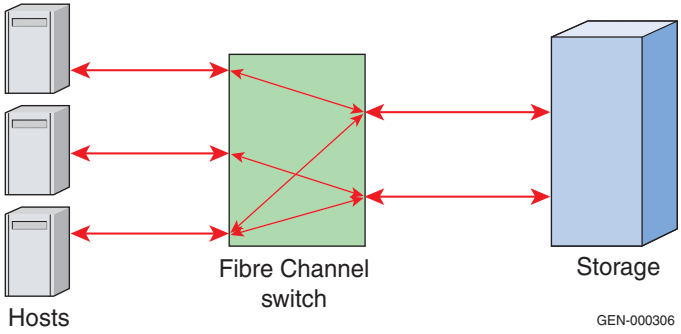


Figure 236 **Switched Fabric example**

Note: A *node* is any hardware device connected to one or more other hardware devices over Fibre Channel. Each node has at least one port that connects to ports in other nodes.

This section discusses:

- "FC-SW terminology" on page 807
- "Switches" on page 813
- "Fabrics" on page 815
- "Build Fabric process" on page 816
- "Preferred Paths / Static assignments" on page 857
- "Trunking" on page 868
- "Congestion and backpressure" on page 877
- "FLOGI" on page 899
- "Nodes" on page 900

FC-SW terminology

This section discusses terms used in FC-SW.

Embedded port

Every FC switch needs to process incoming frames which are destined for a well known address such as FFFFFE (Login Server). Some switch implementations send all requests destined for a well-known address to a special internal port, known as the *embedded port*.

PLOGI

PLOGI, or Port Login, is one port's way of requesting permission to login to another port. With the exception of SCR (State Change Registration), PLOGI needs to be sent and accepted before any

additional requests are made between two N_Ports. For more information, refer to "N_Port Login (PLOGI)" on page 912.

Name Server

Formally known as Distributed Name Server, the Name Server, defined in FC-SW-4 (Section 9.3) is responsible for name registration and management of Nx_Ports that are attached to the switch in which it resides. These entries are stored in a locally resident database. Each switch exchanges its name server information with other switches in the fabric to maintain a synchronized, distributed name service.

These exchanges of information may occur periodically but are usually triggered by a fabric-wide SW_RSCN (switch registered state change notification). Each switch in the fabric must distribute SW_RSCNs throughout the fabric whenever a change takes place in its local name server database. The exchange of information occurs as F-Class traffic.

Fabric controller

Each switch has a local entity called the Fabric controller. The Fabric controller is defined in FC-SW-4 and is responsible for managing and distributing RSCNs to those Nx_Ports that register for them. The Fabric controller is responsible only for the notification that a change has occurred; it is the responsibility of the Nx_Port to re-query the Name Server to identify what changes have occurred. The Fabric controller is also used extensively during zoning changes and the Build Fabricprocess, including the assignment of Domain IDs.

Domain IDs

A Domain ID is a unique identification number provided to each switch in the fabric. The Domain ID and the port number associated with an Nx_Port's connection to the fabric are unique parts of the 24-bit Fibre Channel address for the Nx_Port. Each switch can be assigned a preferred Domain ID, which it uses if there are no Domain ID conflicts in the current fabric.

On some platforms, each preferred Domain ID can also be set to be *persistent*. When Domain ID persistence is set, a switch will not join the fabric unless it can obtain its preferred Domain ID.

When merging two fabrics, if the same Domain ID is being used by multiple switches, a Domain ID overlap is said to exist and the switches being joined will segment from each other. When switches segment, no class F traffic or user data passes between them and the fabrics remain separate. It is possible to allow disruptive rebuilds to automatically take place when a Domain ID overlap exists, but this is not recommended since the Nx_Ports on the switch with the altered Domain ID will end up with different 24-bit FC IDs. The addition of a

switch with a unique Domain ID to a fabric causes a *build* fabric. The disruptive rebuild to fix the overlap condition is referred to as a *reconfigure fabric*.

Domain ID negotiation

EMC recommends that you set all switches to have a unique persistent preferred Domain ID, to limit the chances of incurring problems resulting from duplicate Domain IDs. An example of a problem that can occur if unique persistent preferred Domain IDs are not used is following a power event. If not all of the switches come up in a timely manner, two different fabrics may form, each using the same set of Domain IDs. Unique persistent preferred Domain IDs can also shorten the time it takes for a switch to join / rejoin the fabric, because less negotiation is necessary when the fabric is rebuilding.

Since the Domain ID is used in the addressing of the components on the switch, Domain IDs cannot be changed unless the switch is taken off line first. Taking a switch that has a Domain ID conflict off line and then setting it back on line will allow it to automatically negotiate with the principal switch for a new unique Domain ID as long as the Domain ID is not persistent. In this instance it will not be able to use its preferred Domain ID. For this reason even if switches are not currently in the same fabric, EMC recommends that all switches under your administration be given unique Domain IDs. This will allow you to merge them into a fabric at a later date without any manual intervention.

Principal switch placement and negotiation

One switch in the fabric is responsible for the distribution of Domain IDs, and plays a role in the route creation for fabric management traffic. This switch is known as the *principal switch*. Since this switch must communicate with all other switches and is the basis for fabric traffic routing, this switch should be centrally located. This will assist in the uniform delivery of information in the fabric, and provide a consistent response to fabric build events.

Principal switches are selected during both the creation of the fabric and during fabric reconstruction events. Two pieces of information located on the switch will determine the selection of the principal switch in the fabric.

The first piece of information examined during the selection process is the switch's principal switch priority setting. The switch management applications usually represent these priority settings as **Always**, **Default** and **Never**.

- A switch whose priority is set to **Always** will always be involved in the process of principal switch negotiation.
- A switch set to **Default** will continue to participate in the process if no other switch is set to **Always**.
- A switch set to **Never** will never participate in the principal switch selection process, even if there are no other switches participating.
- A fabric that has no switch capable of participating in the principal switch negotiation will segment.
- If several switches are set to always participate in the principal switch selection, the switch in this group with the lowest World Wide Name (WWN) will become the principal switch.
- If no switches are set to always participate, those switches that are set to **Default** will complete the negotiation for principal switch. The principal switch in this scenario will also be the switch that currently has the lowest WWN.

Connectrix M-series switches have the ability to set a switch priority. EMC recommends that you set selected M-series switches to **Always** based on their location in the fabric. To configure a Brocade director switch in the core of a fabric as the *principal* switch, issue the **fabricprincipal** command on the switch CLI running FOS 4.x and higher. This setting becomes active on the next reboot, or after the Build Fabric(BF) event.

FSPF Fabric Shortest Path First (FSPF) is an algorithm used for routing traffic. This means that, between the source and destination, only the paths that have the least physical hops (described under "Hops" on page 811) will be used for frame delivery.

ISL An ISL (interswitch link) is a link between any two switches in the fabric. ISLs are used to pass both data and management traffic simultaneously. Since each Fibre Channel exchange is unique, there is no limitation on the types of data (RDF, backup, management, or disk traffic) that can be transferred over an ISL simultaneously.

Principal ISL

Principal ISLs construct the path from any switch to and from the principal switch. These ISLs are assigned during the creation of the fabric, and are not necessarily the shortest path between any switch and the principal switch. The algorithm for assigning principal ISLs creates only a single path between any switch and the principal

switch. Failure of the principal ISL would cause a fabric event that would attempt to calculate a new principal ISL between the effected switch and the principal switch.

Primary paths

Primary ISL paths are those that are currently the Fabric Shortest Path First (SFPF) routes between a specific host and its storage, and are being used for data traffic. The failure of all primary ISL paths would cause data traffic to transfer to the available secondary paths that were now the SFPF routes between the hosts and storage. The link table and routing table would also be updated to reflect the changes in the topology.

Secondary paths

Secondary ISL paths are paths between a specific host and storage pair that are not currently the FSPF routes between them and, therefore, not being used for data traffic. A secondary ISL path may be *hot* in the sense that it is acting as a primary ISL path for other server / storage traffic or *cold*, indicating that it is currently not being used for any subscriber data traffic. While EMC recommends that secondary paths to data be provided in the fabric design, the number of cold paths should be evaluated for their opportunity costs, as well as for their possible usage in your sparing model.

Upstream and downstream ISLs

A principal ISL from itself to the principal switch. This principal ISL is also called the upstream ISL. All fabric F-Class traffic that needs to travel from a switch to the principal switch will travel this same path. Traffic that travels from the principal switch to another switch in the fabric will traverse the downstream ISL.

Hops

A hop is the transition between any two ports on the fabric. The number of hops in a path is the number of ISLs that must be traversed to get from the input port to the respective output port. EMC recommends minimizing the number of hops in the fabric, as well as limiting the number of hops between any switch and the principal switch. This can help the fabric to stabilize more quickly after a fabric event by minimizing the time it takes to transmit fabric change events across the fabric.

Routing

The routing algorithm for all switches uses Channels versus networks versus SANs (described under "Channels versus networks versus SANs" on page 695). Every ISL (E_Port) that enters on one port will also be assigned an exit port for each neighboring domain (switch).

The route list is dynamic only in the sense that changes in the fabric topology will cause the routing table to be recalculated. Adding and removing ISLs, as well as adding or removing switches, requires the routing table on each switch to be recalculated. Only those frames that were initially routed to a missing ISL will have to be resent and rerouted if they were lost during the fabric event.

Without events in the fabric, both the routing table and the routing of data exchanges will be static. Only the process of trunking (described under "Trunking" on page 868) will dynamically route frames across ISLs based on traffic patterns and utilization.

Switches use a round-robin approach to loading traffic across multiple same-cost routes to the same destination.

Switches do not set preferences for ISL cables that are physically shorter than others. This means that no matter how much of a distance disparity there is between two ISLs traveling to the same destination, they will still receive the same routing preference (costing). This behavior may be noticeable only when you are working with great disparities in length between similar routes. An example of when this may occur is in a fabric connected by Dense Wave Division Multiplexing (DWDM), or other link-extending equipment. DWDM links usually occur in pairs, and it is common practice to have each link travel a unique physical path to the destination.

While having the links travel in different directions protects the environment from failures caused by physical incidents damaging all of the DWDM links at the same time, this method of cabling also leads to distance disparities between the links in the DWDM ring.

RSCNs

Registered State Change Notifications (RSCNs) are usually sent out by the switch and are destined to an N_Port that has been impacted by a change in the fabric. Ports need to register for state change with the fabric controller in order to receive an RSCN.

SW-RSCNs

Switch Registered State Change Notifications (SW-RSCNs). When an N_Port logs into or out of a switch, the switch will send an SW_RSCN to every other domain in the fabric. The switch RSCNs, in combination with periodic GE_PTs between domains, keeps the distributed name server up-to-date on all switches in the fabric. When a switch receives an RSCN, it is supposed to generate an RSCN to any impacted device which has registered for State Change Notification.

NPIV N_Port ID Virtualization. A Fibre Channel facility allowing multiple N_Port IDs to share a single single physical N_Port. NPIV allows multiple initiators (FC or virtual) to occupy a single physical N_Port. For more information, refer to "NPIV" on page 930.

Switches

One dictionary definition of a switch is *a device for making, breaking, or changing the connections in an electrical circuit*. This basic definition is helpful when working with any type of switch, whether it be for Fibre Channel or some otehr protocol, such as Ethernet. A switch simply makes, breaks, and changes connections. Figure 237 shows the construction of a basic switch.

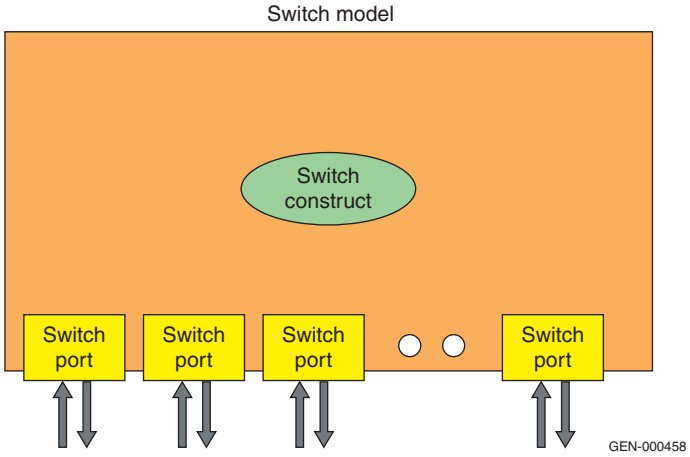


Figure 237 Generic switch construct

Although the physical implementations of switches made by each vendor are unique, they are all elxpected to provide a common set of services as defined in the Fibre Channel Fabric Generic Requirements Standards (also known as FC-FG, a copy of which can be found at http://www.t11.org). Some examples of these services are:

- Fabric F_Port/Login server
- Fabric controller
- Directory server
- Management server

Figure 238 shows a generic switch with services.

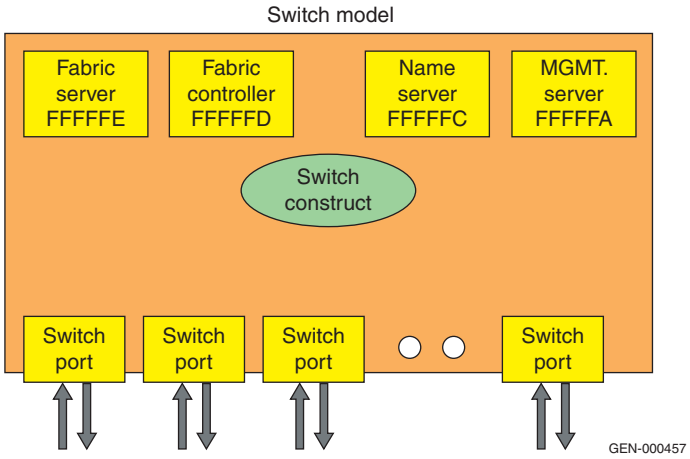


Figure 238 **Generic switch construct with services**

Every switch, regardless of manufacturer, is required to make these services available at certain *well known addresses*, discussed next.

Well-known addresses

Each generic service defined in the Fibre Channel standards FC-PH has a special Destination ID (D_ID) associated with it. A D_ID is another term for a port address. The D_ID for these well-known addresses is in the FFFFF*x* format (where *x* is a value between 0 and F). Every switch, regardless of manufacturer, is expected to perform certain functions on frames that are sent to one of these well-known addresses. Frames that are destined for a well-known address are almost always forwarded to a special internal Fibre Channel port, known as the Embedded port, making it easier for switch vendors to implement.

The fabric server

The fabric server is located at the well-known address of FFFFFE (see Figure 238). Although there are serveral possible requests that could be made to the fabric server, it is mainly used during the first part of the N_Port login process. The specific requests associated with this is actually the extended link service request FLOGI (Fabric Login).

The fabric controller

The fabric controller is located at the well-known address FFFFFD (see Figure 238 on page 814). This addres identifies has special usage

depending on the originator. If the originator is an attached expernal N_Port or NL_Port (attached through an F_Port or FL_Port), then the destination of a frame sent to FFFFFD is the fabric controller of the local switch. If the originator is the fabric controller of the local switch, then the destination of a frame sent to FFFFFD through an ISL is the fabric controller of the remote switch at the other end of the ISL.

The fabric controller also provides services to both N and NL_Ports as well as other switches.

For N_Ports, the fabric controller is usually used for State Change Registration (SCR). When an N_Port sucessfully completes SCR, it is eligible to received Registered State Change Notifications (RSCNs).

RSCNs are generated by the fabric controller and sent to all affected ports (if registered to be notified) whenever there is a change to the environment.

The directory server (or name server)

The directory server, more commonly known as the name server, is located at the well-known address FFFFFC (see Figure 238 on page 814). It is also used during the N_Port initialization process when an N_Port wants to register its FC-4 type and then later query for other FC-4s that are available through zoning and currently logged into the switch.

Route and path selection

Route and path selection is not provided at a well-known address, but is a service provided by all switches. Frames are routed based on their D_ID. (A common misconceptions is that they are routed based on their WWN.) In general, frames are started down the proper route by the ASIC on the receiving port of the switch or by information provided by the receiving port to some other component within the switch. This is also true in a multiswitch fabric.

Fabrics

A fabric is a collection of switches that have been linked together through a Fibre Channel cable, also known as an ISL (interswitch link). The term fabric is generally used to indicate a group of switches, although it is technically correct to refer to a single switch that is not connected to other switches as a *single switch fabric*. Each switch can effectively be thought of as its own domain and, in fact, each switch has a unique Domain ID associated with it.

Figure 239 shows an example of a switched fabric.

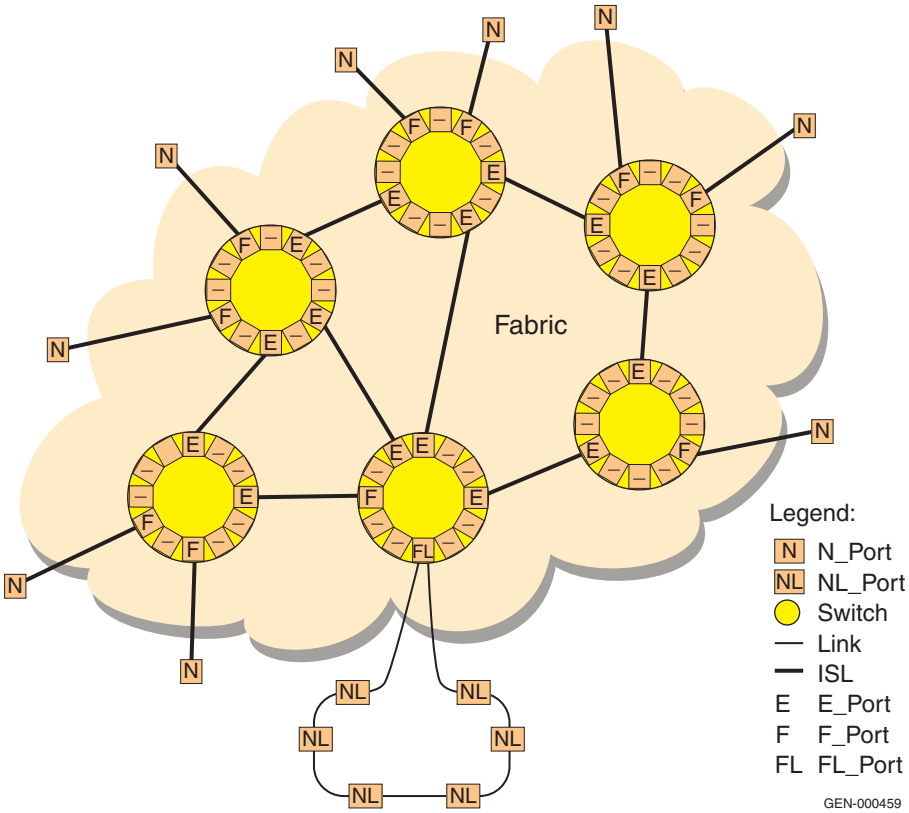


Figure 239 Switched fabric example

Build Fabric process

There are numerous books and reference materials available (including the *Fibre Channel Standard FC-SW*) which provide thorough explanations of the Build Fabric process (also referred to as Fabric Configuration), down to the bit level. The information provided in this section will include only the major concepts of the Build Fabric process with the intention of both familiarizing the reader with the process of fabric creation and facilitating future investigation.

Topology

Throughout this section, the topology shown in Figure 240 on page 817 will be used to illustrate the Fabric Configuration process. The architecture is intended to maximize the readers' understanding

of the process and is not intended to be an endorsement of the topology itself.

Figure 240 Fabric Configuration process

In this example, we will assume that a fiber cable has just been connected between switch D and E. The first step that must be taken is link initialization.

Negotiate port type

When two FC ports are connected, they will negotiate speed and bring the link up to the active state (Figure 241).

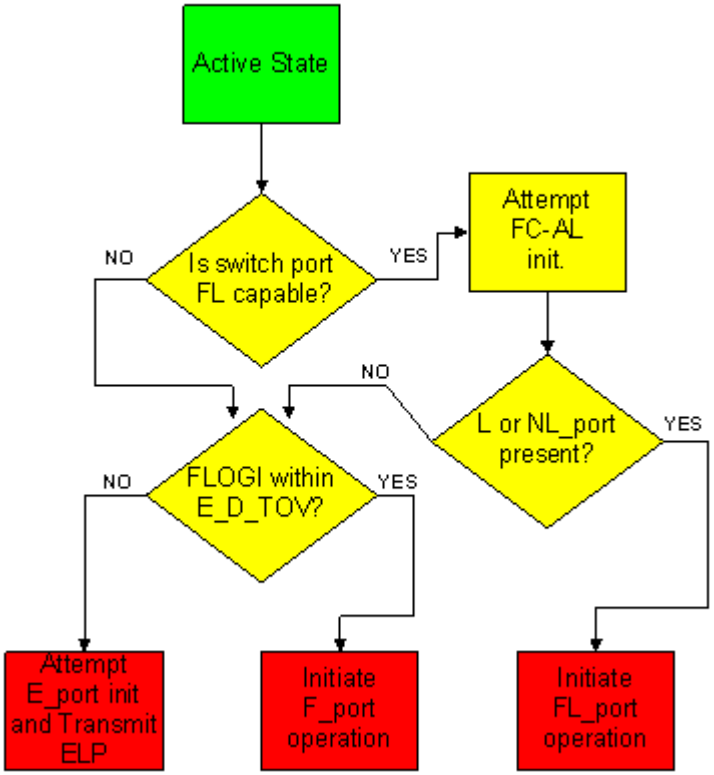


Figure 241 **Switch port initialization**

In the case of an HBA connecting to a switch, the HBA will send FLOGI which indicates to the switch that a host is connecting to the switch. Therefore, the port on the switch will become an F_Port.

The same is not true when two switches are connected together. In this case, the switches usually wait a period of time for a FLOGI. If one is not received and the switch is configured to allow E_Port connections on the port that just came up, the switch will send out an ELP (Exchange Link Parameters) frame and try to initialize the port as an E_Port. If the ELP is accepted by the port on the other end of the link, then the switch will initialize the port as an E_Port and the Fabric Configuration process begins.

Fabric Configuration (Build Fabric) process

The purpose of the Fabric Configuration process is to allow for the formation of ISLs (Interswitch links) between different domains and to facilitate Nx_Port connectivity between them.

Note: For the sake of clarity, the word "domain" will be used in place of "switch" throughout this section .With the advent of VSANs and partitions, many domains can reside on the same switch physical switch.

The FC standard breaks up the Fabric Configuration (also referred to as the Build Fabric) process into the following logical chunks:

- Exchange Link Parameters
- Principal switch selection
- Domain ID acquisition
- Zoning Merge
- Path selection

For clarity, we will slightly reorganize them as follows:

- "ELP – Exchange Link Parameters" on page 820
- "Perform Link Reset (LR)" on page 823
- "ESC - Exchange Switch Capabilities" on page 823
- "EVFP – Exchange Virtual Fabric Parameters" on page 823
- "Principal switch selection" on page 825
- "EFP – Exchange Fabric Parameters" on page 825
- "Build Fabric (BF)" on page 826
- "EFP – Exchange Fabric Paramters (Selecting the principal switch)" on page 828
- "Domain ID acquisition" on page 831
- "Zoning merge" on page 835
- "Path selection and the FSPF protocol" on page 837
- "Assigning egress ports" on page 855

Each of these will be further discussed in this section.

Note: The frames that are sent between switches used a class of service called Class F. This means they all begin with the Start of Frame Fabric (SOFf) delimiter and are all Acknowledged by the receiving switch, like Class 2. For simplicity, the ACK frames have been left out of all the examples in this section, but it is important to realize that they are used throughout the process.

ELP – Exchange Link Parameters

Once the link has been initialized, the first frame that a switch will transmit is the ELP frame (refer to Figure 242 on page 822). The ELP is transmitted from one Fabric Controller to another (FFFFFD to FFFFFD) and contains the following information:

- R_A_TOV (Resource Allocation Time Out Value) — This value needs to be the same on both switches or the ISL will isolate (segment). This typically appears in the user interface as either an ELP failure or as incompatible flow control parameters.

  Note: Originally, the value of R_A_TOV was supposed to be equal to the value of E_D_TOV plus two times the maximum amount of time that a frame can exist in the fabric. However, today it is always left at a value of 10000 ms (10 seconds) and is not tied to how long a switch will hold onto a frame.

  It is important to note that the amount of time a frame can exist in the fabric depends on the number of hops between two NX_Ports (the number of ISLs that need to be crossed) and the value of the hold timer used by the switch being crossed. Currently, the hold timer value can vary from 500 ms up to 2000 ms. The number of hops can be up to 3 in EMC-supported configurations. See "FLOGI" on page 899 for more information.

- E_D_TOV (Error Detect Time Out Value) — This value needs to be the same on both switches or the ISL will isolate. This typically appears in the user interface as either an ELP failure or incompatible flow control parameters.

  The value of E_D_TOV is used for many different purposes, but in each case the E_D_TOV timeout value is used for detecting an error condition. An example that frequently occurs is when a port has zero transmit BB_Credit for E_D_TOV; the port is supposed to transmit the LR primitive to reset the BB_Credit counters to their login values.

- WWPN (World Wide Port Name) — The unique 64 bit address of the port transmitting the ELP.
- Switch name — The unique 64 bit address of the Domain to which the port transmitting the ELP belongs.
- Receive data field size — The maximum payload that is allowed by the switch.
- Flow control parameters —
  - Type of flow control being supported by the switch, i.e., R_RDY, VC_RDY.
  - Number of BB_Credits being granted by the switch — In configurations where distance is going to be a consideration (generally > 10 km), the BB value exchanged between the two FC ports has a significant impact on the maximum sustainable throughput rate.

Figure 242 shows the Exchange Link Parameters (ELP) process.

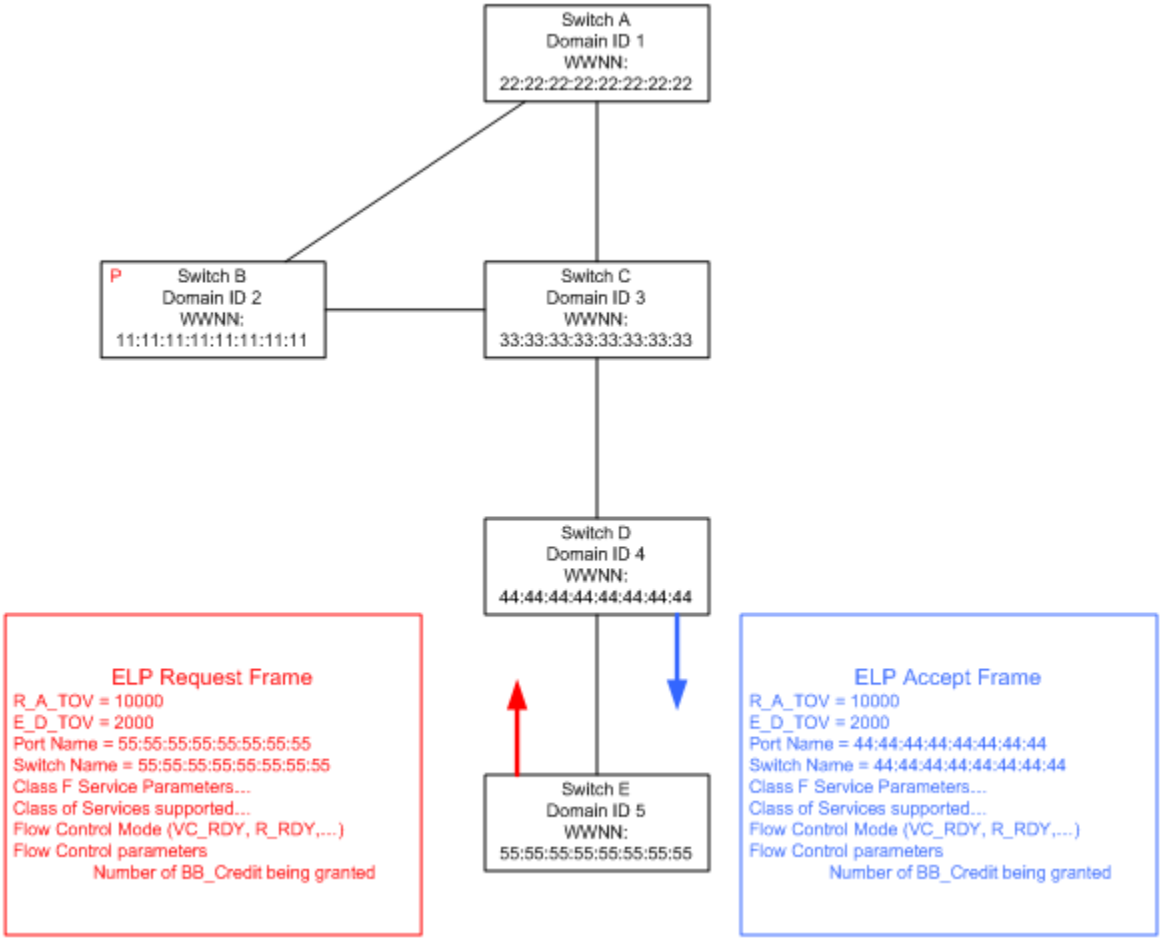


Figure 242 ELP Exchange Link Parameters process

If the port receiving the ELP at the other end of the link is a switch port and is configured to allow E_Ports, the ELP will be processed by the receiving switch. Depending on the switch vendor, features enabled, basic security settings (like fabric binding, switch binding, etc.), the ELP will either be *accepted* (if everything checks out) or *rejected* (if there is a mismatch that cannot be resolved).

If the ELP is accepted, then an ELP Accept frame is sent from the receiving switch back to the ELP originator. As with all of the Exchange requests, the contents of the ELP accept are identical in

type to the ELP request, except it is populated with information specific to the switch accepting the ELP.

If the ELP is rejected, both ports (one on each end of the physical link) should be shown as isolated (segmented).

Perform Link Reset (LR)

As mentioned in "ELP – Exchange Link Parameters" on page 820, two of the link parameters exchanged during ELP are:

- the type of flow control to be used, and
- the number of BB_Credits being granted by each switch port.

Once this information has been exchanged, the Link Reset (LR) protocol is initiated to initialize each port to use these login values. The Link Reset protocol is kicked off by one switch transmitting Link Reset primitive sequences in place of IDLEs. When the other port receives and recognizes the LRs, it resets its credit value to the login value and starts transmitting Link Reset Response (LRR), primitive sequences instead of IDLEs. When the port transmitting LRs receives and recognizes the LRRs, it resets its credit value to the login value and begins to transmit IDLEs again. Once the IDLEs are received by the other port, it begins to transmit IDLEs and both ports are back at the active state. The Link Reset protocol can be invoked for many other reasons as well.

ESC - Exchange Switch Capabilities

After the ELP is completed, the switches have the option to exchange switch capabilities information. Outside of vendor-specific information, the primary use of the ESC request is to determine if the switches support virtual fabrics. If a switch supports virtual fabrics, it will list this capability in the ESC payload. If the responding switch also supports virtual fabrics, it will do the same. If both switches support virtual fabrics, the next frame to be transmitted by either switch will be the Exchange Virtual Fabric Parameters (EVFP). Otherwise they will continue on with the EFP (Exchange Fabric Parameters).

EVFP – Exchange Virtual Fabric Parameters

If both switches list support for virtual fabrics in their ESC / ESC accept, the EVFP will be used to share which virtual fabric features will be supported. The two features that will be discussed in this section are *Virtual Fabric Identifiers* and *Virtual Fabric Tagging*.

Currently there are two basic types of virtual fabric implementations: those that support virtual fabric tagging and those that do not.

An example of an implementation that supports virtual fabrics but does *not* support virtual fabric tagging is the Brocade MI10k. Starting with Firmware EOSn 9.0, users can configure each port on the Brocade Mi10k to have its own unique Virtual Fabric Identifier or VF_ID. Ports in the same partition with the *same* VF_ID *can* all communicate with each other. Ports with *different* VF_IDs *cannot* communicate with each other. If two Brocade MI10ks running EOSn 9.x or higher are connected to each other in an attempt to form an ISL, they will send EVFP,. If they have the same VF_ID they will continue along the Fabric Configuration process. If they do not have the same VF_ID, they will segment from each other with a virtual fabric conflict.

One thing to note about this implementation is that for every VF_ID you want to have communicating between the two Brocade Mi10ks, you will need to have a separate ISL assigned exclusively to each VF_ID. They cannot share the ISL between the different VF_IDs since they do not support VF tagging so there is no way for the switch port at the other end of the ISL to determine which virtual fabric to route the incoming frame to. This is why virtual fabric tagging and the Virtual Fabric Tagging Header (VFT_Header) becomes important. Figure 243 shows a normal FC Frame and another that has the 8 byte VFT_Header prefix added to it.

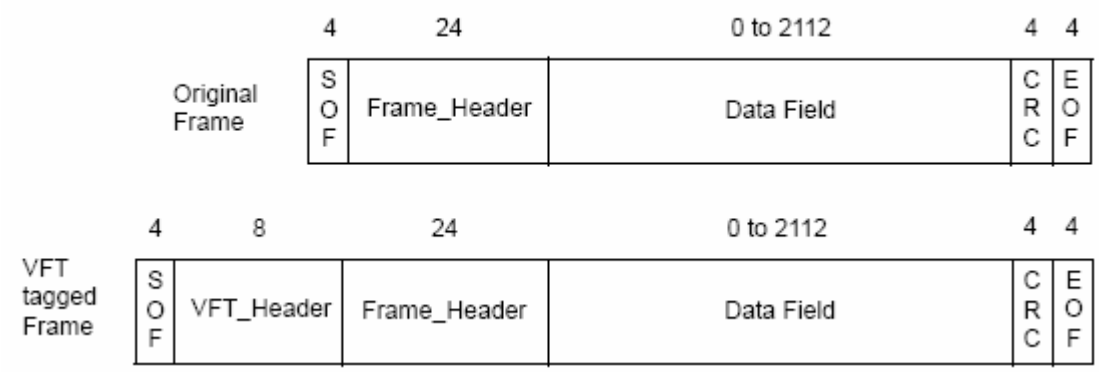


Figure 243 Normal FC Frame and a VFT_Header Frame

Implementations that use the VFT_Header can accommodate multiple virtual fabrics simultaneously using the same ISL. An example is Cisco and their VSAN trunking feature.

Once each switch has shared its supported virtual fabric features, the ISL will be in one of the following three states:

- **Isolated** — VF_ID is mismatched.
- **Active without VFT Headers** – Class F traffic is allowed to flow between the two domains and the Fabric Configuration process can continue.
- **Active with VFT Headers** (i.e., EISL) – Class F traffic from multiple VF_IDs (VSANs) are allowed and the Fabric Configuration processes can continue.

Principal switch selection

During the Fabric Configuration process, each fabric selects one switch to act as the principal switch. The role of the principal switch is to assign Domain IDs to the rest of the switches in the fabric. The principal switch selection process begins with the Exchange of Fabric Parameters (EFP).

EFP – Exchange Fabric Parameters

The EFP exchange is used several times during the Fabric Configuration process as follows:

- After an ISL has been established and before the Build Fabric request is sent out.

  Note: The purpose of this initial EFP is to ensure that a Domain ID overlap does not exist between the two fabrics being joined.

- After the Build Fabric is transmitted.
- After the principal switch is selected.
- After each switch is assigned a Domain ID.

This section covers the initial use of EFP. The subsequent EFPs are covered in the following sections: "Build Fabric (BF)" on page 826 and "EFP – Exchange Fabric Paramters (Selecting the principal switch)" on page 828.

The EFP request payload consists of the principal switch name (WWNN) and its priority as well as a list of Domain IDs already in use. The list of Domain IDs are each associated with a specific switch name. In this way, during the initial EFP, Domain ID overlaps can be detected by comparing the list of Domain IDs and ensuring that if the same Domain ID exists on both sides of the ISL, that it is associated with the same switch name. If the same Domain ID exists but is

associated with a different switch name, the ISL will isolate and no additional Class F traffic will be exchanged on that ISL, in other words, the Fabric Configuration process ceases on that ISL.

In Figure 244, since there is no Domain ID overlap, the Fabric Configuration process continues with Build Fabric.

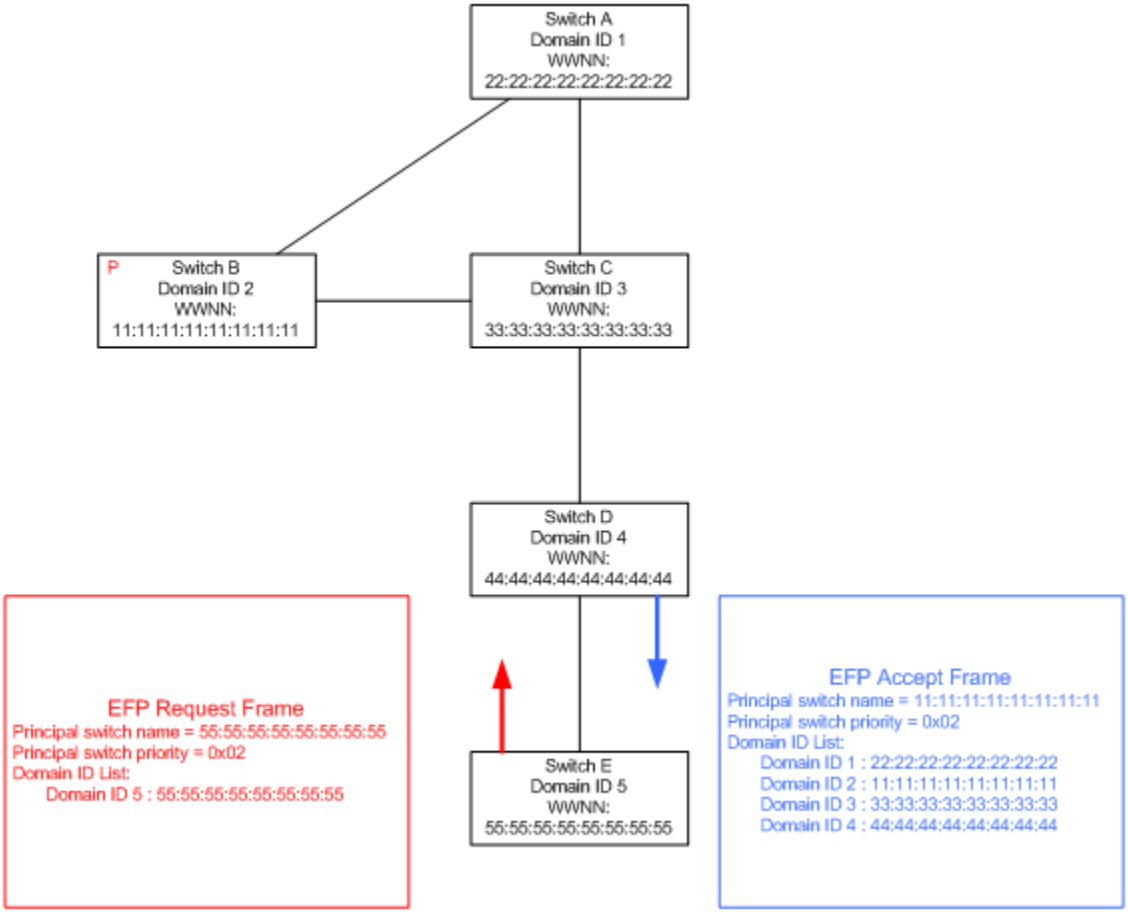


Figure 244 **EFP request payload**

Build Fabric (BF)

If the initial EFP has completed and has not resulted in an isolated ISL, the Fabric Configuration process continues with Build Fabric (BF) being transmitted on all E_Ports that have not already received a Build Fabric from the switch on the other end of the ISL. The Build

Fabric request payload consists of a single 4 byte word which contains the Build Fabric command itself.

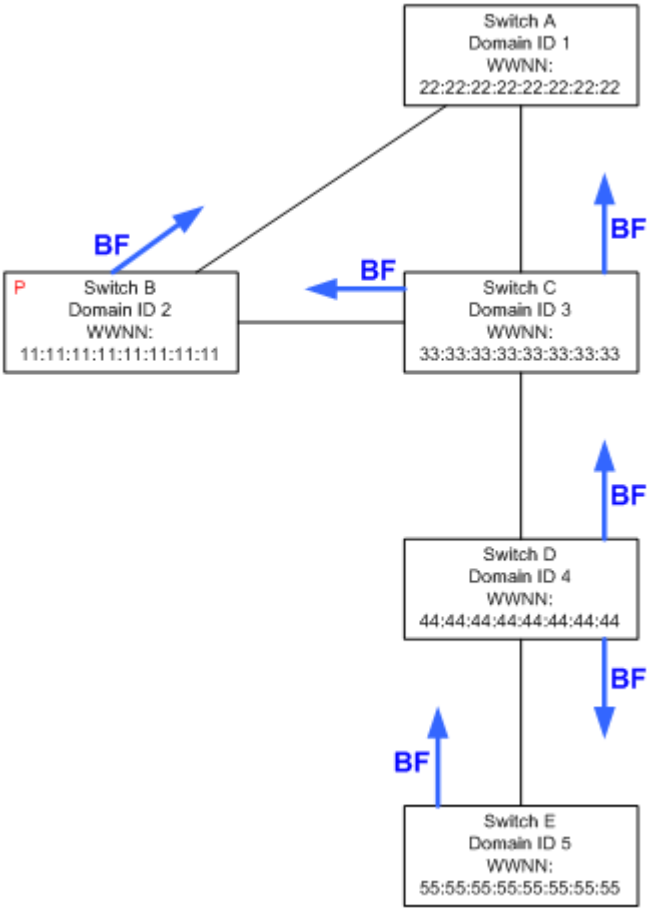


Figure 245 Build fabric

In Figure 245, BF is sent from switch D and E at the same time. Switch D sent the BF to both C and E and switch E sent the BF to switch D. All three switches should accept the BF from the other switches. When switch C receives the BF, it will accept it and then transmit BF out all E_Ports that it has not yet received a BF on. Finally switch B will transmit BF to switch A.

As soon as the switch transmits the BF, it must wait a period of time called F_S_TOV (Fabric Stability Time Out Value) before continuing

on with the next phase of Fabric Configuration. F_S_TOV is defined as 5 seconds.

EFP – Exchange Fabric Paramters (Selecting the principal switch)

Once F_S_TOV has expired, the switch can start initiating and responding to EFP requests. These EFPs are used to select the principal switch. Many different starting conditions and ending configurations are discussed in the EFP section of the *Fibre Channel Standard FC-SW*. A few of the more important concepts are:

- When a switch initially transitions from offline to online, it does *not* have a Domain ID assigned to it. As a result, if this switch were connected to other switches in the fabric while it was coming online, it would send out an EFP containing a Domain ID list with zero entries in it. In this case, the switch coming online is not eligible to become the principal switch unless there is no switch in the fabric capable of becoming the principal. The switch coming online will take note of the first ISL that an EFP was received on and treat it as the principal ISL.
- When two switches are joined, and each has a non-zero Domain ID list as well as a switch priority other than 0xFF, both switches are candidates to become the principal switch. The principal switch is determined by each switch exchanging EFPs with every other switch it is attached to and initially comparing its own value of switch priority and switch name with the other switch to which it is attached.

  When a switch discovers another switch with a lower priority, or a switch with the same priority but a lower switch name, it starts sending out this other switch's priority and switch name instead of its own in subsequent EFPs. This process continues for a time equivalent to 2 x F_S_TOV. The switch that has retained its own switch priority and switch name at the end of 2 x F_S_TOV is declared to be the principal switch.

  During this process, whenever a switch discovers a switch priority / switch name combination lower than its own, it takes note of the port it was received on as a potential candidate for principal ISL. At the end of this process, 3 x F_S_TOV (15 seconds) has elapsed since the start of the Fabric Configuration process. The initial F_S_TOV was after sending BF, while the second and third were during the principal switch selection process in which EFP was used for the second time.

Example of EFP usage during the principal switch selection process

After the BFs were sent and F_S_TOV expired, all switches in the fabric will begin exchanging EFPs. One of the rules for accepting a BF is that the switch will clear its Domain ID list and replace the principal switch priority and principal switch name fields with its own values in subsequent EFP requests or responses.

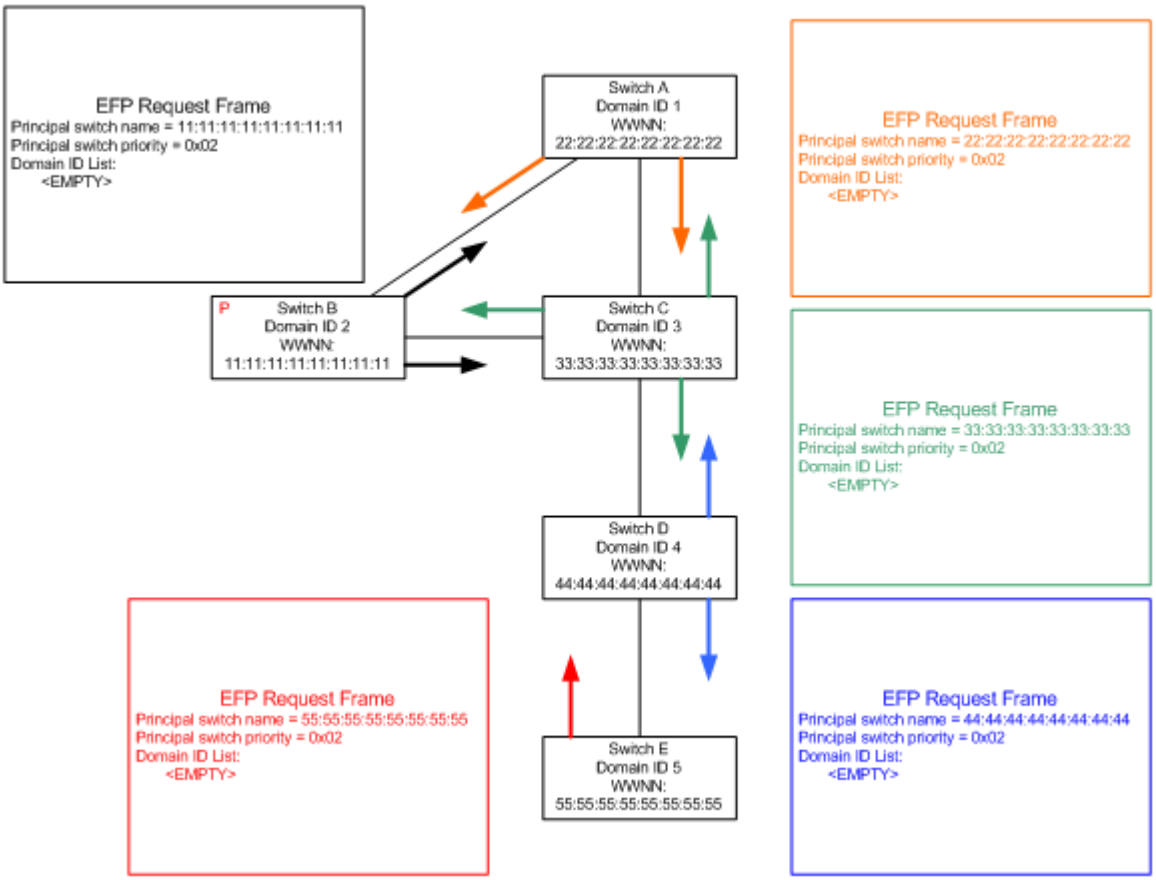


Figure 246 **Principle switch selection process 1**

In Figure 246, the colors of the arrows illustrate what data is being sent in the EFP request. The EFP request is shown in the corresponding colored box.

As shown in Figure 247 on page 831, when switch B receives the EFP from switches A and C, it will compare their priority and principal switch name with its own. Since switch B's priority is the same, but its switch name is lower, switch B will continue to transmit EFPs and EFP accepts with its own switch name and priority.

When switch A or C receive the EFP from switch B, they will compare the priority and principal switch name with their own. Since switch B's priority is the same but switch name is lower than either switch A or C, they will replace their value of principal switch name and principal switch priority with that of switch B and transmit that value in subsequent EFP and EFP Accepts.

When switch D receives the EFP from switch C and E, it will notice that switch C's switch name is lower and begin to transmit switch C's switch name in subsequent EFP requests or accepts.

When switch E receives the EFP from switch D, it will begin to transmit switch D's switch name since it is lower than E's.

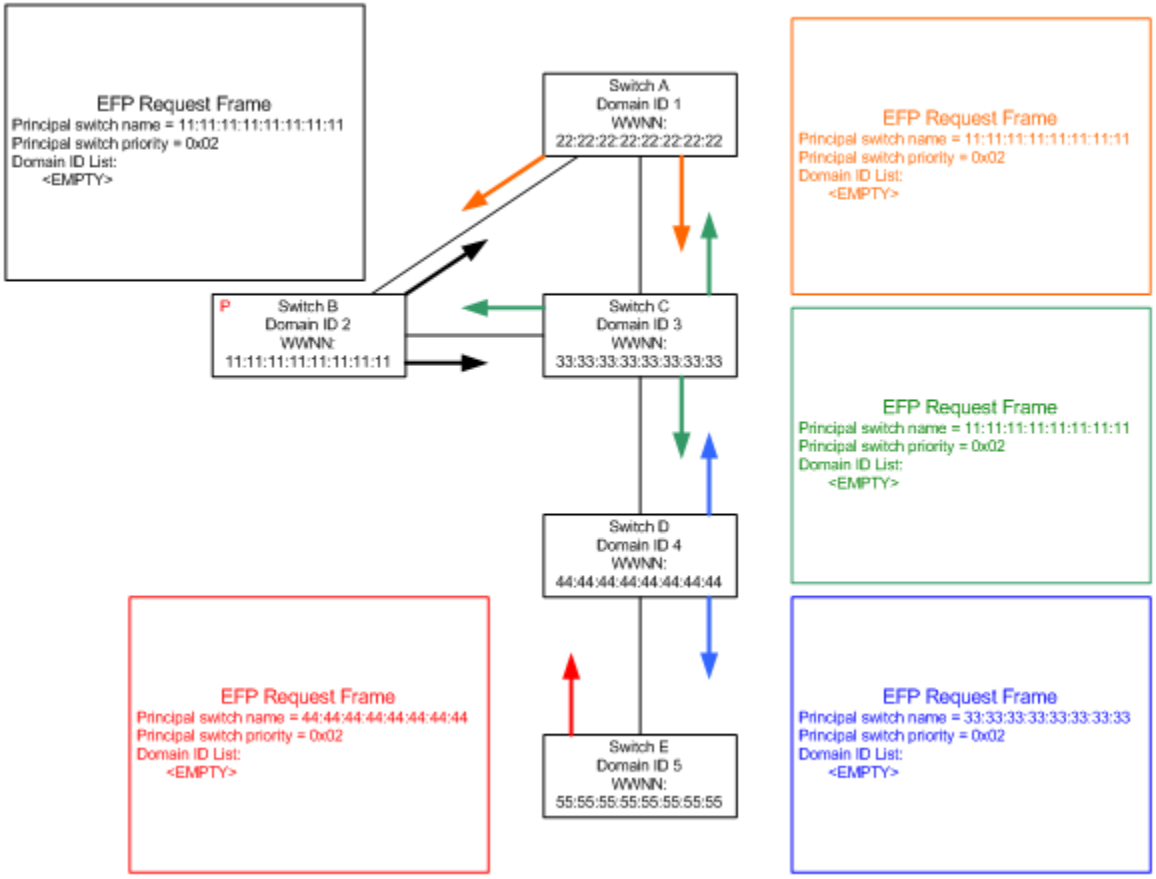


Figure 247 **Principle switch selection process 2**

Finally, switch D will receive the EFP that contains switch B's switch name and transmit it to switch E, which will then begin to transmit switch A's switch name.

After all the EFPs have been exchanged, all switches will agree that the switch with a switch name of 11:11:11:11:11:11:11:11 is the principal switch.

Domain ID acquisition

Once the principal switch has been selected, it will:

1. Assign itself a low switch priority value (0x02).
2. Clear its local copy of the Domain ID list.

3. Grant itself a Domain ID, (typically the last one it was using).
4. Transmit a Domain Identifier Assigned (DIA) on all E_Ports.

DIA - Domain Identifier Assigned

The DIA indicates that a principal switch has been selected and that the upstream switch has been assigned a Domain ID[1].

Once the principal switch has been selected and the appropriate amount of time has transpired, the principal switch will send out DIA on all E_Ports (refer to Figure 248 on page 833). Once the receiving switch accepts the DIA, it can send an RDI (Request Domain ID) to the principal switch via the principal upstream ISL. (The principal upstream ISL was determined when this switch received the EFP with the lowest switch priority / switch name.) The DIA request consists only of the switch name of the switch that initiated the DIA.

After sending the DIA, and before sending or responding to any EFPs, the principal switch will assign itself a Domain ID, typically the one it had previously. Once this is done, the principal will generate an EFP to inform the other switches in the fabric of the new switch in the fabric.

1. FC-SW-4 section 6.1.6

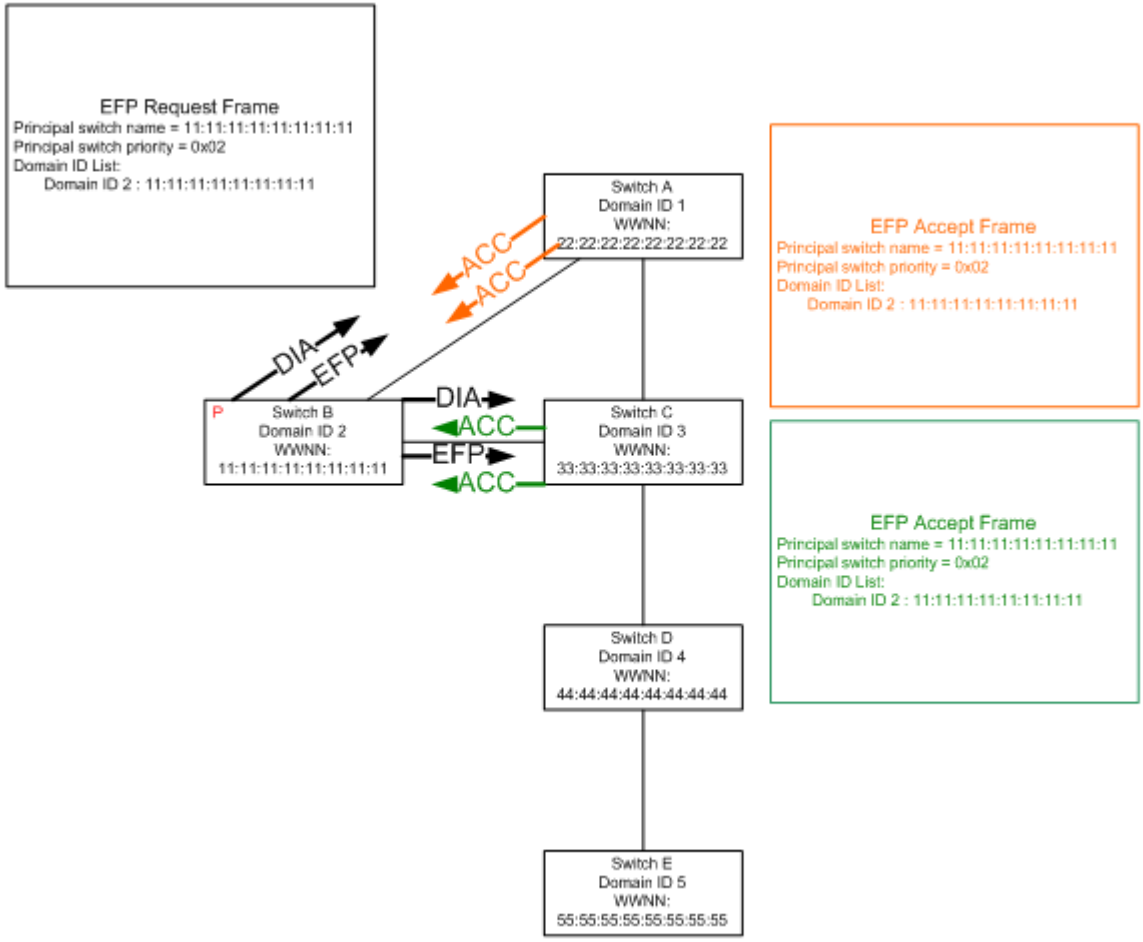


Figure 248 **Domain Identifier Assigned (DIA)**

RDI – Request Domain Identifier

The RDI allows a switch to request a specific Domain ID or let the principal switch just assign one to it (refer to Figure 249 on page 834). When the principal switch receives the RDI, it takes note of the port on to which it was received to use in the future as the principal downstream ISL. At this point, the principal switch will process the RDI and assign a Domain ID to it, if possible. If a Domain ID is available, an *Accept* will be sent back to the switch that sent the RDI via the principal downstream ISL. Once the Accept has been sent, the principal switch will transmit an EFP containing the updated list of

Domain IDs being used. Assuming this is the first RDI to be processed, the Domain ID list in this EFP will contain the principal switches Domain ID and switch name as well as the Domain ID and switch name of the switch which was just assigned the Domain ID. This process will be repeated for each RDI that is processed by the principal switch. The purpose of this EFP is to notify all switches in the fabric of the other Domains and their switch names as they come into the fabric.

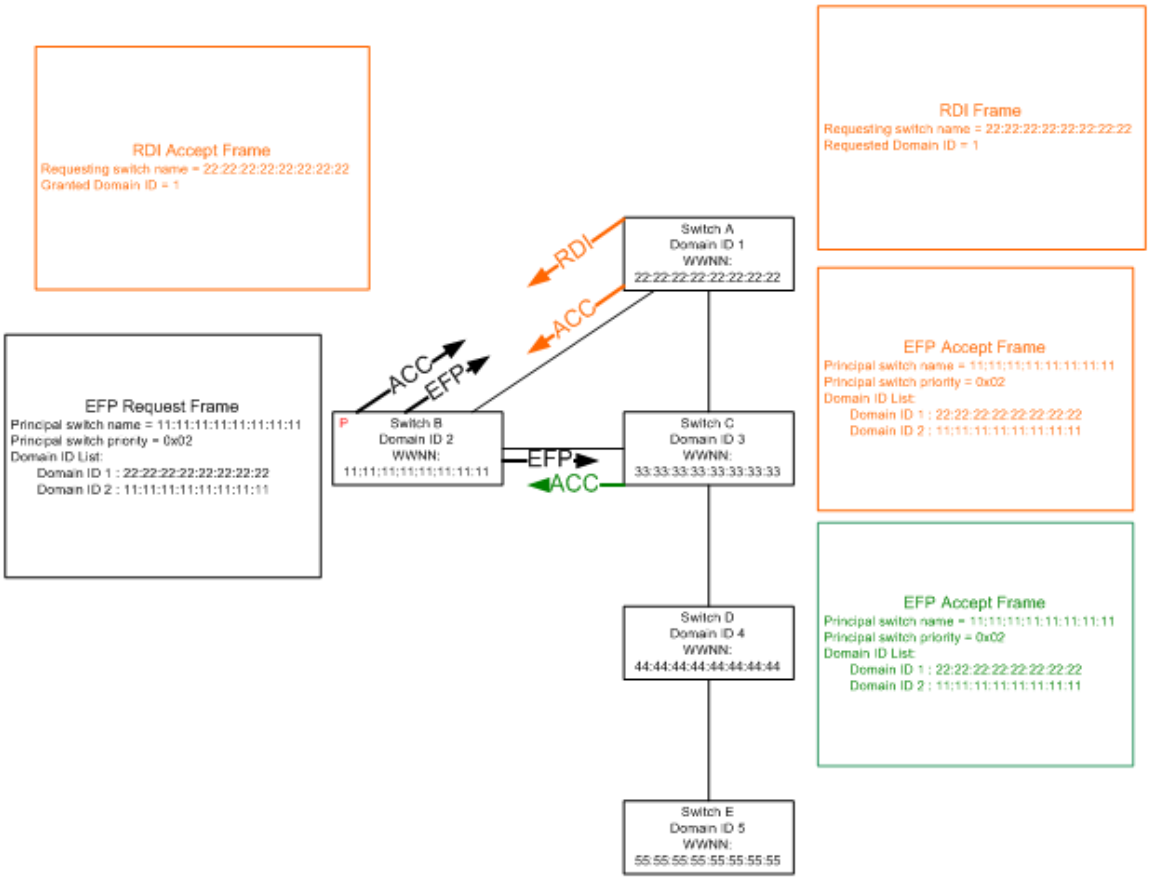


Figure 249 Request Domain Identifier (RDI)

Once the switch has been assigned a Domain ID, it will transmit DIA on all E_Ports other than the principal upstream ISL. These downstream switches will handle the DIA in a manner similar to what is described earlier. However, the RDI from downstream

switches will be forwarded to the principal via the principal upstream ISL and the Accept from the principal will be forwarded back to the originator via the principal downstream ISL. The principal will also generate another EFP to indicate another switch has been assigned an EFP. After the switch receives the Accept to its RDI, it will transmit a DIA on all E_Ports other than the principal upstream ISL.

The process of granting Domain IDs continues until all Domains have been assigned and an additional time period of F_S_TOV expires. The next step in the process is the merge of zoning information.

Zoning merge

After all of the Domain IDs have been assigned, the merge of zoning information is performed. The method in which the merge is completed, and the type of information shared between two switches, varies widely from one platform or operating mode to another. As a result, the information in this section will not be applicable to all environments. However, the diagram from FC-SW-4 and the descriptions listed in Figure 250 on page 836 should give you an idea of how it is performed.

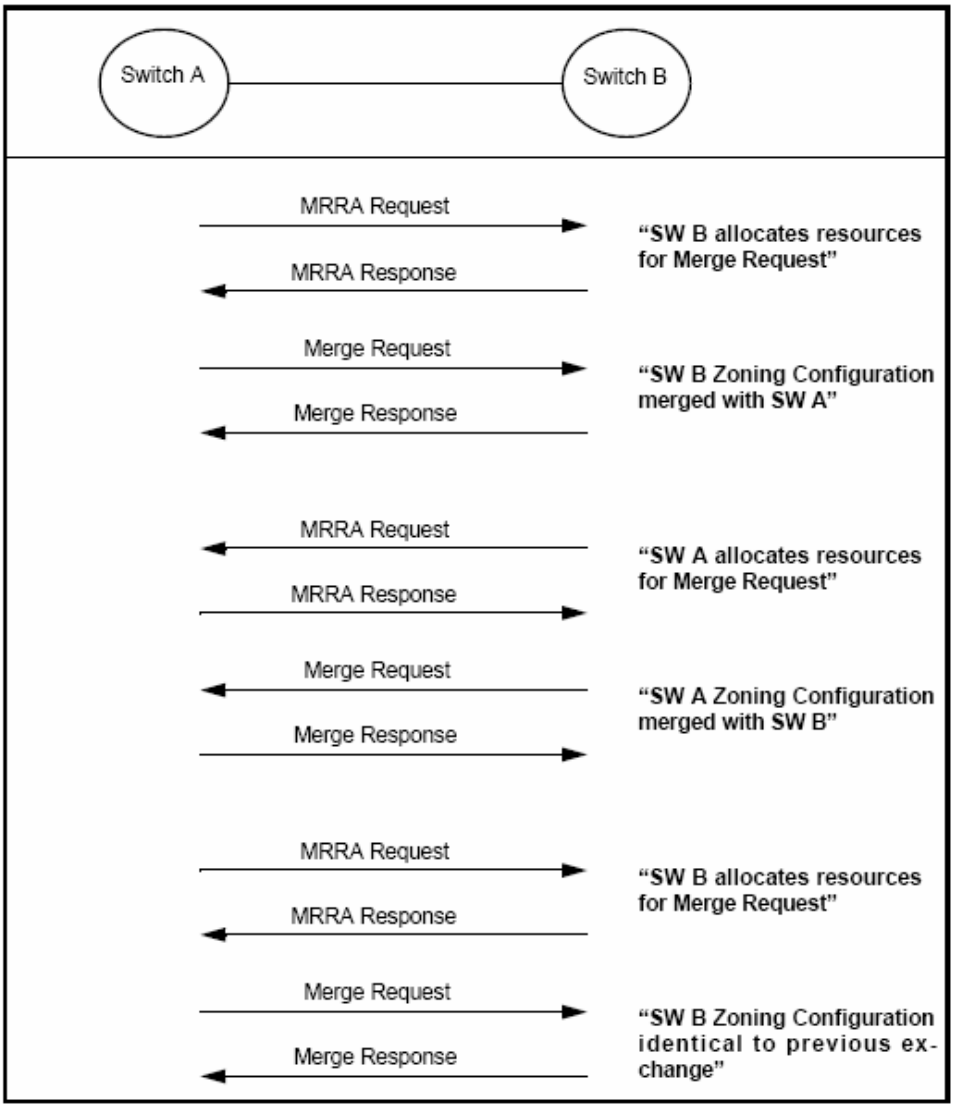


Figure 250 **Exchange of zoning information**

After the ISL has formed between switch A and B and they have both been assigned a Domain ID, the merge process begins with one switch sending another a Merge Request Resource Allocation (MRRA).

MRRA – Merge Request Resource Allocation

The purpose of the MRRA is to ensure that the switch on the other end of the ISL has enough memory available to store the incoming zoning information. When the MRRA is accepted, an Merge Request (MR) is sent from the switch that requested the resource to the switch that accepted the resource request.

MR – Merge Request

The MR contains the actual zoning information, which can consist of differing pieces of information, depending on if enhanced zoning is supported or not. At a minimum, the Merge Request will contain the zone names and zone members for each zone.

Once the zoning information is identical between the two switches, the Fabric Configuration process can continue and path selection begins.

Path selection and the FSPF protocol

When all switches have been assigned a Domain ID and the zoning information has been exchanged, the path selection process can begin. The topology shown in Figure 251 on page 838 will be used for the rest of this example.

Please note that additional ISLs have been added between switch C and D as well as between switch D and E. Also notice that each end of all ISLs have port numbers associated with them. The purpose for these additions is to help facilitate the readers' understanding not only of the path selection process but also to help when we discuss "Trunking and ISL aggregation" on page 869.

Please note that in the next few topology diagrams, the switches icons are represented as different colors.

- **Grey** switches represent switches in an **unknown** FSPFstatus.
- **Red** switches are **unsettled** from an FSPF algorithm point of view.
- **Green** switches are **settled** from an FSPF algorithm point of view.

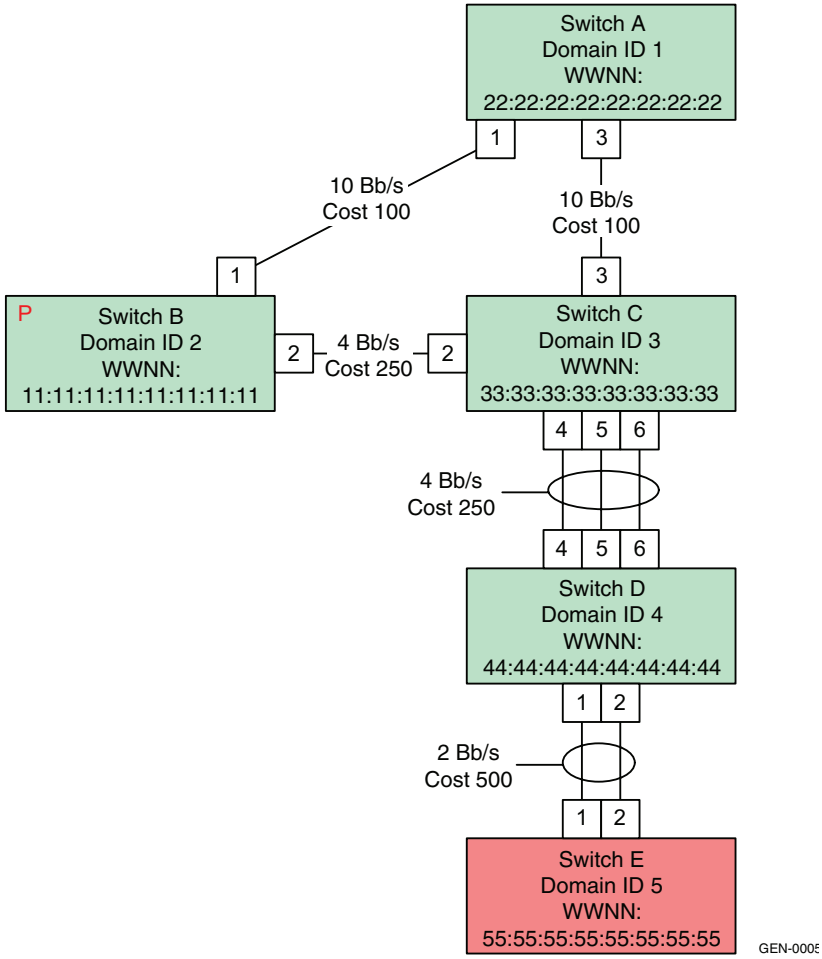


Figure 251 **Path selection topology**

For the sake of simplicity, the path selection process is broken down into five phases, each further discussed in this section:

- "HLO – Hello initial exchange" on page 839
- Database exchange

Note: Refer to "LSU – Link State Update" on page 840 and "LSA – Link State Acknowledgement" on page 842

- "Dijkstra's algorithm for determining the shortest path" on page 843
- "The distributed Name Server, SW_RSCN, and GE_PT" on page 853
- Port assignment

Note: Refer to "Adding Nx_Ports and setting up routes" on page 854.

HLO – Hello initial exchange

Path Selection begins with each switch transmitting HLO out every E_Port. The format of the HLO frame is displayed below on the left and the actual data in the initial HLO from Switch E transmitted out port 1 is shown on the right. In this section, we have tried to avoid showing frame payload contents wherever possible, but the following three commands **HLO**, **LSU**, and **LSA** are used so frequently, and the contents displayed in so many different places in the UIs, that it is advisable to become familiar with how they are laid out.

bits word	31 … 24	23 … 16	15 … 8	7 … 0	
	SOFf				
0	FT-1 Header				
1					
2					
3					
4					
5					
6	Command				FSPF HEADER
7	Version	AR #	Auth Type	Reserved	
8	Originating Domain ID				
9	Authentication				
10					
11	RESERVED				
12	Hello Interval				
13	Dead Interval				
14	Recipient Domian ID				
15	RESERVE	Originating Port Index			
	CRC				
	EOFn				

bits word	31 … 24	23 … 16	15 … 8	7 … 0
	BC	B5	58	58
0	02	FF	FF	FD
1	00	FF	FF	FD
2	22	38	00	00
3	54	00	00	00
4	00	D7	FF	FF
5	00	00	00	00
6	14	00	00	00
7	02	00	00	00
8	00	00	00	05
9	00	00	00	00
10	00	00	00	00
11	00	00	00	00
12	00	00	00	14
13	00	00	00	50
14	FF	FF	FF	FF
15	00	00	00	01
	67	75	4E	D9
	BC	95	D5	D5

Figure 252 Hello frame format and payload

The fields to take note of are:

Command	This is a 4 byte field. For HLO frames it is always set to 14 (00000014).
Originating Domain ID	This is also a 4 byte field and since this HLO is from switch E, it will be set to 05 (00000005).
Hello interval	The length of time in seconds that should elapse between the transmission of each HLO frame. In this case it is set to 0x14 or 20 seconds.
Dead Interval	The length of time in seconds that the switch should wait for an HLO frame before declaring the link dead and removing the link as a valid path from its Link State Database. In this case it is set to 0x50 or 80 seconds.
Recipient Domain ID	Since this is the first HLO being sent between the two switches, neither switch knows the Recipient Domain ID. As a result, it will be set to FFFFFFFF. Once the switch receives a HLO from the other end of the link, it will update this field with the other switches' Domain ID.
Originating Port Index	The port index of the port that transmitted the HLO.

After the first HLOs have been exchanged between all of the switches, they can begin transmitting LSU and perform initial database synchronization.

LSU – Link State Update

The LSU is used to push each switch's Link State Record (LSR) out to the rest of the fabric. There are two basic types of LSUs:

- Database Synchronization — All known LSRs are put into an LSU frame and shared with every other switch that is directly attached.

- Periodic synchronization — Only updated LSRs are put into an LSU frame and shared with every other switch that is directly attached.

In both cases, when a switch receives an LSU from another switch, it will compare the payloads and update its internal Link State Record Database with any LSRs that it either does not have, or are newer than ,the one it currently has for that Domain. The format of the LSU is shown on the *left* in Figure 253; the data from the payload of the first LSU to be transmitted from switch E is shown to the *right*.

SOFf					BC	B5	58	58
FT-1 Header					02	FF	FF	FD
					00	FF	FF	FD
					22	38	00	00
					4C	00	00	00
					00	E8	FF	FF
					00	00	00	00
Command				FSPF HEADER	15	00	00	00
Version	AR #	Auth Type	Reserved		02	00	00	00
Originating Domain ID					00	00	00	05
Authentication					00	00	00	00
					00	00	00	00
Reserved			Flags		00	00	00	01
Number of Link State Records					00	00	00	01
LSR Type	Reserved	LSR AGE		Link State Record Header # 1 (LSR Type 01)	01	00	00	01
Reserved					00	00	00	00
Link State Identifier					00	00	00	05
Advertising Domain ID					00	00	00	05
Link State Incarnation number					80	00	00	01
Checksum		LSR Length			C4	8D	00	3C
Reserved		Number of Link Desc			00	00	00	02
Link ID				Link Descriptor # 1	00	00	00	04
Reserved	Output Port Index				00	00	00	01
Reserved	Neighbor Port Index				00	00	00	01
Link Type	Reserved	Link Cost			01	00	01	F4
Link ID				Link Descriptor # 2	00	00	00	04
Reserved	Output Port Index				00	00	00	02
Reserved	Neighbor Port Index				00	00	00	02
Link Type	Reserved	Link Cost			01	00	02	F4
CRC					06	50	D2	E1
EOFn					BC	B5	D5	D5

Figure 253 Link State Update frame format and payload

The LSR consists of the Link State Record Header as well as any Link Descriptors associated with the header. In Figure 253, the LSU consists of a single LSR which in turn contains two Link Descriptors. After all of the LSRs have been exchanged, there will eventually be *one* LSR for *each* switch in the fabric (i.e., 5 switches and 5 LSRs). Each of the Link Descriptors under the LSR represent one of the ISLs on the

switch that sourced the LSR. For example, a switch with 64 E_Ports would have 64 Link Descriptors in its LSR.

The fields that we need to focus on in order to determine the shortest paths to the other domains are:

Number of link state records	As mentioned previously, there is one LSR for each Domain in the fabric and each Domain is responsible for generating and updating its own LSR. In the example above, since this is the first LSU being sent out by switch E, there is only one LSR.
Advertising Domain ID	The Domain that is responsible for creating this LSR.
LSR Age	This field in used by each switch to determine if it has the most recent copy of every LSR.
Number of Link Descriptors	The number of active E_Ports on the Domain that created this LSR.
Link ID	The Domain ID of the switch on the other end of the link.
Output Port Index	The port number on the switch that is connected to the other switch.
Neighbor Port Index	The port number on the other switch that is connected to the port on this switch.
Link Cost	The cost of the ISL as determined by: S * (1.0625e12 / Signaling Rate) or 0x03E8 = 1000 = 1 Gb/s = 100 MB/s 0x01F4 = 500 = 2 Gb/s = 200 MB/s 0x00FA = 250 = 4 Gb/s = 400 MB/s 0x0064 = 100 = 10 Gb/s = 1200 MB/s

Once the LSU has been sent out, it will be acknowledged with ACK but not accepted with a normal Accept. Instead, the switch receiving the LSU will send a Link State Acknowledgement (LSA) frame.

LSA – Link State Acknowledgement

When a switch receives an LSU, it will compare the contents with its copy of the Link State Database. If the LSU contains an LSR for a particular domain that is more recent than the LSR it has locally for

that domain, it will replace the existing copy with the new one and forward the new LSR in all subsequent LSUs. Whether or not a newer copy of an LSR is detected, the switch that received the LSU will respond by sending an LSA. Inside of the LSA will be an FSPF header and a copy of each of the Link State Record Headers being acknowledged. Once all of the LSRs have been distributed throughout the fabric, each switch will have a local copy of all LSRs and will be able to determine the shortest path to any destination domain in the fabric.

The algorithm that a switch uses to determines the shortest path is not specified by the FC-SW standard, but it does provide an example of Dijkstra's algorithm, discussed next.

Dijkstra's algorithm for determining the shortest path

Note: The method in which Dijkstra's algorithm presented below is based on an article by Renaud Waldura and can be found at http://renaud.waldura.com/doc/java/dijkstra/

Another good source of information is wikipedia, at http://en.wikipedia.org/wiki/Dijkstra's_algorithm

Dijkstra's algorithm allows the user to find the shortest path between two points. In a Fibre Channel fabric, this means the shortest distance between two domains. The result of the LSU and LSA (discussed earlier in this section) will have resulted in a LSR database. For the purposes of our example, a sample LSR database along with the fabric topology in this example is shown in Figure 254 on page 844.

Source Domain ID	Out Port	Destination Domain ID	In port	Cost
1	1	2	1	100
1	3	3	3	100
2	1	1	1	100
2	2	3	2	100
3	2	2	2	100
3	3	1	3	100
3	4	4	4	250
3	5	4	5	250
3	6	4	6	250
4	1	5	1	500
4	2	5	2	500
4	3	3	3	250
4	4	3	4	250
4	5	3	5	250
5	1	4	1	500
5	2	4	2	500

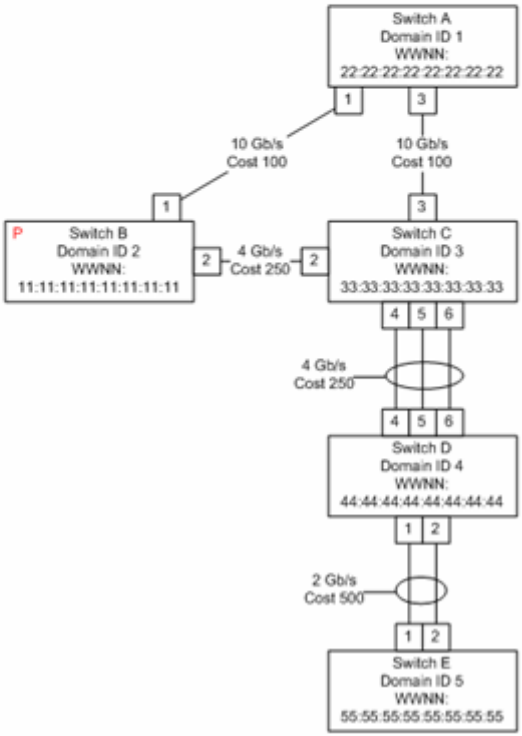


Figure 254 Sample LSR database and fabric topology

To make this interesting, we will watch the process of Switch B applying Dijkstra's algorithm to the LSR database shown in Figure 254 on page 844. To better understand the process, we will borrowethe idea of implementing the algorithm in pseudocode from Renaud Waldura.

The data structures used in the example are as follows:

cost	Stores the best estimate of the lowest cost (shortest path) from the source Domain to each destination domain.
predecessor	Stores the predecessor (previous domain) of each domain on the shortest path from the source.
source	The source domain. In this example, this would be switch B or Domain 2.
S	The set of settled domains, the domains whose shortest distances from the source have been found.
Q	The set of unsettled domains, the domains whose shortest distance from the source have NOT been found.

Start Pseudocode:

```
//initialize cost to infinity
cost = ( 8 )
//initialize predecessor to empty
predecessor = ( )
//initialize S and D to empty
S = Q = ( )

//add source to the set of unsettled domains
add source to Q
//set cost to the source equal to zero since there are no ISLs from the source
to itself
cost(source) = 0

//Loop for as long as there are unsettled domains
while Q is not empty
{
//Get the shortest paths from Q and assign to u
     u = getLowestCostDomain(Q)
//Add the shortest paths to S
add u to S
calculateDistanceToNeighbors(u)
}
```

The functions are listed below:

```
calculateDistanceToNeighbors(u)
{
   //check out each adjacent domain to u as long as the domain isn't settled.
   for each domain d adjacent to u, d not in S
   {
      //if true, a lower cost path (shorter distance) exists
      if cost(d) > ( cost(u) + cost(u-v) )
      {
         cost(d) = ( cost(u) + cost(u-v) )
      }
         predecessor(d) = u
         add d to Q
   }
}

getLowestCostDomain(Q)
{
   find the lowest (in terms of cost) domain in Q
   remove the lowest cost domain from Q
   return the lowest cost domain
}
```

End Pseudocode:

Example: You should follow along with the pseudocode as you read through this example.

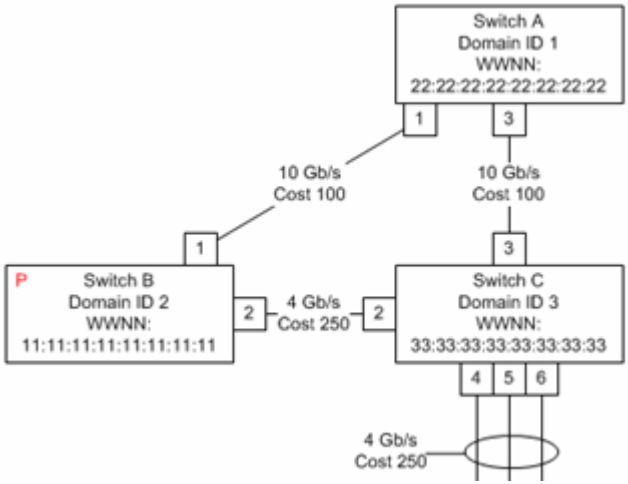


Figure 255 Switch B determining the shortest path

We begin by adding our source switch **B** to the set **Q** (refer to Figure 256). Since Q is not empty, we extract the minimum cost domain B (the only entry in Q, which also has zero cost). We add **B** to **S**, then calculate the distance to its neighbors.

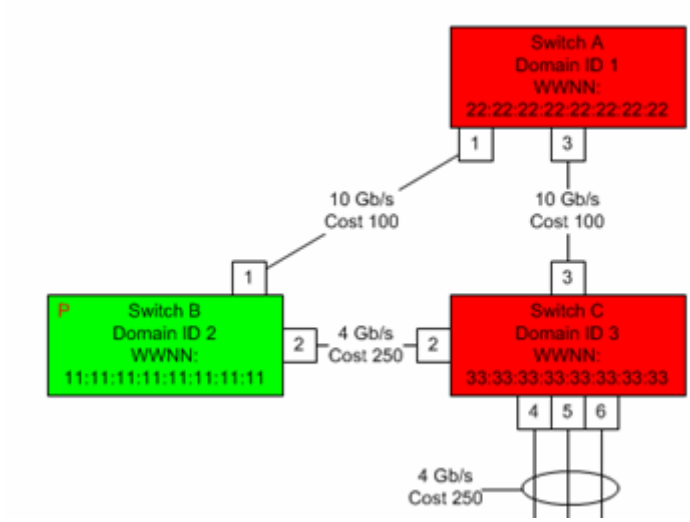


Figure 256 **Switch B is settled**

As can be seen in the LSR database in Figure 257, those neighbors or domains adjacent to B are **A** and **C** (shown in in red in Figure 256 and Figure 257).

Source Domain ID	Out Port	Destination Domain ID	In port	Cost
1	1	2	1	100
1	3	3	3	100
2	1	1	1	100
2	2	3	2	100
3	2	2	2	100
3	3	1	3	100
3	4	4	4	250
3	5	4	5	250
3	6	4	6	250
4	1	5	1	500
4	2	5	2	500
4	3	3	3	250
4	4	3	4	250
4	5	3	5	250
5	1	4	1	500
5	2	4	2	500

Figure 257 LSR database

We first compute the best distance estimate from **B** to **A**. Since distance (A) was initialized to infinity, we calculate the distance B to A as follows:

cost(A) = cost(B) + cost(B-A)

cost(A) = 0 + 100

cost(A) = 100

Next, predecessor(A) is set to B and we add A to Q (the set of unsettled domains). Following the same process for C, we assign cost(C) to 250 and predecessor(C) to B.

The second time around, Q contains A and C. As seen in Figure 258 , **A** is the domain with the current lowest cost of 100. It is extracted from the queue and added to **S**, the set of settled domains.

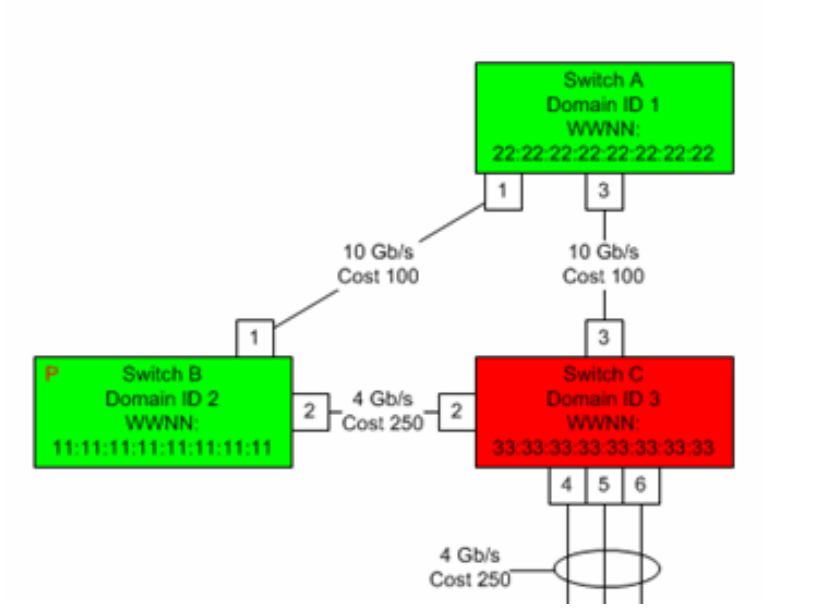


Figure 258 Switch A is settled

Next, we calculate the distance to A's neighbors, **B** and **C**, using the information in the LSR database shown in Figure 259 on page 850.

.

Source Domain ID	Out Port	Destination Domain ID	In port	Cost
1	1	2	1	100
1	3	3	3	100
2	1	1	1	100
2	2	3	2	100
3	2	2	2	100
3	3	1	3	100
3	4	4	4	250
3	5	4	5	250
3	6	4	6	250
4	1	5	1	500
4	2	5	2	500
4	3	3	3	250
4	4	3	4	250
4	5	3	5	250
5	1	4	1	500
5	2	4	2	500

Figure 259 LSR database

Domain B (Domain ID 2) is ignored because it is found in the settled set. The first pass of the algorithm had determined that the lowest cost path from B to C was direct. However, looking at A's neighbor, **C**, we realize that the following statement evaluates to true and therefore a lower cost path to C must exist:

cost(C) > cost(A) + cost(A-C)

250 > 100 + 100

250 > 200

In fact, we have found that a shorter path going through A exists between B and C. The value of distance(C) is updated to 200 and predecessor(C) is set to **A**. **C** is added again to **Q**.

The unsettled domain with the lowest cost is now **C** and it is extracted from the queue and added to **S** (refer to Figure 260 on page 851).

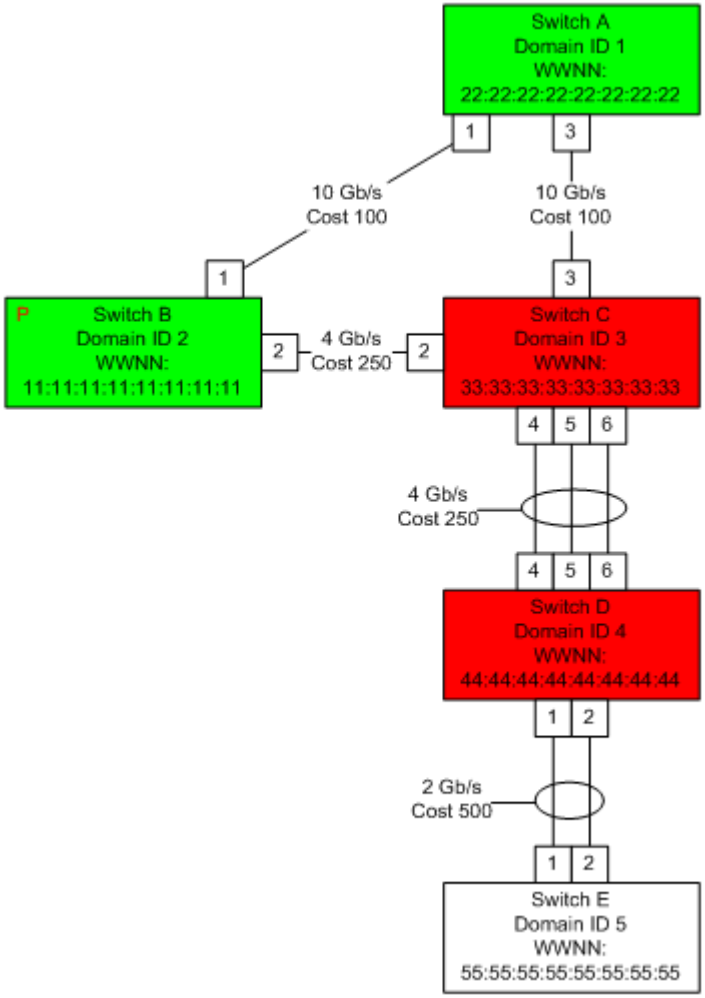


Figure 260 **Shortest path**

The adjacent switches are determined by referring to the LSR database as shown in Figure 261:

Source Domain ID	Out Port	Destination Domain ID	In port	Cost
1	1	2	1	100
1	3	3	3	100
2	1	1	1	100
2	2	3	2	100
3	2	2	2	100
3	3	1	3	100
3	4	4	4	250
3	5	4	5	250
3	6	4	6	250
4	1	5	1	500
4	2	5	2	500
4	3	3	3	250
4	4	3	4	250
4	5	3	5	250
5	1	4	1	500
5	2	4	2	500

Figure 261 LSR database

Domains B (Domain ID 2) and A (Domain ID 1) are ignored because they are found in the settled set. Domain D, however, is not in the settled set so the cost from B to D is evaluated as follows:

cost(D) > cost(C) + cost(C-D)

250 > 1200 + 250

250 !> 450

Since it evaluates to false, a shorter path does not exist to D. Next, predecessor(D) is set to **C** and **D** is added to **Q**. Finally, the unsettled domain with the lowest cost is **E** and since all of its neighbors are settled, the algorithm ends. Routes can now be determined and the routing tables within the switch can now be programmed.

To view portions of the LSR database on each switch, use the following commands:

- For Brocade FOS — **uRouteShow**
- For Brocade M-EOS — **show fabric topology**
- For Cisco SAN-OS — **show fspf database**

Note that there are the multiple equal cost paths between domains C and D, as well as between D and E. In Fibre Channel, all of these equal cost paths will be used to transfer data. There are multiple ways that this is accomplished , further discussed in "Trunking and ISL aggregation" on page 869. However, the Build Fabric process is not quite complete at this point since the name servers have not yet been synchronized.

The distributed Name Server, SW_RSCN, and GE_PT

In Fibre Channel Fabrics, the Name Server is distributed. This means that every switch in the fabric is responsible for maintaining its own copy of the global name server database. Each switch database is kept in sync in three ways:

- Initial synchronization
- Event driven updates
- Periodic updates:

Initial synchronization takes place when a switch is coming up or joining a fabric for the first time. When this happens, the switch will generate a GE_PT (Get Elements based on Port Type) to every switch in the fabric. It does this by specifying the destination switch in the Destination ID of the GE_PT request frame. For example, if Switch A (Domain ID 1) was sending a GE_PT to Switch B (Domain ID 2), the DID / SID would be FFFC02 / FFFC01. In this way each switch can query every other switch in the fabric for its Name Server contents. It is important to realize that when a switch sends a GE_PT to another switch, it has to specify the port type that it is looking for. Typically, a port type of Nx_Port is used, but there have been discovery problems related to the incorrect port type being specified by the requesting switch.

Event driven updates occur when there is a change in the login status of a port. When this happens, the local switch will send a SW_RSCN to each switch in the fabric. It does this by specifying the destination switch in the Destination ID of the switch RSCN (like the GE_PT) and also including in the payload the port ID that changed as well as its current status (offline, online). When the switch receives the RSCN, it may initiate a GE_PT to the switch that sent the SW_RSCN. There have been heterogeneous interoperability issues related to incorrect handling of SW_RSCNs. For an example of this occurring, see the "Adding Nx_Ports and setting up routes" on page 854.

Periodic updates using GE_PT are performed by each switch to ensure that they have the latest copy and did not miss a switch RSCN during the process.

One of the reasons for distributing the name server is so that each switch can service most name server queries locally without needing to initiate a request to another switch in the fabric to satisfy the request. However, on some switch platforms, not all name server database information is shared between all switches in the fabric. An example of this is Symbolic Port Name and Symbolic Node Name information. In some cases when an Nx_Port queries the name server for the Symbolic Port Name or Symbolic Node Name of another port in a different domain, the switch needs to generate a request to the domain in the fabric where that port is logged in. An example of a command that could be used when specific information is needed about a port in another domain is **GA_NXT** (Get All Next).

Adding Nx_Ports and setting up routes

Up until this point, the Build Fabric process has focused on what happens to the switches in the fabric as the Fabric Configuration process is taking place. Now that the Fabric Configuration process has completed, hosts and storage ports can be added and routes can be setup between them.

In Figure 262 on page 855, a single target and initiator are shown. For the sake of this example, we will assume that the target has fully logged in with Switch A and that the initiator has just completed logging in and registered with the name server. Notice that the name server database on both switches contains the entry for target 1 - 10:00:00:00:00:05:00:00 but only the name server on switch B contains both entries. This is because switch B has not transmitted the SW_RSCN yet and subsequently, switch A has not performed the GE_PT.

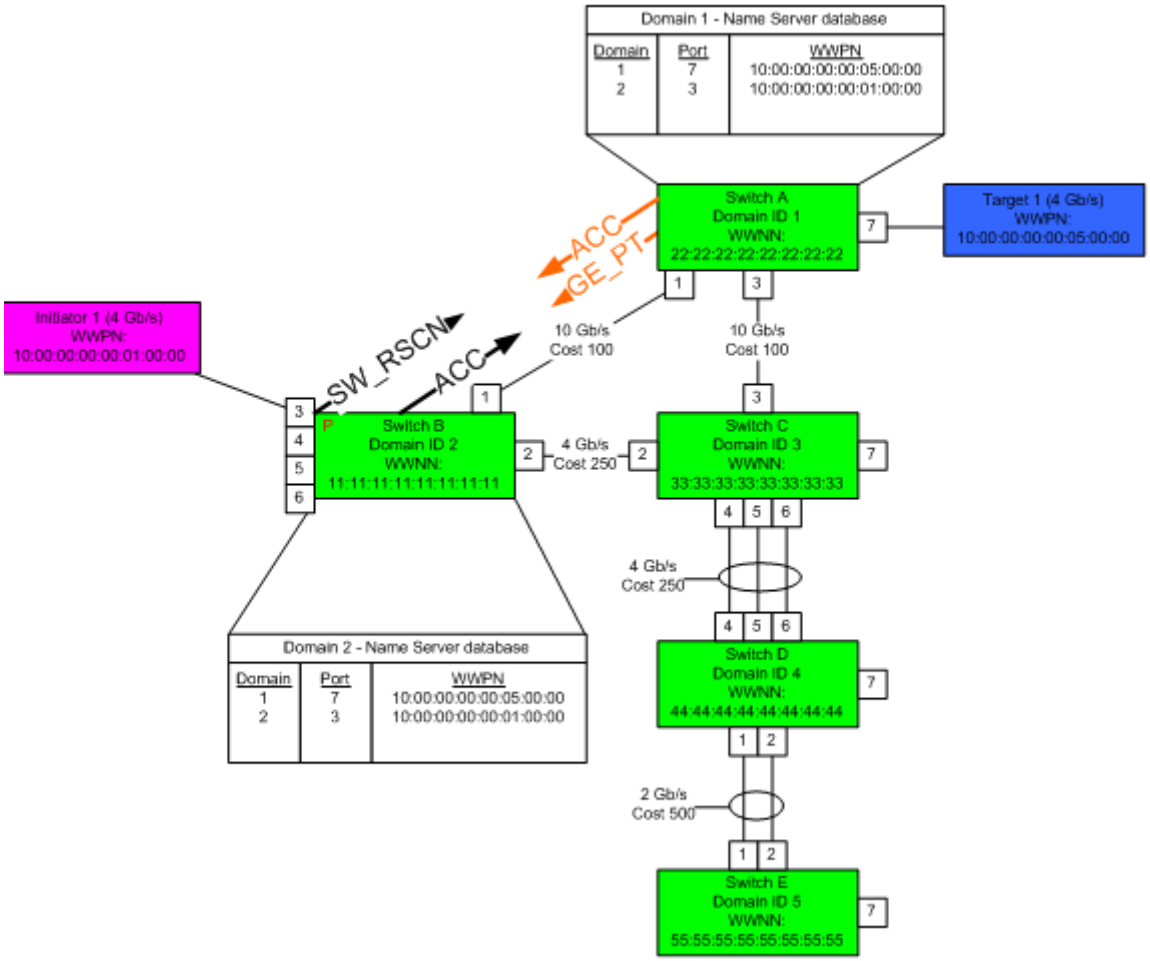


Figure 262 **Adding hosts and storage ports and setting up routes**

Once the SW_RSCN has been accepted and the GE_PT responded to, both of the name server databases will be in sync. Next, routes can be setup between the two switches.

Assigning egress ports

In "Dijkstra's algorithm for determining the shortest path" on page 843, the LSR database, was used to determine the shortest paths. At the end of this process, we were able to determine that the shortest path off of switch B to any destination in the fabric was via port 1. This would make routing decisions simple since any frame coming

into the switch destined for another Domain ID would simply be routed to port 1 and sent off to switch A. If this frame was then received by switch A and was destined to a port on switch A, then it would simply route it to the correct local port. However, if the frame were destined to a domain other than A, and you were to ask what the shortest path off of switch A is, you would have to answer that it depends upon the destination. If the destination were switch B, then the shortest path is via port 1. If the destination were switch C, D or E, then the shortest path would be via port 3.

Determining the shortest path for a frame to take is a decision that each switch must make millions of times per second. One way to make this process as efficient as possible is via the use of Content Addressable Memory (CAM). Regardless of how it is going to be implemented in hardware, some kind of routing table needs to be created that takes into consideration all possible ingress ports (ports where frames are coming into the switch) and egress ports (ports where frames are leaving the switch).

Frequently, switches will have at least a few different routing tables. For example, one routing table would handle frames being received with a destination Domain ID that is different than the local domain ID. The other table may handle frames being received with a Domain ID the same as the local Domain ID. A simple example of a routing table that could be used on switch A in the above example fabric is shown in Figure 263.

Ingress Port	Destination Domain	Egress Port
1	3	3
	4	3
	5	3
3	1	1
7	1	1
	3	3
	4	3
	5	3

Figure 263 Routing table example

In the simplest case where there is no dynamic load sharing or Exchange-based routing on the ISLs, if there were multiple egress port to the same domain, some number of ingress ports would be

assigned by the switch to use each egress port. This presents a problem since this approach does not take into account the utilization of each ingress port before making the assignment. Because the utilization of each ingress port is not taken into account, there is a very high probability that they are distributed in a manner that does not allow for all of the bandwidth to be used on the egress ports. It should be noted that there are several features in place to help alleviate these problems, all of which will be discussed in the next few sections.

Preferred Paths / Static assignments

As discussed in "Assigning egress ports" on page 855, frequently a switch does not assign ingress ports so that they take full advantage of egress port bandwidth. This is not usually a problem that needs to concern customers, but in some circumstances the unpredictability of the latency through a fabric caused by congestion can cause problems to applications which require large amounts of bandwidth. Because of this, a customer may want to dedicate specific ISLs to service a particular application. This would allow the customer to guarantee isolation and the dedication of bandwidth for the applications that need it.

It could also be that a customer wants to dedicate specific ISLs for EMC replication applications, such as SRDF or Mirrorview, or simply differentiate a specific business unit's traffic, such as 'human resources' or 'financing' from the rest of the environment. One way to accomplish these goals is to statically assign each ingress port to use a specific egress port and manually configure all or a subset of the routes used by each storage and target.

While it is possible to do this, it is *not* a recommended best practice for several reasons. First, it is a very time consuming process and very prone to human error. Another concern is failover. If an ISL were to fail, the traffic that was previously isolated on that ISL will be moved to any available ISL. Finally, and most importantly, in the vast majority of cases this is a band-aid approach. In a properly architected SAN, the problems that are trying to be overcome by using this approach will not be an issue in the first place. Rather than try to configure preferred paths to overcome a problem, it would be better to make an architecture change instead.

With the above in mind, if you really need to setup static routing, the following information will describe the layout and help you to understand the ramifications and costs of such a solution.

Connectrix B example

General information

On Connectrix B switches, the user has the option to use *port- based routing*. With port-based routing, FSPF is used to determine the shortest path and the switch will define egress ports cost of the link, S_ID and D_ID.

Once the link is established between the S_ID and D_ID, it remains the same throughout the life of the login session. This can introduce the possibility of all high I/O devices being allocated to one ISL since all egress port assignments are not redistributed on every login. Port-based routing does a "best-effort" job of distributing I/O by balancing source port routes but since the switch has no way to predict which routes will be "jammed" when setting up routes, there is a chance of getting multiple high bandwidth ports stacked up on the same ISL. Again, until a switch goes offline, or the fabric changes; all paths will remain the same.

Configuring port-based routing and FSPF

Brocade uses the standard link cost calculations for FSPF with the exception of 2Gb and 4Gb port settings in FOS 4.4x. In general the link cost for each setting is:

1 Gb = 1000

2 Gb/4 Gb = 500

10 Gb = 100.

4 Gb is typically rated at 250 and is configurable through Brocade FOS, but to keep all things equal, by default 2 Gb/4 Gb were set to 500 to better guarantee the balance of I/O across available ISLs.

It is possible to reconfigure the link speed if one can determine that the cost change will positively impact the flow of traffic in the SAN. The practice of configuring a static route via the port-based policy via the command line interface is shown below:

```
linkcost: [SlotNumber/]PortNumber Cost
```

From Web Tools you can set the link cost and the route, as well as whether or not frames are set to in-order delivery. See Figure 264.

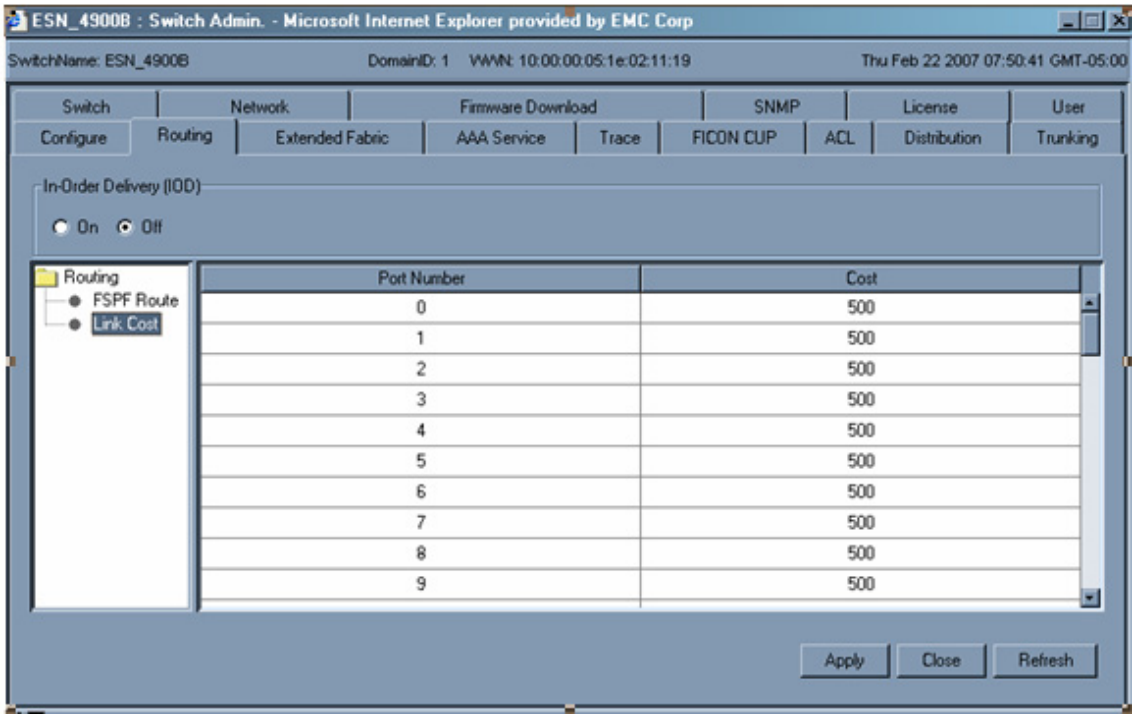


Figure 264 **Routing tab**

To configure static routes in the switch through CLI:

IMPORTANT

This is a disruptive activity and must be pre-planned.

1. Disable the switch with **switchDisable** command:

   l82c7219:admin> **switchDisable**

2. Turn on port-based routing using the **aptpolicy** command:

   l82c7219:admin> **aptpolicy**

   Current Policy: 3 [default policy is always exchange-based routing]

   3: Default Policy

1: Port Based Routing Policy
3: Exchange Based Routing Policy

l82c7219:admin> **aptpolicy 1**
Policy updated successfully.

3. Configure your static route using the command:

   <urouteconfig> [in port], [destination domain], [out port]

   Example:
   l82c7102:admin> urouteconfig 5 219 0
   l82c7102:admin> urouteconfig 6 219 0

   This has mapped the ingress ports 5 and 6 to destination domain ID 219 via Port 0.

4. To display your routes use the **<urouteshow>** command.

Brocade Fabric OS 5.x and higher user manual(s) are good resources for the various settings and configurations that may be used beyond the scope of this paper.

Connectrix MDS example

General information

Preferred Path routing in Connectrix MDS was introduced in v3.0(3). Static Routing can be configured through Device Manager. Route control has been integrated in SAN-OS since v1.x.

Configure Fibre Channel routes

Configure a static route

The first thing that needs to be understood about creating a Fibre Channel route through Device Manager is that FSPF is *disabled*. Static routing in MDS Series fabrics is accomplished through an algorithm that forwards the frame based on the FC ID. By using FC ID for a specific interface and domain, customers can configure a specific route.

To configure a specific route:

1. Log in to Device Manager through **Start> Programs> MDS9000> Device Manager**.
2. From the **Main** menu navigate to **FC> Advanced> Routes**.
3. Click **Create** to add a static route
4. Select the VSAN ID that is going to contain the route.

5. Add the destination address and destination mask for the device being used to configure a route.
6. Select the interface (port) that is going to be used as the egress port to the destination.
7. Select either local or remote options. (Local means the next hop is the final destination; remote means that the next hop is not the final destination.)
8. Click **Create** to save the changes.
9. Click **Close** to abort the operation.

To view all of the Fibre Channel Routes, use Fabric Manager.

Preferred Path routing

The ability to specify routes for various applications can be specified in the Cisco MDS-Series switches in the form of preferred path routing. It provides a method to route traffic without the consideration of FSPF. Paths are chosen based on frames received on a selected interface or based on source FC ID specified. This feature is *not* recommended for best practices, yet it may make sense where logical for isolating applications for business purposes or strategic reasoning.

Preferred paths are defined by route maps and configured on a VSAN basis. The criteria by which routes can be configured are based on S_ID and D_ID and specific to egress traffic only, per switch. This means that a "static" mapped preferred path per switch needs to be configured for traffic to travel from one direction to another. Priorities can be specified per path as well, from 1 to 5 (1 being the highest priority).

IMPORTANT

Setting preferred path is a disruptive action and it is recommended to set the priorities when the switch is offline.

Connectrix M example

General information

The layout for this section is subjective to the display screens throughout the Use Case. For this section, the general topology is two DS-32M2s and a Symmetrix RDF pair.

Configuring static routing through Preferred Path and Prohibit Dynamic Connectivity Mask (PDCM)

Preferred Path is a standard offering with E/OS 6.01.00. There is no license key requirement.

Preferred Path provides granular route control for data transfer across a multi-switch fabric. The user may create a route between two switches, or over a fabric with multiple ISL hops. Each node can have only **one** Preferred Path between two switches. That is, a node is permitted only one "route controlled exit port." For example, Figure 265 on page 863 displays the route of an HBA in a one hop fabric. The first row under the **Out Port #** column identifies port 20 as the ISL used by this node to get to the next domain.

If there is more than one ISL between switches, the "Out Port" may change if the route tables are recalculated because of a fabric event, such as ISL removal or domain addition. Setting a Preferred Path will maintain a static path even if the switch reboots.

The Preferred Path is configured on a per switch basis. That is, in a multi-hop fabric, the user must configure the "Out Port" (or exit por) for every switch between the initiator/target pair.

If a node with a Preferred Path fails for whatever reason (e.g., optic failure, ISL disconnect, switch failure), the routes will be recalculated and the node will be given a new path. When the Preferred Path is reestablished, the node will resume usage of that path.

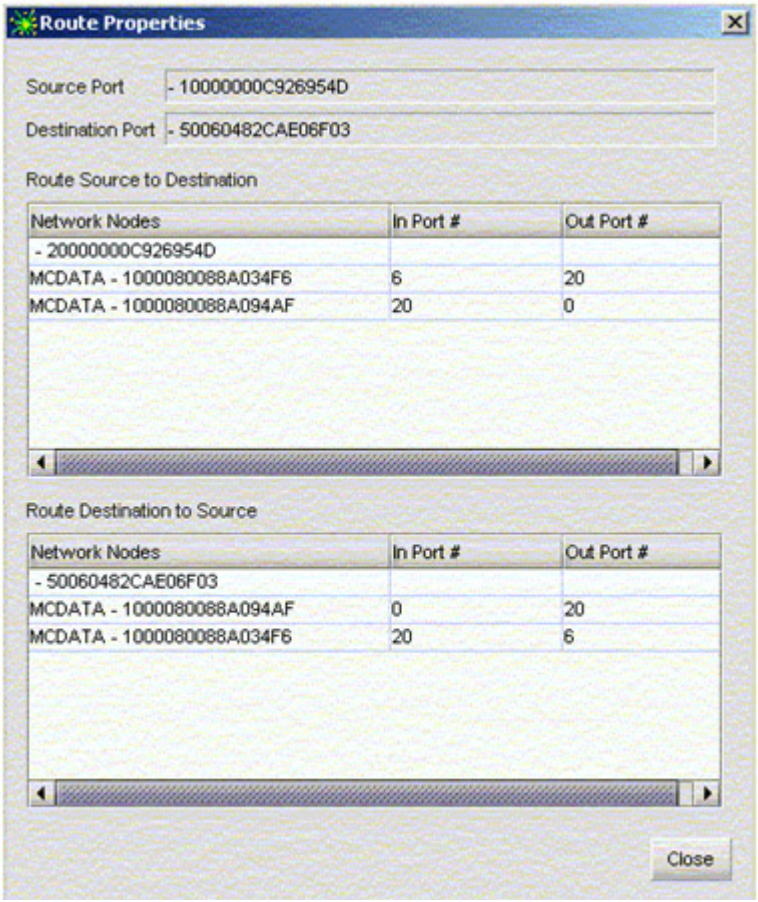


Figure 265 **Route properties**

IMPORTANT

When planning a Preferred Path, it is important to look at the path from the host and storage perspective. That is, if you want an initiator to have a Preferred Path, you should configure the target to use the same path. This makes for easier management and troubleshooting. It must be configured on both the initiator's switch and the target's switch.

Setting a Preferred Path for one node does not automatically prevent other nodes in the fabric from using the same path. To isolate a node on a Preferred Path, it is necessary to use PDCM (Prohibit Dynamic Connectivity Mask) for E_Ports as well.

Unlike Preferred Path, you may permit or deny a node (initiator or target) from accessing multiple ISLs with PDCM. This provides greater flexibility for the user to restrict the ISLs used by a node. Preferred Path can assign a node to a specific ISL while PDCM will prevent other nodes from accessing that ISL. This has the advantage of dedicating bandwidth to particular nodes.

Configuring Preferred Path and PDCM

Using Connectrix Manager 9.01 or McDATA switch CLI, you can configure a Preferred Path. For CLI commands, please consult the appropriate switch user's guide.

To configure a Preferring Path:

1. Open **Connectrix Manager 9.1** and go to **Start > Program Files > Connectrix Manager 9.1 > Connectrix Manager 9.1**.
2. Log in with appropriate username and password.
3. Navigate to the desired switch that you want to support Preferred Path and open up **Element Manager** either by double-clicking or by right -clicking **Element Manager**.
4. Using the **Main** menu, choose **Operating Parameters > Switch Parameters > Check Domain ID to "Insistent"**.

   This assures that the Domain ID remains *consistent* through reboots or fabric events and is a *nondisruptive* action.

   Note: ED-10000M configures the Insistent Domain through the **Configure > Domain** tab.

5. Activate the change and close the GUI screen.
6. Add a Preferred Path.
7. In the Element Manager for the switch, go to **Configure > Preferred Path**.
8. Check the **Enable Preferred Path** box.
9. Click **Add**.

Preferred Path will be configured by port number and not WWN, so the physical port will always be the same unless changed manually in Preferred Path.

10. You have three options:

    - Source Port
    - Exit Port
    - Destination Domain ID

    (Exit port is the E_Port for the ISL to the other domain. D_ID is the next hop. Assure that the correct exit port and Domain ID numbers are used for Preferred Path. See Figure 266.)

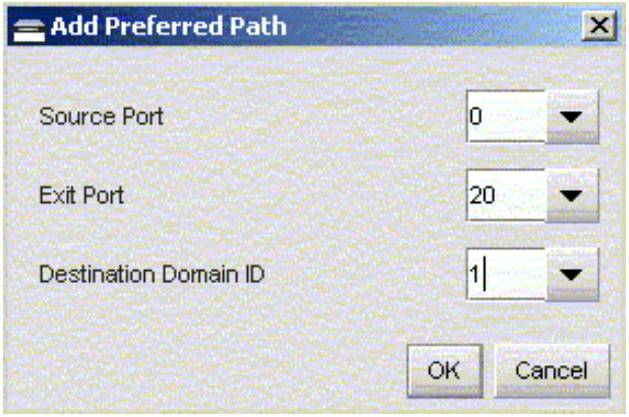


Figure 266 **Add Preferred Path**

In this example, shown in Figure 267 on page 866, a new route will be created for the Symmetrix SRDF pair between two DS-32M2s. The RDF pair is 6F13 on DS-32M2_bottom and 2ED3 on DS-32M2_top. The columns in the **Product List** window in Connectrix Manager can be manipulated to display the relevant port and domain information. Using the **Show Route** function, the end nodes will be displayed in the **Topology** window. 2ED3 is attached to port 13 on domain 23. 6F13 is attached to port 13 on domain 1.

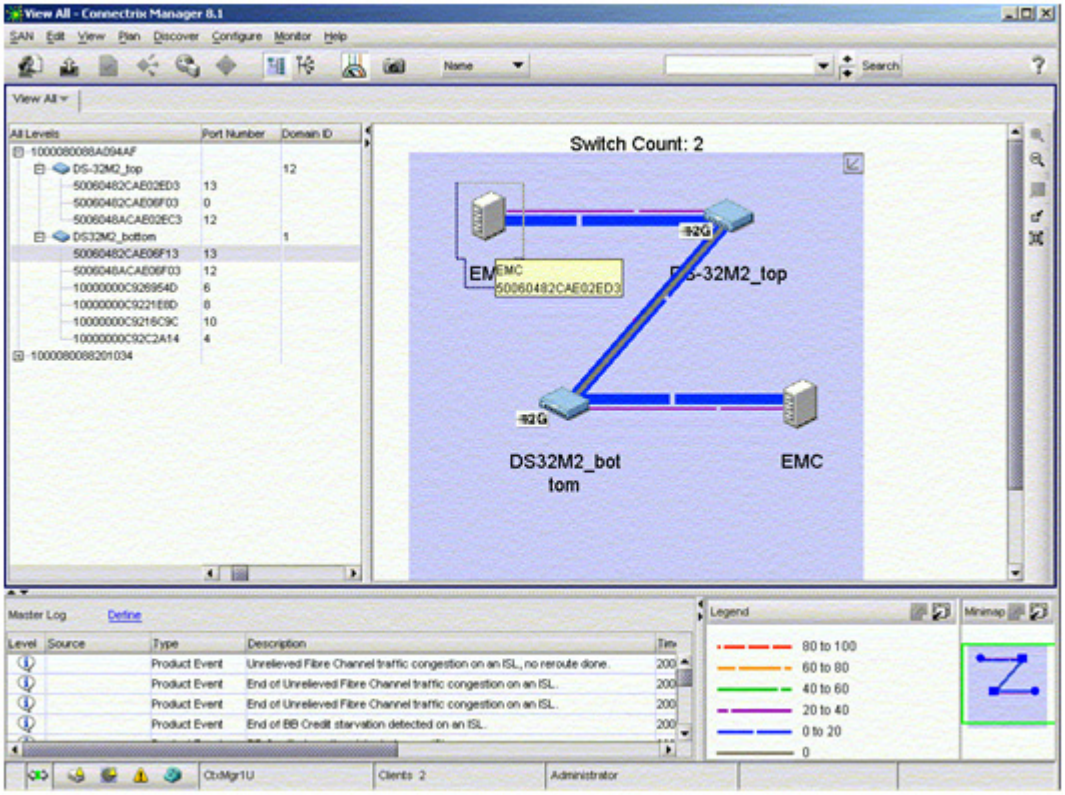


Figure 267 Topology window

The current E_Port used by RDF port 2ED3 can be identified by right-clicking on the link between the node and switch icons. Figure 268 on page 867 shows the route properties before the Preferred Path is configured. It shows that the current path for 2ED3 and 6F13 are using the ISL on port 20 on their respective switches.

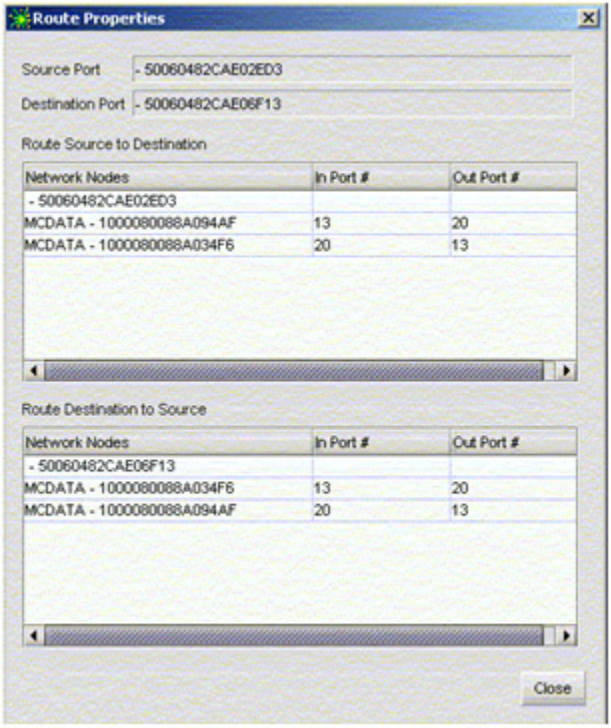
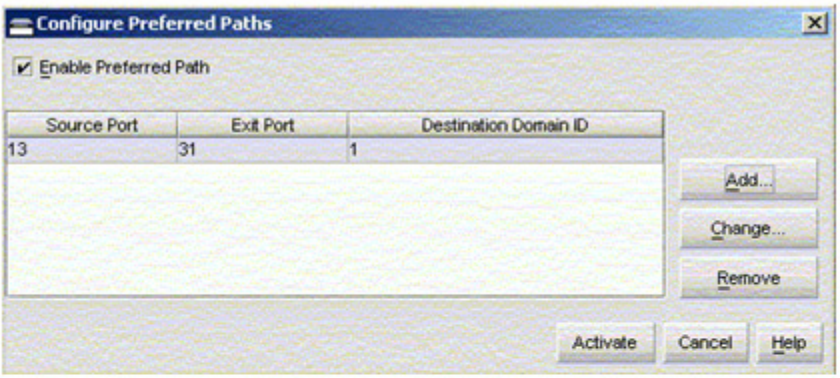


Figure 268 **Route properties dialog box**

This example will change that to the ISL on port 31. Figure 269 on page 868 shows the newly configured Preferred Path windows from DS-32M2_top (top screenshot) and from DS32M2_bottom (bottom screenshot). Because of bi-directionality, Preferred Path is configured on both switches for the SRDF pair..



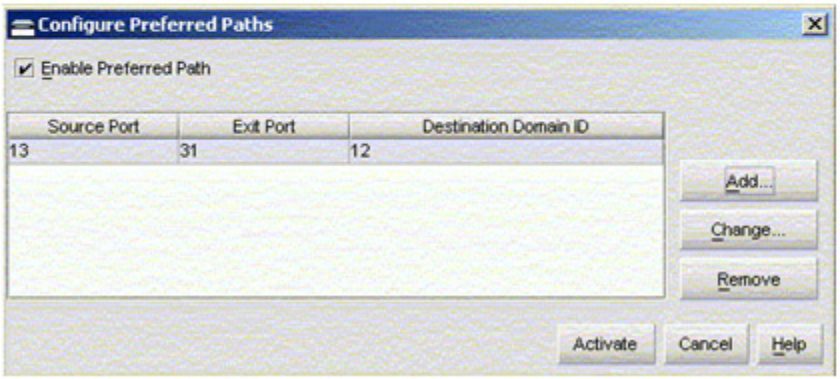


Figure 269 Newly configured Preferred Path windows

Trunking

Trunking is a generic term using to describe functionality that allows for more efficient use of ISL bandwidth. "Trunking and ISL aggregation" is discussed on page 869 in this section.

There are three main types of trunking: Flow-based, Frame-based, and Exchange-based, briefly discussed next. These are further developed, including the pros and cons, along with information on other types of trunking, under "Other types of trunking" on page 873.

Main types of trunking

Flow-based trunking

This type of trunking was developed by Brocade M-Series and is currently marketed as *Open Trunking*. It works by making periodic changes to the hardware routing tables based on the amount of time

being spent at zero BB transmit Credit and a percent utilization measurement. Since the changes are being made to the routing table of the switch transmitting the frames, it can be deployed in any environment and in any interop mode.

Frame-based trunking

Frame-based trunking is available on Brocade B-series (FOS) switches. It involves the aggregation of several physical ISLs between any two adjacent switches into one logical unit so that frames may be evenly distributed across all of the ISLs in a trunk. Once the frames are received on the other side they are reassembled in the proper order and sent to the destination.

Note: For more detailed information on the pros and cons of frame-based trunking, refer to "Frame-based trunking" on page 873.

Exchange-based trunking

Exchange based trunking is available on Brocade, Cisco, and QLogic switches. It works by distributing exchanges evenly over a group of ISLs. One of the limitations of Fibre Channel is that for the most part, Nx_Ports do not handle getting Out Of Order Frames (OOOFs) within a sequence or exchange very well. Typically, the receiving port will abort the exchange and force the entire exchange to be re-driven. This is unfortunate for trunking implementations, because a trunking solution that could evenly transmit the same number of bytes over each ISL would be very efficient.

The next best thing would be one that could transmit the same number of frames over each ISL. The downside to doing this is that special hardware is needed on the receive side to ensure that the frames are reassembled into the proper order (to prevent OOOFs). By implementing trunking at the exchange level, exchange-based trunking prevents frames within an exchange from getting out of order by routing all of the frames for the exchange down the same ISL.

Note: For more detailed information on the pros and cons of Exchange-based trunking, refer to "Exchange-based trunking" on page 875.

Trunking and ISL aggregation

As previously discussed, the FSPF protocol is used to find the shortest path between two domains. Once this is done, a routing table can be created so that a frame received by an FC switch can be forwarded on to the next hop on its way to its destination.

When Fibre Channel switches were first released, the process stopped here, and similar to Brocade's port based routing when Dynamic Load Sharing is disabled, as discussed in the "Connectrix B example" on page 858, once an assignment was made the route was never moved unless something major in the fabric changed. The main problem with this approach is that as new ports are added, they are assigned to an egress port in a round-robin fashion. This means that as ports are added and removed, the ISLs can become imbalanced. For example, in the configuration shown in Figure 270, there are 4 initiators and 4 targets. Each initiator on switch B is zoned to see one of the targets on switch A.

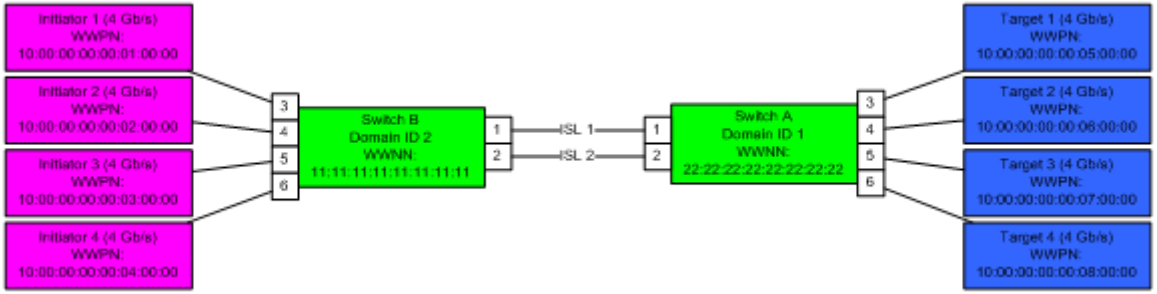


Figure 270 Configuration example

When the initiators and targets come up for the first time, they will be assigned to egress ports in a round-robin fashion, meaning that the first port will be assigned to ISL 1, the second port to log in will be assigned to ISL 2, the third port to log in will be assigned to ISL 1, and the final port will be assigned to ISL 2. See Figure 271 where the arrows indicate which ISL each port is assigned to.

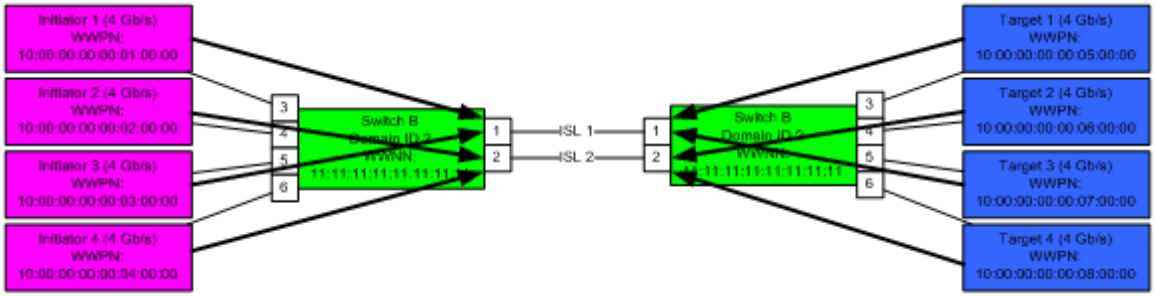


Figure 271 Port assignment example

In Figure 271, the configuration is completely balanced, but what would happen if one of the initiators were to be rebooted? The answer is that after the reboot, as the initiator was coming up, it would be assigned an exit port of ISL 1 since the last assignment was to ISL 2.

The configuration after initiator 2 is rebooted is displayed in Figure 272.

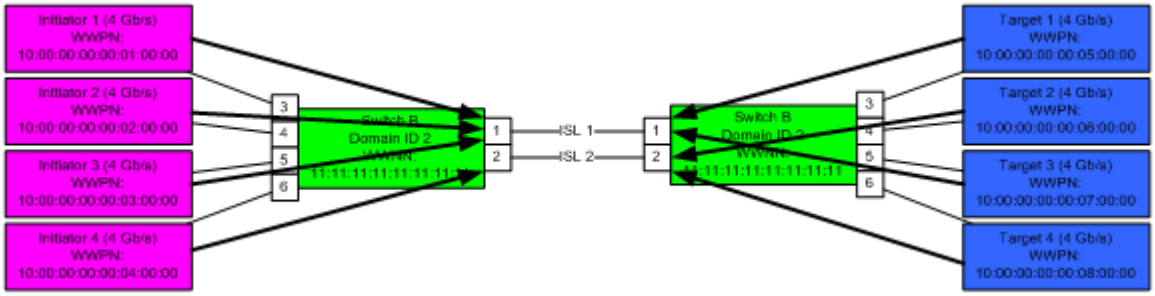


Figure 272 Configuration after initiator 2 is rebooted

One can easily see how this sort of approach could rapidly lead to imbalances in ISL utilization. As a result, this type of functionality did not last for very long, at least not unless a dynamic rebalancing feature is implemented, like Brocade's Dynamic Load Sharing (refer to "Dynamic Load-Sharing (DLS)" on page 871 for more information on DLS).

Round-Robin login with dynamic reassignment

The next major improvement in terms of assigning egress ports was to have the egress ports recalculated whenever a port logged out of a switch. This dynamic reassignment is done in such a way as to prevent Out of Order Frames (OOF's) from occurring. This approach also takes into consideration how many ports are already assigned to each ISL.

This feature is in use on Brocade M-EOS products today and is not configurable. On Brocade FOS products this feature is known as Dynamic Load Sharing (DLS).

Dynamic Load-Sharing (DLS)

As mentioned previously, egress port assignment is generally based on the incoming port and the destination domain. This means that all the traffic coming in from a port (either an E_Port or an Fx_Port)

directed to the same remote domain is routed through the same output E_Port.

If DLS is turned **off** (using **dlsReset**), load sharing is performed only at boot time or when an Fx_Port comes up. By disabling DLS, the possibility of dropped frames is eliminated every time a change in the fabric occurs. A change in the fabric is defined as an E_Port going up or down, or an Fx_Port going up or down.

If DLS is turned **on** (using **dlsSet**), when there are multiple equivalent paths to a remote switch traffic is shared among all the paths. Load sharing is recomputed when a switch is booted up or every time a change in the fabric occurs. With DLS enabled, traffic on existing ISL ports might be affected when one or more new ISLs are added between the same two switches. Specifically, adding the new ISL might result in dropped frames as routes are adjusted to take advantage of the bandwidth provided by the new ISL.

Enabling DLS optimizes fabric routing. For example, if an Fx_Port goes down, another Fx_Port might be rerouted from one E_Port to a different E_Port. The switch minimizes the number of routing changes, but some are necessary to achieve optimal load sharing. This is the factory default on all Brocade branded switches.

To viewing and changethis parameter.

1. Enter the **dlsshow** command to view the current DLS setting.

   One of the following messages appears:

   - *DLS is set.*

     This message means that the DLS option is turned on. Load sharing is reconfigured with every change in the fabric.

   - *DLS is not set.*

     This message means that the DLS option is turned off. Load sharing is only reconfigured when the switch is rebooted or an Fx_Port comes up.

2. Enter the **dlsset** command to *enable* Dynamic Load Sharing.
3. Enter the **dlsReset** command to *disable* Dynamic Load Sharing.

It is important to note that by default, this approach does not take into consideration how much bandwidth each ingress port is consuming. However, add-on functionality is available to enable it to look not only at utilization, but also the amount of time spent at zero BB_Credit.

Other types of trunking

Frame-based trunking

Soon after dynamic reassignment had been introduced into FC Fabrics, Brocade released their first version of trunking at the frame level. The approach requires that some of the ISLs between two domains be grouped onto the same quad of ports and then as frames come in assign them to an output port in round-robin fashion. The frames are then reassembled into the proper order on the other side of the ISLs. This approach ensures that the maximum ISL utilization is achieved by spreading the load out evenly across all ISL. Recently, frame trunking was enhanced between platforms that both support 4 Gb/s. Instead of distributing frames evenly over the ISLs in the trunk, it fills a link and then moves to the next link. While frames are not evenly distributed across all ISLs in a 4 Gb based frame trunk, the end result is the same — congestion of the frame trunk does *not* occur until *all* of the ISL's bandwidth is used..

Note: Due to HA considerations, it is better to have two smaller trunks between two domains than one large trunk. For example, if your environment requires 1600 MB/s of bandwidth between two domains, you should configure two 800 MB/s trunks instead of a larger 1600 MB/s trunk.

Another aspect of frame based trunking is that from an FSPF point of view; there only appears to be a single ISL between the two switches per frame trunk.

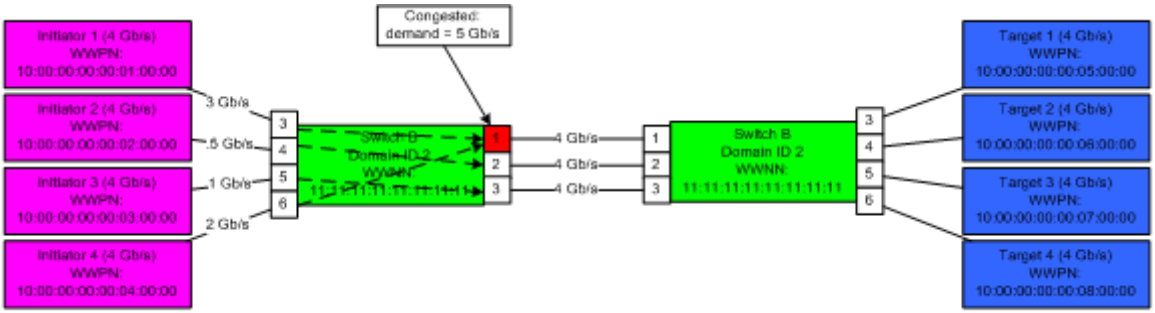


Figure 273 **Before frame-based trunking**

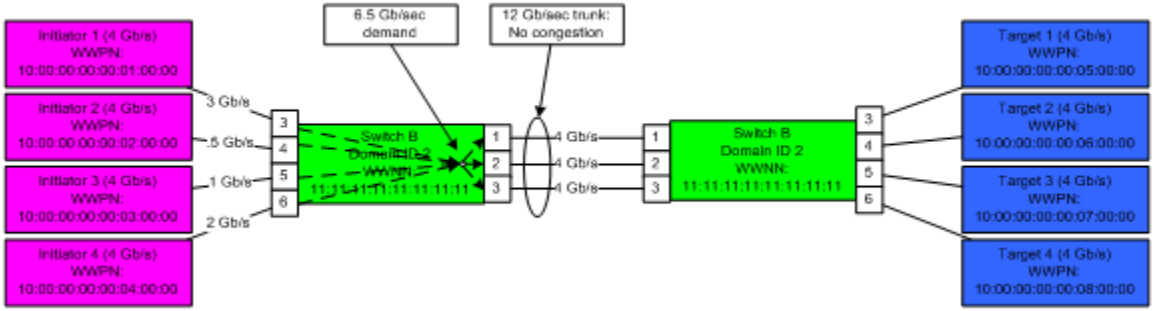


Figure 274 **After frame-based trunking**

Observe the following criteria for standard distance trunking :

- There must be a direct connection between participating switches.
- Trunk ports must reside in the same port group.
- Trunk ports must run at the same speed (either 2 Gb / s or 4 Gb / s).
- Trunk ports must be set to the same ISL mode (L0 is the default).
- Trunk ports must be E_Ports or EX_Ports.
- For optimal performance, cable lengths of no more than 30 meters difference are recommended.
- The switch.interopMode parameter must be set to 0.
- The port ISL mode must be disabled (using the portCfgIslMode command).

Note: For more infomration on trunking, refer to the *Brocade FOS 5.3 Administrators Guide.*

Through the use of special hardware to assist with the re-assembly of frames into the proper order, fame level trunking does not increase the likelihood of receiving OOFs..

PROs:

- Maximum potential utilization of ISLs
- Easy to configure

CONs:

- All ports need to be located on the same port group in order for a trunk to form
- Only works in Brocade Native environments

Exchange-based trunking

The next best thing would be one that could transmit the same number of frames over each ISL. The downside to doing this is that special hardware is needed on the receive side to ensure that the frames are reassembled into the proper order (to prevent OOOFs). By implementing trunking at the exchange level, exchange-based trunking prevents frames within an exchange from getting out of order by routing all of the frames for the exchange down the same ISL without requiring hardware-assisted exchange reassembly.

When Cisco first released their first FC switch, it came with a new kind of trunking that distributed I/O across ISLs based on the Source ID, Destination ID, and Exchange ID. Since its introduction, this method has been adopted by QLogic as well. The concept is very similar to Brocade's frame level trunking but instead of distributing the *frames* equally over all ISLs, *exchanges* get distributed over all ISLs. (Exchanges consist of sequences; sequences consist of frames.)

PROs:

- Works on heterogeneous ISLs (different vendors on each end)
- Does not require that each port be located on the same port group

CONs:

- Downstream congestion can potentially impact all members of the trunk.

IMPORTANT

This type of trunking only balances traffic on egress ports. For this reason, if one of the switches does not support a trunking functionality of some kind, the load will still be optimized, but only in one direction.

Port Channels

Another feature that Cisco released with their first director is the Port Channel. While a Port Channel is not actually a form of trunking (since exchange-based trunking can be done with or without a Port Channel being present), it is important to mention. This is a form of aggregating ISLs so they appear to be a single high-bandwidth ISL from the perspective of FSPF.

PROs:

- Ease of use — Logically grouping all ISLs between two domains enhances the users ability to administer all of them simultaneously.

CONs:

- Slightly more complicated to configure than normal ISLs. If multiple Port Channels exist between two domains, the user needs to ensure that the FSPF cost is the same or else the higher cost Port Channels will not be used.
- Since the default cost of a port channel is determined by <Link speed>/number of ISLs, the user needs to be mindful of how the introduction of a low cost path will impact their topology.
- If a single ISL in a Port Channel goes offline, all frames queued up to use that Port Channel will be discarded. These events are typically very rare and the Upper Layer Protocols will be able to re-drive the I/O that was impacted.

Brocade DPS — Enhanced Exchange-based trunking

DPS is similar to the Exchange-based trunking done by Cisco and QLogic with the exception that instead of making routing decisions based only on the hash of the D_ID, S_ID and OXID, another parameter, RxPort, is used. This is significant because by adding a fourth piece of data to the hash, the chance of encountering collisions is reduced, theoretically increasing the efficiency of the feature.

PROs:

- Can work on top of ISLs and/or frame trunks.
- Works on heterogeneous ISLs (different vendors on each end)
- Does not require that each port be located on the same port group

CONs:

- Downstream congestion can potentially impact all members of the trunk.

Downstream congestion can potentially impact all members of the trunk.

IMPORTANT

This type of trunking only balances traffic on egress ports. For this reason, if one of the switches does not support a trunking functionality of some kind, the load will still be optimized, but only in one direction.

Open trunking

The last type of trunking to be introduced into FC Fabrics was McDATA's Open trunking. Open trunking works by making periodic updates to switch routing table. The updates are made based on the percentage of ISL utilization, the amount of time spent at zero BB_Credit, and the source ports utilization percentage. Open trunking attempts to find the best mix on each ISL to minimize congestion. Because updates are made to the switch routing table, an entire ingress port to a given domain (also known as a flow) is moved instead of one exchange or frame. On the BrocadeMi10k, the feature was improved to not only consider the ingress port and destination ID. but also the port on the destination ID.

PROs:

- Works on heterogeneous ISLs (different vendors on each end)
- Does not require that each port be located on the same port group
- All ports do not need to be located on the same port group in order for trunking to work.
- Downstream congestion will not impact all members of the trunk. See "Congestion and backpressure" on page 877.

CONs:

- The periodic changes to the switch routing table will occasionally result in a small number of OOFs.
- Open Trunking not capable of fully saturating all ISLs since the level of granularity only goes down to the flow level and not to the exchange or frame.

IMPORTANT

This type of trunking only balances traffic on egress ports. For this reason, if one of the switches does not support a trunking functionality of some kind, the load will still be optimized, but only in one direction.

Congestion and backpressure

The terms *congestion* and *backpressure* are sometimes used interchangeably, but although they are closely related, they are very different.

- *Congestion* occurs at the point of restriction.
- *Backpressure* is the effect on the environment leading up to the point of restriction.

To help explain the concepts of congestion and backpressure, we will use a mass transit system.

For the sake of this backpressure example, assume that a bus leaves from terminal "A" destined for terminal "B" once per minute and each bus is capable of carrying 40 passengers at a time (Figure 275). As long as the number of passengers arriving at terminal "A" does not exceed 40 passengers a minute, the system will work perfectly fine. The busses will carry less than their full load of 40 passengers per minute and there will be no passangers waiting in the queue.

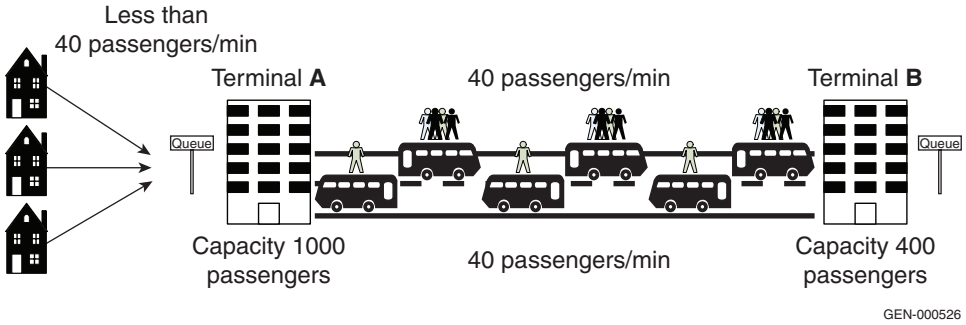


Figure 275 Backpressure example using mass transit system

However, when the number of people arriving at terminal "A" exceeds 40 passengers a minute, (even by a small number, like 42 passengers a minute), a number of things happen (see Figure 276 on page 879).

- The first impact on the environment is an increase in latency through the system. In other words, it will take more time for a passenger to get from terminal "A" to terminal "B"since each passenger will have to wait in a queue for some period of time.
- Occurring in parallel with this increase in latency is an increase in the amount of "buffer" space being consumed by the number of passengers waiting for a bus. Since each passenger consumes some amount of "buffer" space, and each passenger is waiting longer due to the increased latency, the amount of buffer space used increases over time.

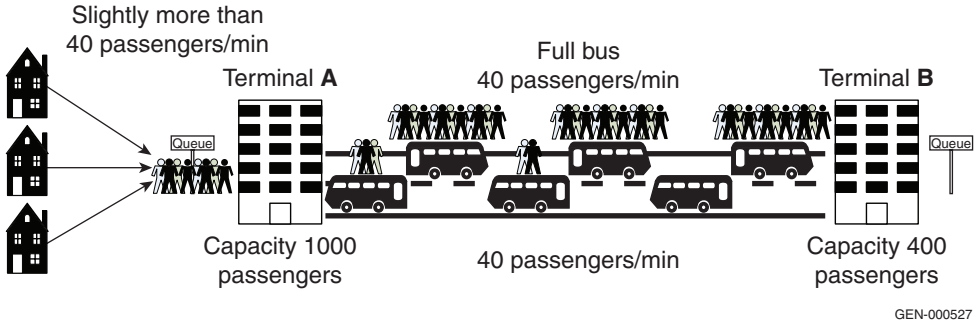


Figure 276 Increase in consumption of buffer

As shown in Figure 277, if passengers continue to be arrive at a rate faster than 40 per minute, the "buffer" would eventually fill up and overflow (people would have to wait outside the terminal).

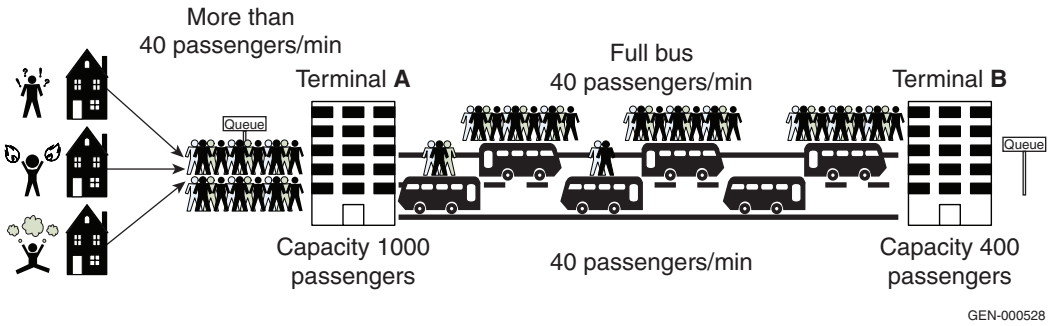


Figure 277 Buffer filled causing overflow

In the case of a Fibre Channel link, because of the use of buffer-to-buffer flow control the buffer does not overflow. Instead, a phenomenon known as *backpressure* is experienced. Think of an terminal employee notifying people planning to come to the terminal that there is no space available and you have an idea of how flow control works.

For the rest of this section, the cause and effects of congestion and backpressure is investigated as well as how to judge if your environment is inherently susceptible to these issues. In addition, information on how to detect these conditions are provided.

To make things easier to explain, the topology shown in Figure 278is used.

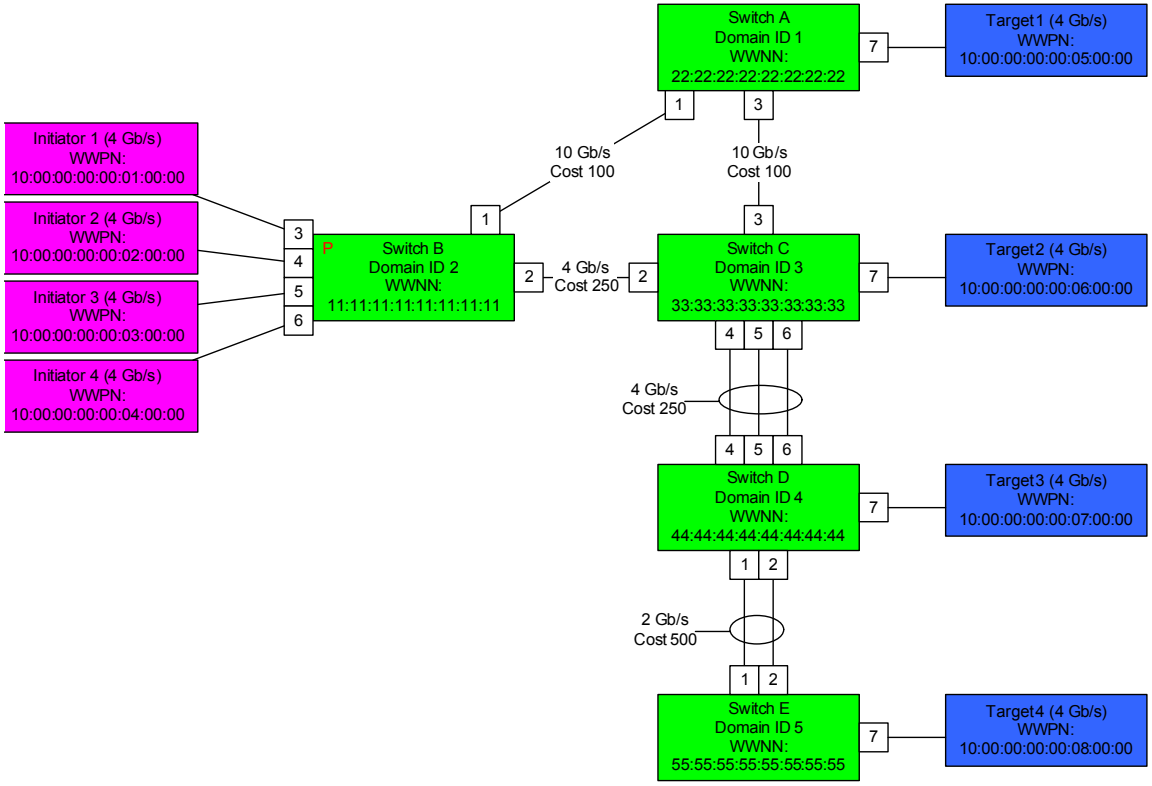


Figure 278 Topology example

For the rest of this example, assume that each initiator is zoned so that it only has access to one target; for example:

- Initiator 1 — Target 1
- Initiator 2 — Target 2
- Initiator 3 — Target 3
- Initiator 4 — Target 4

If we take this configuration and apply the same sort of situation to it as we did in the terminal example, you will gain a better understanding of flow control in an FC fabric works. Before we do this, we will first drill down into the case where only Initiator 1 is sending frames to Target 1 at 4 Gb/s.

Buffer-to-Buffer flow control

Buffer-to-Buffer Flow control is the mechanism used to ensure that a transmitter does not overwhelm the receiver with too many frames. It relies on the use of credits which are exchanged at login time (FLOGI for Nx_Ports, ELP for E_Ports). In the case of an Initiator or Target (Nx_Port), the port will grant some number of credits (typically between 8 and 64) to the switch. It does this by specifying the number of credits to be granted to the switch in the FLOGI frame. If the switch accepts the FLOGI request, the number of credits that the switch will be extending to the Nx_Port will be contained in the FLOGI Accept frame. An example of the login and credit initialization process is shown in Figure 279.

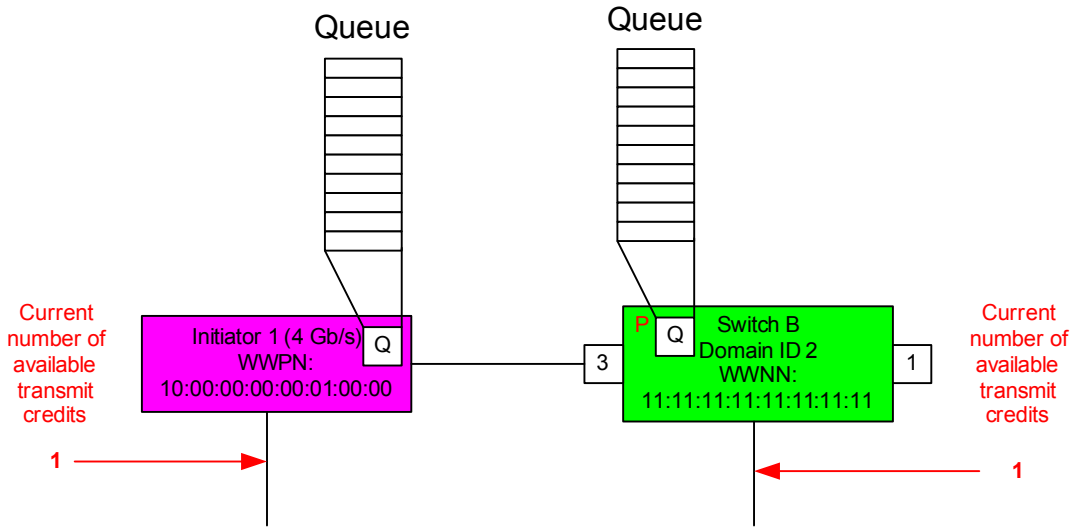


Figure 279 Login and credit initialization process example

Figure 279 shows an initiator and a Fibre Channel switch. Both the Initiator and the switch are shown as having a queue, each containing 10 buffers. Each buffer is capable of storing a single FC frame. All incoming frames will be stored in a buffer until they are ready to be processed by the switch or transmitted by the egress port. After the fiber cable is attached, the link will gain sync and eventually enter the ACTIVE state where the ports are transmitting IDLEs. After the ports are transmitting IDLEs, but before the FLOGI is transmitted, both ports have a single credit by default. This is so each port will be able to transmit a single frame, login, and negotiate credit.

"Receive" versus "Transmit" Buffer-to-Buffer Credits

Ideally, each port should keep track of two credit counts:

- The number of BB_Credits or Buffers that it has remaining available to receive incoming frames (RX BB_Credit)
- The number of BB_Credits or Buffers that are available on the port on the other side of the link (TX BB_Credit)

Each are further explained next.

- Receive Buffer-to-Buffer Credit (RX BB_Credit):

  The RX BB_Credit count is simply a count of how many buffers the local queue has remaining to accept incoming frames. When this number is zero, no additional Fibre Channel Frames should be received. However, in practice this is not always the case. Sometimes due to a bug, the transmitter will send more Frames than for which it has been granted credit. This condition is called a *buffer overflow* and how it is handled is implementation dependent.

  Some ports, especially switch ports, will allow an overflow to occur and even route the FC Frames received when RX BB_Credit is zero to their proper destination. They do this by keeping some number of spare buffers available for conditions like buffer overflow and storing these extra frames in these spare buffers.

  Other implementations will discard the frame and reset the link with LR (Link Reset Primitive) to re-initialize the BB_Credit counts to the login values.

  Both of these are acceptable, with the former being more forgiving and the latter being more rigid but closer to the spirit of the standard. In either case, as long as the transmitter is not constantly overflowing the buffer, both approaches work.

  If the transmitter is constantly overflowing the receive buffer, we enter a "works fine with vendor A (allows overflows) but not with vendor B (resets the link)" kind of interoperability problem.

  On the port on the other end of the link, the RX BB_Credit count is kept track of as the Transmit Buffer-to-Buffer Credit (TX BB_Credit) count.

- Transmit Buffer to Buffer Credit (TX BB_Credit):

  In addition to keeping track of how many receive buffers a port has left available, the number of receive buffers left on the port on the other end of the link needs to be kept track of as well. This is done through the use of the TX BB_Credit count. The count is

kept track of by the transmitter by decrementing the TX BB_Credit count by one for every frame it transmits and incrementing it by one for every Receiver Ready (R_RDY) primitive that it receives.

Receiver Ready (R_RDY)

The R_RDY is a 4 byte primitive sequence that indicates a buffer has been freed on the port that transmitted the R_RDY. It is important to note that a port that transmitted a frame and received the R_RDY is unable to determine which frame the R_RDY was sent for. Also, the R_RDY does not indicate that the frame was received in good condition or that it was forwarded successfully. It just indicates that a buffer on the receiver was freed and is ready to store another frame. The transmission of an R_RDY is also known as "releasing a credit."

BB_Credit loss

Occasionally, it is possible for the TX BB_Credit on a transmitting port to get out of sync with the RX BB_Credit on the receiving port. This typically happens for two reasons: Frame corruption or R_RDY corruption, each discussed next.

- **Frame corruption –** Due to a bit error on the link, a frame is corrupted in such a way so that it is not recognizable as a frame on the other end of the link. This is typically due to a corrupted **Start of Frame** delimiter. Since the receiving port does not know that a frame is being received, it would just detect non-frame data and depending on the implementation, handle it in a number of different ways. However, it would not release a credit since it did not get a frame.
- **R_RDY corruption –** Due to a bit error on the link, an R_RDY primitive is corrupted in such a way so that it is not recognized on the other end of the link. Since the port never receives an R_RDY, it never increments the TX BB_Credit count.

In either case, if the transmitter started with 10 TX BB_Credits and then decremented this count by one after sending the frame, the counts would be out of sync since the transmitter would think it only had 9 TX BB_Credits left while the receiver would think that 10 were still available. If this were to happen often enough, a transmitter could lose most of the TX BB_Credit and poor performance, particularly over distance, would result. See the "Poor performance example," next, for an actual case study with a recommendation how to detect and resolve the problem.

Poor performance example

Ignoring the long roundtrip times in distance applications and the error recovery mechanisms that may be invoked by the initiator or

target, which are by no means trivial, a single bit error can have a surprisingly large impact on the throughput rate. For example, take a 4 GB/s link with a single bit error. Many host and storage ports will re-drive an entire exchange upon a single bit error, so for the sake of this example, assume that it is a 64k write (usually 32 2k frames). In this example, you would need to re-send these 32 frames, which means that during the simple second they are being transmitted, your maximum data transfer rate is 412.25 MB/s minus .064 MB/s = 412.17 MB/s.

This may sound insignificant, but when you consider that 1 bit represents .0000000235% of the total bits transmitted in a second, and it cuts the throughput by 412.17/412.25 = .24%, you are talking about a difference of seven orders of magnitude. If you have 10 bit errors per second, your effective throughput will be decreased by 2.4%. If there are 100, it will be decrease by 24%. Again, this does not consider the impact that invoking error recovery mechanisms will have on the environment.

There is another more insidious problem when bit errors are being logged: should one happen in an R_RDY, the link will lose a credit that will not be recovered until the next Link Reset. To discover the odds of having a bit error impact an R_RDY, read the next section, "Buffer-to-Buffer Credit issue."

Buffer-to-Buffer Credit issue

As previously mentioned, a Fibre Channel link consists of a transmitter and a receiver. Each receiver includes a set of frame buffers. Buffer-to-Buffer Credit is the flow control mechanism that ensures transmitted frames will be received.

Whenever a transmitter sends a frame it decrements its credit counter. The transmitter will continue to send frames and decrement the counter until the counter is zero. When a receiver succeeds in releasing one of its buffers by moving the frame to the next port in the route, the receiver sends an R_RDY to the transmitter. The transmitter increments its Buffer-to-Buffer Credit counter for each R_RDY it receives.

When the link is short and the route is uncongested, the counter is expected to stay near its maximum value. When a link is congested, this value may fluctuate heavily and spend time at zero credit. This is a normal and expected behavior of a Fibre Channel environment. As soon as the congestion is relieved, the frames are forwarded and

Credit is recovered. The source of this backpressure may be a few hops away, originating with an N_Port.

What can cause the Buffer-to-Buffer Credit issue to trigger?

Performance of an FC link depends on the BB_Credit counter being incremented by the receipt of an R_RDY for every frame it transmitted. If an R_RDY is not received, the counter does not increment. If this happens repeatedly, the credit will be decremented to zero and the link will stop functioning.

In a DWDM environment servicing a distance application, BB_Credit is necessary to keep the link fully utilized. The loss of credit can lead to ìdroopî, a condition where there is insufficient credit to support the bandwidth of the link. A 30 km 2 GB/s link will start to droop if available credit falls below 30. The link will continue to operate but at reduced performance. This credit leakage scenario can be caused by two events:

- R_RDYs are corrupted.

  One side thinks credit has been given back, but the other side never receives the credit. The credit counter will not be incremented.

- One side sends a frame and decrements its credit, but the other side never receives the frame.

  Under this scenario the receiver will not launch an R_RDY. The same effect occurs if the SOF (Start_of_Frame) field is corrupted.

The R_RDY and SOF field represent a small percentage of the consumed bandwidth. R_RDY and SOF are four bytes each. A typical data frame will consume 2104 bytes. Given the corruption is randomly distributed, for every corrupted R_RDY or SOF frame you would expect 263 data frames (2104/8) to have been corrupted. For every data frame corrupted, the FC-SCSI Exchange will have to go through an error recovery that can significantly impact response time.

In the case of SRDF, there is a three-second recovery timer. An IO that was supposed to be completed in 10 milliseconds or less could see a 300x slowdown for each corrupted exchange. In order for a 30 km link with 60 credits to lose 30 credits and cause the link to drop, it is likely 7,890 frames (263x30) will have been corrupted and each one will have impacted response time.

As a result, it is best to have a link quality monitoring strategy that monitors overall frame integrity. Monitoring or automatically recovering BB_Credit would mask the real problem, subjecting the environment to ongoing response time penalties. In addition it is very difficult for the switch to distinguish between lost credit and normal congestion events in a timely fashion.

How to resolve with minimal intrusion:

EMC recommends a two-prong monitoring strategy:

- Characterize the nature of the corruption events on both sides of the link by using a monitoring tool to monitor and record link quality metrics.
- Implement an alerting system on both sides of the link to trigger before the risk of a business impact is high.

The most sensitive means of tracking link quality is with the Invalid Transmission Word (ITW) counters in the switch.

EMC also recommends setting up two Counter Threshold Alerts, one for short bursts and one for multiple single occurrences distributed in time.

- The short burst alert would be 40 ITWs in 5 minutes.
- The single occurance alert would be 100 ITWs in 24 hours.

These values are an initial recommendation and may need adjusting to avoid false alarms. These values are expected to trigger long before 30 credits are lost and performance is impacted.

The performance impact level calculated earlier in this section was 7,890 corruption events. The thresholds of 40 and 100 should provide adequate early warning. A third trigger could be created with a threshold of 4000 ITWs. When encountered, this would indicate to the user that they should manually execute a Port Reset to recover any lost credit.

The most critical statistics to monitor are:

- Invalid Transmission Words
- Class 3 Discards
- CRC Errors

Although setting thresholds and monitoring links for bit errors can help detect the loss of credit, it would be helpful if two devices were able to detect this condition themselves and recover from the loss of a credit. Recently, a few vendors have started to implement BB_Credit Recovery which can help to do just that.

BB_Credit Recovery

BB_Credit Recovery is defined in FC-FS-2 and is only implemented on a few of the products that E-Lab currently qualifies. It is expected to be very useful in long distance applications. It works by sending special primitives at intervals which are agreed upon at login. The intervals are defined by the number of frames transmitted and the number of R_RDYs transmitted. If a mismatch is detected then an appropriate number of R_RDYs are transmitted in order to resolve the mismatch. A port specifies that it is capable of supporting BB_Credit Recovery at login time.

Login and credit initialization

Note: To reduce the number of diagrams that would otherwise be needed in this section, vertical bars will be drawn under each SAN component and horizontal (slightly diagonal) arrows will be inserted between the vertical bars to indicate the transmission of a frame between components. Individual frames, such as the FLOGI, will not be shown wherever possible.

As shown in Figure 280, once the initiator transmits the FLOGI, it will have zero transmit credit left and will be unable to send any other frames. An accept to the FLOGI is needed before the initiator can use the fabric.

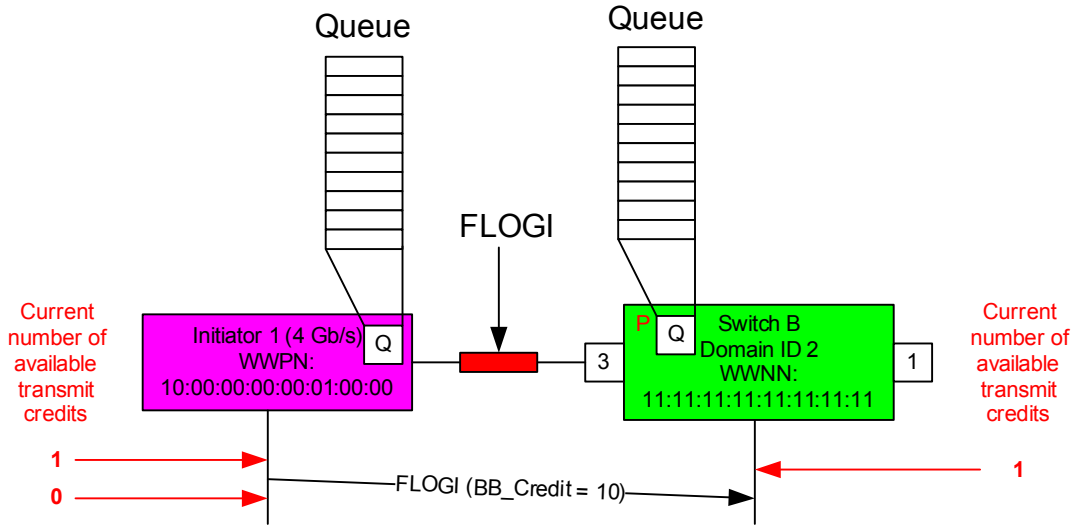


Figure 280 **Login and credit initialization**

As shown in Figure 281, once the switch receives the FLOGI, it will store it in a buffer until the switch is ready to process it. The processing begins when the FLOGI frame is forwarded to the Fabric Server logic within the switch.

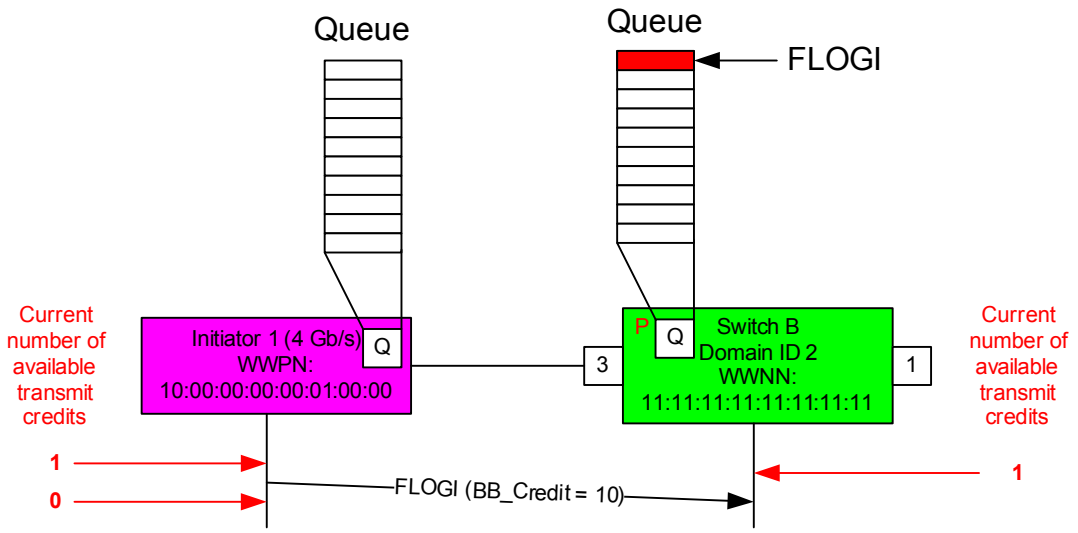


Figure 281 Switch receives the FLOGI

As shown in Figure 282, once the switch processes the FLOGI, it will have 10 transmit credits to send frames back to the initiator with. Next, the FLOGI Accept frame is transmitted by the switch which decrements its number of transmit credits by one. Once the FLOGI Accept frame is received by the initiator, it will update its transmit credits to the value in the FLOGI Accept and then it will release a credit by transmitting an R_RDY.

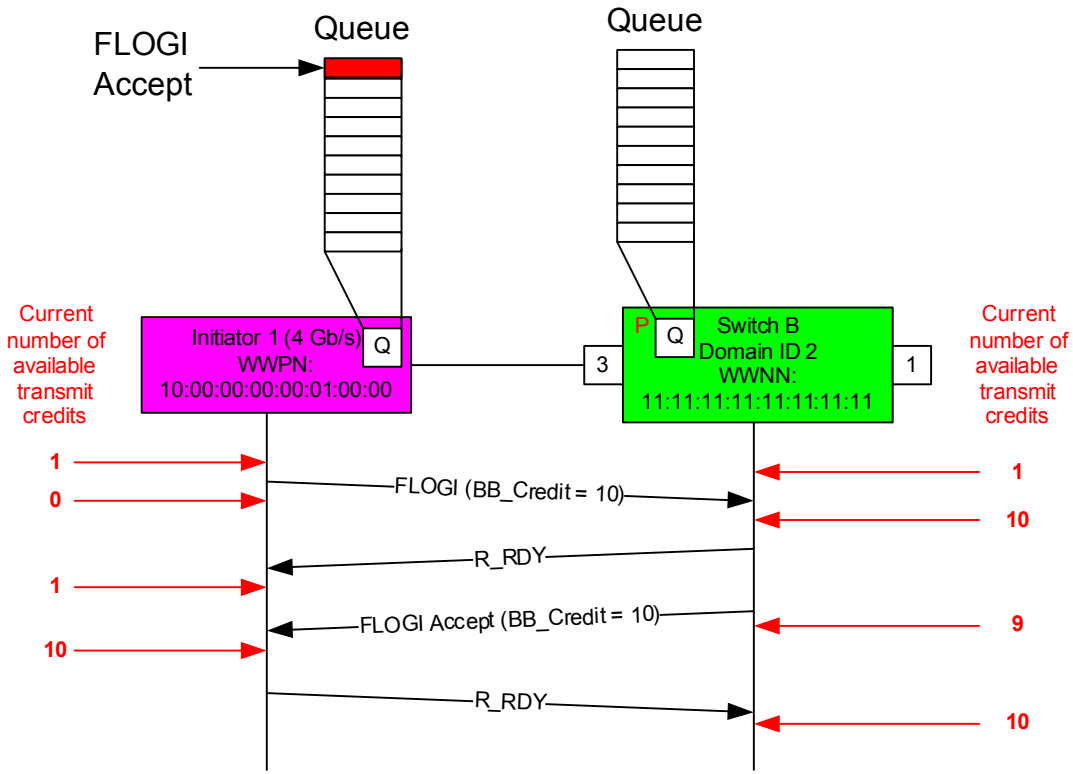


Figure 282 **Switch processes the FLOGI**

Queue types

In Figure 283, Initiator 1 is attached to Switch B. Above Switch B is a *queue*. The queue is where any incoming frames are stored until they

are ready to be transmitted. Each queue has 10 buffers and each buffer is capable of storing a single FC frame.

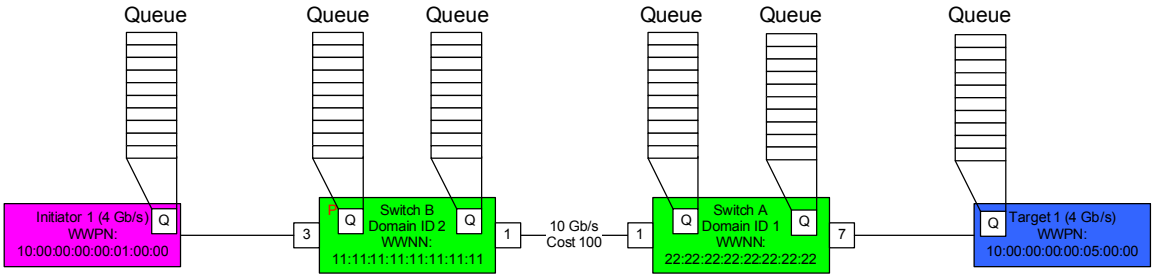


Figure 283 Queues

For the purposes of this discussion, we will consolidate the different queuing mechanisms in use into two groups:

- Shared memory
  Shared memory buffers lump all of the buffers used by a switch or set of ports into one common buffer area and then the transmitting ports read frame data from this buffer area as they are transmitting the frame.

- Virtual Output Queues

  Virtual Output Queues are typically created at a switch ingress port. They work by grouping frames for a given switch egress port (output port) on the switch into the same virtual queue and then these frames are transmitted in the same order they were received. This ensures that frames will not be delivered out of order while at the same time ensuring that there are no head of line blocking problems.

  Virtual Output Queues have finite resources so, although Virtual Output Queues to a given egress port can grow and shrink independently of one another, they are limited to the total number of buffers available in the Queue. In other words, it is possible for the Queue to be completely consumed by frames all destined to the same egress port. The other Virtual Output Queues will still be there but there will be no space left available in the queue.

The type of queue that is used by a switch can vary on a product-by-product basis. Each different type of queuing has it pros and cons but at some level they all do the same thing: *they store frames*.

When a switch detects the Start Of Frame (SOF) delimiter, it stores the SOF, all of the data that follows (up to a certain amount, usually about 2K bytes), and the End Of Frame (EOF) delimiter into a queue (see Figure 284).

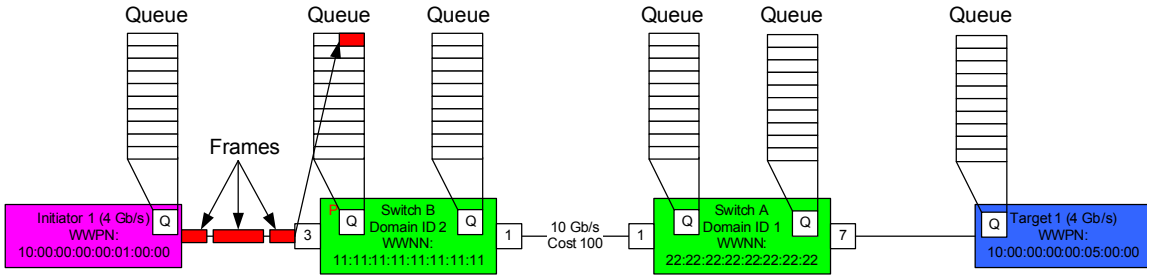


Figure 284 Queue example

What happens next depends on the type of queue being used. In any case, once a switch is ready to transmit a frame, the transmitting port will start pulling the frame out of the queue and begin transmitting it, as shown in Figure 285.

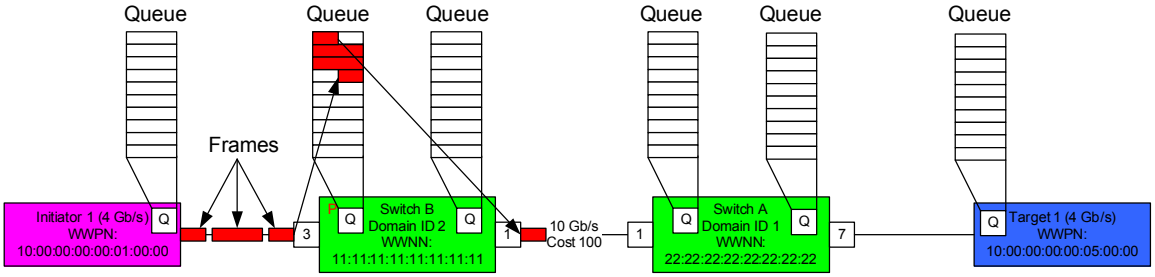


Figure 285 Transmit example

In Figure 285 there are parts of 4 frames in the queue. Depending on the switch architecture, when there is no contention for transmit resources (as implied above) it may be possible to immediately start transmitting the frame without needing to store the entire frame first. This is sometimes referred to as *cut through routing* .

Some switches that do not support true cut through routing do have low enough latency so that they are not considered *store and forward. Store and forward* architectures are those that may not start transmitting a frame until the entire frame has been received.

As shown in Figure 286Figure 286, once Switch B has transmitted the frame, the buffer is freed and a credit is released via R_RDY.

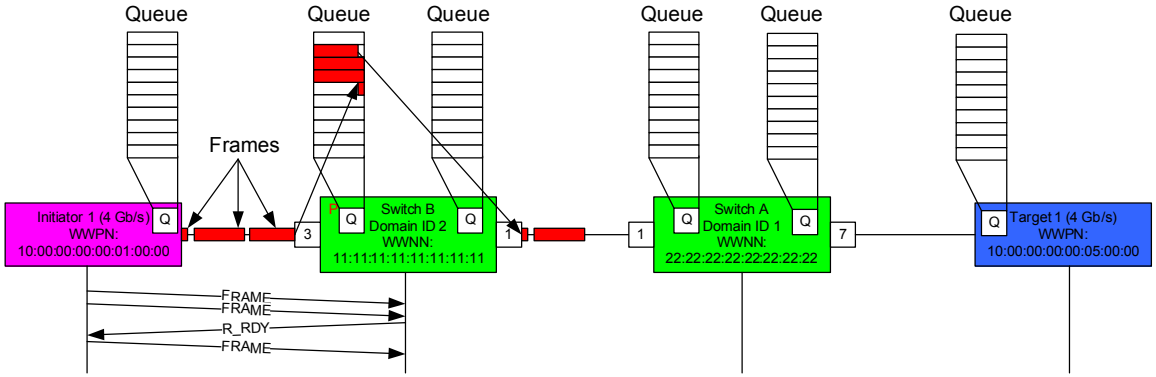


Figure 286 Released credit

Sources of backpressure and congestion

The example of a bus station was provided at the beginning of "Congestion and backpressure" on page 877, to illustrate the concepts of congestion and then backpressure. It was stated that as long as passengers left the terminal at the same rate as they entered, the transportation system would function properly. The same is true for a Fabric. As long as average number of frames being transmitted from the switch is equal to the average number of frames being received by the switch, congestion and backpressure will be minimal. In some cases however, it is not always possible for frames to be transmitted as quickly as they are received and, as is the case with the transportation system, the Fabric will experience congestion. If the congestion persists long enough, backpressure will result. There are two primary reasons for congestion:

- A slow drain device
- Oversubscription

Both are further explained in this section.

Slow drain device

Occasionally, due to an architectural limitation, overall system load, poor volume layout, or a malfunctioning internal component, it is not possible for a destination Nx_Port to process frames at the same rate as they are being received from the fabric. An example would be a target port without enough cache. As shown in Figure 287, if the average number of frames being sent to a target port exceeds its ability to process those frames, Queues will start to fill up. Notice that

the difference in transmission and reception rates does not have to be very large. A difference of 60 KB/sec would be enough to cause the congestion secenario illustrated in Figure 287.

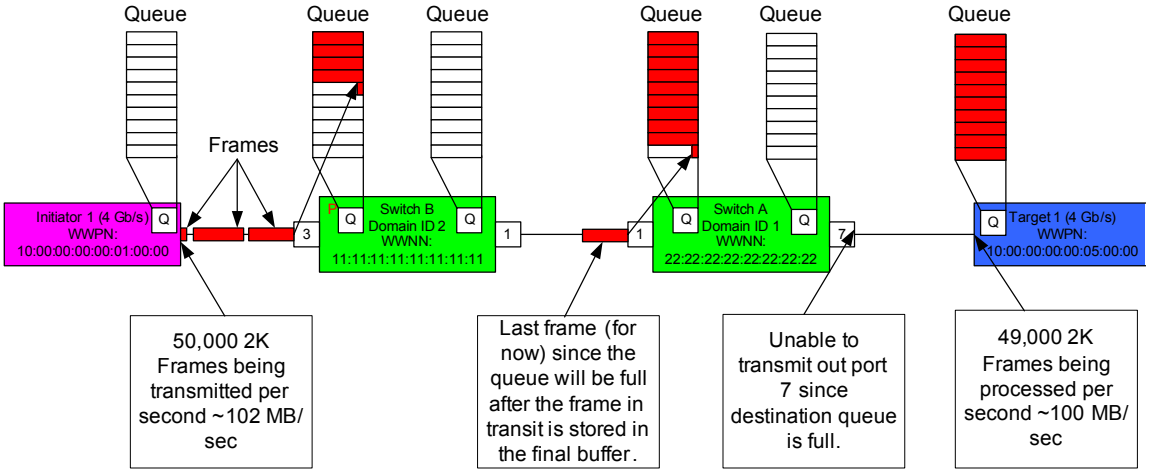


Figure 287 **Slow drain example**

In Figure 287, multiple points are experiencing congestion and backpressure. The congestion at target 1 is causing backpressure to be experienced by Switch A port 7 since it is unable to transmit. This in turn causes congestion at port 1 on Switch A which results in backpressure at switch B port 1. Port 3 on switch B is not congested (yet) and as a result the initiator is not feeling any backpressure.

In this scenario, only the initiator and target being shown would be impacted by this mismatch but as will be shown shortly, this is not the case in a real-world environment containing multiple initiator and target pairs.

Two initiators and targets no congestion and backpressure

Figure 288 shows an uncongested environment containing multiple initiators, each one transmitting to their own target.

Note: All Queues should be considered to have the same number of buffers even though they are not displayed that way in the illustration below. Although the Queue located near each port is intended to indicate a virtual output queue, the same type of issues can also be experienced in environments utilizing shared memory types of buffers.

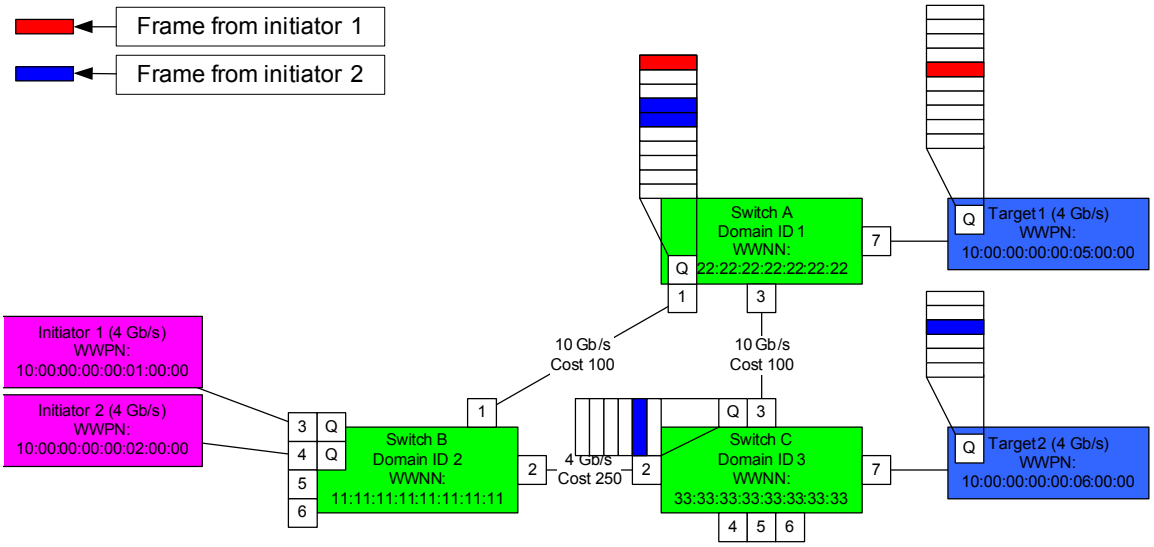


Figure 288 **Uncongested environment**

Two Initiators one slow drain causes congestion or backpressure:

Figure 289 on page 895 through Figure 291 on page 896 illustrate what can happen when a single slow drain device is present in a fabric. As shown in Figure 289, the queue on target 1 is full. This can happen for any number of reasons, as explained in "Slow drain device" on page 892.

Note: For the sake of this example, assume that both initiators are transmitting at the same rate but that Target 1 is handling frames at a rate which is less than they are being transmitted by the initiator.

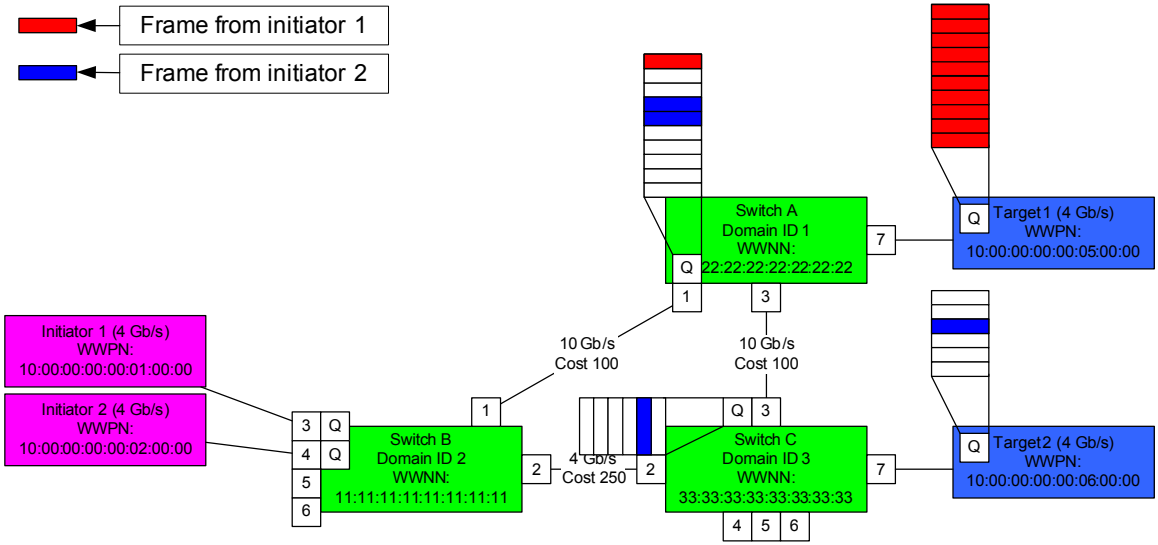


Figure 289 **Impact of a slow drain port**

The impact this slow drain port will have on the rest of the fabric will first be felt on Switch A, as shown in Figure 289. Since the Queue on Target 1 is full, port 7 will have no transmit credit and will be unable to transmit any of the frames in Switch A, Port 1's Virtual Output Queue for port 7. However, since the Queue on Target 2 is not full, little of the Virtual Output Queue for port 3 (within the Switch A Port 1's queue) will be consumed.

If this condition persists more and more of the total number of buffers on Switch A, port 1 will be consumed by the Virtual Output Queue for port 7 and fewer and fewer buffers will be available for port 3 (see Figure 290 on page 896).

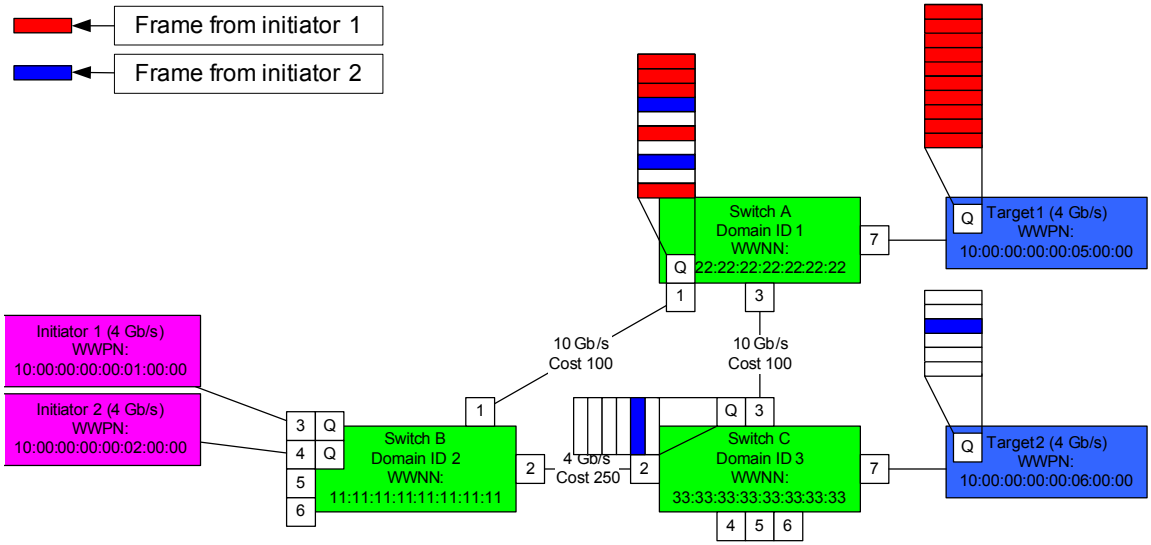


Figure 290 Virtual Output Queue for port 7 continues to grow

As shown in Figure 291, Switch A, port 1's Queue has been completely consumed by frame destined for Target 1.

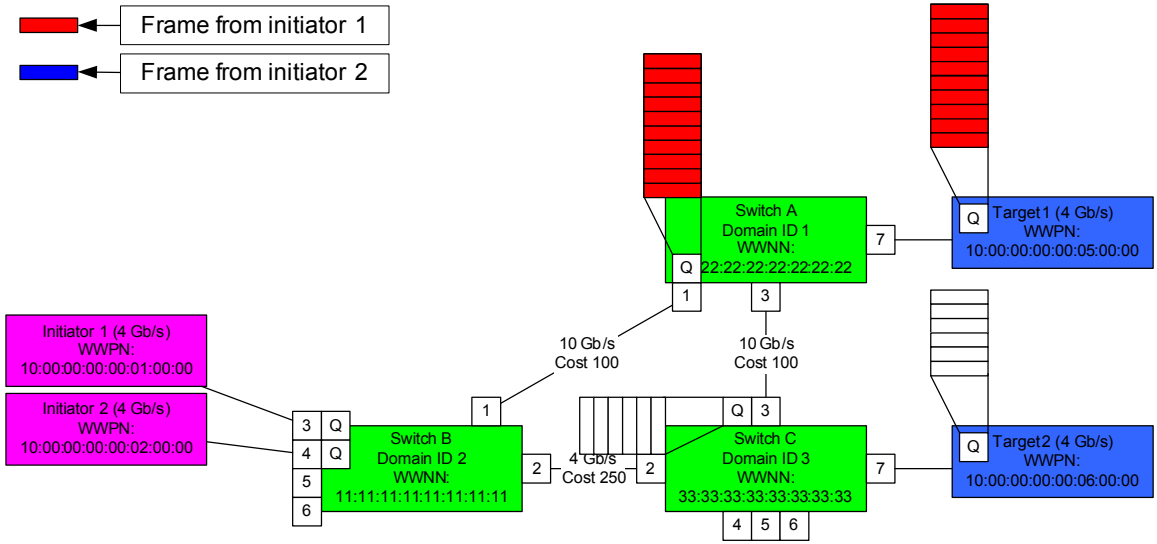


Figure 291 Virtual Output Queue for port 7 consumes entire Queue on Switch A, port 1

This condition cannot persist indefinitely for two reasons: the hold timer and the Link Reset Protocol, each discussed next.

- Hold timers

  The hold timer is the amount of time that a switch will allow a frame to sit in a queue without being transmitted. When the timer expires, the frame is discarded. Hold timers for Connectrix B Series, Connectrix M Series, Connectrix MDS, and QLogic are shown in Table 44.

Table 44 Hold timers

Vendor	Firmware	Hold Timer (milleseconds)
Connectris B Series FOS	All	500
Connectris M Series M-EOS	5.x and lower	500
Connectris M Series M-EOS	6.x	1000
Connectris M Series M-EOS	7.x and higher	1600
Connectrix MDS	All	500
QLogic	All	2000

  This is not really a solution to our problem with the slow drain since average frame latency time through a switch is between 600 ns to 20usec. The shortest hold timer above is 25,000 times greater than the longest average latency. The end result is that the hold timer has limited value for this particular scenario.

- Link Reset Protocol

Note: This is a worst case scenario and happens infrequently. Typically, a slow drain is only a slow drain periodically and this will result in occasionally drops in total throughput. Finding a slow drain, especially without the use of a protocol analyzer, is one of the most challenging performance problems to troubleshoot.

If Target 1 does not release a credit within E_D_TOV (2 seconds), port 7 on switch A will transmit a LR (Link Reset) forcing the BB_Credit values back to the login values. Upon receiving the LR, the receiver will discard all frames in the local queue. This will allow the queue on switch A, port 1 to be drained by Switch A, port 7 since it will now have transmit credit (at least temporarily). Since switch A, port 1 will now have buffer space, both Initiator 1 and 2 will be able to resume transmitting frames.

Oversubscription

There are a few different types of oversubscription that need to be considered.

- The first type trys to move more data across a link than it is capable of carrying. This is the classic case of having x capacity and needing 2x capacity.
- The second type is congestion due to a lack of BB_Credit. This is typical in a distance extension environment and can be architected around by proper planning. See Chapter 3, ”Extended Distance,” for more information.

For the case where not enough capacity is available, it may be possible to architect around this issue through proper planning.

However, when an ISL fails, as shown in Figure 292, it is still possible to encounter congestion and backpressure.

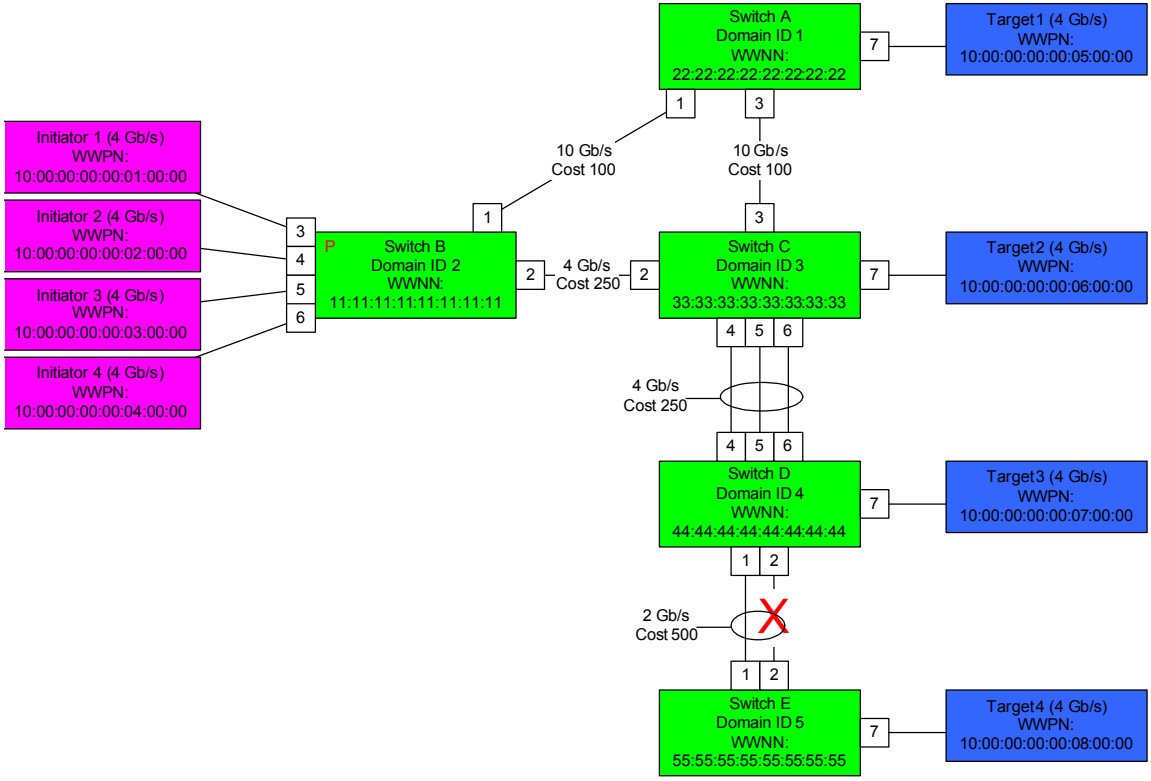


Figure 292 **ISL failure**

In Figure 292, an ISL has failed between Switch D and Switch E. If Initiator 4 was transmitting to Target 4 at something greater than 2 Gb / sec, congestion would occur at Switch D, port 1 and backpressure would begin to spread back towards the initiator. As shown in the previous examples in this section, the impact this could have on the entire environment could be dramatic, so planning for this kind of failure is extremely important.

FLOGI

Once the link comes up to the active state, FLOGI (Fabric Login) is the first frame transmitted by a N_Port when it is attempting to log into a fabric. The N_Port uses the FLOGI frame to ask for permission to use the fabric and to request an FCID. The FLOGI request contains

the N_Ports Node WWN and Port WWN as well as the classes of service supported. It also contains the number of credits being extended by the N_Port to the switch.

The FLOGI ACC from the switch back to the N_Port contains the N_Ports FCID in the D_ID field, the switches WWPN, the WWNN, the supported classes of service, and the number of credits being extended by the switch to the N_Port.

There is no requirement that the BB_Credit values used by the N_Port and the switch be equal to each other. If an N_Port does not send the FLOGI within two seconds of coming up to the active state, the switch may attempt to initialize as an E_Port and transmit the ELP (Exchange Link Parameters) SW_ILS request. In this case, you could expect to see a segmentation reason for ELP failure or E_Port initialization failure.

Fore more information, refer to "Fabric Login (FLOGI)" on page 911.

Nodes

A Fibre Channel environment consists of two or more devices connected together by an interconnection topology. In Fibre Channel devices such as personal computers, workstations, disk array controllers and disk and tape drives are all referred to as nodes. Each node is a source or destination of information for one or more nodes. In EMC's case the node would be the Symmetrix system.

Each node requires one or more ports to provide a physical interface for communicating with other nodes through their ports. The port is a hardware attachment that allows the node to send or receive information using the physical interface. Some devices have these ports integrated and other have pluggable ports such as Host Bus

Adapters for flexibility. In EMC's case, the port would be the port on the Symmetrix FA adapter.

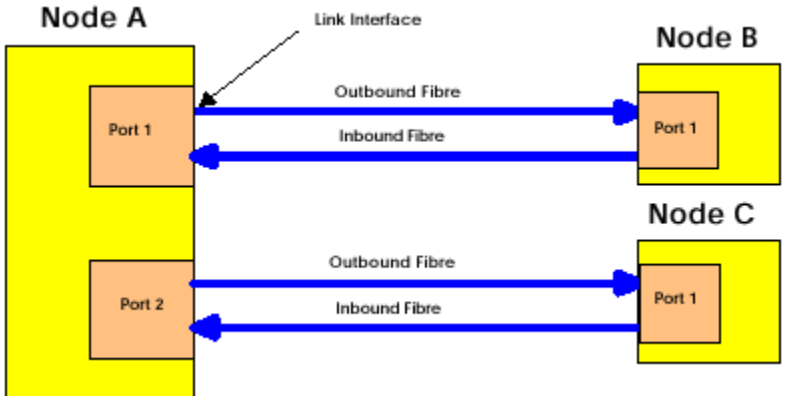


Figure 293 **Nodes**

Flow control

When a port wants to send frames to another port the frames are sent from a buffer at the sending port and received into a buffer at the receiving end. However while a frame is being received at the receiver end if another is transmitted at the sender side then more than one buffer is needed to handle this at the receiver side. If there is no receive buffer space then that port cannot accept any more frames and this leads to busy responses and the possibility of frame loss.

To prevent this type of behavior and regulate the availability of receive buffers Fibre Channel uses two different flow control mechanisms to pace the rate at which the sender is allowed to transmit frames. The first is a link level flow control and is called Buffer-to-Buffer flow control (discussed next). The second is a source and destination based flow control named End-to-End flow control (refer to "End-to-End Credit" on page 902). Both of these methods of flow control use a credit-based agreement between nodes to regulate the flow of frames between them. The credit value is the number of frames a receiving port has allowed a sending port to transmit to it.

For more information, refer to "Buffer-to-Buffer flow control" on page 881.

End-to-End Credit

End-to-End credit (EE_Credit) is the maximum number of data frames a source port can send to a destination port without receiving an acknowledgement frame (ACK). This credit is granted during N_Port login and is replenished with the return of an ACK response frame, as shown in Figure 294.

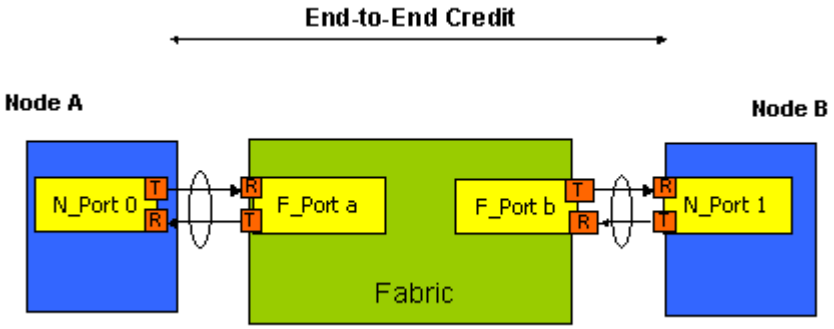


Figure 294 **End-to-End Credit**

EE_Credit is between the source and destination port and is only used in Class 1 and Class 2 circuits which are described further in the section Class of service.

Figure 295 shows a summary of the mechanics of flow control and the main points to be considered for both End-to-End and Buffer-to-Buffer. It should be noted that Symmetrix operates in a class 3 circuit and thus uses Buffer-to-Buffer flow control.

	End to End flow control	Buffer to Buffer flow control
Regulate frame flow from	Source to destination N-Ports.	Transmitting N/F port to receiving N/F port over a link.
Class of service	1 , 2	2 , 3
Credit Counter	EE_Credit	BB_Credit
Signal	ACK	R_RDY
Frame Type	Data Frame only	All frames

Figure 295 **Summary of flow control**

Error detection and recovery

In Fibre Channel, error detection occurs at a number of different levels, further discussed below:

- At the FC-0 level errors associated with signal quality where primitive signals are used to reestablish link integrity.
- At the FC-1 level where it checks for invalid transmission characters or disparity issues decides whether the frame is valid or invalid.
- At the FC-2 level frames are checked for various conditions such as CRC error detected, invalid transmission word detected or the case where no receive buffer is available for the frame.
- At the FC-2 level when all active sequences have had a sequence time-out a link time-out is detected and a Link Reset is initiated.
- At FC-2 where during transmit or receive of a sequence the sequence initiator or recipient has detected an error and an action needs to be taken. For example detecting missing frames in a sequence or corrupted frames. Corrupted frames are discarded and the resulting error is detected and recovered at the sequence level. A missing frame is detected at the receiver by missing SEQ_CNT values and at the initiator by a time-out value being exceeded (E_D_TOV).

Time-outs provide a mechanism to detect where a certain operation did not occur within a certain time or did not occur at all and prevent ports hanging indefinitely. For example, if an HBA sends a write command to the Symmetrix FA and it does not send a transfer ready back then after a certain time-out period an error will be detected and the Upper Layer Protocol will initiate recovery actions. Two of the most common time-out values are Error_Detect time-out value (E_D_TOV) and Resource Allocation time-out value (R_A_TOV).

E_D_TOV

The E_D_TOV time-out value is the timer for transmission of consecutive data frames and responses at the sequence level. Basically this is a short value and indicates how long a sequence can take to complete. It is a time-out value for communicating between two N_Ports that is negotiated at login time. Typical value for this is 2 seconds.

As can be seen in Figure 296 on page 904, if, for example, sequence 2 was a Class 3 sequence and the correct number of frames were not received at node 1 (frame 2 missing) a time-out period of ED_TOV will be allowed before a missing frame error is flagged. As this is Class 3 which has no Acknowledgements or Rejects it is up to the sequence recipient (Node1 in this case) to flag the error and it can then discard all frames in that sequence. Once all the frames in the

sequence have been discarded it can then update the Upper Layer Protocol and this will notify the sequence initiator who will initiate the recovery actions which can be to either to re-drive the entire sequence or exchange.

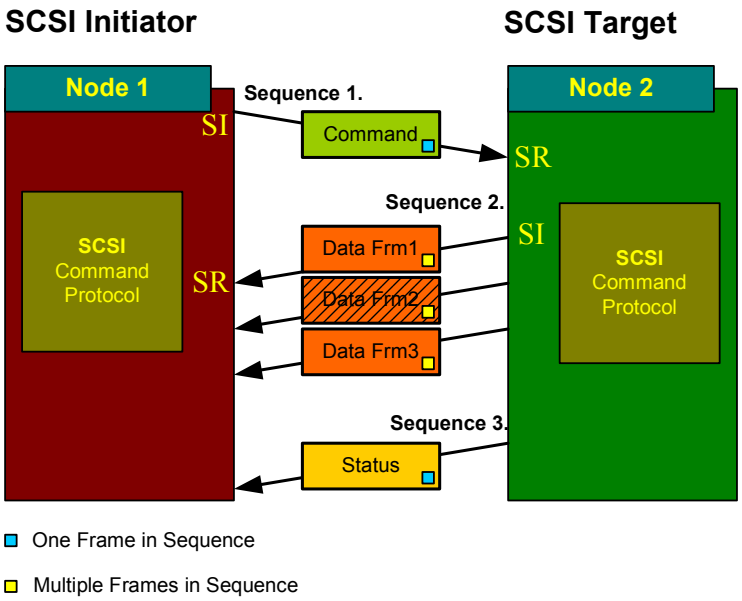


Figure 296 **ED_TOV**

RA_TOV

The R_A_TOV time-out value is a longer time-out value and defines how long before resources associated with an operation that has failed can be used again. For example the time before a sequence initiator shall wait before it can reuse sequence qualifiers such as Sequence ID and sequence count. This value is basically how long an Exchange can stay active until the ULP decides to retry the operation. A typical value for this would be 10 seconds.

In order delivery

Environmental overview

Figure 297 on page 905 shows a three-switch fabric. The host is in the middle of sending data to the storage, as would happen during a write command. The blue boxes represent frames which are all part of the same exchange. The number inside each frame is the sequence count in hex. The majority of the frames are queued while waiting to

be transmitted, as would be the case when the fabric is experiencing some form of congestion.

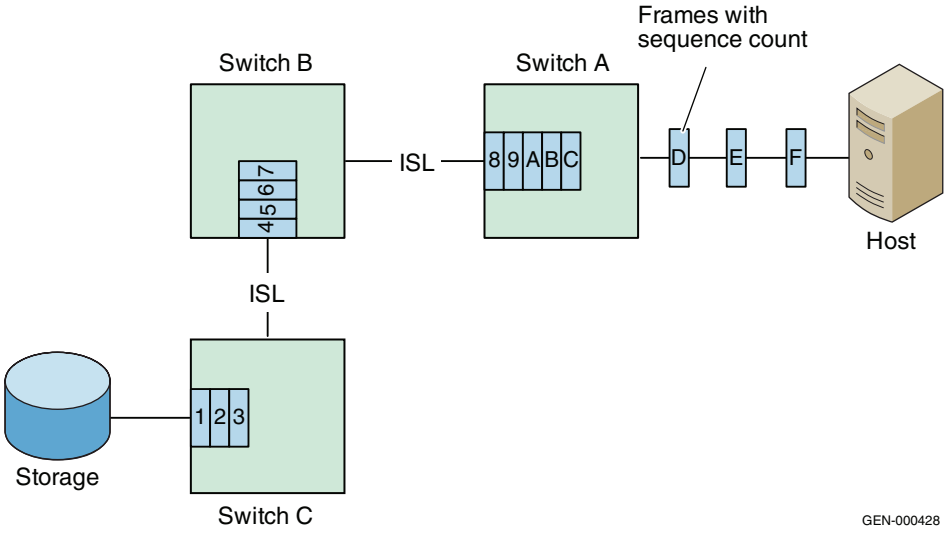


Figure 297 **Host and storage separated by two hops**

Addition of an ISL impact to FSPF cost

As shown in Figure 297, frames have to cross two hops to get from the fabric ingress port on Switch A to the fabric egress port on Switch C. Each hop has a cost associated with it. This cost is used by the switches to calculate the shortest route from one point in a fabric to another. Determining the shortest path is the responsibility of the FSPF (Fibre-shortest-path-first) protocol.

For the sake of this example, assume that each of these ISLs is running at 4 Gb/s and has an FSPF cost of 250. This means that the total cost for the route from switch A to switch C is 500. Since there is only one route from Switch A to Switch C, it is also the shortest path. If another 4 Gb/s ISL were to be added between switch A and C and it had an FSPF cost of 250 as shown in Figure 298 on page 906, a shorter path to switch C would then exist.

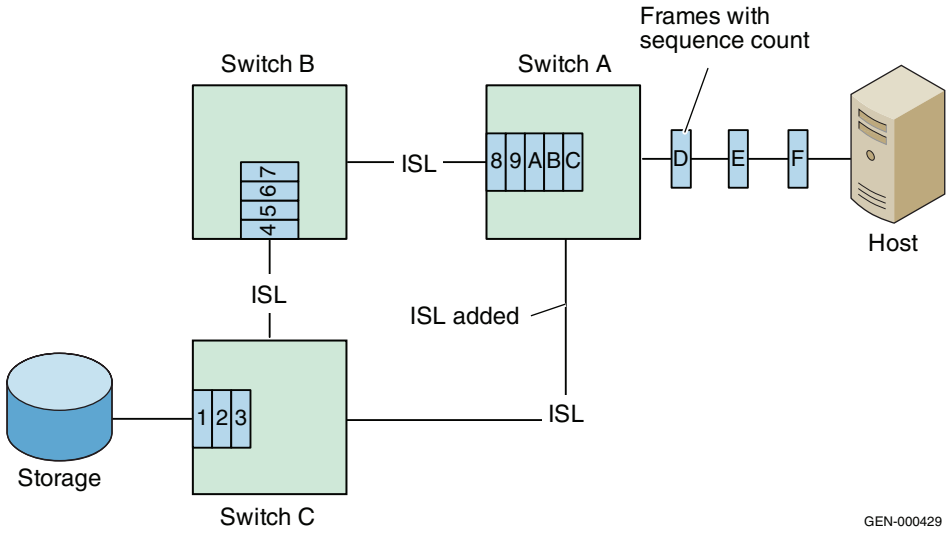


Figure 298 New ISL added. Shorter path introduced

The problem that the new shorter path introduces is that all frames received by the switch after the new ISL is established will take the new shorter path. If in order delivery is not set on switch A, these new frames will be immediately sent down the shorter path as shown in Figure 299 and Figure 300 on page 907.

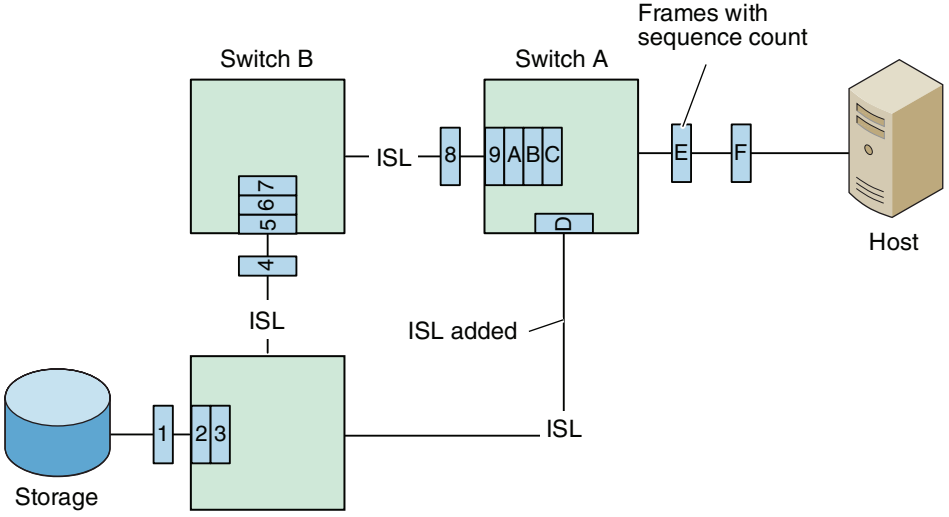


Figure 299 New frames queued for transmit down shorter path

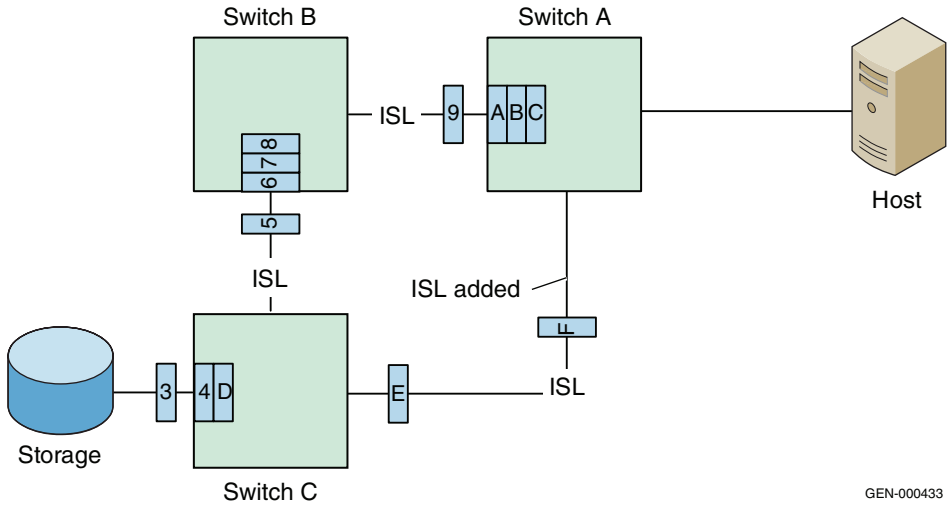


Figure 300 Sequence ID "D" received before Sequcne ID "5"

The problem with the frames being received out of order by the storage is that many hosts, and especially storage ports, do not support the re-ordering of frames. Instead, they will typically abort the exchange and force a retry.

If IOD is set, as shown in Figure 301 on page 908, then new frames received will be held at the new ISL for some period (typically 2 seconds) while the rest of the frames on the older and longer path are allowed to drain.

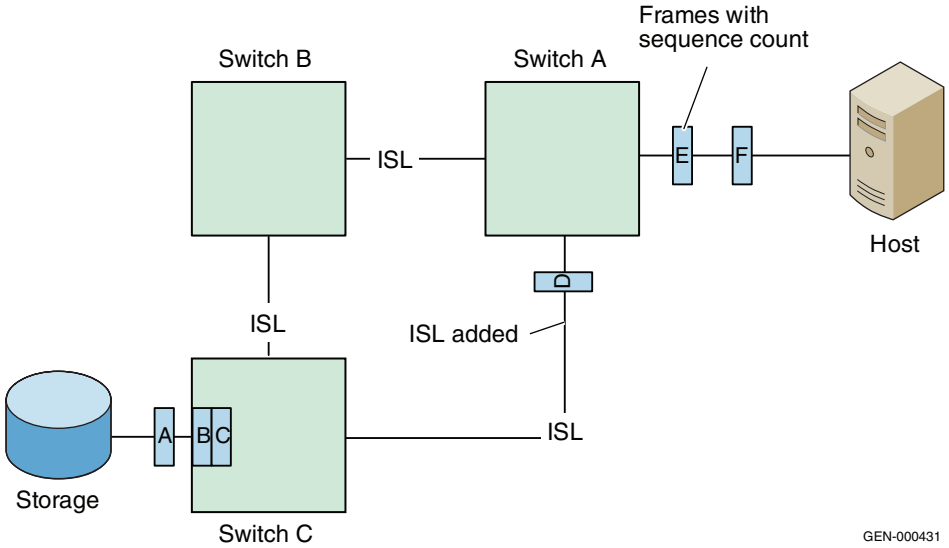


Figure 301 **IOD is set. Sequence ID "D" is held and will be received in proper order**

The advantage to setting IOD is that you are "guaranteed" that all frames will be delivered in order. The disadvantage is that you are guaranteed that all ports in the fabric which will use the new ISL will be paused for the hold time (typically 2 seconds). However, in most cases like the one described above, only a few exchanges will actually have frames arrive out of order and those exchanges are retried and completed well within the time it would have taken to guarantee that all of the frames arrived in order by setting IOD.

Inter-Switch Link (ISL)

An ISL (Inter-Switch Link) allows for two or more Fibre Channel switches to be connected together to form a single ,but larger, fabric. For more information on how ISLs form, refer to "Build Fabric (BF)" on page 826.

Frame services in Fibre Channel

This section provides information on frame services in Fibre Channel:

- "Basic link service" on page 909
- "Extended link service" on page 909

- "Fabric Login (FLOGI)" on page 911
- "N_Port Login (PLOGI)" on page 912
- "Process Login" on page 912
- "Flow control" on page 901

Basic link service

Basic link service frames are link data frames that provide basic control functions within an exchange. These include:

- **No Operation (NOP)**. This basic link service has no data field and no meaning itself but is used to carry control bits between a sequence initiator and recipient.
- **Abort Sequence (ABTS)**. This is used to abort a sequence or an entire exchange.
- **Remove Connection (RMC)**. This is used to request immediate removal of a class 1 connection
- **Basic Accept (BA_ACC)**. This is a response frame to a basic link service command indicating the command has been completed and the applicable information is being returned.
- **Basic Reject (BA_RJT)**. This is a response frame to notify the sender of a basic link service command that the command has been rejected for a certain reason. The reason code is supplied in the first four bytes of the payload.
- **Preempted (PRMT)**. This is used in Class 1 to notify an N_Port that the connection it is using has been removed.

Extended link service

Extended link services are used by an N_Port to provide a specific function at another port. They normally consist of a command (or request) sequence followed by a reply which ends the exchange. Examples of extended link services are Fabric Login (FLOGI), N_Port Login (PLOGI), and Process Login, each discussed in this section. For a full list of the extended link services, see the table in Figure 302 on page 910.

R_CTL Type = 01	Payload Word 0 Byte 0	Name of Extended Link service	Abbr
22 Request	03	N_Port Login	PLOGI
	04	F_Port Login	FLOGI
	05	Logout	LOGO
	06	Abort Exchange	ABTX
	07	Read Connection Status	RSC
	08	Read Exhange Status Block	RES
	09	Read Sequence status Block	RSS
	0A	Requext sequence Initiative	RSI
	0B	Establish Streaming	ESTS
	0C	Estimate Credit	ESTC
	0D	Advise Credit	ADVC
	0E	Read Time-Out Value	RTV
	0F	Read Link Status	RLS
	10	Echo	ECHO
	11	Test (FCAL initialise commands)	Test
	12	Reinstate Recovery Qualifier	RRQ
	20	Process Login	PRLI
	21	Process Logout	PRLO
	22	State Change Notification	SCN
	23	Test Process Login State	TPLS
	30	Get Alias ID	GAID
	31	Fabric Activated Alias ID	FACT
	32	Fabric Deactivated Alias ID	FDACT
	33	N_Port Activated Alias ID	NACT
	34	N_Port Deactivated Alias ID	NDACT
	40	Quality od Service Request	QoSR
	41	Read Virtual Circuit Status	RVCS
	50	Discover N_Port Service Parm	PDISC
23 Response	01	Link Service Reject	LS_RJT
	02	Accept	ACC

Figure 302 Extended link services

As can be seen in Figure 302, there are quite a number of extended link services provided by Fibre Channel.

This section will expand on the following three services, as these are fundamental in understanding how the login process works:

- Fabric Login (FLOGI)
- N_Port Login (PLOGI)
- Process Login

Fabric Login (FLOGI)

Fabric Login (FLOGI) is used in Fibre Channel by an N_Port to detect the presence of a fabric and, if there is a fabric present, to establish a connection by exchanging certain parameters. The Fabric Login process also ensures the attached N_Port is assigned an address by the fabric. Fabric Login takes place following link initialization and is initiated by the N_Port sending a FLOGI extended link service to the switch well-known address 0xFFFFFE. At this time the N_Port does not have a valid fabric address and uses 0x000000 as the source identifier on the FLOGI frame. The N_Port sends its service parameters (operating parameters it supports) in the FLOGI frame and the *Accept* that is sent back by the fabric contains its service parameters and also a fabric address that will be used by that port as its source identifier.

Payload Word	
0	Command Code 04000000 FLOGI / 02000000 ACC
1 to 4	Common Service Parameters
5 to 6	N_Port Port Name
7 to 8	Node Name
9 to 12	Class 1 Service parameters
13 to 16	Class 2 Service parameters
17 to 20	Class 3 Service parameters
21 to 24	Reserved in FC PH Class 4 Service parameters in FC PH2
25 to 28	Vendor version level

Figure 303 FLOGI and Accept frame payload

Figure 303 shows the FLOGI and Accept frame payloads that will be seen during fabric login. The common service parameters and class specific service parameters are basically the FC-2 capabilities of the port.

A number of responses to the initial FLOGI frame other than Accept are possible. If the port receives F_BSY (fabric is busy, try again later) or F_RJT (the fabric did not like the FLOGI frame, class of service not

supported or invalid S_ID) it may need to try attempt again to send the FLOGI frame.

N_Port Login (PLOGI)

N_Port Login (PLOGI) is used by an N_Port to send its service parameters to another port and to request the service parameters of the receiving port back in the Accept. The payload of the PLOGI request and the response (Accept) are identical except for the command code and is performed after fabric login and prior to any upper layer FC-4 type parameters are exchanged.

Payload Word	
0	Command Code 03000000 PLOGI / 02000000 ACC
1 to 4	Common Service Parameters
5 to 6	N_Port Port Name
7 to 8	Node Name
9 to 12	Class 1 Service parameters
13 to 16	Class 2 Service parameters
17 to 20	Class 3 Service parameters
21 to 24	Reserved in FC PH Class 4 Service parameters in FC PH2
25 to 28	Class 4 Service parameters
26 to 29	Vendor version level

Figure 304 PLOGI and accept payload

Figure 304 shows the structure of the payload of a PLOGI frame and the Accept. As mentioned in "Fabric Login (FLOGI)" on page 911, it is possible to have responses other than the Accept. The destination port could be busy in a fabric in which case a P_BSY response would be sent back from the N_Port or also the PLOGI could be rejected with a P_RJT if the class of service being requested was not supported by the receiving port.

Process Login

The Process Login (PRLI) is required to establish communication between two FC-4 layer processes (in our case SCSI) existing at two different N_Ports. The Process Login provides a way of allowing these two ports exchange SCSI information known as *service page* information. The Process Login is normally sent from initiator to target to establish SCSI FCP operating features between both. Again, the mechanism used is for the initiator to send a PRLI request frame and the corresponding Accept contains the service page information on the target.

Payload Word			
0	Cmd code 20 / 02	Page Length =0x10	Total Payload length
1 to 4	Service Parameter page		

Figure 305 **Process Login and Accept frame payload**

Figure 305 shows the payload of the Process Login and the Accept containing the service parameters which are in word 1 to 4 of the payload.

Frame structure in Fibre Channel

All information transferred in Fibre Channel is packaged into frames. The structure of the frame is similar to packets used in networking and consists of a frame header, data field and CRC and the frame beginning is marked with a Start of Frame delimiter and the frame end is marked with an End of Frame delimiter, shown in Figure 306.

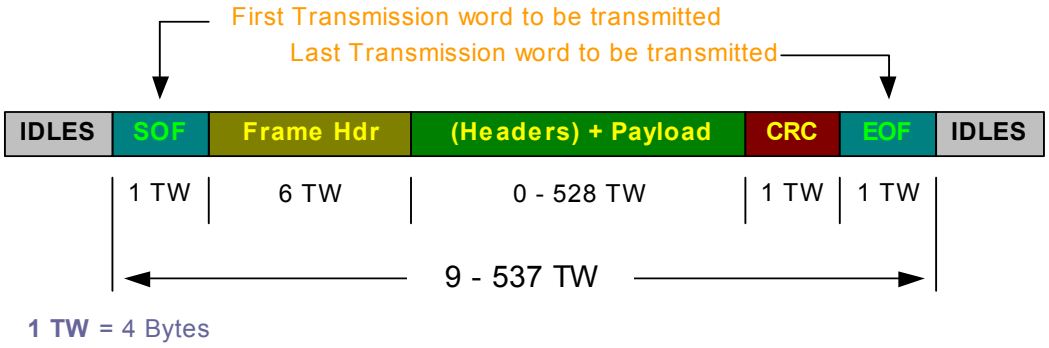


Figure 306 **Frame format**

The following are further discussed in this section:

- "Frame types", next
- "Start-of-frame delimiter" on page 914
- "Frame header" on page 915
- "Data/Payload" on page 922
- "Frame CRC" on page 923
- "End-of-frame delimiter" on page 923

Frame types

There are two different frame types in Fibre Channel each of which is identified by certain bits being set in the Routing Control (R_CTL)

field in the frame header. They are Frame Type FT-0 and FT-1, discussed next.

Frame type FT-0

Frame Type FT-0 is defined as a link control frame with a data field length of zero bytes. They are used for link control functions such as Acknowledgements, Busy and Rejects which are used in Class 1 and Class 2 frame transmission. The R_CTL bits for this type of frame would be R_CTL= 0xCn where n=0 to 7.

Frame type FT-1

Frame Type FT-1 is defined as a data frame and the data field may be any length from 0 to 2112 bytes. Examples of data frames could be frames containing data from a SCSI operation such as a read or write and also frames containing extended link data associated with the login process. Figure 307 shows a table containing information on the various different frame types.

Type	Info category	R_CTL value	Link Service	Notes
Link Frame No Payload No Ack FT-0	Acknowledge (ACK)	C0-C1	Ack_0,Ack_1, Ack_N	N is the number of frames to be acknowldged
	Link Response	C2-C6	Busy F_Busy P_Busy Reject F_RJT P_RJT	Fabric unable to deliver frame Frame has been rejected
	Link Command	C7	LCR	
Data Frame FT-1	FC-4 device data	00 - 07	SCSI,ESCON etc	Data Transfer
	FC-4 Video data	40 - 47	Reserved	
	Basic Link Service	80 - 86	Basic Link Services	
	Link Data	22 - 23	Extended Link Services Plogi,Flogi,Logo,	

Figure 307 **Frame types**

Start-of-frame delimiter

The start of frame delimiter is an ordered set that identifies the beginning of a frame. As mentioned in "Ordered sets" on page 769, additional information is also provided by the start of frame delimiter, such as the class of service associated with the frame, as follows:

- When the connection should be established for classes of service that establish a connection between ports.
- Whether the frame is the first frame of a sequence of frames and whether the frame is initiating the sequence or is a normal frame with a sequence.

Frame header

The Frame header immediately follows the start of frame delimiter and is a 24 byte (6 word) structure which contains control and addressing information. It is divided into sections (Figure 308).

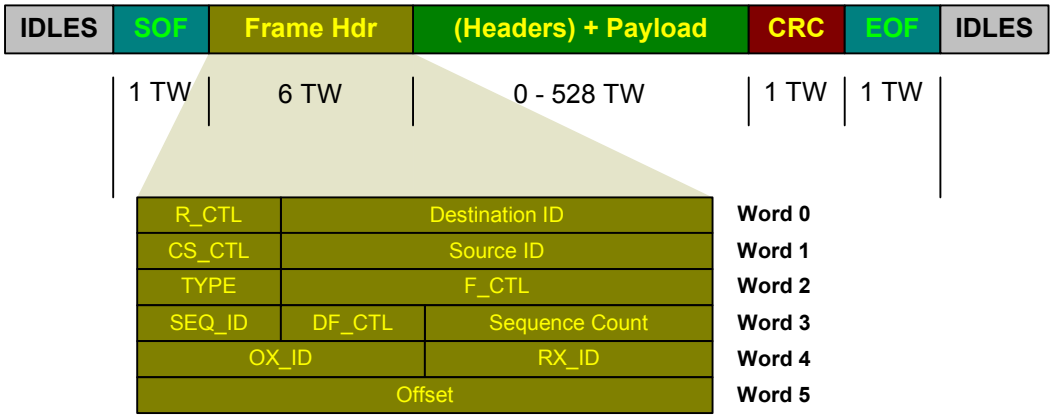


Figure 308 **Frame header**

The following sections shown in the frame header in Figure 308 are further described in this section:

- "Routing Control (R_CTL)", next
- "Destination ID (D_ID)" on page 916
- "Source ID (S_ID)" on page 916
- "Class Specific Control (CS_CTL)" on page 917
- "Frame Type (TYPE)" on page 917
- "Frame Control field (F_CTL)" on page 918
- "Sequence ID (SEQ_ID)" on page 920
- "Data Field Control (DF_CTL)" on page 920
- "Sequence Count (SEQ_CNT)" on page 920
- "Originator Exchange ID (OX_ID)" on page 921
- "Responder Exchange ID (RX_ID)" on page 921
- "Offset / parameter field" on page 922

Routing Control (R_CTL)

The R_CTL field (Word 0 bits 31 - 24) consists of two fields, the Routing bits and the Information field:

- The **Routing bits** (Bits 31 to 28) identify the routing of the frame (how it will be processed) and their general usage. For example if the R_CTL value is 0x0n where n=0 to 7 then you know the frame contains FC-4 device data.
- The **Information field** (Bits 27 to 24) provides further definition for that category of frame. The interpretation is dependant on the Routing bits. For example if the R_CTL value id 0x07 then this decodes to FC_4 device data and Command status. This indicates the frame payload contains command status information which in our case this would be SCSI status. For further information on the R_CTL fields and meanings please refer to Figure 323 on page 924.

Destination ID (D_ID)

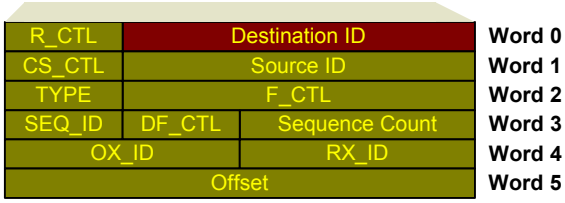


Figure 309 **Destination ID**

This is a three byte field that contains the destination address for the frame. Every N_Port needs to have an address before it can receive frames. This address is acquired during the login process which is different for each topology.

- In Arbitrated Loop the address is acquired during the Loop Initialization stage. Only the first byte of the address is used in this topology and the other two bytes are filled with zero's.
- In Fabric Switch the address is acquired during Fabric Login where the address is assigned by the fabric and all three bytes are used.
- In Point-to Point topology the address is acquired during the N_Port login process.

Source ID (S_ID)

This field again is a three byte value containing the address of the source port of a frame. Whenever a port transmits a frame it places it

address in the S_ID field and the destinations port address in the D_ID field of the frame header.

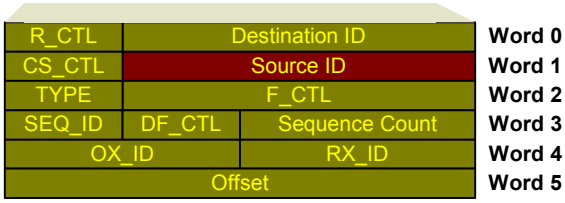


Figure 310 **Source ID**

Class Specific Control (CS_CTL)

This one byte field is only used in Class 1 and Class 4 connections for specifying particular link speeds and link operations. This field is reserved for Class 2 and Class 3 frames.

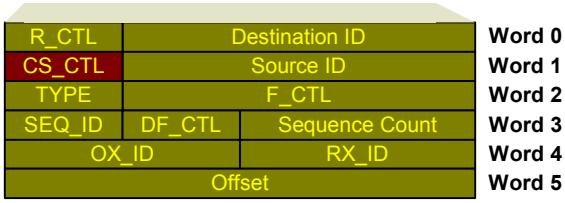


Figure 311 **Class specific control**

Frame Type (TYPE)

The Type field is a one byte field that is used to identify the frame depending on whether the frame is a link control frame (FT-0) or a data frame (FT-1).

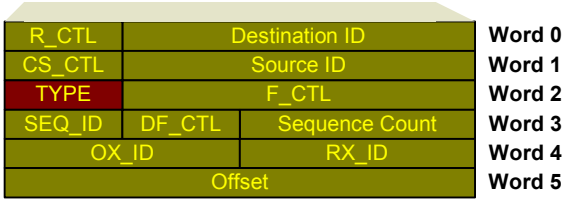


Figure 312 **Frame type**

If the frame is a link control frame (FT-0) this field is reserved except where the frame is a fabric busy (F_BSY) sent in response to a link control frame. The Type field in this case is used to indicate a reason for the busy and reconstruct the original link control frame for retransmission.

When the frame is a data frame (FT-1) the type field is used to identify the protocol being carried by the frame. For example, a type field of “08” indicates SCSI FCP.

For more information on these frame types, refer to “Frame types” on page 913.

Frame Control field (F_CTL)

This is a three byte field that contains control information relating to the frame content. This portion of the frame header is used to identify Exchange control, Sequence control, Acknowledgement policy, and Abort Sequence actions.

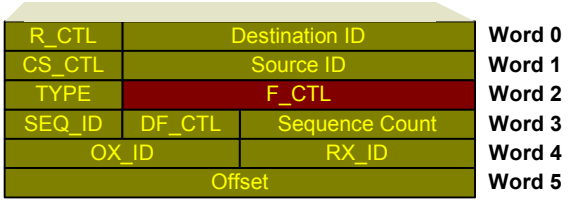


Figure 313 **Frame Control field**

A breakdown of all the bits in the F_CTL field is shown in Figure 314.

F_CTL Bits	Control field	Description
23	Exchange_Context	0 = Originator of Exchange 1 = Responder of Exchange
22	Sequence_Context	0 = Sequence Initiator 1 = Sequence Recipient
21	First_Sequence	0 = Not the first sequence of the exchange 1 = First sequence of the exchange
20	Last_Sequence	0 = Not the last sequence of the exchange 1 = Last sequence of the exchange
19	End_Sequence	0 = Not the last data frame of the sequence 1 = Last data frameof the sequence
18	End_Connection Class 1	0 = Connection active 1 = End class 1 connection pending
	Deactivate_Class_4_Circuit	0 = Class 4 circuit active 1 = Deactivate class 4 circuit

Figure 314 **Frame Control (F_CTL) field**

F_CTL Bits	Control field	Description
17	Chained_Sequence Reserved	0 = No Chained_Sequence 1 = Chained-Sequence active
16	Sequence_Initiative	0 = Sequence Initiative held 1 = Sequence Initiative transferred
15	X_ID reassigned	0 = X_ID assignment retained 1 = X_ID reassigned
14	Invalid X_ID	0 = X_ID assignment retained 1 = Invalid X_ID
13 ,12	Ack_Form	00 = No assistance provided 01 = Ack_1 required 10 = Ack_n required 11 = Ack_0 required
11	Data Compression (reserved)	0 = Uncompressed frame payload 1 = Compressed frame payload
10	Data Encryption (reserved)	0 = Unencrypted frame payload 1 = Encrypted frame payload
9	Retransmitted_Sequence	0 = Original sequence transmission 1 = Sequence retransmission
8	Unidirectional Transmit Class1	0 = Bidirectional transmission 1 = Unidirectional transmission
	Remove Class 4 circuit (Class 4)	0 = Retain or deactivate class 4 circuit 1 = Remove class 4 circuit
7,6	Continue Sequence Condition	Last data frame - sequence initiator 00 = next sequence to follow -no information 01 = next sequence to follow - immediately 10 = next sequence to follow soon 11 = next sequence to follow delayed
5,4	Abort Sequence Condition	In an ACK or RJT frame from sequence recep 00 = Continue sequence 01 = Abort sequence ,perform ABTS 10 = Stop sequence 11 = Immediate sequence retransmission requ

Figure 314 (continued) **Frame Control (F_CTL) field**

Sequence ID (SEQ_ID)

The Sequence ID field is a one byte field used to track all frames in a particular sequence between two ports. All frames in the sequence will have the same SEQ_ID value in the frame header. This value is assigned by the sequence initiator and must be unique between the two ports while the sequence is open.

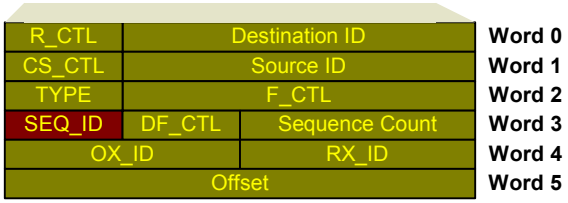


Figure 315 **Sequence ID**

Data Field Control (DF_CTL)

The Data Field Control is a one byte field that defines whether optional headers will exist at the beginning of the data field of the frame. These optional headers are used by some applications and protocols but are not present in all frames. They basically are a means of extending the frame header into the data field for applications that require additional header information. For further information, please see Figure 323 on page 924.

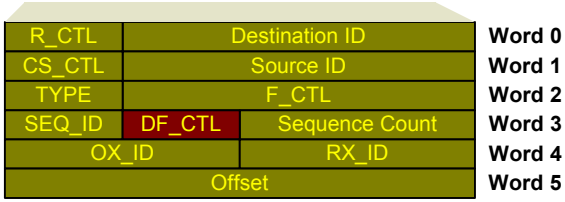


Figure 316 **Data Field Control**

Sequence Count (SEQ_CNT)

The Sequence Count field is a two byte field that indicates the sequential order of a frame within a sequence. The SEQ_CNT value is included in every frame header and is incremented by one from the previous frame in the sequence.

This field is used in determining the order frames are received, that all frames in a sequence have been received and to detect lost or missing frames.

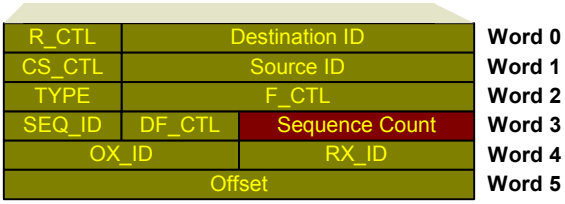


Figure 317 **Sequence Count**

Originator Exchange ID (OX_ID)

The Originator Exchange ID is a two byte field used to identify all frames that are part of an Exchange. Once an OX_ID is assigned to an exchange it is used in every frame of the exchange whether the frame is sent by either the originator or the responder.

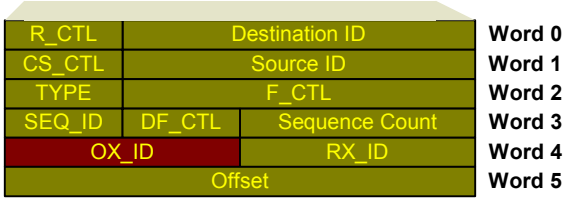


Figure 318 **Originiator Exchange ID**

Responder Exchange ID (RX_ID)

The Responder Exchange ID is also a two byte field used to identify frames that are part of a specific Exchange. This value is assigned by the Exchange responder and is put in every frame sent by the responder which in most cases will be the Symmetrix FA and is also present in every frame echoed back by the exchange originator.

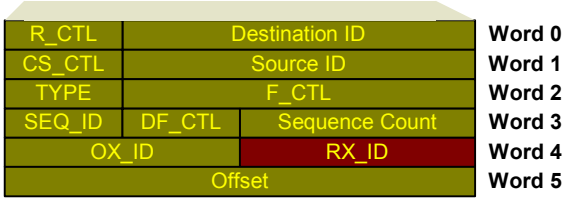


Figure 319 **Responder Exchange ID**

Offset/parameter field

This four byte field in the frame header is used for different things dependant on the frame type.

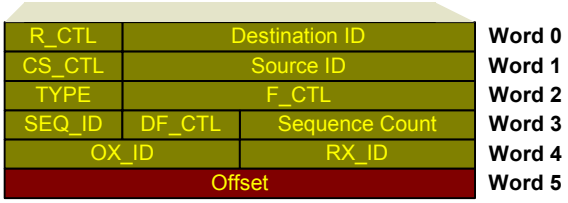


Figure 320 **Offset/parameter field**

If the frame is a Link Control frame such as an Acknowledge (ACK) the parameter field indicates the number of frames to be acknowledged. This field is also used to provide information for Port Busy and Port and Fabric rejects.

If the frame is a data frame then the parameter field contains a value indicating the offset of the first byte of data in the payload relative to the ULP defined buffer for that I/O operation. The existence of a relative offset is indicated by bit 3 in the F_CTL field. The relative offset is used in reassembly of sequences at the receiver end as it provides an offset as to where the payload of each frame needs to be placed in the ULP buffer to reassemble sequences/information units.

Data/Payload

This part of the Fibre Channel frame is a variable length field and can contain optional header information (as mentioned in the DF_CTL field), SCSI data and in certain cases fill words. The maximum size of the data field a frame is 2112 bytes. The contents of the frame payload is determined by the FC-4 layer. In our case this will be SCSI information such as command, data or status information being transported between a source (HBA) and destination N_Port (Symmetrix FA).

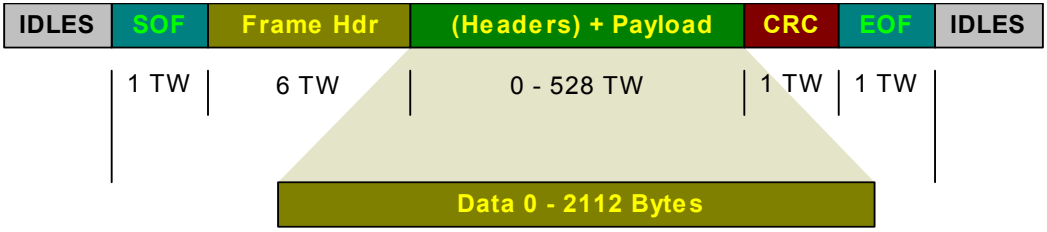


Figure 321 **Payload**

Frame CRC

This four byte value is used to ensure that the contents of the frame header and data field information in the frame have been received correctly. The CRC field follows the data field and is before the EOF delimiter. This value is calculated prior to encoding for transmission and after decoding at reception. SOF delimiters and EOF delimiters are not used in the calculation of the CRC value.

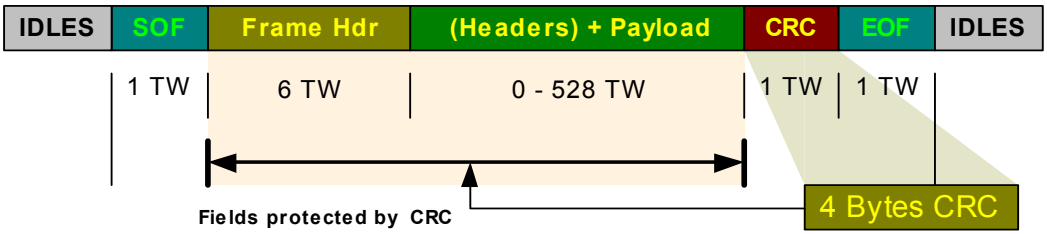


Figure 322 **CRC protection**

End-of-frame delimiter

The end-of-frame delimiter is an ordered set that identifies the end of a frame to the receiving port. It also removes the connection for the classes of service that establish a connection such as class 1 and also signifies whether the frame is the last frame in a sequence or a normal frame within a sequence.

Figure 323 on page 924 further defines the information contained in frames.

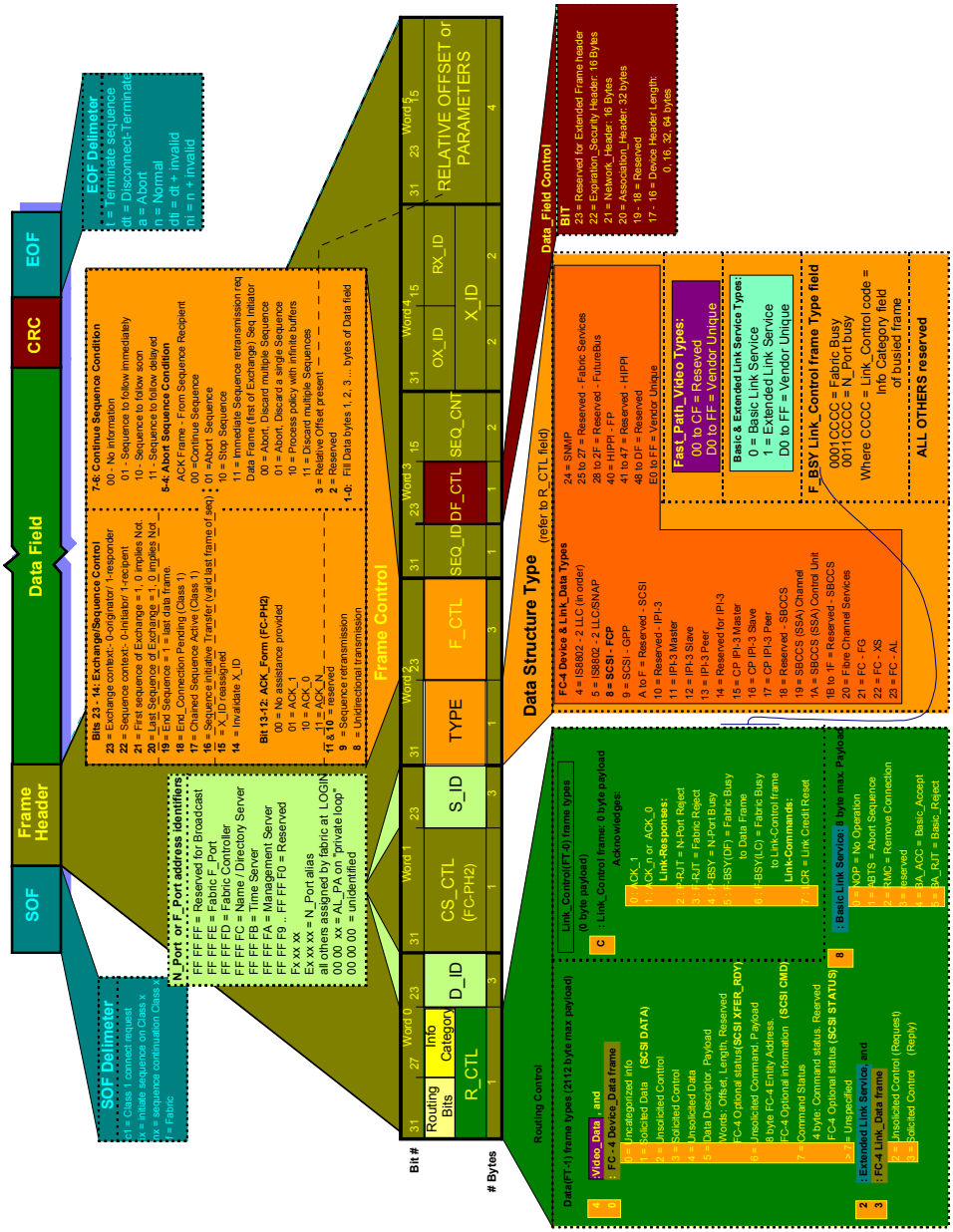


Figure 323 Frame definition

Class of Service (C.O.S.)

Fibre Channel defines different delivery options for frame transmission and these are known as a particular Class of Service. Class 1, 2 and 3 are the different services supported by Fibre Channel and are basically defined with regards to connection, in-order delivery, confirmation of delivery and which type of flow control is used. These classes of service are independent of the topology used.

Class 1

Class 1 is a connection-oriented service that provides a dedicated connection between two ports allowing them to use the maximum available bandwidth. Class 1 provides confirmation of delivery and notification of non- delivery. As shown in Figure 324, a connection is established at the start of the sequence with the SOF(C1) delimiter and the return of the ACK from the recipient. BB_Credit is only used on the request frame and the ACK. Once the connection is established, frame flow is controlled with End-to-End credit.

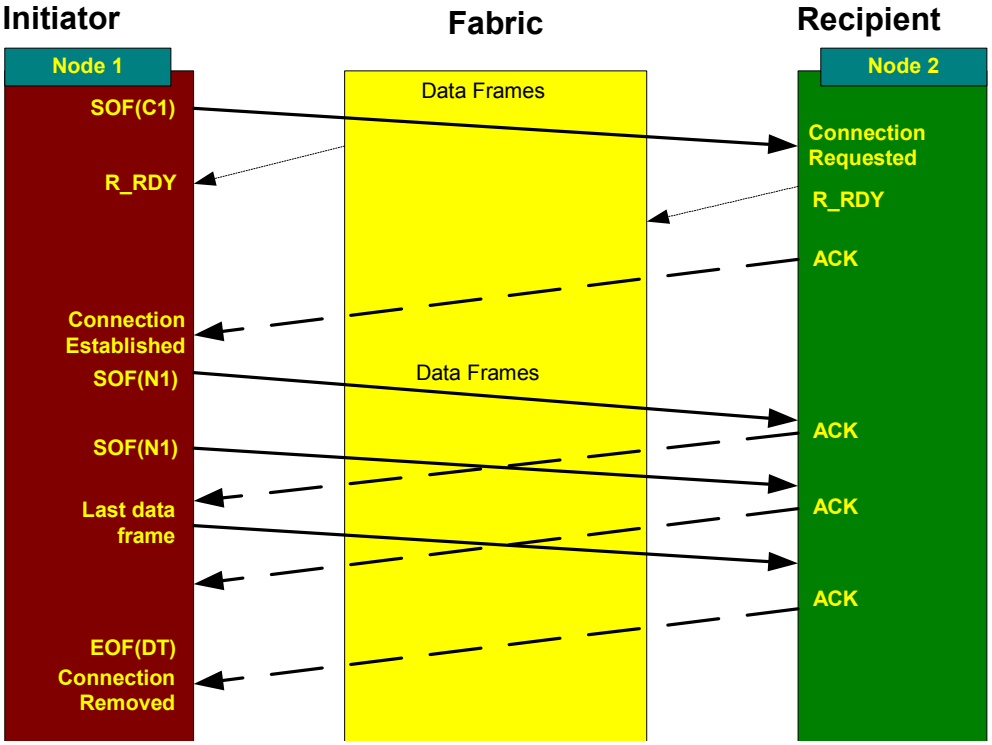


Figure 324 **Class 1 Dedicated connection**

When the connection is no longer needed the port which is sequence initiator at the time sets the End_Connection bit in the F_CTL of the last data frame and the recipient replies with an ACK that has the End of frame disconnect terminate EOF(DT) delimiter set.

This removes the Class 1 connection. Class 1 thus provides guaranteed bandwidth, in-order delivery and confirmation of delivery.

Class 2

Class 2 is a connectionless class of service that provides confirmation of delivery and non-delivery of frames. No connection is established between the source and destination ports and it is up to the fabric to deliver the frames, possibly over different paths to the recipient. Frames will be transmitted in order but depending on the fabric may arrive at the recipient out of order. Frames may be sent to different destinations one after the other and it is possible to be transmitting a frame to one destination while receiving a frame from another which means Class 2 multiplexes on a frame by frame basis.

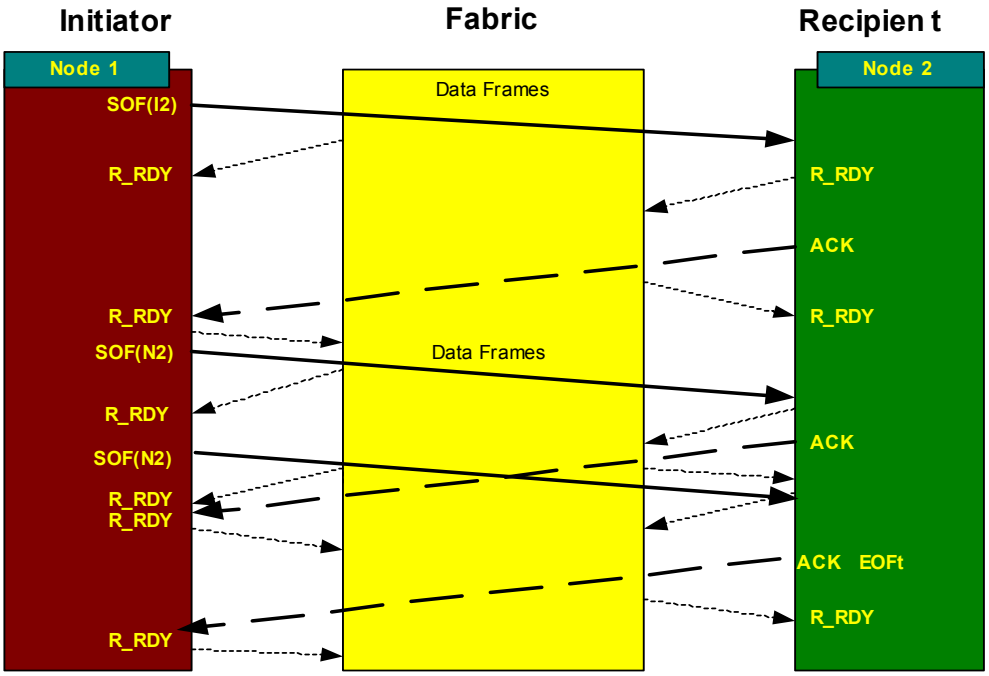


Figure 325 Class 2 operation

Figure 325 shows the Buffer-to-Buffer (R_RDY) and End-to-End (ACK) flow control in Class 2 frame flow. The sequence is started with a Start of Frame Initiate Class 2 SOF (I2) delimiter and there is no disconnect as there is in Class 1 but the last frame in the sequence uses a End of Frame terminate (EOFt) delimiter. The initiator sets the End_Sequence in the F_CTL field of the frame header of the last data frame and the recipient sets the End_Sequence and the EOFt delimiter on the last ACK of the sequence.

Class 3

Class 3 is again a connectionless class of service but with no confirmation of delivery or non-delivery of frames. No connection is established and, like Class 2 frames, Class 3 frames may arrive out of order. The biggest difference is that the delivery of frames is unacknowledged in that the destination port does not send any link control frame (ACK) on receipt of valid data frames. The flow control used in Class 3 is Buffer-to-Buffer which uses the R_RDY primitive signal.

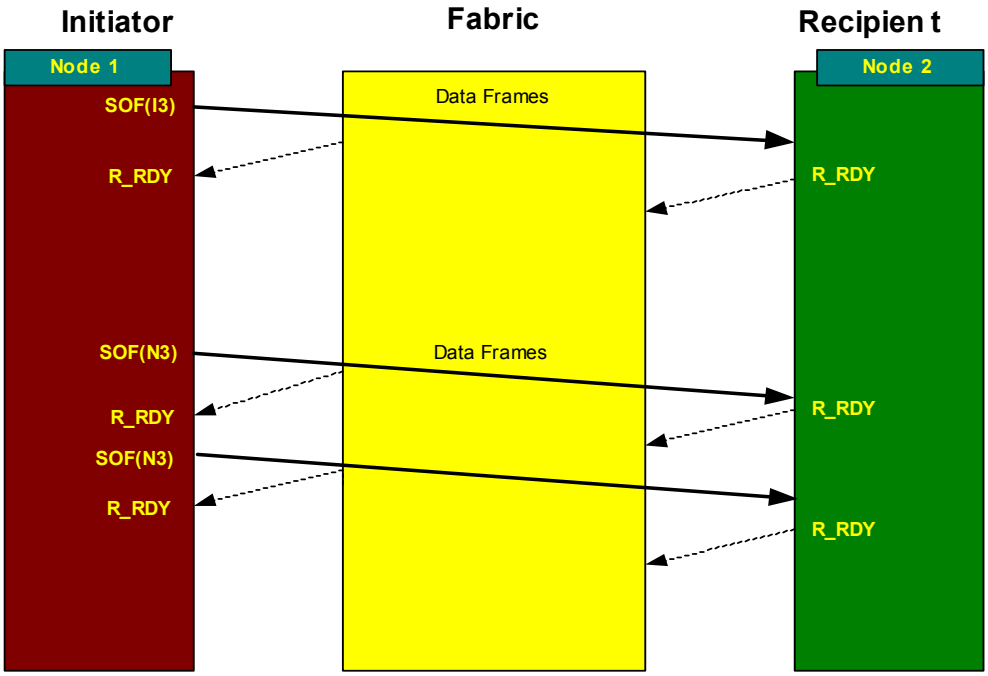


Figure 326 **Class 3 Flow control**

The fact that there is no confirmation of delivery of frames in Class 3 Fibre Channel means error recovery must be handled at the Upper Layer Protocol. The confirmation is provided by the protocol being transported which in our case will be SCSI.

Class 3 is the class of service used by Symmetrix and as SCSI itself is an acknowledged protocol it can handle error detection and recovery.

Figure 327 shows a class of service summary.

	Class 1	Class 2	Class 3
Connection type	Dedicated Connection	Connectionless	Connectionless
Flow Control	End-to-End Credit	End-to-End Credit Buffer -to-Buffer	Buffer-to-Buffer credit
Frame delivery	In order delivery	Order not guaranteed	Order not guaranteed
Frame acknowledged	Acknowledged	Acknowledged	Unacknowledged
Route used	Circuit established / removed	Frame by Frame routing	Frame by Frame routing
Multiplexing	No	Yes	Yes
Bandwidth	Poor bandwidth utilization	High bandwidth utilization	High bandwidth utilization

Figure 327 Class of service summary

Buffer-to-buffer credit (BB_Credit)

Fibre Channel uses the BB_Credit (buffer-to-buffer credit) mechanism for hardware-based flow control. This means that a port has the ability to pace the frame flow into its processing buffers. This mechanism eliminates the need of switching hardware to discard frames due to high congestion. EMC testing has shown this mechanism to be extremely effective in its speed and robustness.

For more information on BB_Credit, refer to "Data buffering and flow control" on page 451. For information on buffer-to-buffer flow control, refer to "Buffer-to-Buffer flow control" on page 881.

Zoning

Zoning is the process of grouping initiator and targets into zones. Initiators and targets placed in the same zone are allowed to communicate by the fabric. Zoning also plays a critical role in event isolation, i.e., RSCN (Registered State Change Notifications) distribution. In order for a Fabric to work properly, it *must* be properly zoned. Currently, a few types of zoning are supported, each discussed briefly below:

- World Wide Port Name (WWPN) zoning

  By far, the most common form of zoning is by WWPN. You select the unique 64 bit addresses of the initiator, its target(s), and place them in a common zone. Its advantage is that no matter where you attach the WWPN, as long as it is in the same fabric, it will always be able to discover ,and be discovered, by other ports to which it has been zoned to have access.

- Domain, Port (D,P) zoning

  This is another common form of zoning . You specify an initiator and target by their physical location in the fabric. The main advantage is perceived to be security, and in some cases it is more secure. Refer to Chapter 4, ”Secure SANs,” for more information.

- World Wide Node Name (WWNN) zoning

  Although it has not been used much up to this point, there is potential for it to be used extensively in clustered environments.

Storage

There are many different types of storage available on the market today including tape, disk, and WORM (Write Once Read Many). For each of these types there are numerous ways to implement and present the storage to the end user. A tape, for example, can be presented directly as a tape drive on the SAN or it can be front-ended with a disk-based library, such as CDL or SDL. The same is true for disk-based storage. It can be presented as a physical disk, as is the case with a JBOD (Just a Bunch Of Disks) solution, or as a virtual disk, as is the case with most array-based storage solutions.

This section provides a high-level overview of the major storage concepts, as follows:

- Disk

  A disk is the physical media that data will be stored on. Disks are typcially carved up or concatenated to make logical volumes of the desired size.

- Volume (aka Logical Volume)

  A volume is an entity which can be made available for use on a front-end port at a specific LUN address. It can also be thought of as a virtual disk.

- LUN (Logical Unit Number)

  A LUN is the address at which a volume is presented to the user on a front-end port.

- Front-end port

  A front-end port is an interface, either FC or IP, on which LUNs are presented for customer use.

- Array controller

  An array controller is the collection of hardware and software features used to convert physical disks into addressable volumes. The array controller also provides a cache for temporarily storing data while it is waiting to be written to disk.

- Cache

  Cache is a high bandwidth/low response time temporary data storage area provided by the array controller.

- Back-end port

  A back end port is an interface, either FC or IP, that the array controller uses to push data to the disk.

NPIV

Traditional N_Port initialization works as follows (refer to Figure 328 on page 931):

- N_Port sends FLOGI to Fabric Server (0xFFFFFE) to obtain a valid address.

- Fabric Server responds with an Accept containing the assigned address.
- N_Port sends PLOGI to the Name Server (0xFFFFFC) to start the registration process.
- N_Port sends SCR to the Fabric Controller (0xFFFFFD) to register for state change notifications (RSCNs).

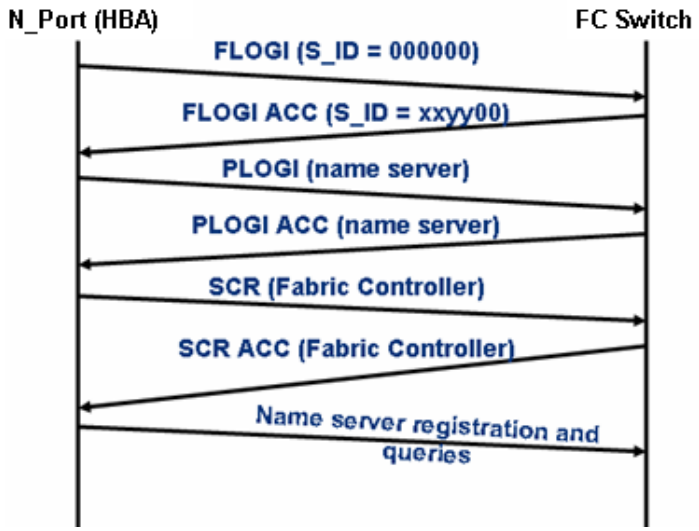


Figure 328 **Traditional NPIV initionalization**

NPIV-capable N_Port initialization works as follows (refer to Figure 329 on page 932):

- N_Port sends FLOGI to Fabric Server (0xFFFFFE) to obtain a valid address.
- Fabric Server responds with Accept with the assigned address.
- N_Port sends FDISC to Fabric Server (0xFFFFFE) to obtain a virtual address.
- Fabric Server responds with Accept with the assigned virtual address N_Port.
- To acquire additional address for NPIV session, repeat FDISC as additional NPIV sessions is needed.
- Each virtual ID has its own WWN.
- Virtual IDs can be zoned as any other normal N_Port.

- Virtual IDs can register and query the name server.

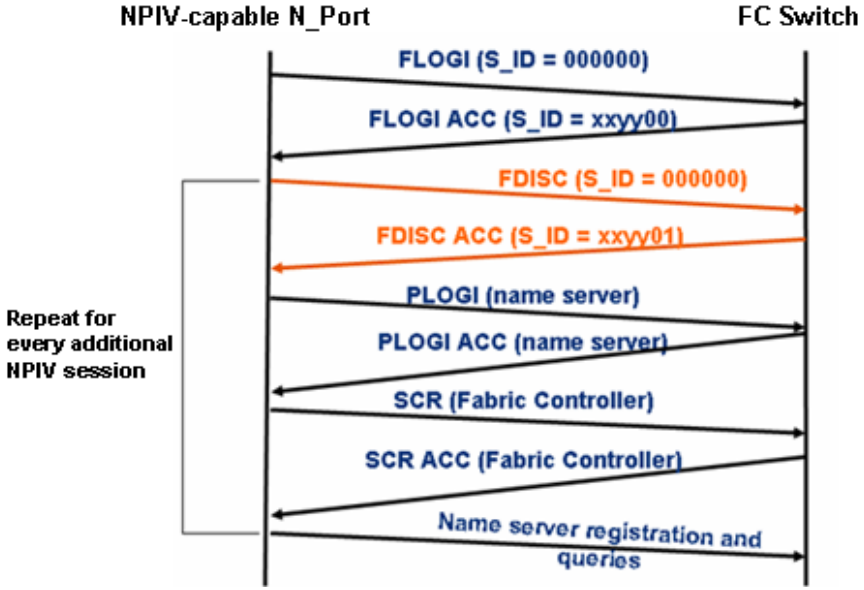


Figure 329 NPIV-capable N_Port initialization

NPIV is useful in environments with Blade Servers, Multi ID devices (tape, encryption) and Server Virtualization, each discussed in more detail in this section.

Blade servers

Blade servers which utilize Fibre Channel Switch Modules present a challenge to SAN scalability and interoperability as each chassis contains two FC switch modules (domains). See Figure 330.

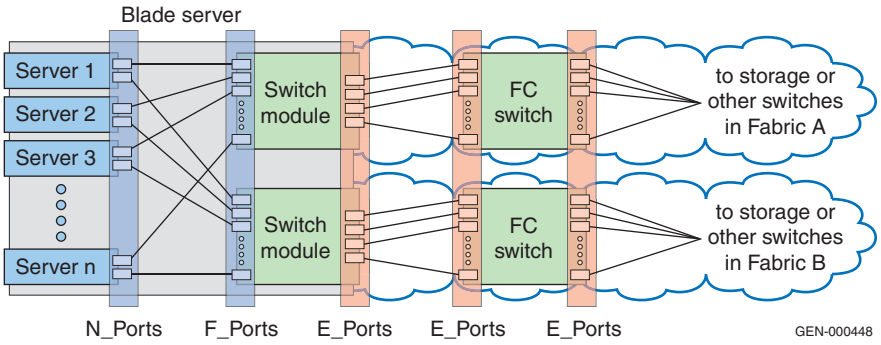


Figure 330 Blade Servers

It is important to stress that each of the FC switch modules embedded in the blade server is a ful-blown FC switch with a name server, Domain ID, etc. Because of this, due to FC switch interoperability concerns, each blade server vendor needs to provide a switch module capable of connecting to a SAN consisting of switches from each of the switch vendors. For example, if a Blade Server were going to be attached to a Brocade fabric, a Brocade switch module would be the best choice to embed in the Blade Server, and not a QLogic module.

A drawback to having two switch modules in each chassis is that two domain IDs are consumed for each chassis. Since most vendors only support between 31 and 50 Domain IDs, this can present a scalability problem. In fact, depending on how many blade servers are added, this can easily constrain the total fabric size to less than 800 ports, which is far short of the minimum number of supported N_Ports in today's fabrics of 2048.

Another consideration is that by embedding the FC switch into the Blade Server, you are also choosing to embed higher risk interoperability points into the server chassis. The E_Ports are a higher risk due to the nature of FC-SW interop (see EMC Knowledgebase article EMC149735 for more details on current FC-SW interop problems).

To eliminate these problems, the switch modules can be replaced with NPIV gateways (Figure 331). This improves fabric scalability and interoperability since each NPIV gateway presents only additional N_Ports to the fabric and not E_Ports or additional domains.

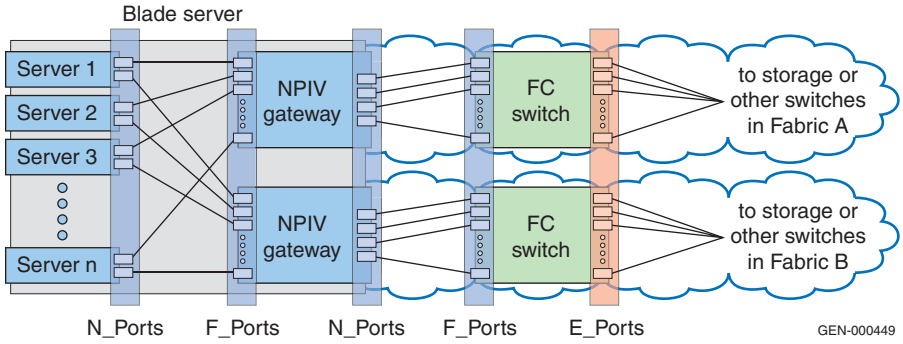


Figure 331 NPIV Gateways

Another advantage is that the higher risk interoperability points are moved out of server chassis. Finally, server blade HBAs use standard drivers and this, in combination with the NPIV gateway's level of transparency, should allow for a higher level of interoperability.

Multi ID devices

Another application for NPIV is as a replacement for the arbitrated loop functionality used by some multi ID devices, such as the CLARiiON Disk Library or Decru DataFort. These devices use public loops to present multiple IDs into the fabric. Outside of the operational inefficiencies of using a loop during normal operation, the largest drawback to loops is that the addition of a new ID (as happens during failover on dual-headed CDL configurations) requires a reset of the loop and is very disruptive to I/O. Another drawback is that it requires loop support on the FC switch.

Figure 332 through Figure 335 on page 936 highlight the difference between failover on CDL when Aribitrated loop is used versus NPIV.

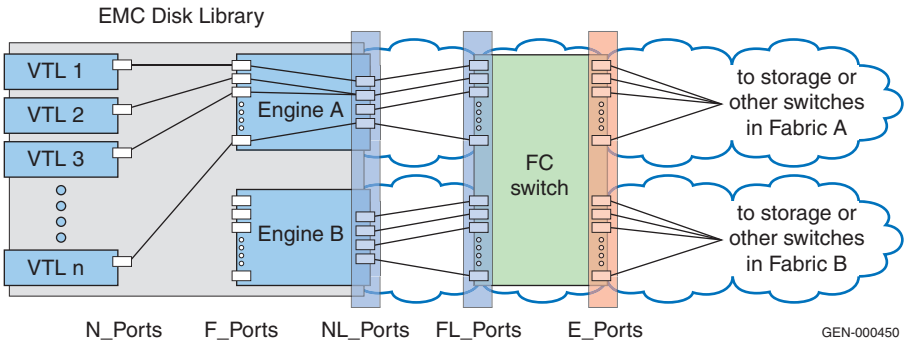


Figure 332 Normal operation with FC-AL

In Figure 333, Engine A fails.

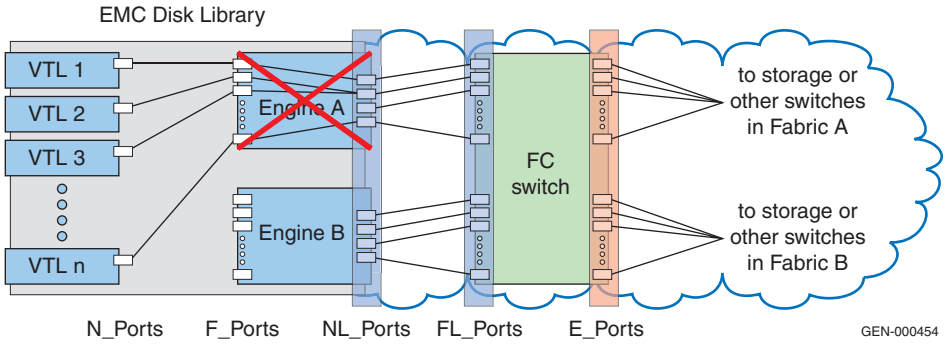


Figure 333 Engine A failure

Failover to Engine B results in LIP generation and disruption to established I/O (Figure 334).

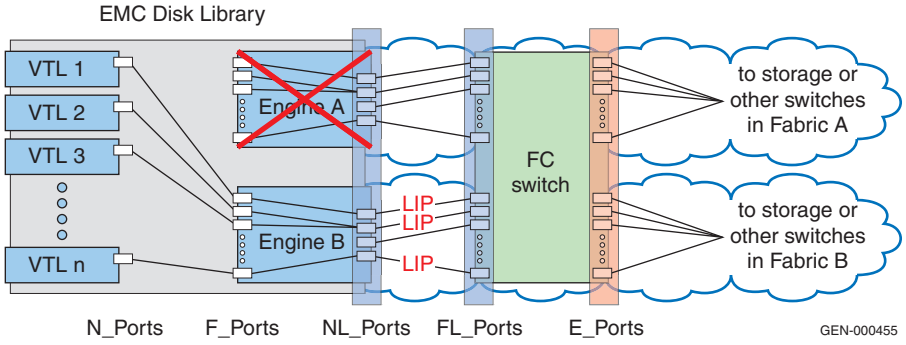


Figure 334 Failover to Engine B

When NPIV is used under the same type of failure, FDISC are sent instead of LIP. The transmission of FDISC does not impact existing I/O (Figure 335).

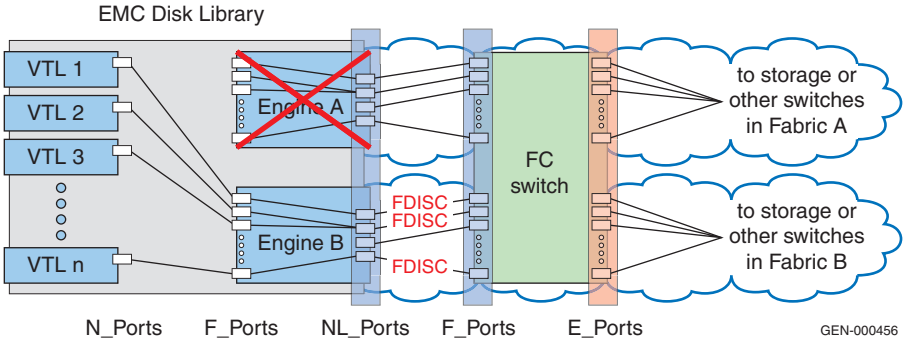


Figure 335 **Using NPIV**

Server virtualization

NPIV can also be used in conjunction with server virtualization. Today this is limited to Mainframe z/Linux environments, but this may eventually be expanded to include support for VMware. The idea is that each virtual machine have its own WWN. This is useful when SAN Management is an entirely separate discipline and organization from Server Management because existing SAN Management tools can be used for storage provisioning.

NPIV challenges

NPIV has the potential to simplify the SAN, but it does have some tradeoffs, the primary one being that there are not many implementations utilizing NPIV today. In spite of the idea having been around for a while, the adoption has been slow until at least early 2007. It is suspected that with the release of NPIV gateways, NPIV's popularity will increase or decrease in proportion to the number of new Blade Server deployments.

Another important consideration is that even though there are no FC-SW interop concerns, this does not mean there are no interoperability concerns. Access Gateway from Brocade, along with similar implementations, still present devices as N_Ports so there are inherent bandwidth sharing concerns that will have an unknown impact.

Finally, once NPIV has been deployed in your environment, you will have eliminated the domain count restriction, but the maximum

number of N_Ports supported in a single fabric (2048-4096) may come soon, so scalability concerns are deferred, but not resolved. To resolve the scalability concerns you will need to deploy Fibre Channel Routers.

Fibre Channel Routing

Fibre Channel Routing is a Fibre Channel facility which allows ports to be shared across separate SANs without merging them. It is defined in FC-IFR (Fibre Channel Inter-Fabric Routing), which is a Fibre Channel standard in *working-draft proposed* status. The purpose of FC-IFR is to reduce the scalability limits imposed by the name server by only sharing certain resources between fabrics. By not collapsing all of the N_Ports into one large name space, Domain ID limits are avoided, as are the current maximum number of N_Port limitations. It provides solutions for configurations with FC-SW interoperability, resource consolidation, distance extension, or scalability concerns, each discussed further in this section. This section will also discuss "Brocade SAN Routing — FCR" on page 944.

As shown in Figure 336 on page 938, routing spans many different topologies and protocols. The simplest concept is FC to FC routing where two Fabrics connect to the same router. A bit more complicated is connectivity over two FC routers which are connected by an IP link or a backbone fabric.

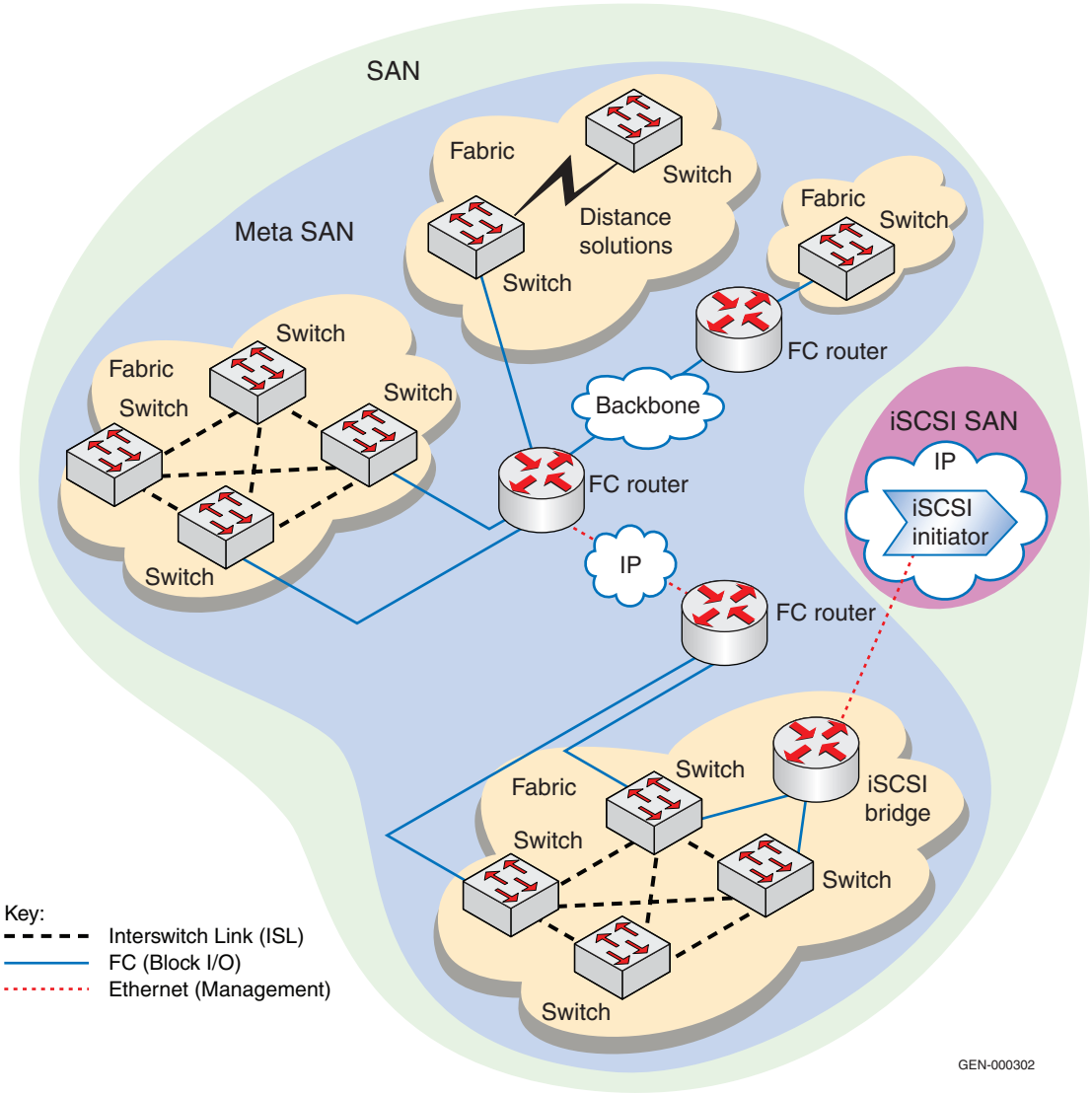


Figure 336 Fibre Channel Routing

Interoperabiliity

Figure 337 shows four different fabrics, none of which are compatible with each other due to Domain ID overlap, different operating modes, and even different firmware revisions.

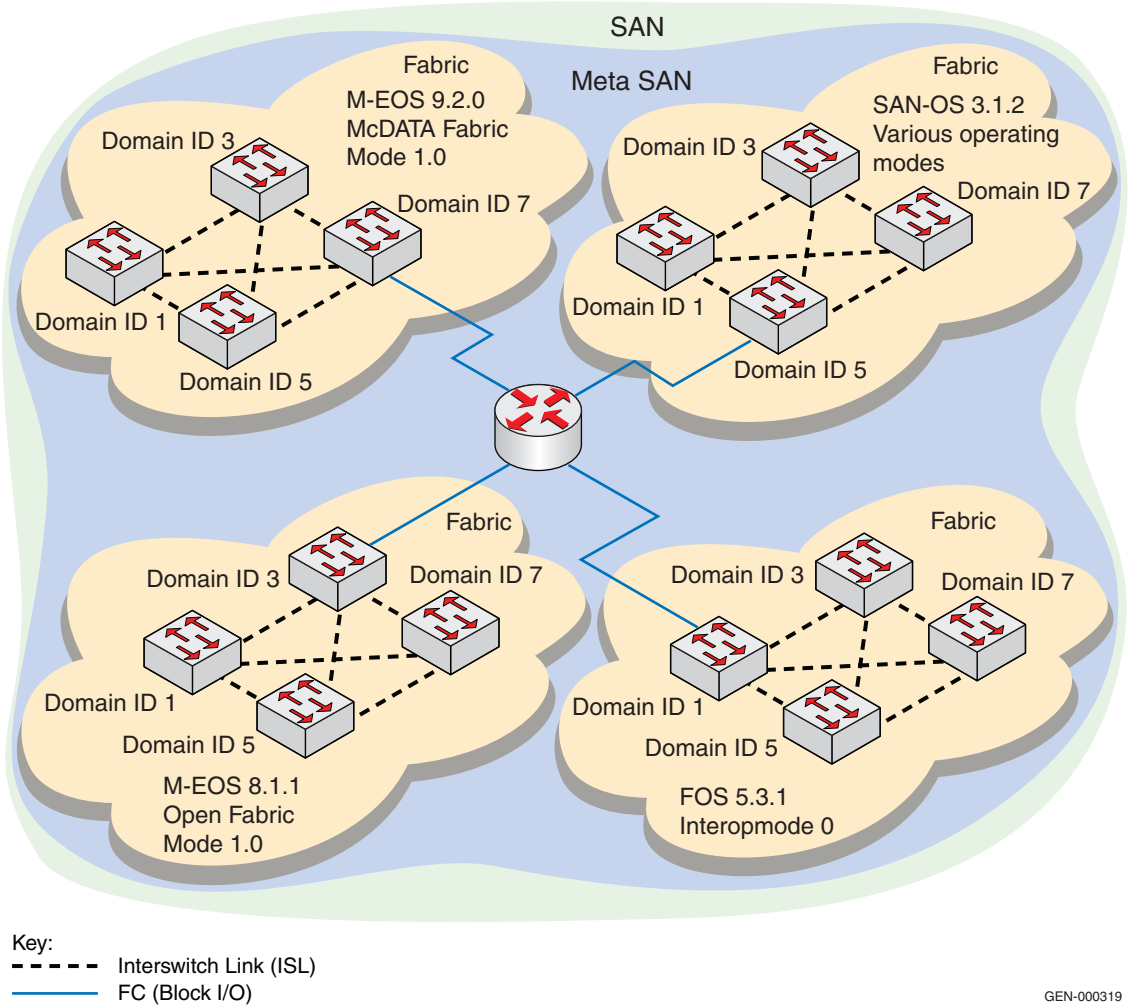


Figure 337 **Interoperarability in Fibre Channel Routing**

Sometimes, due to uptime requirements, it is not possible to take a fabric down to change Domain IDs or operating modes, or to upgrade switch firmware because of interoperability concerns between older, but established, OS / HBA / driver and storage combinations.

It is also possible that you will be required to establish connectivity between devices that reside on each of these fabrics. In this case, you will need to use a router to act as a bridge between the different fabrics.

Resource consolidation

A specific example of why interoperability between different fabrics would be desirable is an environment that needs to deploy tape backup. In Figure 338 on page 941, tape (i.e., CDL) is connected to the McDATA fabric while two of the servers that need to access it are connected to other fabrics with different operating modes and overlapping domain IDs. You may be able to just attach the backup servers and the CDL to the same fabric as long as the storage that needs to be backed up can also be moved to the new fabric. Odds are, though, that this type of a move will *not* be possible in your environment.

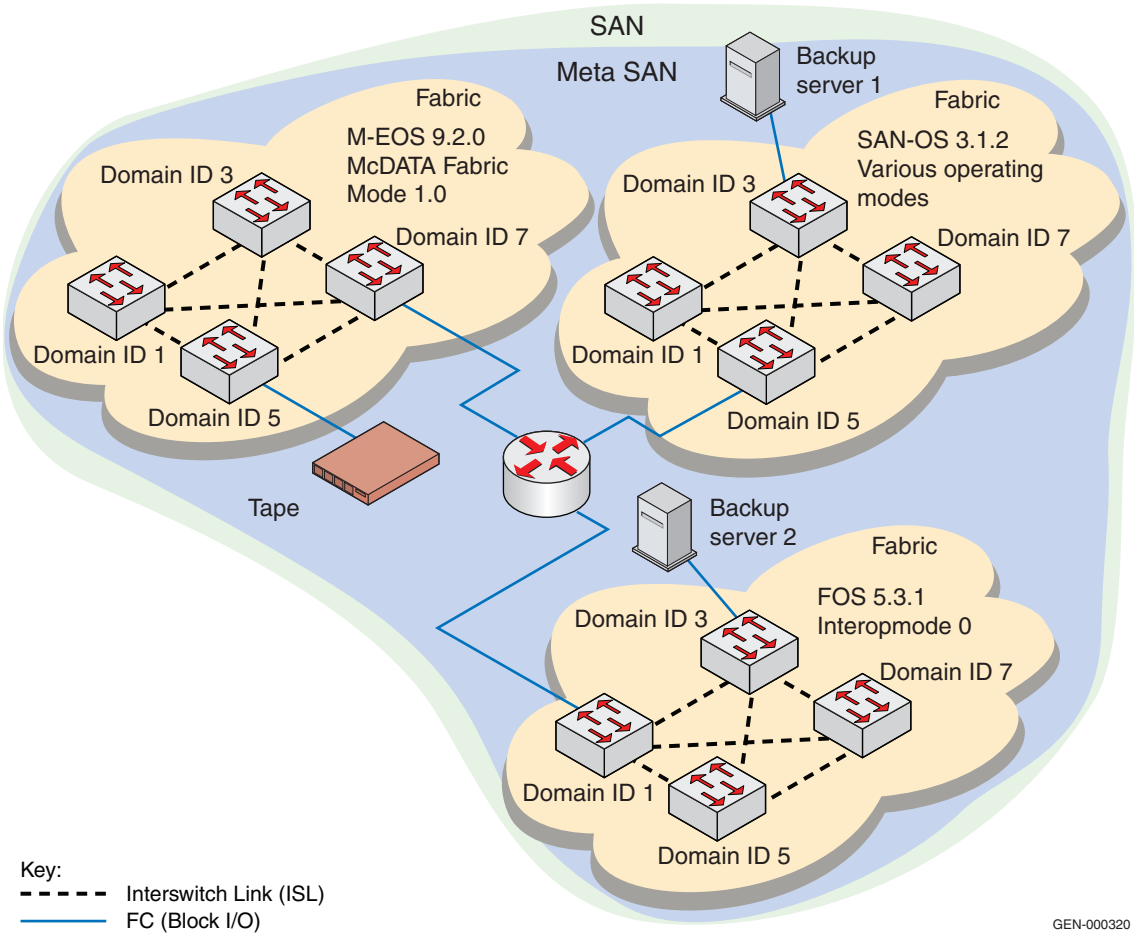


Figure 338 Resource consolidation in Fibre Channel Routing

Distance extension

Another example of where routing is beneficial is when you have a need for a distance extension solution (Figure 339).

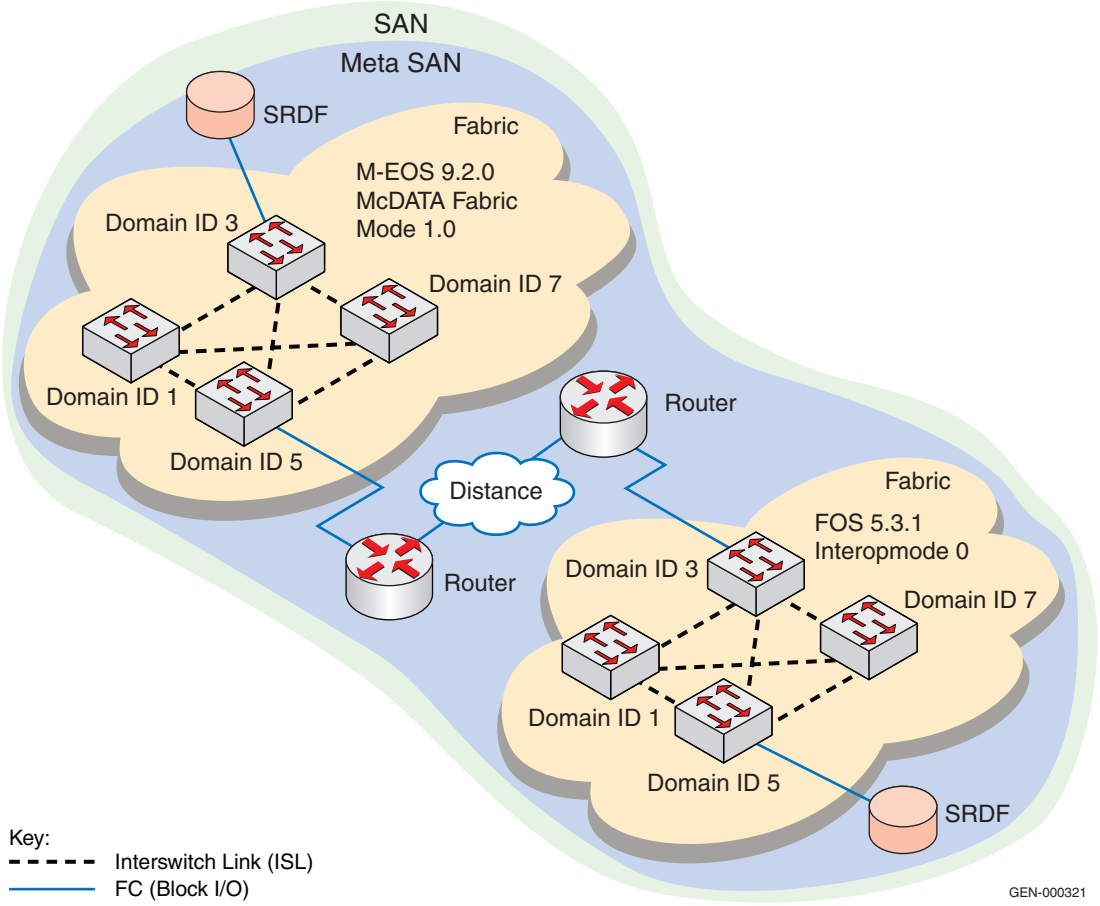


Figure 339 **Distance extension in Fibre Channel Routing**

There are many solutions available, such as FCIP and DWDM, that will enable connectivity over long distances for the purposes of supporting SRDF or MirrorView. The majority of these solutions will simply extend the ISL and allow Fibre Channel to handle buffering and error recovery. When using native Fibre Channel for long distance links, the performance of the link is limited both by the efficiency of the application using the link and the number of buffer-to-buffer credits available. Introducing Fast Write and increasing the buffer count can overcome both of these limitations.

Error recovery, however, cannot be handled so easily. When ISLs are added to or removed from a fabric, especially ones that connect two fabrics over distance, performance throughout the entire fabric can be *negatively* impacted. To minimize the impact that long distance link disruptions have on the fabric, routers can be deployed at each edge of the fabric to ensure that each fabric is isolated from these temporary disruptions.

Scalability

As discussed in "Maximum number of Nx_Ports in the fabric" on page 748, there is a limit to the number of ports that a single fabric can support. With today's firmware revisions, this number is in the range of 2048-4096 ports, depending upon the platform. Figure 340 is intended to illustrate many large fabrics connected by a backbone fabric. Configurations of this type have been successfully deployed by several customers.

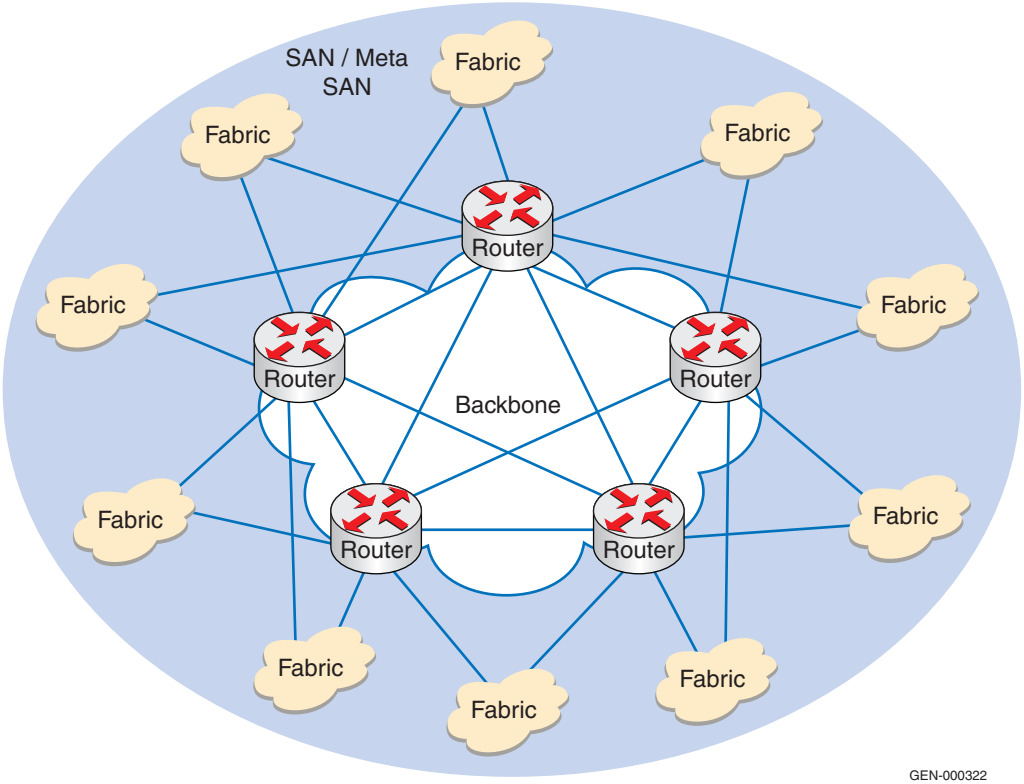


Figure 340 Scalability in Fibre Channel Routing

The advantage to this configuration is that individual devices can be shared between fabrics without running into scalability limits.

Limitations

Consider the following limitations when using routing:

Routing is an excellent solution if you are encountering compatibility issues or experiencing scalability limits. However, mostly due to a lack of good management tools, it adds a great deal of complexity to a fabric. This causes the fabric to be difficult to service and troubleshoot, especially when dealing with end-to-end connectivity problems and Network Address Translation is being used.

Another consideration is manageability. A routed environment is more difficult to manage. Simple tasks, such as zoning, are more difficult to accomplish when each fabric connecting to a router needs to have its zone manually updated. New products are being developed, such as Connectrix Manager which is capable of automatically activating the zone set on the required fabrics, but this product currently only works on the MP-2640M and MP-1620M models.

Although scalability limitations are minimized, connectivity is minimized as well. Routing does not provide any-to-any connectivity. This limitation will be masked somewhat once the management tools are enhanced, but it is important to realize that there are limitations to the number of ports that can be shared between any two fabrics.

Brocade SAN Routing — FCR

The Fibre Channel routing service allows nodes in two or more separate fabrics to communicate without merging those fabrics. The Fibre Channel routing service can be simultaneously used as a SAN router and for SAN extension over wide area networks (WANs) using FCIP.

Fibre Channel routing (SAN routing and SAN extension) supports interoperability with Connectrix M-series fabrics (either in McDATA mode (native mode) or Open Fabric mode).

Note: The SAN router MP-7500B and SAN router FR4-18i blade are referred to as a "SAN router” throughout this chapter.

Note: In this chapter, the term "MetaSAN" is used for the collection of all SANs interconnected with Fibre Channel routers (see "MetaSAN" on page 947).

SAN routing concepts

SAN routing introduces the following concepts:

Logical Storage Area Networks (LSANs)

An LSAN is defined by specially named zones (LSAN zones) in two or more edge or backbone fabrics that contain the same device pWWNs. These LSAN zones enable Fibre Channel zones to cross physical SAN boundaries without merging the fabrics while maintaining the access controls of zones (see Figure 341).

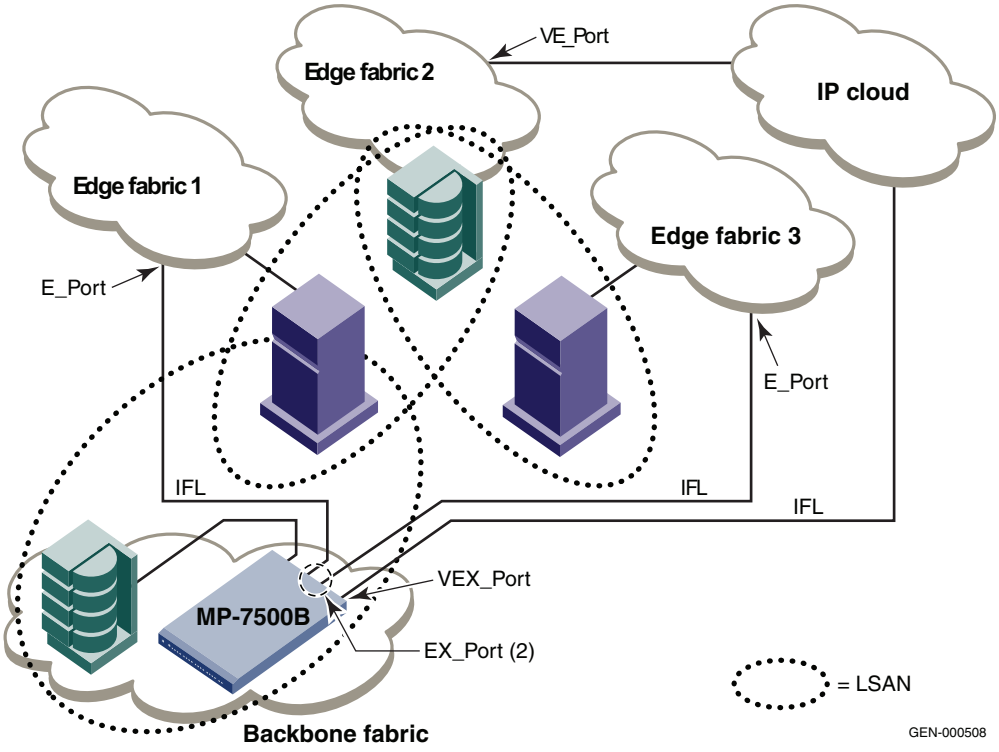


Figure 341 **MetaSAN with edge-to-edge and backbone fabrics**

Figure 341 shows a MetaSAN with a backbone consisting of one SAN router (Connectrix MP-7500B) connecting hosts in Edge Fabric 1 and 3 with storage in Edge Fabric 2 and the backbone through the use of

LSANs. There are devices shared between the backbone and Edge Fabric 1, between Edge Fabric 1 and Edge Fabric 2, and between Edge Fabric 2 and Edge Fabric 3.

EX_Port

A special type of port, called an EX_Port, functions similar to an E_Port, but terminates at the switch and does not propagate fabric services or routing topology information from one edge fabric to another. The link between an E_Port and an EX_Port is called an *interfabric link* (IFL). You can configure multiple IFLs from a SAN router MP-7500B and a Connectrix ED-48000B with the SAN router FR4-18i blade using chassis configuration option 5.

Note: The SAN router MP-7500B and SAN router FR4-18i blade are referred to as a "SAN router" throughout this chapter.

Fibre Channel routing Services support multiple EX_Ports connected to the same edge fabric. Figure 342 shows a MetaSAN consisting of three edge fabrics connected through a Connectrix ED-48000B containing an FR4-18i with interfabric links.

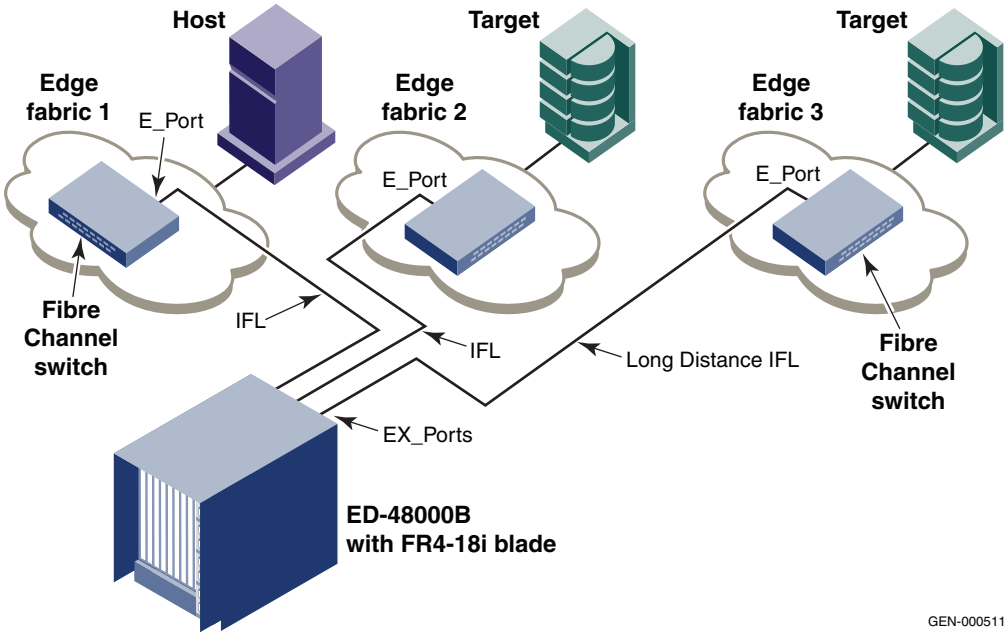


Figure 342 **MetaSAN with interfabric links (IFLs)**

Edge fabric

An edge fabric is a Fibre Channel fabric with targets and / or initiators connected through Fibre Channel routers via an EX_Port.

Backbone fabric

A backbone fabric is an intermediate network that connects one or more edge fabrics. A backbone fabric consists of at least one SAN router and possibly a number of Connectrix B Fibre Channel switches. A backbone fabric enables hosts and targets in one edge fabric to communicate with devices in the other edge fabrics connected to different routers in the backbone (see Figure 343 on page 949).

Fabric ID (FID)

Every EX_Port uses the FID property to identify the edge fabric. All of the EX_Ports attached to the same edge fabric must have the same FID. The FID for every edge fabric *must* be unique from each backbone fabric's perspective.

When two different backbones are connected to the same edge fabric, the backbone FIDs are different but each edge fabric has the same FID. Configuring the same FID for two separate backbone fabrics is invalid. In this configuration, an RAS log message displays a warning about a fabric ID overlap. However, if two backbone fabrics are not connected to the same edge, they can have the same FID.

MetaSAN

A metaSAN is a collection of SAN devices, switches, edge fabrics, Logical Storage Area Networks (LSANs), and Brocade Fibre Channel routers that comprise a physically connected, but logically partitioned, storage network.

A simple MetaSAN can be constructed using a SAN router. Additional SAN routers can be used to increase the available bandwidth between fabrics and to provide redundancy.

Note: In this chapter, the term “MetaSAN” is used for the collection of all SANs interconnected with Fibre Channel routers.

Proxy device

A proxy device is a virtual device presented into a fabric by a SAN router and represents a real device on another fabric. When a proxy device is created in a fabric, the real Fibre Channel device is considered to be imported into this fabric. The presence of a proxy device is required for interfabric device communication. The proxy

device appears to the fabric as a real Fibre Channel device, has a name server entry, and is assigned a valid port ID. The port ID is only relevant on the fabric in which the proxy device has been created.

Proxy PID

The proxy PID is the port ID of the proxy device.

Phantom domains

The SAN router emulates two types of phantom domains: front domain and translate domain.

Front domain

There is one front domain from a SAN router to an edge fabric, even if there are multiple IFLs from that SAN router to the same edge fabric.

Translate domain

The SAN router creates a translate domain for devices exported from one edge fabric to another. The translate domain has logical ISL connections to the front domains in each EX_PORT that connect to an edge fabric. If the translate domain is in a backbone fabric, then it is topologically present within the backbone because there is no front domain for a routed node within a backbone fabric. The translate domain is a proxy domain used to contain devices imported from another fabric. Devices in different edge fabrics can communicate without merging the edge fabrics using this translate and front domain architecture.

Translate domains are sometimes referred to as *xlate* domains. If a SAN router is attached to an edge fabric using an EX_Port, it will create translate domains in the fabric corresponding to the imported edge fabrics with active LSANs defined. If you import devices into the backbone fabric, then a translate domain is created in the backbone device.

If you lose connectivity to the edge fabric due to link failures or the IFL being disabled, translate domains remain visible. This prevents unnecessary fabric disruptions caused by translate domains repeatedly going offline and online when IFL connections are disrupted. To remove the translate domain in the backbone, disable all EX_Ports through which the translate domains were created.

Figure 343 shows another MetaSAN consisting of a host in edge fabric 1 connected to storage in edge fabric 2 through a backbone fabric connecting two Connectrix ED-48000B each containing FR4-18i SAN routing blades.

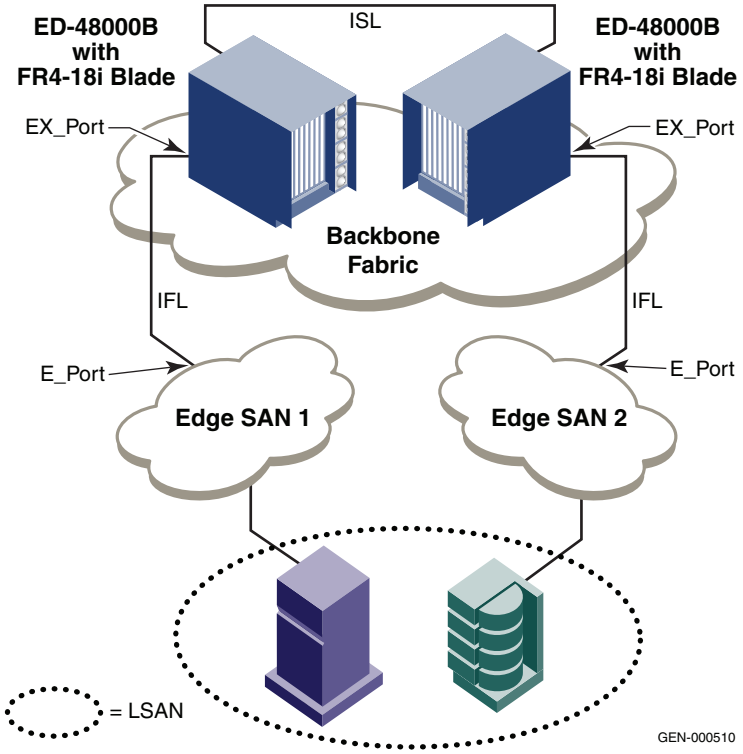


Figure 343 **Edge fabrics connected through a backbone fabric**

Supported configurations and platforms

In an edge fabric that contains a mix of Administrative Domain (AD) capable switches and switches that are not aware of AD, a device which supports Fibre Channel routing (FCR) should be connected directly to the AD-capable switch.

The supported configurations are:

- SAN router connected to a Brocade FOS fabric
- SAN router connected to a Brocade Secure Fabric OS fabric

  Note: EMC does not support Brocade Secure Fabric OS.

- SAN router connected to a Connectrix M Open mode fabric

- SAN router connected to a Connectrix M McDATA Fabric mode (native mode) fabric
- SAN router connected to a FOS and Secure Fabric OS fabrics with EX_Port ISL Trunking enabled
- First generation Fibre Channel routers
  - AP-7420B with Fabric OS v5.1 edge fabrics

Proxy devices

A SAN router achieves interfabric device connectivity by creating proxy devices (hosts and targets) in attached edge fabrics that represent real devices in other fabrics once administrators create the appropriate zones in each edge fabric. For example, a host in Fabric 1 can communicate with a target in Fabric 2 as follows:

- A proxy target in Fabric 1 represents the real target in Fabric 2.
- A proxy host in Fabric 2 represents the real host in Fabric 1.

The host discovers and sends Fibre Channel frames to the proxy target. The SAN router receives these frames and delivers them to the destination fabric for delivery to the target.

The target responds by sending frames to the proxy host. Hosts and targets are exported from the edge fabric to which they are attached and, correspondingly, imported into the edge fabric reached through Fibre Channel routing. Figure 344 on page 951 illustrates this concept.

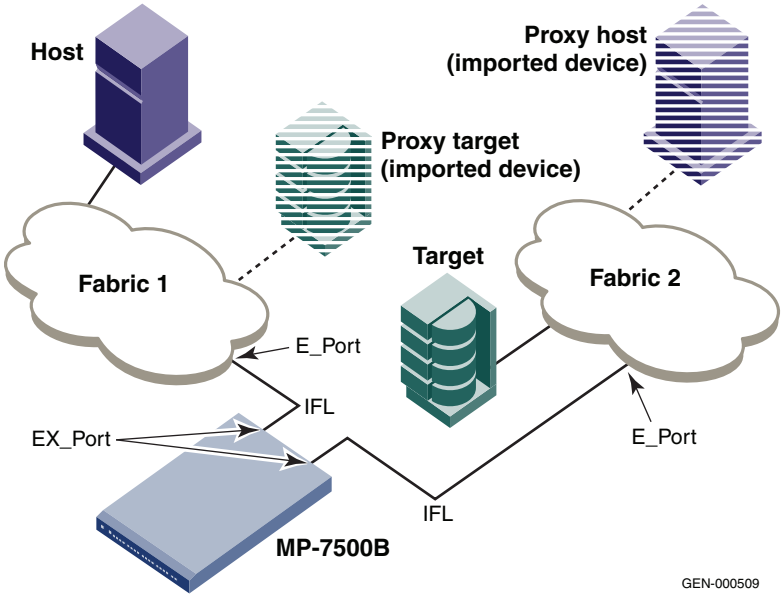


Figure 344 **MetaSAN with imported devices**

Routing types

The following routing types are discussed in this section:

- "Edge-to-edge," next
- "Backbone-to-edge" on page 951
- "Fibre Channel NAT and phantom domains" on page 952

Edge-to-edge

Fibre Channel routing between edge fabrics occurs when devices in one edge fabric communicate with devices in another edge fabric through one or more Fibre Channel routers.

Backbone-to-edge

This is when one or more Fibre Channel routers form a common fabric, known as a *backbone* fabric (see "Backbone fabric" on page 947). A backbone fabric is created with a standalone SAN router or with SAN routers connected by E_PORTS. SAN routers also enable hosts and targets in edge fabrics to communicate with devices in the backbone fabric. This is known as "backbone-to-edge routing." From the edge fabric's perspective, the backbone fabric is just like any other edge fabric. For the edge fabric and backbone fabric devices to communicate, the shared devices need to be presented to each other's

native fabric using the LSAN zoning process. To accomplish this, at least one translate domain is projected into the backbone fabric. This translate domain represents the entire edge fabric. The shared physical device in the edge has a corresponding proxy device attached to the translate domain.

Each edge fabric has only one translate domain for the backbone fabric. The backbone fabric device communicates with the proxy devices whenever it needs to communicate with the shared physical device in the edge. The SAN routing Service receives the frames from the backbone fabric destined to the proxy device, and redirects the frame to the actual physical device. As with an edge fabric, the translate domain switch can never be the principal switch of the backbone fabric. Only translate domains are created in the backbone fabric. (Front domains are not created for backbone-to-edge routing.)

Devices are exported from the backbone fabric to one or more edge fabrics using LSANs.

Fibre Channel NAT and phantom domains

Within an edge fabric or across a backbone fabric, the standard Fibre Channel FSPF protocol determines how frames are routed from the source device to the destination device. The source or destination device can be a proxy device.

Fibre Channel requires that all ports (including EX_Ports) be identified by a unique PID. In a single fabric, FC protocols guarantees that domain IDs are unique, and a PID formed by a domain ID and area ID is unique within a fabric. However, the domain IDs and PIDs in one fabric might be duplicated within another fabric.

SAN routing performs Fibre Channel network address translation (FC-NAT) so the proxy devices in a fabric can have different PIDs than the real devices that they represent. This allows overlapping PIDs in different edge fabrics to communicate and edge fabrics with common domain IDs to have nodes routed between them.

All EX_Ports connected to same edge fabric from one physical FC router present a single front domain and one additional translate (xlate) domain for each edge fabric accessed. All EX_Ports connected to an edge fabric use the same xlate domain ID number for an imported edge fabric. This value persists across switch reboots and fabric reconfigurations. Xlate domains are presented as being connected topologically behind one or more front domains. Each Fibre Channel router presents one front domain per edge fabric. This

allows redundant paths to remote fabrics to present redundant paths to proxy devices to an edge fabric.

Both front and translate domains appear to an edge fabric as logical switches. When an EX_PORT is attached to an edge fabric, the edge fabric sees a switch trying to join the fabric via an E_Port. The edge fabric provides a domain ID to the front domain and adds it to its local FSPF routing table. Similarly, the translate domain is added to the edge fabric when it is created. The combination of front domains and translate domains allows routing around path failures, including path failures through the routers. The multiple paths to a translate domain provide additional bandwidth and redundancy.

There are some differences in how the translate domain is presented in the backbone fabric. The backbone translate domains are topologically connected to SAN routers and participate in FC protocol in the backbone. Front domains are not needed in the backbone fabric. As in the case of translate domains for an edge fabric, backbone translate domains provide additional bandwidth and redundancy by being able to present itself being connected to single or multiple SAN routers with each SAN router capable of connecting multiple IFLs to edge fabrics.

DWDM

Dense Wavelength Division Multiplexing (DWDM) is a process in which different channels of data is carried at different wavelengths over one pair of fiber-optic links. This is in contrast with a conventional fiber-optic system in which just one channel is carried over a single wavelength traveling through a single fiber.

Note: For more information on DWDM, refer to "DWDM" on page 433.

CWDM

Coarse Wave Division Multiplexing (CWDM), like DWDM, uses similar processes of multiplexing and de-multiplexing different channels by assigning different wavelengths to each channel. CWDM is intended to consolidate environments containing a low number of channels at a reduced cost.

Note: For more information on DWDM, refer to "CWDM" on page 436.

Fast Write

The Fast Write feature uses the nature of SCSI I/O (specifically the transfer ready phase) to reduce the total command completion time by tricking the initiator into sending data earlier than it normally would.

Figure 345 depicts how a typical "local" SCSI I/O is performed.

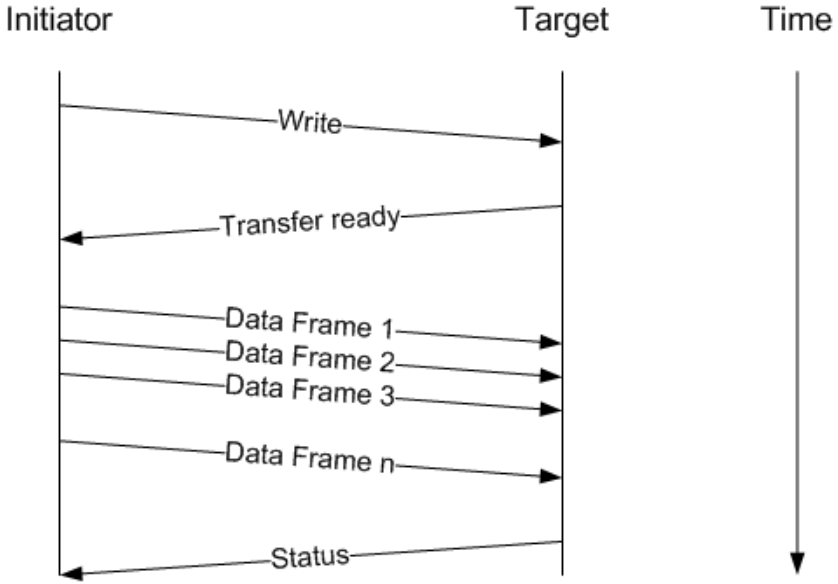


Figure 345 Typical SCSI write

The write command is sent and the target responds with *transfer ready*. The amount of data that the target allows the initiator to send is specified inside the transfer ready.

Figure 346 depicts a SCSI I/O being performed over a longer distance. Note that the long delay from the time the command is sent until the time the first data is sent.

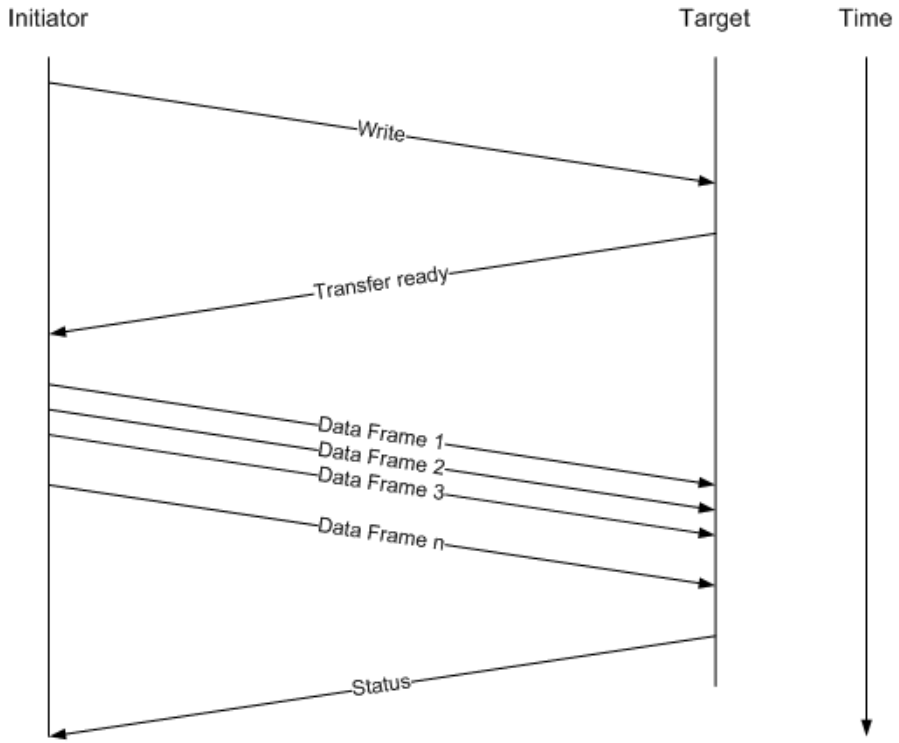


Figure 346 SCSI write over distance without Fast Write

Figure 347 depicts a SCSI I/O being performed over the same distance as in Figure 346, but in this case an appliance that supports Fast Write is inserted into the data path. Note that the amount of time saved in this I/O is approximately 1/2 the roundtrip time.

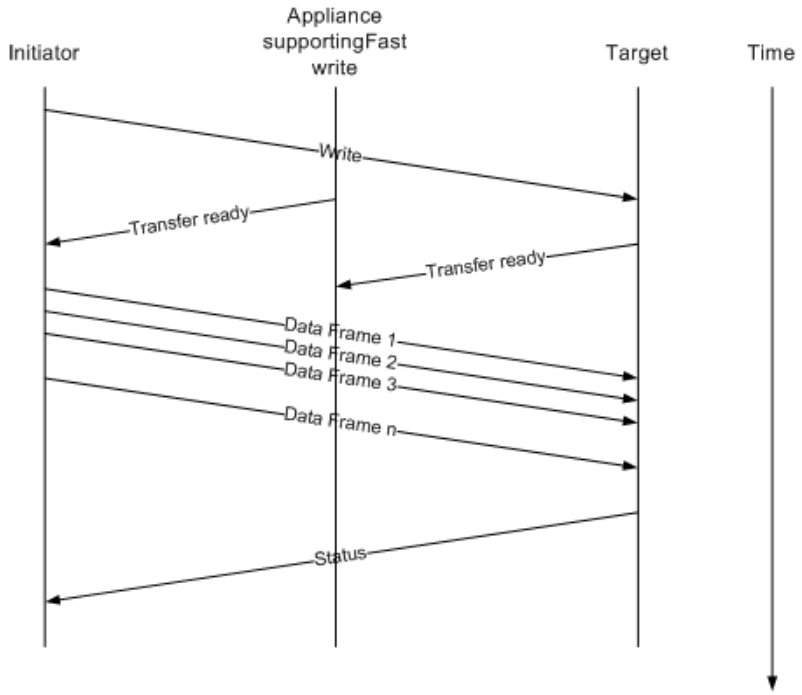


Figure 347 Fast Write over distance with appliance

The amount of time saved can be significant, but before any time savings can be realized, the latency added by the device needs to be overcome. This means that for short distances, enabling Fast Write may *negatively* impact performance. The performance benefit that your environment will realize from using Fast Write depends not only on the distance, but on the applications using the link as well.

Contrary to some information being provided by a switch vendor, a Fast Write feature has not been built into SRDF. However, due to the nature of the I/O being performed, some types of SRDF will see a different result when using an appliance that supports Fast Write.

SRDF over FC and Mirrorview

With SRDF over FC, there are always two round trips per write. With MirrorView, degradation due to distance can be more pronounced since each write results in multiple I/Os being created. In both cases, some benefit will be realized using Fast Write at intermediate and long distances:

- 30-60 km with Cisco's write acceleration using their Cisco MDS w/ SSM blade
- 100 km and greater with Brocade M-2640 or M-1620's Fast Write feature

Note: The reason Cisco provides a benefit starting at 30-60 km and McDATA does not until 100 km, relates to differences in internal efficiencies.

Adaptive Copy

All environments using Adaptive Copy will benefit from a Fast Write appliance. Although improvement increases with distance, a good average is 20-30%.

SRDF/A

Fast Write helps to minimize the impact of gaps in data over long distances. Since SRDF/A stream writes back-to-back, there normally are no gaps. As a result, since it never waits for a response before putting more data on the link, SRDF/A *rarely* benefits from Fast Write. EMC does, however, still recommend enabling Fast Write for SRDF, because it has significant benefit during both initial R1-R2 syncs and during a resync after SRDF/A suspension (planned or unplanned) and it does not affect SRDF/A performance.

The exceptions would be small writes or writes to only few volumes, where SRDF/A performance is throttled by RF max I/O rate or the limit of 32 concurrent writes per volume. In these cases, there would be data gaps, and command response time (roundtrip) directly affects throughput.

Vendor-specific features

This section contains information about some of the more important vendor-specific features: partitions, virtual switches, VSANs, IVR (Inter-VSAN Routing), and IVR-NAT (Inter-VSAN Routing Network Address Translation).

Partitions

Partitions were first introduced on the ED-10000M. Each partition in a chassis can be thought of as a separate physical switch with a couple of exceptions:

- They share the same CTPs — This can become a problem if a single port in the same or another partition were to malfunction and inundate the switch with Login or Name Server requests, all partitions would equally share the load and this could result in elongated service times on partitions that would otherwise be isolated from fabric events on another partition.
- They share the same SWMs — This can be a problem in high usage environments using LMQs. The partition a port belongs to is not considered in the fairness algorithm. Therefore all ports in the same quadrant share the available bandwidth equally regardless of the Partition they are in.
- They share the same power and cooling components — Potential availability issue.
- They must all be running the same version of firmware.

There can be up to four partitions in a single chassis with each one consisting of some number of line cards. Each partition must contain at least one line card. No internal routing function exists between two partitions. This means that in order for two ports in different partitions to access each other, an ISL needs to be connected between them.

Virtual switches

Virtual switches were first introduced by Brocade M-Series on the ED-10000M with firmware 9.1. A virtual switch is very similar to a VSAN with a few notable exceptions:

- Only four virtual switches are supported per partition with 9.1 and 9.2.
- Port Channels are not supported.

VSANs

VSANs were first introduced by Cisco on their 9509 and 9216 products with Firmware 1.x. A VSAN is a group of hosts or storage ports that communicate with each other using a virtual topology defined on the physical SAN. Using VSANs, you can build a single topology containing switches, links, and one or more VSANs. Each VSAN in this topology has the same behavior and property of a SAN.

A VSAN has the following additional features:

- Multiple VSANs can share the same physical topology.
- The same Fibre Channel IDs (FC IDs) can be assigned to a host in another VSAN, thus increasing VSAN scalability.
- Every instance of a VSAN runs all required protocols such as FSPF, domain manager, and zoning.
- Fabric-related configurations in one VSAN do not affect the associated traffic in another VSAN.
- Events causing traffic disruptions in one VSAN are contained within that VSAN and are not propagated to other VSANs.
- VSANs require that both switches have the VSAN defined on them in order for them to access each other across the ISL. In other words, if Cisco switch A has 10 N_Ports in VSAN 1000, and Cisco switch B has 10 N_Ports in VSAN 1 (default), these N_Ports will not be able to access each other until the N_Ports in VSAN 1 are moved into VSAN 1000, or vice versa, or some flavor of IVR is used.

IVR (Inter-VSAN Routing)

Initiators and targets are allowed to access each other on different VSANs without merging VSANs into a single fabric. Switch-to-switch control frames (HLO, LSU, BF) do not flow between VSANs, nor can initiators access any resource across VSANs aside from the designated ones.

Virtual SANs (VSANs) improve storage area network (SAN) scalability, availability, and security by allowing multiple Fibre Channel SANs to share a common physical infrastructure of switches and ISLs. These benefits are derived from the separation of Fibre Channel services in each VSAN and isolation of traffic between VSANs. Data traffic isolation between the VSANs also inherently prevents sharing of resources attached to a VSAN, such as a CLARiiON disk library. Using IVR, you can access resources across VSANs without compromising other VSAN benefits.

Data traffic is transported between specific initiators and targets on different VSANs without merging VSANs into a single logical fabric. Fibre Channel control traffic does not flow between VSANs, nor can initiators access any resource across VSANs other than the designated ones. Valuable resources, such as tape libraries, are easily shared across VSANs without compromise.

IVR is not limited to VSANs present on a common switch. Routes that traverse one or more VSANs across multiple switches can be established, if necessary, to establish proper interconnections.

For information on n IVR, refer to "Cisco Inter VSAN Routing (IVR) in a heterogeneous environment" on page 351.

IVR-NAT (Inter-VSAN Routing Network Address Translation)

IVR-NAT allows you to set up IVR in a fabric without needing unique Domain IDs on every switch in the IVR path. When IVR-NAT is enabled, the virtualized end device that appears in the native VSAN uses a virtual Domain ID that is unique to the native VSAN. For more information on IVRR-NAT, refer to "Configure IVR with Network Address Translation (NAT)" on page 168.

Port fencing

Port fencing is a policy-based feature that allows you to protect your SAN from repeated operational or security problems experienced by switch ports. Port Fencing allows you to set threshold limits on the number of specific port events permitted during a given time period. If the port generates more events during the specified time period, the Connectrix Manager (Port Fencing feature) blocks the port, disabling transmit and receive traffic until you have time to investigate, solve the problem, and manually unblock the port.

Threshold alerts

When you launch the Element Manager from a core switch icon and select **Configure > Threshold Alerts**, the **Threshold Alerts** dialog box displays data on threshold alerts received and transmitted. It also provides full ability to activate, deactivate, create, edit, and delete threshold alerts.

When the Element Manager is launched from a virtual switch icon and you select **Configure > Threshold Alerts**, the **Threshold Alerts** dialog box displays data on threshold alerts received and transmitted. However, function buttons are replaced with only a **View** button. You cannot activate, deactivate, create, edit, or delete threshold alerts. Selecting the **View** button displays the **Threshold Alerts** dialog box, but you can only view configuration settings for current threshold alerts. You cannot change any of these settings.

Management

For more information on SAN management, refer to Chapter 7, ”SAN Management.”

Public versus private

Today, the decision to put a switch on a public or private network is more dependent on customer requirements than anything else. Some customers like to have direct access to their switches so they can get SNMP traps directly from the switches, while others like the security that a private network affords.

Historically, EMC recommended that all switches be placed on a private network that is isolated from the rest of the data center LAN. This was recommended for a number of reasons, including limiting the scope of testing that needed to be done for each release, as well as being conservative and protecting the switch from the unknown.

There have been cases where all FC switches on a single public data center LAN segment rebooted due to an ARP storm. Although most of the FC switches have been specifically hardened against network problems of this type, it is still recommended to put environments with multiple FC switches either onto separate network segments or onto a single private network that is shielded by the Service Processor.

IP SAN concepts

For this discussion, IP can be described in two elements, each discussed in this section:

- "Internetworks, TCP, and routing", next
- "IP addressing" on page 965

Internetworks, TCP, and routing

Internetworks

An internetwork consists of individual networks joined by routers, as pictured in Figure 348. Networks A and B are connected to create internetwork A/B.

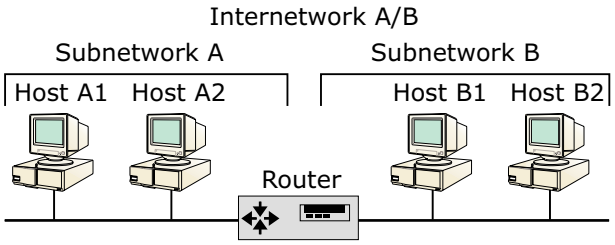


Figure 348 Internetwork example

The Internet is a collection of autonomous networks, each of which contains its own internal network of routers and subnetworks. These autonomous networks are connected to each other using gateway routers and external connections.

Each autonomous network has its own IP network address or range of addresses. Routers are like policemen directing traffic at the network interconnection point. They know the IP network addresses of connected networks and forward packets appropriately.

Every computer on the Internet must have a unique address. IP supports unique addresses by way of a hierarchical addressing scheme. An IP address is a unique number that contains two parts: a network address and host address. The network address is used when forwarding packets across interconnected networks. It defines the destination network and routers along the way know how to forward the packet based on the network address. When the packet

arrives at the destination network, the host portion of the IP address identifies the destination host.

IP can be thought of as an addressing scheme for interconnected networks, similar in function to a postal ZIP code which is a sort of overlay-addressing scheme. A ZIP code identifies a specific area, much like an IP address identifies a specific network in a group of interconnected networks. Just as the street address on an envelope defines the destination for a letter, the host portion of the IP address identifies the destination computer.

A design goal of the early Internet was to create a very basic network connectivity model that would provide fast and efficient data exchange among distributed systems without restrictions on the host systems. If applications needed additional communication services, those services would be provided by end systems, and not by the network itself. In fact, the Internet protocols will work on just about any type of underlying network technology, including Ethernet, wireless, and optical.

Transmission Control Protocol (TCP)

No discussion on IP can be complete without at least a reference to the TCP, a connection-oriented transport protocol that sends data as an unstructured stream of bytes. By using sequence numbers and acknowledgment messages, TCP provides a transmitting node with delivery information about packets sent to a receiving node. Where data has been lost in transit, TCP can retransmit the data until either a time-out condition is reached or until successful delivery has occurred. TCP can also identify duplicate messages and will discard them as necessary. If the sending computer is transmitting too fast for the receiving computer, TCP can employ flow-control mechanisms to slow data transfer. TCP can also communicate delivery information to the upper-layer protocols and applications it supports

As shown in Figure 349 on page 964, TCP and UDP (User Datagram Protocol) are layered on top of IP and take advantage of IP's datagram delivery services. UDP is a communications protocol that offers a limited amount of services when messages are exchanged between computers in a network that uses IP. If an application does not need TCP's services, it goes through UDP.

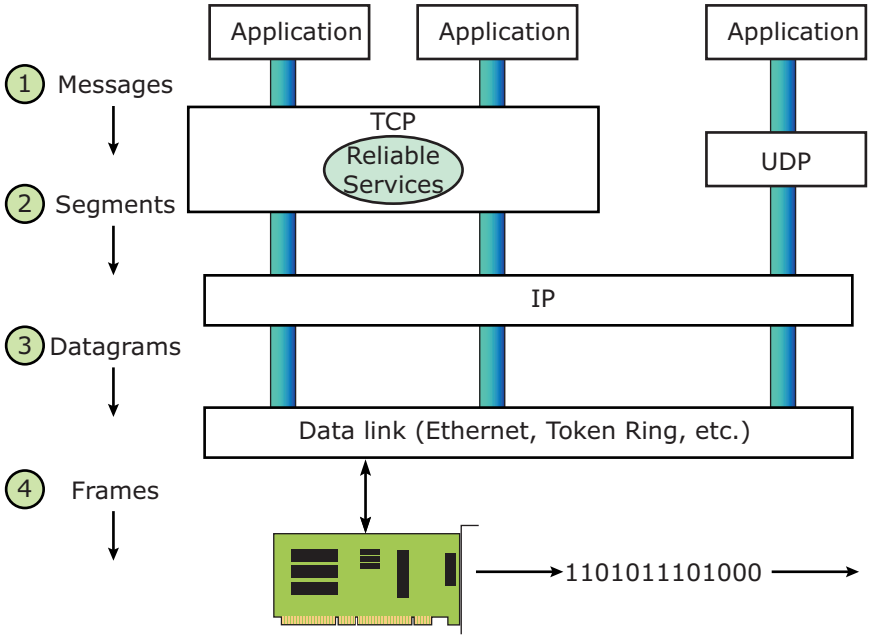


Figure 349 TCP and UDP

Applications make calls to TCP or UDP for network services. Application data or messages are encapsulated in TCP segments or UDP datagrams and then passed down to IP. IP passes the datagrams down to the underlying network, which frames the data into *blocks* for transmission.

Some applications, such as file transfers, require the absolute guarantees that TCP provides in delivering packets. It takes time and additional resources to signal that a packet must be retransmitted, but the delay is worth it to ensure data integrity.

Internetwork routing

Routers (and switches) interconnect different networks through their unique IP address. Routers determine the best way to forward incoming packets. However, a router needs to determine only the best next hop toward a destination, not the complete path to the destination. This is like getting directions while walking in a city. A person might point you in the right direction to an intersection. At the next intersection, another person points you in the right direction. Eventually, you get to where you want to go—not by knowing the exact path from the start, but by being directed along the way.

At one time, routers were called *gateways* because they provided a path or connection into another network. The term is still used to specify which router on the network serves as a gateway to another network. For example, assume an enterprise LAN is connected by one router to another enterprise LAN, and by another router to the Internet. When configuring an Internet connection, the Internet-connected router is specified as the *primary Internet gateway.*

IP addressing

This section discusses IP address structure, classes, subnet masks, and IP storage networking protocol.

IP address structure

An IP address is a 32-bit binary number containing two different pieces of information:

- Network identifier — Indicates the network (a group of computers). Networks with different identifiers can be interconnected with routers.
- Host identifier — Indicates a specific computer on the network.

While computers work with IP addresses as 32-bit binary values, humans normally use the dotted-decimal notation. A binary address and its dotted-decimal equivalent are shown below. Note that the 32-bit address is divided into four 8-bit fields called *octets.*

Example: **11000000.10101000.00001010.00000101** = **192.168.10.5**

The abbreviated form makes it easier to discuss the different addressing schemes.

IP address classes

In the early days of the Internet, the 32-bit IP address space was allocated into three classes of addresses:

- Class A — The first 8 bits identify the class and the network, the remaining 24 bits identify hosts. The 24-bit host address space identifies 16,777,214 hosts per each of the 126 networks. Most class A network schemes were assigned to U.S. government agencies, educational institutions, research organizations and large companies in the early days of the Internet.
- Class B — The first 16 bits identify the class and the network, the remaining 16 bits identify hosts. This scheme defines 16,384 networks and 65,534 hosts per network.

- Class C — The first 24 bits identify the class and the network, the remaining 8 bits identify hosts. This scheme defines 2,097,152 networks and 254 hosts per network.

A class D scheme also exists for multicasting, which is outside the scope of this overview.

The class system would be all but phased out by now except that so many organizations *own* class-based blocks of addresses and many will not voluntarily give them up.

Subnet masks

A subnet mask is an IP address feature that serves as a template to indicate which bits in the IP address define the network and which bits define the host. All devices on the same IP network *must* use the same subnet mask.

A subnet is a logical subsection of an IP network. Subnets are created to separate groups of hosts for many reasons, including security and traffic control. Subnets are usually utilized within enterprise networks to create departmental networks. A service provider with a large block of IP addresses can use subnets to allocate blocks of IP addresses to subscribers. Like networks on the Internet, routers are required at subnet boundaries to transmit packets from one subnet to another.

In terms of addressing, subnetting is equivalent to adding a third level to the Internet addressing hierarchy. The normal levels include all the networks of the Internet, and then the hosts on those networks. Subnetting creates a sublevel of networks within each network. An enterprise or an ISP may configure subnetting to divide a large network into two or more smaller networks that are easier to manage and that match the physical and transmission requirements of the underlying networks.

IP storage networking protocol (iSCSI)

iSCSI (Internet Small Computer System Interface) is an IP-based storage networking standard for linking data storage facilities developed by the Internet Engineering Task Force. By transmitting SCSI commands over IP networks, iSCSI can facilitate block-level transfers over intranets (as opposed to the file sharing found in NAS).

The iSCSI architecture is similar to that of a client/server architectural model. In this case the client is an initiator that issues an I/O request and the server is a target (such as a device in a storage system). The iSCSI initiator issues commands to the target, which fulfills the client's request. The target typically has one or more logical unit numbers (LUNs) that process the initiator's commands.

These commands are contained in a Command Descriptor Block (CDB) that has been issued by the host operating system and translated by the iSCSI driver. In the case of a *read*, the target LUN begins transferring the requested blocks back to the initiator. The iSCSI driver then translates the data into a format the host operating system will recognize.

In a typical EMC environment, for example in mid-range Windows 2000 environments, iSCSI will be used to increase access to hosts that may not have justified the cost of a Fibre Channel host bus adapter. Figure 350 illustrates operations in a typical iSCSI implementation in an EMC environment with existing devices.

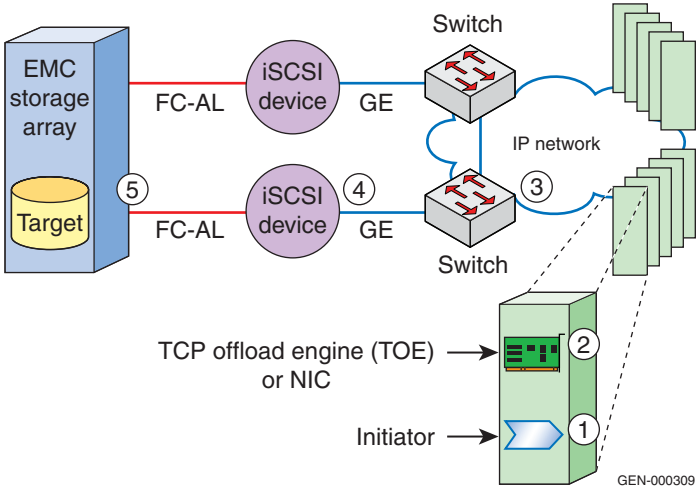


Figure 350 iSCSI implementation example

The process shown in Figure 350 is:

1. Customer application initiates I/O request to the iSCSI *drive*.
2. ISCSI driver encapsulates CDB transaction and transmits I/O over a 10/100 or GE NIC card/TOE engine.
3. Network switch routes packet to appropriate iSCSI router.
4. ISCSI device decodes packet and routes I/O to appropriate LUN.
5. LUN services application request.

You can find a more in-depth review of iSCSI at:

http://www.snia.org/tech_activities/ip_storage/iSCSI_Technical_whitepaper.PDF

IP over Symmetrix directors

This section describes IP connectivity using channel directors in Symmetrix 8000 and DMX series systems:

- "SRDF over GigE remote director" on page 968
- "iSCSI using Symmetrix multiprotocol channel director" on page 971

SRDF over GigE remote director

Symmetrix DMX series systems provide native IP support through Multiprotocol Channel Directors (MPCDs). Symmetrix 8000 series systems provide native IP support through Gigabit Ethernet (GigE) remote directors. The MPCD and GigE remote director provide comparable functionality (with the exception of data compression, which is a feature of the MPCD only).

GigE support for SRDF on Symmetrix systems is based on gigabit Ethernet technology that enables direct Symmetrix-to-IP network attachment. This increases the options for Symmetrix-to-Symmetrix connectivity. It also allows a Symmetrix system to connect to existing Ethernet infrastructure and directly access high-speed data transmission conduits through IP.

Unless otherwise specified, the following information applies to both Symmetrix 8000 series and Symmetrix DMX GigE remote director configurations.

Connectivity

GigE remote director ports are 1000base SX, and can connect to the IP network with 50 micron or 62.5 micron multimode cable. Symmetrix 8000 series systems use dual SC-style connectors; Symmetrix DMX series systems use SFP LC-style connectors.

A Symmetrix system may be connected through GigE remote connections in a one-to-one or a many-to-many (also referred to as fan-in / fan-out) configuration.

When planning an SRDF-over-IP network deployment, it is important to understand the network infrastructure to be used for Symmetrix SRDF connectivity. Network infrastructure can impact TCP connectivity and performance.

TCP connections

SRDF throughput scales per TCP connection; therefore, it is desirable to create more TCP connections through a switched network infrastructure. The number of TCP connections created depends on the network infrastructure between the Symmetrix systems and the mappings of SRDF groups to remote IP targets created in the Symmetrix **IMPL.bin** files.

In the simplest case of a single point-to-point physical connection between a pair of Symmetrix ports (shown in Figure 351), a single TCP connection is created. Scaling to two point-to-point physical connections would allow only two TCP connections.

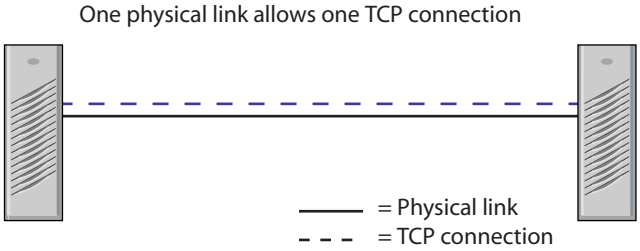


Figure 351 **TCP connection example: One TCP connection**

With a switched network infrastructure, each Symmetrix port can establish a connection to all remote ports configured as targets. For example, a configuration with two source ports and two remote targets configured (as shown in Figure 352) creates four TCP connections.

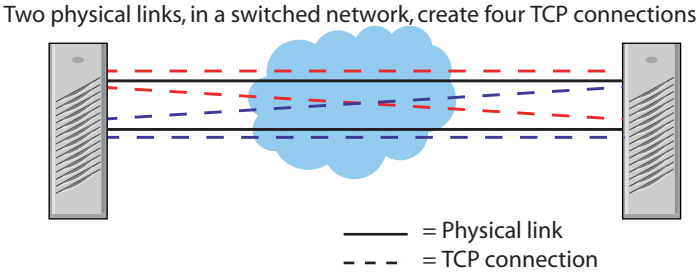


Figure 352 **TCP connection example: Four TCP connections**

Compression

The Symmetrix DMX MPCD compresses SRDF data prior to transmission on the link. Depending on the compressibility of the data, this can lower network bandwidth requirements. The SRDF data is compressed prior to encapsulation in IP packets.

The MCPD supports compression on a per TCP connection basis. The compression feature should be enabled in the file **IMPL.bin** for source directors and target IP addresses where compression functionality is desired.

Once configured in **IMPL.bin**, the compression feature can be temporarily disabled and re-enabled using Symmetrix **Inlines** commands. To successfully negotiate a TCP connection with compression enabled, both source and target ports must have compression enabled. If compression is not enabled on either end, compression is negotiated to *disabled*.

Symmetrix DMX-to-Symmetrix 8000 connectivity

Symmetrix DMX MPCDs may connect to and establish SRDF connections with Symmetrix 8000 series GigE remote directors. The only feature not available in these configurations is data compression. If data compression is enabled on the MPCD, it will recognize that the peer cannot support compression and the TCP connection will negotiate to compression *disabled*. Symmetrix MPCDs cannot establish a TCP connection with third-party compression devices.

Figure 353 shows both Symmetrix DMX and 8000 series switches connecting to an IP network through switches or routers.

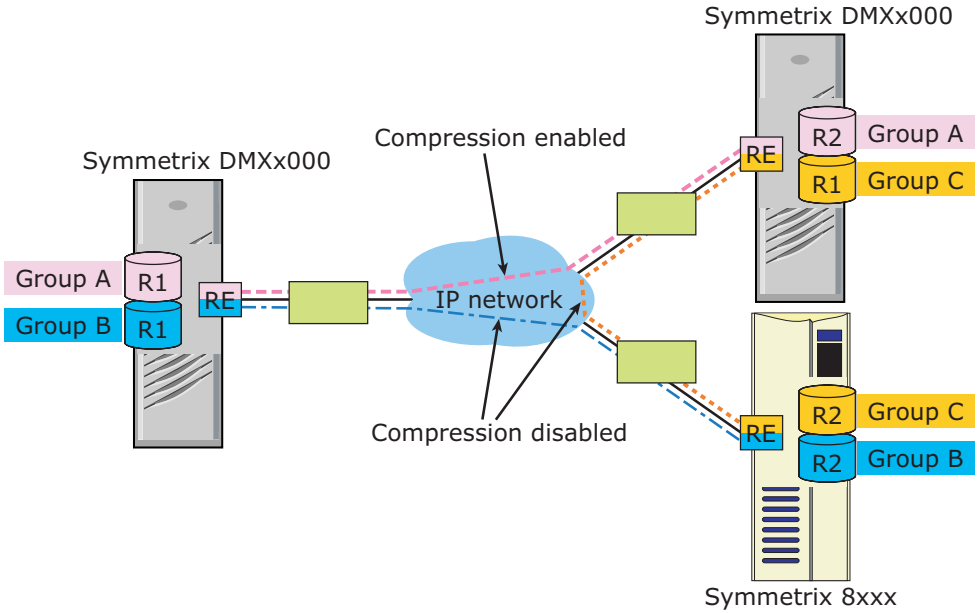


Figure 353 Connectivity: Symmetrix DMX series to Symmetrix 8000 series

iSCSI using Symmetrix multiprotocol channel director

Symmetrix DMX1000, DMX2000, and DMX3000 systems provide high-performance native iSCSI support through the combination of the multiprotocol channel director (MPCD) and the GigE (gigabit Ethernet) back adapter. Symmetrix DMX800 systems also provide native iSCSI support through MPCD to work with DMX800 front-end / back-end adapters (FEBEs).

Supportable configurations

Multiple Ethernet media access types

Multiple media access types are supported through conversion to the optical LC connector over 850 nm multimode shortwave.

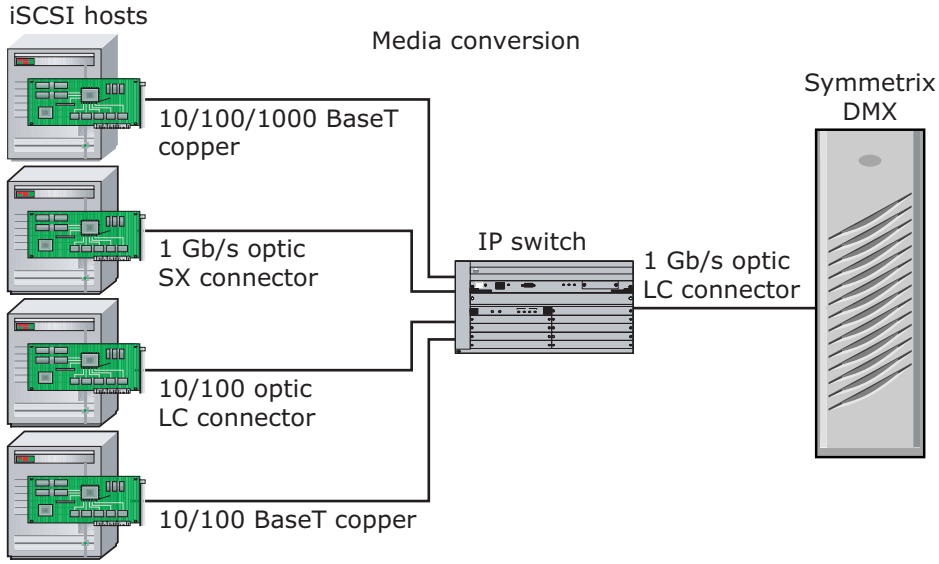


Figure 354 **Media conversion**

Multiple TCP/IP network topology types

Figure 355 and Figure 356 show the supported TCP/IP network configurations.

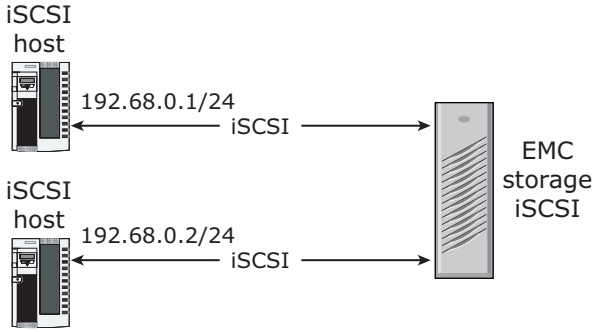


Figure 355 Direct connection of host NIC (from multiple media types) to iSCSI MPCD

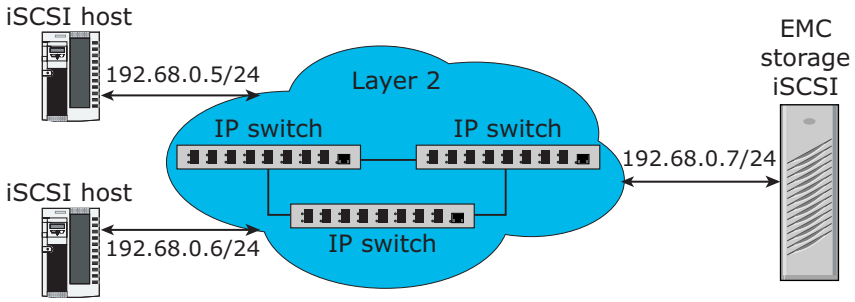


Figure 356 Switched layer 2 (single subnet) to iSCSI MPCD

Security implementation

Authentication is implemented through unidirectional CHAP (Challenge Handshake Access Protocol). The credentials and the secret are stored inside the EMC device-masking database.

The EMC *Solutions Enabler Symmetrix Device Masking CLI Product Guide* provides more information.

Data integrity

Symmetrix iSCSI MPCD supports the header digest and data digest using the CRC-32C method. The digest is negotiable by the host that logs in to the Symmetrix system. The *DMX iSCSI Director Technical Notes* (EMC P/N 300-100-170) provide more information.

B

EMC Connectrix Products Data

This appendix contains data on EMC Connectrix products.

- Overview 974
- Connectrix B-Series directors and switches 976
- Connectrix MDS-Series directors and switches 1068
- Connectrix M-Series directors and switches 1095
- Connectrix EC-1700 cabinet system 1153
- Connectrix EC-1500 cabinet system 1158
- Connectrix EC-1200 cabinet system 1166

Overview

This appendix provides data information for the following EMC Connectrix directors and switches, including descriptions, system architecture, and management:

- "Connectrix B-Series directors and switches" on page 976
- "Connectrix MDS-Series directors and switches" on page 1068
- "Connectrix M-Series directors and switches" on page 1095

This material is intended for technical consultants, solutions architects, implementation specialists, end users, or anyone interested in learning more about the features of the EMC Connectrix director and switches.

Just as the primary function of a disk array is *storage capacity*, measured in gigabytes (Gb), the primary function of a Connectrix director or switch is to provide connectivity. Connectivity capacity, or bandwidth capacity, is currently measured in gigabits per second (Gb/s) or terabits per second (Tb/s) per second.

Fibre Channel directors and switches can be built with many different architectures. The bandwidth comparisons can be complicated by ASIC architectures and front-end against back-end bandwidth characteristics. As a result of these complexities, E-Lab recommends normalizing the comparison of communication capacity between Fibre Channel Director products with a chassis-level metric of ports oversubscribed at a given line-rate. This strategy is an algebraic proxy for a gigabit per second metric, but it speaks more directly to the administrative complexities created by an oversubscribed Director.

For example, a product able to achieve no *oversubscription* when scaled to its maximum capacity will require the minimum amount of bandwidth monitoring and fewest corrective actions. A product with a high degree of oversubscription, in a response-time-sensitive environment may require a large amount of planning, monitoring, and unscheduled attention. A premium price for a product with no oversubscription may yield large administrative dividends.

The definition of oversubscription can vary by vendor. E-Lab's definition is: the ratio of bandwidth required to bandwidth available. When all ports, associated pair-wise, in any random fashion, cannot sustain full duplex at full line-rate, the switch is oversubscribed.

E-Lab's oversubscription analysis of the current Connectrix Director product set is provided in Table 45.

Table 45 Oversubscription

		Chassis			
	Measurements	**ED-10000M**	**ED-140M**	**ED-48000B**	**MDS 9513**
2 Gb/s	Maximum non-oversubscribed port count	256	140	256	264
	Maximum oversubscribed port count	N/A	N/A	N/A	528
	Oversubscription ratio at max config	N/A	N/A	N/A	2:1
4 Gb/s	Maximum non-oversubscribed port count	128	70	128	132
	Maximum oversubscribed port count	256	140	256	528
	Oversubscription ratio at max config	2:1	2:1	16:8	4:1

This appendix also provides installation guidelines and cabinet configuration examples for the following Connectrix cabinets:

- "Connectrix EC-1500 cabinet system" on page 1158
- "Connectrix EC-1200 cabinet system" on page 1166

Connectrix B-Series directors and switches

This section contains information on the following Connectrix B-Series application platform, directors, and switches.

- "Switch and blade ID numbers," next
- "Connectrix AP-7600B" on page 978
- "Connectrix AP-7420B" on page 989
- "Connectrix ED-48000B" on page 998
- "Connectrix ED-24000B" on page 1009
- "Connectrix ED-12000B" on page 1015
- "Connectrix MP-7500B" on page 1019
- "Connectrix DS-5000B" on page 1033
- "Connectrix DS-4900B" on page 1042
- "Connectrix DS-16B3" on page 1051
- "Connectrix DS-8B3" on page 1054
- "Connectrix DS-32B2" on page 1058
- "Connectrix DS-16B2" on page 1060
- "Connectrix DS-16B and DS-8B" on page 1065

Switch and blade ID numbers

This section includes the switch and blade ID numbers for the Connectrix B-Series Diretors and switches. Table 46 on page 977 lists the switch ID numbers. Table 47 on page 977 lists the blade ID numbers.

Each Connectrix B-Series Director, switch and blade type can be identified by its ID number. To locate the Director or switch ID number, use the **switchshow** command. The ID is shown next to "switchType". To determine the blade type installed in director class products, use the **slotshow** command. The ID number is shown in the type column. Once the ID number has been located, use Table 46 and Table 47 on page 977 to decode it.

Table 46 Switch ID numbers

Swtich ID	EMC switch model	Brocade switch model	ASIC
1		1000	Stitich/Flannel
26	D-16B3	3850	Bloom2
27	DS-8B3	3250	Bloom2
32	DS-4100B	4100	Condor
34	DS-220B	200E	Goldeneye
38	AP-7420B	AP7420 Multi-protocol router	Figero
42	ED-48000B	48000	Condor
44	DS-4900B	4900	Condor
46	MP-7500B	7500	Condor
58	DS-5000B	5000	Condor
	AP-7600B	7600	Condor

Table 47 Blade ID numbers

Blade ID	EMC blade name	Brocade blade name	Description	EMC Director
1		CP1	Control Processor	ED-12000B
5		CP2	Control Processor	ED-24000B
16		CP4	Control Processor	ED-48000B
2		FC-16	16 port 2 GB FC	ED-12000B/24000B
4		FC2-16	16 port 2 GB FC	ED-24000B/48000B
17		FC4-16	16 port 4 GB FC	ED-24000B/48000B
18		FC4-32	32 port 4 GB FC	ED-48000B
24	PB-48K-18i	FR4-18i	16 FC 4 GB/2 GigE	ED-48000B
31	PB-48K-16IP	FC4-16IP	iSCSI 8 4 GB/8 GigE	ED-48000B
36	PB-48K-48	FC4-48	48 port 4 GB FC	ED-48000B
33	PB-48K-AP	FA-18	Application Blade	ED-48000B

Connectrix AP-7600B

This section introduces the components of the the EMC Connectrix AP-7600B.

Description

The EMC Connectrix B Series AP-7600B is a 1U Switch that is capable of supporting storage virtualization and a fabric hosted data and storage management solution. The switch has 16 1/2/4 Gbps Fibre Channel ports and 2 physical Gigabit Ethernet (GbE) ports. It includes the Fabric Operating System and is compatible with other switches. It can operate independently or in a fabric containing multiple switches. In addition to 16 Fibre Channel ports supporting FCIP, the switch can have up to 16 virtual ports (VE or VEX ports) available through the two physical GbE ports. These internal ports are managed by virtualization software.

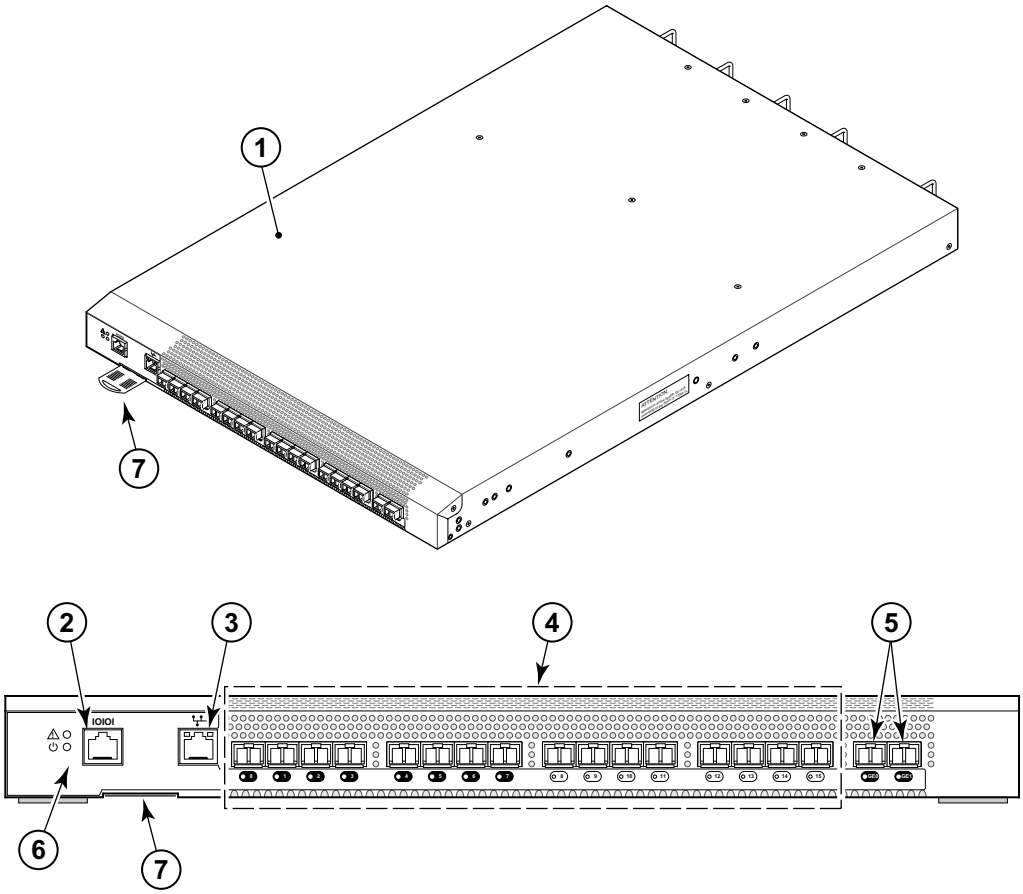
Key features of the AP-7600B

The new application platform switch, the AP-7600B, is supported with firmware v5.3.0a and up.

The AP-7600B application platform is a fixed configuration, 1U high chassis that is capable of enabling fabric intelligence by hosting storage applications that provide storage virtualization, data mobility, and data protection functionality. The AP-7600B is designed to provide the same high availability and RAS features that exist in all Connectrix B-Series platforms utilized in an enterprise class SAN.

AP-7600B port side view

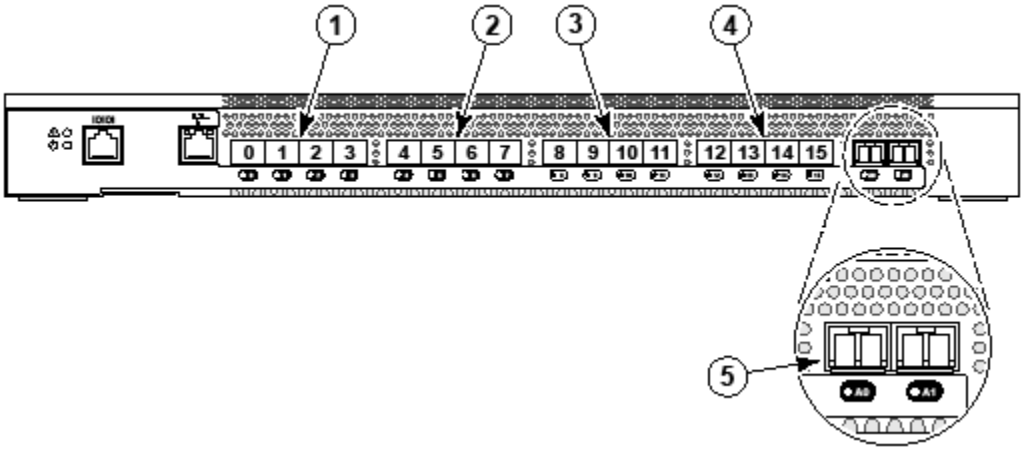
Figure 357 shows the port side of the switch.



1	Console management (serial) port	5	System status LED
2	Ethernet management port	6	Power status LED
3	Fibre ports (0-15)	7	IP address tab location
4	GbE ports (2)		

Figure 357 Port side view

The Fibre Channel ports are numbered from left to right, in eight-port groups, and are also numbered on the faceplate (see Figure 358).



1	Fibre Channel Ports 0 through 3	4	Fibre Channel Ports 12 through 15
2	Fibre Channel Ports 4 through 7	5	Gbe ports (2)
3	Fibre Channel Ports 8 through 11		

Figure 358 Port numbering in the switch

The port side also displays the system status LED, power status LED, and port status LEDs.

Power

Two power supplies, 12 V, 300 W with dual speed fans.

The power supplies are hot-swappable FRU, allowing 1+1 redundant configurations. The unit is universal power supply capable of functioning worldwide without voltage jumpers or switches. The fully enclosed, self-contained unit has internal fans to provide cooling and is autoranging in terms of accommodating input voltages.

If one power supply fails, the switch can continue to operate but the failed power supply should be replaced as soon as possible.

The power supply provides three DC outputs (5V standby and 12V), providing a total output power of 300 maximum usable watts. Each power supply plugs directly into the enclosure from the port side of the unit, mating to internal blind connectors that connect both the DC outputs and the interface signals to the system backplane. An integral on / off switch, input filter, and power indicator are provided in each

power supply, as well as a serial EEPROM device that provides identifying information.

Multiprotocol ports

There are 16 multiprotocol ports (numbered 0 through 15, left to right). Each of the 16 multiprotocol ports can be equipped with an SFP (optional). The SFPs are hot-swappable and use industry-standard local channel connectors. In Fibre Channel mode, each port provides ISL and fabric (E, EX, and F respectively) type connectivity that is automatically sensed and requires no administration to identify the port type.

In addition to the 16 Fibre Channel ports, 16 internal ports (VE and VEX), which are referred to as virtual or hidden ports are available. These internal ports are managed by virtualization software. In Gigabit Ethernet mode, SFP fiber optic transceivers that convert electrical signals to optical signals (and optical signals to electrical signals) are used. Capable of transmitting at both 1, 2, and 4 Gbit/sec speeds, each SFP fiber optic transceiver supports 850 nm SWL on multimode fiber optic cable and 1310 nm LWL on single-mode fiber optic cable. These miniature optical transceivers provide the high port density available in the switch and deliver twice the port density of standard removable GBIC transceivers.

Management ports

The switch provides 10/100 BaseT for switch management and two 10/100/1000 Ethernet ports for virtualization management. When a device is connected to the port, both ends negotiate to determine the optimal speed. The Ethernet port uses an RJ-45 connector. The TCP/IP address for each port can be configured from the serial port.

Serial port

An RS-232 serial port is provided on the switch. The serial port uses an RJ-45 connector. The serial port's parameters are fixed at 9600 baud, 8 data bits, and no parity, with flow control set to None. This connector is for initial IP address configuration and for recovery of the switch to its factory default settings, should flash memory contents be lost. The serial port connection is not intended for normal administration/maintenance functions.

Serial port specifications

The serial port can be used to connect to a workstation to configure the switchIP address before connecting the switch to a fabric or IP network. The serial port's parameters are fixed at 9600 baud, 8 data

bits, and no parity, with flow control set to **None**. Table 52 lists the serial cable pinouts.

Table 48 Serial cable pinouts

PIN	Signal	Description
1	Not supported	NA
2	Not supported	NA
3	UART1_TXD	Transmit data
4	GND	Logic ground
5	GND	Logic ground
6	UART1_RXD	Receive data
7	Not supported	NA
8	Not supported	NA

Fibre Channel port specifications

The Fibre Channel ports in the Switch are compatible with SWL, LWL, and ELWL SFP transceivers. The strength of the signal is determined by the type of transceiver in use.

The ports meet all required safety standards.

The ports are capable of operating at 1-, 2-, or 4-Gbit/sec and are able to auto-negotiate to the maximum link speed.

GbE port specifications

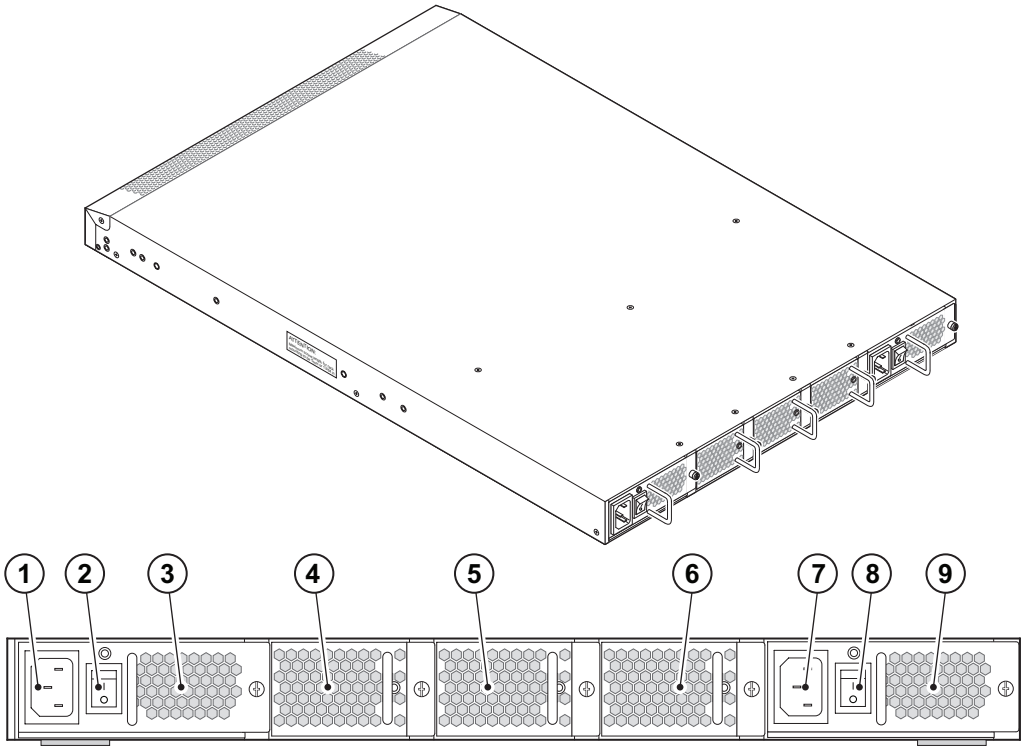
The GbE ports in the Switch are IEEE-compliant RJ-45.

The ports meet all required safety standards.

The GbE ports are capable of operating at 1-Gbit/sec.

AP-7600B nonport side

Figure 359 shows the nonport side of the switch, which contains the power supplies (including the AC power receptacle and AC power switch) and fans.



1	Power plug	6	Fan assembly 1
2	Power supply on/off switch	7	Power plug
3	Power supply 1	8	Power supply on/off switch
4	Fan assembly 3	9	Power supply 2
5	Fan assembly 2		

Figure 359 Nonport side of the switch

AP-7600B supported software features

The AP-7600B is shipped with the following features pre-installed.

- SAS — Storage Application Services
- Zoning
- Web Tools
- Full Fabric
- Extended Fabric

The following optional software is available with the purchase of a specific license key:

- Fabric Watch
- ISL Trunking
- Performance Monitoring

For information on these features, see the *Connectrix B Series Fabric OS Administrator's Guide.*

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

Components

The switch includes the following components:

- 1U chassis designed to be mounted in a standard cabinet space, with forced-air cooling that flows from the nonport side of the Switch to the port side.
- 16 Fibre Channel ports, compatible with short wavelength (SWL), long wavelength (LWL), and extended long wavelength (ELWL) SFP transceivers.
- 2 GbE ports (1-Gbit / sec), IEEE-compliant RJ-45.
- One IEEE-compliant RJ-45 serial cable, 10-ft. (approximately 3 m) long. An RJ-45 to DB9 adaptor is also provided with the Switch.
- Four IEEE-compliant RJ-45 connectors, one serial, one 10 / 100 Mbits / sec, and two GbE ports on the port side of the Switch for use with a serial console and 10 / 100 MB / sec Ethernet.
- Switch status and management LEDs: 1 switch power LED, 1 switch status LED, 6 Ethernet LEDs, 2 power supply LEDs, and 3 fan LEDs.
- Port status LEDs: 16 Fibre Channel port status LEDs.
- 2 universal AC input and redundant power supplies with AC switches and built-in fans.
- Three fan assemblies containing two fans each. The fans have two speeds, which are set automatically and cannot be modified. They default to high speed upon boot, then switch to low speed as Fabric OS comes online, returning to high speed only as required.

Memory specifications

The MP-7500B has three types of memory devices:

- Boot flash — 4 Mb

- Compact flash — 1 Gb
- Main memory (DDR SDRAM) — 1 Gb

System architecture

Performance

- 1.063 Gbit/sec line speed, full duplex
- 2.125 Gbit/sec line speed, full duplex
- Autosensing of 1 Gbit/sec and 2 Gbit/sec port speeds
- Optionally programmable to fixed port speed
- Speed matching between 1 Gbit/sec, 2 Gbit/sec port, and 4 Gbit/sec ports

Ports

- 16 ports with 16 internal ports (VE or VEX ports) which are virtual or hidden ports for use with virtualization software

Port types

- F_Port
- E_Port
- EX_port
- Self-discovery based on switch type (U_Port)
- Gigabit Ethernet

For vitualization:

- VE_port
- VEX_Port

Media types

- Small form-factor plugable (SFP) laser

Maximum size frame

- 2112-byte payload FC
- 1518-byte payload Gigabit Ethernet

Aggregate bandwidth

- 128 Gbit/sec routed full duplex large frame bandwidth

Fabric latency

- Storage application dependent

Class of service

- Class 3

Fabric services

- Simple Name Server
- Registered State Change Notification (RSCN)
- Zoning
- Exchange-Based Trunking
- Web Tools

Options

- SFP media

Management

Management software supported

- Telnet
- SNMP
- Web Tools
- Fabric Manager (optional)

Management access

- 10/100 for switch management
- Two 10/100/1000 Ethernet ports (RJ-45) for storage application management

Diagnostics

- POST
- Embedded online/offline diagnostics

Mechanical specifications

- Width = 429 mm (16.89 in.)
- Height = 1U = 42.44 mm (1.67 in.)
- Depth = 635mm (25.0 inches)
- Weight (with two power supplies and three fan assemblies installed) = 12.43 kg (27.4 lbs)

Environment

Table 53 lists the acceptable environmental ranges for both operating and nonoperating (such as during transportation or storage) conditions.

Table 49 Environmental requirements

Condition	Acceptable during operation	Acceptable during non-operation
Ambient Temperature	0° to 40° C (32° to 104° F)	-25° to 70° C (-13° to 158° F)
Humidity	10% to 85% RH non-condensing, at 40° C (104° F), with maximum gradient of 10% per hour	10% to 90% RH non-condensing, at 70° C (158° F)
Altitude	0 to 3 km (10,000 ft) above sea level	0 to 12 km (40,000 ft) above sea level
Shock	20 G, 6 ms, half-sine wave	15 G, 12-18 milliseconds, trapezoid
Vibration	0.5 G, 5-500 Hz	2.0 G, 5-500 Hz
Air Flow	31 cu ft/min (52.7 cu m/hr)	NA

Power

The power supplies are universal and capable of functioning worldwide without voltage jumpers or switches. They meet IEC 61000-4-5 surge voltage requirements and are autoranging in terms of accommodating input voltages and line frequencies. Each power supply has a built-in fan for cooling, pushing air towards the port side of the switch.

Table 57 lists the power supply specifications for the MP-7500B switch.

Table 50 Power supply specifications (1 of 2)

Specification	Value
Outlet	The outlet must be a correctly wired, primary with earth ground
Maximum output	300 Watts
System power consumption	300W Max, 215W typical
Input voltage	100-240 VAC, Universal
Input line frequency	47 - 63 Hz
Harmonic distortion	Active power factor correction

Table 50 Power supply specifications (2 of 2)

Specification	Value
BTU rating at 80% efficiency	240W X 3.412 BTU/hr/Watts = 820 BTU/Hr
Inrush current	Maximum of 15 amps for period between 10-150 ms at 50 degrees C (122 degrees F), hot or cold start
Input line protection	Fused in both hot and neutral lines, using independent fuses

General requirements

Table 58 lists the general specifications for the AP-7600B switch.

Table 51 General specifications

Specification	Description
Configurable port types	The Fibre Channel ports can be E_Ports, FL_Ports, or F_Ports.
System architecture	Nonblocking shared-memory switch
System processor	PowerPC 440GX, 800 MHz CPU
ANSI Fibre Channel protocol	FC-PH (Fibre Channel Physical and Signalling Interface standard)
Modes of operation	Fibre Channel Class 2 and Class 3
Fabric initialization	Complies with FC-SW-3 Rev. 6.6
Aggregate Switch I/O bandwidth	128 Gbit/sec if all 16 FC ports are running at 4 Gbit/sec, full duplex
Port-to-port latency	Less than 2 microseconds with no contention (destination port is free)

Standards compliance

This release conforms to the Fibre Channel Standards in a manner consistent with accepted engineering practices and procedures. In certain cases, Brocade might add proprietary supplemental functions to those specified in the standards. For the list of standards conformance, visit this Brocade website:

http://www.brocade.com/sanstandards

Further reading

For more information, refer to the following documents located on Powerlink:

- *EMC Connectrix B AP-7600B Hardware Reference Manual*
- *EMC Connectrix B Fabric OS Command Reference Guide*
- *EMC Connectrix B Fabric OS Administrator's Guide*

- *EMC Connectrix B Fabric OS MIB Reference Guide*
- *EMC Connectrix B Fabric OS Error Message and Reference Guide*
- *EMC Connectrix B Fabric OS Administrator's Guide*
- *EMC Connectrix B Web Tools Administrator's Guide*

Connectrix AP-7420B

This section introduces the hardware components of the EMC Connectrix AP-7420B Application Platform and some of the features of the XPath OS, including the Multiprotocol SAN Routing Services (MSRS) and XPath Fibre Channel switch support, and the features inherent to the system architecture.

Description

The Multiprotocol Router is a multiple board design. Below the system board in the chassis is a DC-power printed circuit board (PCB) that provides the required system voltages. These voltages are derived from and regulated by the Multiprotocol Router redundant power supply units. This regulated power is bused via a connector to the main system PCB.

Mounted on top of the system board is a daughter board with a highly integrated PowerPC-family CPU (800 MHz PowerPC 745x). It contains a high-performance RISC processor core with SDRAM controller, PCI bus interface, peripheral local bus for external ROM and peripherals, DMA, I^2C interface, and general-purpose I/O.

The system uses four types of memory devices in the design—SDRAM, kernel flash, compact flash (user flash), and boot flash. The fabric application and switching section of the system board, the XPath per-port processing ASIC and memory chipsets, the XPath fabric ASIC, and the SFP media are the key components to provide high-speed data manipulation and movement.

SFP media interface to external devices and support any combination of short wavelength (SWL) and long wavelength (LWL) optical media.

The system chassis is a 2U-high enclosure with space to contain two hot-swappable 1+1 power supply units, the motherboard, the control system daughter card, the power regulator board, and dual hot-swappable fan assemblies. Refer to "AP-7420 port side view" on page 990 and "AP-7420B nonport side view" on page 993 for additional information about the hardware components.

Figure 360 shows the Connectrix B AP-7420.

Figure 360 Connectrix B AP-7420

Key features of the AP-7420B

- Ports per chassis
  - 16 ports, universal Fibre Channel (E, F and EX)

  Note: EX_Ports are not supported on the AP-7420 when it is hosting an Invista instance.

- Port speed
  - 2 Gb
- Media types
  - Small Form Factor Pluggable (SFP) laser
  - Short-wave up to 300m
- Connectivity

  Operating as a Fabric Application Platform, it is linked via ISLs (neither IFLs nor EX_Ports are supported when operating as a fabric application platform) to a Brocade SAN
- Operations
  - Fibre Channel router
  - Fabric Application Platform

AP-7420 port side view

The port side of the AP-7420B (see Figure 361 on page 991) includes two 320-watt removable, redundant power supplies with hot-swappable capability, Fibre Channel ports, management ports, and a serial port.

Refer to the following sections for information about each port side component:

- "Power supplies" on page 991

- "Multiprotocol ports" on page 992
- "Management ports" on page 992
- "Serial port" on page 992

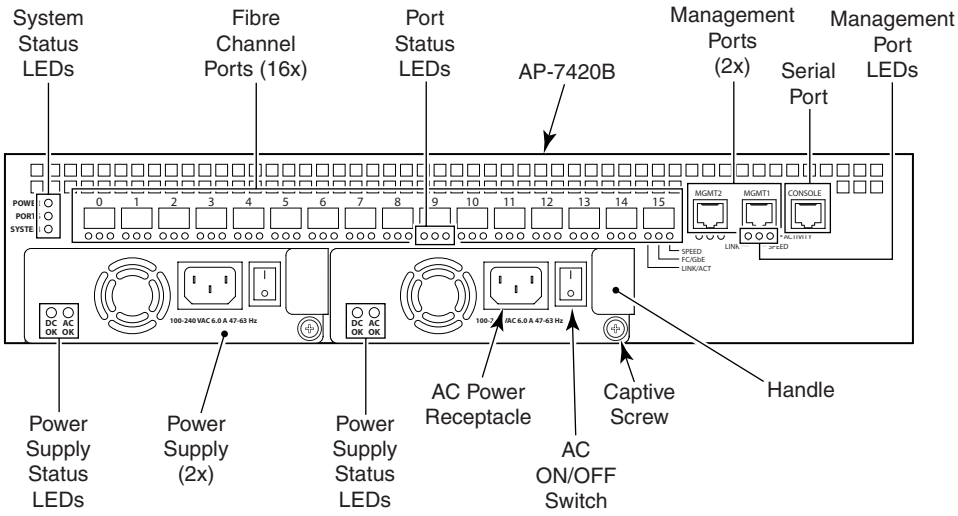


Figure 361 **AP-7420B (Port Side)**

Power supplies

The AP-7420B power supply is a hot-swappable FRU, allowing 1+1 redundant configurations. The unit is a universal power supply capable of functioning worldwide without voltage jumpers or switches. The fully enclosed, self-contained unit has internal fans to provide cooling and is auto-ranging in terms of accommodating input voltages.

The power supply provides three DC outputs (5V standby, 12V, and 48V), providing a total output power of 320 maximum usable watts. The power supplies plug directly into the enclosure from the port side of the unit, mating to internal blind connectors that connect both the DC outputs and the interface signals to the system backplane. An integral on / off switch, input filter, and power indicator are provided in the power supply, as well as a serial EEPROM device that provides identifying information.

Multiprotocol ports

The Multiprotocol Router has 16 multiprotocol ports (numbered 0 through 15, left to right).

Each of the 16 multiprotocol ports are equipped with an SFP. Each ASIC in the Multiprotocol Router switch connects to a single SFP media. The SFPs are hot-swappable and use industry-standard local channel connectors.

In Fibre Channel mode, each port provides ISL, loop, and fabric (E, F, and FL, respectively) type connectivity that is automatically sensed and requires no administration to identify the port type.

Additionally, the ports can be configured to provide EX_Port connectivity for use when routing traffic between SANs. To use the ports as EX_Ports, you must manually configure them. Refer to the *XPath OS Administrator's Guide* for details about port configuration.

Optionally, multiprotocol ports can be configured to act as IEEE-compliant optical Gigabit Ethernet ports, running at 1 Gbit/sec speed. When configured this way, the optional GigE ports are used as part of the Brocade FCIP and iSCSI implementation.

Management ports

The AP-7420B provides dual, fully IEEE-compliant 10/100 Base-T Ethernet ports for management purposes. When a device is connected to the port, both ends negotiate to determine the optimal speed. The Ethernet port uses a RJ-45 connector. There are three externally visible LEDs for each Ethernet port. One LED indicates link, one indicates transmit/receive activity, and one indicates speed (10 Mb/s or 100 Mb/s). The TCP/IP address for each port can be configured from the serial port.

Serial port

An RS-232 serial port is provided on the AP-7420B. The serial port uses an RJ-45 connector. The serial port's parameters are fixed at 9600 baud, 8 data bits, no parity, and flow control is set to *None*. This connector is for initial IP address configuration and for recovery of the switch to its factory default settings should flash memory contents be lost. The serial port connection is not intended for normal administration/maintenance functions.

AP-7420B nonport side view

The nonport side of the AP-7420B (Figure 362) includes dual hot-swappable cooling fan assemblies with LEDs (containing three fans each) and system status LEDs. .

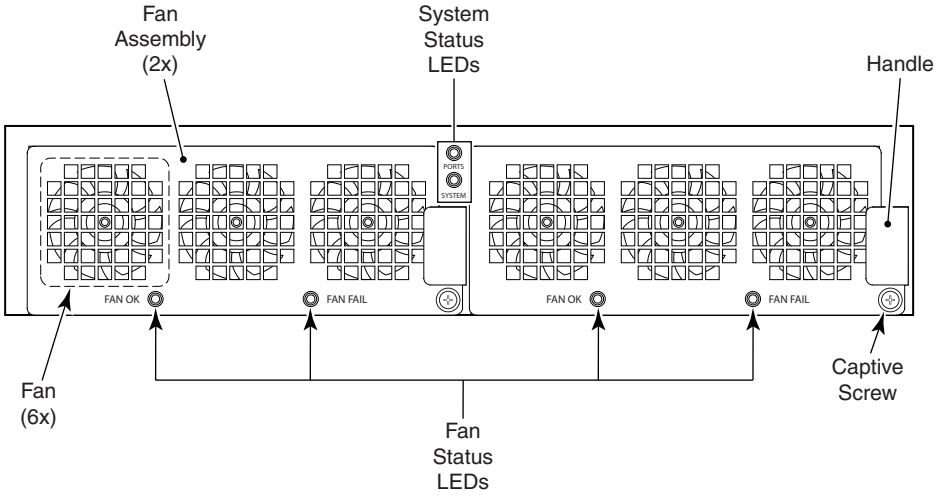


Figure 362 **AP-7420B Multilprotocol Router (Nonport Side)**

Supported features

The XPath OS runs on the Connectrix AP-7420B Application Platform. It provides full Fibre Channel switch capability, FC-FC Routing Service, an iSCSI to Fibre Channel gateway, and Fibre Channel over IP tunneling.

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

You can configure individual ports on the AP-7420B to support the services or switch functionality you need. Refer to the following sections for additional software information:

- "Fibre Channel switch support" on page 994
- "Multiprotocol SAN routing services" on page 994
- "User interfaces" on page 995

Fibre Channel switch support

The XPath OS includes the following Fibre Channel switch features:

- Name Server support
- Zone Server support
- ISL exchange-based trunking

Refer to the *Connectrix B Series XPath OS Administrator's Guide* located on Powerlink for additional information on these features.

Multiprotocol SAN routing services

Multiprotocol SAN routing services are available on the AP-7420B. These licensed services include the FC-FC Routing Service, iSCSI Gateway Service, and the FCIP Tunneling Service.

- FC-FC routing service

  The FC-FC routing service provides connectivity to devices in different fabrics without merging the fabrics. A major component of this service is the multiprotocol router. It allows data to flow across it while it controls the visibility of devices behind it. FC-FC routing allows the creation of logical storage area networks (LSANs). An LSAN can span multiple fabrics, allowing Fibre Channel zones to cross physical SAN boundaries without merging the fabrics yet still maintaining access control of the zones. FC-FC routing also allows you to share devices, such as tape drives, across multiple fabrics without the associated administrative problems that can result from merging the fabrics—including change management, network managment, scalability, reliability, availability, and serviceability.

  Refer to the *Connectrix B Series XPath OS Administrator's Guide* located on Powerlink for additional FC-FC routing information.

- iSCSI gateway service

  Note: At the time of this publication, EMC does not support the iSCSI features of the XPath OS. Please refer to the *EMC Support Matrix* for the most up-to-date information.

  The iSCSI gateway service provides connectivity to Fibre Channel targets for servers using iSCSI. Servers use an iSCSI adapter or an iSCSI driver and Ethernet adapter to connect to a Fibre Channel fabric over IP. Refer to the *Connectrix B Series XPath OS Administrator's Guide* located on Powerlink for additional iSCSI gateway information

- FCIP tunneling service

  The FCIP tunneling service enables passive Fibre Channel frames to “tunnel” through TCP/IP networks by encapsulating them in TCP packets and then reconstructing them at the other end of the link.

  Note: XPath OS only supports FCIP between two AP-7420B Application Platforms.

  Refer to *Connectrix B Series XPath OS Administrator’s Guide* located on Powerlink for additional FCIP tunneling information.

User interfaces

The XPath OS provides both a graphical user interface (Web Tools-AP Edition) and a command line interface (CLI). The graphical user interface is introduced in the *Web Tools-AP Edition Administrators Guide*. The CLI commands are described in the *Connectrix B Series XPath OS Command Reference Manual* located on Powerlink.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architecture

Performance

Performance features include:

- 1.063 Gb/s line speed, full duplex
- 2.125 Gb/s line speed, full duplex
- Autosensing of 1 Gb/s and 2 Gb/s port speeds
- Optionally programmable to fixed port speed
- Speed matching between 1 Gb/s and 2 Gb/s ports

Port types

- FL_Port (restricted), F_Port, and E_Port
- Self-discovery based on switch type (U_Port); Gigabit Ethernet

Dual-mode ports

- 16 ports, Fibre Channel (E, F) and Gigabit Ethernet

Media types

- Small form-factor plugable (SFP) laser
- Shortwave up to 500 m (1,640 feet)

- Long wave up to 10 km (6.2 miles)
- Extended long wave up to 25 km (15.5 miles)

Maximum size frame

- 2112-byte payload FC
- 1518-byte payload Gigabit Ethernet

Aggregate bandwidth

- 42 Gb/s routed full duplex large frame bandwidth

Fabric latency

- 42 Gb/s routed full duplex large frame bandwidth

Class of service

- Class 3

Fabric services

- Simple Name Server
- Registered State Change Notification (RSCN)

Optional fabric services include:

- Zoning
- Exchange-Based Trunking
- Web Tools-AP Edition

Other optional services include:

- Fibre Channel Routing
- FCIP

Options

- SFP media

Switch interoperability

- Connectrix B DS-8B and DS-16B running Fabric OS v2.6.1 or higher.
- Connectrix B DS-8B2 and DS-16B2 running Fabric OS v3.1 or higher.
- Connectrix B DS-32B2 and ED-12000B running Fabric OS v4.1 or higher.
- Connectrix B DS-8B3, DS-16B3, and ED-24000B running Fabric OS v4.2 or higher.

- Connectrix B DS-220B and ED-48000B running Fabric OS v5.0.1or higher.
- SilkWorm 2xxx running Fabric OS V2.6.1 or higher.
- SilkWorm 3200 and 3800 running Fabric OS v3.1 or higher.
- SilkWorm 3900 and 12000 running Fabric OS v4.1 or higher.
- SilkWorm 3250, 3850, and 24000 running Fabric OS v4.2 or higher.
- SilkWorm 200E and 48000 running Fabric OS v5.x or higher.

Management

Management software supported

- Telnet
- SNMP (FE MIB, FC Management MIB)
- Web Tools-AP Edition (optional)
- Fabric Manager (optional)

Mangement access

- 10 / 100 Ethernet port (RJ-45)
- Serial port (RS-232)
- In-band via Management Server

Diagnostics

- POST and embedded online / offline diagnostics.

Mechanical specifications

- Width = 16.8 inches (42.67 cm)
- Height = 3.46 inches (8.79 cm)
- Depth = 25.0 inches (63.5 cm)
- Weight = 35 pounds (15.88 kg); 1 power supply and 1 filler panel
- 40 pounds (18.14 kg); 2 power supplies

Environment

- Temperature
  - Operating: 10º C to 40º C (50ºF to 104ºF)
  - Non-operating: -25º C to 70º C (-13ºF to 158ºF)
- Humidity
  - Operating: 20% to 85%, noncondensing
  - Non-operating: 20% to 85%, noncondensing
- Altitude
  - Operating: 3 km

- Non-operating: 3 km

- Shock
  - Operating: 105 G, 2.5 ms, half-sine
  - Non-operating: 40 G, 13 ms, trapezoidal
- Vibration
  - Operating: 0.5 G (5-500-5 Hz)
  - Non-operating: 2.0 G (5-500-5 Hz)

Power

The power supply is auto-ranging and supports from 90 VAC to 250 VAC and from 47 Hz to 63 Hz. The operating range and current rating for typical operating conditions are shown below:

Imput voltage	100 to 120 VAC	200 to 240 VAC
AC line frequency	47 to 63 Hz	47 to 63 Hz
Current rating	6A	3A

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix ED-48000B

This section provides a description of the hardware and features for the EMC Connectrix ED-48000B. The information provided comes from Brocade data sheets and testing experience in the EMC IOP qualification labs. More information for the ED-48000B can be found in the *Connectrix B Series ED-48000B Hardware Reference Manual* located on Powerlink.

Description

The Connectrix ED-48000B is a Director class switch which can contain up to 256 concurrently active 4 Gb / s full-duplex ports in a single domain. It is based on the Brocade Silkworm 48000 Condor ASIC-based architecture. The ED-48000B Director is a 256-port non-blocking switch. The suggested use of the ED-48000B is in the core of the SAN where high availability, high port count, and robustness are needed. The ED-48000B chassis is a 14U enclosure with up to 512 ports in a single Connectrix cabinet. Two ED-48000B Directors can be interconnected to create a high port-count solution.

The ED-48000B utilizes a Linux 2.4.x modified kernel to load the fabric OS version 5.x and greater.

Note: For specific versions of supported firmware, as well as fabric topology constraints associated with the ED-48000B, refer to the *EMC Support Matrix*.

Key features of the ED-48000B

- Up to 256 ports in a single chassis.
- A single logical switch, that encompasses all port blades in the chassis for ease of maintenance.
- Support for two new high-performance port blades running at 1, 2, or 4-Gb/s
- 16-port blade (FC4-16)
- 32-port blade (FC4-32)
- Dual-redundant control processors provide high availability and enable non-disruptive software upgrades
- Redundant and hot swappable CPs, power supplies, and blower assembly
- Forward and backward compatibility with all Connectrix B series directors
- Supports 1-, 2-, and 4-Gb/s auto-sensing Fibre Channel ports.
- Trunking groups up to eight ports to create high performance 32 Gb/s ISL trunks between switches
- Universal ports self-configure as E_Ports, F_Ports, or FL_Ports.

ED-48000B Port side view

Figure 363 displays the portside of the ED-48000B director.

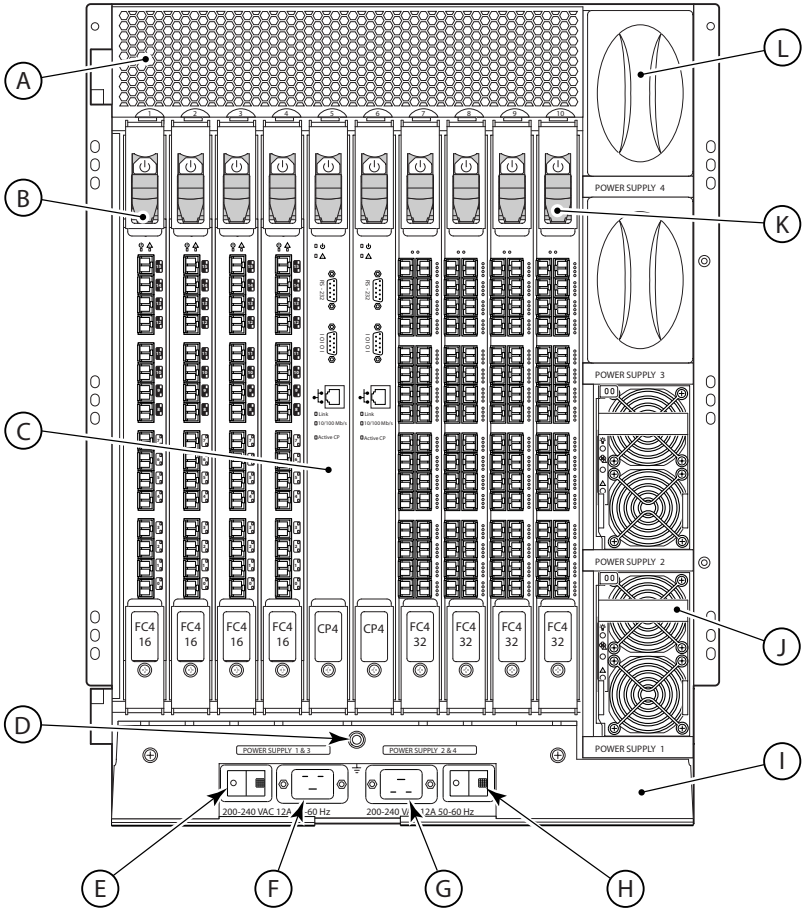


Figure 363 Port side of the ED-48000B Director

Legend:

A	Exhaust Vent	**G**	AC Power Connector (for Power Supplies 2 and 4)
B	FC4-16 Port Blade	**H**	AC Power Switch (for Power Supplies 2 and 4)
C	CP4 Blade (Control Processor Blade)	**I**	Cable Management Tray
D	Grounding Strap Connectort	**J**	Power Supply #1
E	AC Power Switch (for Power Supplies 1 and 3	**K**	FC4-32 Port Blade
F	AC Power Connector (for Power Supplies 1 and 3)	**L**	Power Supply Filler Panel

Nonport Side of the ED-48000B

Figure 364 displays the nonport side view of the ED-48000B.

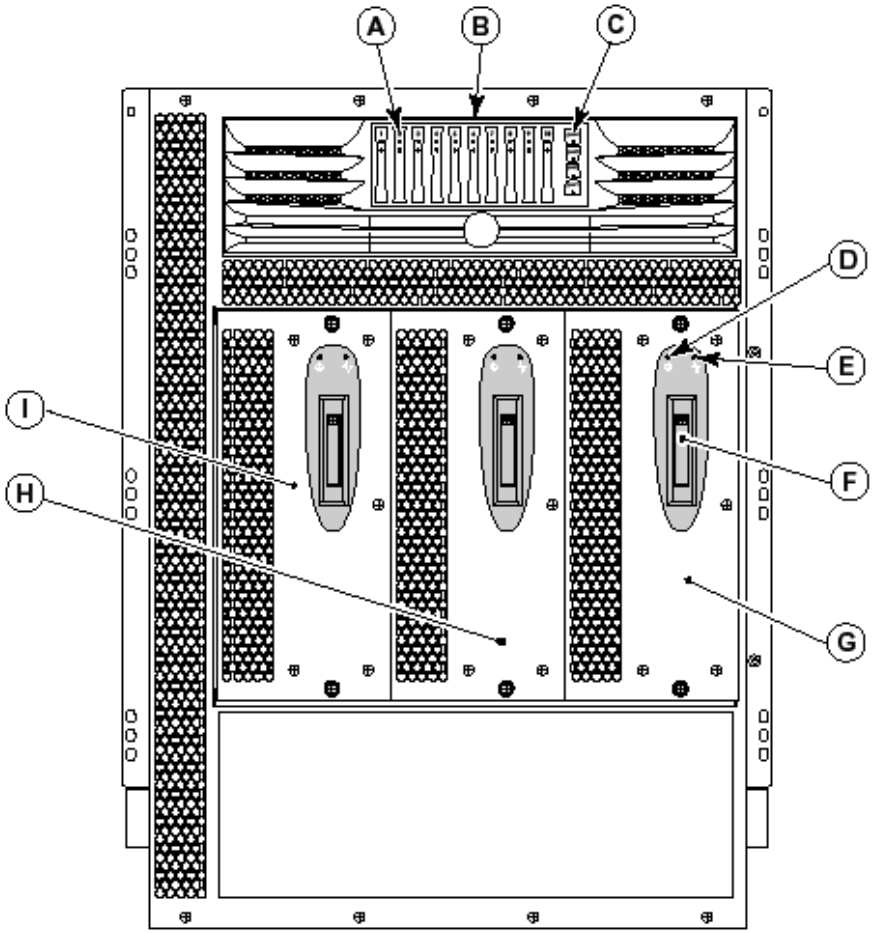


Figure 364 Nonport side of the ED-48000B director

Legend:

A	Port Blade and CP Blade LEDs	**F**	Blower Handle
B	WWN Bezel	**G**	Blower Assembly #3
C	Power Supply LEDs	**H**	Blower Assembly #2
D	Blower Power LED	**I**	Blower Assembly #1
E	Blower Fault LED		

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architecture

Performance

Performance features include:

- 1.063 Gb/s line speed, full duplex
- 2.125 Gb/s line speed, full duplex
- 4.25 Gb/s line speed, full duplex
- Auto-sensing of 1, 2, and 4 Gb/s port speeds
- Optionally programmable to fixed port speed
- Speed-matching between 1,2, and 4 Gb/s ports

Extended ISL mode buffer allocation

Mode	Buffer allocation			Distance @ 1 Gb/s	Distance @ 2 Gb/s	Distance @ 4 Gb/s	Extended Fabric license required?
	1 Gb/s	2 Gb/s	4 Gb/s				
L0	5(26)[a]	5	5	10k m	5 km	2k m	No
LE	11	16	26	10 km	10 km	10 km	No
L0.5	18	31	56	25 km	25 km	25 km	Yes
L1	31	56	106	50 km	50 km	50 km	Yes
L2	56	106	206	100 km	100 km	100 km	Yes
LD[b]	Auto	Auto	Auto	Auto (500 km max)	Auto (250 km max)	Auto (100 km max)	Yes
LS[c]	User defined	User defined	User defined	User defined (500 km max)	User defined (250 km max)	User defined (100 km max)	Yes

a. For each data channel (in this case, there are 4) there are 5 credits, plus 6 extra credits.

b. The dynamic long-distance mode (LD) automatically configures the number of buffer credits required, based on the actual link distance.

c. The static long-distance mode (LS) allocates the number of buffer credits based on the user-specified distance.

World Wide Name Card (WWN)

WWN Card maintains chassis-specific information such as WWNs, IP addresses, and summary status information of each port blade assembly and power supply through LEDs.

Control processor

Redundant (active/standby) control processor modules.

Fibre Channel ports

- Up to eight hot-swappable port blade assemblies, which can be configured in a single chassis, delivering up to 256 Fibre Channel ports
- Two slots for control processor (CP) blade assemblies (slots 5 and 6):
  - A single active CP blade can control all 256 ports in the chassis
  - The standby CP blade assumes control of the switch if the active CP fails
- Up to eight Fibre Channels port cards, either 16 or 32 ports
- Up to 512 ports per 42U rack

Port types

- Universal ports (U_port), which self-configure as E_ports, F_ports, or FL_ports; G_port functionality is determined by the type of attached device
- FL_Port Port Module support for SFP optical transceivers, supporting any combination of SWL and LWL optical media.
- F_Port
- E_Port

Port numbering

The Connectrix ED-48000B uses the following port numbering schemes:

- For the FC4-16 port blade, ports are numbered from 0 through 15 from bottom to top
- For the FC4-32 port blade, ports are numbered from 0 through 15 from bottom to top on the left set of ports and 16 through 31 from bottom to top on the right set of ports
- Slots are numbered 1 through 10, from left to right when facing the portside of the director

Refer to Figure 363 for ED-48000B port numbering.

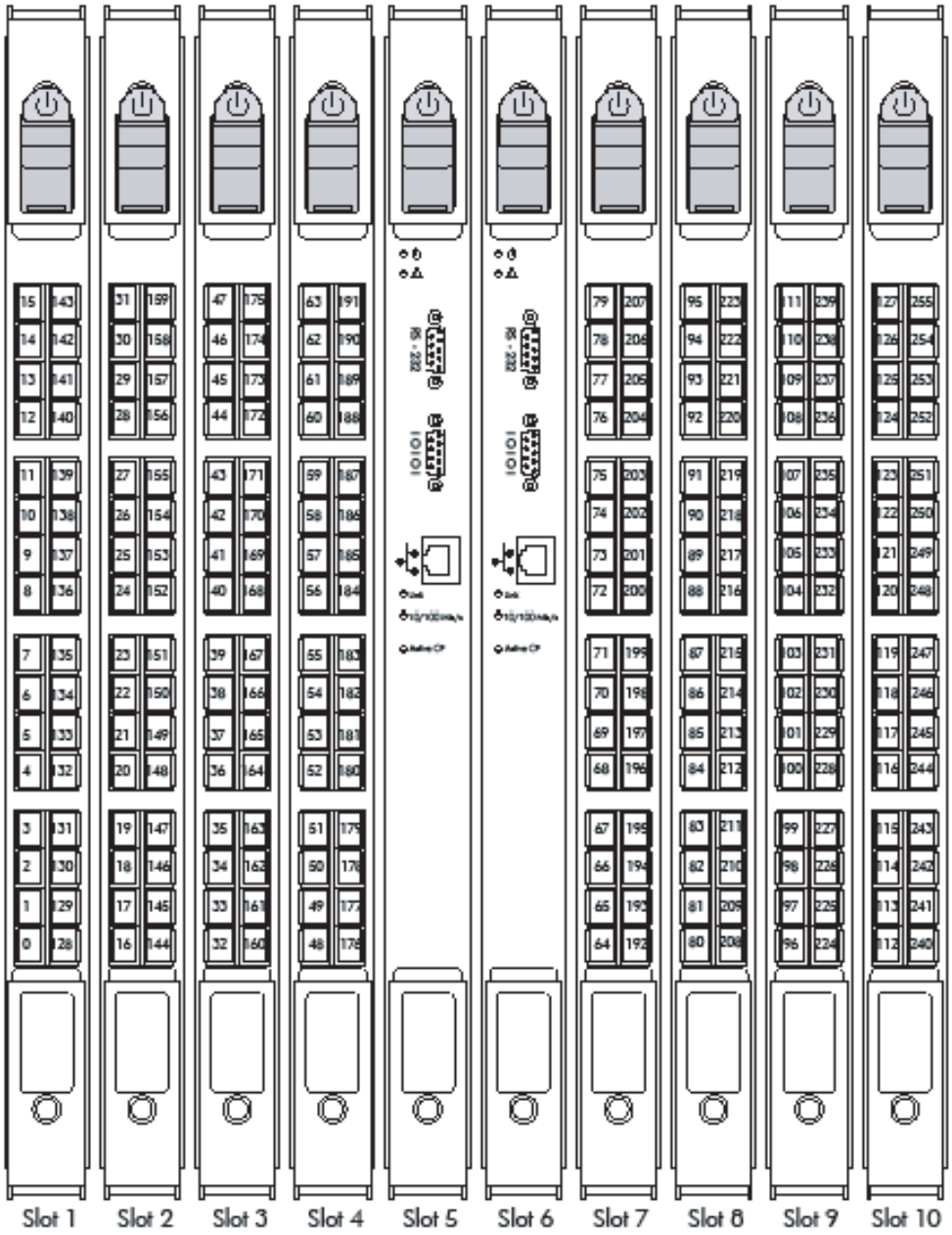


Figure 365 Fully populated ED-48000B with ports numbered

ISL Trunking

- Up to eight 4.25 Gb/s ports per ISL trunk
- Up to 32 Gb/s per ISL trunk
- Up to two 8-port trunk groups supported on 16 port card
- Up to four 8-port trunk groups supported on 32 port card
- ISL trunking at 2 Gb/s for compatibility with prior B series Connectrix products

Aggregate bandwidth

- 2.176 Tb/s

Switch latency

- For the FC4-16 blade:
  - <1.2 µsec any port to any port at 4 GB/s, cut-through routing, within the same blade
- For the FC4-32 blade:
  - <3.6 µsec any port to any port at 4 GB/s, cut-through routing

Oversubscription

- 128 port configuration
  - 1.088 Tb/s at 4 Gb/s port speed
- 256 port configuration
  - 1.088 Tb/s at 2 Gb/s port speed
  - 2.176 Tb/s at 4 Gb/s port speed
- FC4-16 blade:
  - No oversubscription at 2 Gb
  - No oversubscription at 4 Gb
- FC4-32 blade:
  - No oversubscription at 2 Gb
  - 16:8 oversubscribed at 4 Gb

Maximum frame size

- 2112-byte payload

Frame buffers

- 1024 per blade
- Dynamically allocated up to 255 per port

Classes of service

- Class 2, Class 3, and Class F (inter-switch frames) Fibre Channel protocol support

Fabric services

- Simple Name Server
- Registered State Change Notification (RSCN)
- Alias Server (multicast)
- Zoning
- Web Tools
- Fabric Watch
- Extended Fabrics
- ISL Trunking
- Performance Monitoring.

High Availability

- Redundant, hot-swappable components
  - Chassis (including backplane, blower and power supply backplane, AC harness, and blower harness)
  - 16 and 32 port Fibre Channel switch modules
  - WWN Card
  - Small Form-factor Pluggable (SFP) Short-Wavelength Laser (SWL) transceivers
  - Small Form-factor Pluggable (SFP) Long-Wavelength Laser (LWL) transceivers
- Redundant power supply and blower assembly subsystems
  - The ED-48000B requires a minimum of two operating blower assemblies at al times. To ensure continuous adequate cooling, maintain three operating blower assemblies at all times except for the brief period when replacing a blower assembly.
  - The port blades automatically shut down if the internal temperature range is exceeded (up to 75 degrees Celsius).
- Enhanced data integrity on all data paths
- Fabric Shortest Path First (FSPF) rerouting around failed links
- Integration with Simple Network Management Protocol (SNMP) managers

- Automatic control processor failover
- Non-disruptive “hot” software code loads and activation
- Easy configuration, save, and restore
- Hot-swappable World Wide Name (WWN) card
- Four AC-DC power supply modules
- 4N redundancy

Cooling

- Three blowers, providing cooling, allowing continuous operation even if one blower fails
- Rear panel-to-door airflow

Management

Management software supported

- EMC ControlCenter
- Fabric Manager
- RADIUS
- Embedded web server to provide single-point management using the Web Tools GUI
- Centralized configuration and management of fabric using Client / Server architecture and CLI
- Support for 10 / 100 Mb Ethernet connections to CP for out-of-band management
- In-band over Fibre Channel (requires fabric)
- SNMP (FE MIB, FC Management MIB)
- Telnet / SSH

Management access

- One 10 / 100 Mb Ethernet RJ-45 management port per Control Processor module
- Two DB-9 serial ports per Control Processor module

Diagnostics

- Post and embedded online/offline diagnostics

Mechanical specifications

- Width = 43.74 cm (17.2 in)
- Height = 61.24 cm (24.22 in) for 14U

- Depth = 70.90 cm (27.90 in) without door

  74.20 cm (29.20 in) with door
- Weight = 95 kg (210 lb) for 128-port configuration

  (8 X FC4-16, without media)

  98 kg (216 lb) for 256-port configuration

  (8 X FC4-32, without media)

Environment

- Temperature
  - Operating: 0º C to 40º C (32º F to 104º F)
  - Non-operating: -25º C to 70º C (-13º F to 158º F)
- Humidity
  - Operating: 5% to 85% non-condensing at 40º C (104º F)
  - Non-operating and storage (non-condensing): 0% to 93%
- Altitude
  - Up to 3000 meters (9800 feet)
- Shock
  - Operating: 20G, 11 ms, half sine 1G p-p, 5-500Hz, 1 octave min
  - Non-operating: 33G, 11ms, half sine 2.4G p-p, 5-500Hz, 1 octave min
- Vibration
  - Operating: 5G p-p, 0 to 3kHz at 1.0 octave min
  - Non-operating: 10G p-p, 0 to 5 kHz at 1.0 octave min
- Heat dissipation
  - 915 W or 3115 BTU (eight FC4-32 blades and two CP4 blades)
  - 710 W or 2425 BTU (eight FC4-16 blades and two CP4 blades)

Power

- Supported power range
  - Nominal: 200 to 240 VAC nominal, 5.0 A, single phase
  - Operating: 180 to 264 VAC auto-sensing

Note: 256-port configuration requires a maximum of 750 Volt-Amps.

- In-rush current
  - 40 Amps maximum, peak
- Frequency
  - 47 to 63 Hz
- Recommended power connector: IEC 320, EN60320 C19-Angled, 16A / 250VAC

Protocols

- FC
- FICON

Standards compliance

This release conforms to the Fibre Channel Standards in a manner consistent with accepted engineering practices and procedures. In certain cases, Brocade might add proprietary supplemental functions to those specified in the standards. For the list of standards conformance, visit this Brocade website:

http://www.brocade.com/sanstandards

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix ED-24000B

The Connectrix ED-24000B Director contains 128 concurrently active 2 Gb / s full-duplex ports in a single domain. The ED-24000B chassis delivers up to 128 ports in a single domain in a 14U enclosure with up to 256 ports in a single Connectrix cabinet, and can be interconnected to create a high port-count solution.

The ED-24000B utilizes a Linux 2.4.x modified kernel to load the fabric OS version 4.x. The Fabric OS includes all basic switch and fabric support software with an optional software license, which is enabled using license keys. EMC recommends the dual-fabric SAN

deployment for the ED-24000B, to ensure alternate paths to data during upgrades or failures.

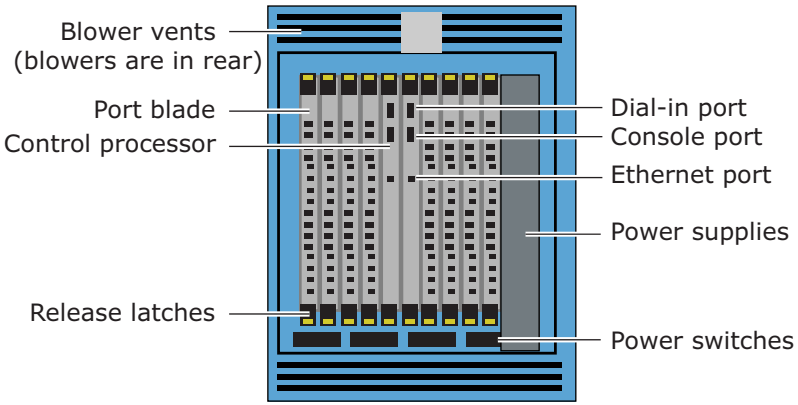


Figure 366 **ED-24000B, front view**

Some components of the ED-24000B are:

- Port Blades — House the SFPs.
- Control Processors — Provide Fabric OS services, I/O traffic routing, and monitoring.
- WWN Card — Stores the unique World Wide Name serial ID for the director, and multiple serial numbers for identification of the chassis, backplane, and the WWN card itself.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

Connectivity

Connectivity features include:

- 128 ports, universal (E_Port, F_Port, and FL_Port); up to 256 ports per 40U Connectrix enclosure.
- Up to eight hot-swappable 16-port switch blade assemblies.
- Modular hot-swappable Field Replaceable Units (FRUs):

- Chassis (including backplane, blower and power supply backplane, AC harness, and blower harness).
- 16-port Fibre Channel switch modules.
- Control Processors.
- WWN Card.
- Two AC-DC power supply modules; 2N redundancy.
- Three blower modules (each includes blower, control board, and housing); at least two must be operational.
- Power supply filler panel.
- SFP SWL transceivers.
- SFP LWL transceivers.

◆ Two Control Processors (CPs), each with an Ethernet port and two IP addresses.

  Each CP has two serial ports; the top serial port is an RS-232 port used for EMC Support connection, and the bottom port is for console access.

◆ Capacity on each switch blade assembly for 16 auto-sensing 1 Gb/s to 2 Gb/s Fibre Channel ports.

◆ Media types: Hot-pluggable, industry-standard Small Form-factor Pluggable (SFP) LC connector; Short-Wavelength Laser (SWL) up to 500 m (1,640 ft) for 1 Gb and 300 m (984 ft) for 2 Gb; Long-Wavelength Laser (LWL) up to 35 km (21.7 mi).

◆ Port Module support for SFP (small form-factor pluggable) optical transceivers, supporting any combination of short wavelength (SWL) and long wavelength (LWL) optical media.

◆ Universal ports, which self-configure as E_Ports, F_Ports, or FL_Ports; G_Port functionality is determined by the type of attached device.

◆ Fabric services: Simple name server; Registered State Change Notification (RSCN); Alias Server (multicast); Brocade Advanced Zoning, Web Tools, Fabric Watch, Extended Fabrics, ISL Trunking, and Advanced Performance Monitoring.

  Note: Some features require an additional license.

◆ Available BB_Credits (buffer-to-buffer credits):

  - 16 per F_Port.

- 26 per E_Ports.
- 27 for L1 mode.
- 60 for L2 mode.

Note: L1 mode requires the Extended Fabric License.

- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.
- Support for wide range of value-added applications, including:
  - Fabric Access.
  - Fabric Watch.
  - Advanced Web Tools.
  - Zoning.
- Aggregate bandwidth 512 Gb/s end to end.
- 2112 byte maximum frame size.
- 108 frame buffers per four-port group, dynamically allocated.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the ED-24000B.

System architecture

Performance

Performance features include:

- 1.063 Gb/s to 2.125 Gb/s line speed, full duplex; auto-sensing of 1 Gb/s and 2 Gb/s port speeds; optionally programmable to fixed port speed; speed matching between 1 Gb/s and 2 Gb/s ports.
- Up to four 2.125 Gb/s ports per ISL trunk; up to 8.5 Gb/s per ISL trunk.
- Full duplex to 200 MB/s data rate per Fibre Channel port.

Management

Management features include:

- Centralized configuration and management of fabric through Client/Server architecture and CLI.
- Management software: Telnet; SNMP (FE MIB, FC Management MIB, Brocade switch-specific MIBs); Web Tools; Fabric Watch; Fabric Access layer.

- Embedded Apache Web server to provide a single-point management using the Web Tools GUI. (The switch can also be managed through Telnet CLI or EMC Fabric Manager.)
- Interactive and non-interactive (nondisruptive) firmware upgrades.
- Support for 10/100 Mb Ethernet connections to CP (control processor) for out-of-band management.
- Support for SNMP, and Telnet CLI.
- 10/100 Ethernet (RJ-45), in-band over Fibre Channel (requires fabric); two serial ports (DB-9) per control processor module.

Mechanical specifications

Mechanical specifications include:

- Rear panel-to-door airflow
- Width = 43.74 cm (17.22 in)
- Height = 61.24 cm (24.11 in) for 14U
- Depth = 70.90 cm (27.90 in) without door, 74.20 cm (29.20 in) with door
- System weight = 88 to 96 kg (193 to 212 lb)

Environment

Mechanical factors include:

- Operating temperature: 0° C to 40° C (50° F to 104° F); non-operating temperature: -25° C to 70° C (-13° F to 158° F)
- Operating humidity: 5% to 85% non-condensing at 40° C (104° F)
- Altitude: up to 3000 meters (9800 feet)
- Operating shock: 20 g, 6 ms duration, half-sine
- Operating vibration: 0.5 g p-p, 5 to 500 to 5 Hz; non-operating vibration: 2.0 g, 5 to 500 Hz
- Heat dissipation: 740 W, 2525 BTU/hr fully loaded

Power

Power specifications include:

- Nominal power range: 200 to 240 VAC, 5.0 A, single phase
- Operating power range: 180 to 264 VAC auto-sensing, maximum 2300 Volt-Amps, maximum 12 Amps
- In-rush current: 40 Amps maximum, peak

- Frequency: 47 to 63 Hz

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Power-on self test (POST), diagnostic tests on hardware.
- Error detection and fault isolation (EDFI), such as cyclic redundancy checking (CRC), parity checking, checksum, and illegal address checking.
- Environmental monitoring and control.
- Modular design with hot-swappable components.
- Redundant Flash memory that stores two firmware images per Control Processor.
- Modem for remote diagnostics and status monitoring through serial port for dial-in support.

  Note: Direct modem connections to the switch are for EMC personnel use only. Consult EMC Customer Support before attaching a modem.

- Memory scrubber, self test, and bus ping to determine if a bus is not functioning.
- Watchdog timers to monitor processes, providing recovery mechanism in event of failure.
- Status LEDs to indicate faulty blades, CPs, or reduced connectivity speeds.
- Predictive diagnostics analysis using Fabric Watch and Error logging.
- SNMP integration with higher-layer managers.
- Persistent logging through Errorlog and Eventlog, which record information on the active CP only.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix ED-12000B

The Connectrix ED-12000B chassis contains dual 64-port Fibre Channel Enterprise Core fabric switches. The dual-switch capability allows use of either one or two 64-port switches per chassis. One or two ED-12000B chassis can be rack-mounted in a Connectrix EC-1230B cabinet, and interconnected together to create a high port-count solution. The ED-12000B utilizes a Linux 2.4.x modified kernel to load the Fabric OS. version 4.0. The Fabric OS includes all basic switch and fabric support software with an optional software license, which is enabled using license keys. EMC recommends the dual-fabric SAN deployment for the ED-12000B to ensure alternate paths to data during upgrades or failures.

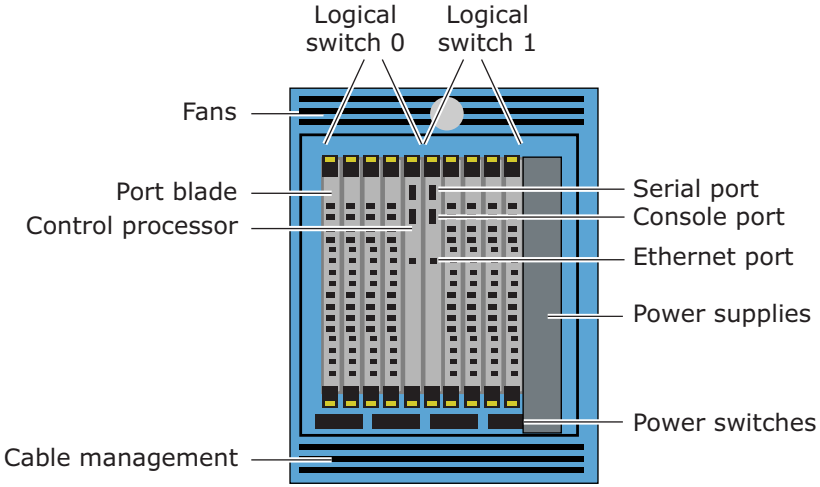


Figure 367 ED-12000B, front view

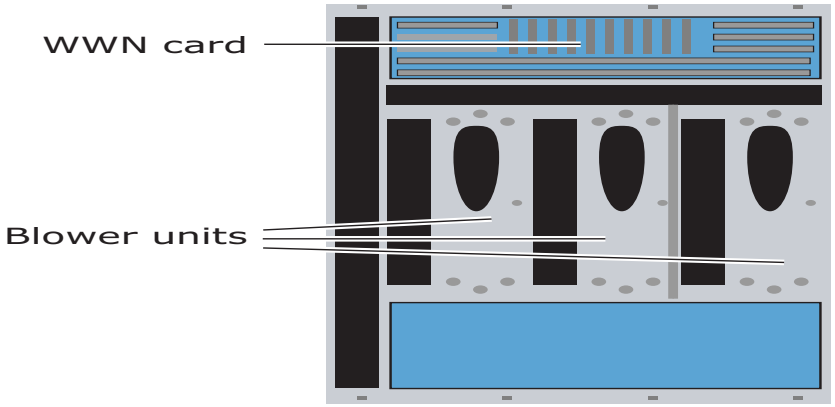


Figure 368 **ED-12000B, rear view**

Some components of the ED-12000B are:

- Port Blade assembly — Houses the SFPs.
- Control Processors assembly — Provides fabric OS services and monitoring.
- WWN Card — Stores the unique World Wide Name serial ID for each logical switch, and multiple serial numbers for identification of the chassis, backplane, and the WWN card itself.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

Connectivity

Connectivity features include:

- Up to eight hot-swappable 16-port switch blade assemblies.
- Up to two 64-port Fibre Channel switches in a single chassis; each 64-port switch uses four blade assemblies.
- Modular hot-swappable Field Replaceable Units (FRUs):
  - Chassis (including backplane, blower and power supply backplane, AC harness, and blower harness)

- 16-port Fibre Channel switch modules
- Control Processors
- Power supply
- Blower assembly (includes blower, control board, and housing)
- Power supply filler panel
- SFP SWL transceivers
- SFP LWL transceivers

◆ Two slots for Control Processor (CP) blade assemblies. A single active Control Processor can control both logical switches in the chassis.

◆ Ethernet port on each Control Processor.

◆ Capacity on each switch blade assembly for 16 auto-sensing 1 Gb/s to 2 Gb/s Fibre Channel ports; blades use SFP optical transceivers.

◆ Two serial ports on each Control Processor. The top serial port is an RS-232 port used for remote dial-in, and the bottom port is for console access.

◆ Availability in 32-port to 128-port configurations: 16 ports per blade, 64 per logical switch.

◆ Port Module support for SFP (small form-factor pluggable) optical transceivers, supporting any combination of short wavelength (SWL) and long wavelength (LWL) optical media.

◆ Universal ports, which self-configure as E_Ports, F_Ports, or FL_Ports; G_Port functionality is determined by the type of attached device.

◆ Full duplex to 200 MB/s data rate per Fibre Channel port.

◆ Trunking technology, which allows groups of up to four ports to create high-performance 8 Gb/s ISL trunks between switches.

◆ Hardware and name server-enforced zoning.

◆ Available BB_Credits (buffer-to-buffer credits):

- 16 per F_Port
- 26 per E_Ports
- 27 for L1 mode
- 60 for L2 mode

Note: L1 and L2 mode require the Extended Fabric License.

- Class 2 and Class 3 Fibre Channel protocol support.
- Support for wide range of value-added applications, including:
  - Fabric Access
  - Fabric Watch
  - Remote Switch
  - Advanced Web Tools
  - Zoning

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the ED-12000B.

Management

Management features include:

- Centralized configuration and management of fabric using Client/Server architecture and CLI.
- Embedded Apache Web server to provide a single-point management using the Web Tools GUI. (The switch can also be managed using Telnet CLI.)
- Both interactive and non-interactive firmware upgrades.
- Support for 10/100 Mb Ethernet connections to CP (control processor) for out-of-band management.
- Support for SNMP, and Telnet CLI.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Power-on self test (POST), diagnostic tests on hardware.
- Error detection and fault isolation (EDFI), such as cyclic redundancy checking (CRC), parity checking, checksum, and illegal address checking.
- Environmental monitoring and control.
- Modular design with hot-swappable components.
- Redundant Flash memory that stores two firmware images per Control Processor.
- Modem for remote diagnostics and status monitoring through serial port for dial-in support.

- Memory scrubber, self test, and bus ping to determine if a bus is not functioning.
- Watchdog timers to monitor processes, providing recovery mechanism in event of failure.
- Status LEDs to indicate faulty blades or reduced connectivity speeds.
- Predictive diagnostics analysis through Fabric Watch and Error logging.
- SNMP integration with higher-layer managers.
- Persistent logging through Errorlog and Eventlog, which record information on the active CP only.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix MP-7500B

This section introduces the components of the the EMC Connectrix MP-7500B.

Description

The MP-7500B is a 1U switch with 16 Fibre Channel SFP ports and 2 physical Gigabit Ethernet (GbE) ports. It includes the Fabric Operating System and is compatible with the entire EMC Connectrix B product family. It can operate independently or in a fabric containing multiple switches.

The MP-7500B has three types of memory devices:

- Boot flash - 4 MB
- Compact flash - 1 G
- Main memory (SDRAM) - 512 MB

This switch is intended as a platform for Fibre Channel Routing Services and Fibre Channel over IP (FCIP). Refer to the *Connectrix B Series Fabric OS Administrrator's Guide* for information on configuring these features.

More information for the MP-7500B can be found in the *Connectrix B Series MP-7500B Hardware Reference Manual* available on Powerlink.

Key features of the MP-7500B

The MP-7500B provides the following features:

- Rack mountable 1U chassis
- 2 redundant, hot-swappable power supplies
- 3 internal temperature sensors
- 3 redundant, hot-swappable fan FRUs. Each fan FRU has two fans (for a total of 6 fans). Only one fan speed is displayed per FRU using the **fanShow** command
- 16 Fibre Channel SFP ports supporting Fibre Channel Routing Services with link speeds up to 1-, 2-, or 4-Gbit/sec
- 2 GbE ports supporting the FCIP and Fibre Channel Routing Services features with link speeds up to 1-Gbit/sec:
  - Each GbE port can support up to 8 FCIP tunnels
  - Each FCIP tunnel is represented and managed as an Fibre Channel E_Port
  - Fibre Channel Routing Services functionality can be used over the FCIP link
  - Fabrics connected through FCIP merge if the ports are configured as VE_Ports, and do not merge if they are configured as VEX_Ports. If VE_Ports are used in an Fibre Channel Routing Services backbone fabric configuration, then the backbone fabric merges but the EX_Port attached to edge fabrics do not merge. For more information see the *Connectrix B Series Fabric OS Reference Manual.*

Once the switch is configured, the **switchShow** command displays 32 Fibre Channel ports (port numbers 0 through 31) and 2 GbE ports. The first 16 Fibre Channel ports are physical ports on the MP-7500B, Ports 16-23 are virtual ports associated with the GE0 physical GbE link and ports 24-31 are virtual ports associated with GE1 physical GbE link. The GbE ports are displayed as **ge0** and **ge1** and are not assigned port numbers or area numbers.

Port side of the MP-7500B switch

Figure 369 shows the port side of the MP-7500B.

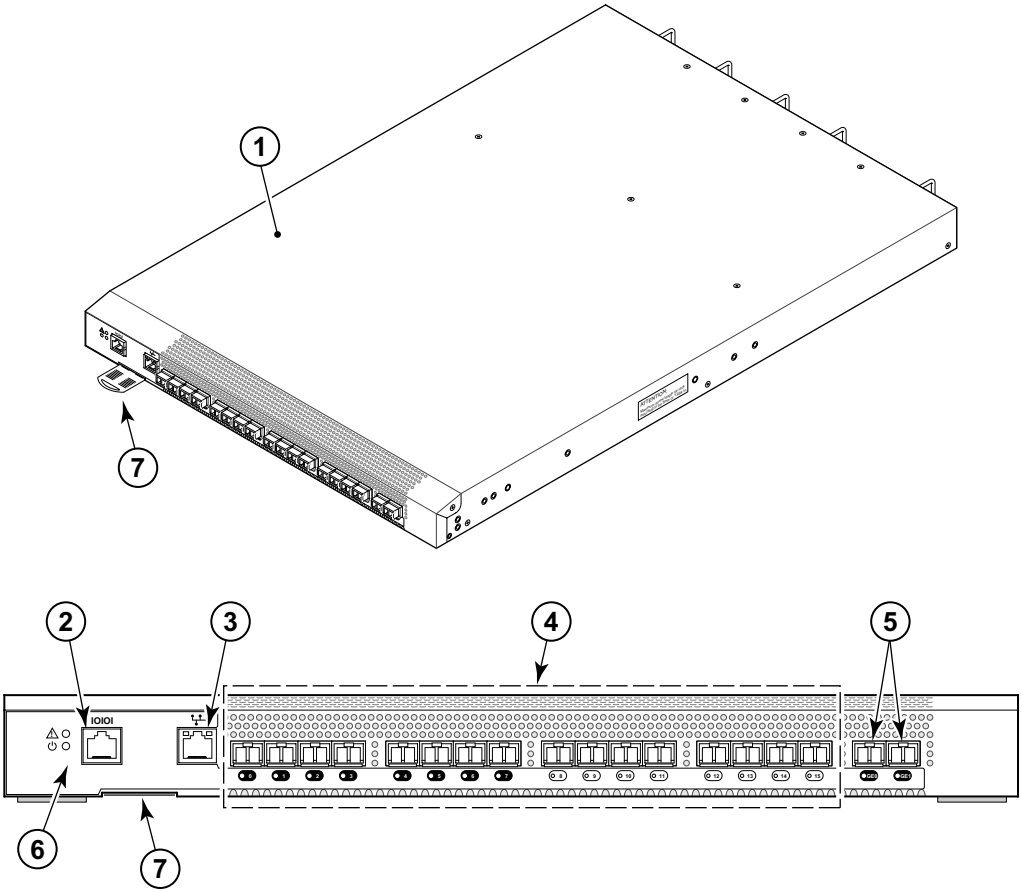


Figure 369 Port side view of the MP-7500B

1	MP-7500B	5	GbE ports (2)
2	Console Managment Port	6	System Status LED (top)
3	Ethernet Management Port		System Power LED (bottom)
4	Fibre Channel Ports (16)	7	IP Address pull out tab

The Fibre Channel ports are numbered from left to right, in eight-port groups, and are also numbered on the faceplate (see Figure 370).

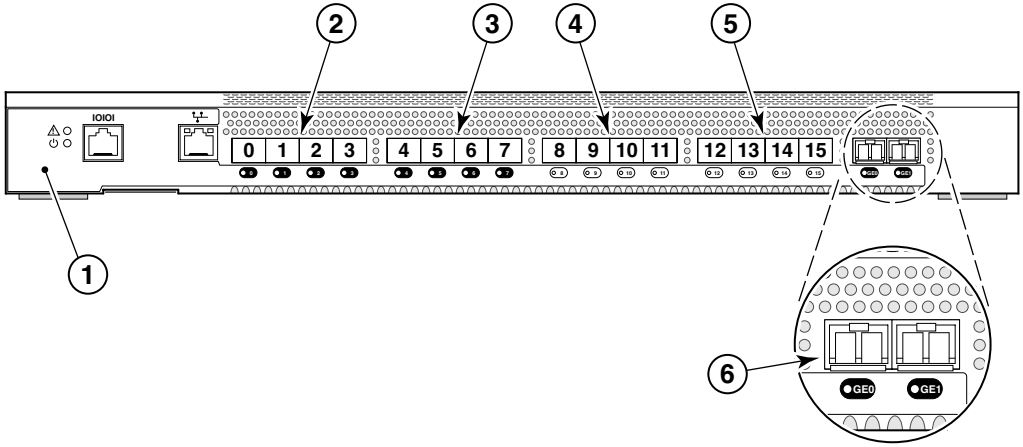


Figure 370 Port numbering in the MP-7500B

1	MP-7500B	4	Fibre Channel Ports 8 through 11
2	Fibre Channel Ports 0 through 3	5	Fibre Channel Ports 12 through 15
3	Fibre Channel Ports 4 through 7	6	GbE ports (2)

The port side of the MP-7500B also displays the system status LED, power status LED, and port status LEDs.

Nonport side of the MP-7500B

Figure 371 shows the nonport side of the MP-7500B, which contains the power supplies (including the AC power receptacle and AC power switch) and fans.

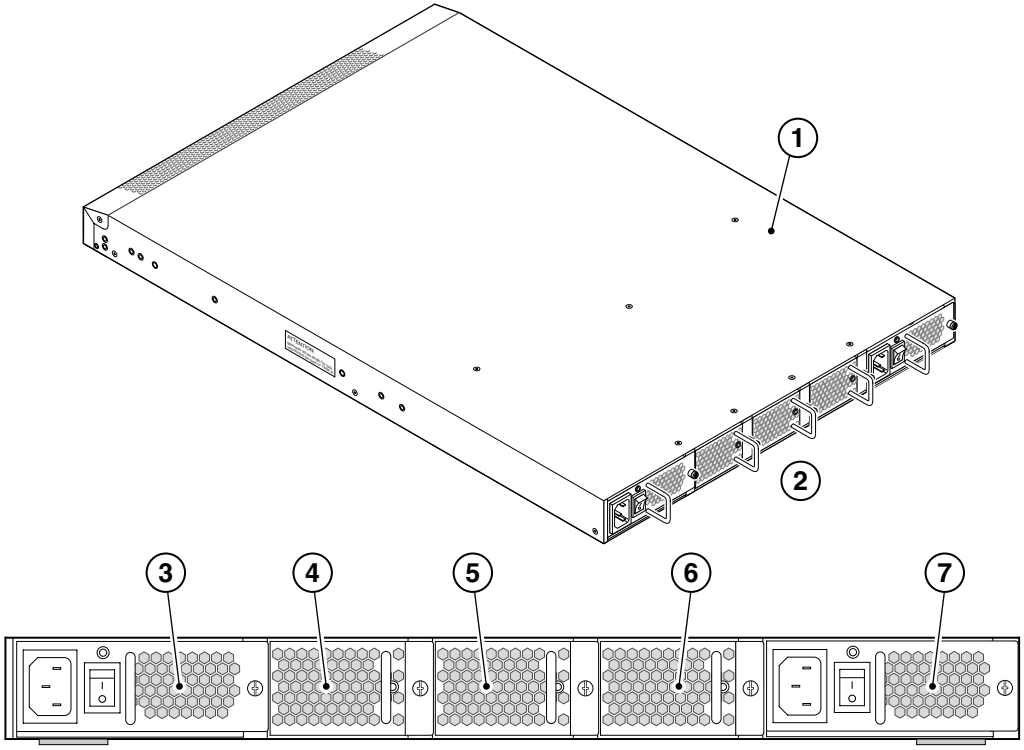


Figure 371 Nonport side of a MP-7500B

1	MP-7500B	5	Fan Assembly #2
2	Nonport Side of Switch	6	Fan Assembly #1
3	Power Supply #2	7	Power Supply #1
4	Fan Assembly #3		

Fibre Channel port specifications

The Fibre Channel ports in the MP-7500B are compatible with SWL, LWL, and ELWL SFP transceivers. The strength of the signal is determined by the type of transceiver in use.

The ports meet all required safety standards. For more information about these standards, see the “Regulatory Compliance” section in the *Connectrix B Series MP-7500B Hardware Reference Manual*.

The ports are capable of operating at 1-, 2-, or 4-Gbit / sec and are able to auto-negotiate to the maximum link speed.

GbE port specifications

The GbE ports in the MP-7500B are compatible with SWL, LWL, and ELWL SFP transceivers. The strength of the signal is determined by the type of transceiver in use.

The ports meet all required safety standards. For more information about these standards, see the “Regulatory Compliance” section in the *Connectrix B Series MP-7500B Hardware Reference Manual*.

The GbE ports are capable of operating at 1-Gbit / sec. The 8 virtual FCIP Fibre Channel links over each physical GbE connection share this bandwidth.

Serial port specifications

The serial port is located on the port side of the switch. The MP-7500B uses an RJ-45 connector for serial port. An RJ-45 to DB9 adaptor is also provided with the switch.

Note: To protect the serial port from damage, keep the cover on the port when not in use.

The serial port can be used to connect to a workstation to configure the switch IP address before connecting the switch to a fabric or IP network. The serial port’s parameters are fixed at 9600 baud, 8 data bits, and no parity, with flow control set to None.

Table 52 lists the serial cable pinouts.

Table 52 Serial cable pinouts

PIN	Signal	Description
1	Not supported	NA
2	Not supported	NA
3	UART1_TXD	Transmit data
4	GND	Logic ground
5	GND	Logic ground

Table 52 Serial cable pinouts (continued)

PIN	Signal	Description
6	UART1_RXD	Receive data
7	Not supported	NA
8	Not supported	NA

Supported features

The section includes supported software features.

IMPORTANT

References in the Fabric OS 5.2.1 documentation set specific to features/functionality available on the PB-48K-18i blade, such as FCR/FCIP functionality and port configurations, also apply to the MP-7500B.

The MP-7500B supports the following features that are included with the firmware:

- Web Tools
- Fibre Channel Routing Services
- Zoning
- Full Fabric

The following optional software is available with the purchase of a specific license key:

- FC over IP SAN Extension (FCIP)
- Performance Monitoring
- Fabric Watch
- Extended Fabric
- Trunking

For information on these features, see the *Connectrix B Series Fabric OS Administrator's Guide.*

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architechture

Nonblocking shared-memory switch.

Performance

Performance features include:

Fibre Channel:

- 1.063 Gb/s line speed, full duplex
- 2.125 Gb/s line speed, full duplex
- 4.250 Gb/sec line speed, full duplex
- Autosensing of 1 Gb/s, 2 Gb/s , and 4 Gb/s port speeds
- Optionally programmable to fixed port speed
- Speed matching between 1 Gb/s, 2 Gb/s, and 4 Gb/s ports

Ethernet:

- 10/100/1000 Mb support

Port types

- Self-discovery based on switch type (U_Port)

The GbE ports can be:

- VE_Ports
- VEX_Ports

The Fibre Channel ports can be:

- E_Ports
- EX_Ports
- FL_Ports
- F_Ports

Dual-mode ports

- 16 ports, Fibre Channel (E, F) and Gigabit Ethernet

Media types

- Small form-factor plugable (SFP) laser
- Shortwave up to 500 m (1,640 feet)
- Long wave up to 10 km (6.2 miles)

- Extended long wave up to 25 km (15.5 miles)

Maximum size frame

- 2112-byte payload FC
- 2250-byte payload Gigabit Ethernet
- 2048-byte payload for Fibre Channel routed networks

Aggregate switch I/O bandwidth

A total 132 Gbit/sec:

- 128 Gbit/sec if all 16 FC ports are running at 4 Gbit/sec, full duplex
- 4 Gbit/sec if both GbE ports are running at 1 Gbit/sec, full duplex

Fabric latency

- < 8 microseconds (FC-to-FC routed traffic)
- 30 microseconds (FCIP)

Class of service

- Class 2
- Class 3

Fabric services

- Simple Name Server
- Registered State Change Notification (RSCN)
- Fibre Channel Routing

Optional fabric services include:

- Zoning
- Exchange-Based Trunking
- Web Tools-AP Edition

Optional service includes:

- FCIP Tunneling

Switch interoperability

- Connectrix B DS-8B and DS-16B running Fabric OS v2.6.1 or higher.
- Connectrix B DS-8B2 and DS-16B2 running Fabric OS v3.1 or higher.

- Connectrix B DS-32B2 and ED-12000B running Fabric OS v4.1 or higher.
- Connectrix B DS-8B3, DS-16B3, and ED-24000B running Fabric OS v4.2 or higher.
- Connectrix B DS-220B and ED-48000B running Fabric OS v5.0.1or higher.
- SilkWorm 2xxx running Fabric OS V2.6.1 or higher.
- SilkWorm 3200 and 3800 running Fabric OS v3.1 or higher.
- SilkWorm 3900 and 12000 running Fabric OS v4.1 or higher.
- SilkWorm 3250, 3850, and 24000 running Fabric OS v4.2 or higher.
- SilkWorm 200E and 48000 running Fabric OS v5.x or higher.

Note: Refer to the E*MC Support Matrix* for the most up-to-date information on supported firmwares and EOL listings.

Options

- Redundant power supply
- SFP media

Management

Management software supported

- Telnet
- SNMP (FE MIB, FC Management MIB)
- Web Tools (optional)
- Fabric Manager (optional)
- RADIUS

Mangement access

- 10 / 100 Ethernet port (RJ-45)
- Serial port (RJ-45)
- In-band via Management Server

Diagnostics

- The MP-7500B switch runs POST by default each time it is powered on; it typically requires from 1 to 3 min to boot and complete POST.
- POST can be skipped after subsequent reboots by entering the **fastBoot** command. For more information about this command, see the *Connectrix B Series Fabric OS Administrator's Guide.*

Mechanical specifications

- Width = 429 mm (16.89 in.)
- Height = 1U = 42.44 mm (1.67 in.)
- Depth = 635mm (25.0 inches)
- Weight (with two power supplies and three fan assemblies installed) = 12.43 kg (27.4 lbs)

Environment

Table 53 lists the acceptable environmental ranges for both operating and nonoperating (such as during transportation or storage) conditions.

Table 53 Environmental requirements

Condition	Acceptable during operation	Acceptable during non-operation
Ambient Temperature	0° to 40° C (32° to 104° F)	-25° to 70° C (-13° to 158° F)
Humidity	20% to 85% RH non-condensing, at 40° C (104° F) , with maximum gradient of 10% per hour	10% to 85% RH non-condensing, at 70° C (158° F)
Altitude	0 to 3 km (10,000 ft) above sea level	0 to 12 km (40,000 ft) above sea level
Shock	20 G, 6 ms, half-sine wave	15 G, 12-18 milliseconds, trapezoid
Vibration	0.5 G, 5-500 Hz	2.0 G, 5-500 Hz
Air Flow	31 cu ft/min (52.7 cu m/hr)	NA

Power

The power supplies are universal and capable of functioning worldwide without voltage jumpers or switches. They meet IEC 61000-4-5 surge voltage requirements and are autoranging in terms of accommodating input voltages and line frequencies. Each power supply has a built-in fan for cooling, pushing air towards the port side of the switch.

Table 57 lists the power supply specifications for the MP-7500B switch.

Table 54 Power supply specifications

Specification	Value
Outlet	The outlet must be a correctly wired, primary with earth ground
Maximum output	300 Watts
System power consumption	240W Max, 215W typical
Input voltage	85-264 VAC, Universal
Input line frequency	47 - 63 Hz
Harmonic distortion	Active power factor correction
BTU rating at 80% efficiency	240W X 3.412 BTU/hr/Watts = 820 BTU/Hr
Inrush current	Maximum of 15 amps for period between 10-150 ms at 50 degrees C (122 degrees F), hot or cold start
Input line protection	Fused in both hot and neutral lines, using independent fuses

Standards compliance

This release conforms to the Fibre Channel Standards in a manner consistent with accepted engineering practices and procedures. In certain cases, Brocade might add proprietary supplemental functions to those specified in the standards. For the list of standards conformance, visit this Brocade website:

http://www.brocade.com/sanstandards

General specifications

Table 55 lists the general specifications for the MP-7500B switch.

Table 55 General specifications

Specification	Description
EMC (electromagnetic compatibility)	Emissions An operating MP-7500B conforms to the emissions requirements specified by the following regulations: ◆ FCC Rules & Regulations, Part 15 subpart B, Class A ◆ CSA C108.8 Class A ◆ VCCI Class A ITE ◆ CISPR 22 Class A ◆ EN55022 Class A ◆ Korean EMC Requirements ◆ BSMI Standard CNS 13438 ◆ EMC Directive 89/336/EEC ◆ EN50082-2/EN55024: 1998 ◆ ICES-003 Class A Immunity ◆ IEC 61000-4-2 Severity Level 3 for Electrostatic Discharge ◆ IEC 61000-4-3 Severity Level 3 for Radiated Fields ◆ IEC 61000-4-4 Severity Level 3 for Fast Transients ◆ IEC 61000-4-5 Severity Level 3 for Surge Voltage ◆ IEC 61000-4-6 Conducted Emissions ◆ IEC 61000-4-11 Voltage Variations ◆ EN 61000-4-12 Oscillatory Waves Immunity ◆ EN 61000-3-2 Limits for Harmonic Current Emissions ◆ EN 61000-3-3 JEIDA
System processor	PowerPC 440GX, 800 MHz CPU
ANSI Fibre Channel protocol	FC-PH (Fibre Channel Physical and Signalling Interface standard)
Fabric initialization	Complies with FC-SW-3 Rev. 6.6
FC-IP (IP over Fibre Channel)	Complies with FC-IP 2.3 of FCA profile

Facility requirements

To ensure correct operation of the switch, the facility where the switch is in use must meet the following requirements:

- Electrical:
  - Primary AC input 100-240 VAC (switch autosenses input voltage), 47-63 Hz.
  - Correctly wired primary outlet, protected by a circuit breaker and grounded in accordance with local electrical codes.
  - Adequate supply circuit, line fusing, and wire size, as specified by the electrical rating on the switch nameplate.
  - Electrical interference must be less than the levels stated in the standards listed in Table 55 on page 1031, under "EMC (electromagnetic compatibility), Immunity."
- Thermal:
  - Air flows from the non-port side to the port side. A minimum air flow of 56.5 cu ft / min (96 cu m / hr) available in the immediate vicinity of the switch.
  - Ambient air temperature must not exceed 40° C (104° F) while the switch is operating.
- Environmental: Specifications are listed in Table 53 on page 1029.
- Cabinet:
  - Cabinet space of one rack unit in a 19-in. cabinet.
  - All equipment in cabinet must be grounded through a reliable branch circuit connection.
  - The additional weight of the switch must not exceed the cabinet's weight limits.
  - The cabinet must be secured to ensure stability in case of unexpected movement.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-5000B

This section introduces the components of the the EMC Connectrix DS-5000B.

Description

The DS-5000B switch is a 1U Fibre Channel switch with 32 fixed Fibre Channel SFP ports that supports link speeds up to 1, 2, or 4 Gbit/sec. It includes the Fabric Operating System and is compatible with the entire EMC Connectrix B product family. The DS-5000B can operate in a fabric containing multiple switches or independently.

The switch's enclosure has forced-air cooling, with the fans pushing the air from the rear part intake through the enclosure, and exhausting across the SFP devices through venting holes in the front panel (port side). The SFP media and integrated power supply/fan assembly Field Replaceable Units (FRUs) are hot-swappable so they can be removed and replaced without interrupting the system power.

Power supply/fan assembly units are the only FRUs in the DS-5000B. There are two power supply/fan assembly units in the DS-5000B. They are hot-swappable and redundant, and capable of functioning universally without voltage jumpers or switches. The FRU units are identical and interchangeable. The front panel of the FRUs has a status LED to indicate status of the unit.

The DS-5000B has three types of memory devices:

- Boot flash: 4 MB
- Compact flash: 1 GB
- Main memory (SDRAM): 256 MB

More information for the DS-5000B can be found in the *Connectrix B Series DS-5000B Hardware Reference Manual* available on Powerlink.

Key features of the DS-5000B

The DS-5000B switch includes the following features:

- Cabinet-mountable 1U chassis designed to be mounted in a 19-inch cabinet space, with forced-air cooling that flows from the nonport side of the switch to the port side.
- 32 Fibre Channel ports, compatible with short wavelength (SWL), long wavelength (LWL), and extended long wavelength (ELWL) SFP transceivers.
- One RS-232 serial port on the port side of the switch.

- One IEEE-compliant RJ-45 connector on the port side of the switch for use with 10/100 MB/sec Ethernet.
- 32 port LEDs, 1 switch power LED, 1 switch status LED, 2 Ethernet LEDs, and 2 power supply LEDs.
- Two universal AC input and redundant power supplies with AC switches and built-in fans.

DS-5000B port side view

Figure 372 shows the port side of the DS-5000B.

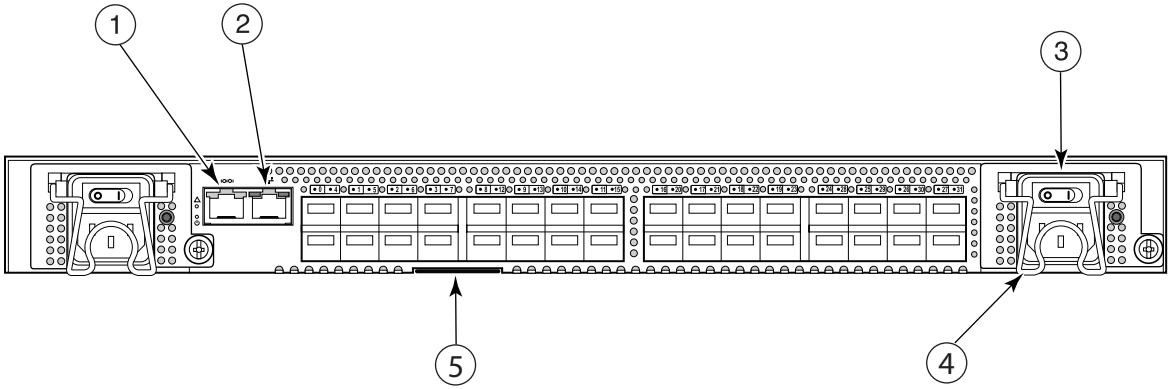


Figure 372 Port side view of the DS-5000B

1 System console port
2 System Ethernet port
3 Power supply/fan assembly Field Replaceable Unit (2x)
4 Power cord retainer (2x)
5 Switch ID pull out tab

On the port side of the unit there are two port connections:

- **Ethernet Port**: The DS-5000B provides a fully IEEE-compliant 10/100BaseT Ethernet port for switch management console interface. When a device is connected to the port, both ends negotiate to determine the optimal speed. The DS-5000B adopts a 1x2 RJ45 connector to provide Ethernet and serial ports to the outside. The Ethernet connection uses one of two RJ-45 ports. There are two integrated visible LEDs for Ethernet port (see Figure 372). One indicates transmit/receive activity and one

indicates speed (10 Mbps or 100 Mbps). The TCP/IP address for the port can be configured from the serial port or directly from the Ethernet port itself.

- **Serial Port**: An RS-232 DTE terminal port is provided on the DS-5000B switch. The serial console uses the other RJ45 port in the 1x2 RJ45 connector. The serial port parameters are fixed at 9600 baud, 8 data bits, no parity and no hardware flow control (except during boot up for the console port). This connector is for initial IP address configuration and for recovery of the switch to its factory default settings should Flash memory contents be lost. The serial port connection is not intended for performance of normal administration/maintenance functions.
  The Fibre Optic cables, Ethernet cables and Serial port cables connect in to the port side of the switch. AC power input cables and the power supply/fan assembly FRUs are inserted and removed from the port side of the switch.

The DS-5000B can be mounted in a 1U 19-inch Electronic Industries Association (EIA) rack, with a height of 1U. Because of the shallow depth, no rail kits are required for a rack mount, however the switch can be installed using the optional rack mount kit. The DS-5000B can also be used in a tabletop configuration.

The Fibre Channel ports are numbered from left to right, in eight-port groups, and are also numbered on the faceplate (see Figure 373).

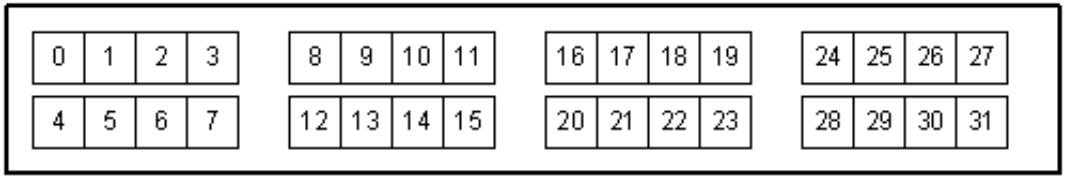


Figure 373 Port numbering in the DS-5000B switch

Note: ISL Trunking is an optionally licensed software that allows you to create trunking groups of ISLs between adjacent switches. For more information about ISL Trunking, refer to the *EMC Connectrix B Series Fabric OS Administrator's Guide*.

The port side of the DS-5000B switch also displays the system status LED, power status LED, and port status LEDs.

Supported features

The DS-5000B is shipped with the following features pre-installed:

- Web Tools
- Zoning
- Full Fabric

The DS-5000B supports theoptional software, which you can activate by purchasing a corresponding license key:

For more information , refer to the *EMC Connectrix B Fabric OS Administrator's Guide*.

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architecture

Nonblocking shared-memory switch.

Performance

Performance features include:

- 1.063 Gb/s line speed, full duplex
- 2.125 Gb/s line speed, full duplex
- 4.25 Gbit/s line speed (full duplex)
- Auto-sensing of 1 Gb/s, 2 Gb/s, and 4 Gb/s port speeds
- Optionally programmable to fixed port speed
- Speed matching between 1, 2, and 4 Gb/s ports

Port types

- F_Port
- FL_Port
- E_Port
- M_port (Mirror Port)
- Self-discovery based on switch type (U_Port)

Media types

- Hot-pluggable, industry-standard Small Form-factor Pluggable (SFP), LC connector
- Short-Wavelength Laser (SWL) up to 500 meters (1640 feet)
- Long-Wavelength Laser (LWL) up to 10 km (6.2 mi)
- Extended Long-Wavelength Laser (ELWL) up to 40 km (24.8 mi), and 80 km (49.6 mi) at 2 Gbit / sec

Note: Distance depends on fiber-optic cable and port speed, CWDM SFPs (8 lambdas).

Maximum size frame

- 2112-byte payload FC

Aggregate switch I/O bandwidth

- 256 Gbit / sec if all 32 ports are running at 4 Gbit / sec, full duplex

Fabric latency

- 800 nano-seconds port-to-port latency

Class of service

- Class 2
- Class 3
- Class F (interswitch frames)

Fabric services

Fabric services include:

- Simple Name Server
- Registered State Change Notification (RSCN)
- Zoning
- Web Tools

Optional fabric services include:

- Fabric Watch
- Extended Fabrics
- ISL Trunking (available between devices running FOS)
- Performance Monitoring
- FICON CUP

Note: Performance Monitoring and FICON CUP are not available in M-EOS environments.

Switch interoperability]

- Connectrix B DS-8B and DS-16B running Fabric OS v2.6.1 or higher.
- Connectrix B DS-8B2 and DS-16B2 running Fabric OS v3.1 or higher.
- Connectrix B DS-32B2 and ED-12000B running Fabric OS v4.1 or higher.
- Connectrix B DS-8B3, DS-16B3, and ED-24000B running Fabric OS v4.2 or higher.
- Connectrix B DS-220B and ED-48000B running Fabric OS v5.0.1or higher.
- SilkWorm 2xxx running Fabric OS V2.6.1 or higher.
- SilkWorm 3200 and 3800 running Fabric OS v3.1 or higher.
- SilkWorm 3900 and 12000 running Fabric OS v4.1 or higher.
- SilkWorm 3250, 3850, and 24000 running Fabric OS v4.2 or higher.
- SilkWorm 200E and 48000 running Fabric OS v5.x or higher.
- Refer to the *EMC Support Matrix* for the most up-to-date information on supported firmwares and EOL listings.

Management

Management software supported

- Brocade EFCM
- EMC Connectrix Manager
- Fabric Manager (optional: FOS environments only)
- Web Tools
- Telnet
- SNMP v3 (FE MIB, FC Management MIB)
- SSH
- HTTPS / SSL
- RADIUS
- Third-party applications utilizing the SMI Agent

Mangement access

- 10/100 Ethernet port (RJ-45)
- Serial port (RS-232)
- In-band via Management Server

Diagnostics

- The switch performs POST by default each time it is powered on or rebooted or the system is reset. Boot time with POST is a minimum of three minutes.
- POST can be skipped after subsequent reboots by entering the **fastBoot** command. For more information about this command, refer to the *Connectrix B Fabric OS Command Reference Manual.*

Mechanical specifications

- Width = 428.75 mn (16.88 inches)
- Height = 1U = 43.5 mm (1.71 inches)
- Depth = 264 mm (10.39 inches)
- Weight (with two power supplies/fan assemblies installed, no SFPs) = 10.8 lbs (4.9 kg)

Environment

Table 56 lists the acceptable environmental ranges for both operating and nonoperating (such as during transportation or storage) conditions.

Table 56 Environmental requirements

Condition	Acceptable during operation	Acceptable during non-operation
Ambient Temperature	0° to 40° Celsius (32° to 104° Fahrenheit)	-25° to 70° Celsius (-13° to 158° Fahrenheit)
Humidity	10% to 85% RH non-condensing, at 40° Celsius (104° Fahrenheit) , with maximum gradient of 10% per hour	10% to 90% RH non-condensing, at 70° Celsius (158° Fahrenheit)
Altitude	3,000 meters (9,842 feet) above sea level	0 to 12 kilometers (39,370 feet) above sea level
Shock	20 G, 6 ms, half-sine wave	Half sine, 33G 11ms, 3/eg Axis
Vibration	0.5 G sine, 0.4 grms random, 5-500 Hz	2.0 G sine, 1.1 grms random 5-500 Hz
Air flow	High speed, 9300 RPM, 20.8 cubic feet/minute Low speed, 7200 RPM, 15.4 cubic feet/minute	None required.

Power

The power supplies are universal and capable of functioning worldwide without voltage jumpers or switches. They meet IEC 61000-4-5 surge voltage requirements and are autoranging in terms of accommodating input voltages and line frequencies. Each power supply has a built-in fan for cooling, pushing air towards the port side of the switch. Table 57 lists the power supply specifications for the DS-5000B.

Table 57 Power supply specifications

Specification	Value
Outlet	The outlet must be a correctly wired, primary with earth ground
Maximum output	70 Watts
System power consumption	70 Watts maximum, 60 Watts typical
Input voltage	90 - 264 VAC, Universal
Input line frequency	47 - 63 Hz
BTU rating at 80% efficiency	56 Watts / 0.8 X 3.412 BTU/hr/Watts = 239 BTU/hr
Inrush current	Maximum of 15 amps for period between 10-150 ms at 50 degrees Celsius (122 degrees Fahrenheit), hot or cold start

Standards compliance

This release conforms to the Fibre Channel Standards in a manner consistent with accepted engineering practices and procedures. In certain cases, Brocade might add proprietary supplemental functions to those specified in the standards. For the list of standards conformance, visit this Brocade website:

http://www.brocade.com/sanstandards

General specifications

Table 58 lists the general specifications for the DS-5000B.

Table 58 General specifications

Specification	Description
EMC (electromagnetic compatibility)	Emissions: Immunity IEC 61000-4-2 Severity Level 3 for Electrostatic Discharge IEC 61000-4-3 Severity Level 3 for Radiated Fields IEC 61000-4-4 Severity Level 3 for Fast Transients IEC 61000-4-5 Severity Level 3 for Surge Voltage IEC 61000-4-6 Conducted Emissions IEC 61000-4-11 Voltage Variations EN 61000-4-12 Oscillatory Waves Immunity EN 61000-3-2 Limits for Harmonic Current Emissions EN 61000-3-3 JEIDA
System processor	PowerPC 440GP, 333 MHz CPU
ANSI Fibre Channel protocol	FC-PH (Fibre Channel Physical and Signalling Interface standard)
Fabric initialization	Complies with FC-SW-3 Rev. 6.6
FC-IP (IP over Fibre Channel)	Complies with FC-IP 2.3 of FCA profile

Facility requirements

To ensure correct operation of the switch, the facility where the switch is in use must meet the following requirements:

- Electrical:
  - Primary AC input 90-264 VAC (switch autosenses input voltage), 47-63 Hz.
  - Correctly wired primary outlet, protected by a circuit breaker and grounded in accordance with local electrical codes.
  - Adequate supply circuit, line fusing, and wire size, as specified by the electrical rating on the switch nameplate.
  - Electrical interference must be less than the levels stated in the standards listed in Table 58 on page 1041, under “Immunity.”
- Thermal:
  - Air flows from the non-port side to the port side. A minimum air flow of 47 cubic feet / minute (79.8 cubic meters / hour) available in the immediate vicinity of the switch.

- Ambient air temperature must not exceed 40° Celsius (104° Fahrenheit) while the switch is operating.
- Environmental: The specifications listed in Table 56 on page 1039
- Cabinet:
  - Cabinet space of one rack unit in a 19-inch cabinet.
  - All equipment in cabinet must be grounded through a reliable branch circuit connection.
  - The additional weight of the switch must not exceed the cabinet's weight limits.
  - The cabinet must be secured to ensure stability in case of unexpected movement.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-4900B

This section provides a description of the hardware and features for the EMC Connectrix DS-4900B. The information provided comes from Brocade data sheets and testing experience in the EMC IOP qualification labs. More information for the DS-4900B can be found in the *Connectrix B Series DS-4900B Hardware Reference Manual* available on Powerlink.

Description

The Connectrix DS-4900B is a Departmental class switch which can contain up to 64 concurrently active 4 Gb/s full-duplex ports in a single domain. It is based on the Brocade Silkworm 4900 Condor ASIC based architecture. The DS-4900B is a 64-port non-blocking switch housed in a 2U enclosure. The DS-4900B is ideal for consolidating existing 1 and 2 Gb/s SANs into a single high-availability switch to improve performance, reduce the ISL count, and simplify management. The DS-4900B can also serve as an edge switch in a core-edge topology.

The DS-4900B utilizes a Linux 2.4.x modified kernel to load the fabric OS version 5.1.x and later.

Note: For specific versions of supported firmware, as well as fabric topology constraints associated with the DS-4900B, refer to the E-Lab Navigator.

Key features of the DS-4900B

- Base configuration of 32 ports with the ability to increase port count 16 at a time with a Ports On Demand (POD) license.
- Non-disruptive firmware upgrades.
- Redundant and hot swappable power supplies and fans.
- Forward and backward compatibility with all Connectrix B series switches.
- Supports 1-, 2-, and 4-Gb/s auto-sensing Fibre Channel ports. Trunking groups up to eight ports to create high performance 32-Gb/s ISL trunks between switches.
- Universal ports self-configure as E_Ports, F_Ports, or FL_Ports

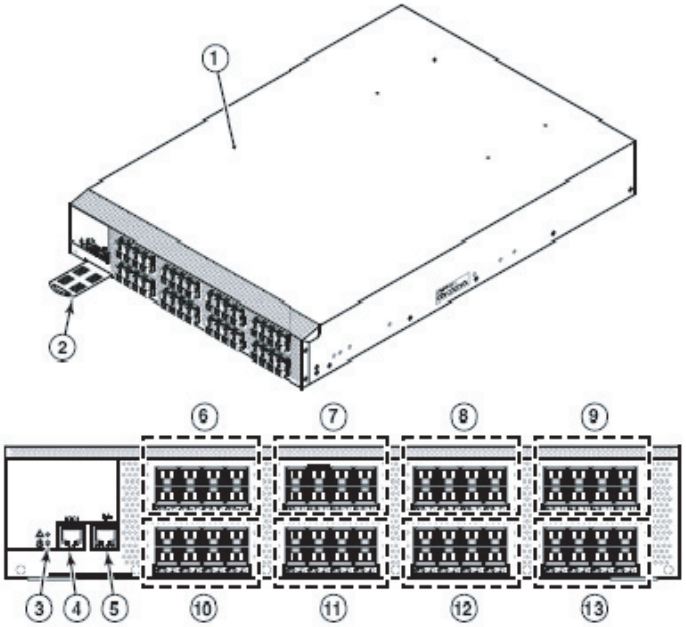


Figure 374 DS-4900B Port side view

Legend:

1	Connectrix DS-4900B	**8**	FC Ports 16 – 23
2	Switch ID Pull-out Tab	**9**	FC Ports 24 – 31
3	System Status LED (top) System Status LED (bottom)	**10**	FC Ports 32 – 39
4	Console Port	**11**	FC Ports 40 – 471
5	Ethernet Port	**12**	FC Ports 48 – 55
6	FC Ports 0 – 7	**13**	FC Ports 56 – 63
7	FC Ports 8 – 15		

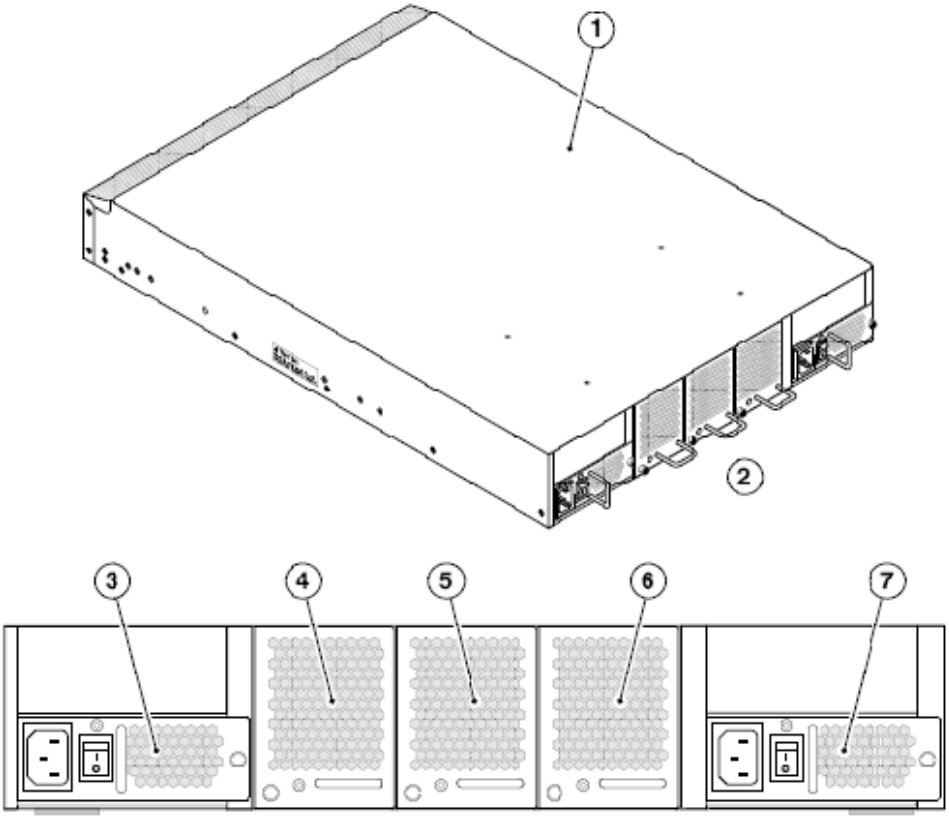


Figure 375 DS-4900B Non-port side view

Legend:

1	Connectrix DS-4900B	**5**	Fan Assembly #2
2	Nonport Side of Switch	**6**	Fan Assembly #1
3	Power Supply #2)	**7**	Power Supply #1
4	Fan Assembly #3		

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architecture

Performance

- 1.063 Gb/s line speed, full duplex
- 2.125 Gb/s line speed, full duplex
- 4.25 Gb/s line speed, full duplex
- Auto-sensing of 1, 2, and 4 Gb/s port speeds
- Optionally programmable to fixed port speed
- Speed-matching between 1, 2, and 4 Gb/s ports

Extended ISL Mode Buffer Allocation

Mode	Buffer allocation			Distance @ 1 Gb/s	Distance @ 2 Gb/s	Distance @ 4 Gb/s	Extended Fabric license required?
	1Gb/s	2Gb/s	4Gb/s				
L0	5(26)[a]	5(26)	5(26)	10 km	5km	2 km	No
LE	11	16	26	10 km	10 km	10 km	No
L0.5	18	31	56	25 km	25 km	25 km	Yes
L1	31	56	106	50 km	50 km	50 km	Yes
L2	56	106	206	100 km	100 km	100 km	Yes
LD[b]	Auto	Auto	Auto	Auto (500 km max)	Auto (250 km max)	Auto (100 km max)	Yes
LS[c]	User defined	User defined	User defined	User defined (500 km max)	User defined (250 km max)	User defined (100 km max)	Yes

a. For each data channel (in this case, there are 4) there are 5 credits, plus 6 extra credits.

b. The dynamic long-distance mode (LD) automatically configures the number of buffer credits required, based on the actual link distance.

c. The static long-distance mode (LS) allocates the number of buffer credits based on the user-specified distance.

Fibre Channel ports

- Up to 64 Fibre Channel ports which are housed in a 2U enclosure.

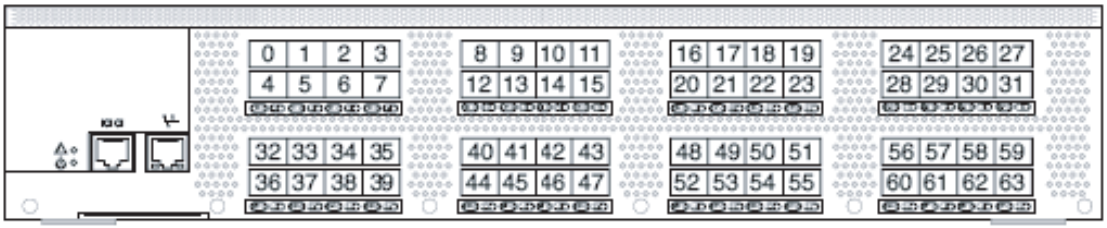
Port types

- Universal ports (U_Port), which self-configure as E_Ports, F_Ports, or FL_Ports; G_Port functionality is determined by the type of attached device
- FL_Port Port Module support for SFP optical transceivers, supporting any combination of SWL and LWL optical media.
- F_Port
- E_Port

Port numbering

The Connectrix DS-4900B uses the following port numbering schemes:

- Ports are numbered from left to right, in eight-port groups.
- Port numbers are also located on the faceplate.



ISL Trunking

- Up to eight 4.25 Gb/s ports per ISL trunk
- Up to 32 Gb/s per ISL trunk
- ISL trunking at 2 Gb/s for compatibility with prior B series Connectrix products.

Aggregate bandwidth

- 512 Gb/s

Switch latency

- ~700 nanoseconds within a locally switch group and <2.4 µsec with no contention, cut-through routing at 4 Gb/s between locally switched groups.

Oversubscription

- 64 port non-oversubscribed switch at 4 Gb

Maximum frame size

- 2112-byte payload

Frame buffers

- 584 buffers per switch
- Dynamically allocated up to 255 per port

Classes of service

- Class 2, Class 3, and Class F (inter-switch frames) Fibre Channel protocol support.

Fabric services

- Simple Name Server
- Registered State Change Notification (RSCN)
- Alias Server (multicast)
- Zoning
- Web Tools
- Fabric Watch
- Extended Fabrics
- ISL Trunking
- Performance Monitoring.

High Availability

- Redundant, hot-swappable components
  - Three fan assemblies
  - Power supplies.
  - Small Form-factor Pluggable (SFP) Short-Wavelength Laser (SWL) transceivers.
  - Small Form-factor Pluggable (SFP) Long-Wavelength Laser (LWL) transceivers.
- Redundant power supply and fan assembly
  - The DS-4900B requires a minimum of two operating blower assemblies at al times. To ensure continuous adequate cooling, maintain three operating blower assemblies at all times except for the brief period when replacing a blower assembly.
- Enhanced data integrity on all data paths
- Fabric Shortest Path First (FSPF) rerouting around failed links

- Integration with Simple Network Management Protocol (SNMP) managers
- Non-disruptive “hot” software code loads and activation
- Easy configuration, save, and restore
- Two hot swappable 300W power supply modules providing N + 1 redundancy

Cooling

- Three fan assemblies containing two fans each. The fans have two speeds, which are set automatically and cannot be modified. They default to high speed upon boot, then switch to low speed as Fabric OS comes online, returning to high speed only as required.
- Non-port side to port-side airflow

Management

Management software supported

- EMC ControlCenter
- Brocade Fabric Manager (v5.1.x and greater)
- Embedded web server to provide single-point management using the Web Tools GUI.
- Centralized configuration and management of fabric using Client / Server architecture and CLI.
- Support for 10 / 100 Mb Ethernet connections for out-of-band management.
- SNMP (FE MIB, FC Management MIB)
- Telnet/SSH

Management access

- 10 / 100Mb Ethernet port (RJ-45)
- Serial port (RJ-45)

Diagnostics

- Post and embedded online/offline diagnostics

Mechanical specifications

- Rear panel-to-door airflow
- Width = 42.87 cm (16.88 in)
- Height = 8.60 cm (3.39 in) for 2U
- Depth = 61.0 cm (24.02 in)

- Weight = 13.7 kg (30.2 lb) with dual power supply, no SFP

Environment

- Temperature
  - Operating: 0º C to 40º C (32º F to 104º F)
  - Non-operating: -25º C to 70º C (-13º F to 158º F)
- Humidity
  - Operating: 20% to 85% non-condensing at 40º C (104º F)
  - Non-operating and storage (non-condensing): 0% to 93%
- Altitude
  - Up to 3000 meters (9800 feet)
- Shock
  - Operating: 20 G, 11 ms, half sine 1 G p-p, 5-500 Hz, 1 octave min
  - Non-operating: 33 G, 11ms, half sine 2.4G p-p, 5-500 Hz, 1 octave min
- Vibration
  - Operating: 20 G, 6 ms half-sine
  - Non-operating: 15 G, 12 to 18 ms trapezoid

Power

- Maximum output
  - 300 Watts

    Note: 256-port configuration requires a maximum of 750 Volt-Amps.
- In-rush current
  - Maximum of 15 Amps for period between 10 to 150 ms at 50ºC (122ºF), hot or cold start
- System power consumption
  - Nominal system draw 175 Watts
- Nominal input voltage
  - 90 to 264 VAC, 47 to 63 Hz
- Minimum input voltage
  - 90 VAC
- Maximum input voltage

- 264 VAC
- Frequency
  - 47 to 63 Hz
- BTU rating (80% efficiency)
  - 220 Watts / 0.8 X 3.412 BTU / hr / Watts = 756 BTU / hr

Protocols

- FC
- FICON

Standards compliance

This release conforms to the Fibre Channel Standards in a manner consistent with accepted engineering practices and procedures. In certain cases, Brocade might add proprietary supplemental functions to those specified in the standards. For the list of standards conformance, visit this Brocade website:

http://www.brocade.com/sanstandards

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-16B3

The DS-16B3 Departmental Switch offers 16 universal ports capable of auto-sensing 1 Gb / s and 2 Gb / s.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-16B3.

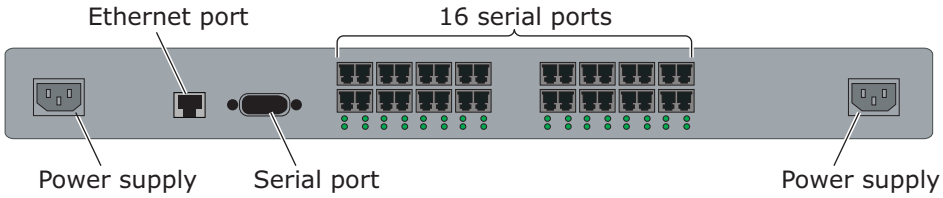


Figure 376 Connectrix DS-16B3

Connectivity

Connectivity features include:

- 16 ports, universal (E_Port, F_Port, and FL_Port).
- ISL Trunking to other B series switches: Up to four ports per trunk; up to 8.5 Gb/s per ISL trunk.
- Data traffic types: unicast, multicast (256 groups), and broadcast.
- Media types: Hot-pluggable, industry-standard Small Form-factor Pluggable (SFP) LC connector; Short-Wavelength Laser (SWL) up to 500 m (1,640 ft); Long-Wavelength Laser (LWL) up to 10 km (6.2 mi). (Distance depends on fiber-optic cable and port speed.)
- Universal ports, which self-configure as E_Ports, F_Ports, or FL_Ports.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architecture

Performance

Performance features include:

- 2.125 Gb/s line speed, full duplex; full duplex to 200 MB/s data rate per Fibre Channel port.

Management

Management features include:

- Centralized configuration and management of fabric using Client/Server architecture (Web Tools) and CLI.
- In-band or out-of-band management using Telnet or Web Tools.
- Management software: Telnet; SNMP (FE MIB, FC Management MIB, Brocade switch-specific MIBs); Web Tools; EMC Fabric Manager.
- Embedded Apache Web server to provide a single-point management using the Web Tools GUI. (The switch can also be managed using Telnet CLI or EMC Fabric Manager.)

- Interactive and non-interactive (nondisruptive) firmware upgrades.
- Support for 10/100 Mb Ethernet connections for out-of-band management.
- 10/100 Ethernet (RJ-45); serial port (RS-232); in-band through Management Server.
- Persistent centralized logging: Event logs, Session logs, and SNMP support.

Mechanical specifications

Mechanical specifications include:

- Width = 42.9 cm (16.9 in)
- Height = 4.3 cm (1.7 in)
- Depth = 30.7 cm (12.1 in)
- Weight = 4.2 kg (9.4lb) with dual power supplies, no media

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Power-on self test (POST) and embedded online/offline diagnostics.
- Automatic error notification using Fabric Watch utility, syslogd, and Web Tools.
- MIB compliance: MIB system group, Interface group, and SNMP group integration with higher layer managers, Fabric Element MIB, Vendor-specific MIBs, Standard Generic traps, Enterprise-specific traps.
- Fan, temperature, and port status are easily viewable from GUI Web Tools management tool.
- Dynamic statistics display performance data for each online port using GUI or Telnet CLI.
- Persistent logging through Event logs and Errorlog enhance troubleshooting to provide rapid error source identification.
- Automatic communication to central support facility through modem connection.
- Pre-emptive hardware health checking and recovery.
- Power-on self test (POST) and commands performs diagnostic tests on hardware.

- Error detection and fault isolation (EDFI), such as cyclic redundancy checking (CRC), parity checking, checksum, and illegal address checking.
- Environmental monitoring and control.
- Modem for remote diagnostics and status monitoring through serial port for dial-in support.
- Watchdog timers monitor processes, providing recovery mechanism in event of failure.
- Predictive diagnostics analysis through Fabric Watch and Error logging.
- Four field-replaceable fan units.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-8B3

The DS-8B3 Departmental Switch offers eight universal ports capable of auto-sensing 1 Gb/s and 2 Gb/s.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-8B3.

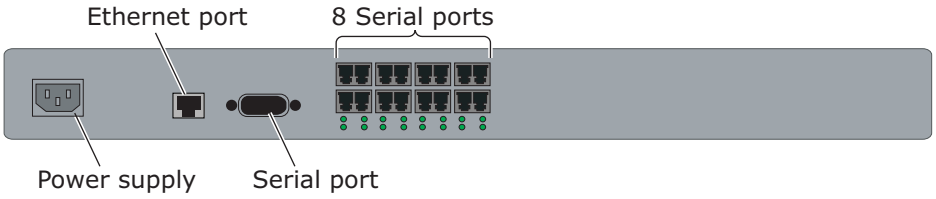


Figure 377 **Connectrix DS-8B3**

Connectivity

Connectivity features include:

- 8 ports, universal (E_Port, F_Port, and FL_Port).
- ISL Trunking to other B series switches: Up to four ports per trunk; up to 8.5 Gb/s per ISL trunk.
- Data traffic types: unicast, multicast (256 groups), and broadcast.

- Media types: Hot-pluggable, industry-standard Small Form-factor Pluggable (SFP) LC connector; Short-Wavelength Laser (SWL) up to 500 m (1,640 ft); Long-Wavelength Laser (LWL) up to 10 km (6.2 mi). (Distance depends on fiber-optic cable and port speed.)
- Universal ports, which self-configure as E_Ports, F_Ports, or FL_Ports.
- Fabric services: Simple name server; Registered State Change Notification (RSCN); Alias Server (multicast); Brocade Zoning optional services: Fabric Watch, Extended Fabrics, ISL Trunking, Advanced Performance Monitoring, Secure Fabric OS, Remote Switch.

  Note: Some features require an additional license.

- Available BB_Credits (buffer-to-buffer credits):
  - 16 per F_Port
  - 26 per E_Ports
  - 27 for L1 mode
  - 60 for L2 mode

  Note: L1 mode requires the Extended Fabric License.

- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.
- Support for wide range of value-added applications, including:
  - Fabric Access
  - Fabric Watch
  - Advanced Web Tools
  - Zoning
- Aggregate bandwidth 32 Gb/s end to end.
- 2112 byte maximum frame size.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

System architecture

Performance

Performance features include:

- 2.125 Gb/s line speed, full duplex; full duplex to 200 MB/s data rate per Fibre Channel port.

Management

Management features include:

- Centralized configuration and management of fabric using Client/Server architecture (Web Tools) and CLI.
- In-band or out-of-band management using Telnet or Web Tools.
- Management software: Telnet; SNMP (FE MIB, FC Management MIB, Brocade switch-specific MIBs), Web Tools, EMC Fabric Manager.
- Embedded Apache Web server to provide a single-point management using the Web Tools GUI. (The switch can also be managed using Telnet CLI or EMC Fabric Manager.)
- Interactive and non-interactive (nondisruptive) firmware upgrades.
- Support for 10/100 Mb Ethernet connections for out-of-band management.
- 10/100 Ethernet (RJ-45); serial port (RS-232); in-band through Management Server.
- Persistent centralized logging: Event logs, Session logs, and SNMP support.

Mechanical specifications

Mechanical specifications include:

- Width = 42.9 cm (16.9 in)
- Height = 4.3 cm (1.7 in)
- Depth = 30.7 cm (12.1 in)
- Weight = 4.2 kg (9.4lb) with no media

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Power-on self test (POST) and embedded online/offline diagnostics.

- Automatic error notification using Fabric Watch utility, syslogd, and Web Tools.
- MIB compliance: MIB system group, Interface group, and SNMP group integration with higher layer managers, Fabric Element MIB, Vendor-specific MIBs, Standard Generic traps, Enterprise-specific traps.
- Fan, temperature, and port status are easily viewable from GUI Web Tools management tool.
- Dynamic statistics display performance data for each online port through GUI or Telnet CLI.
- Persistent logging through Event logs and Errorlog enhance troubleshooting to provide rapid error source identification.
- Automatic communication to central support facility through modem connection.
- Pre-emptive hardware health checking and recovery.
- Power-on self test (POST) and commands that perform diagnostic tests on hardware.
- Error detection and fault isolation (EDFI), such as cyclic redundancy checking (CRC), parity checking, checksum, and illegal address checking.
- Environmental monitoring and control.
- Modem for remote diagnostics and status monitoring using serial port for dial-in support.
- Watchdog timers monitor processes, providing recovery mechanism in event of failure.
- Predictive diagnostics analysis through Fabric Watch and Error logging.
- Three field-replaceable fan units.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-32B2

The DS-32B2 Departmental Switch is a 32-port Fibre Channel gigabit switch that supports link speeds up to 2 Gb/s. The DS-32B2 is one rack unit (1 U) in height, offers features and functionality similar to the DS-16B2 (described under "Connectrix DS-16B3" on page 1051), and utilizes the same firmware as the ED-12000B (described under "Connectrix ED-12000B" on page 1015).

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-32B2.

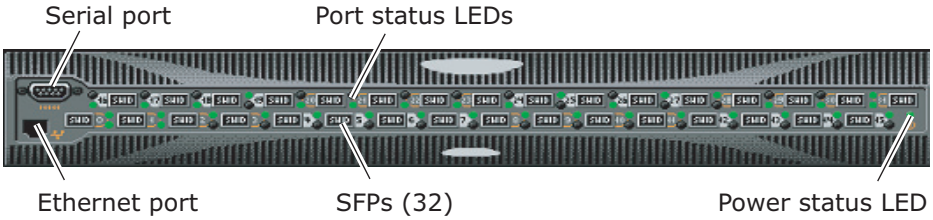


Figure 378 **Connectrix DS-32B2**

Connectivity

Connectivity features include:

- 32 ports available; auto-sensing of 1 Gb/s and 2 Gb/s port speeds; port types include: FL_Port, F_Port, and E_Port; self-discovery based on switch type (U_Port); optional port type control.
- Media: Small Form-Factor Pluggable SFP on each port enables connectivity.
- In-band or out-of-band management available through Telnet or Web-based Web Tools GUI.
- 16 buffer-to-buffer (BB_Credits) available per F_port, 26 buffer-to-buffer credits for E_ports, 27 for buffer-to-buffer credits for L1 mode, 60 buffer-to-buffer credits for L2 mode, which requires the Extended Fabric License.
- Support for Class 2, Class 3, Class F (inter-switch frames) protocols.
- ISL Trunking, up to four ports per trunk; up to 8.5 GB/s per ISL trunk.
- Laser support: Shortwave, longwave, and extended longwave.
- ISL to a B-series fabric of switches or directors.

- Full duplex to 200 MB/s data rate per Fibre Channel port.
- LC (Small Form Factor) switch port connectors.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

Management

The DS-32B2 can be configured and centrally managed using remote and local access methods. Fabric management can be performed in band or out of band using either Telnet, serial port, or Web-based server (Web Tools) by connecting to the 10/100BaseT Ethernet port on the switch.

Management features include:

- Centralized configuration and management of fabric using Web Tools or command line.
- In-band or out-of-band management available using Telnet or Web Tools.
- Embedded Apache Web Server to provide a single-point management using the Web Tools GUI. (The switch can also be managed via Telnet CLI.)
- Support for both interactive and non-interactive firmware upgrades.
- 10/100 Mb Ethernet connections for out-of-band management.
- Persistent centralized logging: event and session logs, as well as SNMP support.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Automatic error notification through Fabric Watch utility (sent using syslogd) or Web Tools.
- MIB compliance: MIB system group, Interface group, and SNMP group integration with higher layer managers, Fabric Element MIB, vendor-specific MIBs, Standard Generic traps, Enterprise-specific traps.

- Dual redundant power supplies.
- Hot-swappable fans, power supplies and Small Form Factor Plug (SFP) media.
- Easily viewed status of fan, temperature, and ports from Web Tools management tool.
- Dynamic statistics display, showing performance data for each online port through GUI or Telnet CLI.
- Persistent logging through Event logs and Error log, for enhanced troubleshooting to provide rapid error source identification.
- Automatic communication to central support facility through modem connection.
- Preemptive hardware health checking and recovery.
- Power-on self test (POST), which performs diagnostic tests on hardware.
- Error detection and fault isolation (EDFI), such as cyclic redundancy checking (CRC), parity checking, checksum, and illegal address checking.
- Environmental monitoring and control.
- Modem for remote diagnostics and status monitoring through serial port for dial-in support.
- Watchdog timers, which monitor processes, providing recovery mechanism in event of failure.
- Predictive diagnostics analysis through Fabric Watch and Error logging.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-16B2

The DS-16B2 (Brocade SilkWorm 3800) Departmental Switch is a 16-port Fibre Channel gigabit switch that supports link speeds up to 2 Gb / s. The DS-16B2 is compatible with the DS-16B / DS-8B product family, and can operate as a single-switch fabric or in a multi-switch fabric.

The DS-16B2 is one rack unit (1U) in height, has an air-cooled chassis, and can be set up as a stand-alone unit or mounted in a 19-inch rack.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-16B2.

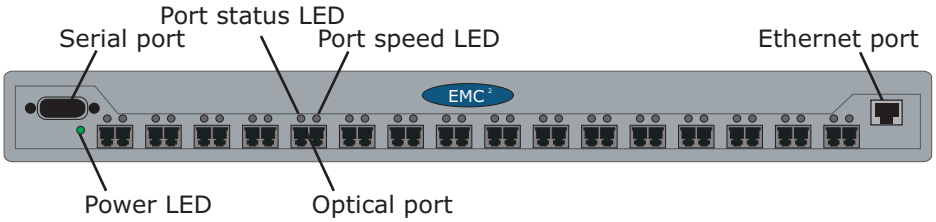


Figure 379 Connectrix DS-16B2

Connectivity

Connectivity features include:

- 16 optical ports, compatible with small form factor pluggable media (SFPs).
  - Switch is capable of 2 Gb/s.
  - Ports are self-configuring, and are capable of individually becoming a fabric-enabled port (F_Port) or an expansion port (E_Port).
  - Each port can automatically negotiate to the highest common speed of all devices connected to the port.
- Optional *trunking* capability between interconnected DS-16B2 ports.
  - Allows up to four ISL connections (E_Ports) between two DS-16B2 switches to be merged into one link.
  - Provides up to 8 Gb/s between DS-16B2 switches.
- One RS-232 serial port (DB-9 connector) on the front panel.
- One IEEE-compliant RJ-45 connector on the front panel for use with 10/100 Mb/s Ethernet.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

Management

The DS-16B2 can be managed using several remote and local access methods. Fabric management can be performed out of band using eithr ESNM, Telnet, or the Web by connecting to the 10/100BaseT Ethernet port on the switch. Additionally, Fabric Manager, a Client/Server-based application, can be installed on a PC with Ethernet access to the switch, and can be used to perform fabric management.

Management features include:

- Centralized configuration and management of fabric using Client/Server architecture
- Supported Clients: Windows 95, Windows 98, Windows NT 4.0, Windows 2000, Sun Solaris
- Out-of-band management available using ESN Manager, Telnet or Web-based tool
- Centralized management of switch
- 10/100 Mb Ethernet connections for out-of-band management
- Extensive centralized logging: event, audit, and session logs, as well as SNMP support

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic error notification using optional FabricWatch utility
  - Automatic error notification sent through **syslogd**
  - Automatic error notification using Web Tool or Fabric Manager
  - MIB compliance: MIB-II system group, Interface group, and SNMP group, Fabric Element MIB, vendor-specific MIBs, standard generic traps, enterprise-specific traps
- Dual redundant power supplies
- Highly integrated, reliable, multifunction (ASIC) devices throughout the switch

- Fan, temperature, and port status easily viewable from GUI management tool
- Dynamic statistics display performance data for each online port through GUI or telnet CLI
- Event logs for streamlined troubleshooting and rapid error source identification

Trunking

Note: For a complete description of trunking, refer to the *Departmental Switch Model DS-16B2 Interswitch Link (ISL) Trunking User Guide*, available on Powerlink.

The ISL trunking feature allows up to four ISL connections between two switches through an E_Port to merge logically into a single link and aggregate their throughput. For example, at 2 Gb/s speeds, a four-ISL trunk delivers an aggragated throughput of up to 8 Gb/s.

ISL trunking dynamically performs load sharing, at the frame level, across a set of available links between two adjacent switches to establish a trunking group. Ports that belong to a trunking group are called trunking ports. One port is used to assign traffic for the group, and is referred to as the trunking master.

The primary task of ISL trunking is to route data and edge switches that aggregate connections to servers and storage.

Benefits of trunking technology

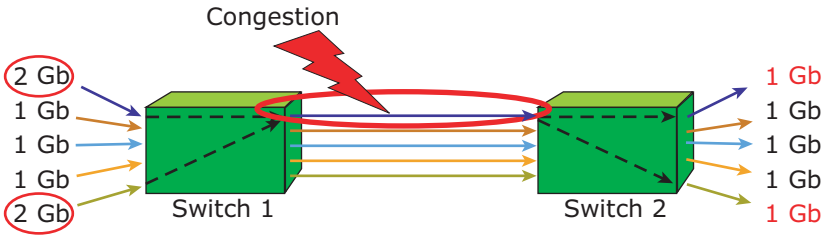


Figure 380 **ISLs with no trunking**

With no trunking:

- Single ISL is capable of 2 Gb/s maximum.
- During fabric build process, multiple nodes may be assigned the same ISL.

- In this case, two nodes pushing 4 Gb/s combined are competing for a single 2 Gb/s pipe.
- Congestion results.

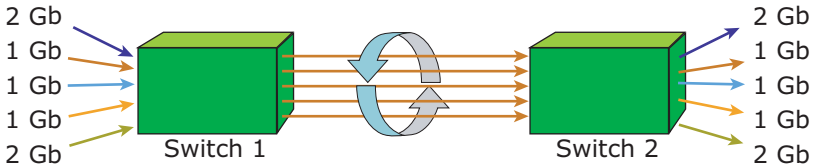


Figure 381 **ISLs with trunking**

With trunking:

- Trunking all four ISL creates one logical ISL.
- Fibre Channel frames from initiators and targets are sent across the first available ISL.
- Trunking relieves the congestion.
- Trunking utilizes bandwidth more efficiently.

Trunking restrictions

Note these restrictions:

- Ports must be capable of being configured as E_Ports.
- Ports must reside in contiguous four-port groups; for example: 0 through 3, 4 through 7, 8 through 11, 12 through 15.
- Ports must be set to run at the 2 Gb speed.
- The cable used in a trunking group should not be longer than 300 meters, and individual cable lengths in a trunking group must be within 30 meters of each other.

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-16B and DS-8B

The DS-16B (16-port) and DS-8B (8-port) (Brocade SilkWorm 2xxx) Departmental Switches are Fibre Channel gigabit switches. The modular design of these switches (with an LCD front panel display) allows for easy installation and setup. The switch's modular construction gives the switch a range of flexibility in creating, upgrading, maintaining, and configuring a Fibre Channel fabric. Connectrix DS-16B and DS-8B handle ESN configurations, from small workgroups to large departmental and enterprise solutions. The DS-16B and DS-8B meet the needs of broad range of ESN applications by allowing multiple-switch fabric devices to connect to multiple Fibre Channel ports.

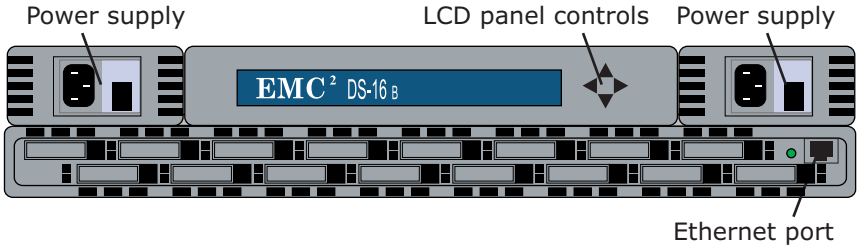


Figure 382 **Connectrix DS-16B**

These versions are offered:

- DS-16B with 16 short wave GBICs and dual power supplies
- DS-16B-02 with 14 short wave GBICs and 2 long wave GBICs and dual power supples
- DS-8B with 8 short wave GBICs and dual power supplies

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-16B and DS-8B.

Connectivity

Connectivity features include:

- Up to 8 or 16 ports available (depending on model)
- GBIC modules supported are the short- wavelength (SWL) and long- wavelength (LWL) fiber-optics
- Supports Class 2, Class 3 and Class F Fibre Channel protocols
- 16 buffer-to-buffer credits (BB_Credits) available per port excluding extended fabrics

- Minimum aggregate routing capacity of 4,000,000 frames / second is specified for Class 2, Class 3, and Class F frames
- All ports auto-negotiate to provide either switched N_Port or E_Port connections
- Full duplex 100 MB / s data rate per Fibre Channel port
- Supported port types include: F_Port, FL_Port, and E_Port

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix B Series release notes and product documentation on Powerlink.

Management

The DS-16B and DS-8B Departmental Switches can be managed using several remote and local access methods. Fabric management can be performed in-band or out-of-band through Telnet, or through the Web by connecting to the 10 / 100BaseT Ethernet port on the switch. Additionally, Fabric Manager, a Client / Server-based application, can be installed on a PC with Ethernet access to the switch, and can be used to perform fabric management.

Management features include:

- Centralized configuration and management of fabric using Client / Server architecture
- Supported Clients: Windows 95, Windows 98, Windows NT 4.0, Windows 2000, Sun Solaris
- In-band or out-of-band management available using Telnet or Web-based tool
- Centralized management of switch
- 10 / 100 Mb Ethernet connections for out-of-band management
- Extensive centralized logging: Event, Audit, Session logs, and SNMP support

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic error notification using FabricWatch utility

- Automatic error notification sent through **syslogd**
- Automatic error notification through Web tool or Fabric Manager
- MIB compliance: MIB-II system group, Interface group, and SNMP group, Fabric Element MIB, Vendor-specific MIBs, Standard Generic traps, Enterprise-specific traps

◆ Dual redundant power supplies

◆ Highly integrated, reliable, multifunction (ASIC) devices are used throughout the switch

◆ Fan, temperature, and port status are easily viewable from GUI management tool

◆ Dynamic statistics display performance data for each online port using GUI or Telnet CLI

◆ Event logs streamline troubleshooting and provide rapid error source identification

Further reading

Several related EMC Connectrix B manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix MDS-Series directors and switches

This section contains information on the following Connectrix MDS-Series Storage Service Module, directors, and switches.

- "Connectrix MDS Storage Service Module (SSM)," next
- "Connectrix MDS 9513" on page 1069
- "Connectrix MDS 9509" on page 1076
- "Connectrix MDS 9506" on page 1080
- "Connetrix MDS 9216" on page 1091

Connectrix MDS Storage Service Module (SSM)

Connectrix MDS storage service module

The Connectrix MDS Storage Services Module (SSM) is an open, standards-based module designed specifically to support intelligent fabric applications. The SSM provides investment protection as well as more rapid access to storage application innovations. The module provides thirty-two Fibre Channel ports and has embedded ASICs designed to deliver performance improvements for third-party intelligent fabric applications, based on open standards, such as Fabric Application Interface Standard (FAIS).

Figure 383 shows the Connectrix MDS storage service module.

Figure 383 Connectrix MDS storage service module (SSM)

Key features of the SSM

- Ports per blade

  32 ports. ~4:1 oversubscribed
- Port speed

  2 Gb
- Media types
  - Small Form Factor Pluggable (SFP) laser
  - Short-wave up to 300m
- Operations
  - Oversubscribed line card for host connectivity
  - Platform for Invista virtualization services.

  Both functions can be performed simultaneously

For more information on overall capabilities of the Cisco SSM, go to:

http://www.cisco.com/en/US/products/ps5991/products_data_sheet0900aecd802496f7.html

At this URL, you will find data information, including:

- Specification of the blade
- Additional services of the SSM
- Network management
- Environmental information
- Distance capabilities

Connectrix MDS 9513

The MDS 9513 chassis is a Director Class switch capable to support up to 528 ports in a single chassis capable of 1/2/4 Gb speeds. MDS 9513s are installed with Supervisor-2 modules only, and support all of the Cisco line cards Generation 1 and 2 with the caveat that utilizing one or the other, or subsequent mixture of line cards will reduce or increase the number of ports capable of being used.

All line cards supported (12, 16, 24, 32, 48) are capable of 1/2/4 Gb FC and 10 Gb FC These line cards are also supported in the MDS 9200 and MDS 9500 switches. IPS and 14/2 blades are also supported in the director class MDS-9500 series. The following information will help guide implementers design SAN environments and topologies by utilizing the features provided by Cisco MDS platform. For more information on the MDS hardware can be found at www.cisco.com.

Key features

Key features of the MDS 9513 include:

- Up to 528 ports in a single chassis
- Generation 2 line cards
  - 12-port supports 1/2/4 Gb/s
    - Dedicated bandwidth at 1/2/4 Gb/s
  - 24-port supports 1/2/4 Gb/s speeds
    - Dedicated bandwidth at 1/2 Gb/s
    - 2:1 shared (oversubscribed) bandwidth at 4 Gb/s
  - 48-port supports 1/2/4 Gb/s speeds
    - Dedicated bandwidth at 1 Gb/s
    - 2:1 shared (oversubscribed) bandwidth at 2 Gb/s
    - 4:1 shared (oversubscribed) bandwidth at 4 Gb/s
  - 4-port supports 10 Gb/s speed
- For 24-port and 48-port modules, Port Bandwidth Reservation feature allows dedicating specific bandwidth to each particular port.. This provides the optimal allowed bandwidth allocation for any types of applications as well as high performance ISLs within logical appropriated reason. Each line card has 4-port groups that accommodates 12.8 Gb/s of shared bandwidth. When using dedicated bandwidth on specific ports, remaining ports within a group will share the remaining available bandwidth.
  - 24-port Module
    4 groups of 6 ports share 12.8 Gb/s of bandwidth
    Port Groupings:
    - Group1: Ports 1-6
    - Group2: Ports 7-12
    - Group3: Ports13-18
    - Group4: Ports 19-24
  - 48-port Module
    4 groups of 12 ports shares 12.8 Gb/s of bandwidth
    Port Groupings:
    - Group1: Ports 1-12
    - Group2: Ports 13-24
    - Group3: Ports 25-36
    - Group4: Ports 37-48
- 4095 Extended Buffer credits on all the Generation-2 line cards.

- Supports Generation-1 switching Modules and all of their Features
  - Fibre Channel Switching Modules
  - IP Storage Services Modules
  - Multiprotocol Service Module
  - Storage Services Module
- Up to 20 VSAN per switch
- Redundant supervisors also providing non-disruptive upgrade
- Redundant power supplies
- Up to 16 links in a Port-Channel providing up to 64 Gb of bandwidth

- Port mode support of E, TE, FL, F and SD (Span Destination)

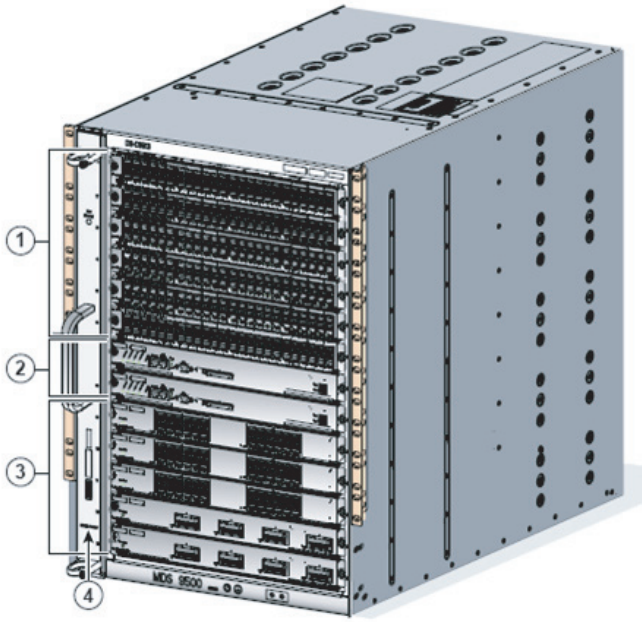


Figure 384 MDS 9513 chassis front panel view1

Legend	
1	Switching or services modules in slots 1-6
2	Supervisor-2 modules in slot 7 and slot 8
3	Switching or services modules in slots 9-13
4	Fan Tray

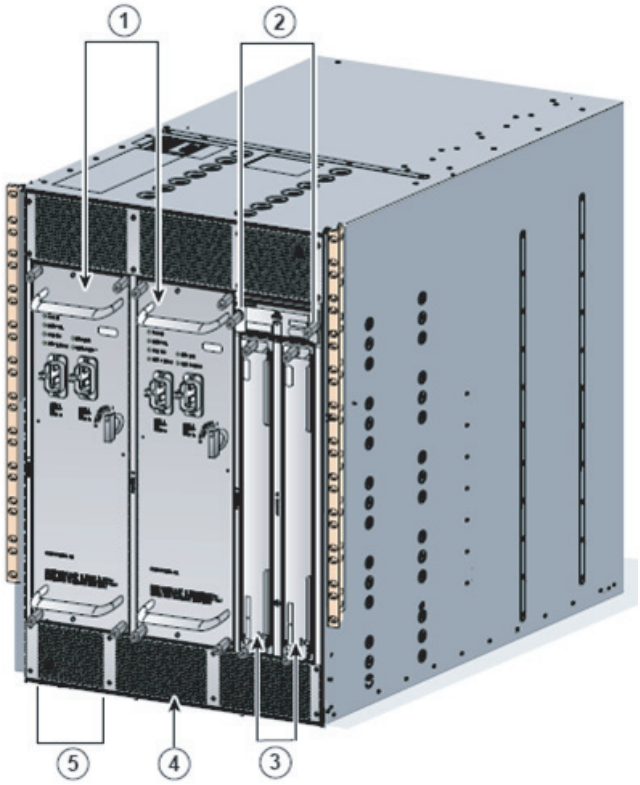


Figure 385 **MDS 9513 chassis rear panel view**

Legend	
1	Power Supplies
2	Crossbar Module Fans
3	Crossbar Modules
4	Clock Module - Clock modules are located inside the air vent panel. You must remove the air vent panel to access the clock modules
5	Air Vent Panels

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*,

refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the *EMC Connectrix MDS 9000 SAN-OS Release Notes* release notes on Powerlink.

System architecture

Generation-1 Switching Modules	Generation-2 Switching Modules
16-port FC Module (1/2 Gb)	12-port FC Module (1/2/4 Gb)
32-port FC Module (1/2 Gb)	24-port FC Module (1/2/4 Gb)
Storage Services Module (SSM)	48-port FC Module (1/2/4 Gb)
IP Services 4-port GE Module (IPS4)	4-port FC Module (10 Gb)
IP Services 8-port GE Module (IPS8)	
Multi-Protocol Services Module (MPS – 14 FC + 2 GE)	

With second generation line cards able to be used on first generation chassis, such as the MDS 9506 and MDS 9509, Table 59 on page 1085 shows capabilities for all three chassis

Management

- EMC Control Center
- Fabric Manager
- Device Manager
- Web-Based GUI
- Scriptable CLI
- Policy Base Alerts
- Fabric API
- End-to-End Performance Monitoring
- On-line Fault Isolation

Mechanical specifications

- Height = 24.5 inches (14RU)
- Weight = 17.37 inches
- Depth = 28.0 inches

- Power

Product ID	Description	Power (W)
MDS 9513 Common Equipment (Use EMC ids for this column)	MDS 9513 common equipment includes: • DS-C9513 Chassis (1) • DS-13SLT-FAN-F (1) : Front Fan Tray • DS-13SLT-FAN-R (1) : Rear Fan Tray • DS-13SLT-FAB1 (2) : Crossbar Fabric • DS-C9513-CL (2) : Clock Module • DS-X9530-SF2-K9 (2) – Supervisor/Fabric-2	696
DS-X9112	MDS 9000 12-port 1/2/4 Gb/s FC Module	132
DS-X9124	MDS 9000 24-port 1/2/4 Gb/s FC Module	147
DS-X9148	MDS 9000 48-port 1/2/4 Gb/s FC Module	185
DS-X9704	MDS 9000 4-port 10 Gb/s FC Module	172

- Weight

MDS 9513 chassis component	Weight (pounds)
Chassis (includes fans and clock modules)	100.0
Power Supply (6000W AC)[a]	32.5
Front Fan Tray	18.0
Rear Fan Tray	2.25
Crossbar Fabric Module[a]	5.75
Supervisor-2 Module[a]	7.25
12-Port 1/2/4 Gb/s Fibre Channel Module	7.75
24-Port 1/2/4 Gb/s Fibre Channel Module	7.75
48-Port 1/2/4 Gb/s Fibre Channel Module	11.0
Line Card Blank Panel	0.5

a. 2 of such units are needed for a normal HA configuration of the switch.

Protocols

- FC
- FICON
- IP-FC
- FCIP
- iSCSI

Further reading

For details on MDS platform features and functionality, refer to Connectrix M release notes on Powerlink.

Other white papers, data sheets, and configuration guides can be accessed at www.cisco.com in the Technical Support and Documentation section.

Connectrix MDS 9509

The MDS 9509 chassis is a Director class switch capable to support up to 336 ports in a single chassis capable of 1/2/4 Gb speeds. MDS 9509s are installed with Supervisor-1 or Supervisor-2 modules only and support all of the Cisco line cards Generation 1 and 2 with the caveat that utilizing one or the other, or subsequent mixture of line cards, will reduce or increase the number of ports capable of being used. All line cards supported (12, 16, 24, 32, 48) are capable of 1/2/4 Gb FC and 10 Gb FC These line cards are also supported in the MDS 9200 and MDS 9500 switches. IPS and 14/2 blades are also supported in the director class MDS-9500 series. The following information will help guide implementers design SAN environments and topologies by utilizing the features provided by Cisco MDS platform. For more information on the MDS hardware can be found at www.cisco.com.

Key features

Key features of the MDS 9509 include:

- Up to 336 ports in a single chassis
- Generation 2 line cards
  - 12-port supports 1/2/4 Gb/s
    - Dedicated bandwidth at 1/2/4 Gb/s
  - 24-port supports 1/2/4 Gb/s speeds
    - Dedicated bandwidth at 1/2 Gb/s
    - 2:1 shared (oversubscribed) bandwidth at 4 Gb/s
  - 48-port supports 1/2/4 Gb/s speeds
    - Dedicated bandwidth at 1 Gb/s
    - 2:1 shared (oversubscribed) bandwidth at 2 Gb/s
    - 4:1 shared (oversubscribed) bandwidth at 4 Gb/s
  - 4-port supports 10 Gb/s speed
- For 24-port and 48-port modules, Port Bandwidth Reservation feature allows dedicating specific bandwidth to each particular port. This provides the optimal allowed bandwidth allocation for any types of applications as well as high performance ISLs

within logical appropriated reason. Each line card has 4-port groups that accommodates 12.8 Gb/s of shared bandwidth. When using dedicated bandwidth on specific ports, remaining ports within a group will share the remaining available bandwidth.

- 24-port module
  4 groups of 6 ports share 12.8 Gb/s of bandwidth
  Port groupings:
  - Group1: Ports 1-6
  - Group2: Ports 7-12
  - Group3: Ports13-18
  - Group4: Ports 19-24
- 48-port module
  4 groups of 12 ports shares 12.8 Gb/s of bandwidth
  Port groupings:
  - Group1: Ports 1-12
  - Group2: Ports 13-24
  - Group3: Ports 25-36
  - Group4: Ports 37-48

◆ 4095 extended buffer credits on all the Generation-2 line cards.

◆ Supports Generation-1 switching modules and all of their features

- Fibre Channel Switching Modules
- IP Storage Services Modules
- Multiprotocol Service Module
- Storage Services Module

◆ Up to 20 VSAN per switch

◆ Redundant supervisors also providing non-disruptive upgrade

◆ Redundant power supplies

◆ Up to 16 links in a port-channel providing up to 64 Gb of bandwidth

◆ Port mode support of E, TE, FL, F and SD (Span Destination)

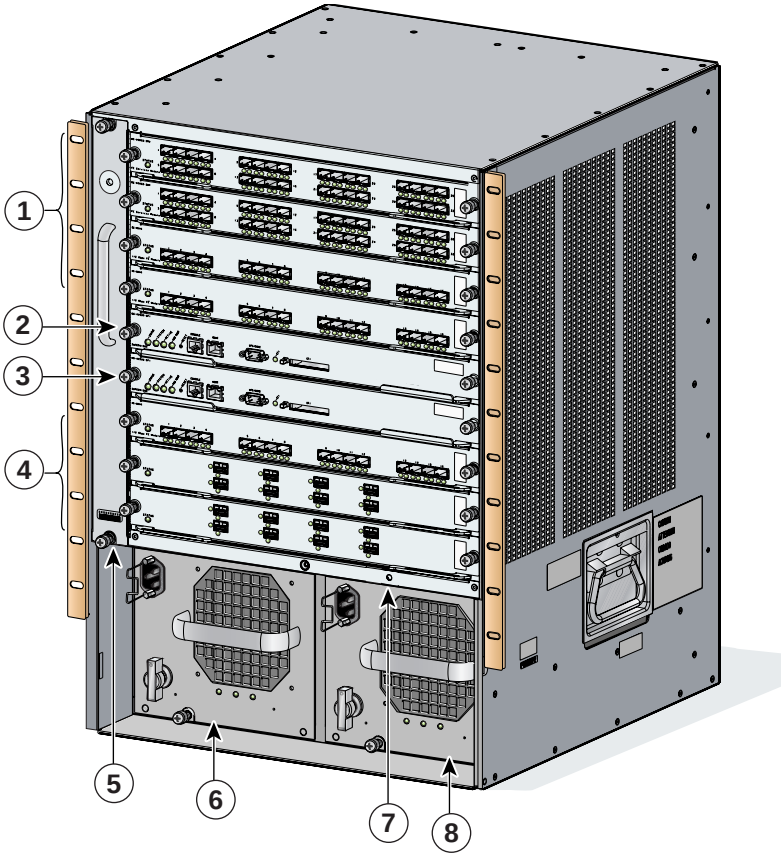


Figure 386 MDS 9509 chassis front panel view1

Legend				
1	Switching or services modules in slots 1-4		5	Fan module
2	Supervisor module in slot 5		6	Power supply 1
3	Redundant supervisor module in slot 6		7	ESD socket
4	Switching or services modules in slots 7-9		8	Power supply 2 (redundant)

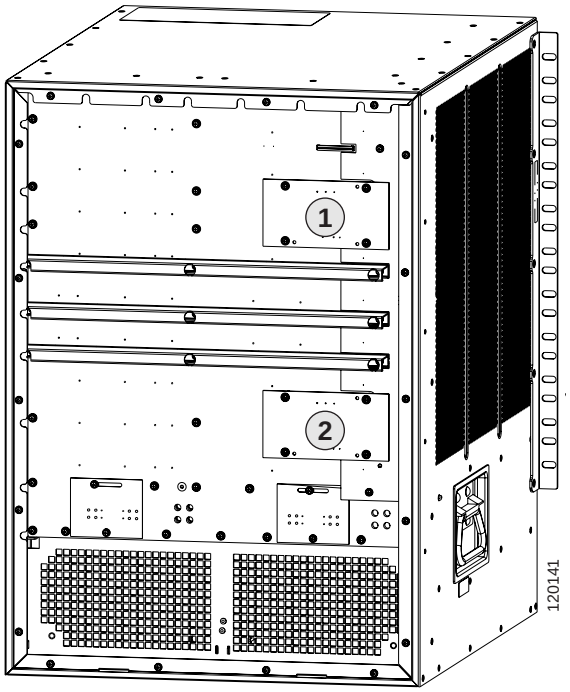


Figure 387 **MDS 9509 chassis rear panel view**

Legend	
1	Clock A (CLK A)
2	Clock B (CLK B)

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

MDS 9509 supports SAN OS 3.0(2) and subsequent releases moving forward. For the most up-to-date information on unsupported features, consult the *EMC Connectrix MDS 9000 SAN-OS Release Notes* release notes on Powerlink.

System architecture

.

Generation-1 Switching Modules	Generation-2 Switching Modules
16-port FC Module (1/2 Gb)	12-port FC Module (1/2/4 Gb)
32-port FC Module (1/2 Gb)	24-port FC Module (1/2/4 Gb)
Storage Services Module (SSM)	48-port FC Module (1/2/4 Gb)
IP Services 4-port GE Module (IPS4)	4-port FC Module (10 Gb)
IP Services 8-port GE Module (IPS8)	
Multi-Protocol Services Module (MPS – 14 FC + 2 GE)	

With second generation line cards able to be used on first generation chassis, such as the MDS 9506 and MDS 9509, Table 59 on page 1085 shows chassis support and capabilities.

Further reading

For details on MDS platform features and functionality, refer to Connectrix M release notes on Powerlink.

Other white papers, data sheets, and configuration guides can be accessed at www.cisco.com in the Technical Support and Documentation section.

Connectrix MDS 9506

The MDS 9506 chassis is a Director class switch capable to support up to 192 ports in a single chassis capable of 1/2/4 Gb speeds. MDS 9506s can be installed with Supervisor-1 or 2 modules and support all of the Cisco line cards Generation 1 and 2 with the caveat that utilizing one or the other, or subsequent mixture of line cards, will reduce or increase the number of ports capable of being used. All line cards supported (12, 16, 24, 32, 48) are capable of 1/2/4 Gb FC and 10 Gb FC. These line cards are also supported in the MDS 9200 and MDS 9500 switches. IPS and 14/2 blades are also supported in the director class MDS-9500 series. The following information will help guide implementers design SAN environments and topologies by utilizing the features provided by Cisco MDS platform. For more information on the MDS hardware can be found at www.cisco.com.

Key features

Key features of the MDS 9506 include:

- Up to 192ports in a single chassis
- Generation 2 line cards
  - 12-port supports 1/2/4 Gb/s
    - Dedicated bandwidth at 1/2/4 Gb/s
  - 24-port supports 1/2/4 Gb/s speeds
    - Dedicated bandwidth at 1/2 Gb/s
    - 2:1 shared (oversubscribed) bandwidth at 4 Gb/s
  - 48-port supports 1/2/4 Gb/s speeds
    - Dedicated bandwidth at 1 Gb/s
    - 2:1 shared (oversubscribed) bandwidth at 2 Gb/s
    - 4:1 shared (oversubscribed) bandwidth at 4 Gb/s
  - 4-port supports 10 Gb/s speed
- For 24-port and 48-port modules, Port Bandwidth Reservation feature allows dedicating specific bandwidth to each particular port. This provides the optimal allowed bandwidth allocation for any types of applications as well as high performance ISLs within logical appropriated reason. Each line card has 4-port groups that accommodates 12.8 Gb/s of shared bandwidth. When using dedicated bandwidth on specific ports, remaining ports within a group will share the remaining available bandwidth.
  - 24-port module
    4 groups of 6 ports share 12.8 Gb/s of bandwidth
    Port groupings:
    - Group1: Ports 1-6
    - Group2: Ports 7-12
    - Group3: Ports13-18
    - Group4: Ports 19-24
  - 48-port module
    4 groups of 12 ports shares 12.8 Gb/s of bandwidth
    Port groupings:
    - Group1: Ports 1-12
    - Group2: Ports 13-24
    - Group3: Ports 25-36
    - Group4: Ports 37-48
- 4095 extended buffer credits on all the Generation-2 line cards.

- Supports Generation-1 switching modules and all of their features
  - Fibre Channel switching modules
  - IP Storage services modules
  - Multiprotocol service module
  - Storage services module
- Up to 20 VSAN per switch
- Redundant supervisors also providing non-disruptive upgrade
- Redundant power supplies
- Up to 16 links in a port channel providing up to 64 Gb of bandwidth
- Port mode support of E, TE, FL, F and SD (Span Destination)

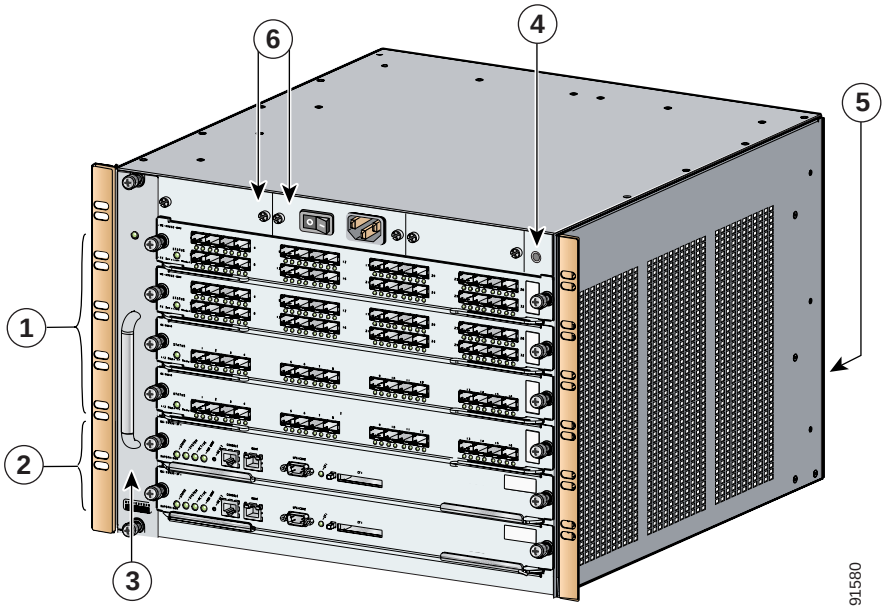


Figure 388 MDS 9506 chassis front panel view1

Legend				
1	Switching or services modules in slots 1-4		4	ESD Socket
2	Supervisor modules in slot 5 and slot 6		5	Power supplies (in back)
3	Fan module		6	Location of power entry modules (PEMs) — one PEM shown and one filler panel shown.

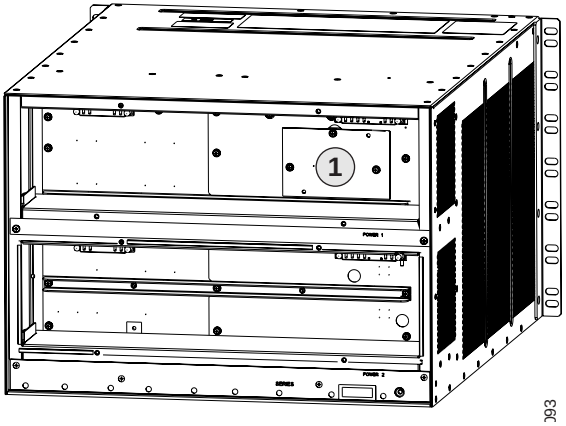


Figure 389 **MDS 9506 chassis rear panel view**

Legend	
1	Clock module

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

MDS 9506 supports SAN OS 3.0(2) and subsequent releases moving forward. For the most up-to-date information on unsupported features, consult the *EMC Connectrix MDS 9000 SAN-OS Release Notes* release notes on Powerlink.

System architecture

Generation-1 Switching Modules	Generation-2 Switching Modules
16-port FC Module (1/2 Gb)	12-port FC Module (1/2/4 Gb)
32-port FC Module (1/2 Gb)	24-port FC Module (1/2/4 Gb)
Storage Services Module (SSM)	48-port FC Module (1/2/4 Gb)

Generation-1 Switching Modules	Generation-2 Switching Modules
IP Services 4-port GE Module (IPS4)	4-port FC Module (10 Gb)
IP Services 8-port GE Module (IPS8)	
Multi-Protocol Services Module (MPS – 14 FC + 2 GE)	

With second generation line cards able to be used on first generation chassis, such as the MDS 9506 and MDS 9509, Table 59 shows capabilities for all three chassis.

Table 59 Chassis capabilities

	MDS 9506	MDS 9500	MDS 9513
Rack space	7 RU	14 RU	14 RU
Ports per Chassis (Generation 1 modules)	192	252	252
Ports per Chassis (Generation 2 modules)	192	336	528
Ports per Chassis Mixed (Generation 1and Generation 2 modules)	192	252	252
Fibre Channel Ports (Generation 1 Switching Modules)			
Port Speed	1 Gb/s/2 Gb/s autosensing	1 Gb/s/2 Gb/s autosensing	1 Gb/s/2 Gb/s autosensing
Long Wave SFP	10 Km	10 Km	10 Km
CWDM	Yes, 100 Km	Yes, 100 Km	Yes, 100 Km
DWDM	Yes	Yes	Yes
Fibre Channel standard port Types E, F, FL, B, Private loop (B Port and Private loop not supported)	Yes	Yes	Yes
Fibre Channel standard port Types SD, ST, TE	Yes	Yes	Yes
FC Ports per blade	14, 16, or 32	14, 16, or 32	14, 16, or 32

Table 59 Chassis capabilities (continued)

	MDS 9506	MDS 9500	MDS 9513
GE Ports per blade (4-port GigE module discontinued)	2 or 8	2 or 8	2 or 8
Port blades per chassis	4	7	11
Fibre Channel Ports (Generation 2 Switching Modules)			
Port Speed	1 Gb/s/2 Gb/s autosensing	1 Gb/s/2 Gb/s autosensing	1 Gb/s/2 Gb/s autosensing
CWDM	Yes, 100 Km, 1 & 2 Gb/s only	Yes, 100 Km, 1 & 2 Gb/s only	Yes, 100 Km, 1 & 2 Gb/s only
DWDM	Yes	Yes	Yes
Fibre Channel standard port Types E, F, FL	Yes	Yes	Yes
Fibre Channel standard port Types SD, ST, TE	Yes	Yes	Yes
FC Ports per blade	12, 24, 48 1/2/4 Gb/s ports, 4 10 Gb/s ports	12, 24, 48 1/2/4 Gb/s ports, 4 10 Gb/s ports	12, 24, 48 1/2/4 Gb/s ports, 4 10 Gb/s ports
Port blades per chassis	4	7	11
Performance (Generation 2)			
Fabric Bandwidth	1.4 Tbps	1.4 Tbp	2.2 Tbp
Fully Non-Blocking	Yes	Yes	Yes
Virtual Output Queues	Yes	Yes	Yes
Data Rate 12-Port Module	Full line rate at all supported speeds	Full line rate at all supported speeds	Full line rate at all supported speeds
Data Rate 24-Port Module	Full line rate at 1/2 Gb/s speeds 2:1 Oversubscribed at 4 Gb/s	Full line rate at 1/2 Gb/s speeds 2:1 Oversubscribed at 4 Gb/s	Full line rate at 1/2 Gb/s speeds 2:1 Oversubscribed at 4 Gb/s

Table 59 Chassis capabilities (continued)

	MDS 9506	MDS 9500	MDS 9513
Data Rate 48-Port Module	Full line rate at 1/2 Gb/s speeds 2:1 Oversubscribed at 2 Gb/s 4:1 Oversubscribed at 4 Gb/s	Full line rate at 1/2 Gb/s speeds 2:1 Oversubscribed at 2 Gb/s 4:1 Oversubscribed at 4 Gb/s	Full line rate at 1/2 Gb/s speeds 2:1 Oversubscribed at 2 Gb/s 4:1 Oversubscribed at 4 Gb/s
Ethernet Ports			
Port Speed	1 Gb/s	1 Gb/s	1 Gb/s
Long Wave SFP	10 Km	10 Km	10 Km
CWDM	Yes, 100 Km	Yes, 100 Km	Yes, 100 Km
DWDM	Yes	Yes	Yes
FCIP Support	Yes	Yes	Yes
iSCSI Support	Yes	Yes	Yes
EtherChannel Support	Yes	Yes	Yes
Virtual ISL per port (FCIP)	3	3	3
Ports per blade	2 or 8	2 or 8	2 or 8
Ports per chassis	3	6	10
Max Ports per chassis	24	48	80
Features			
Virtual SANs (VSAN)	Yes	Yes	Yes
Port Channel with Multipath Load Balancing	Up to 16 ports per Port Channel across switching modules	Up to 16 ports per Port Channel across switching modules	Up to 16 ports per Port Channel across switching modules
QoS	Yes	Yes	Yes

Table 59 Chassis capabilities (continued)

	MDS 9506	MDS 9500	MDS 9513
Max BB_Credits per port (Without using Licensed Extended Buffer Credit feature)	255 per port in the 16-Port Module 12 per port in the 32-Port Module 255 per port in MPS-14/2-Module 250 per port in the 12-Port Module 250 per port in dedicated mode for 24-Port and 48-Port Module	255 per port in the 16-Port Module 12 per port in the 32-Port Module 255 per port in MPS-14/2-Module 250 per port in the 12-Port Module 250 per port in dedicated mode for 24-Port and 48-Port Module	255 per port in the 16-Port Module 12 per port in the 32-Port Module 255 per port in MPS-14/2-Module 250 per port in the 12-Port Module 250 per port in dedicated mode for 24-Port and 48-Port Module
Extended Buffer Credit (Licensed feature)	Up to 4095 Buffer Credits per port. Supported on all Generation-2 switching modules. Up to 3500 per port supported on MPS-14/2 module.	Up to 4095 Buffer Credits per port. Supported on all Generation-2 switching modules. Up to 3500 per port supported on MPS-14/2 module.	Up to 4095 Buffer Credits per port. Supported on all Generation-2 switching modules. Up to 3500 per port supported on MPS-14/2 module.
Zoning			
N_Port WWN	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
F_Port WWN	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
N_Port FCID	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
Port	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
Broadcast	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
Default	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
LUN	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced
Read Only	Yes, HW enforced	Yes, HW enforced	Yes, HW enforced

Table 59 Chassis capabilities (continued)

	MDS 9506	MDS 9500	MDS 9513
Security			
Multi-level Management Access Control	Yes, Role-based	Yes, Role-based	Yes, Role-based
VSAN-Based Access Control	Yes, release 1.2 and higher	Yes, release 1.2 and higher	Yes, release 3.0
Port Security	Yes, release 1.2 and higher	Yes, release 1.2 and higher	Yes, release 3.0
RADIUS Support	Yes	Yes	Yes
TACACS Support	Yes, release 1.3 and higher	Yes, release 1.3 and higher	Yes, release 3.0
SNMPv3	Yes	Yes	Yes
RBAC	Yes	Yes	Yes
FC-SP for switch-to -switch and host-to-switch authentication	Yes	Yes	Yes
SSH	Yes	Yes	Yes
SFTP	Yes	Yes	Yes
Platform High Availability Features			
Hot-swap Supervisor Modules	Yes, non-disruptive	Yes, non-disruptive	Yes, non-disruptive
Hot-swap Optics	Yes	Yes	Yes
Hot-swap Power	Yes	Yes	Yes
Hot-swap Fan Assemblies	Yes	Yes	Yes
Hot-swap Line Cards	Yes	Yes	Yes
Hot-swap Cross-bars	Yes	Yes	Yes
Dual OOB Management Channels	Yes	Yes	Yes
Dual System Clocks	Yes	Yes	Yes
Stateful Process Restartability	Yes	Yes	Yes
Serviceability			
Non-disruptive processor failover	Yes	Yes	Yes
Non-disruptive code load	Yes	Yes	Yes

Table 59 Chassis capabilities (continued)

	MDS 9506	MDS 9500	MDS 9513
Non-disruptive code activation	Yes	Yes	Yes
Non-disruptive line card replacement	Yes	Yes	Yes
SNMP traps for alerts	Yes	Yes	Yes
Configuration file management	Yes	Yes	Yes
Network Boot	Yes	Yes	Yes
Call Home	Yes	Yes	Yes
Diagnostics			
POST	Yes	Yes	Yes
FC Ping	Yes	Yes	Yes
FC Traceroute	Yes	Yes	Yes
Port Beacon	Yes	Yes	Yes
Online Port Loopback	Yes	Yes	Yes
Port Level Statistics	Yes	Yes	Yes
On-Line System Health	Yes	Yes	Yes
Real Time Protocol Debug	Yes	Yes	Yes
SPAN	Yes	Yes	Yes
Remote SPAN	Yes	Yes	Yes
Embedded Protocol Analyzer	Yes	Yes	Yes
Embedded CLI Audit	Yes	Yes	Yes
Syslog	Yes	Yes	Yes

Further reading

For details on MDS platform features and functionality, refer to Connectrix M release notes on Powerlink.

Other white papers, data sheets, and configuration guides can be accessed at www.cisco.com in the Technical Support and Documentation section.

Connetrix MDS 9216

The MDS 9216 fabric switch has a fixed non-redundant *supervisor* card coupled with a built-in 16-port Fibre Channel switch, dual system clocks, a single field-replaceable power supply, and a non-redundant, hot-swappable fan assembly. An empty port card slot can accomodate a 16-port or 32-port line card, providing up to 48 ports in the 3U switch.

Figure 390 **Connectrix MDS 9216**

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the MDS 9216.

Supervisor module

The supervisor module is a fixed non-redundant module with a built-in 16-port Fibre Channel switch.

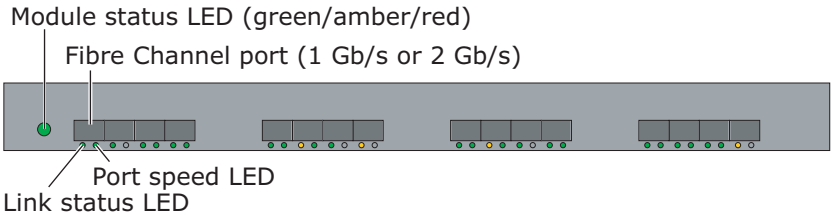


Figure 391 **MDS 9216 supervisor module**

The supervisor module functions the same as MDS 9500 supervisor module, and has dual system clocks.

Interface module

The interface module is a fixed, non-redundant module.

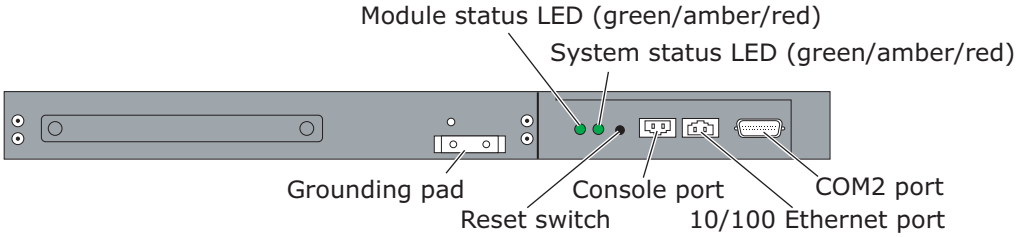


Figure 392 MDS 9216 interface module

Fibre Channel line cards

The ports are auto-configurable to 1 GB/s to 2 Gb/s, and can carry standard port types, as well as SFP optical interfaces.

16-port line card

The 16-port line card defaults with Fx_Ports configured for 16 BB_Credits, but the card can be set for up to 255 BB_Credits using TE_Port or E_port configurations.

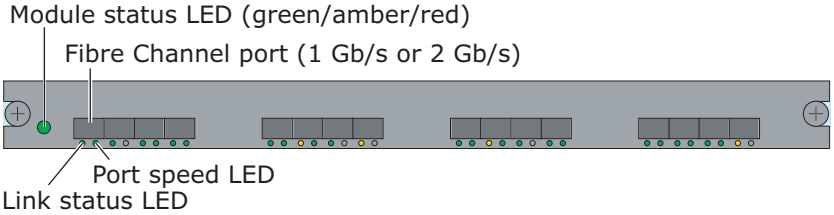


Figure 393 MDS 9216 16-port line card

32-port line card

The 32-port line card defaults with Fx_Ports, and is fixed with 12 BB_Credits per port.

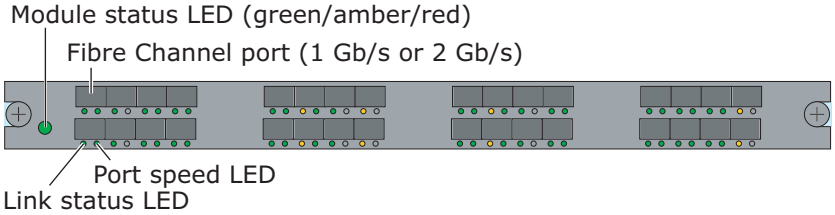


Figure 394 MDS 9216 32-port line card

MDS management

Management features include:

- **Command line interface (CLI)**:
  - Adheres to the syntax of Cisco IOS CLI
  - Multiple connection options and protocols:
    - Direct or serial link — VT100
    - Secure shell access — SSH (encrypted)
    - Terminal Telnet — TCP / IP over Ethernet or Fibre Channel
    - Over 1700 commands based on Cisco IOS
    - Exec mode (operations other than config)
    - Configuration mode (all parameters)
    - Help facility (help, errors, and history)
- **Device Manager** — SNMP-based GUI device management tool:
  - Switch-embedded Java application
  - Downloaded through web browser
  - Installed and updated automatically by Java Web Start
  - Runs on end user's workstation
- **Fabric Manager** — SNMP-based GUI fabric management tool for easy management across multiple switches and fabrics:
  - Switch-embedded Java application
  - Downloaded through web browser
  - Installed and updated automatically by Java Web Start
  - Runs on end user's workstation
- **SNMPv3 secures communications with switches**
- **Management access**:
  - SSHv2
  - RADIUS
  - SNMPv1,2,3
  - Role-based access

Fabric Manager troubleshooting tools

- Zone Merge Analysis
- End-to-End Connectivity Analysis
  - Verifies all fabric connections with *ping*
  - Flags operations exceeding user-defined latency value
- Switch Health Analysis Checks
  - CPU, memory, and flash utilization greater than 90%
  - Cards, power supplies, fan failures

- Over temperature conditions
- Switch port failures (40+ checks)
- VSAN and Fibre Channel services

Troubleshooting with the SysLog

- Chronological log of system messages
- Details of SysLog depend on configuration
- Eight severity levels from debug to critical events
- User can limit severity levels per specific service
- Syslog messages can be stored on disk, server, or sent to a console for immediate view by the user
- Radius Accounting Records to SysLog service
- Syslog analyzer tool
- Filters logged messages
- Displays explanations of probable causes and recommended actions
- Manual parsing of Syslog files

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the *EMC Connectrix MDS 9000 SAN-OS Release Notes* release notes on Powerlink.

Further reading

For details on MDS platform features and functionality, refer to Connectrix M release notes on Powerlink.

Other white papers, data sheets, and configuration guides can be accessed at www.cisco.com in the Technical Support and Documentation section.

Connectrix M-Series directors and switches

This section lists information on Connectrix M directors and switches. The following Connectrix M-Series directors and switches are discussed in this section. Also included is information on the QPM, XPM, and Open Trunking feature.

- "Switch model numbers," next
- "Connectrix ED-10000M" on page 1096
- "Connectrix ED-140M" on page 1108
- "Connectrix M Quad Port Module (QPM)" on page 1111
- "XPM port card" on page 1119
- "Connectrix ED-64M" on page 1122
- "Connectrix ED-1032" on page 1124
- "Connectrix DS-4700M" on page 1127
- "Connectrix DS-4400M" on page 1134
- "Connectrix DS-24M2" on page 1141
- "Connectrix DS-32M, DS-32M2" on page 1143
- "Connectrix DS-16M, DS-16M2" on page 1145
- "Connectrix M Series (Brocade M-Series) Open Trunking feature" on page 1147

Switch model numbers

This section includes the Connectrix M and Brocade M model names and ASICs.

Table 60 EMC and Brocade models and ASICs

EMC switch model	Brocade M switch model	ASIC
DS-16M	ES-3016	Viper
D-16M2	ES-3216	Viper 2
DS-24M2	ES-4500	Poseidon, Poseidon II (4500v2)
DS-32M	ES-3032	Viper

Table 60 EMC and Brocade models and ASICs (continued)

EMC switch model	Brocade M switch model	ASIC
DS-32M2	ES-3232	Viper 2
DS-4400M	ES-4400	Atlas
DS-4700M	ES--4700	Pegasus
ED-64M	ED-6064 or Intrepid 6064	Viper, Viper 2
ED-140M	ED-6140 or Intrepid 6140	Viper 2
ED-10000M	ED-10000 or Intrepid 10000	SPP, FLC, SNX, SPPe, FNX
ED-1032	ED-5000	Stitch
MP-1620M	Eclipse 1620	Dione 2B, Tzar 2B, Regent 2B
MP-2640M	Eclipse 2640	Yaff, Dione 3, Tzar 3, Regent 3

Connectrix ED-10000M

The Connectrix ED-10000M is a director class 256-port Fibre Channel switch that provides dynamic switched connections between Fibre Channel servers and devices in a SAN environment. 1 Gb/s, 2 Gb/s, 4 Gb/s, and 10 Gb/s port modules are available.

For 1 Gb/s, 2 Gb/s, and 4 Gb/s devices, the chassis may be configured from 32 to 256 ports. 10 Gb/s ports may also be mixed into the configurations as E_Ports for ISLs to other switches supporting 10 Gb/s.

The ED-10000M (shown in Figure 395 on page 1097) is installed in a Connectrix EC-1500-M cabinet, which also houses a Connectrix service processor. The cabinet can contain up to two ED-10000M Directors.

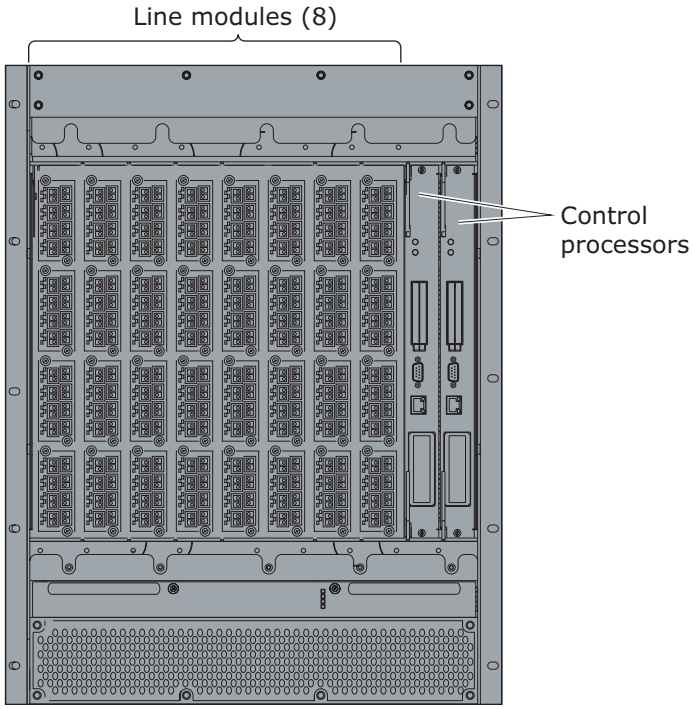


Figure 395 **Connectrix ED-10000M**

Hardware

Hardware components include:

- Two Control Processors (CTPs)
- Eight Line Modules (LIMs)
- Four switching modules
- Two sets of dual power supplies
- Four fan trays
- One chassis
- One midplane

Physical specifications

Dimensions and weights

- Height — 24.50 in (62.2 cm) (14 U)
- Width — 17.45 in (44.3 cm)
- Depth — 29.51 in (75.0 cm)

- Weight:
  - Chassis and midplane only (no FRUs) — 101 lb (45.8 kg)
  - Fully loaded — 333 lb (151.1 kg)

Power requirements

- Two power switches
- Input voltage — 180 to 270 VAC
- Input current — 11.4 A at 208 VAC
- Input frequency — 47 to 63 Hz
- Consumption — 2400 Watts maximum

Operating environment

- Temperature — 32° F to 104° F (0° C to 40° C)
- Relative humidity — 8% to 80%
- Maximum wet-bulb temperature — 81° F (27° C)
- Altitude — 10,000 ft (3,048 m)
- Inclination — 10° maximum
- Heat Output — 8192 BTUs / Hr

Oversubscription

On an ED-10000M, bandwidth is shared by all ports connected to the same SPP / SPPe (Port ASIC).

Bandwidth required:

Each Port ASIC provides bandwidth to 16 front-end ports. To determine the bandwidth required, sum the negotiated front-end port speeds. For example, if 16 ports are all negotiated to 2 Gb / s, then the bandwidth required is 32 Gb / s. If 10 ports are negotiated to 1 Gb / s, then the bandwidth required is 10 Gb / s.

Bandwidth available:

Each Port ASIC also receives a total of 10 Gb / s of bandwidth from each SWM / SWX (switch module) installed. To calculate the bandwidth available, multiply 10 Gb / s by the number of switch module installed. For example, in a switch containing 4 switch modules, 40 Gb / s of bandwidth is available.

Table 61 provides inforamtion on oversubscription.

Table 61 ED-10000M Oversubscription

	2 Gb		4 Gb	
	128	256	128	256
2 SWMs	None	2::1	2::1	4::1
4 SWMs	None	None	None	2::1

Protocols

The ED-10000M supports:

- FCP (SCSI over Fibre Channel)
- FICON

Note: Intermixing FCP and FICON is supported. FICON cascading is also supported.

Zoning

The ED-10000M supports zoning by:

- World Wide name (hardware enforced at the ingress port)
- Physical port number

Note: Zone sets may be stored on both Connectrix Manager and the ED-10000M. However, Connectrix Manager is the recommended tool for zone management.

Firmware

The supported firmware for the ED-10000M differs from all other Connectrix M series models. However, just as with other M series models, code loads are nondisruptive to I/O.

CTP

The active Control Processor (CTP) initializes and configures the director after power-on, and contains the microprocessor and associated logic that coordinate director operation.

The backup CTP takes over operation if the active CTP fails. Failover from a faulty CTP to the backup CTP is transparent to attached devices.

CTPs are hot-swappable.

Line modules

There are two types of line modules supported on the ED-10000M:

- LIM
- LMQ

Common features

Common features are listed below. Refer to "LIM" on page 1101 and "LMQ" on page 1102 sections for additional information on each feature.

- Up to thirty-two ports per line module
- Hot Pluggable and non-disruptive insertion into ED-10000M chassis
- Hot pluggable optical transceivers
- Inserts into any ED-10000M LIM slot
- Can fully populate ED-10000M chassis (up to 8 per i10K for a total of 256 ports)
- Longwave (1310nm) optical transceivers available for 4 km and 10 km links over standard 9 micron singlemode fiber
- Brocade M-Series OpenTrunking compatible
- Configured and managed from CLI or Connectrix Manager.

Each line module has two front-end port ASICs. There are 16 SPPs in a chassis with 8 LIMs. Each SPP is responsible for sending I/O to and from the optical ports and the SWM. On each LIM, the bottom SPP supports ports 0 through 15, and the top SPP supports ports 16 through 31. The boxes in Figure 396 show the four quadrants of the ED-10000M. Each SPP is part of a different quadrant.

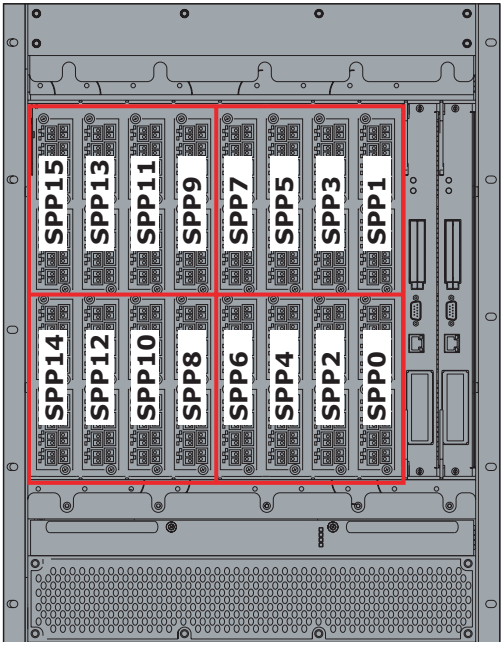


Figure 396 **ED-10000M quadrants**

Port location, addressing, and numbering

The ports are numbered from right to left and from bottom to top, with 0 at the bottom right. The port number is the same as the port address. Port numbers (and addresses) range from 0 to 255. A *port locator* is the physical location of the port. For example, a port locator of 1/0/4 designates LIM location 1, optical paddle location 0, port location 4.

LIM

Each module consists of the optical I/O interface (optical paddle) and the LIM, which connects to the midplane.

The optical paddles and LIMs are hot-swappable.

Paddles

Two types of optical paddles are available:

- 1 Gb or 2 Gb optical paddle
- 10 Gb optical paddle

The paddles are hot-swappable.

A small controller drives LEDs on the paddle and provides the control lines for the SERDES. The paddle houses the optics, and provides speed upgradeability from 2 Gb/s to 10 Gb/s using different paddles.

The paddle does not provide upgradeability to 4 Gb. 4 Gb requires the LMQ.

The paddle contains no port ASICs.

Features

For features common to all ED-10000M line modules, see "Common features" on page 1100.

- Auto-negotiates at 2 Gb/s and 1Gb/s
- Up to 4 paddles per LIM
- SPP ASIC
- Paddles provide either eight 1 and 2 Gb/s ports or two 10 Gb/s ports
- Shortwave (850nm) optical transceivers available for 300m links over standard 50 micron multimode fiber and OM2 cable (500m over OM3 cable) at 2 Gb/s.

Distance

The LIM supports up to 1133 buffer-to-buffer credits per port. At a link speed of 2 Gb/s, an ISL spanning a distance of 1133 km can be supported without experiencing a drop in the transfer rate (refer to "Data transfer rates" on page 800 for more information. Although, the switch supports that distance in regards to BB_Credit, the optical equipment in between must be able to support that distance as well. For example, distances greater than 150 meters require OM3 cable. Greater than 500 meters require longwave optics and single mode fiber, etc. (Refer to Chapter 3, "Extended Distance," for more information.)

LMQ

The LMQ is a line module that is capable of operating at 4 Gb/s. The module consists of optics and the LMQ itself which interfaces to the midplane.

Paddles

Unlike the LIM, there are no removable paddles on the LMQ.

Features

For features common to all ED-10000M line modules, see "Common features" on page 1100.

- Auto-negotiates at 4 Gb/s, 2 Gb/s, and 1 Gb/s
- Up to 256 non oversubscribed ports at 1 or 2 Gb/s
- Up to 256 oversubscribed 4 Gb/s ports (2:1 over-subscription)
- Fully compliant with Fibre Channel physical interface INCITS standard FC-PI-3 rev 1.0
- SPPe ASIC
- Can intermix LMQs with LIMs containing both 2 and 10 Gb/s paddles
- Shortwave (850nm) optical transceivers available for 150m links over standard 50 micron multimode fiber and OM2 cable (300m over OM3 cable)
- Requires E/OS 9.1 or above installed on the ED-10000M
- Compatible with Connectrix Manager 9.0 or above

Distance

On each LMQ, ports 0-3 and 16-19 can support up to 100 km at 4 Gb/s through 200 BB_Credits.

Ports 4-15 and 20-31 can support up to 32 km at 4 Gb/s through 64 BB_Credits. Although, the switch supports these distances in regards to BB_Credit, the optical equipment in-between must be able to support that distance as well. For example, distances greater than 150 meters require OM3 cable, greater than 500 meters require longwave optics and single mode fiber, etc. (Refer to Chapter 3, "Extended Distance," for more information.)

SWM

There are two types of switch modules, the SWM and the SWX. The SWM was the original switching module released with the ED-10000M. The SWX is the second generation switch module which has been optimized for reliability but does *not* introduce any new user features. Both switch modules are capable of supporting LIMs and LMQs. Throughout the ED-10000M and Connectrix Manager, both SWM's and SWXs are referred to generically as SWMs.

The director is delivered with up to four Switch Modules (SWMs). All SWMs are normally active and provide parallel incremental bandwidth to frame traffic across the midplane. With four SWMs, 586

Gb/s full duplex is available per system. Each SWM delivers 140 Gb/s of user-equivalent full-duplex data to all LIMs or 8.7 Gb/s to each of the 16 possible SPPs. With four SWMs populated, each SPP receives 34.9 Gb/s (4 x 8.7 Gb/s) of total bandwidth in support of sixteen 2 Gb/s ports. All midplane traces are high-speed differential traces that carry data at 2.86 Gb/s 8b/10b encoded.

The SWMs are responsible for Fibre Channel frame transmission from any director port to any other director port. Connections are established without software intervention. All SWMs are active. The module accepts a connection request from a port, determines if a connection can be established, and establishes the connection if the destination port is available. If any SWM fails, the remaining SWMs stay in service and provide connectivity and data frame transmission without interruption.

All frames must traverse an SWM between any two ports in the systems. There is no fixed path for data traffic from a host or a storage device. A node will use all SWMs. That is, all Fibre Channel frames are sprayed across all SWMs, thereby making use of all available bandwidth.

In a fully-loaded chassis (256 2 Gb ports), one needs four SWMs to achieve full port-to-port bandwidth, (2:1 oversubscribed with 256 4 Gb/s ports). A fully-loaded chassis with two SWMs has a 2:1 oversubscription ratio (4:1 with 4 Gb/sec ports). In a half-populated chassis (128 2 Gb ports), however, one needs only two SWMs to achieve full port-to-port bandwidth (four SWMs are needed for full bandwidth with 128 4 Gb/sec ports). Four SWMs are recommended to provide the most scalable bandwidth.

SWMs are hot-swappable.

FRU location and numbering

Field-replaceable-units (FRUs) are located by FRU type and chassis position. FRUs are numbered as shown in Figure 397 on page 1105 and Figure 398 on page 1106.

Note that the optical ports can be identified by a physical port locator in the format *x/y/z*, where *x* is the LIM slot, *y* is the paddle number (shown in circles in Figure 397), and *z* is the port on the paddle (also shown in circles in Figure 397).

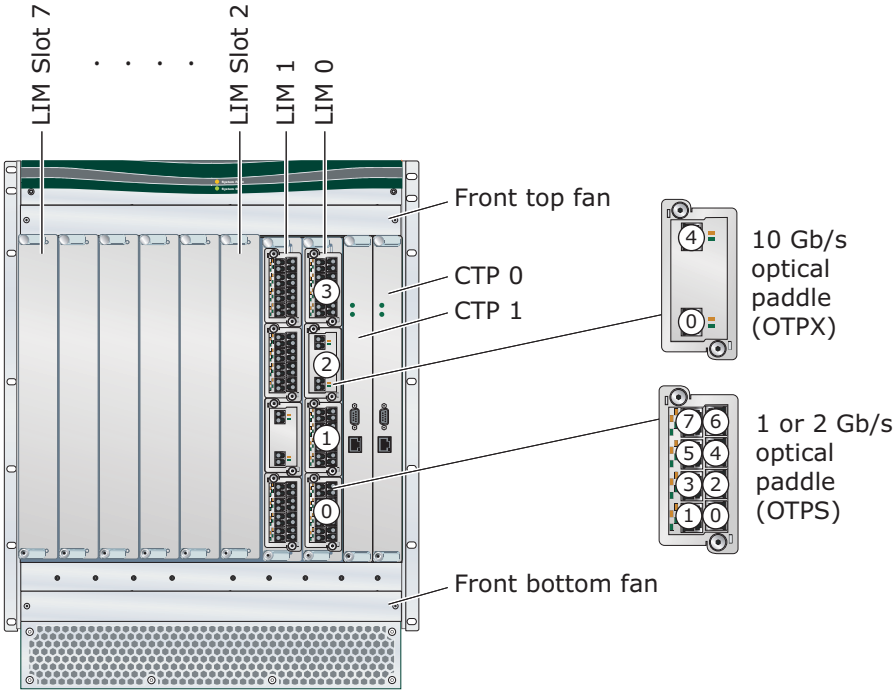


Figure 397 **FRUs on front of ED-10000M**

Note: Paddles are no longer removable on LMQ cards.

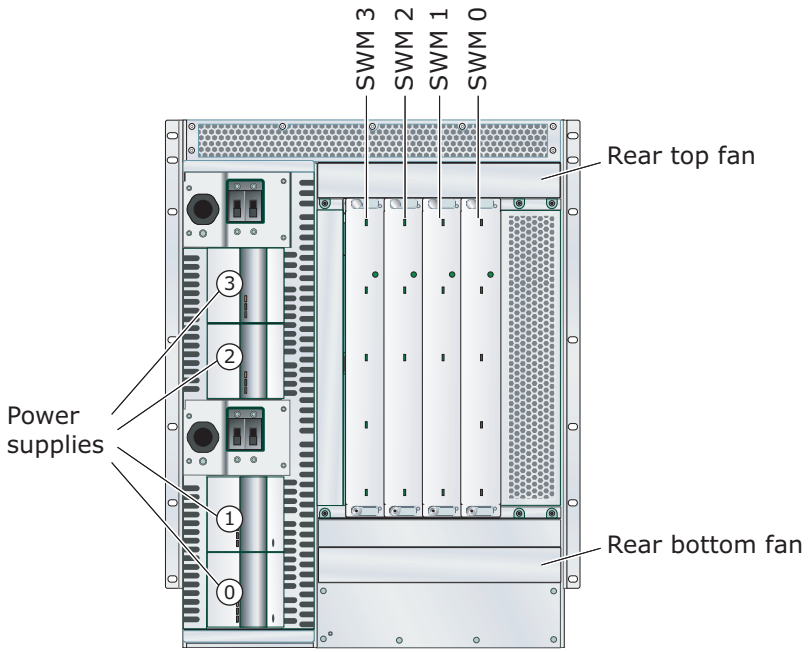


Figure 398 **FRUs on rear of ED-10000M**

Partitioning (FlexPars)

The ED-10000M has the capability to partition one physical director into four virtual directors. With the director partitioned, one can:

- Independently zone and administer each partition and its associated devices.
- Enable separate fabric services.
- Create independent configurations for all ports.

One can dynamically enable, disable, and assign resources (LIMs) to partitions based on applications or shared resource needs. Because each partition is independent, fabric faults are always isolated to a partition and do not degrade the operation of the entire director.

FlexPars lets you increase network utilization by sharing a director between applications or between test, production, and development environments of the same application. One has the ability to dynamically move resources (LIMs) between the director FlexPars.

Nodes in one partition cannot communicate to another partition without an ISL. To keep partitions as individual logical fabrics and

allow nodes to communicate between them (e.g., tape fabric), use an external Fibre Channel routing device such as an MP-2640M.

While each partition is a separate entity, all partitions on the same director share some system resources, such as the microprocessor on the CTP. Because of this resource sharing, switch administrators should keep in mind what each partition is being used for. In addition, for the first release, when configuring multiple partitions, you must ensure that the sum of N_Ports connected to each fabric in which each partition is participating does not exceed 1024.

Operating modes

The ED-10000M can operate in Brocade M-Series Fabric Mode or Open Fabric Mode.

Management

The ED-10000M should be managed by Connectrix Manager 8.06.01 or higher. You can also using Telnet to access the switch's CLI.

The switch can also be monitored using the FibreAlliance or proprietary ED-10000M SNMP MIBs.

Connectrix Manager is available on a Connectrix Service Processor provided by EMC or with a customer-supplied service processor, as described in the appendix in the *Connectrix Manager User's Guide*.

If the director is configured as a single director, all ports are accessible through a single Product Manager. If the director is partitioned (into as many as four FlexPars), ports are managed through multiple Product Managers. They would appear as individual director icons in Connectrix Manager. Only those assigned LIM(s) to a partition would be accessible by the administrator for that FlexPar.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix ED-140M

The Connectrix ED-140M is a Fibre Channel Enterprise Director. The ED-140M is installed in a Connectrix EC-1500-M cabinet, which also houses a Connectrix service processor.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the ED-140M.

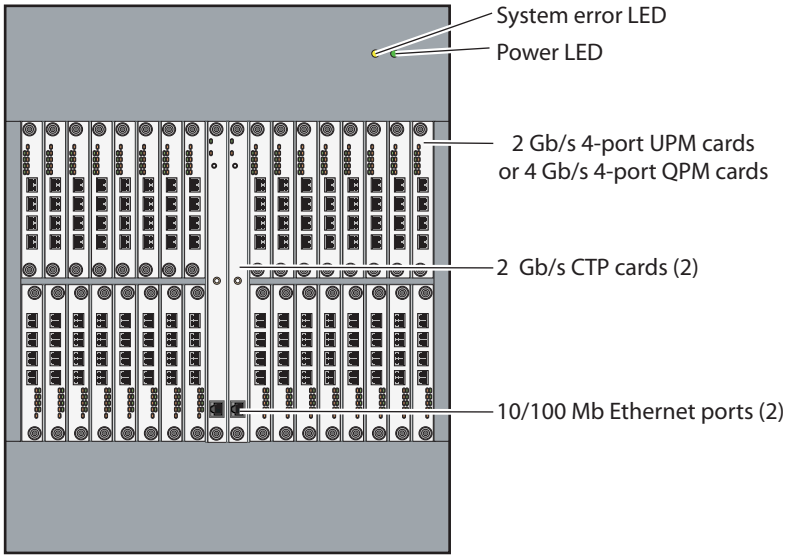


Figure 399 Connectrix ED-140M, front view

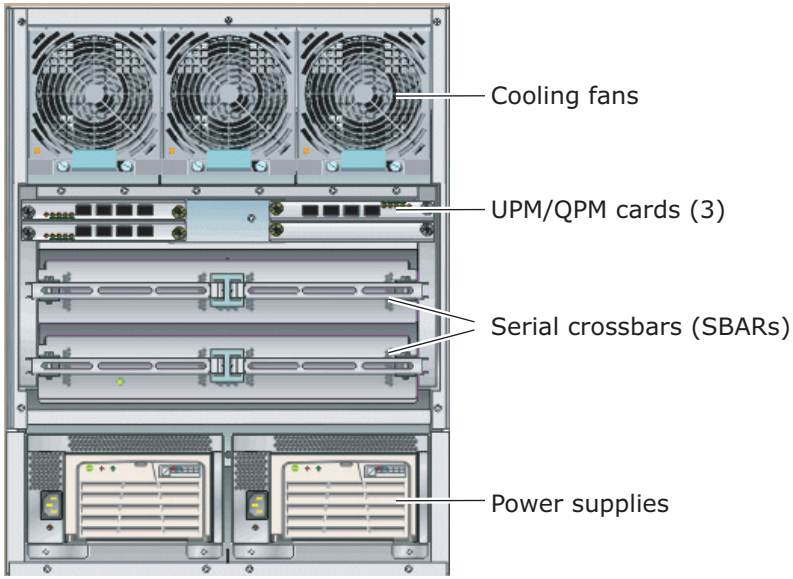


Figure 400 **Connectrix ED-140M, rear view**

Connectivity

Connectivity features include:

- 4 to 140 Fibre Channel ports, in 4-port increments (4 ports per port module):
  - Up to 128 ports in the front of the chassis
  - Up to 12 ports in the rear of the chassis
- 2 Gb/s Universal Port Module (UPM) in multimode fiber (shortwave); upgradeable to longwave.

  Note: 4 Gb/s Quad Port Module (QPM) in multimode fiber (shortwave); upgradeable to longwave.

- Four duplex SFP transceivers on each UPM/QPM.
- 60 BB_Credits available per port on UPM. QPM supports up to 125 BB_Credits.
- Supports Class 2 and Class 3 Fibre Channel protocols.
- All ports auto-negotiate to provide either switched N_Port or E_Port connections; UPM has G_Port functionality, determined by the type of attached device.
- All ports auto-negotiate for 1 Gb/s or 2 Gb/s transfer rate.

- Full duplex to 200 MB/s data rate per Fibre Channel port (at 2 Gb/s).

Management

Management features include:

- Centralized configuration and management of products and fabrics through Connectrix Manager, which is a client/server architecture. The Connectrix service processor laptop management station provides server functionality, and can also act as a client. Clients can be installed onto remote workstations running Windows NT, Windows 2000, Solaris, Linux, HP-UX, or AIX.
- Centralized management of up to 48 units.
- Support for online, non-disruptive code upgrades.
- 10/100 Mb Ethernet connections to CTP (control processor) for out-of-band management.
- Extensive centralized logging: Event, Audit, Hardware, and Session logs.
- Embedded Web Server, SNMP, and CLI support.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic communication to central support facility through modem connection
  - Automatic error notification sent through email
- Pre-emptive hardware health checking and redundant hardware switchover.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix M Quad Port Module (QPM)

Note: QPM support for the ED-140M Director debuted in E/OS 9.0 and was expanded to include the ED-64M in E/OS 9.1. However, due to cooling concerns, do *not* install a QPM in Slot **0** of an ED-64M.

This section provides a description of the hardware and features for the Quad Port Module (QPM). The information provided comes from Brocade M-Series product description documentation and testing experience during eLab's qualification. More information for the Brocade M-Series QPM Port Module can be found in the *EMC Connectrix M Enterprise Operating System (E/OS) Version 9.0 Release Notes*, available on Powerlink.

Because the QPM has only two fiber-optic ports in full 4 Gb/s mode, but consumes four backplane ports, the ED-140M Management interfaces will be modified to recognize a new port type or port attribute corresponding to real, but unmanaged, backplane ports or channels. When in full 4 Gb/s mode, a QPM port will be represented by its base port. For example, a QPM installed in the second slot of a ED-140M chassis will be identified as port numbers 4-7. But if port 4 is in full 4 Gb/s mode (which utilizes backplane channels 4-5), port number 5 will be disabled to the end user.

For distance solutions, QPM can can be configured for long link lengths over dark fiber or DWDM. QPM supports:

- 125 buffer credits per port, allowing Fibre Channel and FICON paths as long as 125 km at 2 Gb/s and 60 km at 4 Gb/s.
- At 150 meters (4.2500 Gb/s) through 50-micron multimode fiber-optic cable
- At 70 meters (4.2500 Gb/s) through 62.5-micron multimode fiber-optic cable

Note: Requires E/OS 9.0 and Connnectrix Manger v8.9.1.

Note: QPM is not supported for ED-64M with E/OSc 9.0.

Description

The Brocade M-Series QPM Port Module is based on Brocade M-Series's Sidewinder ASIC, providing four ports of 4 GB/s connectivity.

Note: With E/OSc 9.00.00, the QPM may only be used in the ED-140M (ED-6140). If a QPM is inserted into an ED-64M (ED-6064) running EOSc 9.00.00, the card will be faulted by the firmware during initialization and will be shown as "failed".

Key features of the QPM Port Module

Key features of the QPM Port Module include:

- Two different modes of operation: *burst* and *sustained*
- Auto-negotiates at 4 Gb/s, 2 Gb/s and 1 Gb/s
- Hot-pluggable and non-disruptive insertion into ED-140M chassis
- Hot-pluggable optical transceivers
- Inserts into any ED-140M Port Module slot
- Can fully populate ED-140M series chassis (Up to 35 per ED-140)
- Can intermix with UPM 2 Gb/s Port Modules and XPM 10 Gb/s Port Modules
- Shortwave (850nm) optical transceivers available for 150m links over standard 50-micron multimode fiber and OM2 cable (300m over OM3 cable)
- Longwave (1310nm) optical transceivers available for 10 km links over standard 9-micron single-mode fiber
- 60 km total path distance with full 4 Gb/s performance at full frame using 125 buffer-buffer credits per port
- Brocade M-Series Open Trunking compatible

Hardware architecture

A new line card called the QPM (4 Gigabit Port Module) is available for the ED-140M. The QPM card consists of a Sidewinder ASIC and 1/2/4 gigabit optics modules that support longwave and shortwave transceivers. Only changing the line card with the associated port logic enables reuse of the existing chassis, power supplies, SBAR, CTP, backplane, etc. The QPM utilizes link aggregation to pass 4 Gb/s data across the 2 Gb/s backplane. The Fibre interface for the QPM is ANSI Fibre Channel compliant and consists of a maximum of two 4 Gb/s links.

Figure 401 identifies port designations. The heavy black bars adjacent to ports 0 and 2 designate these ports as 4 G sustained mode capable.

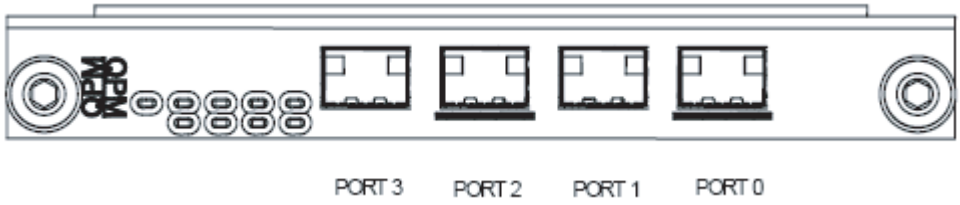


Figure 401 QPM 4 G faceplate

When one port is running in 4 Gb/s sustained mode, the adjacent port will be blocked. When operating at 1,2 or 4 Gb/s burst mode, all four FC ports can be enabled (see Figure 402).

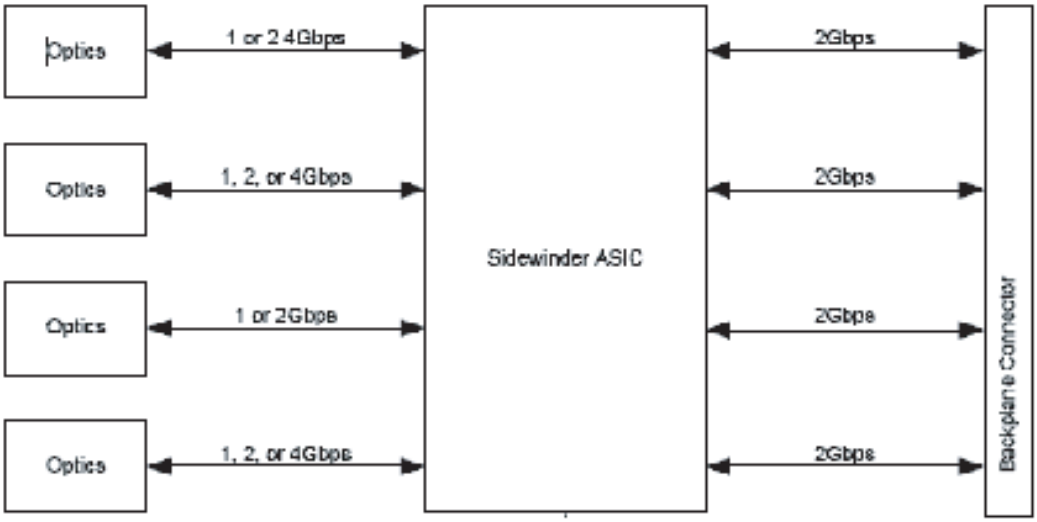


Figure 402 QPM 4 G

Burst versus sustained ports – Backplane data transfer

Although all QPM FC ports support 4 Gb/s data transfer rate, the backend paths only have 2 Gb/s bandwidth. However, two backend paths can be aggregated together to provide a combined 4 Gb/s data transfer rate to the backplane for an FC port.

- An FC port that is configured to utilize two backend paths, it is referred as a 4 Gb/s *sustained* port.

- An FC port with just one backend path is referred as a 4 Gb/s *burst* port since it has 4 Gb/s at FC side, but only 2 Gb/s at backplane.

Mode 1: Sustained ports

The QPM can be configured to have one or two 4 Gb/s sustained ports. In this configuration, each sustained port utilizes two backend paths and is allocated an aggregated bandwidth of 4 Gb/s.

Only ports 0 and 2 on the QPM module can be configured as 4 Gb/s sustained port. The accompanying ports 1 and 3 will be inactive and not available to users.

The user can configure a 4 Gb/s sustained port as either fixed at 4 Gb/s rate or capable of autonegotiation for automatic determination of rate.

In order to configure an E_Port for 4 Gb/s, sustained mode must be used.

Mode 2: Burst ports

The QPM can be configured to have two or four 4 Gb/s burst ports. Burst ports provide the ability to connect to devices in the SAN at 4 Gb/s speeds. In this configuration, each burst port utilizes a single backend path but has the ability to cache frames in the available buffers. This caching feature provides short-term receive bandwidth increases to alleviate network traffic “spikes” or “bursts”. The user can configure a 4 Gb/s burst port as either fixed at 4 Gb/s rate or capable of autonegotiation for automatic determination of rate.

A burst port cannot be configured for use as a G_Port or an E_Port.

Note: EMC recommends that all 4 Gb/s burst ports be configured to have the maximum Rx_BB_Credit value of 125.

Other modes

Any QPM port can also be configured as fixed 1 Gb/s, fixed 2 Gb/s, or 2 Gb/s autonegotiated. 2 Gb/s autonegotiated is the default port mode and provides for automatic determination of rates (i.e., autosensing and configuring to either 1 Gb/s or 2 Gb/s, but not 4 Gb/s).

These modes are identical to modes available on UPM.

QPM configurations

The QPM has four physical FC ports and provides users the flexibility to set QPM module into different hardware configurations for different requirements and desired throughputs. There are four basic configurations available. Users can choose any of the configurations for a QPM, or change its configuration if needed.

Configuration 1

To achieve the 4 Gb/s throughput, a QPM can be configured to have two 4 Gb/s sustained ports (see Figure 403 on page 1116). This configuration automatically results when the user configures both QPM ports as sustained ports. In this configuration, each 4 Gb/s sustained port utilizes two backend paths and is allocated an aggregated bandwidth of 4 Gb/s. The inbound frames are striped across two backend paths while the outbound frames are serialized before being transmitted. Only ports with indices 0 and 2 on the QPM module can be configured as 4 Gb/s sustained port; companying ports with indices 1 and 3 will be inactive.

When using this configuration, users should be aware of possible waste of bandwidth if connecting a 4 Gb/s sustained port to a 1 or 2 Gb/s device.

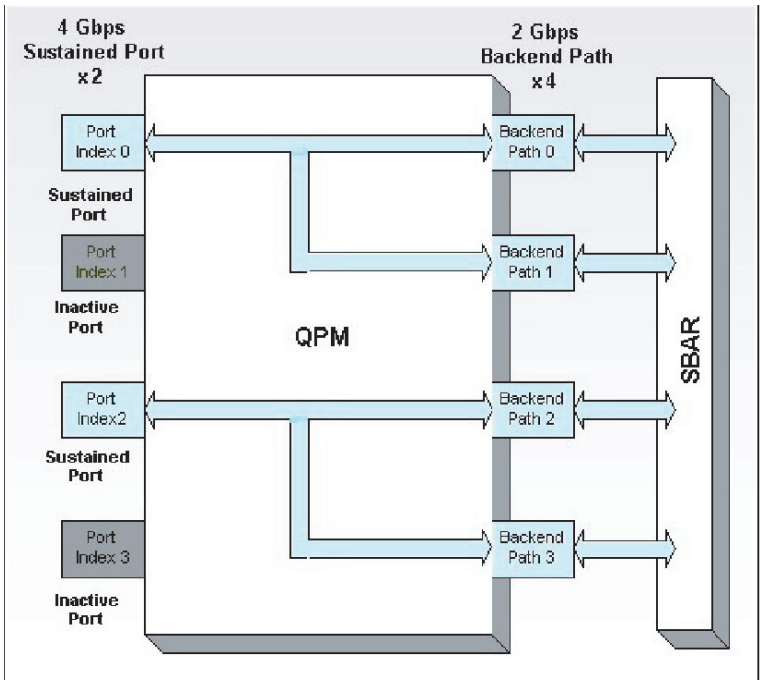


Figure 403 **QPM configuration 1**

Configuration 2

This configuration results when neither port is configured as a sustained port. This configuration may configure all four ports as 4 Gb/s burst ports. In this case, each port uses a single 2 Gb/s backend path (see Figure 404 on page 1117). In this configuration, a QPM works similar to a UPM.

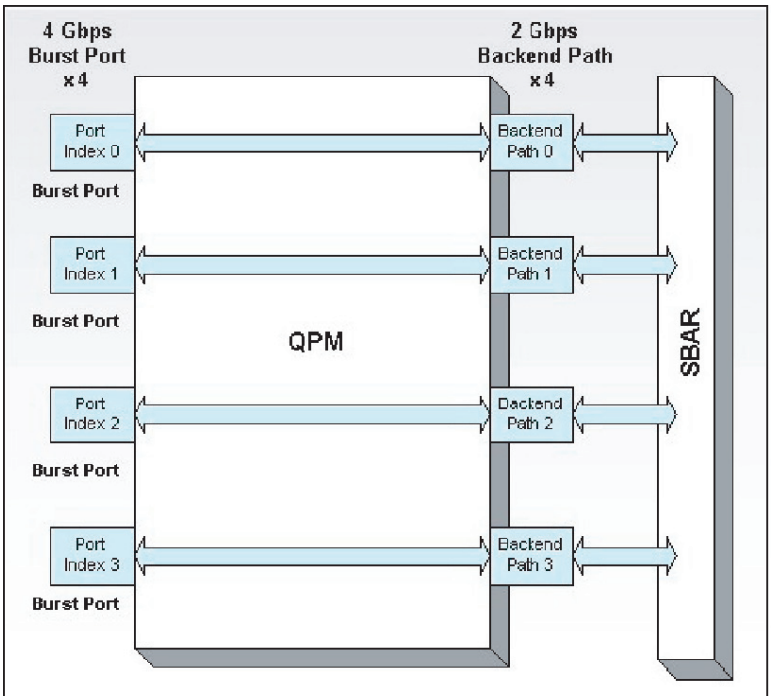


Figure 404 **QPM configuration 2**

Configuration 3 and 4

The QPM also supports the configuration of mixed 4 Gb/s sustained ports and non-sustained ports. As shown in Figure 405 on page 1118, configuration 3 and 4 are logically equivalent and both provides one 4 Gb/s sustained port utilizing combined bandwidth of two backend paths and two other ports each using a single backend path. Again only ports with indices 0 or 2 on the QPM module can be configured as 4 Gb/s sustained port and its companying port, indices 1 or 3, will be inactive.

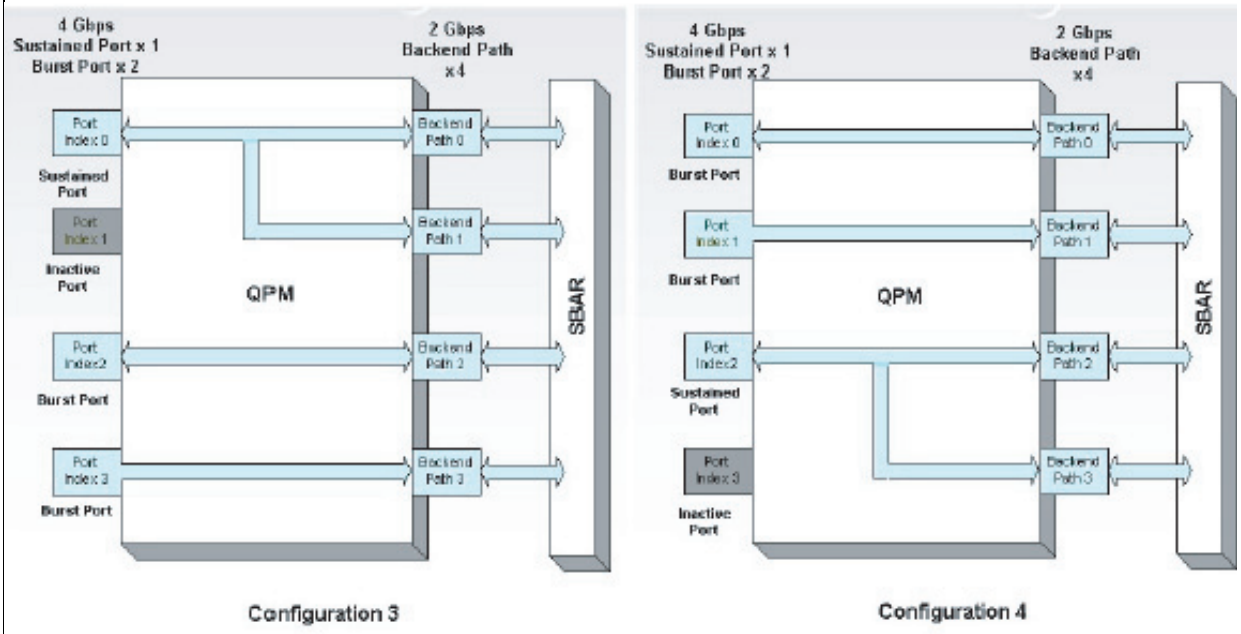


Figure 405 **QPM configuration 3 and 4**

Model numbers

Table 62, next, and Table 63 on page 1119 list the QPM model numbers.

Table 62 **QPM model numbers: New sales**

Model number	New sales
ED-QPM-04	ED-QPM 4 Gb/s 4-port card for ED-140M director, supporting up to 270m with multimode SPF.
ED-QPM-L04	ED-QPM 4 Gb/s 4-port card for ED-140M director, supporting up to 4 km with single-mode SPF.
ED-QPM-L10	ED-QPM 4 Gb/s 4-port card for ED-140M director, supporting up to 10 km with single-mode SPF.

Table 63 QPM model numbers: Upgrades

Model number	Upgrades
ED-QPM-04-U	Field upgrade ED-QPM 4 Gb/s 4-port card for ED-140M directors, supporting up to 270m with multimode SFP.
ED-QPM-L04-U	Field upgrade ED-QPM 4 Gb/s 4-port card for ED-140M directors, supporting up to 4 km with single-mode SFP.
ED-QPM-L10-U	Field upgrade ED-QPM 4 Gb/s 4-port card for ED-140M directors, supporting up to 10 km with single-mode SFP.

XPM port card

Figure 406 shows an XPM port card. An XPM port card is a 10 Gb/s port card for the ED-140M and ED-64M. The XPM contains one 10 Gb/s Fibre Channel port. It allows you to consolidate several 1/2 Gb/s ISLs down to one link. The XPM Blade can be installed in any 2 Gb/s capable ED-140M or ED-64M running E/OS 8.1.1 or higher and managed by Connectrix Manager v8.7.1 or higher. The XPM Blade is hot pluggable and is automatically recognized as a 10 Gb/s device upon insertion and is nondisruptive to system I/O when installed. XPM Blades can be installed in any ED-140M or ED-64M slot and in every slot if desired. Older ED-64Ms running at 1 Gb/s must first upgrade CTP-1s to 2 Gb/s CTP-2s before XPM blades can be installed.

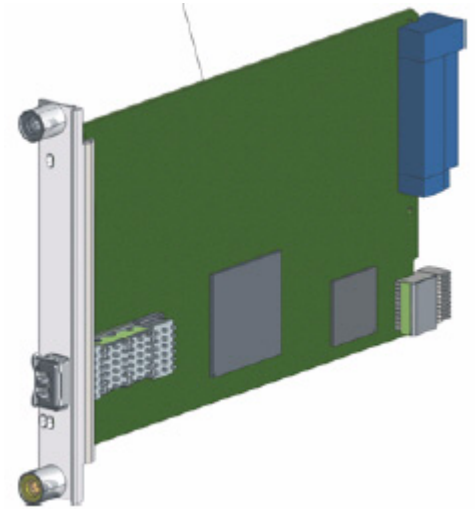


Figure 406 XPM card

Management

The XPM contains one external 10 Gb/s port. When shown in the ED-xxM Element Manager in Connectrix Manager, the port list will increment by four when viewing the port card after the XPM. For instance, if a universal port module (UPM) was in slot 7 of an ED-64M containing ports 28-31 and slot 8 contains an XPM, then the XPM port in slot 8 will be port 32. The ports in slot 9 will start at port 36 since the XPM in slot 8 occupies 4 backend ports. This four-port increment per XPM is reflected in Connectrix Manager. Currently, EMC ControlCenter does not increment the four ports, but rather displays 4 ports per XPM. This does not affect the performance and/or the ability to use the XPM. Only the displayed port count list is inaccurate for the XPM itself; all other port cards are displayed correctly in EMC ControlCenter.

Note: Connectrix Manager v8.7.1 and EOS v8.1.1 or higher must be used for the XPM. Lower versions of Connectrix Manager and EOS will not work.

System architecture

Performance

The full duplex signaling rate for the XPM blade is 10.5 Gb/s (this is the Fibre Channel Standard).

The sustained user data rate for the XPM blade is approximately 740 MBytes/sec (this is limited by the ED-140M/ED-64M backplane).

- This is approximately 97% of the 4 x 2Gb/s UPM ports theoretical data rate.
- XPM 10 Gb/s link is 3.7x the performance of a theoretical 2 Gb/s link.

Distance

The XPM was released Q2 2006. The Shortwave XPM (EMC model# ED-XPM-01) with a shortwave (850nm) optical transceiver will transmit up to 82m over standard 50 micron OM2 multimode fiber. This distance can be extended to 300 meters if high bandwidth 50 micron OM3 cable is used.

Longwave (1310nm) XPMs are available for 10 km links over standard 9 micron single-mode fiber (EMC model# ED-XPM-L10). Either the shortwave XPM or the longwave XPM can be upgraded to support a 40 km link over standard 9 micron single-mode fiber by installing the 40 km 10 gig LW (1550nm) XFP optic (EMC model# DSED-LWKT-XG4).

Since XPM blades have an approximate maximum sustained data rate of 740 MBs per second and have 400 BB_Credits, they can support distances of up to 100 Km with repeaters.

Usage

The XPM blade is used to reduce the number of ISLs between directors. These directors can be running in either Open or Brocade M-Series fabric mode. ISLs can be created between other ED-140Ms or ED-64Ms with XPMs or ED-10000Ms with 10 Gb/s ports. The 10 km or 40 km are ideal candidates for use to link remote sites.

Applications that may be good uses for the XPM blade include the following:

- Reduce number of core ISLs
- Long links between floors, buildings or sites where cost of path is significant compared to cost of ports
- BC/DR links
- Metro WDM – Typically few km to 100km

Note: Currently the ADVA 10 G DWDM has been qualified for use with the XPM. Check the E-Lab Navigator for the latest list of compatible DWDM systems.

The XPM can provide significant economic benefit by allowing the customer with a remote ISL application to reduce the number of leased dark fibers, delivering rapid return-on-investment. The XPM provides 400 BB_Credits which could support a 90 km link at full bandwidth with an average frame size of 1800 bytes or 105 km with a maximum frame size of 2100 bytes.

The XPM is recommended for consolidating multiple core director ISLs down to one 10 Gb/s ISL across one hop. As with 2 Gb/s ISLs, the practice of multiple ISLs remains to protect against any sort of hardware failure, therefore when implementing XPMs a minimum of two 10 Gb/s ISLs should be used between any two domains.

Special attention needs to be heeded to designing the data paths/flows through the fabric. Attempting to push 10 gig traffic to further downstream domains through slower speed ISLs can result in congestion which, if severe enough, could eventually back-pressure the ingress ports. Therefore, XPMs should only be used to transmit data 1 hop across domains from initiators to targets in the fabric.

Intermixing of 10 Gb/s and 4 and/or 2 Gb/s ISL from domain to domain is not recommended because the 10 Gb/s ISL will end up carrying all the traffic due to its lower FSPF cost.

ISLs from domain to domain should be speed-matched for optimum performance; i.e., 2 Gb/s and 2 Gb/s, 4 Gb/s and 4 Gb/s, 10 Gb/s and 10 Gb/s.

Models[a]	
ED-XPM-01	Single Port 10 Gb/s Short Wave Card w 500 meter XFP type Optic
ED-XPM-L10	Single Port 10 Gb/s Long Wave Card w 10 kilometer XFP type Optic
DSED-SWKT-XG	Short Wave Multi-mode 850nm XFP optic
DSED-LWKT-XG1	Long Wave Single-mode 1310nm 10km 10Gb/s XFP optic
DSED-LWKT-XG4	Long Wave Single-Mode 1550nm 40km 10Gb/s XFP optic

a. (**-***-***-U for upgrades)

Connectrix ED-64M

The Connectrix ED-64M (in 1 Gb/s and 2 Gb/s versions) is a Fibre Channel Enterprise Director. The ED-64M (shown in Figure 407) is installed in an EMC cabinet (described under "Connectrix DS-32B2" on page 1058) that also houses a Connectrix service processor.

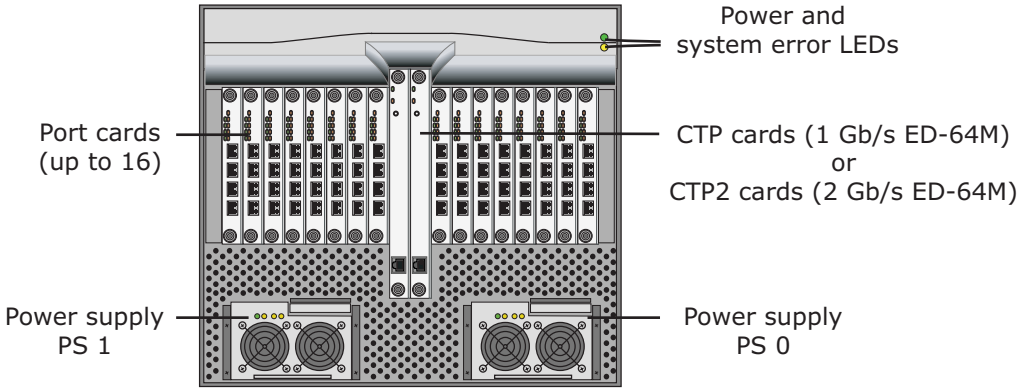


Figure 407 Connectrix ED-64M

Connectivity

Connectivity features include:

- 4 to 64 ports available in 4-port increments (4 ports per port module).
- Fibre Port Module (FPM) available in multimode fiber (shortwave).
- Universal Port Module (UPM) available on 2 Gb/s ED-64M in multimode fiber (shortwave); upgradeable to longwave.
- Single fiber-optic SFP on each FPM; four duplex SFP transceivers on each UPM.
- 60 buffer-to-buffer credits (BB_Credits) available per port.
- Supports Class 2 and Class 3 Fibre Channel protocols.
- All ports auto-negotiate to provide either switched N_Port or E_Port connections; UPM has G_Port functionality, determined by the type of attached device.
- Full duplex to 200 MB/s data rate per Fibre Channel port (in 2 Gb/s ED-64M).

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Management

Management features include:

- Centralized configuration and management of products and fabrics using Connectrix Manager, which is a client/server architecture. The Connectrix service processor laptop management station provides server functionality, and can also act as a client. Clients can be installed onto remote workstations running Windows NT, Windows 2000, Solaris, Linux, HP-UX, or AIX.
- Centralized management of up to 48 units.
- Support for online, non-disruptive code upgrades.
- 10/100 Mb Ethernet connections to CTP (control processor) for out-of-band management.

- Extensive centralized logging: Event, Audit, Hardware, and Session logs.
- Embedded Web Server, SNMP, and CLI support.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic communication to central support facility through modem connection
  - Automatic error notification sent through email
- Pre-emptive hardware health checking and redundant hardware switchover.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix ED-1032

The Connectrix ED-1032 is a 32-port fabric Director. The ED-1032 integrates high-speed Fibre Channel connectivity, advanced fault-tolerant switching technology, consulting, and services to help derive more value from your IT infrastructure. The combination of EMC Connectrix and EMC Enterprise Storage Network (ESN) extends the walls of the traditional data center. This breakthrough changes the entire data center concept, transforming the data center into a virtually limitless, accessible, and invaluable information resource.

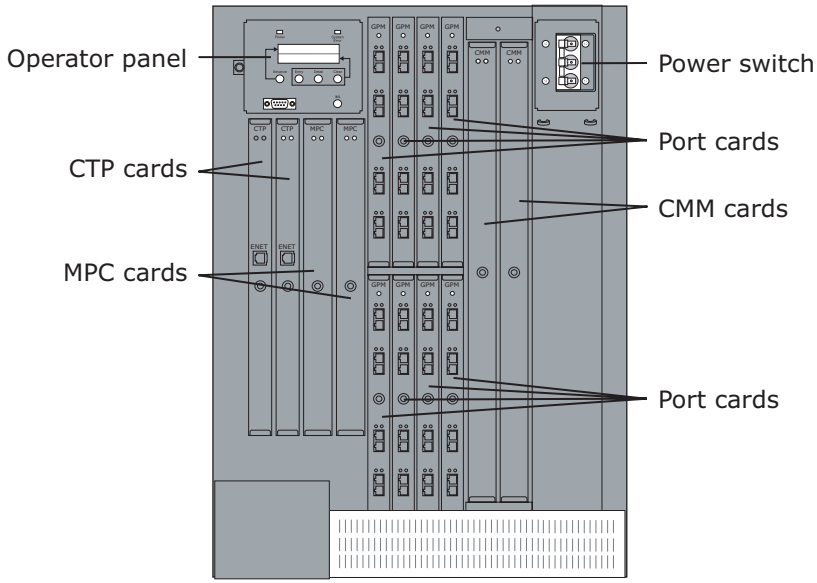


Figure 408 **Connectrix ED-1032**

Key ED-1032 benefits

Some key benefits are:

- Reaches more servers at greater distances from the enterprise.
- Enhances storage-server connectivity, information consolidation, and data center flexibility.
- Maximizes information availability through a fault-tolerant, self-healing architecture for unsurpassed business continuity.
- Centralizes monitoring and control for exceptional information and resource management in an expanding network.
- Promotes substantial business advantage across the entire enterprise using powerful EMC software solutions.
- Ensures thorough needs-analysis, orderly implementation and growth with EMC Professional Services.

Fabric configuration guidelines

The following sections present some general guidelines for setting up a fabric that includes Connectrix ED-1032.

Connections

Each ED-1032 provides support for 32 F_Port connections with the high-availability features. Connectrix provides capability to increase the storage capacity by connecting multiple Symmetrix Fibre Channel director ports to a single Fibre Channel HBA port on a single host. On the other hand, Connectrix provides the capability for more hosts to share a Symmetrix Fibre Channel director port by connecting multiple Fibre Channel HBA port on multiple hosts.

Software support

Connectrix software support is provided the Connectrix Manager. Connectrix Manager is a Java-based GUI that can be invoked as local or remote management software. For more information, refer to the Connectrix ED-1032 Director User Guide.

Diagnostics and maintenance

Repair and system recovery from a Fibre Channel switch failure can be done on line with a properly configured high-availability system.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-4700M

The Connectrix DS-4700M combines up to 32 4 Gb / s ports with high availability and ports-on-demand scalability to support mission-critical storage area networks (SANs). For complete investment protection, the DS-4700M supports 1 Gb / s, 2 Gb / s, and 4 Gb / s, and is available with a flexible architecture of 16, 24, or 32 ports.

The DS-4700M meets the needs of a broad range of networked storage applications, including storage consolidation, backup, and cluster environments and includes a rich optional feature set.

Switch description

The DS-4700M provides Fibre Channel connectivity through 32 generic mixed ports (GX_Ports). Ports operate at 1.0625, 2.1250, or 4.2500 gigabits per second, and can be configured as:

- Fabric ports (F_Ports), which provide direct connectivity for switched fabric devices.
- Expansion ports (E_Ports), which provide interswitch link (ISL) connectivity to fabric directors and switches.
- Fabric loop ports (FL_Ports), which provide connectivity and fabric attachment for Fibre Channel arbitrated loop (FC-AL) devices.

The switch is installed on a table or desktop, mounted in an EC-1500 equipment cabinet, or mounted in any standard 19-inch equipment rack.

The switch provides two port buffer pools of 235 receive BB_Credits each (470 credits total). Ports 0 through 3, 8 through 11, 16 through 19, and 24 through 27 share buffer pool zero. Ports 4 through 7, 12 through 15, 20 through 23, and 28 through 31 share buffer pool one. Each port can be assigned between two and 120 credits, provided the

total credits allocated to all ports in a pool does not exceed 235. The default value is 12 credits per port.

The switch provides a modular design that enables quick removal and replacement of FRUs, including small form factor pluggable (SFP) optical transceivers and power supply assemblies.

Figure 409 shows the front of the DS-4700M.

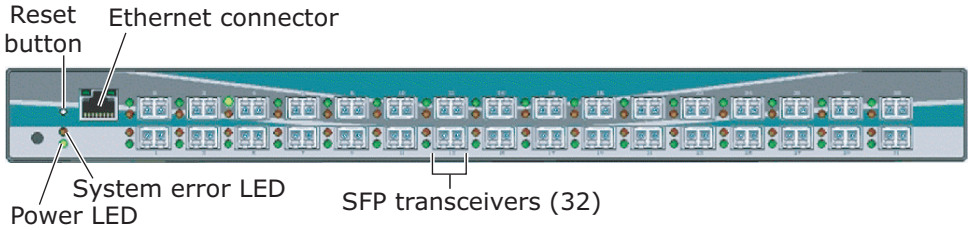


Figure 409 DS-4700M (front view)

Figure 410 shows the rear of the DS-4700M.

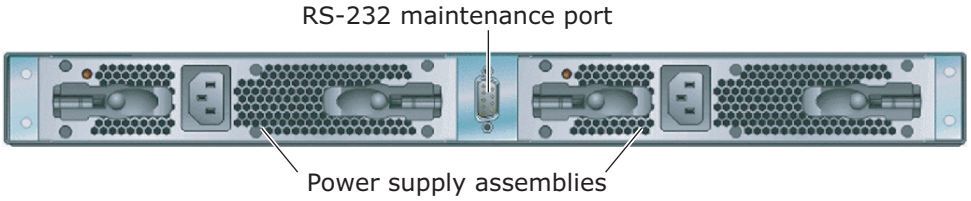


Figure 410 DS-4700M (rear view)

SFP transceiver

Single-mode or multimode fiber-optic cables attach to switch ports through SFP transceivers. The fiber-optic transceivers provide duplex LC connectors, and can be detached from switch ports (through a 10-pin interface) for easy replacement.

The following fiber-optic transceiver types are available:

- Tri-rate shortwave laser (1.0625, 2.1250, or 4.2500 Gb/s) — Shortwave laser transceivers (850 nm) provide connectivity:
  - At 500 meters (1.0625 Gb/s) through 50-micron multimode fiber-optic cable
  - At 300 meters (2.1250 Gb/s) through 50-micron multimode fiber-optic cable

- At 150 meters (4.2500 Gb / s) through 50-micron multimode fiber-optic cable
- At 300 meters (1.0625 Gb / s) through 62.5-micron multimode fiber-optic cable
- At 150 meters (2.1250 Gb / s) through 62.5-micron multimode fiber-optic cable
- At 70 meters (4.2500 Gb / s) through 62.5-micron multimode fiber-optic cable

◆ Tri-rate longwave laser (1.0625, 2.1250, or 4.2500 Gb / s) — Longwave laser transceivers (1310 nm) provide connectivity at 10 kilometers through 9-micron single-mode fiber-optic cable.

The switch also provides a predictive optics monitoring (POM) feature that monitors operation of SFP optical transceivers. Digital diagnostics-enabled optical transceivers report temperature, voltage current, transceiver power, and receiver power to product firmware. Optical transceivers also provide vendor-specific threshold values for these parameters.

Power supply assembly

The DS-4700M contains two power supply assemblies with internal cooling fans. These redundant, load-sharing assemblies step down and rectify facility input power to provide 12 VDC to the control processor (CTP) card. The power supplies also provide input filtering, overvoltage protection, and overcurrent protection.

Either power supply can be replaced while the switch is operational. Each power supply has a separate connection to the CTP card to allow for independent AC power sources. The power supplies are input rated at 90 to 264 VAC.

Three cooling fans integrated in each power supply (six fans total) provide cooling for the power supplies and CTP card, as well as redundancy for continued operation if a single fan fails. Fans are removed and replaced as part of the integrated power supply.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

System architecture

DS-4700M architecture includes:

- Fibre Channel standards — FC-PH, FC-PH-2, FC-PH3, FC-GS-2, FC-FLA, FC-FG, FC-SW2
- Fibre Channel ports — 16 universal ports in base model; up to 32 ports maximum, in 8-port increments.
- Cascade limit — Up to 16 switches, three hops (depending on configuration)

  Note: E-Lab Navigator describes the latest fabric configuration limits.

- Performance — Full line-speed switching 4.25 Gb/s
- Bandwidth — 128 Gb/s
- Switch core — Non-blocking design
- Fabric latency — Less than two microseconds average measured across two electrically farthest ports within the switch product
- Maximum frame size — 2,112 bytes
- Classes of Service — Class 2 and 3
- Code load — Hot code load and activation
- Port types — E_Port, F_Port, and G_Port
- Media types — Hot-pluggable, industry-standard, LC small form-factor pluggable (SFPs); shortwave SFPs standard; longwave optional
- Fabric services — Simple name server, Registered State Change Notification (RSCN), Alias Server (Multicast) and Zoning
- Hot-swappable components — Redundant power supplies, redundant cooling fans, and hot-swappable optics
- Optional features — SANtegrity Security Suite, Open Trunking, Performance Monitoring/Event Management, Group Configuration Manager, FICON Management Server
- Installation options — Standard configuration in EMC EC-1500-M cabinet, CLARiiON rack, or 19-inch EIA-compliant rack

Switch specifications

This section lists physical characteristics, storage and shipping environment, operating environment, and service clearances.

Physical characteristics

Dimensions:

- Height — 1.6 in. (4.1 cm) (1 rack unit)
- Width — 17.2 in. (43.7 cm)
- Depth — 15.5 in. (39.4 cm)
- Weight — 15.0 lb (6.8 kg)

Power requirements:

- Input voltage: 90 to 264 VAC
- Input current: 0.35 amps at 208 VAC
- Input frequency: 47 to 63 Hz

Heat dissipation — 71 W (242 Btu / hr)

Cooling airflow clearances (switch chassis):

- Right and left side — 0.5 in. (1.3 cm)
- Front and rear — 3.0 in. (7.6 cm)
- Top and bottom — No clearance required

Shock and vibration tolerance — 60 Gs for 10 ms without nonrecoverable errors

Acoustical noise — 70 dB “A” scale

Inclination — 10º maximum

Environment

Protective packaging must be provided to protect the switch under all shipping methods (domestic and international).

Enviromental requirements for shipping:

- Temperature — -40º F to 140º F (-40º C to 60º C)
- Relative humidity — 5% to 100%
- Maximum wet-bulb temperature — 84º F (29º C)
- Altitude — 40,000 ft (12,192 m)

Enviromental requirements for storage:

- Temperature — 34º F to 140º F (1º C to 60º C)

- Relative humidity — 5% to 80%
- Maximum wet-bulb temperature — 84º F (29º C)
- Altitude — 40,000 ft (12,192 m)

Enviromental requirements for operation:

- Temperature — 40º F to 104º F (4º C to 40º C)
- Relative humidity — 8% to 80%
- Maximum wet-bulb temperature — 81º F (27º C)
- Altitude — 10,000 ft (3,048 m)

Switch management

The DS-4700M is managed and controlled through a customer-supplied PC platform with Internet communication to the product-resident Connectrix Manager Basic Edition interface. The interface allows performing configuration tasks, view system alerts and related log information, and monitor switch status, port status, and performance. FRU status and system alert information are highly visible.

Note: Connectrix Manager Basic Edition is the revised Embedded Web Server application (also called SANPilot) that provides a Web interface more similar to Connectrix Manager.

The DS-4700M can also be managed by Connectrix Manager v8.7.1 running on the Connectrix 1U service processor. Connectix Manager provides a central point of managment for all the switches and directors in the SAN. Direct management of each managed switch is provided by the Element Manager application within Connectrix Manager.

Reliability, availability, and serviceability

The DS-4700M provides the following error detection, reporting, and serviceability features:

- LEDs on switch FRUs and adjacent to Fibre Channel ports provide visual indicators of hardware status or malfunctions.
- Redundant FRUs (SFP transceivers and power supply assemblies) can be removed or replaced without disrupting switch or Fibre Channel link operation.
- A modular design enables quick removal and replacement of FRUs without the use of tools or equipment.

- System alerts and logs display switch, Ethernet link, and Fibre Channel link status at the Connectrix Manager Basic Edition interface, client communicating with the management server, or customer-supplied server (running a Connectrix Manager application).
- Diagnostic software performs power-on self-tests (POSTs) and port diagnostics (loopback tests).
- An RS-232 maintenance port at the rear of the switch enables installation or service personnel to change the switch's IP address, subnet mask, and gateway address. (Access to the maintenance port is password protected.)

  These parameters can also be changed through a Telnet session, access for which is provided through a local or remote PC with an Internet connection to the switch.
- Data collection through the Connectrix Manager Basic Edition interface or Element Manager application helps isolate system problems. The data includes a memory dump file and audit, hardware, and engineering logs.
- Beaconing assists service personnel in locating a specific port or switch. When port beaconing is enabled, the amber LED associated with the port flashes. When unit beaconing is enabled, the system error indicator on the front panel flashes. Beaconing does not affect port or switch operation.
- An internal modem can be used by support personnel to dial-in to the management server (optional) for event notification and to perform remote diagnostics.
- Automatic notification of significant system events (to support personnel or administrators) occurs through email messages or the call-home feature.

  Note: The call-home feature is not available through the Connectrix Manager Basic Edition.

- SNMP management can be performed using the Fibre Channel Fabric Element MIB, Transmission Control Protocol / Internet Protocol (TCP / IP) MIB-II definition (RFC 1157), or a product-specific private enterprise MIB that runs on the switch. Up to six authorized management workstations can be configured through the Connectrix Manager Basic Edition

interface or Element Manager application to receive unsolicited SNMP trap messages. The trap messages indicate product operational state changes and failure conditions.

- Optional SNMP management can be performed using the Fibre Alliance MIB that runs on the management server. Up to 12 authorized management workstations can be configured through the Connectrix Manager application to receive unsolicited SNMP trap messages. The trap messages indicate operational state changes and failure conditions.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-4400M

The Connectrix DS-4400M combines up to 16 4 Gb/s ports with high availability and ports-on-demand scalability. The DS-4400M is ideal for a range of small-to-midrange applications, including the creation of a simple, entry-level SAN. For pay-as-you-grow scalability, the DS-4400M starts from a base of eight ports and grows in four-port increments. The DS-4400M is the only full-feature Fibre Channel switch available in a half-rack-width configuration. The DS-4400M meets the needs of a broad range of networked storage applications, including storage consolidation, backup, and cluster environments.

Switch description

The DS-4400M provides Fibre Channel connectivity through 16 generic mixed ports (GX_Ports). Ports operate at 1.0625, 2.1250, or 4.2500 gigabits per second, and can be configured as:

- Fabric ports (F_Ports), to provide direct connectivity for switched fabric devices.
- Expansion ports (E_Ports), to provide interswitch link (ISL) connectivity to fabric directors and switches.
- Fabric loop ports (FL_Ports), to provide connectivity and fabric attachment for Fibre Channel arbitrated loop (FC-AL) devices.

The switch is installed on a table or desktop, mounted in an EC-1500 equipment cabinet, or mounted in any standard 19-inch equipment rack.

The DS-4400M provides a port buffer pool of 150 receive BB_Credits. Each port can be assigned between two and 120 credits, provided the total credits allocated to all ports does not exceed 150. The default value is six credits per port.

The switch provides a modular design that enables quick removal and replacement of FRUs, including small form factor pluggable (SFP) optical transceivers and power supply assemblies.

Figure 411 shows the front of the DS-4400M.

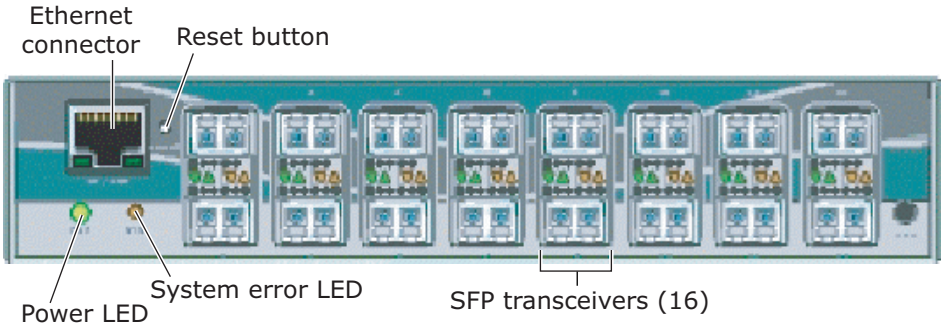


Figure 411 DS-4400M (front view)

Figure 412 shows the rear of the DS-4400M.

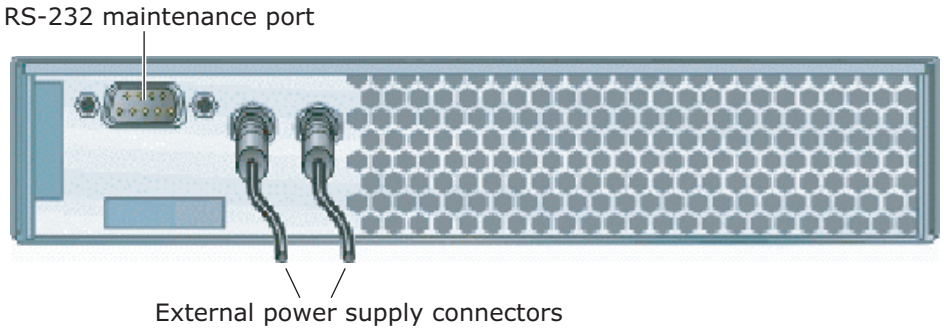


Figure 412 DS-4400M (rear view)

SFP transceiver

Single-mode or multimode fiber-optic cables attach to switch ports through SFP transceivers. The fiber-optic transceivers provide duplex LC connectors, and can be detached from switch ports (through a 10-pin interface) for easy replacement.

The following fiber-optic transceiver types are available:

- Tri-rate shortwave laser (1.0625, 2.1250, or 4.2500 Gb/s) — Shortwave laser transceivers (850 nm) provide connectivity:
  - At 500 meters (1.0625 Gb/s) through 50-micron multimode fiber-optic cable
  - At 300 meters (2.1250 Gb/s) through 50-micron multimode fiber-optic cable
  - At 150 meters (4.2500 Gb/s) through 50-micron multimode fiber-optic cable
  - At 300 meters (1.0625 Gb/s) through 62.5-micron multimode fiber-optic cable
  - At 150 meters (2.1250 Gb/s) through 62.5-micron multimode fiber-optic cable
  - At 70 meters (4.2500 Gb/s) through 62.5-micron multimode fiber-optic cable
- Tri-rate longwave laser (1.0625, 2.1250, or 4.2500 Gb/s) — Longwave laser transceivers (1310 nm) provide connectivity at 10 kilometers through 9-micron single-mode fiber-optic cable.

The switch also provides a predictive optics monitoring (POM) feature that monitors operation of SFP optical transceivers. Digital diagnostics-enabled optical transceivers report temperature, voltage current, transceiver power, and receiver power to product firmware. Optical transceivers also provide vendor-specific threshold values for these parameters.

Power supply assembly

The DS-4400M is delivered with two external power supplies. The power supplies step down and rectify facility input power to provide 12 volts direct current (VDC) to the control processor (CTP) card. The power supplies also provide input filtering, overvoltage protection, and overcurrent protection. Each power supply is input rated at 90 to 264 VAC.

When two power supplies are connected the switch automatically enables high availability (HA) mode. Either power supply can be replaced while the switch is operational. Each power supply has a separate connection to the CTP card to allow for independent AC power sources.

Three internal cooling fans provide airflow for the CTP card, as well as redundancy for continued operation if a single fan fails.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

System architecture

DS-4400M architecture includes:

- Fibre Channel standards — FC-PH, FC-PH-2, FC-PH3, FC-GS-2, FC-FLA, FC-FG, FC-SW2
- Fibre Channel ports — 16 universal
- Cascade limit — Up to 24 switches, three hops (depending on configuration)

  Note: E-Lab Navigator describes the latest fabric configuration limits.

- Performance — Full line-speed switching 4.25 Gb/s
- Bandwidth — 128 Gb/s
- Switch core — Non-blocking design
- Fabric latency — Less than two mcroseconds average measured across two electrically farthest ports within the switch product
- Maximum frame size — 2,112 bytes
- Classes of Service — Class 2 and 3
- Port types — E_Port, F_Port, and G_Port
- Media types — Hot-pluggable, industry-standard, LC small form-factor pluggable (SFPs); shortwave SFPs standard; longwave optional
- Fabric services — Simple name server, Registered State Change Notification (RSCN), Alias Server (Multicast) and Zoning
- Hot-swappable components — Redundant power supplies, redundant cooling fans, and hot-swappable optics
- Optional features — SANtegrity Security Suite, Open Trunking, Performance Monitoring/Event Management, Group Configuration Manager

- Installation options — Standard configuration in EMC EC-1500-M cabinet, CLARiiON rack, or 19-inch EIA-compliant rack

Switch specifications

This section lists physical characteristics, storage and shipping environment, operating environment, and service clearances.

Physical characteristics

Dimensions:

- Height — 1.6 in. (4.1 cm) (1 rack unit)
- Width — 7.8 in. (19.9 cm)
- Depth — 13.1 in. (33.3 cm), plus 2.4 in. (6.1 cm) for external power supplies
- Weight — 8.8 lb (4.0 kg)

Power requirements:

- Input voltage: 90 to 264 VAC
- Input current: 0.18 amps at 208 VAC
- Input frequency: 47 to 63 Hz

Heat dissipation — 36 W (123 Btu / hr)

Cooling airflow clearances (switch chassis):

- Right and left side — 0.5 in. (1.3 cm)
- Front and rear — 3.0 in. (7.6 cm)
- Top and bottom — No clearance required

Shock and vibration tolerance — 60 Gs for 10 ms without nonrecoverable errors

Acoustical noise — 70 dB “A” scale

Inclination — 10º maximum

Environment

Protective packaging must be provided to protect the switch under all shipping methods (domestic and international).

Enviromental requirements for shipping:

- Temperature — -40º F to 140º F (-40º C to 60º C)
- Relative humidity — 5% to 100%
- Maximum wet-bulb temperature — 84º F (29º C)

- Altitude — 40,000 ft (12,192 m)

Enviromental requirements for storage:

- Temperature — 34º F to 140º F (1º C to 60º C)
- Relative humidity — 5% to 80%
- Maximum wet-bulb temperature — 84º F (29º C)
- Altitude — 40,000 ft (12,192 m)

Enviromental requirements for operation:

- Temperature — 40º F to 104º F (4º C to 40º C)
- Relative humidity — 8% to 80%
- Maximum wet-bulb temperature — 81º F (27º C)
- Altitude — 10,000 ft (3,048 m)

Switch management

The DS-4400M is managed and controlled through a customer-supplied PC platform with Internet communication to the product-resident Connectrix Manager Basic Edition interface. The interface allows performing configuration tasks, view system alerts and related log information, and monitor switch status, port status, and performance. FRU status and system alert information are highly visible.

Note: Connectrix Manager Basic Edition is the revised Embedded Web Server application (also called SANPilot) that provides a Web interface more similar to Connectrix Manager.

The DS-4400M can also be managed by Connectrix Manager v8.7.1 running on the Connectrix 1U service processor. Connectix Manager provides a central point of managment for all the switches and directors in the SAN. Direct management of each managed switch is provided by the Element Manager application within Connectrix Manager.

Reliability, availability, and serviceability

The DS-4400M provides the following error detection, reporting, and serviceability features:

- LEDs on switch FRUs and adjacent to Fibre Channel ports provide visual indicators of hardware status or malfunctions.
- Redundant FRUs (SFP transceivers and power supply assemblies) can be removed or replaced without disrupting switch or Fibre Channel link operation.

- A modular design enables quick removal and replacement of FRUs without the use of tools or equipment.
- System alerts and logs display switch, Ethernet link, and Fibre Channel link status at the Connectrix Manager Basic Edition interface, client communicating with the management server, or customer-supplied server (running a Connectrix Manager application).
- Diagnostic software performs power-on self-tests (POSTs) and port diagnostics (loopback tests).
- An RS-232 maintenance port at the rear of the switch enables installation or service personnel to change the switch's IP address, subnet mask, and gateway address. (Access to the maintenance port is password protected.)

  These parameters can also be changed through a Telnet session, access for which is provided through a local or remote PC with an Internet connection to the switch.
- Data collection through the Connectrix Manager Basic Edition interface or Element Manager application helps isolate system problems. The data includes a memory dump file and audit, hardware, and engineering logs.
- Beaconing assists service personnel in locating a specific port or switch. When port beaconing is enabled, the amber LED associated with the port flashes. When unit beaconing is enabled, the system error indicator on the front panel flashes. Beaconing does not affect port or switch operation.
- An internal modem can be used by support personnel to dial-in to the management server (optional) for event notification and to perform remote diagnostics.
- Automatic notification of significant system events (to support personnel or administrators) occurs through email messages or the call-home feature.

  Note: The call-home feature is not available through the Connectrix Manager Basic Edition.

- SNMP management can be performed using the Fibre Channel Fabric Element MIB, Transmission Control Protocol / Internet Protocol (TCP / IP) MIB-II definition (RFC 1157), or a product-specific private enterprise MIB that runs on the switch. Up to six authorized management workstations can be

configured through the Connectrix Manager Basic Edition interface or Element Manager application to receive unsolicited SNMP trap messages. The trap messages indicate product operational state changes and failure conditions.

- Optional SNMP management can be performed using the Fibre Alliance MIB that runs on the management server. Up to 12 authorized management workstations can be configured through the Connectrix Manager application to receive unsolicited SNMP trap messages. The trap messages indicate operational state changes and failure conditions.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-24M2

The Connectrix DS-24M2 is a 2 Gb/s Fibre Channel Departmental Switch. The DS-24M2 is installed in a Connectrix EC-1200 cabinet, which also houses a Connectrix service processor.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-24M2.

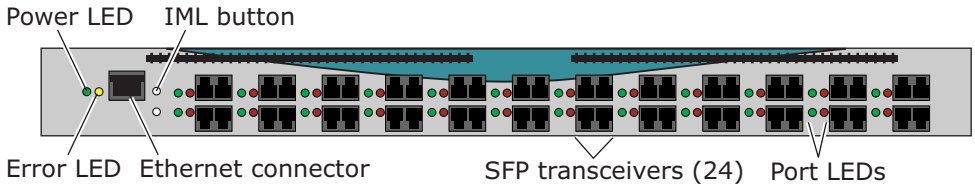


Figure 413 Connectrix DS-24M2

Connectivity

Connectivity features include:

- 24 ports of non-blocking 2 Gb/s Fibre Channel switching in a 1U form factor:
  - F_Ports (fabric ports) to provide direct connectivity for up to 24 switched fabric devices.

- FL_Ports (fabric loop ports) to provide arbitrated loop connectivity and fabric attachment for FC-AL devices. Each FL_Port can theoretically support the connection of 126 FC-AL devices.
- E_Ports (expansion ports) to provide interswitch link (ISL) connectivity to fabric directors and switches.

◆ Extended distance support — Through repeaters or Dense Wave Division Multiplexing (DWDM) equipment, switch ports can be configured for extended-distance operation. The DS-24M2's maximum BB_Credit value is 12, and only on ports 0 through 3. Therefore, use port 0, 1, 2, or 3 for 10 km LW optics.

◆ Supports Class 2 and Class 3 Fibre Channel protocols.

◆ Full duplex 200 MB/s data rate per Fibre Channel port.

◆ LC (Small Form Factor) switch port connectors.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Management

Management features include:

◆ Centralized configuration and management of products and fabrics using Connectrix Manager, which is a client/server architecture. The Connectrix service processor laptop management station provides server functionality, and can also act as a client. Clients can be installed onto remote workstations running Windows NT, Windows 2000, Solaris, Linux, HP-UX, or AIX.

◆ Centralized management of up to 48 units.

◆ Support for online, non-disruptive code upgrades.

◆ 10/100 Mb Ethernet connections to CTP (control processor) for out-of-band management.

◆ Extensive centralized logging: Event, Audit, Hardware, and Session logs.

- Embedded Web Server and SNMP support.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic communication to central support facility through modem connection
  - Automatic error notification sent through email
- Pre-emptive hardware health checking and redundant hardware switchover.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-32M, DS-32M2

The Connectrix DS-32M (1 Gb / s) and DS-32M2 (2 Gb / s) are Fibre Channel Departmental Switches. The DS-32M / DS-32M2 is installed in a Connectrix EC-1200 cabinet, which also houses a Connectrix service processor.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-32M / DS-32M2.

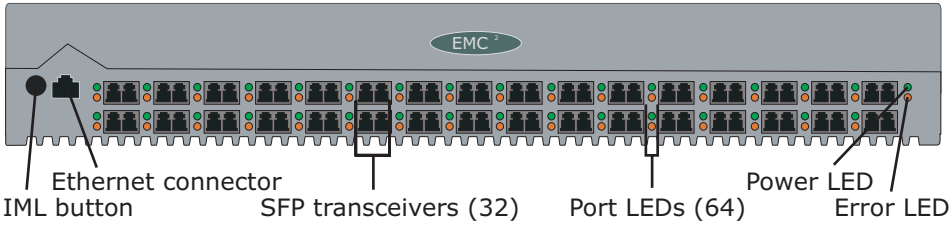


Figure 414 **Connectrix DS-32M, D32M2**

Connectivity

Connectivity features include:

- 32 ports available.
- Single fiber-optic SFP (small form-factor pluggable) on each port enables connectivity.
- Multimode fiber (shortwave) support.

- 60 buffer-to-buffer credits (BB_Credits) available per port.
- Supports Class 2 and Class 3 Fibre Channel protocols.
- All ports auto-negotiate to provide either switched N_Port or E_Port connections.
- Full duplex 200 MB/s data rate per Fibre Channel port in DS-32M2.
- LC (Small Form Factor) switch port connectors.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Management

Management features include:

- Centralized configuration and management of products and fabrics through Connectrix Manager, which is a client/server architecture. The Connectrix service processor laptop management station provides server functionality, and can also act as a client. Clients can be installed onto remote workstations running Windows NT, Windows 2000, Solaris, Linux, HP-UX, or AIX.
- Centralized management of up to 48 units.
- Support for online, non-disruptive code upgrades.
- 10/100 Mb Ethernet connections to CTP (control processor) for out-of-band management.
- Extensive centralized logging: Event, Audit, Hardware, and Session logs.
- Embedded Web Server and SNMP support.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic communication to central support facility through modem connection

- Automatic error notification sent through email
- Pre-emptive hardware health checking and redundant hardware switchover.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware / Platforms Documentation > Connectrix Directors and Switches.

Connectrix DS-16M, DS-16M2

The Connectrix DS-16M (1 Gb / s) and DS-16M2 (2 Gb / s) are Fibre Channel Departmental Switches. The DS-16M / DS-16M2 is installed in a Connectrix EC-1200 cabinet, which also houses a Connectrix service processor.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the DS-16M / DS-16M2.

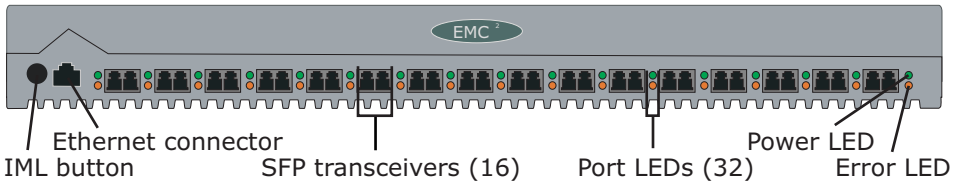


Figure 415 **Connectrix DS-16M, DS-16M2**

Connectivity

Connectivity features include:

- 16 ports available.
- Single fiber-optic SFP on each port enables connectivity.
- Multimode fiber (shortwave) support.
- 60 buffer-to-buffer credits (BB_Credits) available per port.
- Supports Class 2 and Class 3 Fibre Channel protocols.
- All ports auto-negotiate to provide either switched N_Port or E_Port connections.
- Full duplex 200 MB / s data rate per Fibre Channel port in DS-16M2.
- LC (Small Form Factor) switch port connectors.

Supported features

For the most up-to-date information on supported features, consult the *EMC Support Matrix* located on the E-Lab Interoperability Navigator. If you need further instructions on using the *EMC Support Matrix*, refer to "EMC Support Matrix and E-Lab Interoperability Navigator" on page 31.

Unsupported features

For the most up-to-date information on unsupported features, consult the most current version of the appropriate Connectrix M Series release notes and product documentation on Powerlink.

Management

Management features include:

- Centralized configuration and management of products and fabrics through Connectrix Manager, which is a client/server architecture. The Connectrix service processor laptop management station provides server functionality, and can also act as a client. Clients can be installed onto remote workstations running Windows NT, Windows 2000, Solaris, Linux, HP-UX, or AIX.
- Centralized management of up to 48 units.
- Support for online, non-disruptive code upgrades.
- 10/100 Mb Ethernet connections to CTP (control processor) for out-of-band management.
- Extensive centralized logging: event, audit, hardware, and session logs.
- Embedded Web Server and SNMP support.

Reliability, availability, and serviceability

Reliability, availability, and serviceability features include:

- Error notification:
  - Automatic communication to central support facility through modem connection
  - Automatic error notification sent through email
- Pre-emptive hardware health checking and redundant hardware switchover.

Further reading

Several related EMC Connectrix M manuals and release notes are available on Powerlink, under Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Connectrix Directors and Switches.

Connectrix M Series (Brocade M-Series) Open Trunking feature

The Open Trunking feature monitors the average data rates of all traffic flow on ISLs (from a receive port to a target domain), and periodically adjusts routing tables to reroute data flows from congested links to lightly loaded links and optimize bandwidth use.

The objective of Open Trunking is to make the most efficient possible use of redundant ISLs between neighboring switches, even if these ISLs have different bandwidths. Load balancing across the ISLs does not require user configuration, other than enabling Open Trunking. However, you can modify or tweak default settings for congestion thresholds (per port) and low BB_Credit threshold if desired. In particular, you do not need to manually configure ISLs into trunk groups of redundant links where data can be off-loaded. Candidate links for rerouting flow are identified and maintained automatically. This means that flow may be rerouted onto a link that goes to a different adjacent switch, as long as that link is on the least-cost / shortest path to the destination domain ID.

The Open Trunking feature is installed and enabled on each switch / director. The algorithm controls the exits the frames will take.

To fully balance an ISL, Open Trunking must be enabled on both switches connected to the ISL.

Using Open Trunking in interop mode

If a switch with Open Trunking has an ISL to a Connectrix ED-1032, Brocade, or Cisco switch, the Connectrix M series switch uses the Open Trunking algorithms to balance the transmit flows to the interop switch.

Potential impact on environment

Significant improvements to the Open Trunking algorithm have been made with Connectrix E / OS 6.01 to lower the likelihood of a reroute causing frames to arrive out of order at the N_Ports. Out-of-Order frame handling by N_Ports typically requires a retry of the FCP Exchange. In some cases this retry may depend on upper-level protocol timers in the host device driver, which may be as long as 60 seconds. On the other hand, overly congested links can lead to discarded frames. Error recovery for discarded frames often depends on the upper-level protocol timers. Open Trunking helps minimize the likelihood of congested links.

Out-of-order and discarded frames caused by Open Trunking reroutes can cause the following:

- Occasional BF2D, xx1F, AB3E errors could occur on Symmetrix Fibre Channel adapters.
- Some Open Systems hosts might log temporary disk access types of events. (For example, Solaris hosts might log SCSI transport errors in the file **/var/adm/messages**.)
- Windows hosts attached to CLARiiON arrays might see Event 11s.
- In FICON environments, an IFCC error can result from an out-of-order frame.

The storage environment can be monitored for a change in the rate of these errors and events to determine whether Open Trunking improves system throughput.

Supported topologies

Open Trunking is fully supported in all fabric topologies currently supported by EMC. However, it is strongly recommended that the number of hops between server and storage pairs be kept to a minimum (ideally one hop maximum). The total hop count is the number of ISLs that a frame will have to cross when routed between the server and its storage. Not all switches in the fabric need to have Open Trunking enabled.

Example — Single fabric

In Figure 416, no limitations are placed on which host can access which storage array. For all possible combinations, Open Trunking is fully supported.

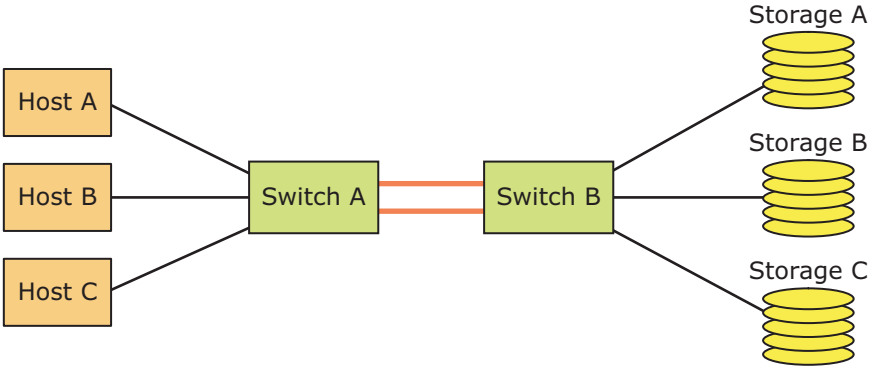


Figure 416 **Open Trunking on single fabric**

Example — Balanced fabrics

In Figure 417, no limitations are placed on which host can access which storage array. For all possible combinations, Open Trunking is fully supported.

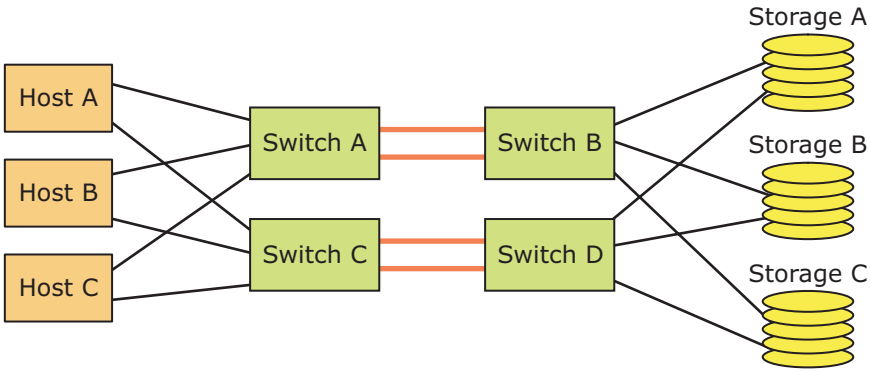


Figure 417 **Open Trunking on balanced fabrics**

Example — Four-corner cross-connect (full mesh)

In Figure 418, no limitations are placed on which host can access which storage array. For all possible combinations, Open Trunking is fully supported.

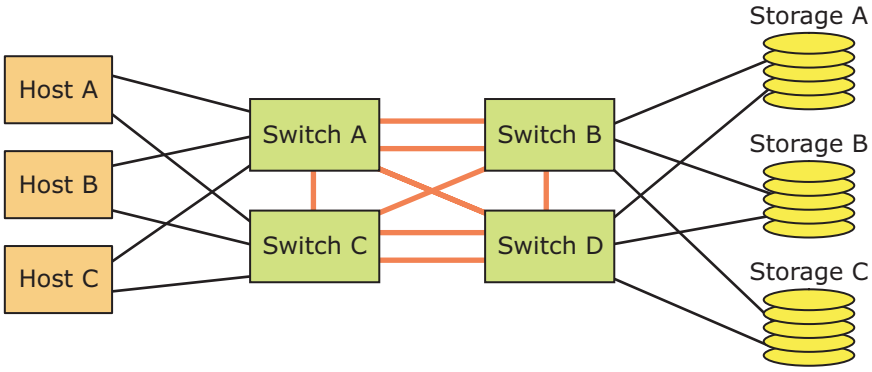


Figure 418 **Open Trunking on full mesh**

Example — Three-hop core/edge; one hop logical

In Figure 419, although there are three hops in the fabric, all hosts logged in to the core switches (A, B, C or D) can access each storage array also logged in to the core switches bwithout crossing more than one hop. Therefore, the hosts are considered to be one logical hop away from each other, and can access each other without restriction.

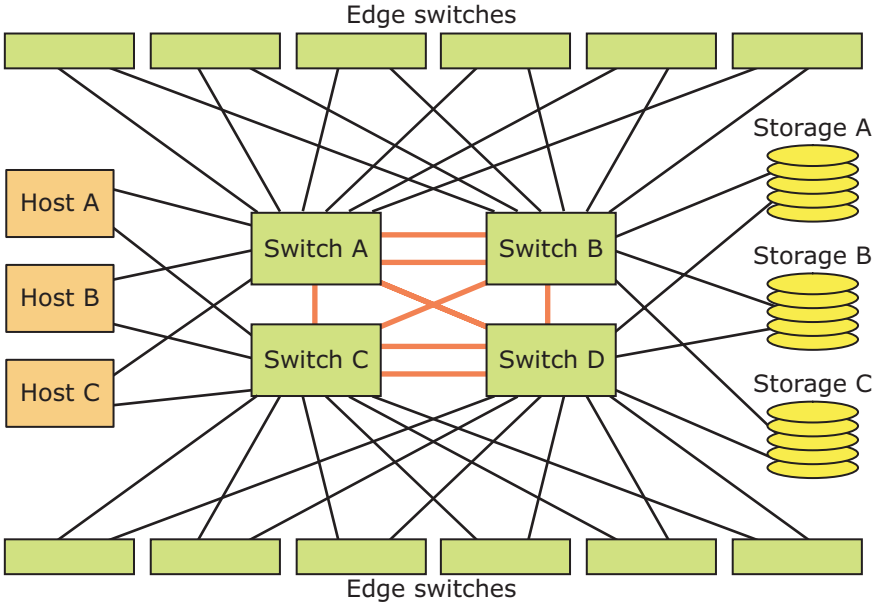


Figure 419 **Open Trunking on three-hop core/edge; one hop logical (Example 1)**

In Figure 420, although Host D is logged into an edge switch, it is still only one hop away from Storage D, even though there are three hops in the environment. Host D and Storage D are still considered to be one logical hop away from each other, and can access each other without restriction.

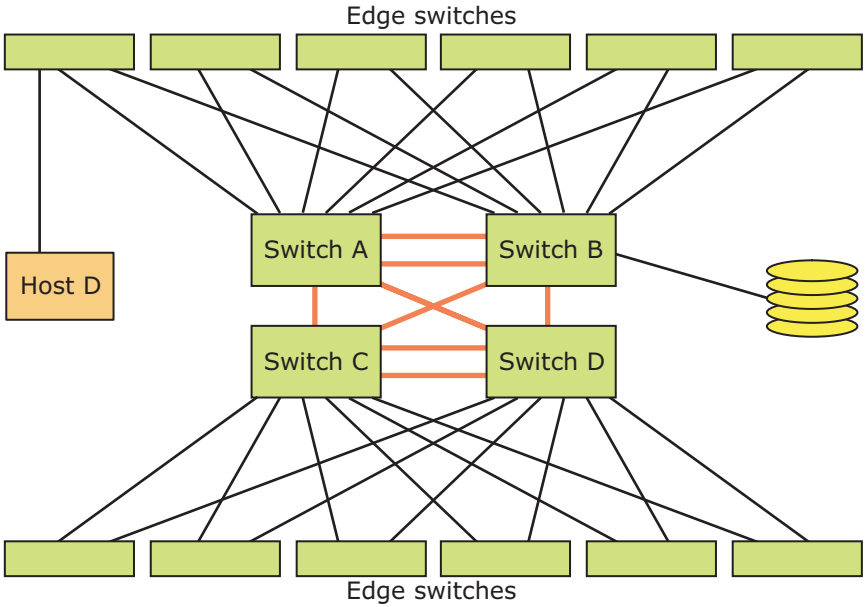


Figure 420 Open Trunking on three-hop core/edge; one hop logical (Example 2)

Example — Three-hop core/edge; three-hop logical

In Figure 421, not only are there three hops in the fabric, but Host E is three hops away from Storage E. This configuration is not recommended, because as the number of hops increases, so does the chance of a frame being delivered out of order when an Open Trunking reroute occurs. Again, it is important to note that if Open Trunking did not cause a reroute, frames might be discarded anyway, due to ISL congestion.

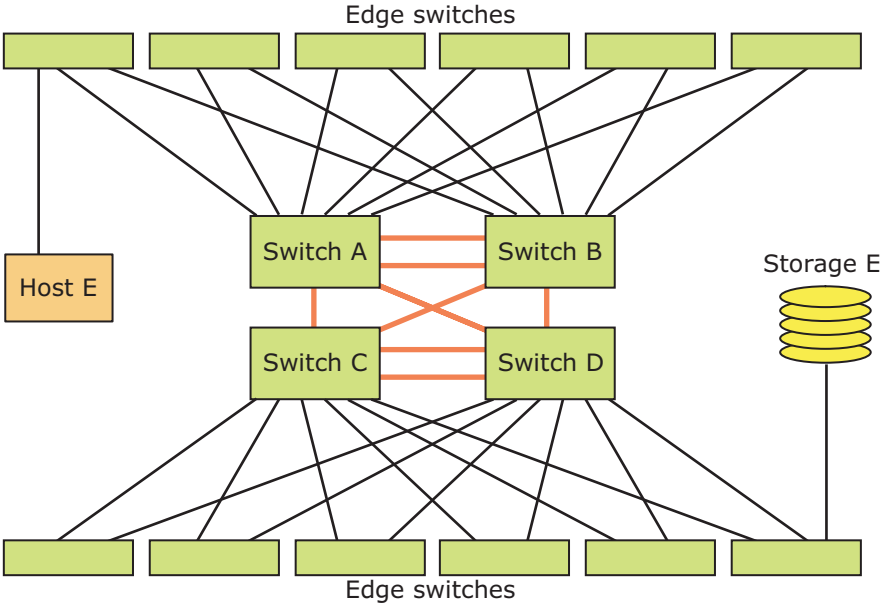


Figure 421 **Open Trunking on three-hop core/edge; three-hop logical**

Connectrix EC-1700 cabinet system

The EMC Connectrix EC-1700 cabinet (Figure 422) provides a central installation platform for mounting Connectrix ED-48000B and MDS Fibre Channel directors.

Figure 422 EC-1700 cabinet front view

EC-1700-B description

The EC-1700-B, is designed to support the ED-48000B director.

Features

Features of the EC-1700-B cabinet include:

- Mechanical, power, and cooling capacity for up to 1100 ports of director-class Fibre Channel switching.
- The EC-1700-B is designed to optimize air-flow and cooling for the ED-48000B. Only two ED-48000B directors can be installed into the EC-1700-B, or one ED-48000B and one previously released ED-24000B or ED-12000B.
- Specially perforated front and rear doors to optimize air flow in and out of the cabinet.
- Fiber cable entry access through floor or ceiling via perforated top panels that are removeable.
- Open floor pan for airflow and cable egress.

EC-1700-B base configuration

The EC-1700-B comes from the factory with cable management fingers to support two fully populated ED-48000B directors, 1000+ fiber cables.

Other base configuration equipment includes four 30amp single phase PDU's for connection of the ED-48000B power cords Table 64 describes the EC-1700-B cabinet.

Table 64 EC-1700-B cabinet models

Model	Description	Products
EC-1700-B	Enterprise cabinet for Connectrix B-Series directors	Directors: ED-48000B

EC-1700-C description

The current release of the EC-1700, the EC-1700-C, is designed to support the newly released MDS 9513 director and the MDS 9509, as described in Table 65.

Table 65 EC-1700-C cabinet models

Model	Description	Products
EC-1700-C	Enterprise cabinet for Connectrix MDS series directors	Directors: MDS 9509, MDS 9513

Features

Features of the EC-1700 cabinet include:

- Mechanical, power, and cooling capacity for up to 1100 ports of director-class Fibre Channel switching.
- The EC-1700-C is designed to optimize air-flow and cooling for the MDS-9513, therefore only two MDS directors can be installed into the EC-1700-C. Either two 9513's, or two 9509s, or one 9513 and one 9509.
- Fiber cable entry access through floor or ceiling.

EC-1700-C base configuration

The EC-1700-C comes from the factory with cable management fingers and trays to support two fully populated MDS-9513 directors, 1000+ fiber cables. Other base configuration equipment includes:

- Specially cut-out vertical NEMA mounting rails to provide air passages to the sides of the MDS directors.
- Pre-loaded directional air-baffles to direct hot exhaust air out the rear and top of cabinet.
- Specially perforated front and rear doors to optimize air flow in and out of the cabinet.
- Pre-loaded L-Brackets for mounting the 9513 or 9509 directors into the EC-1700-C.

Each EC-1700-C cabinet can hold up to:

- Two MDS 9509 directors
- Two MDS 9513 directors
- One 9513 and one 9509

Note: Only two MDS 9500 series directors per cabinet regardless of type and/or configuration.

Guidelines for installing EC-1700 components

This section contains placement guidelines for Connectrix B Series and MDS Series directors. You can find examples of recommended B Series and MDS Series component placement in the *EMC Networked Storage Topology Guide*, available at http://Powerlink.EMC.com, under **Support**, **Document Library**, **Solutions Guides**.

B Series examples

Figure 423 shows some examples of ED-48000B placement.

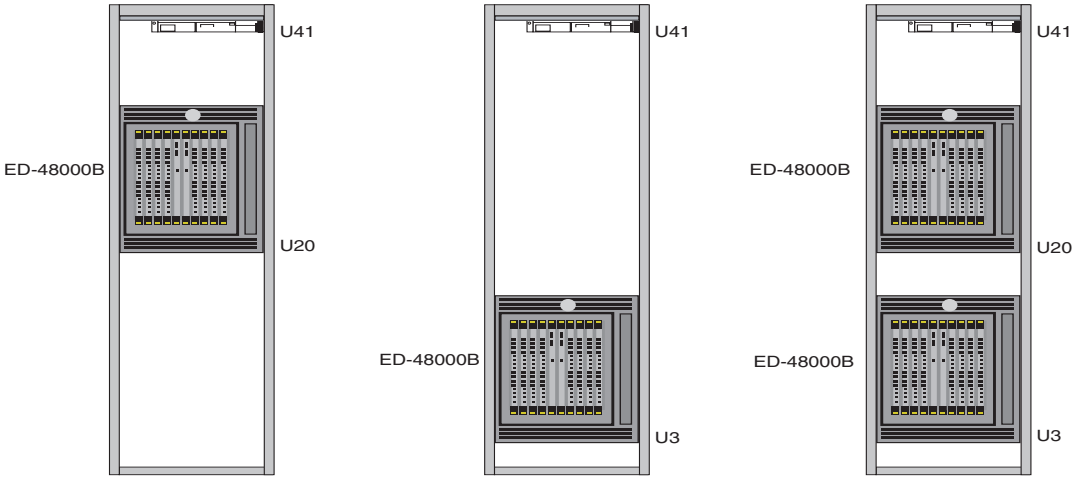


Figure 423 EC-1700-B cabinet configuration examples

MDS Series examples

CAUTION

The Figure 424 configurations on page 1157 show the maximum allowable in order to maintain proper cooling of the MDS-9513 units. Do not exceed the maximum number of units per cabinet or install other equipment in the cabinet with the 9513s or 9509s.

Figure 424 on page 1157 shows some examples of placement of MDS 9509 and MDS 9513 directors.

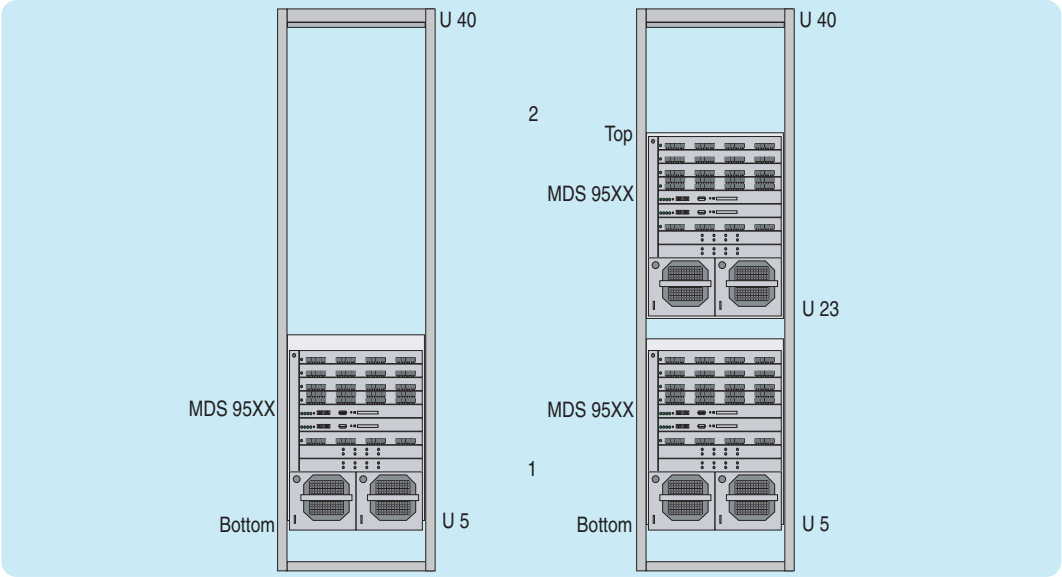


Figure 424 **EC-1700-C cabinet configuration examples 1 – 2**

Note: Only a maximum of two MDS 9513 or 9509 Series directors may be installed into the EC-1700 cabinet, regardless if they are MDS 9509s or MDS 9513s.
Any graphical representation displayed herein of a MDS 9509 could also represent a MDS 9513 and vice versa.

Connectrix EC-1500 cabinet system

The Connectrix EC-1500 cabinet provides a central installation platform for mounting a variety of Connectrix Fibre Channel directors and switches for centralized management and monitoring by the Connectrix Manager.

EC-1500 features include:

- Mechanical, power, and cooling capacity for up to 512 ports of director-class Fibre Channel switching.
- Mechanical, power, and cooling capacity for up to 384 ports of departmental-class Fibre Channel switching.
- Mixing of multiple supported Connectrix switch and director models in the same cabinet. Models can be mixed within Connectrix M series, B, series, or MDS series.
- Integrated rackmount service processor, 10/100 Ethernet, and backup systems for management, maintenance, and serviceability.
- Dual AC power strips and inputs.
- Fiber cable entry access through floor or ceiling.

Note: Connectrix DS-xxM (DS-16M, DS-16M2, DS-32M, DS-32M2) Departmental Switches can also be mounted in the CLARiiON® product cabinet or a standard EIA cabinet. For additional information, refer to the *Connectrix DS-16M2/DS-32M2 Installation and Service Manual* (P/N 014003065).

Guidelines for installing directors and switches

The EC-1500 contains 40 *units* (40 u) of available rack-mount height. (1 u = 1.75 inches.) Director and switch heights are:

Connectrix series	Connectrix model	Height
M Series	ED-1032	18 u
	ED-10000M	14 u
	ED-140M	12 u
	ED-64M	9 u
	DS-32M, DS-32M2	1.5 u
	DS-24M2, DS-16M, DS-16M2	1 u
B Series	ED-12000B	14 u
	DS-32B2, DS-16B, DS-16B2, DS-8B	1 u
MDS Series	MDS 9509	14 u
	MDS 9506	7+2 u [a]
	MDS 9216	5+2 u [a]
	MDS-9140, MDS-9120	1 u

a.Each MDS 9216 and MDS 9506 requires a 2 u mounting bracket assembly.

With this in mind, here are some guidelines for installation:

- Install the heaviest units from the bottom up.
- 1 u and 1.5 u switches should be installed to allow 3 u of rack height per switch, which is necessary to facilitate organizing and managing the fiber-optic cables in the cabinet and to maintain proper airflow and ventilation.
- Unless the cabinet will will be loaded with the maximum configuration of departmental switches, leave adequate space at the bottom of the cabinet for possible future installation of one or more directors.

Cabinet configuration examples

Figure 425 through Figure 432 on page 1165 show some examples of cabinet configurations; numbers in circles represent the recommended order of installation.

Connectrix M series examples

Note: DS-xxMs can also be mounted in the CLARiiON product cabinet or a standard EIA cabinet. The *Connectrix DS-16M and DS-32M Installation and Service Manual* (P / N 014003065) contains more information.

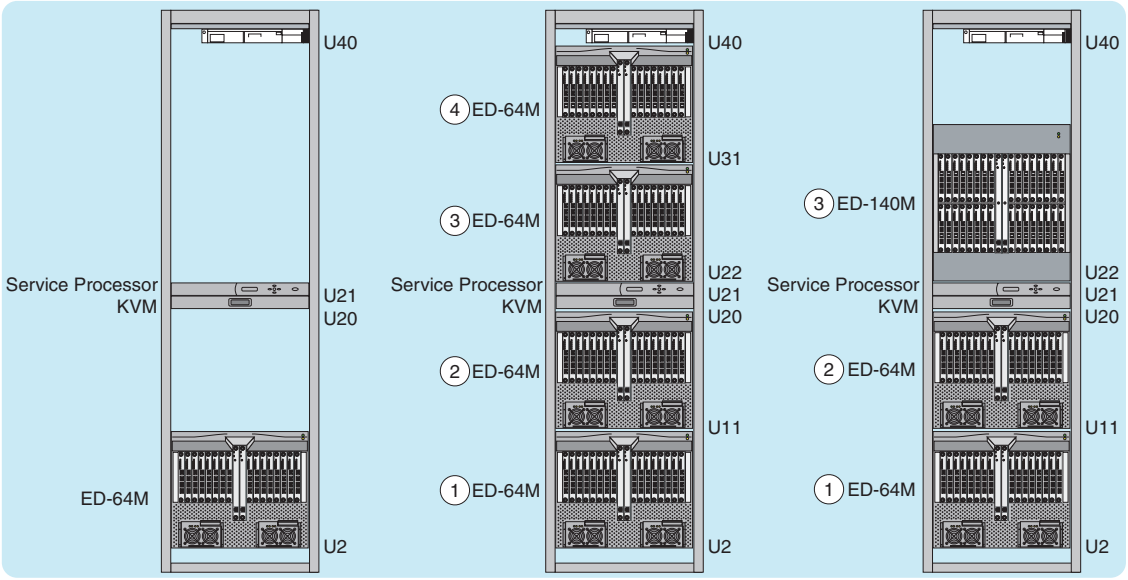


Figure 425 EC-1500 cabinet configuration: M series examples 1 – 3

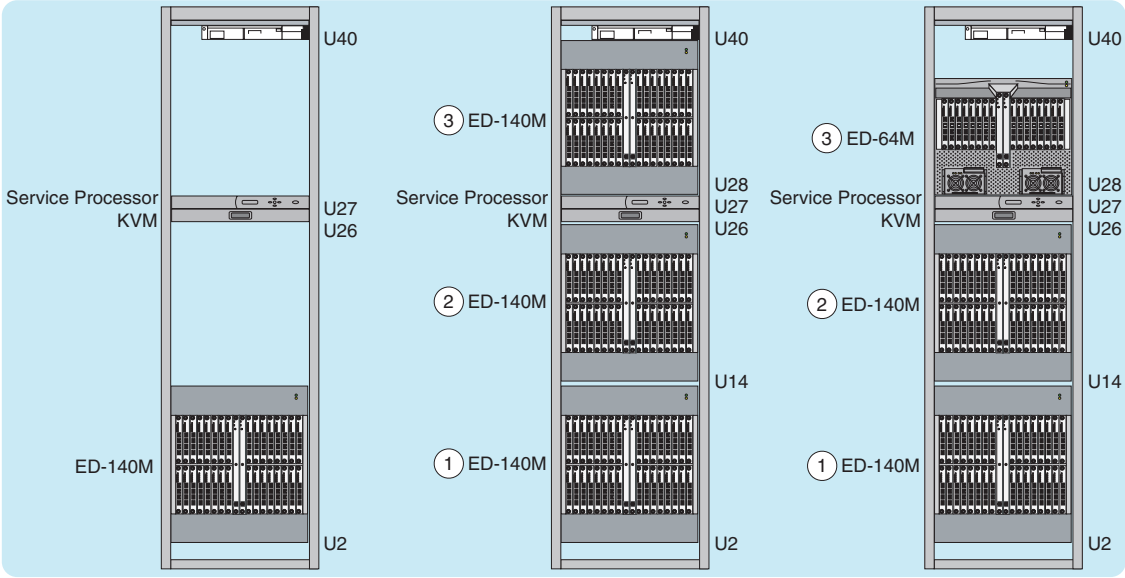


Figure 426 **EC-1500 cabinet configuration: M series examples 4 – 6**

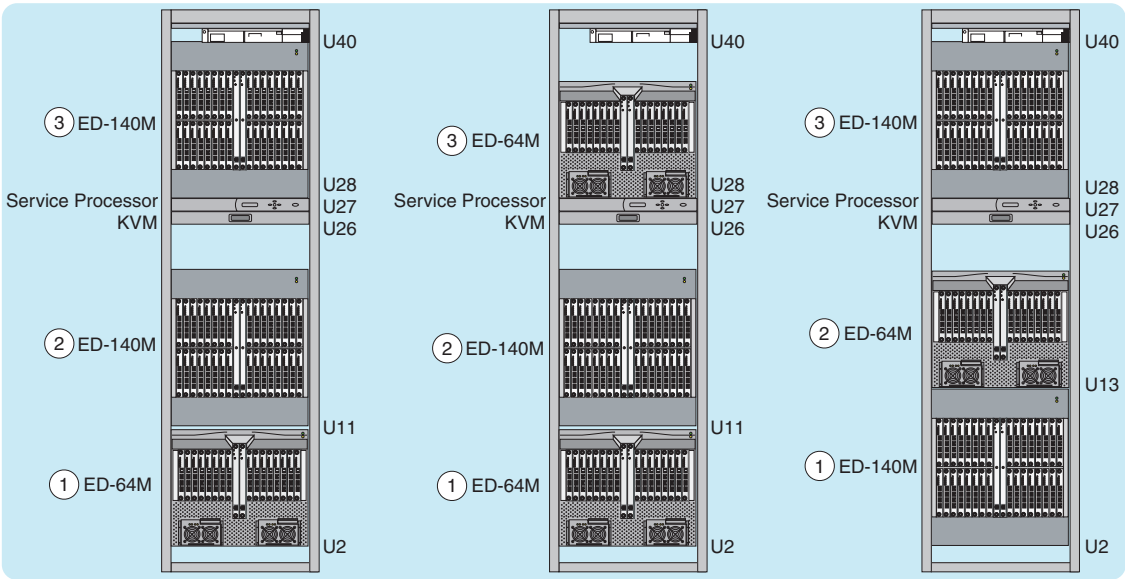


Figure 427 **EC-1500 cabinet configuration: M series examples 7 – 9**

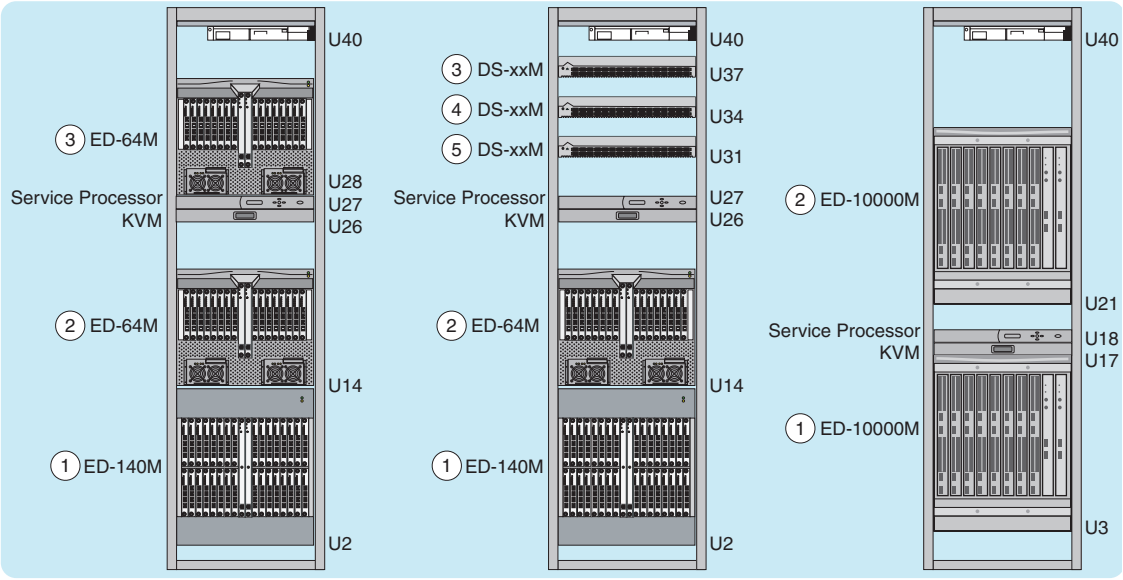


Figure 428 EC-1500 cabinet configuration: M series examples 10 – 12

Connectrix B series examples

Note: The are no rail kits available to support the mounting of DS-xxB series departmental switches into the EC-1500 cabinet.

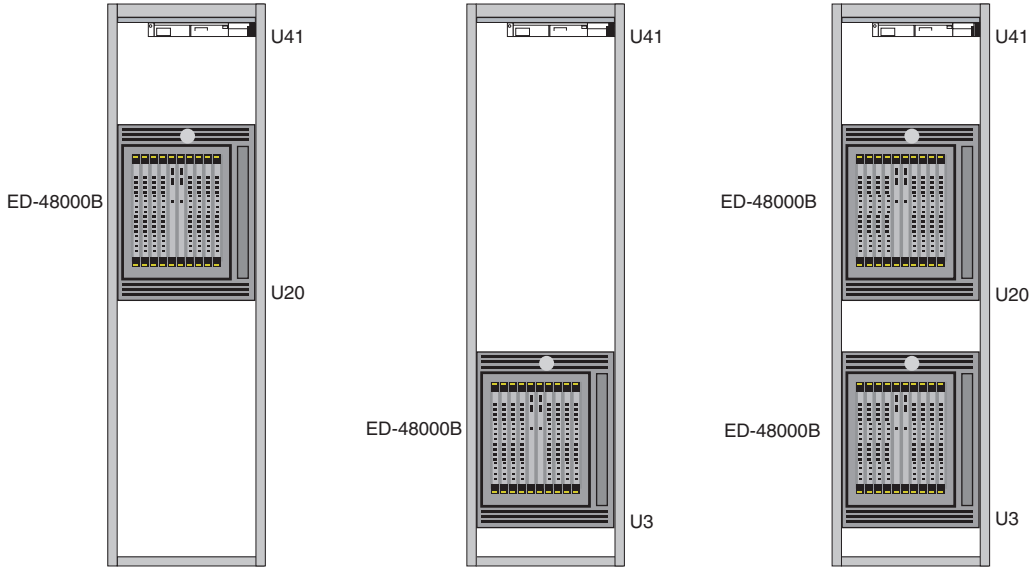


Figure 429 **EC-1500 cabinet configuration: B series examples**

Note: The example on the left in Figure 429 shows the location when the EC-1500 is shipped with one ED-12000B installed. This facilitates field-installation of a second ED-12000B.

Connectrix MDS series examples

All MDS switches are field-installed. As Figure 430 on page 1164 through Figure 432 on page 1165 show, the bottom switch is installed with a 2 u gap below it to facilitate installation, removal, and cable management.

Also note that each MDS 9216 and MDS 9506 requires a 2 u mounting bracket assembly beneath it.

Note: These figures do not show placement of the MDS-9120 or MDS-9140. Guidelines for placing these switches are similar to those for M series departmental switches.

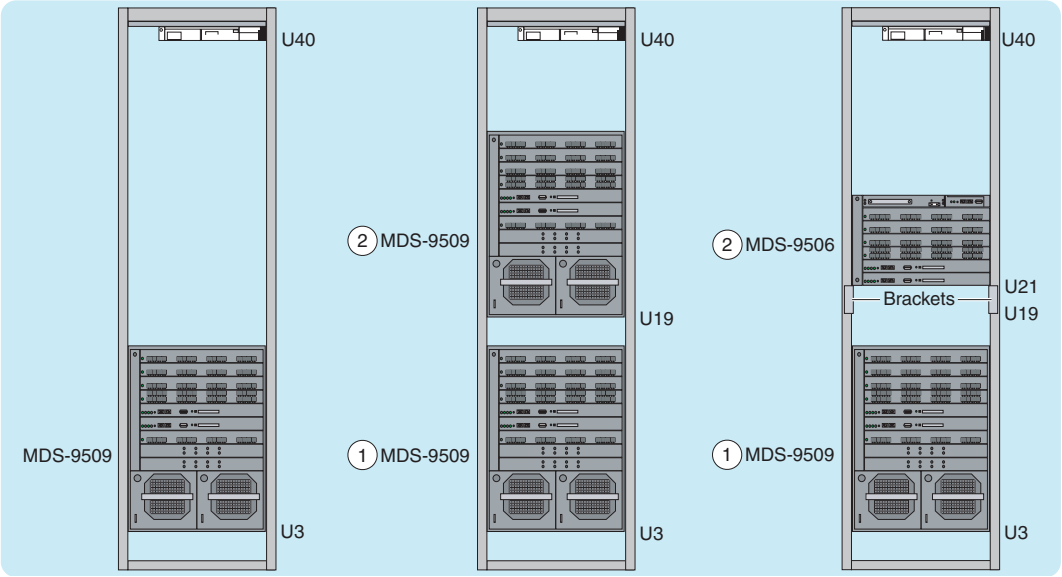


Figure 430 EC-1500 cabinet configuration: MDS series examples 1 – 3

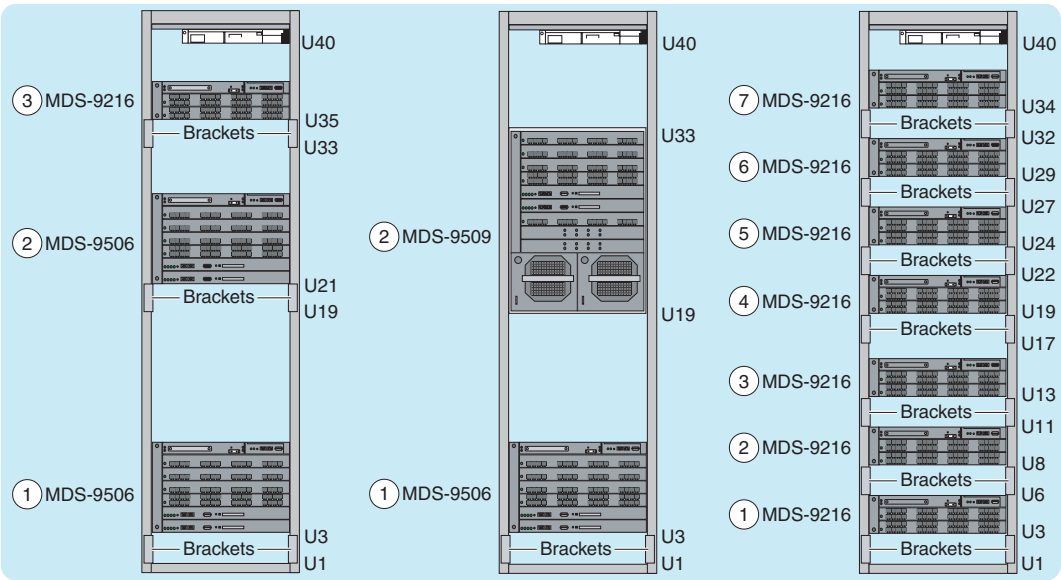


Figure 431 EC-1500 cabinet configuration: MDS series examples 4 – 6

Note: The extra 1 u space shown between U13 and U19 in the right-side configuration in Figure 431 is simply to align the top four switches similar to other configurations. This facilitates upgrading the cabinet to a configuration that replaces some of the smaller switches with a larger switch.

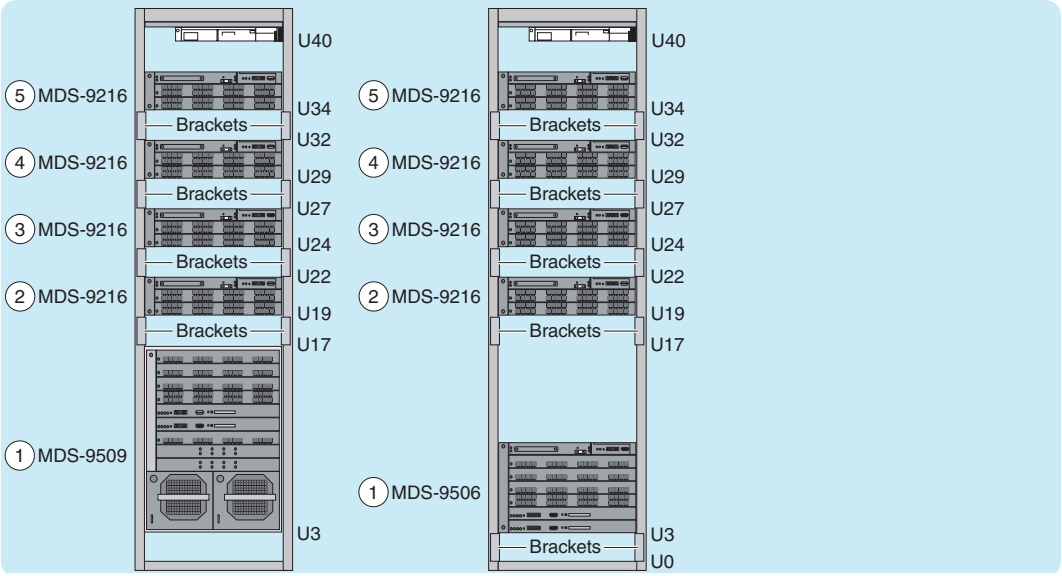


Figure 432 **EC-1500 cabinet configuration: MDS series examples 7 – 8**

Connectrix EC-1200 cabinet system

The Connectrix EC-1200 delivers a third-generation Fibre Channel Connectivity system to the storage application environment.

EC-1200 features include:

- Mechanical, power, and cooling capacity for 420 ports of director-class Fibre Channel switching
- Mechanical, power, and cooling capacity for 384 ports of departmental-class Fibre Channel switching
- Mixing Connectrix ED-xxM Directors, ED-1032 Directors, and DS-xxM Departmental Switches in the same cabinet
- Integrated service processor, 10 / 100 Ethernet, and backup systems for management, maintenance, and serviceability
- Dual AC power strips and inputs
- Fiber cable entry access through floor or ceiling

Note: Connectrix DS-xxM (DS-16M, DS-16M2, DS-32M, DS-32M2) Departmental Switches can also be mounted in the CLARiiON product cabinet or a standard EIA cabinet. The *Connectrix DS-16M2/DS-32M2 Installation and Service Manual* (P / N 014003065) contains more information.

Guidelines for installing directors and switches

The EC-1200 contains 36 *units* (36 u) of available rack-mount height. Director and switch heights are:

Model	Height
ED-1032	18 u
ED-140M	12 u
ED-64M	9 u
DS-32M, DS-32M2	1.5 u
DS-24M2, DS-16M, DS-16M2	1 u

With this in mind, here are some guidelines for installation:

- Install the heaviest units from the bottom up.
- DS-xxMs should be installed to allow 3 u of rack height per switch. 3u per switch is necessary to facilitate organizing and managing the fiber-optic cables in the cabinet and to maintain proper airflow and ventilation.
- Unless the cabinet will will be loaded with the the maximum configuration of 12 DS-xxMs, leave 9 u (or 18 u if possible) of space at the bottom of the cabinet for possible future installation of one or more ED-64Ms or ED-140Ms.

Cabinet configuration examples

Figure 433 through Figure 438 on page 1170 show some examples of cabinet configurations; the numbers in circles represent the recommended order of installation.

Note: DS-xxMs can also be mounted in the CLARiiON product cabinet or a standard EIA cabinet. The *Connectrix DS-16M and DS-32M Installation and Service Manual* (P / N 014003065) contains more information.

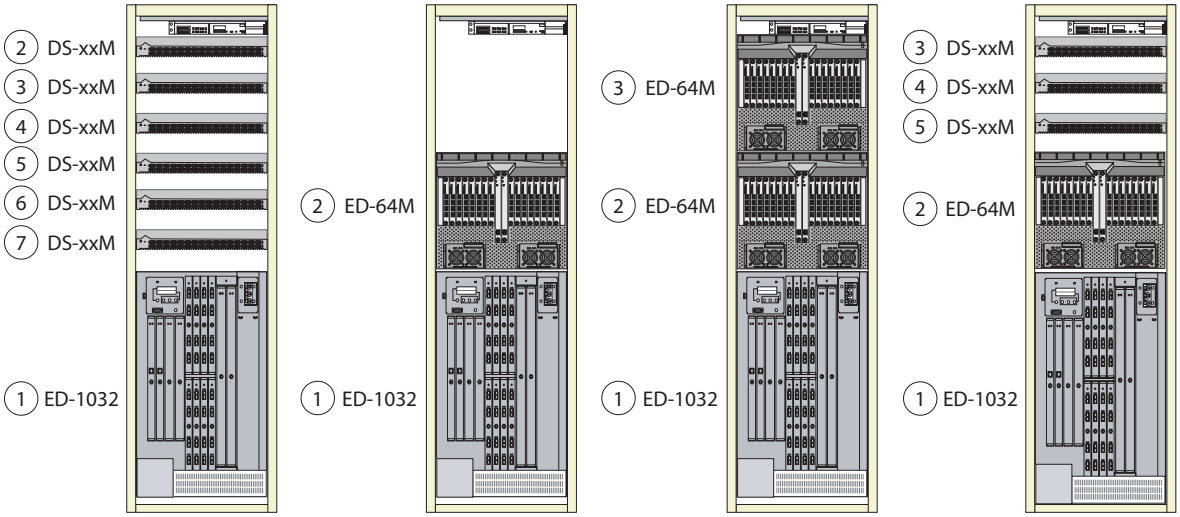


Figure 433 EC-1200 cabinet configuration examples 1 – 4

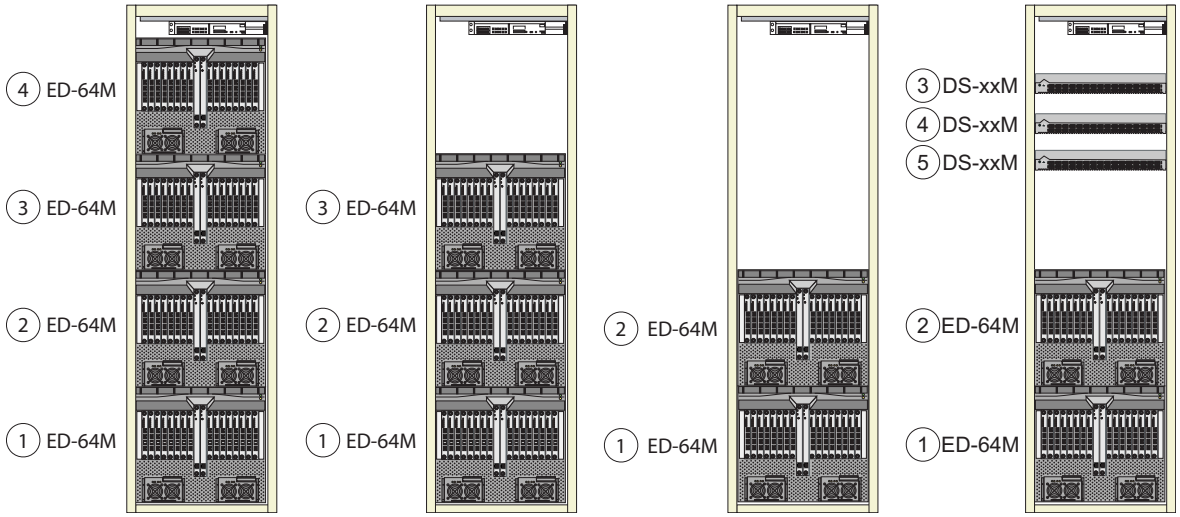


Figure 434 EC-1200 cabinet configuration examples 5 – 8

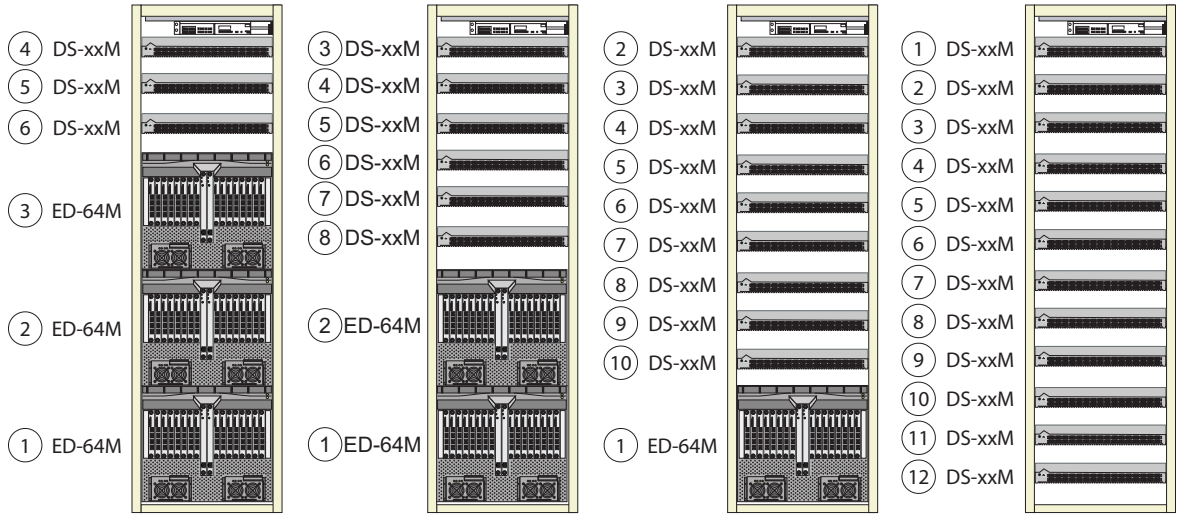


Figure 435 EC-1200 cabinet configuration examples 9 – 12

Note: The order of installation in the two examples on the right in Figure 436 reflects inital installation of fewer than the maximum number of units, allowing space for future installation of additional directors and / or switches in the lower half of the cabinet. If these additonal units are installed, their order would follow that shown in the figure.

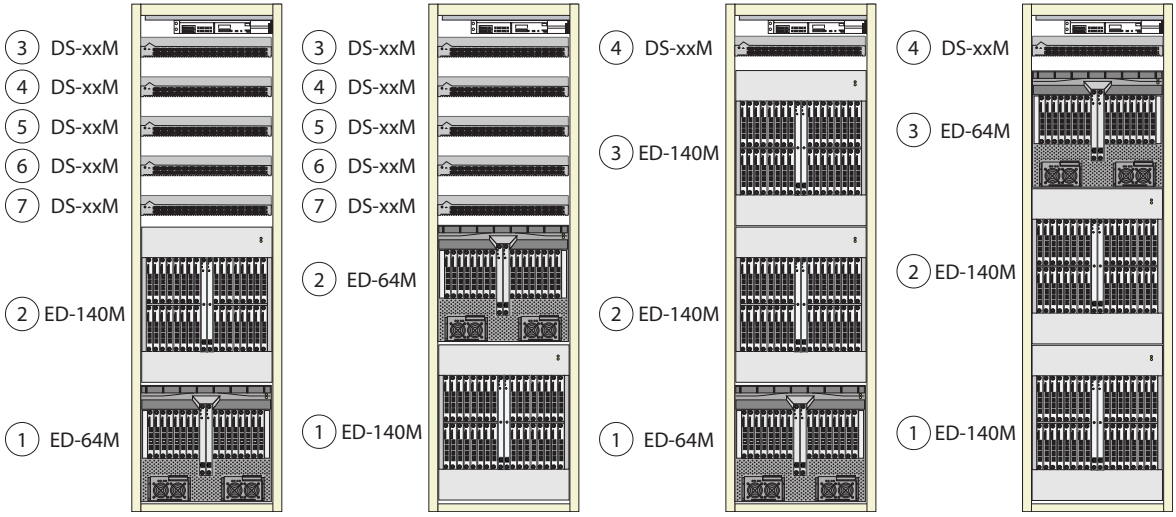


Figure 436 EC-1200 cabinet configuration examples 13 – 16

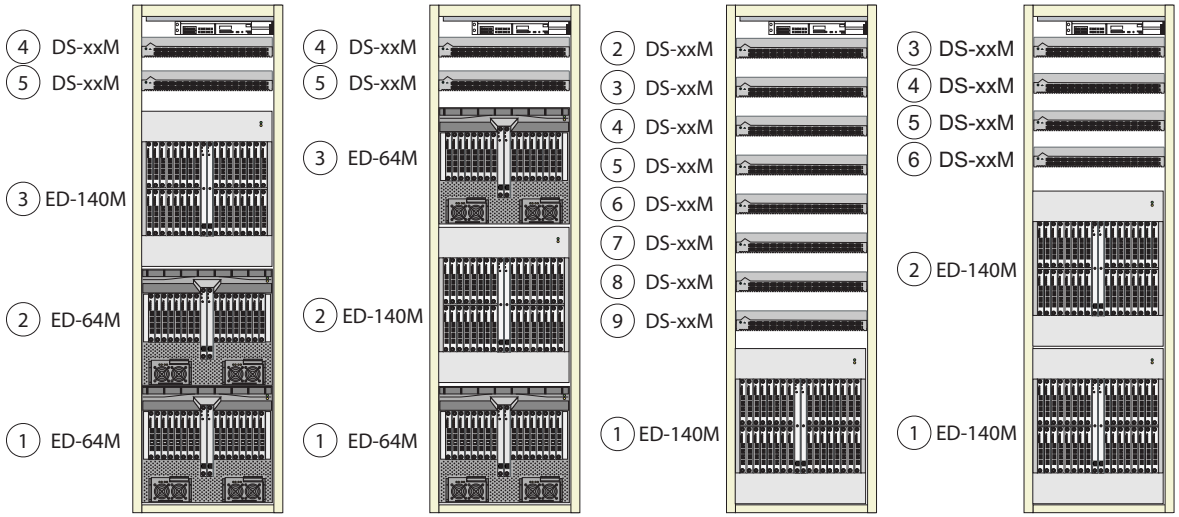


Figure 437 EC-1200 cabinet configuration examples 17 – 20

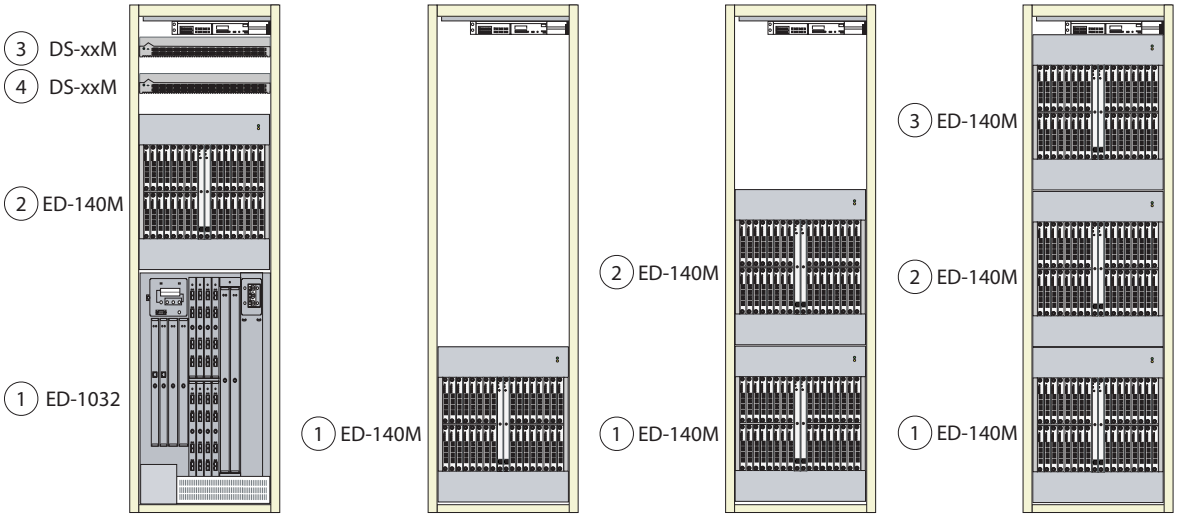


Figure 438 EC-1200 cabinet configuration examples 21 – 24

CAUTION

EMC recommends keeping the Connectrix stabilizer /outrigger brackets (P/N 100-605-014) on the EC-1200 cabinet at all times. When the cabinet is positioned such that it is not immediately adjacent (within two inches) to another equipment cabinet, the stabilizer bracket outriggers must be positioned in their outermost position. If the cabinet is positioned within two inches of another equipment cabinet, the outrigger brackets can be positioned in their innermost position. Any side of the EC-1200 cabinet that is not within two inches of another equipment cabinet should have that side's outriggers positioned in the outermost position. These rules apply not only for initial installation, but also if the EC-1200 switch and/or director configuration is altered at a later date.

C

Non-EMC Products Data

This appendix contains data on non-EMC products.

- CNT/Inrange switches and directors 1174
- Nortel OPTera Metro 1180
- Ciena ONLINE7000 1185
- Ciena CN 4200/CN4200 MC 1195
- Cisco Metro 1500 1210
- Cisco ONS 15540 1214
- Cisco ONS 15454 MSTP 1224
- Finisar FLX-2000 link extender 1261
- SONET overview 1263
- CIENA CN 2000 1266
- Lucent OptiStar EdgeSwitch 1271
- Blade server switch modules 1273
- Blade servers 1296
- Brocade Access Gateway 1314

Note: For information on EMC-qualified third-party products, refer to the *EMC Select* document on Powerlink.

CNT/Inrange switches and directors

This section contains information on CNT/Inrange switches and directors.

CNT/Inrange FC/9000

The CNT/Inrange FC/9000 64 is a Fibre Channel Enterprise Director, scalable from 24 ports to 64 ports. E-Lab Navigator lists the configurations supported by the FC/9000.

The FC/9000 can be used in a high-speed SAN designed to support data-intensive high-availability applications, such as backup and recovery, business continuance, and data and resource sharing.

Note: FICON and distance testing are not complete.

Fabric Management is performed by the IN-VSN Enterprise Manager in a client/server architecture installed on a PC, usually housed in or near the FC/9000 Director cabinet.

E-Lab Navigator lists specific versions of supported firmware, as well as fabric topology constraints associated with the FC9000.

Figure 439 on page 1175 shows the CNT/Inrange FC/9000 Director cabinet.

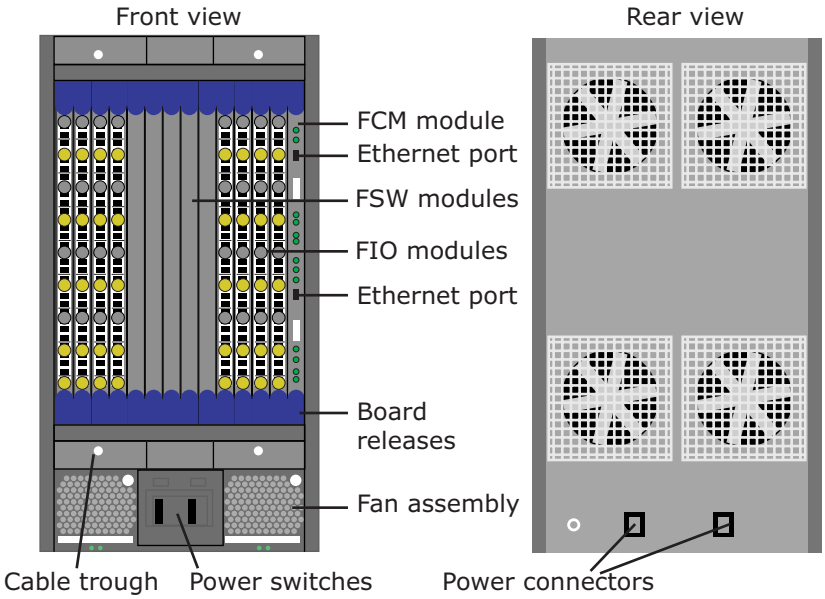


Figure 439 **CNT/Inrange FC/9000**

Component overview

FIO (FC/9000 XCAF or base I/O module)

FIO modules provide the physical connectivity between the FC/9000 backplane and the external devices connected to the FC/9000. The FC/9000 supports a minimum (base) configuration of three base FIO modules and a maximum of eight per chassis.

FIO modules are hot-swappable, containing a single processor, supporting memory, reset button, and front panel indictors to show current status: temperature, heartbeat, logged in/out, and activity (traffic).

The Extended Credit Addressing Facility (XCAF) FIO has two features: the ability to provide 64 buffer-to-buffer credits and FICON addressing.

Blank panels must be inserted where FIO modules are not present.

FSW (FC/9000 switching module)

FSW modules provide the physical and logical connectivity between FIO modules installed in the chassis. These modules provide the middle or cross-connect stage of the switch architecture.

The FC/9000 supports a minimum (base) configuration of four FSW modules and maximum of five (high-availability configuration) per chassis, with only four FSW on line at any time.

FSW modules are hot-swappable, containing a single processor, supporting memory, reset button, and front panel indictors to show current status: temperature, heartbeat, and activity (on line or off line).

Blank panels must be inserted where FSW modules are not present.

FCM (FC/9000 control module)

FCM modules provide the common control interface for the FC/9000 system. This module acts a proxy for all external communication to other modules from the IN-VSN Enterprise Manager. The FC/9000 requires a minimum of one FCM, or can have a maximum two for redundancy.

Each hot-swappable FCM module contains a single processor, supporting memory, reset button, and front panel indictors to show status: link, data, 100 Mb/s, online/offline, primary board, fault, 4-character display window.

Blank panels must be inserted where FCM modules are not present.

Power supply assembly

Power supplies are located in the front left and right sides of the chassis. The power supplies are hot swappable, redundant, and load sharing. The on/off switches are located in the front lower section of the switch.

Fan module assembly

A fan assembly in the rear of the chassis provides cooling for the FC/9000, drawing air through and pushing air out. In the event of a single fan failure, the remaining three fans will accelerate to maintain the cooling process. These fans are replaceable either within a unit or the entire unit (pair).

Backplane module

The backplane provides connectivity among all system modules, including the FIO, FSW, FCM, power supply, and fans. The backplane can be expanded with a special wiring harness and dipswitch configuration to provide connectivity for 128 ports.

The backplane provides:

- Port-to-port bus connection
- Interprocess communication at 100 Mb/s with redundancy
- Connection to expansion interface modules
- Power distribution bus connectivity
- Power supply alarm signals
- *DC fail* status
- *AC fail* status
- **Power supply present** status
- Fan status and control
- Slot geographical addressing
- Miscellaneous status and control

IN-VSN Enterprise Manager

The IN-VSN Enterprise Manager is the GUI used to manage the FC/9000. It is a software application that has two components based on a client/server architecture. The client periodically (every five seconds) polls the server through Ethernet (10/100 Mb/s) to send and retrieve changes to the FC/9000.

The IN-VSN Enterprise Manager Client can:

- Define module and port configurations
- Define zoning
- Monitor alarms
- Monitor system performance
- Invoke system diagnostics
- Implement some director parameters

The IN-VSN Enterprise Manager's client software application operates on Windows NT and Windows 2000 Professional platforms. The IN-VSN Enterprise Manager server software application requires a dedicated PC, operates on Windows NT and Windows 2000 Professional platforms, and has basic hardware compatibility requirements from CNT/Inrange. The server communicates with the FCM module through Ethernet (10/100 Mb/s) to send and retrieve changes to/from the FC/9000.

Features overview

Features of the CNT/Inrange FC/9000 include:

- High availability: fully redundant internal pathing, power, cooling and control; no single point of failure
- Non-disruptive code loads and hot-swappable GBICs
- Auto-discovering, self-configuring 1.0625 Gb ports; arbitrated loop (FC-AL), transitive loop (TL), switched fabric (FC-SW)
- Enterprise Manager SAN fabric management system IN-VSN Enterprise Manager
- Employment of orphan zoning by zoning, hard zoning, port zoning, and broadcast zoning, to guard against losing ports not proactively assigned to a defined zone
- Phone-home and pager direct-dial feature
- Statistical and diagnostic monitoring
- Class 2,3 Fibre Channel environments
- 64 ports available through 8-port I/O modules (8 ports per FIO module)
- GBIC Port Module (SE form factor) available in multimode fiber (shortwave)
- 64 buffer-to-buffer credits (BB_Credits) available per port
- Support for Class 2 and Class 3 Fibre Channel protocols
- Auto-negotiate function on all ports, to provide either switched F_Port or T_Port connections
- Full duplex 100 MB/s data rate per Fibre Channel port
- Supported port types: F_Port, FL_Port, TL_Port

Director Management (IN-VSN Enterprise Manager)

The IN-VSN Enterprise Manager:

- Provides centralized monitoring and control of multiple fabrics and all vital network functions from a single console. Using a Java-based and/or SNMP interface, multiple concurrent users can access levels of fabric information ranging from basic monitoring and configuration information to detailed performance data. Enterprise Manager is a configuration tool, as well as an application for management of SAN configuration, application and performance.
- Allows centralized configuration and management of fabric using client/server architecture.

- PC Management Server allows Server functionality and Client.
- Supports Windows NT 4.0 and Windows 2000 Professional Clients.
- Allows centralized management of Director.
- Provides support for online, nondisruptive code upgrades.
- Features 10/100 Mb Ethernet connections to FCM for out-of-band management.
- Provides extensive centralized logging: Event, Audit, Session logs, and SNMP support.

Availability management

Availability management includes:

- Management system helps you track the status of redundant power, cooling, and control.
- Phone home/email home provides instant notification of system or network issues.
- Front panel display allows quick check of fan operation, temperature, and port status.

Performance management

Performance management includes:

- Dynamic statistics display performance data for each online port.
- Zoning of FC/9000 ports allows efficient and secure communication among nodes.
- Event Log and Audit Log streamline the troubleshooting process and provide rapid error source identification.
- SNMP traps show whether defined limits have been exceeded.

References

Note the following for more information:

- http://www.cnt.com
- *IN-VSM FC9000 Fibre Channel Director Installation Manual*
- *IN-VSM FC9000 Fibre Channel Director Maintenance Manual*
- *IN-VSM FC9000 Fibre Channel Director Installation and Operation Manual*
- *IN-VSM FC9000 Fibre Channel Director Site Planning Guide Manual*

Nortel OPTera Metro

The Nortel Optera Metro platform is a true protocol- and bit-rate-independent fiber-optic transport system. The Optera Metro supports the following protocols: SONET, ATM, Gigabit Ethernet, IP, FDDI, and all optical interfaces (OC-n).

Note these capacities:

- An ETSI- and NEBS-compliant unit, the Optera Metro shelf can accommodate up to 10 Gb/s of capacity in less than 2 cubic feet of space.
- An Optera Metro system can have one to eight pairs of shelves at multiple sites configured in a hubbed ring or point-to-point topology . Each site can have one or more shelves.
- A fully loaded system (16 shelves) can transport up to 32 protected or 64 unprotected channels (wavelengths) over each optical fiber.
- Each channel can operate from 50 Mb/s to 2.5 Gb/s. This allows a total transport capacity of 80 Gb/s.

A *shelf* is a basic building block of a Nortel OPTera DWDM system. A shelf contains a subsystem of components that convert optical signals into electrical, allow *adding* and *dropping* functionality, and multiplex and pass signals through the network.

A Nortel OPTera shelf holds:

- Optical Channel Interface (OCI); provides signal interface card.
- Optical Channel Laser and Detector (OLCD)
- Optical Channel Manager (OCM)
- Optical Multiplexer (OMX); provides add/drop filtering (ADF) to multiplex each OCLD optical wavelength signal onto the single mode fiber.

- Shelf Processor (SP); provides monitoring and control functionality.

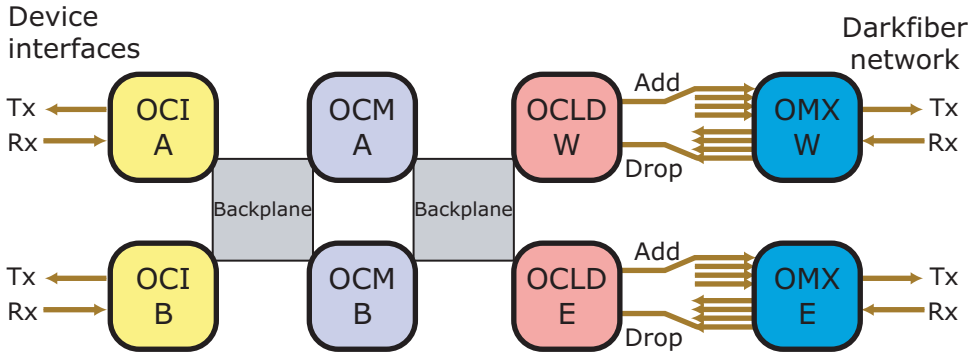


Figure 440 Nortel OPTera shelf diagram

The OCI interface provides the necessary connections to connect the Optera to customer traffic. There are two types of OCI card(s): 1.25 Gb/s and 622 Mb/s. The Optical Channel and Detector (OCLD) receives electrical client signal from the back plane, converts electrical signal to DWDM wavelength, and provides 32-channel fault monitoring and two Fibre Channel optical connectors (connected to Fibre Channel pigtails from OMX modules).

The optical signals from the DWDM network are converted into electrical signals only on the shelf that drops them, creating a logical point-to-point topology between two shelves that carry the same optical wavelength band. (At least two shelves with the same wavelength band in two locations of the network are necessary.) All other bands pass through the shelf's optical filters.

Available Optera Metro topologies

The point-to-point configuration (which is the basis for all other configurations) will include a local and remote site. The data will flow between the different sites using two links. Each link includes a transmit and receive single-mode fiber cable (dark fiber). The two links are usually described as *east-to-west* or *west-to-east*. The cabinets usually contain four shelves or bands, which make up the site.

Figure 441 and Figure 442 show possible Optera Metro DWDM topologies.

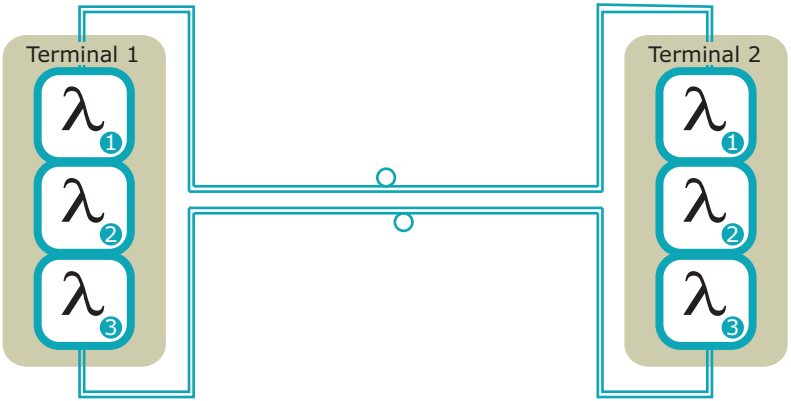


Figure 441 Point-to-point protected topology

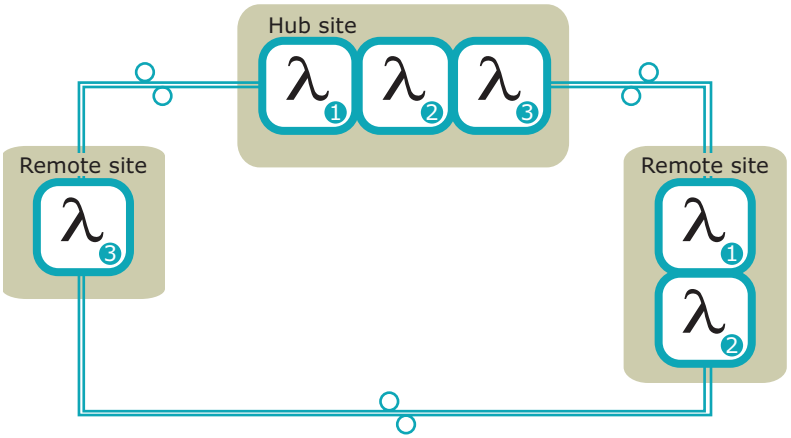


Figure 442 Hubbed ring topology

Nortel OPTera protection scheme

A protected channel connects an attaching device interface by using a single OCI card and two OCLD cards (each having the same wavelength) in one shelf of the shelf pair to two corresponding OCLD cards (each having the same wavelength) and a single OCI card in the second shelf of the shelf pair. The data flow between the

OCI card and two OCLD cards within each shelf is managed by the two Optical Channel Manager (OCM) cards in each shelf.

This scheme creates two data paths inside the DWDM network. One path is active as long as signal integrity is maintained by the physical connections. Any disruption will fail over to the alternate data path using the second OCLD pair.

Note that this configuration neutralizes and single OCI, one on each shelf. A mixture of protected and unprotected channels is available in a single shelf.

Power budget calculations

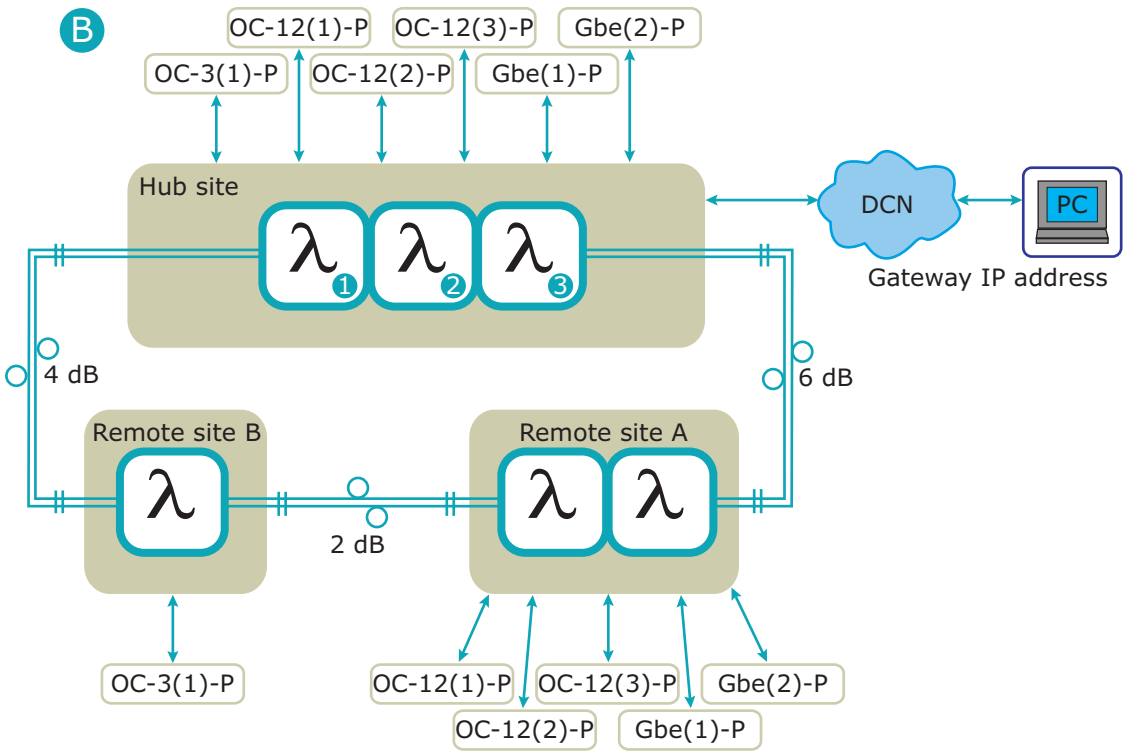


Figure 443 **OPTera protection scheme**

Calculate the power budget as shown in the following example. Note that this is only an approximation, and that a site survey is required before activation:

1. Fiber losses are calculated according to the distances times the specifications of the fiber cable. In this case 0.2 dB per km, so for the 30 km leg (30 km * 0.4 dB / km = 6 dB loss).
2. Each connector has approximately 0.5 dB loss per connector.
3. Add up all of the fiber losses, on all three legs, with the connector losses:

   6 dB + 2 dB + 4 dB + (0.5 dB * 6) = 15 dB
4. Add repair margin (10%): 15dB + 1.5 dB = 16.5 dB.
5. With a Maximum Link Budget (see Nortel OPTera end-of-life chart): 18.3 dB. Subtract total link budget (16.5 dB) from Maximum Link Budget (18.3 dB) and if result is Positive (1.8 dB), then link budget is within parameters.

Diagnostics and maintenance

For configuration, power budget calculation and troubleshooting details, consult the Nortel Optera Metro technical publications.

Using the Symmetrix Fibre Channel director online utilities can complement repair and system diagnostics.

Ciena ONLINE7000

The Ciena ONLINE*7000* platform is a true protocol and bit-rate-independent fiber-optic transport system that supports the following protocols:

- SONET/SDH
- IP
- Ethernet
- GbE
- Fibre Channel
- FDDI
- ESCON
- FICON
- ATM

The ONLINE*7000* backplane is utilized for management (FCAPS capability: Fault Configuration, Administration, Provisioning and Service).

The ONLINE*7000* platform offers the following:

- Up to 33 protected wavelengths and 66 unprotected wavelengths
- Client interface through a Software Provisionable Transceiver: OC-3/12/48 (STM-1/4/16), GbE
- Bitrate Flexible Transceiver: 100 Mb/s to 2.5 Gb/s, including Fibre Channel, FICON, ESCON, D1 Video, HDTV, FDDI, Fast Ethernet, ATM, IP for long distance applications

Figure 444 on page 1186 is a general UPSR protection diagram of a single Network Element (shelf) consisting of a general Main Shelf without Expansion shelf, accompanied by a top-down I/O flowchart (Figure 445 on page 1187) in conjunction with card descriptions used for a two-node point-to-point/two-node ring configuration over extended distance.

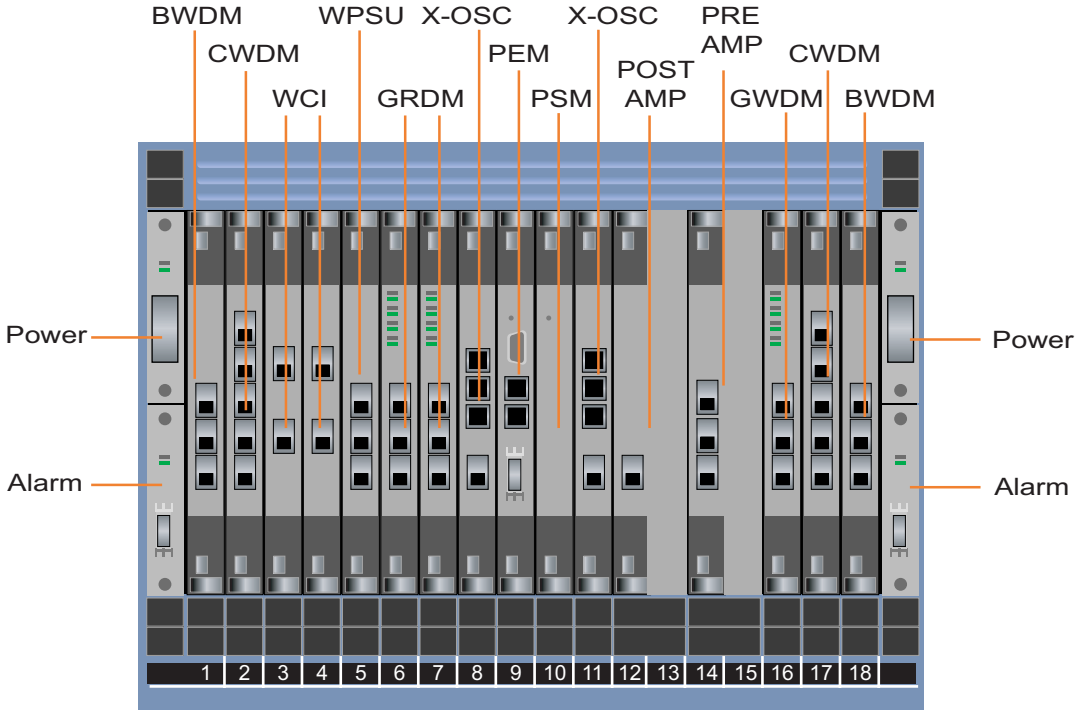


Figure 444 ONLINE*7000* UPSR protection diagram for main shelf

Within the ONLINE*7000* Network Element the cards are housed in at least two areas (Main Shelf and one or more Expansion Shelves) of a single DWDM node. The complexity and density of the multiplexing (number of GRDM cards) circuits determine the number of shelves utilized within the DWDM configuration.

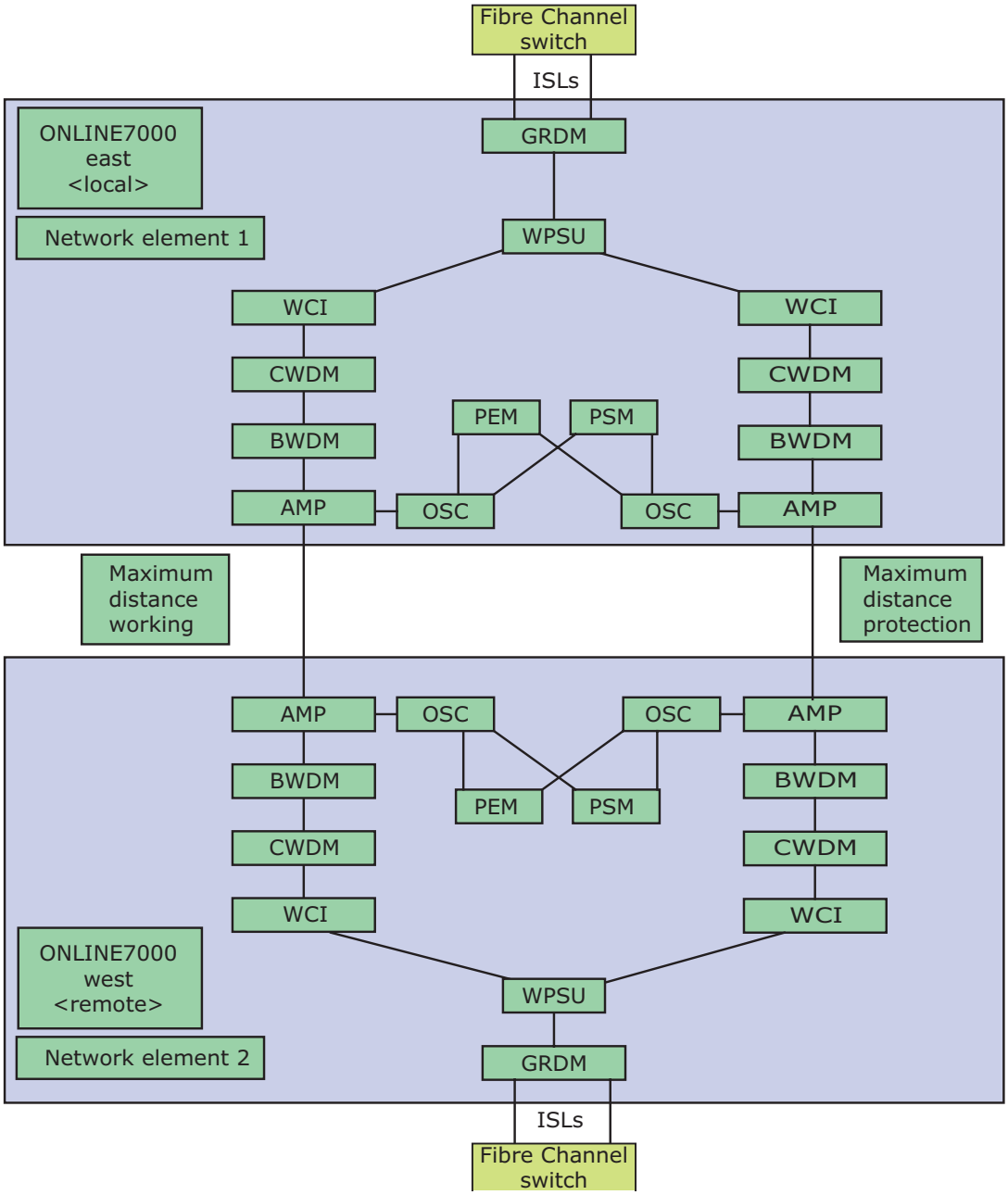


Figure 445 **I/O flowchart: Point-to-point/ring configuration over extended distance**

ONLINE*7000* card types

Note: All ports available on the following cards require MU cables.

GRDM CP: gigabit rate data mux circuit pack

GRDM cards are used for point-to-point traffic connections, and must be used in conjunction with 3RWCI CPs (Wave Channel Interface Circuit Packs responsible for Regeneration, Reshape, Retime of optical signals) by multiplexing two Gibabit Ethernet (GbE) or two Fibre Channel Channels MU ports (or tributaries) into a single (Tx, Rx) wavelength for transport across the ONLINE7000/9000/11000 system.

Conceptually bypassing the GRDM card is allowed, since WCIs support OC-48 bitrate capacity. However, Fibre Channel uses 1.0625 Gb/s, and WCIs accept only one port. GRDM accepts two ports and allows FC/GigE Performance monitoring.

The GRDM card accepts only MU Fibre Channel connections. GRDM will accept shortwave 850nm and longwave 1310nm connections. SC-to-MU cable converters are required in order to attach to EMC-supported switches.

Note: When an 850nm wavelength enters the GRDM, the wavelength out of the GRDM will be increased from 850nm to 1310nm.

Note: Do not mix protocols on the same GRDM card. (For example, Port 1 cannot use GbE if Port 2 uses Fibre Channel.)

WPSU: working protection splitter unit

Note: WPSUs are used in unidirectional path switched ring configurations (redundant path, redundant band with dedicated protection).

UPSR cards use optical couplers that have either one of these two capabilities:

- Splits a single (T_x, R_x) wavelength input from the GRDM card to two identical wavelengths utilized for working and protection lines.
- Combines two optical incoming signals into one output.

Direct connection to a WPSU is possible if the input received is an MU type OC-3 or 12 or OC-48 bandwidth connection.

WCI: wave converter interface (SONET/SDH input if there are no GRDMs)

The WCI converts the optical signal received by the WPSU card to an electrical pulse, and back again to an optical signal. WCI converts the 1310nm impulse from the WPSU to a frequency band (1530 to 1563.1 nm) utilized by the ONLINE*7000*.

Note: WCI supports up to 80 km.

Note: The number of WCIs should be double the number of GRDM cards in a UPSR protection configuration.

Note: WCI3RL is a card with 3R (reshape, regeneration, retime) capability, along with 160 km extended reach going from customer premise equipment to the WCI3RL.

CWDM: channel wave division multiplexer

The CWDM multiplexes up to three different wavelengths received from WCI cards into a single band and reroutes the single band to a CWDM on Network Element 1 of a two-node point-to-point connection. On the other side of the ring (Network Element 2 of a two-node point-to-point connection), CWDM demuxes the single band into a maximum of three different channels with the same band. CWDMs can also demultiplex the input of a BWDM to wavelength outputs to several WCI cards.

BWDM: band wave division multiplexer

This card multiplexes (adds) a band of three wavelengths coming from the CWDM and demultiplexes (drops) a band coming from the pre-amp. Working-line bands are in the range 1 through 5, and protection-line bands are 7 through 11.

Pre-amp: pre-amplifier

The pre-amp amplifies optical signals entering the node. The pre-amp is used in conjunction with a post-amp to compensate for signal loss caused by long spans between nodes, or it is used in conjunction with the Line CP if no post-amp is required in that span. pre-amps are field-replaceable and hot-swappable.

PEM: processor element module

PEMs contain the Software Application/Firmware on the Network Element (ONLINE*7000*). PEMs between multiple Network Elements share the same Global database (user privileges, circuits, Network Element configuration information, etc.) tables.

PSM: persistent storage module

PSMs contain nonvolatile storage of configuration and status information for the node. The PSM data store is implemented in Flash EPROM and appears to the PEM CP as a networked removable storage device.

The PSM serves as a secondary storage for the Network Element configuration data. Also, the connections on the front panel differ for the PSM CP and the PEM CP.

PSMs are field-replaceable and hot-swappable.

OSC: optical supervisory channel

OSCs are utilized for internetwork element communication. This card transfers information stored in the Global database tables of the PEMs. Different versions of OSCs exist.

Functionality of the following OSCs is the same, but the OSCs differ in link budgets and distance coverage:

OSC	Link budget	Distance
OSC	25 dB	80 km
OSCE	35 dB	110 km
X-OSC	35 dB (approximate)	120 km (approximate)

Post-amp: post-amplifier

The post-amp circuit provides the following functions:

- Splits off the Optical Supervisory channel (OSC) 1510 nm wavelength from the other data channels on the line fiber and redirects it for termination at the OSC CP.
- Monitors and adjusts the remaining wavelengths optical power.
- Amplifies the optical channels from the node out to the line as necessary. Amplification is done using an Erbium-Doped Fiber Amplifier (EDFA).
- Allows monitoring EDFA output power through a test-access port.

ONLINE*7000* topologies

The ONLINE*7000* supports these topologies:

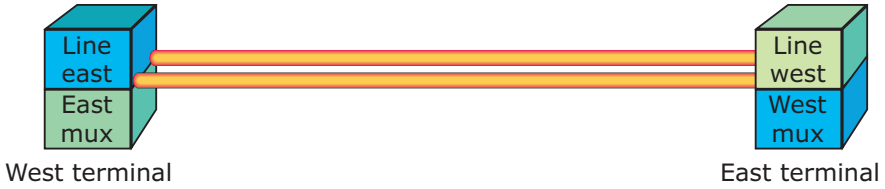


Figure 446 Point-to-point topology

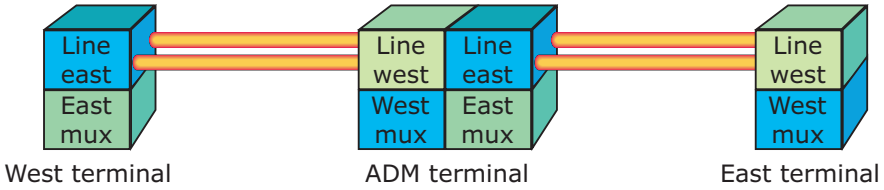


Figure 447 Line drop mode

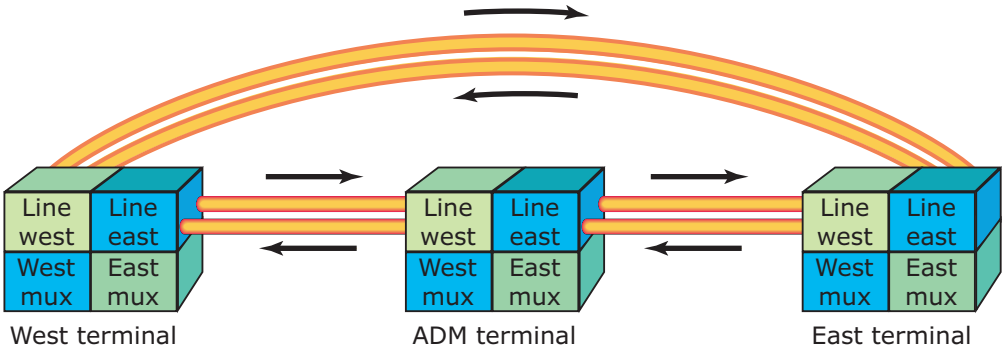


Figure 448 Ring network

ONLINE*7000* protection scheme

The ONLINE*7000* employs O-UPSR (Optical–Unidirectional Path Switched Ring) for data protection. In an O-UPSR ring, traffic is duplicated and sent around both sides of the ring simultaneously.

Note: The generic term for O-UPSR is ODPR (Optical Dedicated Protection Ring).

O-UPSR restoration is performed on a per-channel basis at those nodes where the channel enters or exits the ring (rather than at the intervening nodes). The destination node selects the better of the two signals and forwards the traffic to the subtending equipment.

Typically, the working path is selected unless it has failed or degraded. In the case of a failure or degradation, the destination node performs restoration by selecting the protected path.

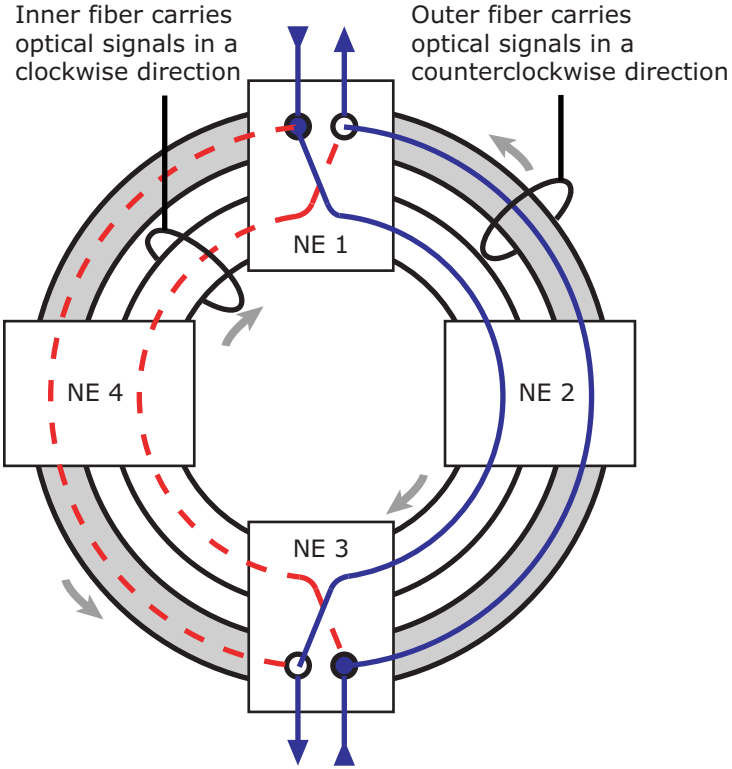


Figure 449 O-UPSR ring configuration in normal mode

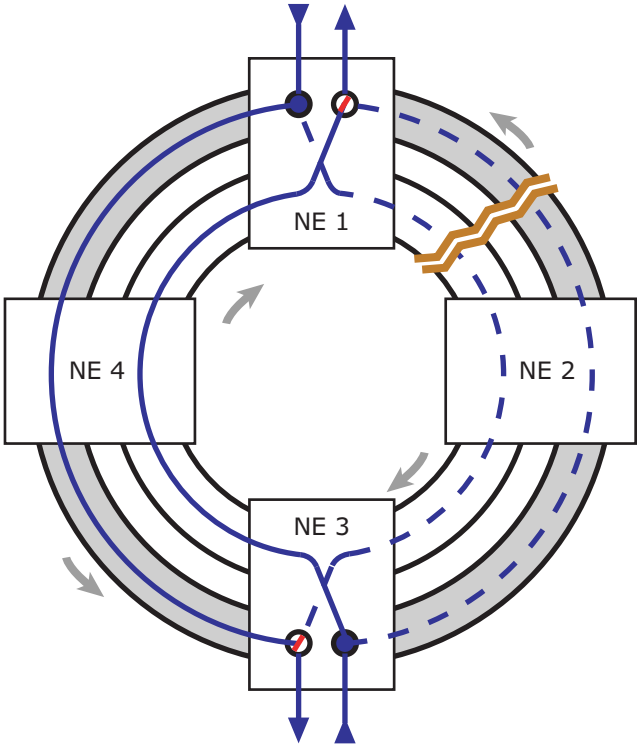


Figure 450 **O-UPSR ring configuration in failure mode**

Note: Optical UPSR dedicated protection: redundant path and bands.

Note: ONLINE7000 supports a maximum of 33 wavelengths utilizing a WPSU splitter on the tributary incoming links.

ONLINE*7000* power budget calculations

The link budget needed for the ONLINE*7000* to function must fall within the range of 33 dB or less. Ciena's power budget calculation is:

<LINK BUDGET> = [0.3 dB/km * <km fiber used>] +
[0.5 * (number of connects)] +
[<km fiber used> * 0.1 dB/km] + 1.0 dB

where:

- <number of connects> = [number of Cross-connects + termination].
- <0.1 dB/km> is used to calculate the Maintenance Margin.

ONLINE*7000* diagnostics and maintenance

For configuration, power budget calculation, and troubleshooting details, consult the Ciena technical publications.

Additional information regarding the ONLINE*7000* is accessible at http://www.ciena.com.

Ciena CN 4200/CN4200 MC

Ciena's CN 4200 and CN 4200 MC Advanced Services Platforms are multiservice switching, aggregation, and transport systems that allow carriers to groom, switch, and transpond a diversity of sub-wavelength client services onto higher-speed OTU1 (2.7 Gb/s) and OTU2 (10.7 Gb/s) transport streams. Using innovative timeslot technology, these service platforms can support a multitude of both like and unlike services. They can also directly transpond a variety of 10G services such as 10 GbE LAN/WAN PHY, 10G FC/FC1200 and OC-192/STM-64 into OTU2.

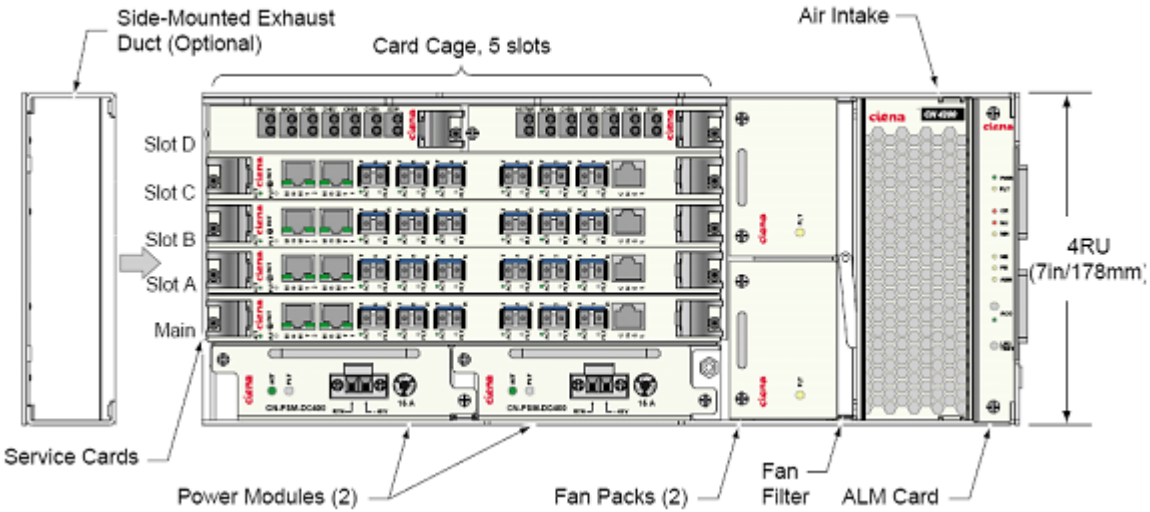


Figure 451 Ciena CN 4200 FlexSelect Advanced Services Platform

The full list of supported services includes:

- 10/100BaseT (supports jumbo frames)
- ESCON
- Fibre Channel, FC100, FC200, FC400, and 10G FC/FC1200
- FICON, both 1G and 2G
- Gigabit Ethernet and 1000BaseT (supports jumbo frames)
- OC-3/12/48/192
- STM-1/4/16/64

- 10 GbE WAN PHY
- 10 GbE LAN PHY (supports jumbo frames1)
- OTU1 and OTU2

For higher optical fiber efficiencies, the CN 4200 and CN 4200 MC support in-chassis CWDM and DWDM filters. These filters are designed in a modular fashion and accommodate growth up to 40 DWDM or 8 CWDM channels without service interruption. DWDM and CWDM channels can even be combined on the same fiber.

The CN 4200 and CN 4200 MC service platforms share a common transport architecture.

- The CN 4200 has four line card slots, each delivering client connectivity using OTU1 and / or OTU2 transport (Figure 452.)
- The smaller CN 4200 MC, which is intended for more focused site service requirements, has two line card slots (Figure 453 on page 1197.)

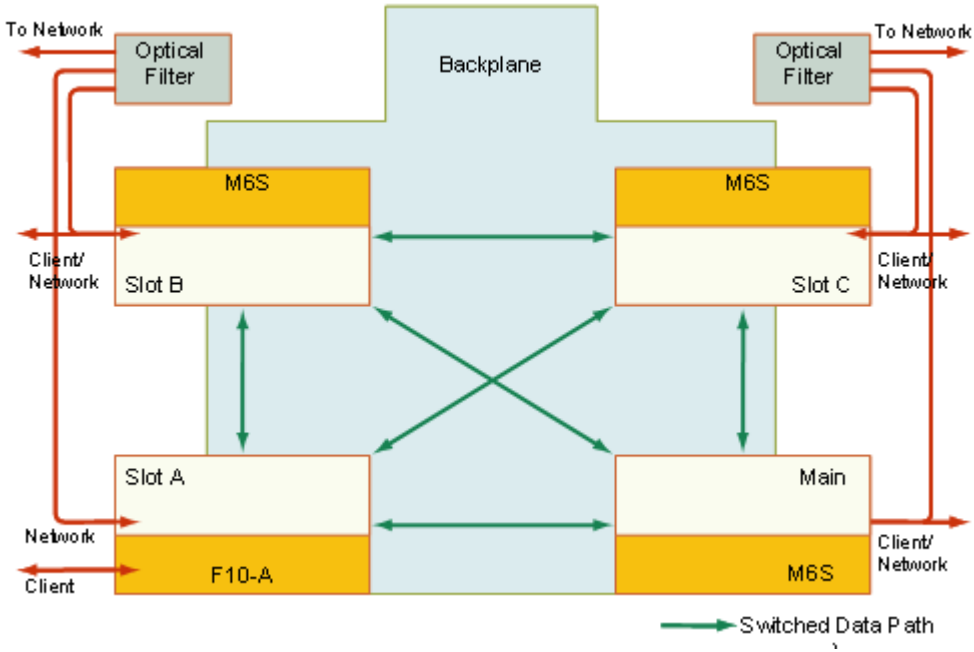


Figure 452 CN 4200 block diagram

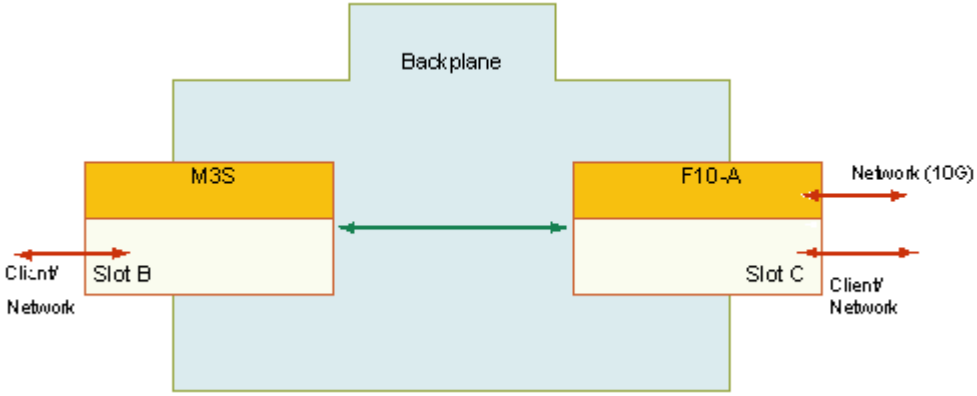


Figure 453 CN 4200 MC block diagram

Both systems employ a distributed switch architecture that provides sufficient grooming capability for stand-alone applications, yet is optimized for aggregation of services onto OTU1 and OTU2 wavelengths. Optional redundancy of interfaces and common elements guarantee operator-class service availability, with less than 25 ms automatic protection switching, faster than SONET/SDH.

Client service identities (e.g., ESCON, 1000BaseT) are software-provisionable rather than determined at the hardware layer, thereby offering a wide mix of services on the same card. The CN 4200 delivers standardized and rigorous performance monitoring metrics for all supported client services.

The CN 4200 and CN 4200 MC systems give operators a scalable solution for delivering multiservice transport and offer the following key features:

- High density multiservice transport platform supporting CWDM/DWDM
- Flexible assignment of ports as client or network ports
- Extensive performance monitoring on all client and network services
- Service-level loopbacks for fault isolation
- Front chassis accessibility for all connections and servicing

Available modules

This section provides information on available modules.

- "M6S / M3S"
- "F10-T" on page 1198
- "F10-A" on page 1199
- "FC4-T" on page 1199
- "Optical Protection Switch (OPS) modules" on page 1200
- "Optical Amplifier (OA) modules" on page 1201
- "Optical Supervisory Channel (OSC) module" on page 1202

M6S/M3S

The M6S is a 6-port multiservice module with programmable interfaces using SFP pluggable optics. Each port can be provisioned with its own service identity. The M3S is a 3-port version of the same module. Each M6S / M3S also contains a switch fabric.

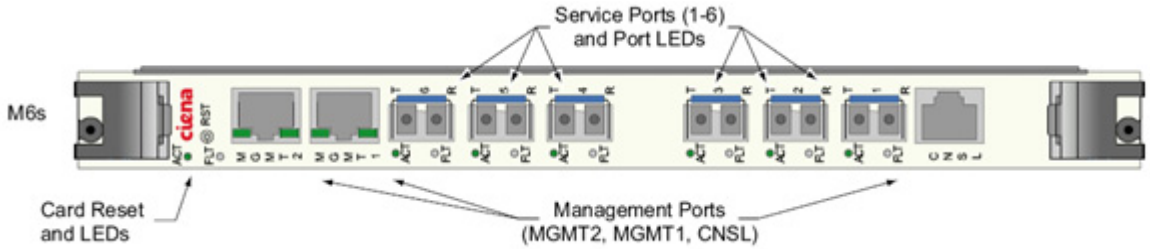


Figure 454 **M6S module**

The three RJ-45 ports (MGMT-1, -2, and CNSL) are used for remote and local management purposes, are accessible through the module's faceplate, and are only active if the module is functioning as the system controller. (Please note that although MGMT-1, -2 and CNSL are present in this figure, this is not an OA card.)

F10-T

The F10-T is a highly-adaptable transponder / regenerator card designed exclusively for 10G services. It is equipped with two full-duplex optical ports, a pluggable XFP transceiver port, and a network port implemented in discrete DWDM optics and intended

for long-range transport. This port can be equipped with either a fixed or tunable wavelength transmitter.

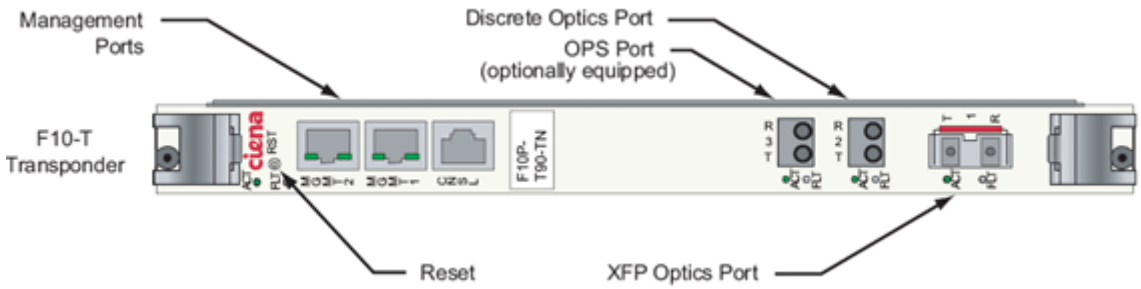


Figure 455 F10-T module

F10-A

The F10-A card is a service aggregation and grooming module that combines 2.5G services (such as ODU1 and OC-48/STM-16), as well as time-slotted services (such as Gigabit Ethernet and FC200), into a G.709 OTU2 stream. External connectivity is provided by two pluggable SFP transceiver ports and a network port implemented in discrete DWDM optics and is intended for long-range OTU2 transport. This port can be equipped with either a fixed or tunable wavelength transmitter. The F10-A switch fabric is similar to that on the M6S.

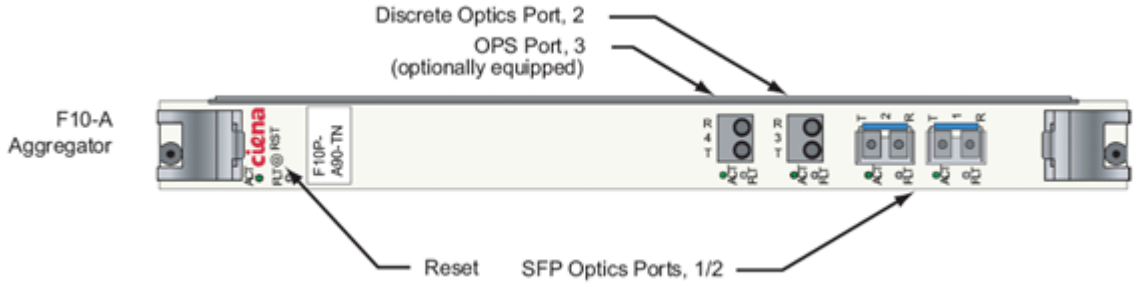


Figure 456 F10-A module

FC4-T

The FC4-T card is a service aggregation muxponder that combines three Fibre Channel client services into a single 10G OTU2 stream. The client services are software provisionable as either all FC200 or all FC400. (It does not support mixed client data rates.) The

aggregated 10G output of an FC4-T card can only be dissembled by another FC4-T. For this reason, FC4-T cards always work in pairs in a point-to-point manner.

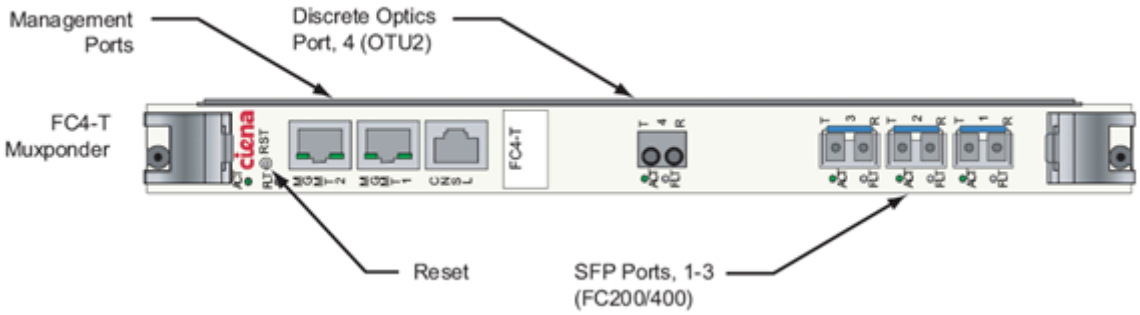


Figure 457 FC4-T module

Optical Protection Switch (OPS) modules

The CN 4200 / CN 4200 MC systems offer three OPS modules for single (OPS-1), dual (OPS-2), and 850nm (OPS-2 850) protection switching, each which can be used for both client-side and network-side protection switching. Optical protection switches protect against line failures by detecting a decrease or loss of optical power in the primary signal source and, if found, switching to the protecting signal source.

OPS-1

The OPS-1 module contains one optical protection switch and supports 1310nm, DWDM, and CWDM wavelengths on SMF.

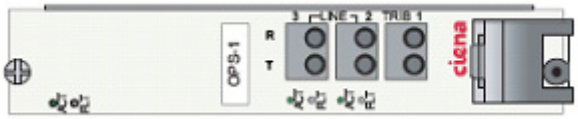


Figure 458 OPS-1 module

OPS-2

The OPS-2 module incorporates two optical protection switches into one module and supports 1310nm, DWDM, and CWDM wavelengths on SMF.

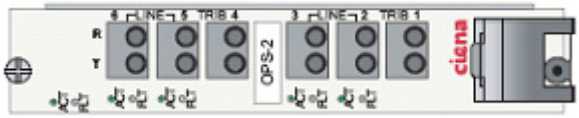


Figure 459 OPS-2 module

OPS-2 850

The OPS-2-850 module incorporates two optical protection switches, but it dedicated solely to support 850 nm services on 50µm and 62.50 µm.

Note: There is no single switch 850nm module.

This module is composted of 50 µm MMF and components. Additional loss is experienced when connected to 62.5 µm MMF.

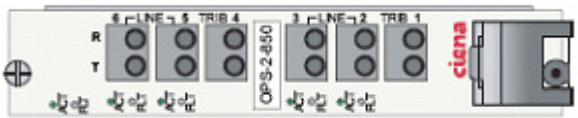


Figure 460 OPS-2 850 module

Optical Amplifier (OA) modules

The Optical Amplifier (OA) product suite comprises of the OAF-00-1-C (a fixed-gain optical amplifier), as well as the OAV-0S-U-C variable-gain optical amplifier. The OAs are C-band amplifiers and accommodate a single optical fiber for unidirectional transmission.

These amplifiers yield a bandwidth of 35nm and have a gain region of about 1530nm to 1565nm to coincide with Ciena's 100GHz and 200GHz DWDM channel plans.

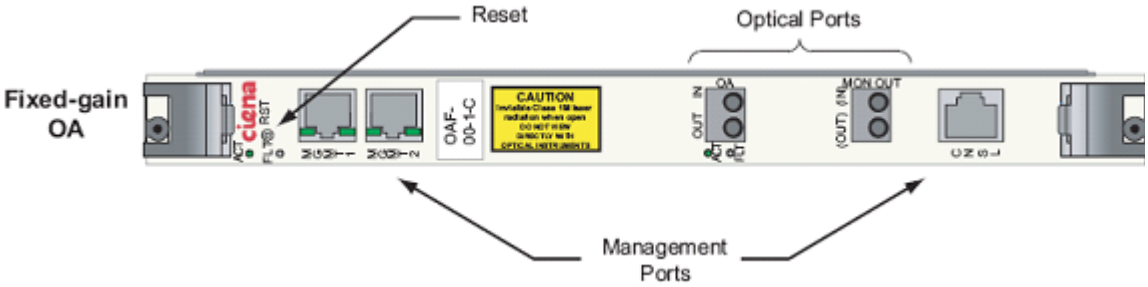


Figure 461 Fixed-gain Optical Amplifier module

The fixed-gain OA provides a constant gain and a flat gain shape across the DWDM C-band transport spectrum. The three RJ-45 ports that are used for remote and local management purposes (labeled MGMT-1, -2, and CNSL in Figure 451 on page 1195) are accessible through the module's faceplate and are only active if the module is functioning as the system controller.

Optical Supervisory Channel (OSC) module

The Optical Supervisory Channel (OSC) module offers OSC connectivity to nodes that have no integrated OSC function. The OSC module provides an out-of-band full-duplex communications channel for remote node management, monitoring, and control, similar in purpose to the DCC of SONET/SDH.

Equipped with its own 1510nm mux/demux filter, the OSC travels the same fiber as the DWDM stream and always terminates on

neighboring nodes. Because it is has a dedicated wavelength, the OSC data is entirely separate from the user data.

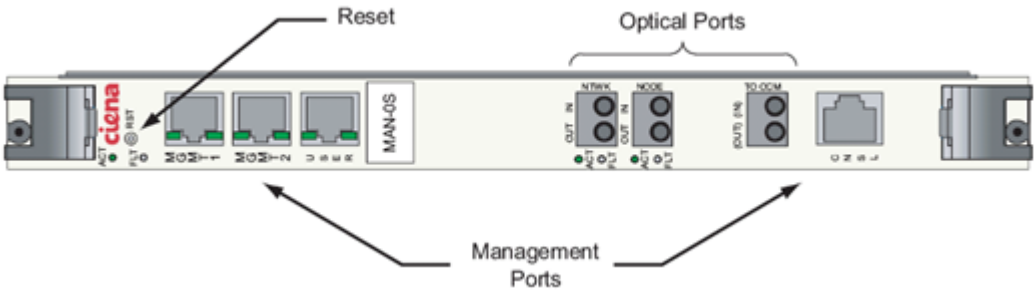


Figure 462 Optical Supervisory Channel (OSC) module

Common optical filters

The CN 4200 / CN 4200 MC chassis accommodates the deployment of Ciena common optical filters for integrated optical multiplexing capabilities. The filter modules come in half-width and full-width form factors. These modules include:

- DWDM, 100GHz ITU grid provides for up to 40 wavelengths
- DWDM, 200GHz ITU grid provides for up to 16 wavelengths
- CWDM, 20nm ITU grid provides for up to 8 wavelengths
- Wide band filters for 1310nm and 1550nm / DWDM band mux / demux.

Service protection schemes

The CN 4200 and CN 4200 MC offer the following types of protection:

- 1+1 redundant path line card protection, which prevents service failures due to both fiber and line cards
- 1+1 switch fabric-based protection, which prevents service outages due to switch fabric, line cards, and optical transceiver failures as well as fiber cuts
- Optical protection switching, which is useful for guarding against service outages in the event of either a fiber break or far-end transmitter failure

Redundant path line card protection for the CN 4200 / CN 4200 MC is used when the client supplies two client signals requiring path protection. In this case, the switching is performed by the client and not the CN 4200 / CN 4200 MC. An example is when a CN 4200 / CN 4200 MC is connected to an Ethernet switch implementing protection using IEEE 802.3ad Link Aggregation.

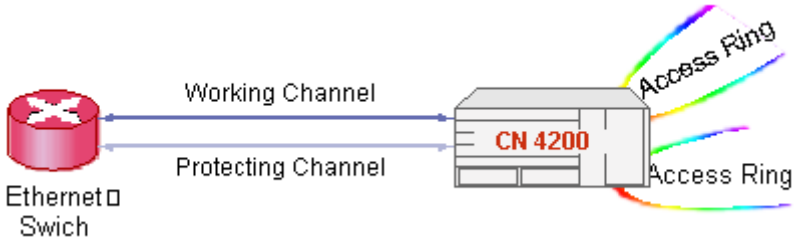


Figure 463 Redundant path line card protection

The advantage of line card protection is that during normal operation, with no failure conditions present, the client is free to use protection bandwidth as well as working bandwidth, doubling available bandwidth for the client.

For enhanced network protection requirements, the CN 4200 / CN 4200 MC distributed switch fabric provides a robust 1+1 network protection solution that prevents service outages due to not only fiber cuts, but line card failures as well.

Using the system's switching capabilities, operators can establish 1+1 protection at both the channel (i.e., timeslot) and ODU path levels. With switch fabric protection, a signal failure on the trunk fiber or on a trunk card causes the system to switch to the protecting signal. Channel layer protection is most appropriate for multiservice OTU1 streams where the sub-wavelength services terminate at different sites. Multichannel ODU-layer protection should be used whenever the sub-wavelength services of a multiservice OTU1 stream terminate at the same sites. Single-channel ODU layer protection should be established for any OTU1 stream that carries a single, directly-mapped service such as an OC-48 or STM-16.

For OTU2 protection, the CN 4200 / CN 4200 MC offers integrated optical switch protection on the 10G transport modules.

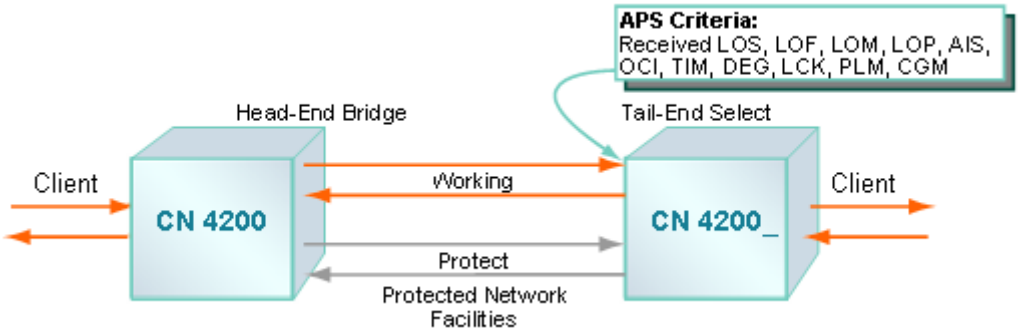


Figure 464 **CN 4200/CN 4200 MC optical protection switching**

Optical protection switching can also be implemented using the OPS modules. The discrete OPS modules allow network designers to deploy optical protection switching at several levels in the network to provide trunk protection, wavelength protection, or client protection. Trunk protection guards against network fiber breaks in the trunk line. Wavelength protection guards against wavelength failures, which may be the result of either a network fiber cut or a far-end transmitter failure. Client protection is positioned on the client interfaces of the transport cards, switching between working and protecting client ports and presenting a single protected channel to the client equipment.

Supported network topologies

This section describes supported network topologies.

Optically managed network topologies (wavelength level)

Using either CWDM or DWDM SFP transceivers and the appropriate optical filter modules, operators can implement any of the following optical transport topologies discussed in this section.

Point-to-point

Figure 465 on page 1206 shows an unprotected configuration that interconnects two systems, residing at separate sites, over a single fiber pair. In this configuration one optical filter positioned at each node muxes/demuxes several WDM OTU1 transport signals.

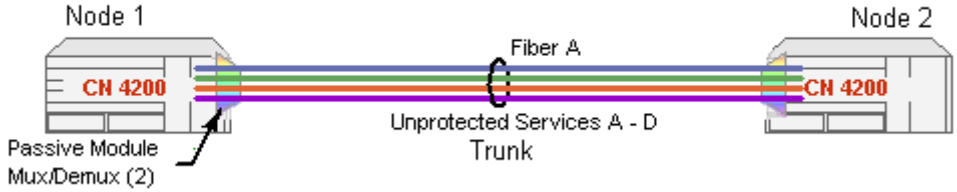


Figure 465 Unprotected point-to-point configuration

The CN 4200CN / CN 4200 MC also supports protected point-to-point topologies for maximum protection of client services. Figure 466 depicts one such configuration.

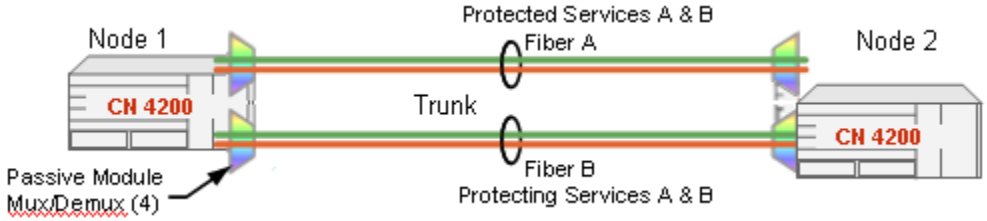


Figure 466 Protected point-to-point configuration

Linear add/drop point-to-point

The CN 4200 / CN 4200 MC can also support a linear optical add / drop point-to-point configuration where two terminal nodes bracket some number of intermediate add-drop nodes. In Figure 467, the terminal nodes are 1 and 4 and the add / drop nodes are 2 and 3.

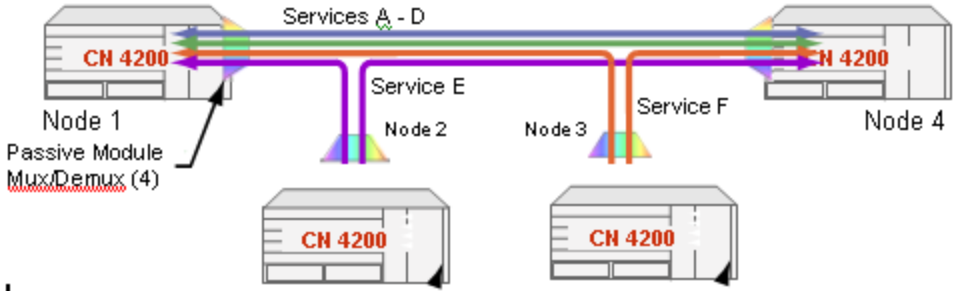


Figure 467 Linear Add/Drop configuration

Hubbed ring

A hubbed ring is composed of a hub and two or more add / drop or satellite nodes. All channels on the ring originate and terminate on the hub node. At add / drop nodes, certain channels are terminated (dropped and added) while channels not being dropped are passed through optically without being electrically regenerated.

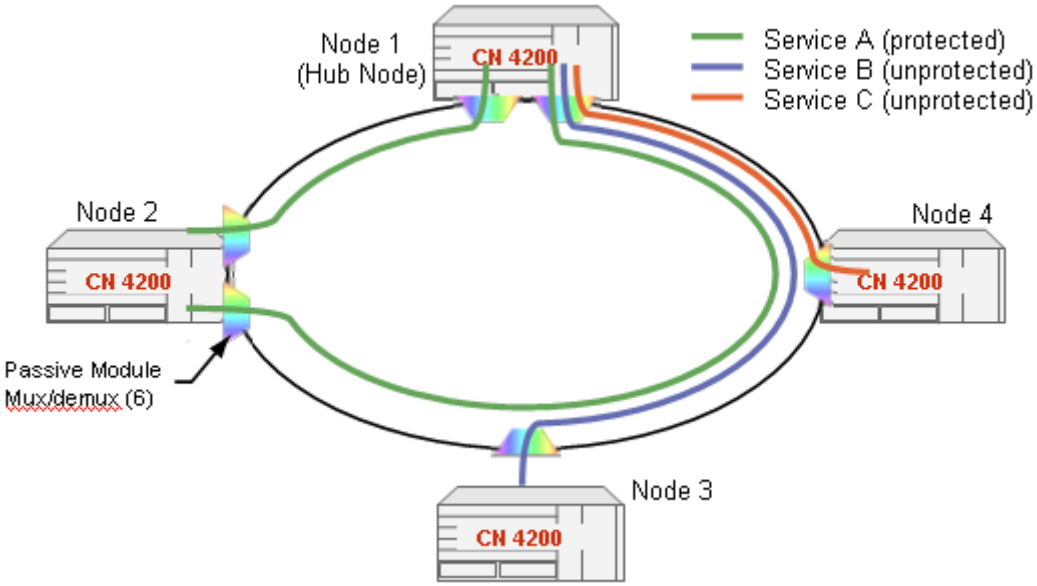


Figure 468 Hubbed ring configuration

Meshed ring

A meshed ring is a physical ring having the logical characteristics of a mesh. While traffic travels on a physical ring, the logical connections between individual nodes are meshed.

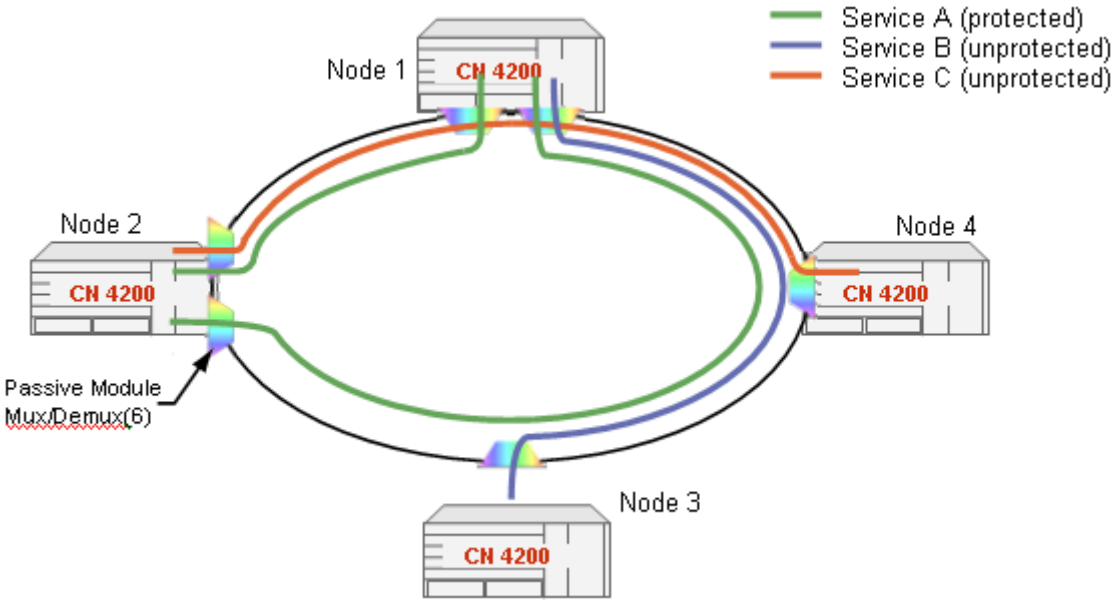


Figure 469 Meshed ring configuration

Electronically managed topologies (sub-wavelength level)

For any given aggregated wavelength, operators can implement the same basic topologies for the encapsulated client services as those described for wavelength-level networking. The logical connections are identical to those shown in the diagrams for the wavelength-level topologies (Figures 465 through 469) except that service aggregation is performed electronically by the distributed switch fabric rather than optically by optical filters.

Power budget calculations

Span margin

Margin is assigned depending on amplifiers present. The variable gain amplifier (OAV) with span loss compensation (see "Span loss compensation" on page 1209) automatically adjusts for span margin. For fixed-gain amplifiers (OAFs), 0.4dB plus 3 percent of span loss needs to be added to the calculation. Ciena component specifications already include connectors.

Span loss compensation

Ciena's span loss compensation feature mitigates the effects of aging spans by dynamically adjusting the gain of a span's receive amplifier to offset any increased loss, such as that which occurs due to fiber aging. Span loss compensation acts on a per fiber basis, thus ensuring independent gain adjustments on each fiber in the span. Ordinarily, the affect of aging spans must be set aside as system margin during initial span engineering, but when this loss is offset through dynamic span loss compensation, much less margin (6 - 8dB for a regional system) need be held in reserve.

Diagnostics and maintenance

For configuration, power budget calculation, and troubleshooting details, consult the Ciena technical publications. Additional information regarding the CN 4200/ CN 4200 MC is accessible at http://www.ciena.com.

Cisco Metro 1500

The Cisco Metro 1500 platform is a true protocol and bit-rate-independent fiber-optic transport system. It supports the following protocols: SONET, ATM, Gigabit Ethernet, IP , FDDI, and all optical interfaces (OC-n).

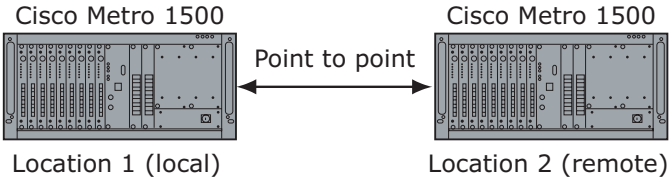
- The Cisco Metro 1500 does not have an internal backplane, and the optical connections are made using external jumpers.
- The system can support up to eight channel cards and has redundant (two) power supplies.
- Three types of channel card can be inserted into the chassis:
  - Low-Speed (LS) (data rate 10 to 200 Mb / s)
  - High-Speed (HS) (data rate 100 to 1250 Mb / s)
  - OC-48 (2.5G) (data rate 2.5 Gb / s)

DWDM Modules are *mux* and *demux* modules. The point-to-point configuration is normally connected between the mux and demux modules. If you require a protected scheme you will have to provision the boxes with RSM modules and use their connectors to run two single-mode (dark) fiber connectors. The RSM module monitors the signals and, if the primary connections fail, switches to the secondary backup connector or link.

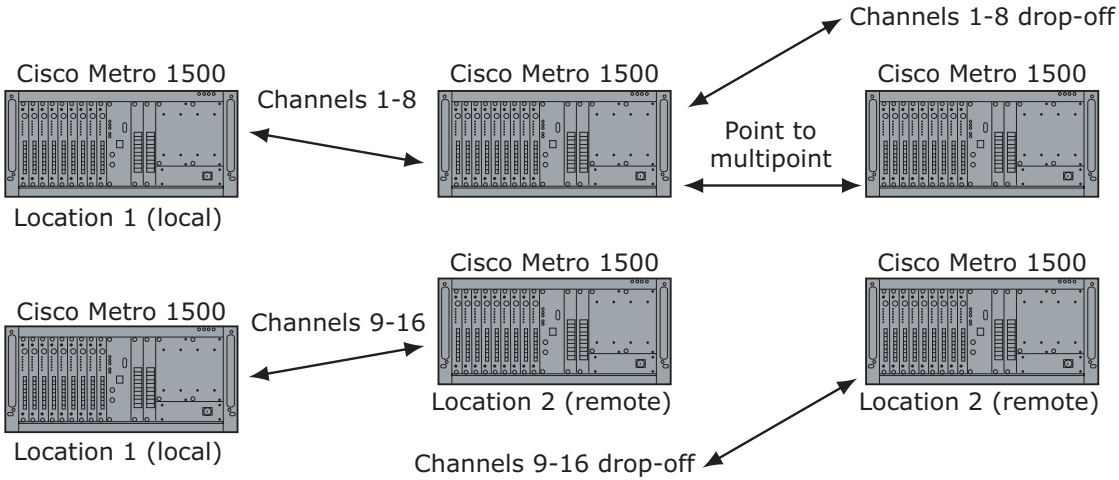
The operator should be aware that the RSM modules are a passive device and cause significant power losses to the system. The result is a significant shorter maximum distance; E-Lab Navigator contains information that can help you determine that distance.

The following configuration(s) can be implemented using the Cisco Metro 1500 system:

- **Point-to-point** — The point-to-point configuration, which is the basis for all other configurations, will include a local and remote site. The data will flow between the different sites using two links. Each link includes a transmit and receive single-mode fiber cable (dark fiber).



- **Point-to-multipoint**



Metro 1500 power budget calculations

Use the following information to calculate the power budget:

- Link loss budget for Cisco Metro 1500 input cards:

Card	Link loss budget
Standard	19 dB
Extended	24 dB
2.5 Gb/s	10 to 19 dB

- Fiber losses are calculated on the distances times the specifications of the fiber cable, in this case .2dB per km (80 km * 0.2 dB/km = 16 dB loss) and each connector has approximately 0.5 dB loss per connector.
- BSM (Band Splitter Module) has an insertion loss of 3dB. Mux insertion loss is approximately 4 dB.
- Demux insertion loss is approximately 4 dB. RSM insertion loss is approximately 3 to 5 dB. BSM insertion loss is approximately 3 dB.

No RSM, no BSM (unprotected mode) extended

Maximum link budget = 24 dB:

1. Calculate total link budget minus mux insertion loss minus demux insertion losses: 24 dB – 2(4) dB = 16 dB.
2. Divide resultant power budget by cable specification: 16 dB divided by 0.2 dB / km = 80 km (50 miles).

RSM, no BSM (protected mode) extended

Maximum link budget = 24 dB:

1. Calculate total link budget minus mux insertion loss minus demux insertion loss minus RSM insertion loss: 24 dB – 2(4) dB – 5 dB = 11 dB.
2. Divide resultant power budget by cable specification: 11 dB / 0.2 dB / km = 55 km (34 miles).

Supported Metro 1500 topologies

Figure 470, next, and Figure 471 on page 1213 show supported Cisco Metro 1500 topologies. (E-Lab Navigator contains distance limitations.)

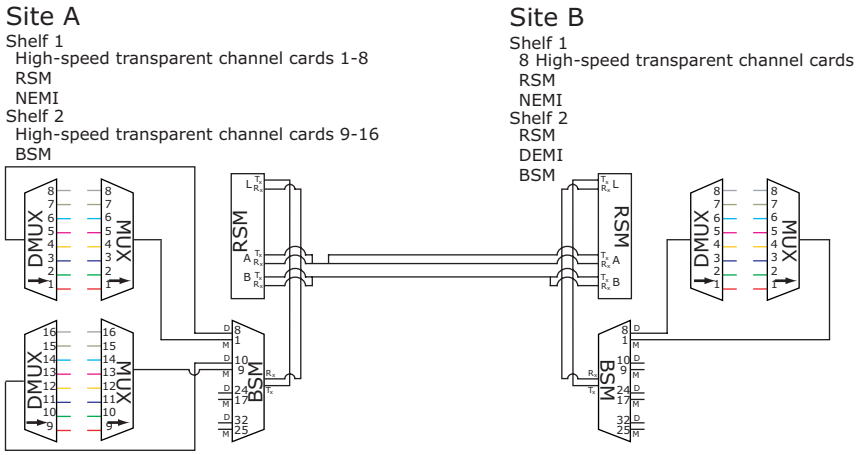


Figure 470 Point-to-point (two-site) configuration

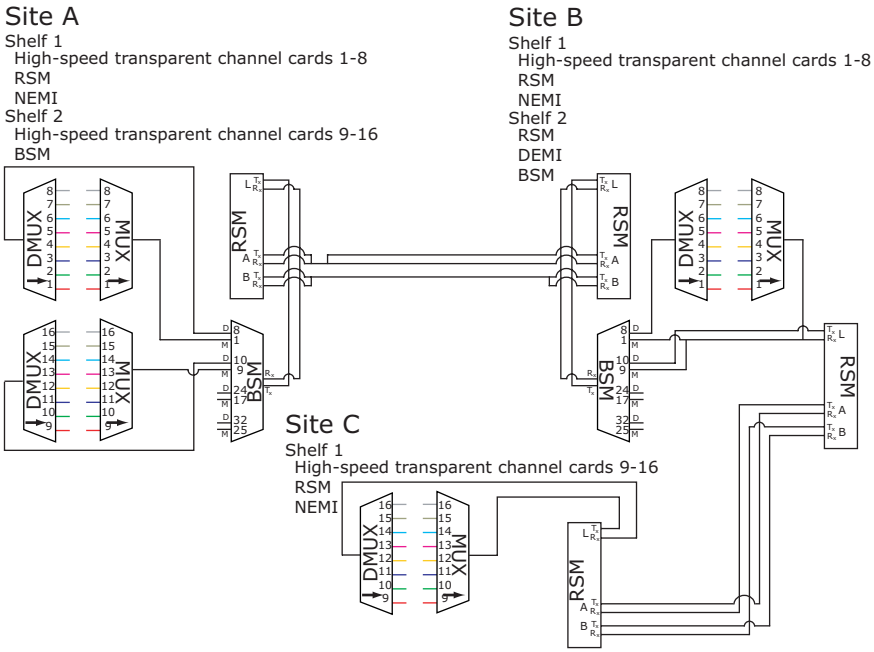


Figure 471 **Multi hop (three-site) configuration**

Metro 1500 diagnostics and maintenance

For configuration, power budget calculation and troubleshooting details consult the Cisco Metro 1500 technical publications.

Using the Symmetrix Fibre Channel director online utilities can complement repair and system diagnostics.

Cisco ONS 15540

The Cisco ONS 15540 platform (shown in Figure 472) is a true protocol and bit-rate-independent fiber-optic transport system. A client's 1310nm single-mode SC Fibre Channel protocol connection is converted from an electrical signal to an ITU G.692 compliant wavelength, and then optically multiplexes it with the other client signals for transmission over a fiber-optic link and demultiplexed from an optical wavelength back to an electrical signal.

The system is capable of supporting 3R (Reshape, Regeneration, Retime) mode with monitoring capabilities are Gigabit Ethernet, SONET (OC-3, OC-12, OC-48), SDH (STM-1, STM-4, STM-16), Fibre Channel (1 Gb/s), ESCON and FICON.

Up to 32 channels are supported on the ONS 15540 and 1310nm single-mode cables are accepted to the ONS 15540 unit.

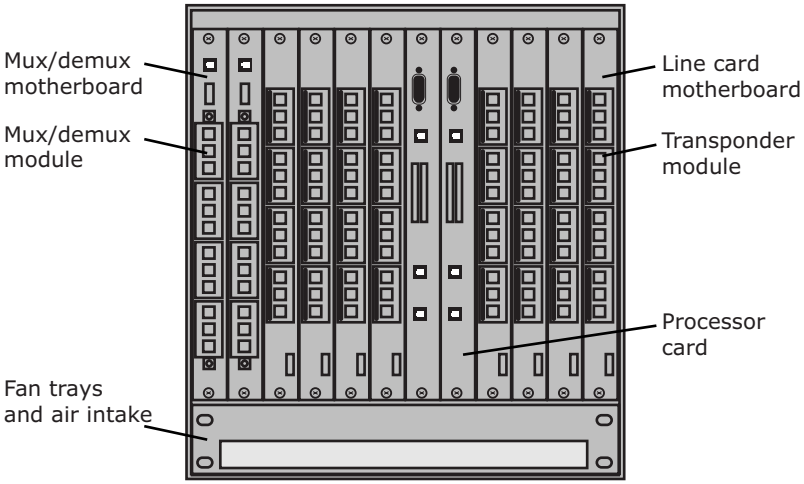


Figure 472 **Cisco 15540**

A 12-slot modular vertical chassis that receives power through redundant –48 VDC inputs and contains a redundant external AC power supply. The two leftmost slots (slots 0 and 1) hold the mux/demux motherboards containing optical mux/demux modules corresponding to the east and west directions. Slots 2 through 5 and 8 through 11 hold the line card motherboards, which are populated with transponder modules. Slots 6 and 7 hold the processor cards.

The system has an optical backplane for carrying signals between the transponders and the optical mux/demux modules and an electrical backplane for system control.

Figure 473 illustrates the principle functions required for transmission of the signal between the client and transport networks, within the Cisco ONS 15540.

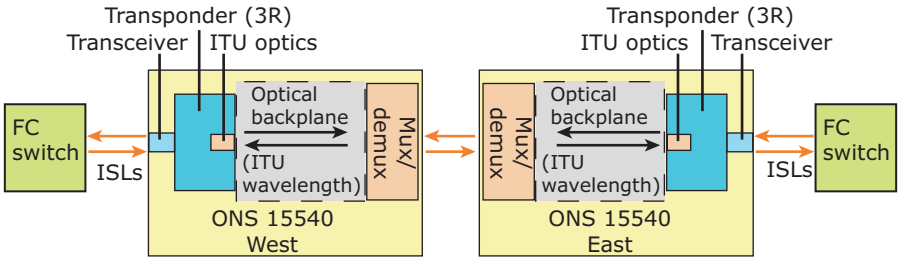


Figure 473 Client/network signal transmission in Cisco ONS 15540

From the *transmit* direction, the following functions occur. (From the *receive* direction the opposite functions occur.)

1. A transceiver receives the 1310 nm input client signal.
2. The client signal is converted to electrical; then 3R-function-signal enhancements are performed in the transponder.
3. A modulated laser diode converts the electrical signal to optical wavelength compliant to ITU laser grid.
4. Transmission of the signal is implemented over the optical backplane.
5. The optical Mux/demux module takes multiple (clients') input signals and multiplexes them into a single DWDM signal.
6. The DWDM signal is placed onto fiber for transmission on the DWDM network side.

ONS 15540 system components

The Cisco ONS 15540 has a modular architecture that allows flexibility in configuration and permits incremental upgrades of the system.

Transponder modules

The transponder modules populate the line card motherboards and have two interfaces; an external interface that connects to client

equipment and an internal interface that connects to the line card motherboard.

Client-side interfaces

Client-side interfaces on the transponder module is protocol transparent and bit-rate transparent and accepts a client signal on the 1310nm wavelength through SC connections.

Both multimode (MM) and single-mode (SM) fiber are supported for client connections on specific protocols.

Encapsulation of client signals is supported on the transponder interfaces in either 3R-enhanced mode, which allows some client protocol monitoring (such as code violations and data errors) 3R mode, where the transponder is transparent to the client data stream or in regular 3R (Fast Ethernet, FDDI, 2 Gb/s Fibre Channel).

Note: For Fibre Channel switch (1 Gb/s) attachment, only 1310 nm wavelength single mode cables are supported on the ONS 15540. 850 nm multimode cables are supported only for serial (ESCON) protocols.

Transport-side interfaces

In the transponder module, the client signal is regenerated and retransmitted on an ITU-compliant wavelength across the optical backplane. The laser on each transponder module is capable of generating one of two wavelengths on the transport side. Thus, there are 16 different transponder modules (for channels 1 and 2, 3 and 4, 5 and 6, and so on, up to 31 and 32) to support the 32 channels.

Transport-software determines which wavelength each module should generate based on whether it is inserted in the upper (subslot 0 or 2) or lower (subslot 1 or 2) of a line card motherboard.

A safety protocol, LSC (laser safety control), is provided to shut the transmit laser down on the transport side when a fiber break or removed connector is detected.

The transponder modules are hot-pluggable, permitting in-service upgrades and replacement.

Line card motherboards

Line card motherboards hold the transponder modules and provide the optical connections from the transponder modules to the optical backplane. The line card motherboards are modular and are capable of being populated based on customer requirements.

One system can hold up to eight line card motherboards, each of which accepts four transponder modules.

Three types of line card motherboards are available: Splitter, East, and West:

- Splitter supports protection against fiber failure by delivering the ITU wavelengths emitted from their associated transponders over the optical backplane to the optical mux/demux modules in both the west and east slots (slots 1 and 2, respectively).
- East and West deliver the ITU wavelengths from their associated transponder modules over the optical backplane to the optical mux/demux modules in either the east or west slot.

Mux/demux motherboards

The mux/demux motherboards hold the optical mux/demux modules. Either slot 0 or slot 1 can be populated with a single mux/demux motherboard for unprotected operation, or both slots can be populated for protected operation. Each motherboard can accept up to four optical mux/demux modules depending upon the type of module used, and can be populated according to user needs.

OSC (optical supervisory channel)

There are two versions of the mux/demux motherboard with and without the OSC. Implemented with a dedicated laser and detector for a 33rd wavelength (channel 0) on the mux/demux motherboard, the OSC is a per-fiber duplex management channel for communicating between Cisco ONS 15540 systems. The OSC allows control and management traffic to be carried without the necessity of a separate Ethernet connection to each Cisco ONS 15540 in the network.

The OSC carries the following types of information:

- CDP (Cisco Discovery Protocol) packets – Used to discover neighboring devices
- IP packets – Used for SNMP and Telnet sessions between nodes
- OSCP (OSC Protocol) – Used to determine whether the OSC link is up

OSC is required for each Cisco ONS 15540 system in order for the NMS to manage it on the network. Without an OSC module, ONS 15540 system must be separately managed by Ethernet or serial connections.

Optical mux/demux modules

Mux/demux modules are responsible for optically multiplexing and demultiplexing a specific band of ITU wavelengths. In the transmit direction, the optical mux/demux modules multiplex signals transmitted by the transponder modules over the optical backplane and provide the interfaces to connect the multiplexed signal to the

transport (DWDM) network. From the receive direction the optical mux/demux modules demultiplex the signals from the transport network side before passing them over the optical backplane to the transponders.

Two types of optical mux/demux modules are used in the ONS 15540, add/drop and terminal. Each module supports a range of channels called a band. In regards to add/drop mux/demux modules, a band contains 4 or 8 channels; in the case of the terminal mux/demux modules, a band contains 16 channels.

ONS 15540 channels	4-channel add/drop mux/demux module	8-channel add/drop mux/demux module	16-channel terminal mux/demux module [a]
1 through 4	Band A	Band AB	Band AD
5 through 8	Band B		
9 through 12	Band C	Band CD	
13 through 16	Band D		
17 through 20	Band E	Band EF	Band EH
21 through 24	Band F		
25 through 28	Band G	Band GH	
29 through 32	Band H		

a. A 16-channel terminal mux/demux module occupies two subslots in a mux/demux slot.

An *add/drop* mux/demux module adds a specified band of channels at a node and passes the other bands through. To support the 32-channel spectrum, there are eight different 4-channel modules and four different 8-channel modules (as shown in the table).

Processor cards

The Cisco ONS 15540 includes two processor cards for redundancy. Each processor is composed of a number of subsystems, including a CPU, a system clock, Ethernet switch for communicating between processors and with the LRC (line card redundancy controller) on the mux/demux and line card motherboards, and a processor redundancy controller.

The active processor controls the node, and all motherboards, and a processor redundancy controller. The active processor controls the node, and all cards in the system make use of the system clock and synchronization signals from the active processor.

Available ONS 15540 topologies

The Cisco ONS 15540 supports the topologies shown in Figure 474, next, through Figure 478 on page 1221. E-Lab Navigator contains information that can help you determine which topologies EMC supports.

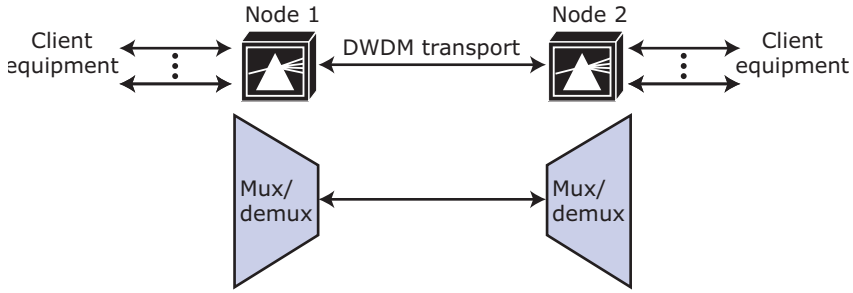


Figure 474 Unprotected point-to-point topology

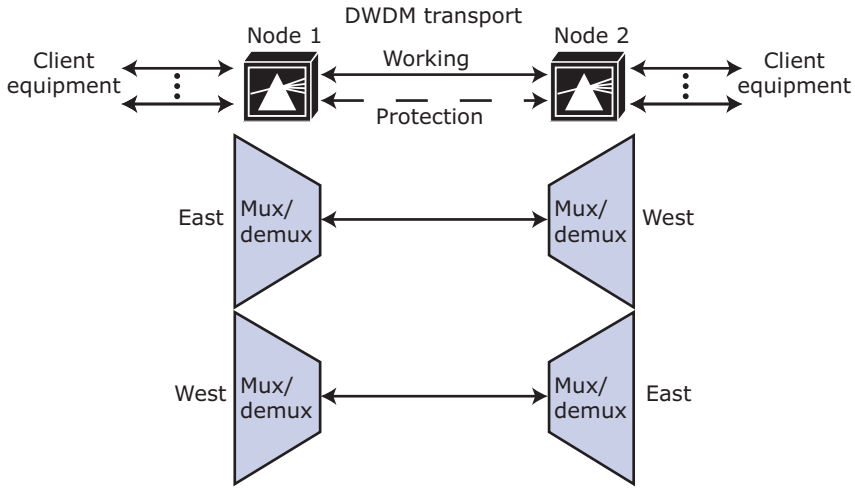


Figure 475 Protected point-to-point topology

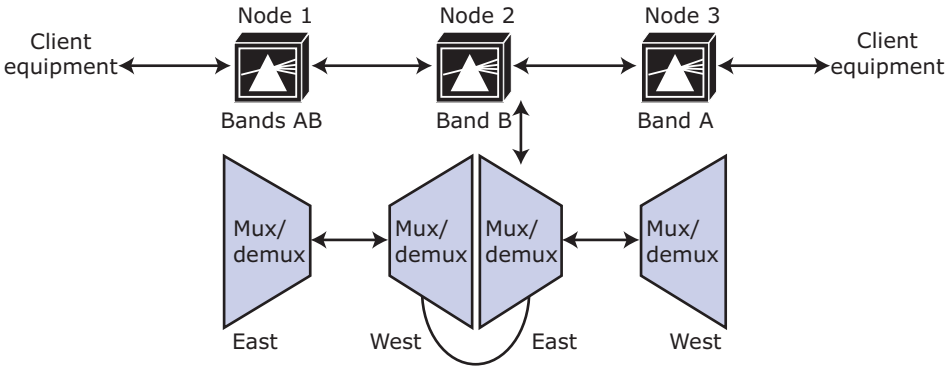


Figure 476 Bus topology

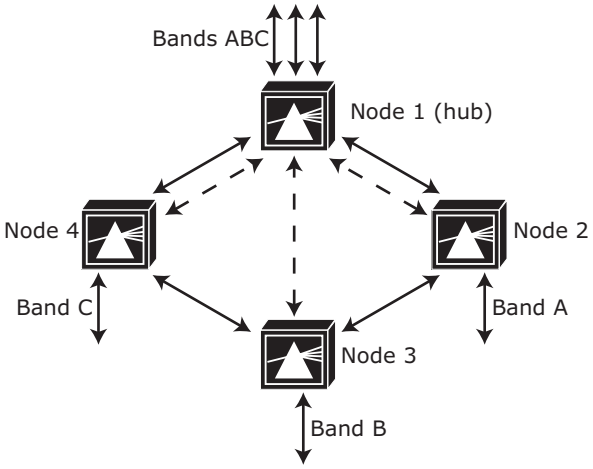


Figure 477 Hubbed-ring topology

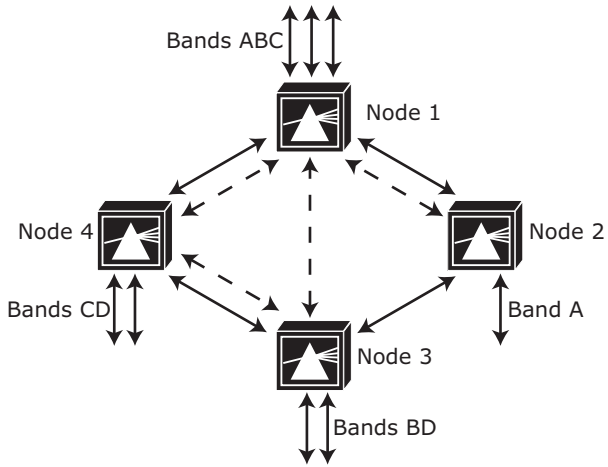


Figure 478 **Meshed ring topology**

ONS 15540 protection scheme

To survive a fiber failure, fiber-optic networks are designed with both working and protection fibers. In the event of a fiber cut or other facility failure, working traffic is switched to the protection fiber. The Cisco ONS 15540 supports such facility protection using a splitter scheme (shown under "Splitter protection considerations" on page 1222) to send the output of the DWDM transmitter on two transport side interfaces.

A splitter on each line card motherboard couples the transponder's DWDM interface across the optical backplane to the internal interfaces on the optical mux/demux modules in the east and west mux/demux slots.

On the transport side, one fiber pair serves as the working connection, while the other pair provides protection. The signal is transmitted on both connections, but in the *receive* direction; an optical switch selects one signal to be the active one.

If a loss of light is detected on the working fiber, a switch to the standby signal is made under control of the LRC (line card redundancy controller).

Assuming, for example that the working signal in the diagram below is on the east interface, a failure of the signal on that fiber would

result in a switchover, and the signal on the west interface would be selected for the *receive* signal.

Splitter protection considerations

The following considerations apply when using splitter protection:

- Splitter protected line card motherboard supports splitter protection. The signal splitter introduces 4.6 dB of loss in the transmit direction. Cisco recommends using the nonsplitter protected line card motherboards (east or west version) for configurations where splitter protection is not required.
- Switchover after a failure under splitter protection is nonrevertive. After a switchover, manual intervention is required to revert to using the previously failed fiber for the working traffic once the fault has been remedied.
- The OSC plays a crucial role in splitter-based protection by allowing the protection fiber to be monitored for a cut or other interruption of service.
- LSC (laser safety control) is not available when splitter protection is used.

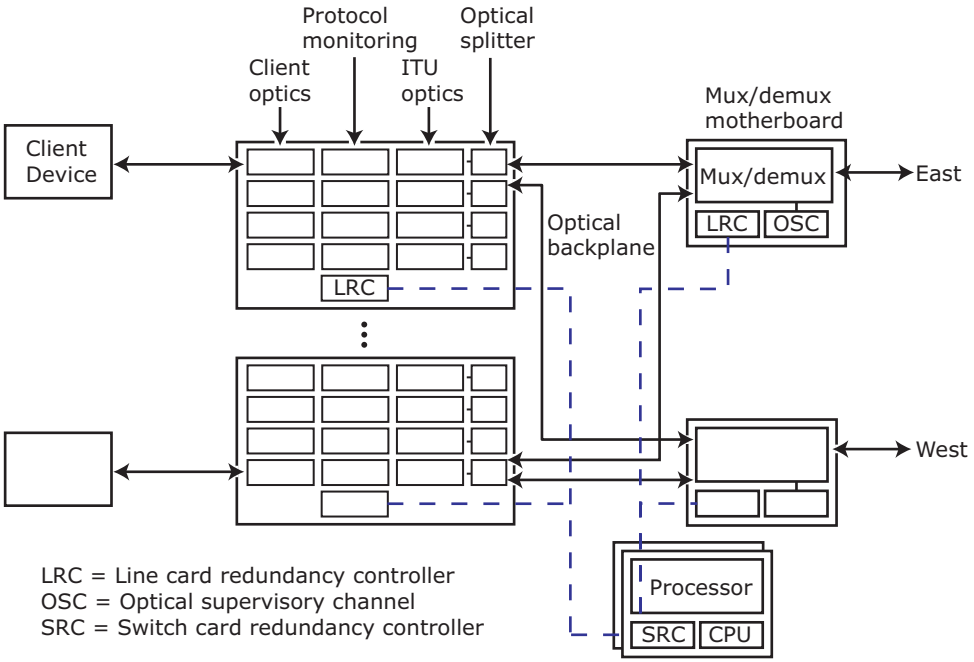


Figure 479 Cisco 15540 splitter protection diagram

ONS 15540 power budget calculations

Use the following information to calculate the power budget:

Total Loss = (Line Card Motherboard Receive Loss) + (Mux/Demux module drop loss) + (Fiber loss) + (Mux/Demux module transmit loss) + (Line Card Motherboard transmit loss)

ONS 15540 overall optical link loss budget

Transmit power and receiver range:

Channel	Transmit power (dBm)	Receiver sensitivity (dBm)
Data Channels	6	-32 to -8
OSC	6	-24 to 0

ONS 15540 diagnostics and maintenance

For configuration, power budget calculation and troubleshooting details, consult Cisco technical publications.

Additional information regarding Cisco ONS 15540 is accessible from http://www.cisco.com.

Cisco ONS 15454 MSTP

This section contains information on the Cisco ONS 15454 MSTP, Multi-Service Provisioning Platform), including:

- "Intended audience," next
- "Description" on page 1224
- "Key features" on page 1225
- "System architecture" on page 1228
- "MSTP SFP support" on page 1242
- "SFP technical details" on page 1253
- "MSTP management" on page 1259
- "MSTP further reading" on page 1260

Intended audience

This material is intended for Technical consultants, solutions architects, implementation specialists, end-users, or anyone interested in learning more about the features of the ONS 15454 Multi-Service Transport Platform (MSTP).

Description

The Cisco ONS 15454 MSTP allows a metro or regional LH DWDM system to become as intelligent and flexible as the Cisco ONS 15454 Multi-Service Provisioning Platform (MSPP), including wide service interface mix, service transparency, flexible topology, completely reconfigurable traffic pattern, and simplified operations. The Cisco ONS 15454 MSTP, with its MSPP capabilities, supports a broad range of standards-based services in a single platform. Figure 480 on page 1225 shows the Cisco ONS 15454 MSTP.

Figure 480 **Cisco ONS 15454 MSTP ANSI and ETSI mechanics**

Key features

The key features mentioned in this section may or may not be supported by EMC. Please refer to the *EMC Support Matrix* for the latest updates on client level support.

The Cisco ONS 15454 MSTP includes the following features:

Node configurations

- Terminal
- Hub
- Line amplifier
- OADM
- 2° Reconfigure Optical Add/Drop Multiplexer (ROADM)
- Multidegree ROADM (3-8°)

Network configurations

- Linear point-to-point
- Open ring, single hub
- Open ring, multiple-hub
- Closed ring, no hub
- Mesh and multiple-ring topologies

Protocol features

- 10 GE LAN PHY
- 10 GE WAN PHY
- 10 G Fibre Channel
- OC-768 / STM-256
- OC-192 / STM-64
- OC-48 / STM-16
- OC-12 / STM-4
- OC-3 / STM-1
- 4 G Fibre Channel / FICON
- 2 G Fibre Channel / FICON
- 1 G Fibre Channel / FICON
- InterSystem Channel-3 (ISC-3) p
- InterSystem Channel-1 (ISC-1) c & p
- 1G Ethernet
- ESCON
- Fast Ethernet
- HDTV
- D1 Video
- DV-6000
- DVB-ASI
- SDI
- Sysplex CLO & Sysplex ETR

Advanced intelligent software features

- Network topology Auto Discovery
- Point-and-click node and network setup and regulation
- Automatic network optical power management and monitoring
- Single management interface (single IP address) for all the shelves in a node
- Network-level alarm correlation for a quick and easy troubleshooting (G.798-based)
- DCN extension to provide the possibility to use any available DCN access (including DCC and GCC bytes) for management of nodes
- Automatic Node turn-up for installation and deployment without the use of Cisco Transport Planner parameters

User interface: Cisco Transport Controller (CTC)

- Integrated node and subnetwork craft GUI
- Layered graphical views: network, wavelength, node, shelf, card
- User-provisionable graphics and fonts
  - Background maps
  - Color schemes

- A-to-Z wavelength circuit routing and creation
- Network autodiscovery with provisionable subnetwork domain control
- System inventory
- PC-based client
- Familiar browser interface-Netscape Navigator or Microsoft Internet Explorer
- Complete Performance Monitoring support
  - 15-minute (32 entries) and 24-hour (two entries)
  - Optical layer
  - SONET/SDH layer
  - ITU-T G.709 layer (including FEC/E-FEC)
  - Client interface type-specific
  - Threshold-crossing alerts threshold setting

For more information on CTC, refer to "Cisco Transport Controller (CTC)" on page 1259.

Alarm monitoring and reporting

- Shelf LEDs - Critical, major, minor, remote
- Card LEDs - Card failure, active/standby state, signal fail
- Cisco Transport Controller craft interface
  - Layered graphical views with real-time alarm text and coloring: network, wavelength, node, shelf, card
- Multiple technology views including DWDM and SONET/SDH with MSTP integration
- Environmental alarm contacts
- 4-alarm output contact closures (standard): critical, major, minor, remote
- Up to 48 provisionable alarm contacts in systems equipped with Alarm Interface Controller (AIC-I)

Network security features

- Four-level user control with provisionable timeout durations: superuser, provisioning, maintenance, retrieve
- Multiple user names and logged-in users
- RADIUS

Maintenance features

- Remote software downloads and in-service, hitless activation
- Loopback
- Database backup and restore
- Lamp test

Timing and synchronization

- Two external timing-source inputs (SONET, T1 and SDH E-1, 2 MHz)
- Line timing
- Two timing-source outputs (recovered from line optics)
- Internal Stratum 3 holdover
- Synchronous status-messaging support

Additional features

- 100-Mbps user data channel (Fast Ethernet) transported on the optical supervisory channel (OSC)
- Front only (ETSI) or front and rear access (ANSI) shelf assembly options
- A and B monitored DC power inputs

Compliance and certifications

- Network Equipment Building Standards (NEBS) Level 3 compliance
- Operations Systems Modification of Intelligent Network Elements (OSMINE) certifications
- Storage-vendor qualification and certifications
- ITU-T and CE Mark compliance
- MEF 9 and MEF 14 certification for GE and 10 GE Xponder units

System architecture

This section includes the following information:

- "Wavelength interfaces" on page 1229
- "Transmission elements" on page 1230
- "Distance extension cards" on page 1231

Wavelength interfaces

Table 66 provides a list of services supported by the different line cards on the ONS 15454 MSTP along with the different transmission elements that make up the ONS 15454 MSTP.

Table 66 Protocol support per service card (1 of 2)

Modules	Supported service interfaces
2.5-Gbps FEC Multirate Transponder	1-Gbps Fibre Channel/FICON 2-Gbps Fibre Channel/FICON ISC-1 ISC-3 ESCON Fast Ethernet (FE) Gigabit Ethernet (GE) T3 (Optical) OC-3/STM-1 OC-12/STM-4 OC-48/STM-16 D1-SDI Video HDTV C-Cor DV-6000 (2.38-Gbps) ETR/CLO
2.5-Gbps Data Muxponder	1-Gbps Fibre Channel/FICON 2-Gbps Fibre Channel/FICON ESCON GE
10-Gbps EFEC Multirate Transponder	10 GE LAN 10 GE WAN OC-192/STM-64 10 Gigabit Fibre Channel
4x 2.5-Gbps/10-Gbps EFEC Muxponder	OC-48/STM-16
10-Gbps EFEC Data Muxponder	1-Gbps Fibre Channel/FICON 2-Gbps Fibre Channel/FICON 4-Gbps Fibre Channel/FICON ISC-1 ISC-3 GE

Table 66 Protocol support per service card (2 of 2)

Modules	Supported service interfaces
GE XPonder	GE
10GE XPonder	10GE LAN
MSPP-On-A-Blade	OC-3 OC-12 OC-48 GE

Transmission elements

Table 67 lists the supported modules on ONS 15454 MSTP.

Table 67 Supported modules on ONS 15454 MSTP (1 of 2)

Module	Unit name
Multiplexer and demultiplexer filters	
40-wavelength multiplexer, 100GHz, C band	40-MUX-C
32-wavelength multiplexer, 100-GHz, C band	32MUX-O
32-wavelength demultiplexer, 100-GHz, C band	32DMX-O
4-wavelength multiplexer/demultiplexer, 100-GHz, C band	4MD-xx.x
Optical amplifier	
Preamplifier, 50-GHz capable, C band	OPT-PRE
Booster amplifier, 50-GHz capable, C band	OPT-BST
Enhanced Booster amplifier, 50-GHz capable, C band	OPT-BST-E
Optical amplifier, 17dB gain, 50-GHz capable, C band	OPT-AMP-17C
Amplifier (can be used as Preamplifier or Booster), L band	OPT-AMP-L
Booster amplifier, 50-GHz capable, L band	OPT-BST-L
Reconfigurable optical add/drop multiplexer	
40-channel wavelength cross connect, 100-GHz, Odd, C band	40-WXC-C
Degree-4 Mesh patch panel	PP-MESH-4
Degree-8 Mesh patch panel	PP-MESH-8
40-channel wavelength selective switch, 100-GHz, Odd, C band	40-WSS-C
40-channel demultiplexer, 100-GHz, Odd, C band	40-DMX-C
40-channel wavelength selective switch, 100-GHz, Even, C band	40-WSS-CE
40-channel demultiplexer, 100-GHz, Even, C band	40-DMX-CE
32-channel wavelength selective switch, 100-GHz, C band	32-WSS
32-channel demultiplexer 100-GHz (for use with 32-WSS), C band	32-DMX
32-channel wavelength selective switch, 100-GHz, L band	32-WSS-L
32-channel demultiplexer 100-GHz (for use with 32-WSS), L band	32-DMX-L

Table 67 Supported modules on ONS 15454 MSTP (2 of 2)

Module	Unit name
Multiring/mesh upgrade unit, C band and L band	MMU
Optical band add/drop multiplexer 1-band, 50-GHz capable, C band 4-band, 50-GHz capable, C band	 AD-1B-xx.x AD-4B-xx.x
Optical channel add/drop multiplexer 1-channel, 100-GHz, C band 2-channel, 100-GHz, C band 4-channel, 100-GHz, C band	 AD-1C-xx.x AD-2C-xx.x AD-4C-xx.x
Optical service channel Standard Integrated combiner and separator	 OSCM OSC-CSM
Dispersion compensation Dispersion-compensation unit shelf assembly (2-slot) Dispersion-compensation units	 DCU-SA DCU-<value>
Y-cable protection modules Shelf assembly Y-cable protection module, single-mode Y-cable protection module, multimode	 YCBL-LC or FL-SA YCM-SM-LC or CS-SM-Y YCM-MM-LC or CS-MM-Y

Distance extension cards

The ONS 15454 MSTP has three cards that focus primarily on the data center for distance extension, each discussed further in this section:

- "2.5 Gb/s data muxponder," next
- "8-Port enhanced data muxponder" on page 1235
- "10 Gb/s multi-rate enhanced transponder" on page 1239

2.5 Gb/s data muxponder

The Cisco ONS 15454 2.5 Gb/s muxponder, shown in Figure 480 on page 1225, can transport up to 2 x 1-Gbps services, 1 x 2-Gbps services, 8 x 200 Mb/s services, or a mix of these services aggregated over a G.709 OTU-2-based, 50-GHz spaced, 50-GHz stabilized, ITU-compliant wavelength. The muxponder card architecture contains eight client interfaces that are mapped to a single line interface.

Figure 481 **2.5 Gb/s Data muxponder**

Each client interface supports Gigabit Ethernet, 1-Gbps or 2-Gbps Fibre Channel/FICON/, or ESCON using a Small Form-Factor Pluggable (SFP) optics module with LC connectors, providing the flexibility to support several protocols, data rates, or fiber types with support for qualified SFP modules (see *Cisco ONS 15454 MSTP SFP Support Matrix*). The 2.5 G data muxponder card supports a mixture of SFP types and also supports in-service insertion or removal without affecting other active ports, allowing networking flexibility and reduced pre-planning activities. The DWDM line interface provides one 2.5-Gbps G.709 OTU-2 digital wrapper, long-reach/long-haul, ITU-compliant, 50-GHz spaced optical interface using LC connectors supporting G.709 OTU-2 digital wrapper interfaces. The DWDM output line interface is tunable across 4 wavelengths across the C band. The muxponder cards are deployable in the 12 multiservice interface card slots of the Cisco ONS 15454 MSTP. The only other common card required for operation is the timing, communications, and control (TCC) card. The data muxponder card provides many carrier-class features and capabilities necessary to deliver SAN and data services, including selectable service mix, wavelength tunability, flexible protection mechanisms, flexible timing options, and management capabilities.

Figure 482 shows the data muxponder both protected and unprotected.

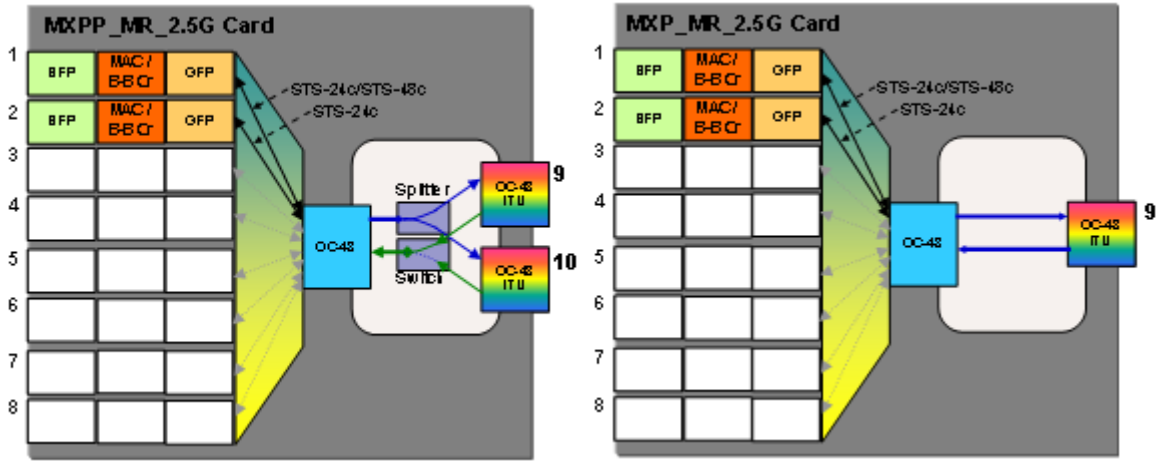


Figure 482 2.5 Gb/s data muxponder (protected and unprotected)

2.5 Gb/s data muxponder optical parameters

- Transmitter (trunk side)
  - Minimum output power: -1 dBm
  - Maximum output power: +3 dBm
  - 50-GHz wave locked
  - Four-channel wavelength tunability at 100-GHz spacing
  - Transmitter maximum return reflectance: -27 dB
  - Chromatic dispersion allowance: 5400 ps/nm, giving an optical power penalty < 2.0 dB
  - Minimum side mode suppression ratio: 30 dB
  - Transmitter is a direct modulated laser
  - Wavelength stability (drift): +/- 25 picometers (pm)
- Receiver (trunk side)

  Table 68 on page 1234 lists the data muxponder receiver trunk side specifications.

Table 68 Data muxponder receiver trunk side specifications

OSNR[a]	FEC Type	Pre-FEC BER	Post-FEC BER	Input power sensitivity	Chromatic dispersion tolerance
17 dB	N/A	< 10 exp - 12	N/A	- 9 to - 23 dBm	—
17 dB	N/A	< 10 exp - 12	N/A	- 9 to - 22 dBm	+/- 1800 ps/nm
17 dB	N/A	< 10 exp - 12	N/A	- 9 to - 21 dBm	+/- 5400 ps/nm
18 dB	N/A	< 10 exp - 12	N/A	- 9 to - 23 dBm	+/- 1800 ps/nm
19 dB	N/A	< 10 exp - 12	N/A	- 9 to - 23 dBm	+/- 5400 ps/nm
21 dB	N/A	< 10 exp - 12	N/A	- 9 to - 30 dBm	—
21 dB	N/A	< 10 exp - 12	N/A	- 9 to - 29 dBm	+/- 1800 ps/nm
21 dB	N/A	< 10 exp - 12	N/A	- 9 to - 28 dBm	+/- 5400 ps/nm
22 dB	N/A	< 10 exp - 12	N/A	- 9 to - 30 dBm	+/- 1800 ps/nm
23 dB	N/A	< 10 exp - 12	N/A	- 9 to - 30 dBm	+/- 5400 ps/nm

a. OSNR defined with 0.1 nm RBW

- Receiver sensitivity -28 dBm, BER 1 * 10 exp - 12
- Receiver overload is equal to or exceeds -8 dBm
- Receiver maximum reflectance of -27 dB

◆ Transmitter (client side)

- Maximum transmitter output power: -1 dBm
- Minimum transmitter output power: -6 dBm
- Center wavelength: 1290 to 1330 nm
- Nominal wavelength: 1310 nm
- Transmitter: DFB laser

◆ Receiver (client side)

- Maximum receiver level: -1 dBm at BER 1 * 10 exp - 12
- Minimum receiver level: -14 dBm at BER 1 * 10 exp - 12
- Receiver: APD
- Link loss budget: 8 dB minimum, at BER = 1 * 10 exp - 12
- Receiver input wavelength range: 1290 to 1605 nm

8-Port enhanced data muxponder

The Cisco ONS 15454 8-Port 10-Gbps data muxponder card can transport up to 8 x 1-Gbps services, 4 x 2-Gbps services, 2 x 4-Gbps services, or a mix of these services aggregated over a G.709 OTU-2-based, 50-GHz spaced, 50-GHz stabilized, ITU-compliant wavelength with selectable Forward Error Correction (FEC) or Enhanced Forward Error Correction (E-FEC).

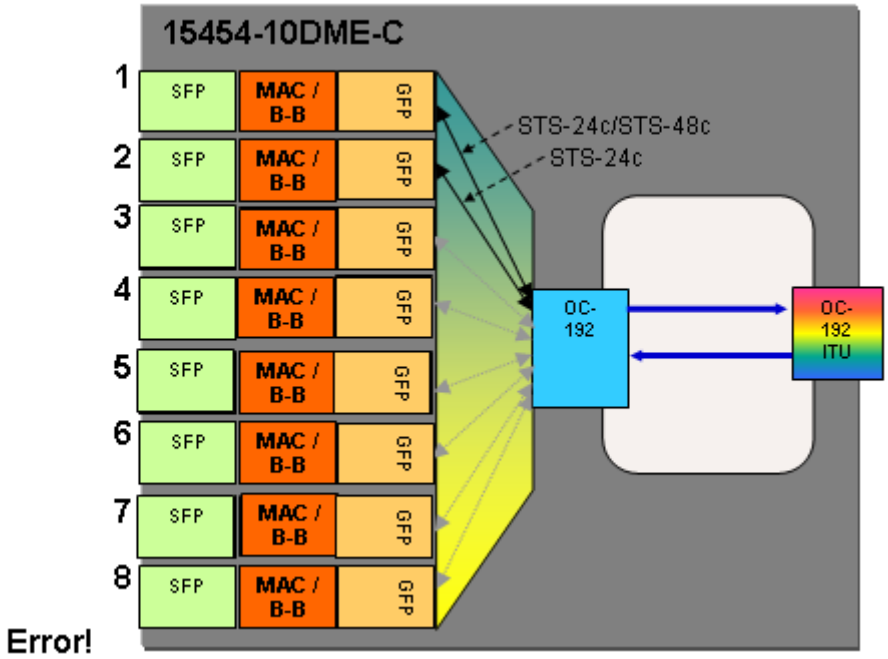
Figure 483 **8-Port enhanced data muxponder**

As shown in Figure 483, the muxponder card architecture contains eight client interfaces that are mapped to a single line interface. Each client interface supports Gigabit Ethernet, 1-Gbps or 2-Gbps Fibre Channel/FICON/ISC, or 4-Gbps Fibre Channel using a Small Form-Factor Pluggable (SFP) optics module with LC connectors, providing the flexibility to support several protocols, data rates, or fiber types with support for qualified SFP modules (see the *Cisco ONS 15454 MSTP SFP Support Matrix*).

The data muxponder card supports a mixture of SFP types and also supports in-service insertion or removal without affecting other active ports, allowing networking flexibility and reduced pre-planning activities. The DWDM line interface provides one 10.70923-Gbps G.709 OTU-2 digital wrapper, long-reach/long-haul, ITU-compliant, 50-GHz spaced optical interface using LC connectors supporting G.709 OTU-2 digital wrapper interfaces. The DWDM output line interface is tunable across the full optical C or L band,

dramatically reducing the inventory of spares. When operated within the outlined specifications, each card will transport the aggregated signals with a maximum bit error rate (BER) of 10E-15. The muxponder cards are deployable in the 12 multiservice interface card slots of the Cisco ONS 15454 MSTP. The only other common card required for operation is the timing, communications, and control (TCC) card. The data muxponder card provides many carrier-class features and capabilities necessary to deliver SAN and data services, including selectable service mix, wavelength tunability, flexible protection mechanisms, flexible timing options, and management capabilities.

Figure 484 shows an unprotected 8-port data muxponder.



Error!

Figure 484 8-Port data muxponder (unprotected)

8-Port data muxponder optical parameters

- Transmitter (trunk side)
  - Minimum output power: +3 dBm
  - Maximum output power: +6 dBm
  - Minimum Single-Mode Suppression Ratio (SMSR): 30 dB

- Minimum optical extinction ratio: 10 dB
- 82 wavelength tunability at 50-GHz spacing
- Receiver maximum return reflectance (Rx return loss): -27 dB
- Chromatic dispersion allowance: 1200 ps / nm, giving an optical power penalty < 2.0 dB
- Minimum side mode suppression ratio: 30 dB
- Wavelength stability (drift): + / - 25 picometers (pm)

◆ Receiver (trunk side)

Table 69 lists the data muxponder receiver trunk side specifications.

Table 69 8 Gb/s data muxponder receiver trunk side specifications

FEC applications	OSNR[a]	Pre-FEC BER	Post-FEC BER	Input power sensitivity	Chromatic dispersion tolerance	Power penalty	OSNR penalty
None	23 dB	<10 exp - 12	—	-8 to -20 dBm	+/- 1200 ps/nm	2 dBm	—
	19 dB	<10 exp - 12	—	-9 to -22 dBm	+/- 1000 ps/nm	2 dBm	—
FEC	10 dB	<10 exp - 5	<10 exp - 15	-8 to -18 dBm	+/- 800 ps/nm		1.5 dB
Enhanced FEC	19 dB	<10 exp - 4	<10 exp - 15	-8 to -26 dBm	+/- 800 ps/nm	2 dBm	2 dB
	8 dB	<10 exp - 4	<10 exp - 15	-8 to -18 dBm	+/- 800 ps/nm	2 dBm	1.5 dB

a. OSNR defined with 0.5 nm RBW

- Receiver: APD
- Link loss budget: 24 dB minimum, with no dispersion or 22 dB optical path loss at BER = 1 * 10 exp - 12 including dispersion
- Receiver input wavelength range: 1529 to 1562 nm

◆ Transmitter (client side)

- Maximum transmitter output power: -1 dBm
- Minimum transmitter output power: -6 dBm
- Center wavelength: 1290 to 1330 nm
- Nominal wavelength: 1310 nm
- Transmitter: DFB laser

◆ Receiver (client side)

- Maximum receiver level: -1 dBm at BER 1 * 10 exp - 12
- Minimum receiver level: -14 dBm at BER 1 * 10 exp - 12
- Receiver: APD

- Link loss budget: 8 dB minimum, at BER = 1 * 10 exp - 12
- Receiver input wavelength range: 1290 to 1605 nm

Table 70 lists the buffer credits supported per port on the 8-port enhanced data muxponder.

Table 70 Buffer credits supported per port on 8-port enhanced data muxponder

Fibre Channel protocol speed	Distance (KM)	Buffer credit per port
1 G FC	1400	700 B2B
2 G FC	700	700 B2B
4 G FC	500	1000 B2B

Note: There is no concept of total B2B on the card. This is determined by the available free memory on the card. When a port on the 8-port enhanced data muxponder is configured for 4 G FC, more memory is allocated to B2B.

Figure 485 shows the 9-port buffer credit spoofing flow chart.

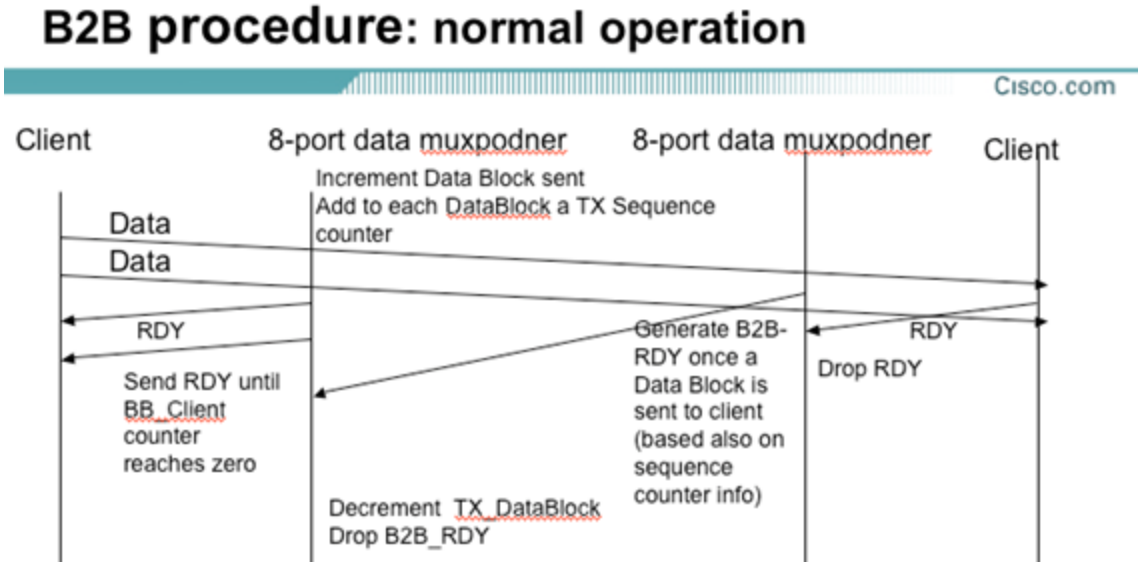


Figure 485 8-Port buffer credit spoofing flow chart

10 Gb/s multi-rate enhanced transponder

The 10-Gb/s multi-rate transponder card can transport 10 Gigabit Ethernet WAN physical layer (PHY) and LAN PHY, 10 Gigabit Fibre Channel, SONET OC-192, and SDH STM-64 services over a 100-GHz spaced, 50-GHz stabilized, ITU-compliant wavelength. The transponder card architecture contains a single client interface that is mapped to a single line interface, as shown in Figure 486.

Figure 486 **10 Gb/s Multi-rate enhanced transponder**

The client interface supports 10 Gigabit Ethernet LAN PHY, 10 Gigabit Ethernet WAN PHY, 10 Gigabit Fibre Channel, SONET OC-192, and SDH STM-64 signals. The interface is based on 10 Gigabit Small Form-Factor Pluggable (XFP) Multisource Agreement (MSA). Currently available XFP optics support a short-reach/intra-office, 1310-nanometer (nm) optical interface using LC connectors supporting fiber distances of up to 2 kilometers (km) (with or without the Y-protection option). The line interface provides one 10-Gbps, long-reach, ITU-compliant, 100-GHz-spaced optical interface using LC connectors supporting OUT-2 G.709 digital wrapper, OC-192, STM-64, 10 Gigabit Fibre Channel, 10 Gigabit Ethernet LAN PHY, or 10 Gigabit Ethernet WAN PHY interfaces. The DWDM line interface provides one 10.70923-Gbps G.709 OTU-2 digital wrapper, long-reach/long-haul, ITU-compliant, 50-GHz spaced optical interface using LC connectors supporting G.709

OTU-2 digital wrapper interfaces. The DWDM output line interface is tunable across the full optical C or L band, dramatically reducing the inventory of spares. When operated within the outlined specifications, each card will transport the aggregated signals with a maximum bit error rate (BER) of 10E-15. The 10 Gbps transponder card incorporates both a client and DWDM line interface on the same card. The 10 Gbps transponder cards are deployable in the 12 multiservice interface card slots of the Cisco ONS 15454 platform. The only required common card is the appropriate timing, communications, and control card (TCC). The 10 Gbps transponder card provides many carrier-class features and advanced capabilities necessary to deliver 10 Gbps services, including the protocol transparency, wavelength tunability, flexible protection mechanisms, flow-through timing, management, and performance monitoring capabilities outlined in this section.

Figure 487 shows a block diagram of the 10 G multi-rate enhanced transponder.

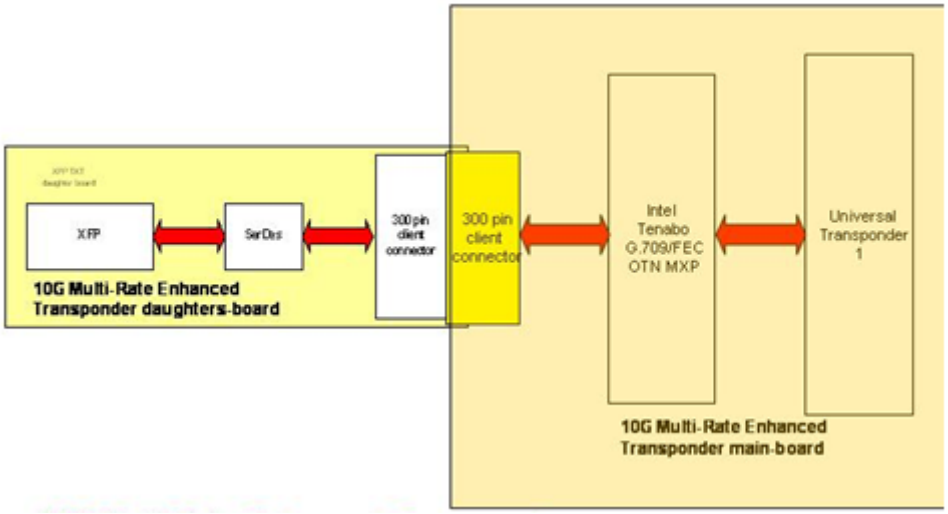


Figure 487 10 Gb/s Multi-rate enhanced transponder block daigram

10 Gb/s multi-rate enhanced transponder optical parameters

- Transmitter (trunk side)
  - Minimum output power: +3 dBm
  - Maximum output power: +6 dBm
  - Transmitter: LN external modulator transmitter

- Wavelength stability (drift): +/- 25 picometers (pm)

◆ Receiver (trunk side)

Table 71 lists the 10 Gb/s multi-rate enhanced transponder receiver trunk side specifications.

Table 71 10 Gb/s multi-rate enhanced transponder receiver trunk side specifications

OSNR[a]	FEC type	Pre-FEC BER	Post-FEC BER	Input power sensitivity[b]	Chromatic dispersion tolerance
30 dB	Off	<10 exp - 12	N/A	-8 to -20 dBm	+/- 1200 ps/nm
26 dB	Off	<10 exp - 12	N/A	-8 to -20 dBm	+/- 1000 ps/nm
26 dB	Off	<10 exp - 12	N/A	-8 to -22 dBm	—
17 dB	Standard	<10 exp - 5	<10 exp - 15	-8 to -18 dBm	+/- 800 ps/nm
15.5 dB	Standard	<10 exp - 5	<10 exp - 15	-8 to -18 dBm	—
14 dB	Enhanced	<7 X 10 exp - 4	<10 exp - 15	-8 to -18 dBm	+/- 800 ps/nm
12 dB	Enhanced	<7 X 10 exp - 4	<10 exp - 15	-8 to -18 dBm	—

a. OSNR defined with 0.1 nm RBW

b. Receiver filter bandwidth greater than or equal to 180 pm (at -3 dBm)

- Receiver: APD
- Link loss budget: 24 dB minimum, with no dispersion or 22 dB optical path loss at BER = 1 * 10 exp - 12 including dispersion
- Receiver input wavelength range: 1529 to 1562 nm

◆ Transmitter (client side)

- Maximum transmitter output power: -1 dBm
- Minimum transmitter output power: -6 dBm
- Center wavelength: 1290 to 1330 nm
- Nominal wavelength: 1310 nm
- Transmitter: DFB laser

◆ Receiver (client side)

- Maximum receiver level: -1 dBm at BER 1 * 10 exp - 12
- Minimum receiver level: -14 dBm at BER 1 * 10 exp - 12
- Receiver: APD
- Link loss budget: 8 dB minimum, at BER = 1 * 10 exp - 12
- Receiver input wavelength range: 1290 to 1605 nm

MSTP SFP support

This section contains tables listing SFP support. For technical details, refer to "SFP technical details" on page 1253.

Table 72 shows the ONS 15454 MSTP SFP matrix.

Table 72 ONS15454 MSTP SFP matrix (1 of 2)

ONS 15454 MSTP Boards									
Product ID	**100M-2.5G MR-TXP**	**2.5G MR DataMux**	**4x2,5G FEC MXP**	**4x2.5G EFEC TXP**	**10G EFEC TXP**	**10G MR DataMux**	**ADM-10G**	**GE-XP**	**10GE-XP**
15454-SFP3-1-IR=	X								
ONS-SI-155-SR-MM=	X								
ONS-SI-155-L2=							X		
15454-SFP12-4-IR=	X								
ONS-SI-622-I1=							X		
15454-SFP-OC48-IR=	X		X	X					
ONS-SE-2G-S1=	X		X	X			X		
ONS-SE-2G-L2=	X			X			X		
ONS-SI-2G-S1							X		
ONS-SI-2G-I1=							X		
ONS-SI-2G-L2=							X		
ONS-SE-Z1=							X		
ONS-SE-ZE-EL=						X		X	
15454-SFP-GE+-LX=	X	X				X			
15454-SFP-GEFC-SX=	X	X				X			
ONS-SE-G2F-SX=	X	X				X	X	X	
ONS-SE-G2F-LX=	X	X				X	X	X	
ONS-SE-GE-ZX=	X	X							
ONS-SI-GE-ZX=						X[a]	X	X	

Table 72 ONS15454 MSTP SFP matrix (2 of 2)

ONS 15454 MSTP Boards									
Product ID	**100M-2.5G MR-TXP**	**2.5G MR DataMux**	**4x2,5G FEC MXP**	**4x2.5G EFEC TXP**	**10G EFEC TXP**	**10G MR DataMux**	**ADM-10G**	**GE-XP**	**10GE-XP**
15454-SFP-200=	X								
ONS-SE-200-MM=	X	X							
ONS-SE-4G-MM=						X			
ONS-SE-4G-SM=						X			
ONS-SC-2G-xx.x=[b]	X			X			X	X	
ONS-SC-Z3-xxxx=[c]	X	X		X			X	X	
ONS-XC-10G-S1=					X		X	X	X
ONS-XC-10G-I2=[d]					X			X	X
ONS-XC-10G-L2=					X [e]				
ONS-XC-10G-xx.x=[f]							X	X	X

a. Supported with SW release 7.0.x but not in 8.0. Support in Release 8.5.

b. DWDM SFPs, xx.x from 30.3 to 60.6. Supporting GE, 1G FC, 2G FC and OC48 pending board rate support. Only P/N version 02 is qualified on 15454 MSTP.

c. CWDM SFPs, xxxx from 1470 to 1610 . Supporting GE, 1G FC, 2G FC and OC48 pending board rate support.

d. Only P/N version 02 is qualified on 15454 MSTP.

e. ONS-XC-10G-L2. 10G TXP when equipped with LR2 XFP need to be placed on High Speed slot for power dissipation constraint if using FTA-3 or FTA-48V. If it is used CC-FTA there is no restirction.

f. DWDM XFPs, xx.x from 30.3 to 61.4.

SONET/SDH SFPs SFPs fully compliant with SONET and SDH standards. Details are shown in Table 73.

Table 73 SONET/SDH SFPs (1 of 2)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
15454-SFP3-1-IR=	SFP - OC-3/D1-SDI - 1310nm IR - SM - LC	10-1828-01	GR253 OC3 IR1 G.957 S-1.1	0 to +70
15454-SFP12-4-IR=	SFP - OC-12 - 1310nm IR - SM - LC	10-1976-01	GR253 OC12 IR1 G.957 S-4.1	0 to +70
15454-SFP-OC48-IR=	SFP - OC-48- 1310nm IR - SM - LC	10-1975-01	GR253 OC48 IR1 G.957 S-16.1	0 to +70
ONS-SE-2G-S1=	SFP - OC48/STM16 - 1310 SR - SM LC	10-2017-01	GR253 OC48 SR G.957 I-16	-10 to +85
ONS-SE-2G-L2=	OC48/STM16, LR2, 1550nm, Small Form Pluggable (SFP), EXT	10-2013-01	GR253 OC48 LR2 G.957 L-16.2	-10 to +85
ONS-SI-155-SR-MM=	SFP - OC3, SR, 1310 NM, MULTI MODE, I-TEMP	10-2279-01	GR253	-40 to +85
ONS-SI-155-I1=	SFP - OC3/STM1 IR1/S-1.1 1310 SFP, ITEMP	10-1938-02	GR253 OC3 IR1 G.957 S-1.1	-40 to +85
ONS-SI-155-L1=	SFP -OC3/STM1 LR, L-1.1, 1310 nm, , ITEMP	10-1957-02	GR253 OC3 LR1 G.957 L-1.1	-40 to +85
ONS-SI-155-L2=	SFP -OC3/STM1 LR-2, L-1.2, 1550 nm, ITEMP	10-1937-02	GR253 OC3 LR2 G.957 L-1.2	-40 to +85
ONS-SI-622-SR-MM=	SFP - OC12, SR, 1310 NM, MULTI MODE, I-TEMP	10-2280-01	GR253	-40 to +85
ONS-SI-622-I1=	SFP -OC12/STM4 and OC3/STM1 IR, S-4.1, S-1.1, 1310 nm, ITEMP	10-1956-02	GR253 OC3/OC12 IR1 G.957 S-4.1/S-1.1	-40 to +85
ONS-SI-622-L1=	SFP -OC12/STM4 LR, L-4.1, 1310 nm, ITEMP	10-1958-02	GR253 OC12 LR1 G.957 L-4.1	-40 to +85

Table 73 SONET/SDH SFPs (2 of 2)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SI-622-L2=	SFP -OC12/STM4 LR, L-4.2, 1550 nm, ITEMP	10-1936-02	GR253 OC12 LR2 G.957 L-4.2	-40 to +85
ONS-SI-2G-S1	SFP - OC48/STM16, SR, 1310nm, ITEMP, LC	10-1992-02	GR253 OC48 SR G.957 I-16	-40 to +85
ONS-SI-2G-I1=	SFP - OC48/STM16, IR, 1310nm, ITEMP, LC	10-1993-02	GR253 OC48 IR1 G.957 S-16.1	-40 to +85
ONS-SI-2G-L1=	SFP - OC48/STM16, LR1, 1310nm, ITEMP, LC	10-2102-02	GR253 OC48 LR1 G.957 L-16.1	-40 to +85
ONS-SI-2G-L2=	SFP - OC48/STM16, LR2, 1550nm, ITEMP, LC	10-1990-02	GR253 OC48 LR2 G.957 L-16.2	-40 to +85

Data SFPs DATA SFPs able to transmit GbE, FC, FICON, and ESCON signal format. Details are shown in Table 74.

Table 74 Data SFPs (1 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SE-100-LX10=	SFP - 100Mbps Long Reach - 1310nm - SM - LC, EXT-TEMP	10-2213-01	100Base LX IEEE-802.3	-10 to +85
ONS-SI-100-LX10=	SFP - 100Mbps Long Reach - 1310nm - SM - LC, ITEMP	10-2294-01	100Base LX IEEE-802.3	-40 to +85
ONS-SE-100-FX=	SFP - 100Mbps Short Reach - 1310nm - MM - LC, EXT-TEMP	10-2212-01	100Base FX IEEE-802.3	-10 to +85
ONS-SI-100-FX=	SFP - 100Mbps Short Reach - 1310nm - MM - LC, ITEMP	10-2350-01	100Base FX IEEE-802.3	-40 to +85
ONS-SE-100-BX10U=	SFP -10/100 BX-U, EXT	10-2353-01	100Base BX-U IEEE-802.3	-10 to +85
ONS-SE-100-BX10D=	SFP -10/100 BX-D, EXT	10-2352-01	100Base BX-D IEEE-802.3	-10 to +85

Table 74 Data SFPs (2 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
15327-SFP-LC-SX=	1000 Base SX LC, SFP	30-1301-01	1000Base SX IEEE-802.3	0 to +70
15327-SFP-LC-LX=	1000 Base LX LC, SFP	30-1299-01	1000Base SX IEEE-802.3	0 to +70
15454-SFP-LC-SX=	1000 Base SX LC, SFP	30-1301-01	1000Base SX IEEE-802.3	0 to +70
15454-SFP-LC-LX=	1000 Base LX LC, SFP	30-1299-01	1000Base LX IEEE-802.3	0 to +70
ONS-SC-GE-SX=[a]	1000Base SX SFP - 850nm - LC - C Temp	10-2301-01	1000Base SX IEEE-802.3	0 to +70
ONS-SI-GE-SX=	SFP - 1000BASE-SX Gigabit Ethernet, 850nm, MM, I-TEMP	10-2295-01	1000Base SX IEEE-802.3	-40 to +85
ONS-SC-GE-LX=[b]	1000Base LX SFP - 1310nm - LC - C Temp	10-2298-01	1000Base LX IEEE-802.3	0 to +70
ONS-SI-GE-LX=	SFP - 1000BASE-LX Gigabit Ethernet, 1310, SM, I-TEMP	10-2300-01	1000Base SX IEEE-802.3	-40 to +85
15454-SFP-GE+-LX= EOS	SFP-GE/1G-FC/2G-FC-1310nm-MM-LC	10-1832-03	1000Base LX IEEE-802.3, 100-M5-SN-I 200-M5-SN-I	-10 to +85
15454-SFP-GEFC-SX= EOS	SFP-GE/1G-FC/2G-FC-850nm-MM-LC	10-1833-02	1000Base SX IEEE-802.3, 100-M5-SN-I 100-M6-SN-I 200-M5-SN-I 200-M6-SN-I	-10 to +85
ONS-SE-G2F-SX=	SFP - GE/1G-FC/2G-FC - 850nm - MM – LC - EXT TEMP	10-2272-01	1000Base SX IEEE-802.3, 100-M5-SN-I 100-M6-SN-I 200-M5-SN-I 200-M6-SN-I	-10 to +85

Table 74 Data SFPs (3 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SE-G2F-LX=	SFP - GE/1G-FC/2G-FC/HDTV - 1310nm - SM – LC - EXT TEMP	10-2273-02	1000Base LX IEEE-802.3, 100-SM-LC-L 200-SM-LC-L	-10 to +85
ONS-SI-GE-ZX=	SFP - 1000BASE-ZX Gigabit Ethernet, 1550, SM, I-Temp	10-2296-01	1000Base ZX IEEE-802.3	-40 to +85
ONS-SE-GE-ZX=	SFP - 1000BASE-ZX Gigabit Ethernet, 1550, SM, Ext-Temp	10-2354-01	1000Base ZX IEEE-802.3	-10 to +85
15454-SFP-200=	SFP-ESCON - 1310nm - MM - LC	10-1750-01	ESCON	0 to +70
ONS-SE-200-MM=	SFP-ESCON - 1310nm - MM – LC - EXT TEMP	10-2248-01	ESCON	-10 to +85
ONS-SE-4G-MM=	4G FC SFP , 850nm, LC, MM - EXT TEMP	10-2259-01	400-M5-SN-I and 400-M6-SN-I	-10 to +85
ONS-SE-4G-SM=	4G FC SFP , 1310nm, LC, SM - EXT TEMP	10-2252-01	400-SM-LC-L	-10 to +85

a. This SFP will replace the old 15327-SFP-LC-SX= and 15454-SFP-LC-SX= that have initiated EoL process.

b. This SFP will replace the old 15327-SFP-LC-LX= and 15454-SFP-LC-LX= that have initiated EoL process

DWDM SFPs DWDM ITU-T compliant SFPs. Details are shown in Table 75.

Table 75 DWDM SFPs (1 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SC-2G-28.7=	OC-48/STM16, SFP, 1528.77, 100 GHz, LC	10-2307-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-30.3=	OC-48/STM16, SFP, 1530.33, 100 GHz, LC	10-2155-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-31.1=	OC-48/STM16, SFP, 1531.12, 100 GHz, LC	10-2156-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-31.9=	OC-48/STM16, SFP, 1531.90, 100 GHz, LC	10-2157-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-32.6=	OC-48/STM16, SFP, 1532.68, 100 GHz, LC	10-2158-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-33.4=	OC-48/STM16, SFP, 1533.47, 100 GHz, LC	10-2306-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-34.2=	OC-48/STM16, SFP, 1534.25, 100 GHz, LC	10-2159-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-35.0=	OC-48/STM16, SFP, 1535.04, 100 GHz, LC	10-2160-02	ITU G694, GR2918	0 to +70

Table 75 DWDM SFPs (2 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SC-2G-35.8=	OC-48/STM16, SFP, 1535.82, 100 GHz, LC	10-2161-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-36.6=	OC-48/STM16, SFP, 1536.61, 100 GHz, LC	10-2162-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-38.1=	OC-48/STM16, SFP, 1538.19, 100 GHz, LC	10-2163-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-38.9=	OC-48/STM16, SFP, 1538.98, 100 GHz, LC	10-2164-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-39.7=	OC-48/STM16, SFP, 1539.77, 100 GHz, LC	10-2165-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-40.5=	OC-48/STM16, SFP, 1540.56, 100 GHz, LC	10-2185-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-41.3=	OC-48/STM16, SFP, 1541.35, 100 GHz, LC	10-2305-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-42.1=	OC-48/STM16, SFP, 1542.14, 100 GHz, LC	10-2166-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-42.9=	OC-48/STM16, SFP, 1542.94, 100 GHz, LC	10-2167-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-43.7=	OC-48/STM16, SFP, 1543.73, 100 GHz, LC	10-2168-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-44.5=	OC-48/STM16, SFP, 1544.53, 100 GHz, LC	10-2169-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-46.1=	OC-48/STM16, SFP, 1546.12, 100 GHz, LC	10-2170-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-46.9=	OC-48/STM16, SFP, 1546.92, 100 GHz, LC	10-2171-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-47.7=	OC-48/STM16, SFP, 1547.72, 100 GHz, LC	10-2172-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-48.5=	OC-48/STM16, SFP, 1548.51, 100 GHz, LC	10-2173-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-49.3=	OC-48/STM16, SFP, 1549.32, 100 GHz, LC	10-2304-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-50.1=	OC-48/STM16, SFP, 1550.12, 100 GHz, LC	10-2186-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-50.9=	OC-48/STM16, SFP, 1550.92, 100 GHz, LC	10-2174-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-51.7=	OC-48/STM16, SFP, 1551.72, 100 GHz, LC	10-2175-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-52.5=	OC-48/STM16, SFP, 1552.52, 100 GHz, LC	10-2176-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-54.1=	OC-48/STM16, SFP, 1554.13, 100 GHz, LC	10-2177-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-54.9=	OC-48/STM16, SFP, 1554.94, 100 GHz, LC	10-2178-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-55.7=	OC-48/STM16, SFP, 1555.75, 100 GHz, LC	10-2179-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-56.5=	OC-48/STM16, SFP, 1556.55, 100 GHz, LC	10-2180-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-57.3=	OC-48/STM16, SFP, 1557.36, 100 GHz, LC	10-2308-02	ITU G694, GR2918	0 to +70

Table 75 DWDM SFPs (3 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SC-2G-58.1=	OC-48/STM16, SFP, 1558.17, 100 GHz, LC	10-2181-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-58.9=	OC-48/STM16, SFP, 1558.98, 100 GHz, LC	10-2182-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-59.7=	OC-48/STM16, SFP, 1559.79, 100 GHz, LC	10-2183-02	ITU G694, GR2918	0 to +70
ONS-SC-2G-60.6=	OC-48/STM16, SFP, 1560.61, 100 GHz, LC	10-2184-02	ITU G694, GR2918	0 to +70

CWDM SFPs CWDM ITU-T compliant SFPs. Details are shown in Table 76.

Table 76 CWDM SFPs (1 of 2)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SE-155-1470=	SFP - OC3/STM1, CWDM, 1470 nm, EXT	10-1996-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1490=	SFP - OC3/STM1, CWDM, 1490 nm, EXT	10-1998-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1510=	SFP - OC3/STM1, CWDM, 1510 nm, EXT	10-1999-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1530=	SFP - OC3/STM1, CWDM, 1530 nm, EXT	10-2000-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1550=	SFP - OC3/STM1, CWDM, 1550 nm, EXT	10-2001-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1570=	SFP - OC3/STM1, CWDM, 1570 nm, EXT	10-2002-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1590=	SFP - OC3/STM1, CWDM, 1590 nm, EXT	10-2003-01	ITU-T G.694.2	-10 to +85
ONS-SE-155-1610=	SFP - OC3/STM1, CWDM, 1610 nm, EXT	10-1997-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1470=	SFP - OC12/STM4, CWDM, 1470 nm, EXT	10-2004-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1490=	SFP - OC12/STM4, CWDM, 1490 nm, EXT	10-2005-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1510=	SFP - OC12/STM4, CWDM, 1510 nm, EXT	10-2006-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1530=	SFP - OC12/STM4, CWDM, 1530 nm, EXT	10-2007-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1550=	SFP - OC12/STM4, CWDM, 1550 nm, EXT	10-2008-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1570=	SFP - OC12/STM4, CWDM, 1570 nm, EXT	10-2009-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1590=	SFP - OC12/STM4, CWDM, 1590 nm, EXT	10-2010-01	ITU-T G.694.2	-10 to +85
ONS-SE-622-1610=	SFP - OC12/STM4, CWDM, 1610 nm, EXT	10-2011-01	ITU-T G.694.2	-10 to +85
ONS-SC-Z3-1470=	SFP - OC48/STM16/GE, CWDM, 1470 nm	10-2285-01	ITU-T G.694.2	0 to +70

Table 76 CWDM SFPs (2 of 2)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-SC-Z3-1490=	SFP - OC48/STM16/GE, CWDM, 1490 nm	10-2286-01	ITU-T G.694.2	0 to +70
ONS-SC-Z3-1510=	SFP - OC48/STM16/GE, CWDM, 1510 nm	10-2287-01	ITU-T G.694.2	0 to +70
ONS-SC-Z3-1530=	SFP - OC48/STM16/GE, CWDM, 1530 nm	10-2288-01	ITU-T G.694.2	0 to +70
ONS-SC-Z3-1550=	SFP - OC48/STM16/GE, CWDM, 1550 nm	10-2289-01	ITU-T G.694.2	0 to +70
ONS-SC-Z3-1570=	SFP - OC48/STM16/GE, CWDM, 1570 nm	10-2290-01	ITU-T G.694.2	0 to +70
ONS-SC-Z3-1590=	SFP - OC48/STM16/GE, CWDM, 1590 nm	10-2291-01	ITU-T G.694.2	0 to +70
ONS-SC-Z3-1610=	SFP - OC48/STM16/GE, CWDM, 1610 nm	10-2292-01	ITU-T G.694.2	0 to +70

Grey XFPs Grey XFPs. Details are shown in Table 77.

Table 77 Grey XFPs

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-XC-10G-S1=	XFP - OC192/STM64/10GE - 1310 SR - SM LC	10-2012-02	ITU G694 I-64.1 GR253 SR-1 10GE BASE LR 1200-SM-LL-L	0 to +70
ONS-XC-10G-I2=	XFP - OC192/STM64/10GE - 1550 IR2 - SM LC	10-2193-02	ITU G694 S-64.2b GR253 IR-2	0 to +70
ONS-XC-10G-L2=	XFP - OC192/STM64 - 1550 LR2 - SM LC	10-2194-02	ITU G959.1 P1L1-2D2 GR253 LR-2	0 to +70

DWDM XFPs DWDM XFPs. Details are shown in Table 78.

Table 78 DWDM XFPs (1 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-XC-10G-30.3=	OC-192/STM64/10GE, XFP, 1530.33, 100 GHz, LC	10-2347-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-31.1=	OC-192/STM64/10GE, XFP, 1531.12, 100 GHz, LC	10-2346-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-31.9=	OC-192/STM64/10GE, XFP, 1531.90, 100 GHz, LC	10-2344-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-32.6=	OC-192/STM64/10GE, XFP, 1532.68, 100 GHz, LC	10-2345-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-33.4=	OC-192/STM64/10GE, XFP, 1533.47, 100 GHz, LC	10-2343-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-34.2=	OC-192/STM64/10GE, XFP, 1534.25, 100 GHz, LC	10-2342-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-35.0=	OC-192/STM64/10GE, XFP, 1535.04, 100 GHz, LC	10-2341-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-35.8=	OC-192/STM64/10GE, XFP, 1535.82, 100 GHz, LC	10-2340-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-36.6=	OC-192/STM64/10GE, XFP, 1530.33, 100 GHz, LC	10-2347-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-37.4=	OC-192/STM64/10GE, XFP, 1531.12, 100 GHz, LC	10-2346-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-38.1=	OC-192/STM64/10GE, XFP, 1531.90, 100 GHz, LC	10-2344-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-38.9=	OC-192/STM64/10GE, XFP, 1532.68, 100 GHz, LC	10-2345-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-39.7=	OC-192/STM64/10GE, XFP, 1533.47, 100 GHz, LC	10-2343-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-40.5=	OC-192/STM64/10GE, XFP, 1534.25, 100 GHz, LC	10-2342-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-41.3=	OC-192/STM64/10GE, XFP, 1535.04, 100 GHz, LC	10-2341-01	ITU G694, GR2918	0 to +70

Table 78 DWDM XFPs (2 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-XC-10G-42.1=	OC-192/STM64/10GE, XFP, 1535.82, 100 GHz, LC	10-2340-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-42.9=	OC-192/STM64/10GE, XFP, 1536.61, 100 GHz, LC	10-2339-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-43.7=	OC-192/STM64/10GE, XFP, 1537.40, 100 GHz, LC	10-2338-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-44.5=	OC-192/STM64/10GE, XFP, 1538.19, 100 GHz, LC	10-2337-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-45.3=	OC-192/STM64/10GE, XFP, 1538.98, 100 GHz, LC	10-2336-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-46.1=	OC-192/STM64/10GE, XFP, 1539.77, 100 GHz, LC	10-2335-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-46.9=	OC-192/STM64/10GE, XFP, 1540.56, 100 GHz, LC	10-2348-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-47.7=	OC-192/STM64/10GE, XFP, 1541.35, 100 GHz, LC	10-2334-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-48.5=	OC-192/STM64/10GE, XFP, 1542.14, 100 GHz, LC	10-2333-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-49.3=	OC-192/STM64/10GE, XFP, 1542.94, 100 GHz, LC	10-2332-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-50.1=	OC-192/STM64/10GE, XFP, 1543.73, 100 GHz, LC	10-2331-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-50.9=	OC-192/STM64/10GE, XFP, 1544.53, 100 GHz, LC	10-2330-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-51.7=	OC-192/STM64/10GE, XFP, 1545.32, 100 GHz, LC	10-2329-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-52.5=	OC-192/STM64/10GE, XFP, 1546.12, 100 GHz, LC	10-2328-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-53.3=	OC-192/STM64/10GE, XFP, 1546.92, 100 GHz, LC	10-2327-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-54.1=	OC-192/STM64/10GE, XFP, 1547.72, 100 GHz, LC	10-2326-01	ITU G694, GR2918	0 to +70

Table 78 DWDM XFPs (3 of 3)

Product ID	Product description	Part number	Applicable standard	Temperature range (°C)
ONS-XC-10G-54.9=	OC-192/STM64/10GE, XFP, 1548.51, 100 GHz, LC	10-2325-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-55.7=	OC-192/STM64/10GE, XFP, 1549.32, 100 GHz, LC	10-2324-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-56.5=	OC-192/STM64/10GE, XFP, 1550.12, 100 GHz, LC	10-2323-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-57.3=	OC-192/STM64/10GE, XFP, 1550.92, 100 GHz, LC	10-2322-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-58.1=	OC-192/STM64/10GE, XFP, 1551.72, 100 GHz, LC	10-2321-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-58.9=	OC-192/STM64/10GE, XFP, 1552.52, 100 GHz, LC	10-2320-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-59.7=	OC-192/STM64/10GE, XFP, 1553.33, 100 GHz, LC	10-2319-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-60.6=	OC-192/STM64/10GE, XFP, 1554.13, 100 GHz, LC	10-2318-01	ITU G694, GR2918	0 to +70
ONS-XC-10G-61.4=	OC-192/STM64/10GE, XFP, 1554.94, 100 GHz, LC	10-2317-01	ITU G694, GR2918	0 to +70

SFP technical details

This section contains tables listing SFP technical details.

SONET/SDH SFPs

The SFPs are compatible with the SONET/SDH standards, and support the digital diagnostic functions specified in the SFF-8742 Multi-Source Agreement (MSA).

Table 79 lists the optical parameters for the Cisco ONS SONET/SDH SFPs.

Table 79 SONET/SDH SFPs optical specification

Product ID	Operating wavelength range (nm)	Transmit power range (dBm)	Receiver power range (dBm)	Maximum dispersion (ps/nm)
15454-SFP3-1-IR= ONS-SI-155-I1=	1261-1360	-15 to -8	-28 to -8	96 (@155.52Mb/s)
ONS-SI-155-L1=	1263-1360	-5 to 0	-34 to -10	NA
ONS-SI-155-L2=	1480-1580	-5 to 0	-34 to -10	NA
ONS-SI-155-SR-MM=	1270-1380	-20 to -14	-30 to -14	NA
ONS-SI-622-SR-MM=	1270-1380	-20 to -14 (50 mm) -24 to -14 (62.5 mm)	-26 to -14	NA
15454-SFP12-4-IR= ONS-SI-622-I1=	1293-1334	-15 to -8	-28 to -8	46 (@ 622.08Mb/s)
ONS-SI-622-L1=	1280-1335	-3 to +2	-28 to -8	NA
ONS-SI-622-L2=	1480-1580	-3 to +2	-28 to -8	NA
ONS-SE-2G-S1= ONS-SI-2G-S1=	1266-1360	-10 to -3	-18 to -3	12
ONS-SI-2G-I1= 15454-SFP-OC48-IR=	1260-1360	-5 to 0	-18 to 0	NA
ONS-SI-2G-L1=	1280-1335	-3 to +2	-27 to -9	NA
ONS-SE-2G-L2= ONS-SI-2G-L2=	1500-1580	-3 to +2	-28 to -9	1200-1600[a]

a. The indicated dispersion range corresponds to the approximate worst-case dispersion for 80 km G.652/G.654 fiber over the wavelength range 1500-1580 nm.

GBIC / DATA SFPs

The ONG Pluggables GBIC / Data SFPs are compatible with the IEEE 802.3, SBCON Single-Byte Command Code Sets CONnection architecture (SBCON) Rev2.3 [ESCON], ANSI INCITS 352-2002 Information technology - Fibre Channel – Physical Interfaces (FC-PI) Rev. 13[1xFC & 2xFC] and support the digital diagnostic functions specified in the SFF-8742 Multi-Source Agreement (MSA).

Table 80 on page 1255 through Table 82 on page 1256 list the optical parameters for the Cisco ONS Data SFPs.

Table 80 Ethernet pluggables optical specification

Product ID	Operating wavelength range (nm)	Transmit power range (dBm)	Receiver power range (dBm)	Maximum dispersion (ps/nm)
ONS-SE-100-LX10=	1260-1360	-15 to -8	-28 to -8	NA
ONS-SE-100-FX=	1270-1380	-20 to -14[a]	-31 to -14	NA
15327-SFP-LC-SX= 15454-SFP-LC-SX= ONS-SC-GE-SX= ONS-SI-GE-SX= 15454-SFP-GEFC-SX= ONS-SE-G2F-SX= ONS-GC-GE-SX=	770-860	-9.5 to 0	-17 to 0[b]	NA
15327-SFP-LC-LX= 15454-SFP-LC-LX/LH= 15454-SFP-GE+-LX= ONS-SC-GE-LX= ONS-SI-GE-LX= ONS-SE-G2F-LX= ONS-GC-GE-LX=	1270-1355	-9.5 to -3	-19 to -3[c]	NA
ONS-GC-GE-ZX= ONS-SI-GE-ZX= ONS-SE-GE-ZX=	1500-1580	0 to +5	-23 to -3	1200-1600[d]

a. 62.5/125 µm, NA = 0.275 Fiber.

b. Minimum Stressed Sensitivity (10-12): -12.5(62.5um) and -13.5(50um) dBm.

c. Minimum Stressed Sensitivity (10-12): -14.4 dBm.

d. The indicated dispersion range corresponds to the approximate worst-case dispersion for 80 km G.652/G.654 fiber over the wavelength range 1500-1580 nm.

Table 81 ESCON SFPs optical specification

Product ID	Operating wavelength range (nm)	Transmit power range (dBm)	Receiver power range (dBm)	Maximum dispersion (ps/nm)
15454-SFP-200= ONS-SE-200-MM=	1280-1380	-20.5 to -15	-14 to -29[a]	NA

a. Based on any valid 8B/10B code pattern measured at, or extrapolated to, 10E-15 BER measured at center of eye.

Table 82 FC/FICON pluggables optical specification

Product ID	operating wavelength range (nm)	Transmit power range (dBm)	Receiver sensitivity (dBm)	Operating distance (m)
1G FC/FICON (1062.5 Kbps)				
15454-SFP-GEFC-SX= ONS-SE-G2F-SX=	770-860	-10 to -3.5	-22	0.5 to 500m (50/125?mm fiber) 0.5 to 300m (62.5/125??m fiber)
ONS-GX-2FC-MMI=	770-860	-10 to -2.5	-22	0.5 to 500m (50/125??m fiber) 0.5 to 300m (62.5/125??m fiber)
15454-SFP-GE+-LX= ONS-SE-G2F-LX=	1270-1360	-10 to -3.5	-22	2 to 10,000
ONS-GX-2FC-SML=	1270-1355	-9 to -3	-23.5	2 to 10,000
2G FC/FICON (2125 Kbps)				
15454-SFP-GEFC-SX= ONS-SE-G2F-SX=	830-860	-10 to -3.5	-20	0.5 to 300
ONS-GX-2FC-MMI=	820-860	-9.5 to -5	-15	0.5 to 300
15454-SFP-GE+-LX= ONS-SE-G2F-LX=	1270-1360	-10 to -3.5	-21	2 to 10,000
ONS-GX-2FC-SML=	1270-1355	-9 to -3	-23.5	2 to 10,000
4G FC/FICON (4250 Kbps)				
ONS-SE-4G-MM=	830-860	-9 to -3.5	-15	0 to 70 (50/125??m fiber) 0 to 150 (62.5/125??m fiber)
ONS-SE-4G-SM=	1270-1355	-8.4 to -3	29??W OMA[a]	2 to 10,000

a. Specified OMA at 4.25 Gb/s is equivalent to an average power of -17.3 dBm at an ER of 9 dB.

XWDM SFPs

ONG is offering a full set of CWDM SFPs for 155Mbps and 622Mbps and DWDM SFPs for 2.5 Gbps application. Table 83 through Table 85 on page 1257 list SFPs optical parameters.

Table 83 CWDM SFPs optical specification

Product ID	Receiver wavelength range (nm)	Spectral width (nm)	Transmit power range (dBm)	Receiver power range (dBm)
ONS-SE-155-xxxx=	1460-1620	1	0 to +5	-34 to -3 (BER 10^{-10})
ONS-SE-622-xxxx=	1460-1620	1	0 to +5	-28 to -3 (BER 10^{-10})
ONS-SC-Z3-xxxx=	1460-1620	1	0 to +4	-28 to -9 (BER 10^{-10})

xxxx from 1470 to 1610.

Table 84 DWDM SFPs optical specification

Product ID	Receiver wavelength range (nm)	Transmitter stability (pm)	Spectral width (pm)	Transmit power range (dBm)
ONS-SC-2G-xx.x=	1260-1620[a]	-100 to +100 (100GHz Spacing)	200	0 to +4

a. Receiver Sensitivity specified over 1528-1561nm only, with 2dB degradation permitted outside of this range.

xx.x from 30.3 to 60.6

Table 85 DWDM SFPs optical performances

Power limited performances			
Input power range	dBm	-9 to -28	At BER=10e-12 with SONET framed PRBS23 at OSNR of 21dB, 0.1nm BW
Dispersion Tolerance	ps/nm	-800 to +2400	Power Penalty=3dB, OSNR=21dB@0.1nmBW (Noise Penalty=0dB)
Noise limited performances			
Input power range	dBm	-9 to -22	At BER=10e-12 with SONET framed PRBS23. at OSNR of 16dB at 0.1nm bandwidth.
Dispersion Tolerance	ps/nm	-800 to +2400	Noise Penalty=3dB, OSNR=19dB@0.1nmBW (Power Penalty=0dB)

Grey XFPs Grey XFPs for 10 Gbps application. Table 86 lists optical parameters.

Table 86 XFPs optical specification

Product ID	Transmitter wavelength range (nm)	Transmit power range (dBm)	Receiver wavelength range (nm)	CD robustness (ps)	Receiver power range (dBm)
ONS-XC-10G-S1=	1260-1335	-6 to -1[a] 8.2 to +0.5 [b]	1260-1565	6.6	-11 to -1[a] -14.4 to +0.5[b, c]
ONS-XC-10G-I2=	1530-1565	-1 to +2	1260-1565	800	-14 to +2
ONS-XC-10G-L2=	1530-1565	0 to +4	1260-1565	1600	-24 to -7

a. SONET/SDH Application.

b. 10 GE/10 G FC Application.

c. Stressed receiver sensitivity (max) in OMA is -10.3 dBm.

DWDM XFPs DWDM XFPs for 10 Gbps application. Table 87 lists optical parameters.

Table 87 DWDM XFPs optical specification

Product ID	Receiver wavelength range (nm)	Transmitter stability (pm)	Spectral width (pm)	Transmit power range (dBm)
ONS-XC-10G-xx.x=	1260-1607[11]	-100 to +100 (100GHz Spacing)	200	-1 to +3

xx.x from 30.3 to 61.4

Table 88 DWDM XFPs optical performances (1 of 2)

Short Wavelength performances			
Input power range	dBm	-7 to -20	At BER=10e-12 (@1310nm ± 20nm) applicable at 9.9G, 10.3G only
Long Wavelength performances C band NO-FEC applications power limited			
Input power range	dBm	-7 to -23	At BER=10e-12 applicable at 9.9G, 10.3G only 23dB OSNR
Input power range		-7 to -20	At BER=10e-12 (-500 to +1600 ps/nm) applicable at 9.9G, 10.3G only – 23dB OSNR
Long Wavelength performances C band NO-FEC applications noise limited			
Input power range	dBm	-7 to -18	At BER=10e-12 applicable at 9.9G, 10.3G only 17dB OSNR

Table 88 DWDM XFPs optical performances (2 of 2)

Input power range	dBm	-7 to -18	At BER=10e-12 (-500 to +1600 ps/nm) applicable at 9.9G, 10.3G only – 20dB OSNR
Long Wavelength performances C band FEC applications noise limited			
Input power range	dBm	-7 to -18	At BER PREFEC <10e-5 applicable at 10.7G, 11.1G only -- 11dB OSNR
Input power range	dBm	-7 to -18	At BER PREFEC <10e-5 (-500 to +1100 ps/nm) applicable at 10.7G, 11.1G only – 12dB OSNR
Long Wavelength performances C band E-FEC applications power limited			
Input power range	dBm	-7 to -27	At BER PREFEC <7*10e-4 applicable at 10.7G, 11.1G only – 23dB OSNR
Input power range	dBm	-7 to -24	At BER PREFEC <7*10e-4 (-500 to +1300 ps/nm) applicable at 10.7G, 11.1G only – 23dB OSNR
Long Wavelength performances C band E-FEC applications noise limited			
Input power range	dBm	-7 to -18	At BER PREFEC <7*10e-4 applicable at 10.7G, 11.1G only – 8dB OSNR
Input power range	dBm	-7 to -18	At BER PREFEC <7*10e-4 (-500 to +1100 ps/nm) applicable at 10.7G, 11.1G only – 9dB OSNR

MSTP management

The ONS 15454 MSTP has several options in which it can be managed:

- Cisco Transport Controller (CTC), discussed briefly in this section
- Cisco Transport Manager (CTM), , discussed briefly in this section
- SNMP v1, v2
- TL1
- Corba

The ONS 15454 MSTP has one software load which distributes the appropriate firmware to the individual line cards. If cases that new firmware is available for a particular line card an administrator is required to activate the new firmware if the new firmware is service disruptive.

Cisco Transport Controller (CTC)

Cisco Transport Controller (CTC) is a provisioning and administration tool located on the TCC2 card and is activated when a web browser is directed to the IP address of the ONS 15454 MSTP.

CTC is a Java application that is uses Corba to communicate with the node to retrieve / provision / monitor a network of MSTP nodes.

Cisco Transport Manager (CTM)

Cisco Transport Manager (CTM) is an intelligent element management system (EMS) that efficiently manages the entire Cisco ONS Family of products as well as the Cisco Carrier Routing System (CRS-1), Cisco XR 12000 Series Router, Cisco 7609, Cisco MGX 8880 Media Gateway, and voice-enabled Cisco MGX 8850 and MGX 8830 Multiservice Switches.

Cisco Transport Manager supports configuration, fault, performance, and security management and serves as a foundation for integration into a larger OSS environment. High availability, self-healing processes, and comprehensive management across different types of Cisco devices give service providers and large enterprises a resilient, adaptable, and integrated carrier-class EMS solution.

Cisco Transport Manager is a powerful GUI-based management system that makes advanced management capabilities such as A-to-Z circuit provisioning easy to learn and use. Its explorer-style navigation and extensive use of wizards, tool tips, legends, and online help allow operators to perform tasks quickly and efficiently. Many otherwise tedious, time-consuming, and error-prone tasks can be fully streamlined using these intuitive tools.

Cisco Transport Manager is designed for continuous operation. High-availability deployment configurations are available to provide local and geographic redundancy options. Cisco Transport Manager can cope with heavy load scenarios such as high circuit-transaction rates, performance-monitoring data collection, alarm storms, and numerous simultaneous clients. Cisco Transport Manager plays a crucial role in the high availability of the network itself. The Cisco Transport Manager fault-management and performance-management capabilities help to ensure that network problems are discovered quickly and accurately so that often they can be addressed even before a customer's service is noticeably affected.

MSTP further reading

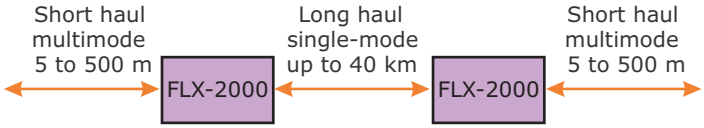
For further reading on the ONS 15454 MSTP the following link has white papers, data sheets, case studies, etc., on the platform.

http://www.cisco.com/en/US/products/hw/optical/ps2006/tsd_products_support_series_home.html

Finisar FLX-2000 link extender

The Symmetrix Fibre Channel fabric implementation supports the Finisar Model FLX-2000-1062-40, a 40 km optical extender.

As illustrated in the figure below, the function of the FLX-2000 is to relay or convert a Fibre Channel multi-mode signal from the source and translate the input to a single-mode long-haul signal. Another FLX-2000 then translates the single-mode signal back into a multi-mode signal into the target device .



Configuration guidelines

The following are some guidelines and limitations for using extenders in a switched Fibre Channel environment with Symmetrix Fibre Channel storage.

Host connection

The FLX-2000 extender can be used between a host Fibre Channel connection and a Connectrix® Fibre Channel switch port connection.

Symmetrix connection

The FLX-2000 extender can be used between a Symmetrix Fibre Channel director (FA) port and a Connectrix Fibre Channel switch port connection.

Switch connections

Multiple extender connections into a Connectrix (a requirement for high-availability environments) require multiple FLX extenders, since each FLX-2000 supports only one through-connection.

Buffer-to-Buffer Credit calculation

In order to achieve maximum utilization of the Fibre Channel link, it is highly advisable that both ports, connected on either side of the long-haul setup provided by the extender, be capable of high BB_Credit counts.

Use the following formula to calculate the approximate BB_Credit required for the specific long haul application:

(Required BB_Credit) = (One way distance in Km) / 2

Cable requirements

A long-haul connection requires a long-haul fiber connector, ST/PC single mode, zero dispersion at 1310 nm, 9/125mm, maximum attenuation 0.4 dB/Km@1310 nm.

Power budget

Power measurements must be taken in order to insure long-haul link integrity. Maximum losses in long-haul medium (cable and connectors) must not exceed 23 dBm:

- Minimum transmission power is -3 dB.
- Minimum receiver sensitivity is -26 dB.

Note: Consult single-mode cable manufacturer specifications.

For reference, *dBm* is a logarithmic power ratio of a power source relative to 1 mW (milliwatt):

dB = 10 x LOG (P1/P2)
dBm = 10 x LOG (P/1E-3)

where:

p2 = 1 mW
1E-3 = 0.001 W

Examples:

0 dBm = 1 mW
-20 dBm = 0.01 mW

Firmware

The FLX-2000 firmware must be revision 3.0 or later.

Symmetrix microcode

The Symmetrix microcode must be 5265.38.24 or later.

Diagnostics and maintenance

Repair and system diagnostics can be performed through the FLX control console, as described in the FLX user manual.

SONET overview

SONET (Synchronous Optical NETwork) is a telecommunications standard developed by the Exchange Carriers Standards Association for ANSI. SONET defines a technology for carrying different capacity signals through a synchronous optical network. The standard defines a byte-interleaved multiplexed transport occupying the physical layer of the OSI model. Synchronization is provided by one principal network element with a very stable clock (Stratum 3), which is sourced on its outgoing OC-N signal. This clock is then used by other network elements for their clock (loop timing).

SONET is useful in a SAN for consolidating multiple low-frequency channels (Client ESCON and 1 Gb Fibre Channel) into a single higher-speed connection. This can reduce DWDM wavelength requirements in an existing SAN infrastructure. It can also allow a distance solution to be provided from any SONET service carrier, saving the expense of running private optical cable over long distances.

The basic SONET building block is an STS-1 (Synchronous Transport Signal), composed of the transport overhead plus a synchronous payload envelope (SPE), totaling 810 bytes. The 27-byte transport overhead is used for operations, administration, maintenance, and provisioning. The remaining bytes make up the SPE, of which an additional nine bytes are path overhead. It is arranged as depicted in Figure 488. Columns 1, 2, and 3 are the transport overhead.

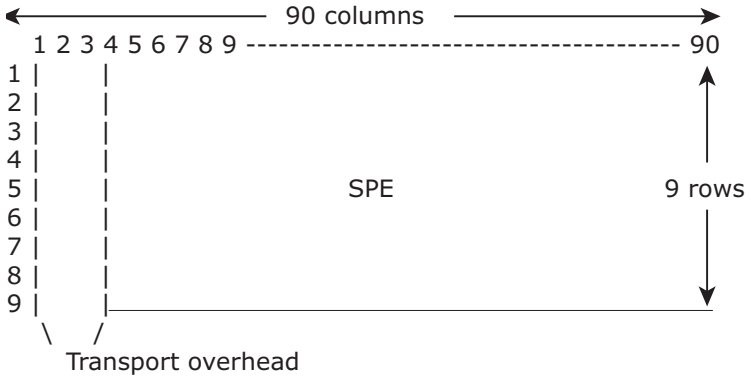


Figure 488 STS-1 organization

An STS-1 operates at 51.84 Mb/s, so multiple STS-1s are required to provide the necessary bandwidth for ESCON, Fibre Channel and Ethernet, as shown in Table 89. Multiply the rate by 95 % to obtain the usable bandwidth in an STS-1 (reduction due to overhead bytes).

Table 89 STS-1s and optical carrier rates

STS	Optical Carrier	Optical Carrier Rate
STS-1	OC-1	51.84 MB/s
STS-3	OC-3	155.52 MB/s
STS-12	OC-12	622.08 MB/s
STS-48	OC-48	2488.32 MB/s
STS-192	OC-192	9953.28 MB/s

As Table 89 shows, one OC-48 can carry roughly 2.5 channels of 1 Gb/s traffic. To achieve higher data rates for client connections, multiple STS-1s are byte-interleaved to create an STS-N. SONET defines this as byte-interleaving three STS-1s into an STS-3, and subsequently interleaving STS-3s.

By definition, each STS is still visible and available for ADD/DROP multiplexing in SONET, although most SAN requirements can be met with less complex point-to-point connections. The addition of DWDM can even further consolidate multiple SONET connections (OC-48), while also providing distance extension.

How STS-1s are interleaved to create an STS-12:

```
STS byte 1            Byte order - A1 is sent first
A1 --------\
A2 -------- -----G1 STS-3   A3  A2  A1-----\
A3 --------/                                \
                                             \
A4 --------\                                 \
A5 --------- ----G2 STS-3   A6  A5  A4 --------\
A6 --------/                                   \
                                                 >--- STS-12 A10 A7 A4 A1
A7 --------\                                   /
A8 --------- ----G3 STS-3   A9  A8  A7 ---------/
A9 --------/                                 /
                                             /
A10 -------\                                /
A11 -------- ----G4 STS-3   A12 A11 A10 ---/
A12 -------/
Entire 12 byte stream:
STS-12  A12  A9  A6  A3  A11  A8  A5  A2  A10  A7  A4  A1
                      |                |                |
                      A3               A2               A1   G1 STS-3
```

CIENA CN 2000

The CIENA CN 2000 is a SONET multiplexer performing the adaptation of Fibre Channel, Gigabit Ethernet, and ESCON to SONET, providing storage application transparent connectivity over SONET or DWDM networks.

The CN 2000 has two models:

- OUSP 2048 runs OUSP software release 2.04.
- OUSP 2048E runs OUSP software release 3.10.

The OUSP 2048 provides a total of six and the OUSP 2048E provides a total of eight configurable connections for Fibre Channel, ESCON, or Gigabit Ethernet into a SONET service (OC-3/OC-12/OC-48 or DS-3, which is only available on OUSP 2048E). These interfaces are then configured on the CN 2000 to reside on dedicated STS-1s (up to 96). The number of STS-1s provisioned determines the bandwidth available to the connection, and should be allocated to meet performance guidelines as required by the application.

The equivalent number of STS-1s for the support of wire rate of each of the protocols is:

- Fibre Channel — 20 STS-1
- Gigabit Ethernet — 24 STS-1
- ESCON — 4 STS-1

In many cases, however, applications require significantly less than wire rate. Therefore, less than the full rate of STS-1 can be provisioned, allowing for more circuits per line interface to be transported.

Note: Table 90 on page 1268 contains examples of sub-rate provisioning.

If the OUSP is configured for service path protection, 96 STS-1s through two OC-48 connections are available, yielding 5 Gb/s total SONET bandwidth. 1+1 protection reduces this to 2.5 Gb/s (48 STS-1s).

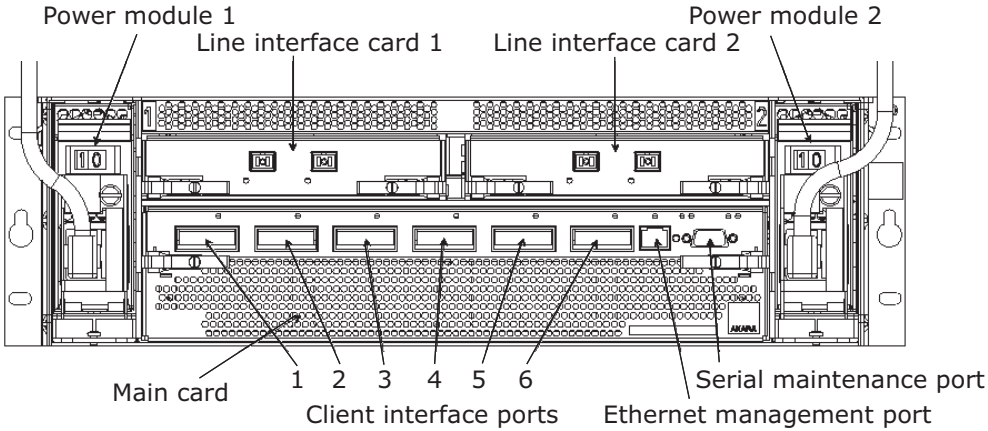


Figure 489 CIENA CN 2000, OUSP 2048 model (rear view)

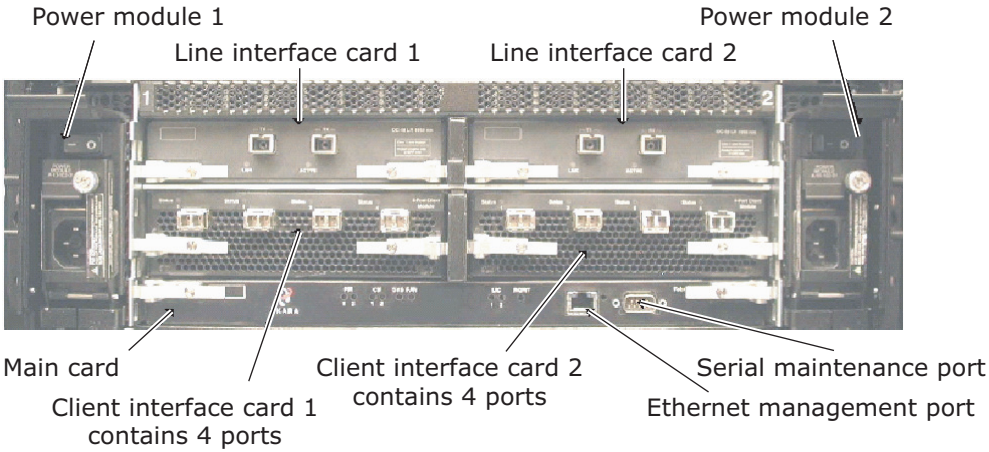


Figure 490 CIENA CN 2000, OUSP 2048E model (rear view)

Once the network connection is known, use Table 90 to allocate the appropriate number of STS-1s.

Table 90 Maximum number of Fibre Channel circuits on the OUSP

Line Interface	Network	Number of STS-1s per Fibre Channel circuit	Maximum number of circuits (1+1 Linear protection)	Maximum number of circuits (Service path protection)
OC-48	DWDM	20	2	4
OC-48	DWDM	10	4	6
OC-48	DWDM	8	6	6
OC-12	DWDM/SONET	12	1	2
OC-12	DWDM/SONET	6	2	4
OC-12	DWDM/SONET	2	6	6
OC-3	SONET	3	1	2
OC-3	SONET	1	3	6

OC-48 is the most com1350mon interface into DWDM networks, with OC-12 being less common. OC-3 is the most common interface for SONET connectivity, with OC-12 being less common.

There are two ways to configure the CIENA CN 2000:

- Start a CLI session using Telnet, or
- Install the OUSM 7000 GUI and server.

The OUSM software requires a Solaris / Windows server and a Windows client. Refer to the CIENA documentation for additional details.

The following are the recommended configuration parameters for the CIENA CN 2000:

- Enable Extended Reach Mode (flow control on) for Fibre Channel and insure that auto negotiation for BB_Credit is configured.
- Enable the Link Reset on Protection switch (**lr-protocol-link-reset**).
- Set one CN 2000's synchronization mode as self-timed, and all others loop-timed.

- ESCON SRDF requires four STS-1s. If configured for fewer than four STS-1s and FarPoint is used, the FarPoint buffers cannot exceed 500 KB.
- (Model OUSP 2048E only) Enable Auto Compression mode to get optimum throughput.

You must also enable Interop mode on Brocade switches.

Extended reach mode distance solution for Fibre Channel

In extended reach mode, the CN 2000 extends the reach of the Fibre Channel protocol by locally terminating the R_RDYs from the attached equipment and providing instantaneous R_RDY responses to received frames. The attached equipment still transparently performs Fibre Channel port login through the CN 2000.

Extended reach mode enables the attached equipment receiving data to control the maximum sustainable throughput over any distance. When an attached device slows the traffic flow by reducing or stopping the transmission of R_RDYs to the CN 2000, the CN 2000 buffers the frames in transit between the connected CN 2000s until the attached device resumes normal traffic flow.

Some frame loss still can occur if the buffers overflow due to extended traffic pauses, depending on the bandwidth provisioned for the service and the distance between the connected CN 2000s. To guarantee zero frame loss, the system must operate within the parameters defined under *Enhanced Distance Limits for Lossless Throughput During Endpoint Congestion*.

Enhanced distance limits for lossless throughput during endpoint congestion

When Fibre Channel extended reach or Gigabit Ethernet perfect Quality of Service is enabled, CIENA guarantees zero frame loss over the distance calculated using the following formula:

$$\text{Distance} \leq \frac{\text{BB_Credit} * 108\ (\text{km} * \text{MB/s})}{\text{Throughput}}$$

where:

- **Distance** the service path length in km.
- **BB_Credit** is **256** for OUSP 2048 and **1792** for OUSP 2048E.
- **Throughput** is the average traffic (in in MB/s) transmitted across the service path.

A service path can be extended indefinitely when the attached client equipment receiving the traffic is not congested. The receiving device

is considered congested if it stops issuing R_RDYs (for Fibre Channel).

When the receiving client equipment is not congested, there is theoretically no limit to how far the service path can be extended.

If the limit defined in the calculation is exceeded and the receiving client equipment is congested, frames could be lost.

The *CN 2000 Installation, Configuration and Engineering Guide* (A0012103-A00 ICE GUIDE.pdf, available from the CIENA Corporation), contains additional information on flow control, transport latency, and data compression.

Figure 491 depicts the client (Fibre Channel and ESCON) and line (OC-N) connections for the CIENA solution.

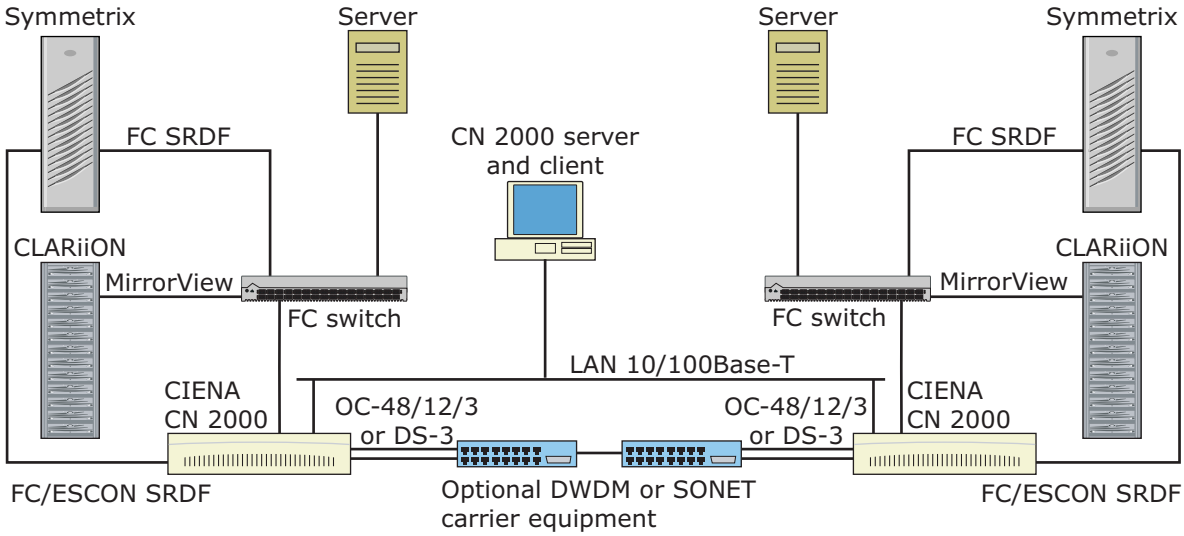


Figure 491 Typical CIENA CN 2000 distance extension setup

Lucent OptiStar EdgeSwitch

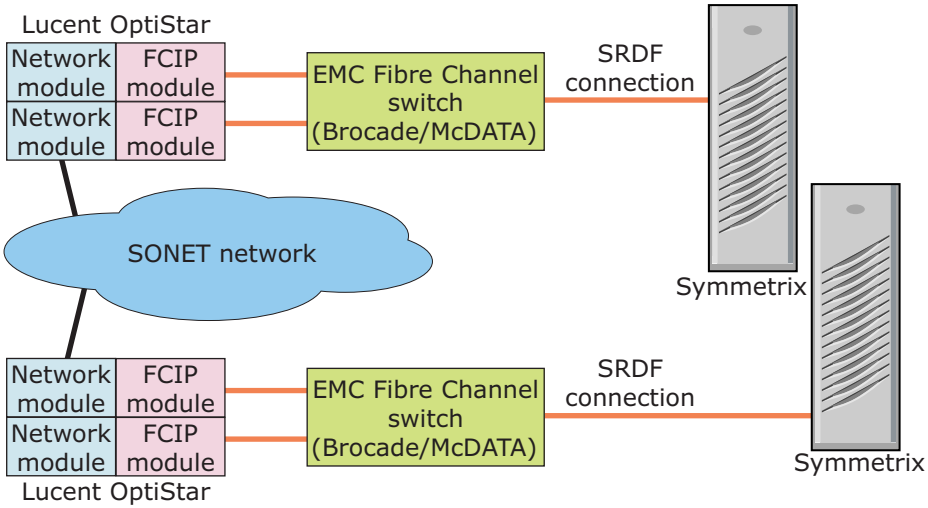


Figure 492 **OptiStar switches over SONET network**

Supportable configurations

Minimum requirements for this setup include the following:

- The OptiStar switches will be used as part of a DR (disaster recovery) SAN only.
- SRDF is the only supported configuration.
- No direct-attach or remote Host I/O will be supported in this configuration.
- Ideally, the total aggregated throughput of SRDF links should not exceed the OptiStar Edgeswitch Network module's bandwidth.

The OptiStar switch will be supported in the following network characteristics (refer to Lucent technical notes):

- 0% packet loss.
- Guaranteed bandwidth; however, the minimum contiguous bandwidth should not be less than 622 Mb/s (an OC-12 link).

As a result of these limitations, the OptiStar will fit best in a leased bandwidth network such as SONET or DWDM.

The DR SAN will be composed of one OptiStar EdgeSwitch plus a Brocade or Brocade M-Series switch at the local site, and the same number and type of switches at the remote site.

Note: E-Lab Navigator describes the latest supported configurations.

Symmetrix setup

Symmetrix SRDF ports should be configured as standard Fibre Channel SRDF ports. The Lucent switch, similar to any Fibre Channel switch, provides Fibre Channel switching services (such as zoning, RSCN, and name services) to the SRDF ports connected to the Brocade or Brocade M-Series switch.

SRDF can be set up to run in any one of the following modes:

- Semi-synchronous Adaptive Copy Disk mode
- Semi-synchronous Adaptive Copy Write Pending mode
- Synchronous mode

Blade server switch modules

Note: For more information on blade servers, refer to "Blade servers" on page 1296.

The Brocade switch modules in this section integrate Brocade technology into the respective blade server architecture.

The Brocade M-Series modules in this section are actually QLogic switch modules with Brocade M-Series firmware.

Table 91 lists switch features for Brocade and Brocade M-Series switches.

Table 91 Switch features

	ASIC family	No. of Internal ports	No. of External ports	Line speed per port (in Gb/s)	ISL Trunking performance	
					Max no. of ports per trunk	Max. speed per ISL trunk (in Gb/s)
BROCADE:						
DELL SW3014	Bloom 2	10	4	2/1	4	8
DELL SW4016	Golden Eye	10	4	4/2/1	4	16
HP A7535A	Golden Eye	8	4	4/2/1	4	16
IBM 26K5601	Bloom 2	14	2	2/1	N.A	N.A
IBM 90P0165	Bloom 2	14	2	2/1	N.A	N.A
IBM 32R1812	Golden Eye	14	6	4/2/1	4	16
IBM 32R1813	Golden Eye	7	3	4/2/1	3	12

Table 91 Switch features (continued)

Brocade M-Series:						
DELL SW4314	Agate	10	4	2/1	4	8
DELL SW4416	Agate-2	10	4	4/2/1	4	16
HP A8001A	Agate	8	2	4/2/1	4	16
IBM 32R1790	Agate	14	6	2/1	4	8
IBM 32R1833	Agate-2	14	6	4/2/1	4	16
IBM 32R1905	Agate-2	7	3	4/2/1	3	12

DELL SW3014 Brocade 4-port 1/2 Gb FC switch module

The SW3014 includes the Brocade Fabric Operating System (Fabric OS) and is compatible with the entire SilkWorm product family. It is designed to work solely within switch module bay 3 or 4 of a Dell Poweredge 1855 chassis.

ASIC family Bloom 2

Key features Connectivity feature includes:

- 14 ports
  - 10 internal fabric ports
  - 4 external 2 Gb / 1 Gb auto-sensing auto-links: Ports numbered 10-13
- At the time of this publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)
  - Alias Server (multicast)
  - Brocade Zoning
  - Fabric Watch
  - Extended Fabrics
  - ISL Trunking
  - Performance Monitoring
  - Secure Fabric OS
- ISL Trunking to other B series switches:
  - Up to four ports per trunk, doing up to 8 Gb/s per ISL trunk.
- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

- 2.125 Gb/s line speed, full duplex per Fibre Channel port.

Management

- Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.
- Management software:
  - Telnet
  - SNMP
  - Web Tools
  - EMC Fabric Manager
  - Interactive and non-interactive (nondisruptive) firmware upgrades.
- RJ-45 connector
  - Supports 10/100 Mb Ethernet connections.

Mechanical specifications

Mechanical specifications include:

- Width = 127 mm (5 in)
- Height = 35 mm (1.375 in)
- Depth = 247.7 mm (9.75 in)

DELL SW4016 Brocade 4-port 1/2/4 Gb FC switch module

The SW3014 includes the Brocade Fabric Operating System (Fabric OS) and is compatible with the entire SilkWorm product family. It is designed to work solely within switch module bay 3 or 4 of a Dell Poweredge 1855 chassis.

ASIC family

Golden Eye

Key features

Connectivity feature includes:

- 14 ports
  - 10 internal fabric ports
  - 4 external 4 Gb/2 Gb/1 Gb auto-sensing auto-links: Ports numbered 10-13
- At the time of this publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)
  - Alias Server (multicast)
  - Brocade Zoning
  - Fabric Watch
  - Extended Fabrics
  - ISL Trunking
  - Performance Monitoring
  - Secure Fabric OS
- ISL Trunking to other B series switches:
  - Up to four ports per trunk, doing up to 16 Gb/s per ISL trunk.
- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

- 4.25 Gb/s line speed, full duplex per Fibre Channel port.

Management

- Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.
- Management software:
  - Telnet
  - SNMP
  - Web Tools
  - EMC Fabric Manager
- Interactive and non-interactive (nondisruptive) firmware upgrades.
- RJ-45 connector
  - Supports 10/100 Mb Ethernet connections.

Mechanical specifications

Mechanical specifications include:

- Width = 127 mm (5 in)
- Height = 35 mm (1.375 in)
- Depth = 247.7 mm (9.75 in)

HP A7535A Brocade 4-port 1/2/4Gb FC switch module

The HP A7535A switch module plugs into the back of the HP Blade System chassis thus resulting in a zero footprint Fibre Channel connection solution like any other switch module.

ASIC family

Golden Eye

Key features

Connectivity feature includes:

- 12 ports
  - 8 internal fabric ports
  - 4 external 2 Gb/1 Gb auto-sensing auto-links: Ports numbered 8-11)
- At the time of this publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)
  - Alias Server (multicast)
  - Brocade Zoning
  - Fabric Watch
  - Extended Fabrics
  - ISL Trunking
  - Performance Monitoring
  - Secure Fabric OS
- ISL Trunking to other B series switches. Up to four ports per trunk, doing up to 16 Gb/s per ISL trunk.
- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

- 4.25 Gb/s line speed, full duplex per Fibre Channel port.

Management

- Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.
- Management software:
  - Telnet
  - SNMP
  - Web Tools
  - EMC Fabric Manager
- Nondisruptive firmware upgrades.
- RJ-45 connector
  - Supports 10/100 Mb Ethernet connections.

Mechanical specifications

Mechanical specifications include:

- Width = 86.6 mm (3.41 in)
- Height = 38.7 mm (1.525 in)
- Depth = 114.12 mm (4.493 in)

IBM PN 26K5601 Brocade 2-port entry-level switch module

This switch is ideal for smaller Brocade Storage Area Networks still providing the option to expand and grow as needed. Integration of SAN switching capabilities reduces complexity and increases SAN manageability.

ASIC family

Bloom 2

Key features

Connectivity feature includes:

- 16 ports
  - 14 internal fabric ports
  - 2 external 2 Gb/1 Gb auto-sensing auto-links: Ports 0 and 15
- Switch module supports connectivity to a maximum of 2 domains per fabric. Thus, this switch is ideal for smaller Brocade Storage Area Networks.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)
  - Alias Server (multicast)
  - Brocade Zoning
  - Fabric Watch
  - Extended Fabrics
  - ISL Trunking
  - Performance Monitoring
  - Secure Fabric OS
- ISL Trunking option is not available since the two external ports are not consecutive ports.
- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

- 2.125 Gb/s line speed, full duplex per Fibre Channel port.

Management

- Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.
- Management software:
  - Telnet

- SNMP
- Web Tools
- EMC Fabric Manager

- Nondisruptive firmware upgrades.
- Support for 10/100 Mb Ethernet connections for out-of-band management.

Mechanical specifications

Mechanical Specifications include:

- Width = 260 mm (10.2 in)
- Height = 112 mm (4.4 in)
- Depth = 29 mm (1.14 in)

IBM PN 90P0165 Brocade 2-port enterprise-level switch module

This switch is ideal for larger Brocade Storage Area Networks. Integration of SAN switching capabilities reduces complexity and increases SAN manageability.

ASIC family Bloom 2

Key features Connectivity feature includes:

- 16 ports
  - 14 internal fabric ports
  - 2 external 2 Gb/1 Gb auto-sensing auto-links: Ports 0 and 15
- At the time of this publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)
  - Alias Server (multicast)
  - Brocade Zoning

- Fabric Watch
- Extended Fabrics
- ISL Trunking
- Performance Monitoring
- Secure Fabric OS

◆ ISL Trunking option is not available since the two external ports are not consecutive ports.

◆ Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

◆ 2.125 Gb/s line speed, full duplex per Fibre Channel port.

Management

◆ Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.

◆ Management software:

- Telnet
- SNMP
- Web Tools
- EMC Fabric Manager

◆ Nondisruptive firmware upgrades.

◆ Support for 10/100 Mb Ethernet connections for out-of-band management.

Mechanical specifications

Mechanical specifications include:

◆ Width = 260 mm (10.2 in)

◆ Height = 112 mm (4.4 in)

◆ Depth = 29 mm (1.14 in)

IBM PN 32R1812 Brocade 6-port SAN switch module

This switch module enables high end-to-end performance with 1, 2, and 4 Gigbits per second SAN solutions for the data center.

ASIC family Golden Eye

Key features

Connectivity feature includes:

- 20 ports
  - 14 internal fabric ports
  - 6 external 4 Gb/2 Gb/1 Gb auto-sensing auto-links: Ports 14-19
- At the time of this publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)
  - Alias Server (multicast)
  - Brocade Zoning
  - Fabric Watch
  - Extended Fabrics
  - ISL Trunking
  - Performance Monitoring
  - Secure Fabric OS
- ISL Trunking to other B series switches. In this case, up to four ports per trunk, doing up to 16 Gb/s per ISL trunk.
- Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

- 4.25 Gb/s line speed, full duplex per Fibre Channel port.

Management

- Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.
- Management software:
  - Telnet
  - SNMP
  - Web Tools
  - EMC Fabric Manager

- Nondisruptive firmware upgrades.
- Support for 10/100 Mb Ethernet connections for out-of-band management.

Mechanical specifications

Mechanical specifications include:

- Width = 260 mm (10.2 in)
- Height = 112 mm (4.4 in)
- Depth = 29 mm (1.14 in)

IBM PN 32R1813 Brocade 3-port SAN switch module

This switch module is physically similar to the IBM PN 32R1812, however it has just 10 functional ports. It enables high end-to-end performance with 1, 2, and 4 Gbs per second SAN solutions for the data center.

ASIC family Golden Eye

Key features Connectivity feature includes:

- 10 ports
  - Seven internal fabric ports
  - Three external 4 Gb/2 Gb/1 Gb auto-sensing auto-links: Ports 14-16

    Note: Hence, in this case, the switch module has access to only 7 blades. If access to more blades is required the 20-port IBM brocade PN 32R1812 must be used.

- At the time of this publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

  For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

Fabric services include:

- Simple name server
  - Registered State Change Notification (RSCN)

- Alias Server (multicast)
- Brocade Zoning
- Fabric Watch
- Extended Fabrics
- ISL Trunking
- Performance Monitoring
- Secure Fabric OS

◆ ISL Trunking to other B series switches: In this case, up to four ports per trunk, doing up to 16 Gb/s per ISL trunk.

◆ Class 2, Class 3, and Class F (interswitch frames) Fibre Channel protocol support.

Performance

◆ 4.25 Gb/s line speed, full duplex per Fibre Channel port.

Management

◆ Centralized configuration and management of fabric using client/server architecture (Web Tools) and CLI.

◆ Management software:

- Telnet
- SNMP
- Web Tools
- EMC Fabric Manager

◆ Nondisruptive firmware upgrades.

◆ Support for 10/100 Mb Ethernet connections for out-of-band management.

Mechanical specifications

Mechanical specifications include:

◆ Width = 260 mm (10.2 in)

◆ Height = 112 mm (4.4 in)

◆ Depth = 29 mm (1.14 in)

DELL SW4314 Brocade M-Series 6-port SAN switch module

The SW4314 can be integrated into the Dell Blade server enclosure and provides native switch support and interoperability in existing Brocade M-Series fabrics as well as standard Fibre Channel interoperability in Open Mode.

ASIC family Agate

Key features Connectivity feature includes:

- 14 ports
  - 4 external ports
  - 10 internal ports (2 Gb/s), 1/2 Gb auto sensing
- Brocade M-Series Fabric Mode

  The Brocade M-Series 4314 may be configured to operate in either open standards or Brocade M-Series mode to enable non-disruptive connectivity of the Brocade M-Series 4314 to existing or planned Brocade M-Series-based fabrics.
- ISL trunking is supported on these modules. Up to 4 ports can be used for an ISL trunk. They need not be sequential ports.

Fabric services include:

- Fabric Monitoring
- Zoning
- FDMI: Fabric Device Management Interface, a Storage Networking Industry Association (SNIA) standard that defines the sharing of information between SAN devices, including HBAs and switches
- FC Security: Complies with T11 FC-GS4 security standards
- Access using SSH / SSL
- Fabric Device Management Interface (FDMI), a Storage Networking Industry Association (SNIA) standard that defines the sharing of information between SAN devices, including HBAs and switches. Some of the features include WWN, firmware, and drive revision level data interchange
- RADIUS is an out-of-band security method that can be used to authenticate devices that are logging into the SAN\ fabric
- Fibre Channel service classes: Class 2 and Class 3

Performance 2.125 Gb/s non-blocking throughput on all external .

Management Non-Disruptive firmware upgrade.

Out of band:

- EFCM
- EMC Connectrix Manager
- Brocade M-Series Embedded Web Server - Embedded Web GUI served up as applet from the switch
- SNMP
- Telnet
- EMC VisualSAN

In band:

- GS-3 management server

Mechanical specifications

Mechanical specifications include:

- Width = 260 mm (10.2 in)
- Height = 112 mm (4.4 in)
- Depth = 29 mm (1.14 in)

DELL SW4416 Brocade M-Series 6-port 1/2/4 Gb FC switch module

The SW4316 can be integrated into the Dell Blade server enclosure and provides native switch support and interoperability in existing Brocade M-Series fabrics as well as standard Fibre Channel interoperability in Open Mode.

ASIC family Agate-2

Key features Connectivity features include:

- 16 ports
  - 10 internal ports: Ports numbers 0-9.
  - 6 external auto-sensing ports: Ports numbered 10-15.
- At the time of publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator.

- Interoperability:

  The Brocade M-Series 4314 may be configured to operate in either open standards or Brocade M-Series mode to enable seamless, non-disruptive connectivity of the Brocade M-Series 4314 to existing or planned Brocade M-Series-based fabrics.

- ISL trunking is supported on these modules. Up to 4 ports can be used for an ISL trunk and they need not be sequential ports.

Fabric services include:

- FC Security
  - Complies with T11 FC-GS4 security standards.
- Fabric Device Management Interface (FDMI) is a Storage Networking Industry Association (SNIA) standard that defines the sharing of information between SAN devices, including HBAs and Switches. Some of the features include WWN, firmware and drive revision level data interchange.
- RADIUS is an out-of-band security method that can be used to authenticate devices that are logging into the SAN fabric.
- Fibe Channel service classes: Class 2, Class 3, and Class F.

Performance

4.25 Gb/s non-blocking throughput on all external ports.

Management

- Telnet
- Secure Shell (SSH)
- Switch Management
- Inband Management
- Secure Socket Layer (SSL)
- Embedded GUI
- Simple Network Management Protocol (SNMP)
- Network Time Protocol (NTP)
- Common Information Model (CIM)
- File Transfer Protocol (FTP)
- Management Server (MS)

- Command Line Interface
- Brocade M-Series Element Manager1 through EFCM
- Non-Disruptive Code Load and Activation

Mechanical specifications

Mechanical specifications include:

- Width = 127 mm (5 in)
- Height = 35 mm (1.375 in)
- Depth = 247.7 (9.75 in)

IBM PN 32R1790 Brocade M-Series 6-port SAN switch module

This 6-port Brocade M-Series Fibre Channel (FC) switch enables 2 Gb SAN solutions. It provides easy integration of the IBM Blade center into Brocade M-Series environments.

ASIC family Agate

Key features Connectivity feature includes:

- 20 ports:
  - 14 internal 2 Gb/s ports
  - 6 external 1/2 Gb/s auto-sensing ports
- At the time of publication, EMC supports connectivity to a maximum of 24 domains for this kind of a switch module. In general, these switch modules have a domain ID range from 1-239.

  Note: For an actual count of the maximum number of domains supported by this switch module, please refer to the E-Lab Navigator..

- Interoperability:

  The I/O module supports Native Brocade M-Series Mode Interoperability.

Fabric services include:

- Security:

  The I/O module provides fabric security and interswitch link security. Fabric security controls management access to the fabric. When fabric security is enabled on all switches in the fabric, you must provide a user name and password to access the fabric. Security is enabled by default.

- Registered State Change Notification (RSCN):

  The I/O module supports RSCN as described in FC-FLA. RSCN enables an agent to register for change events in the fabric and attached devices.

- Error detection: The I/O module supports the following error detection methods
  - Cyclic redundancy check (CRC)
  - 8-byte and 10-byte conversion
  - Parity
  - Long frame and short frame
  - D_ID mismatch
  - S_ID mismatch

- Frame bundling:

  The I/O module provides the following frame bundling methods: - No frame bundling - Intermix frames from different sources at will. - Soft lockdown - Soft lockdown causes the I/O module to wait for either the sequence to be completed or a gap in the frame traffic to occur before servicing requests from a different port.

- Configurable Fabric Address Notification (FAN):

  A FAN, as described in FC-FLA, is sent out by the fabric after an initialization event (usually a loop initialization port) to each attached NL_Port. The purpose of the FAN is to inform the attached NL_Ports of the fabric node name, port name, and fabric address.

- Simple name server implementation:

  The simple name server is implemented as described in Fibre Channel Generic Services (FC-GS-3). The simple name server requests and responses are based on the Common Transport Interface (CTI) as described in FC-GS-3. Name server database objects are defined as follows:

  - Native Port Identifier (P_ID)
  - Port Name (PN)
  - Node Name (NN)
  - Class of Service (CoS)
  - Internet protocol (IP) Address (IP_A)
  - Initial Process Associator (IPA)
  - FC-4 Types (Type) and Port Type (PT)
  - Symbolic Port Name (SPN)
  - Symbolic Node Name (SNN)

Performance

2.125 Gb/s non-blocking throughput on all external ports.

Management

- Simple Network Management Protocol (SNMP)/Management Information Base (MIB) IBM Confidential Version 1.1
- Telnet/Command Line Interface (CLI)
- Web Browser Management Interface
- Standalone Switch Management Application
- API Interface
- IBM Director Deployment Wizard Plug-in
- Support for Non-Disruptive Code Load Activation (NDCLA)

Mechanical specifications

Mechanical specifications include:

- Height: 112mm (4.4 inches)
- Width: 260mm (10.2 inches)
- Depth: 29mm (1.14 inches)

IBM PN 32R1833 Brocade M-Series 6-port SAN switch module

The 6-port Brocade M-Series Fibre Channel (FC) switch enables 4 Gb SAN solutions. It is ideal for the small to medium sized storage area networks and provides easy integration of BladeCenter into Brocade M-Series environments.

ASIC family Agate-2

Key features Connectivity feature includes:

- 20 ports:
  - 14 internal 2/4 Gb/s ports
  - 6 external 1/2/4 Gb/s auto-sensing ports.
- Supports 239 domains.
- Interoperability:

  The I/O module supports Native Brocade M-Series Mode Interoperability.
- ISL trunking is supported on these modules. Up to 4 ports can be used for an ISL trunk and they need not be sequential ports.

Fabric services include:

- Security:

  The I/O module provides fabric security and interswitch link security.
- Registered State Change Notification (RSCN):

  The I/O module supports RSCN as described in FC-FLA. RSCN enables an agent to register for change events in the fabric and attached devices.
- Error detection: The I/O module supports the following error detection methods.
  - Cyclic redundancy check (CRC)
  - 8-byte and 10-byte conversion
  - Parity
  - Long frame and short frame
  - D_ID mismatch
  - S_ID mismatch

- Frame bundling: The I/O module provides the following frame bundling methods:
  - No frame bundling
    Intermix frames from different sources at will.
  - Soft lockdown
    Soft lockdown causes the I/O module to wait for either the sequence to be completed or a gap in the frame traffic to occur before servicing requests from a different port.
- Configurable Fabric Address Notification (FAN):

  A FAN, as described in FC-FLA, is sent out by the fabric after an initialization event (usually a loop initialization port) to each attached NL_Port. The purpose of the FAN is to inform the attached NL_Ports of the fabric node name, port name, and fabric address.
- Simple name server implementation:

  The simple name server is implemented as described in Fibre Channel Generic Services (FC-GS-3). The simple name server requests and responses are based on the Common Transport Interface (CTI) as described in FC-GS-3. Name server database objects are defined as follows:
  - Native Port Identifier (P_ID)
  - Port Name (PN)
  - Node Name (NN)
  - Class of Service (CoS)
  - Internet protocol (IP) Address (IP_A)
  - Initial Process Associator (IPA)
  - FC-4 Types (Type) and Port Type (PT)
  - Symbolic Port Name (SPN)
  - Symbolic Node Name (SNN)

Performance 4.25 Gb/s

Management
- Simple Network Management Protocol (SNMP)/Management Information Base (MIB) IBM Confidential Version 1.1
- Telnet/Command Line Interface (CLI)
- Web Browser Management Interface

- Connectrix Manager
- API Interface
- IBM Director Deployment Wizard Plug-in
- Non-disruptive firmware upgrades

Mechanical specifications

Mechanical specifications include:

- Depth 1.14 in
- Height 4.4 in
- Width 10.2 in

IBM PN 32R1905 Brocade M-Series 3-port SAN switch module

This 3-port Brocade M-Series Fibre Channel (FC) switch enables 4 Gb SAN solutions. It is ideal for the small to medium-sized storage area networks and provides easy integration of BladeCenter into Brocade M-Series environments.

ASIC family Agate-2

Key features Connectivity feature includes:

- 10 ports:
  - 7 internal 2/4 Gb/s ports
  - 3 external 1/2/4 Gb/s auto-sensing ports.
- Supports 239 domains.
- Interoperability:

  The I/O module supports Native Brocade M-Series Mode Interoperability.
- ISL trunking is supported on these modules. Up to 4 ports can be used for an ISL trunk and they need not be sequential ports.

Fabric services include:

- Security: The I/O module provides fabric security and interswitch link security.

- Registered State Change Notification (RSCN):

  The I/O module supports RSCN as described in FC-FLA. RSCN enables an agent to register for change events in the fabric and attached devices.

- Error detection: The I/O module supports the following error detection methods:
  - Cyclic redundancy check (CRC)
  - 8-byte and 10-byte conversion
  - Parity
  - Long frame and short frame
  - D_ID mismatch
  - S_ID mismatch

- Frame bundling:

  The I/O module provides the following frame bundling methods: - No frame bundling - Intermix frames from different sources at will. - Soft lockdown - Soft lockdown causes the I/O module to wait for either the sequence to be completed or a gap in the frame traffic to occur before servicing requests from a different port.

- Configurable Fabric Address Notification (FAN):

  A FAN, as described in FC-FLA, is sent out by the fabric after an initialization event (usually a loop initialization port) to each attached NL_Port. The purpose of the FAN is to inform the attached NL_Ports of the fabric node name, port name, and fabric address.

- Simple name server implementation:

  The simple name server is implemented as described in Fibre Channel Generic Services (FC-GS-3). The simple name server requests and responses are based on the Common Transport Interface (CTI) as described in FC-GS-3. Name server database objects are defined as follows:
  - Native Port Identifier (P_ID)
  - Port Name (PN)
  - Node Name (NN)
  - Class of Service (CoS)
  - Internet protocol (IP) Address (IP_A)
  - Initial Process Associator (IPA)

- FC-4 Types (Type) and Port Type (PT)
- Symbolic Port Name (SPN)
- Symbolic Node Name (SNN)

Performance 4.25 Gb/s

Management

- Simple Network Management Protocol (SNMP)/Management Information Base (MIB) IBM Confidential Version 1.1
- Telnet/Command Line Interface (CLI)
- Web Browser Management Interface
- Connectrix Manager
- API Interface
- IBM Director Deployment Wizard Plug-in
- Non-disruptive firmware upgrades

Mechanical specifications

Mechanical specifications include:

- Depth 1.14 in
- Height 4.4 in
- Width 10.2 in

Blade servers

A blade server is a server chassis housing a number of individual minimally-packaged computer motherboard *blades*, each including one or more processors, computer memory, computer storage, integrated network controllers, an optional fiber channel host bus adapter (HBA) and other input/output (IO) ports, but sharing a common power supply and air-cooling resources. The motivation behind the evolution of blade servers is to allow more processing power in less rack space, simplifying cabling and reducing power consumption.

A blade typically comes with one or two local drives. For additional storage, blade servers can connect to a storage pool facilitated by a Fibre Channel or iSCSI- based storage area network (SAN) or by network attached storage (NAS).

The advantages of blades do not lie merely in housing several servers in a single chassis, but also from the consolidation of associated resources (like storage and network equipment) into a smaller architecture that can be managed through a single interface.

Blade server architecture

This section reviews the generic blade server architecture.

Blade server chassis

The individual blades in a blade server chassis are typically hot-swappable. One of the greatest advantages of blade servers is that they allow the use of a single reliable heavy-duty DC power supply, rather than many small and unreliable power supplies. However, since a single power supply is a single point of failure, manufacturers recommend using a minimum of two power supplies per enclosure.

Blade servers use custom-designed blade boards, rather than commodity PC motherboards, so can be designed to have significantly more efficient air-cooling airflow than a rack of servers.

The server blades connect to vendor-implemented internal buses on a shared backplane or midplane in the chassis. The internal buses serve to provide connectivity between the blades to all of the common functions it requires, such as KVM (Keyboard-Video-Mouse) console access, IP networking, and FC interconnect for storage.

The blades draw DC power using the backplane from shared power-supply modules in the chassis. The cooling function is also consolidated in the chassis through dual-redundant fan modules.

Figure 493 shows a simple model of a blade server with all its components.

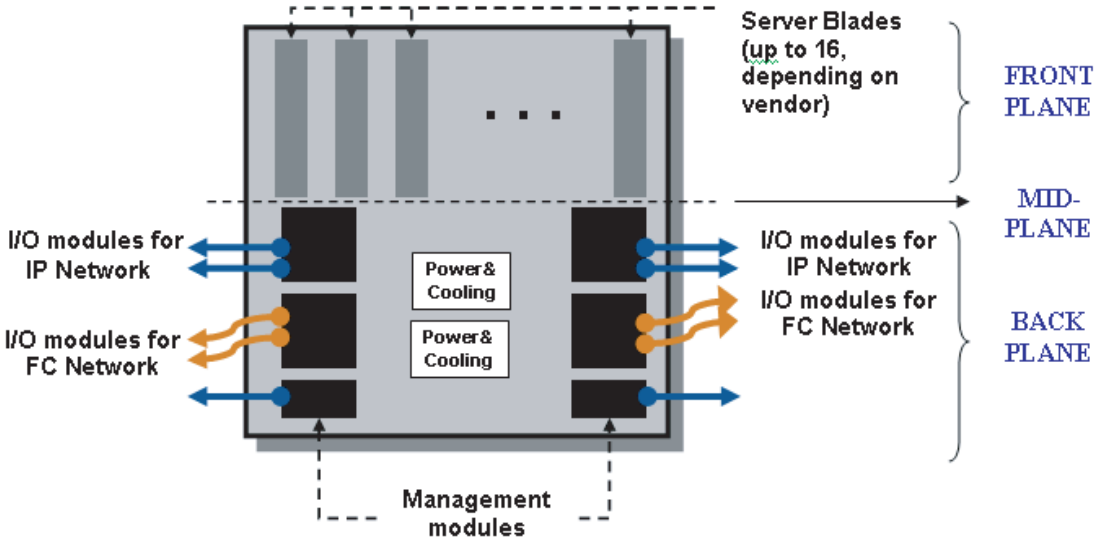


Figure 493 Basic blade server architecture

Figure 494 shows the front panel of a Dell blade server with 10 server blades.

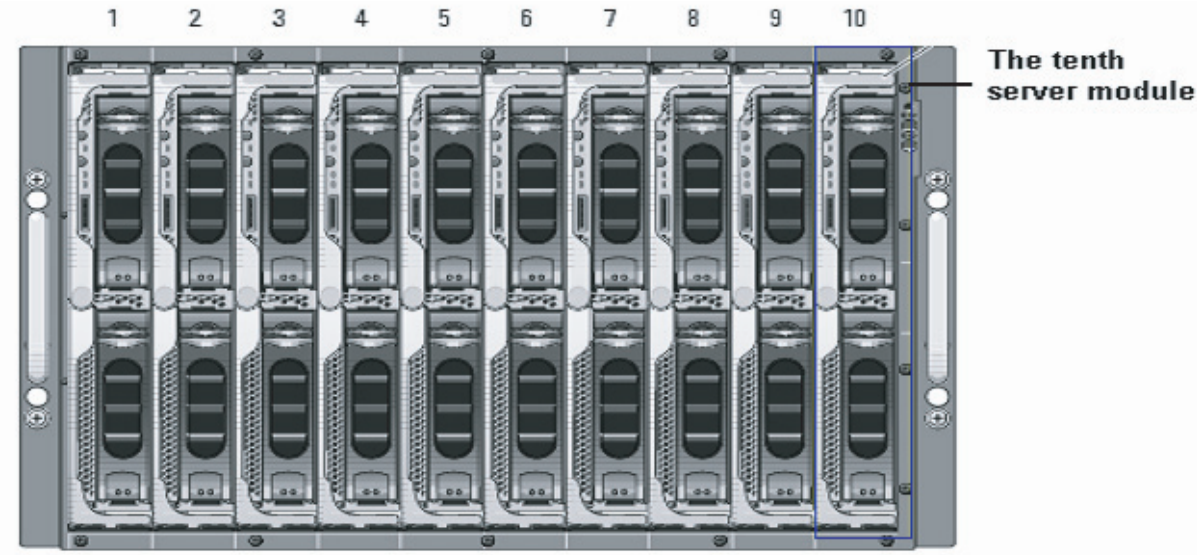


Figure 494 Front plane of a Dell blade server with 10 server blades

Figure 495 on page 1299 shows the back plane of a Dell blade server showing the power supply modules, the Fan modules, I/O bays for the Ethernet and FC switch modules, and the Management module (DRAC/MC module).

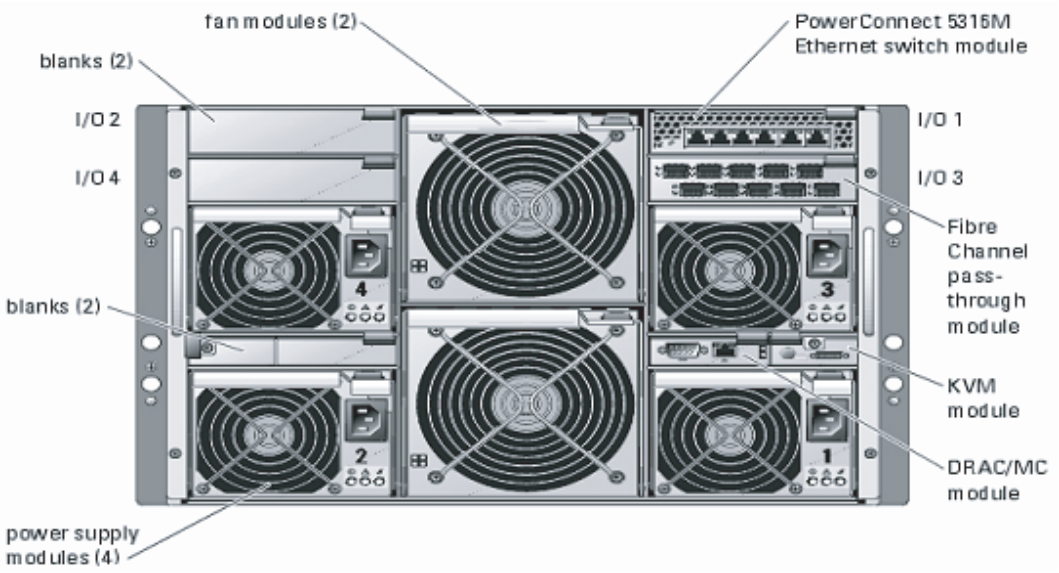


Figure 495 Back plane of a Dell blade server with 10 server blades

Server blades

At the time of this publication, a single chassis accommodates up to 16 processor blades, depending on vendor model. Processor blades available today mostly use Intel or AMD processors (i86 architecture).

Blades can accommodate up to two internal disk drives (either ATA or SCSI). Typically the operating system is installed on two local drives that are hardware-mirrored. It is possible to use a diskless blade if deploying boot over SAN.

A mezzanine card, also referred to as a *daughter card*, may be installed as an option inside the server blade. This card typically goes into a PCI-X or PCI-Express slot in the server and provides two internal ports. Each port connects through internal bus to one of a suitable I/O module pair. A given mezzanine card has the ability to provide either a GigE networking functionality, Infiniband, or FC Connectivity, in which case it is equivalent to an FC HBA.

Figure 496 is an example of a server blade showing a highlighted view of some ports, indicators, and buttons generally seen on a blade.

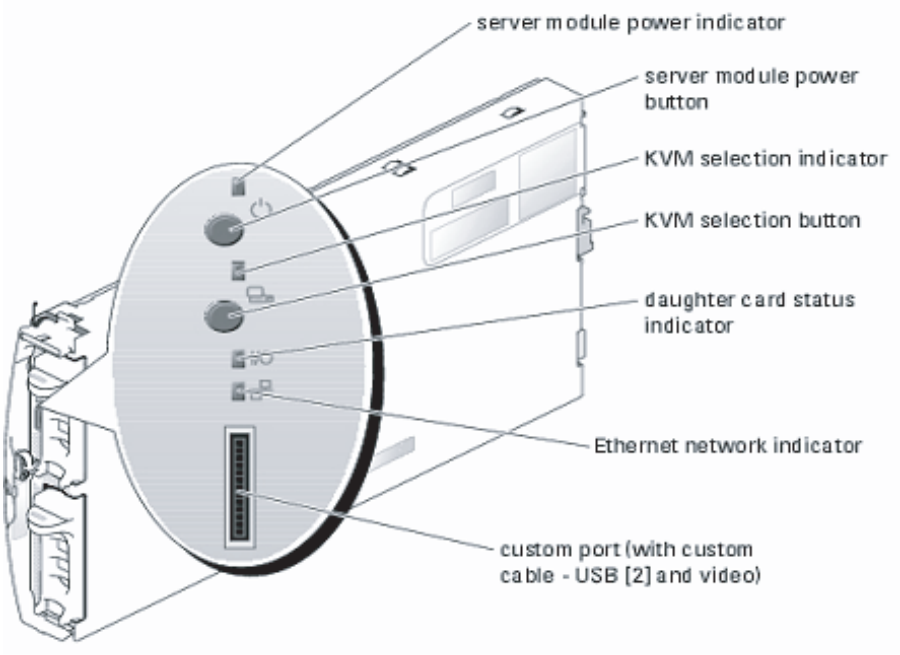


Figure 496 Server blade example

Table 92 on page 1301 explains the functionality or indicator codes of the different buttons and indicators which are common to most of the EMC supported blade servers. Please note that the ones that are specific to the example in the figure have not been listed as a part of this publication.

Table 92 Indicator codes for blade servers

Indicator	Icon	Activity indicator	Indicator code
Server module power indicator	⏻	Off	Power is not available to the server module, the server module is not turned on, or the server module is installed incorrectly.
		Green	The module is turned on.
		Green blinking fast	The module power is on and there is a fault with the server module.
		Amber	The module power is off, but the system power is on.
		Amber blinking fast	The module power is off and there is a fault with the server module.
Server module power button	N/A	None	Turns server module power off and on. • If you turn off the module using the power button and the module is running an ACPI-compliant operating system, the module can perform an orderly shutdown before the power is turned off. • If the module is not running an ACPI-compliant operating system, power is turned off immediately after the power button is pressed. • Press and hold the button to turn off the server module immediately. The button is enabled in the System Setup program. When disabled, you can only use the button to turn on the server module.
KVM selection indicator		Off	The server module is not selected by the KVM.
		Green	The server module is selected for the KVM.
		Amber blinking	The server module is not selected by the KVM and a power fault exists.
		Green/amber blinking	The server module is selected for the KVM and a power fault exists.
KVM selection button	N/A	None	Selects the server module for use with the KVM located on the back of the system.
Daughter card status indicator	I/O	Off	Power is off or the signal is lost.
		Green	Power is on or signal is online.
		Green blinking	A firmware error exists.

Table 92 Indicator codes for blade servers (continued)

Indicator	Icon	Activity indicator	Indicator code
Network indicators	OffIndicates that the server module does not have a link to the Ethernet switch or pass-through module.	Off	Indicates that the server module does not have a link to the Ethernet switch or pass-through module.
		Green on	Indicates that the server module has a valid link to the network switch module.
		Green blinking	Indicates network activity between the server module and the network switch module. **Note:** External network activity is not reported by this indicator.

Management and KVM modules

A management module provides a common console (KVM: keyboard-video-mouse) access to each of the individual blades. It also allows for chassis connection to the external network. In some designs, IP configuration for server blades and for installed I/O modules must be done using chassis utilities, which can be accessed through the management module. They are usually installed as a dual-redundant pair. There are vendor-specific applications for chassis management.

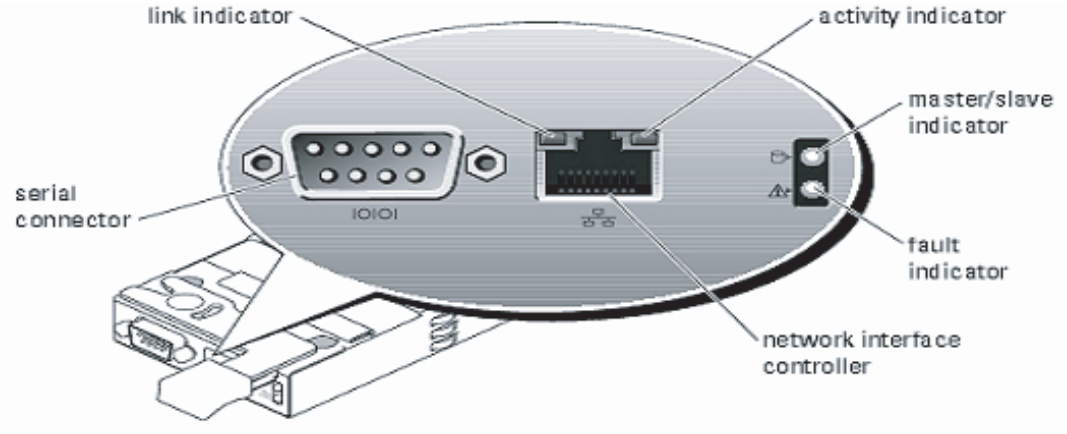


Figure 497 Management module example

Note: Some of the indicators shown in Figure 497 on page 1302 are specific to the example in the figure and do not apply to all management modules in general. Thus, their functionality and indicator codes have not been listed as a part of this publication.

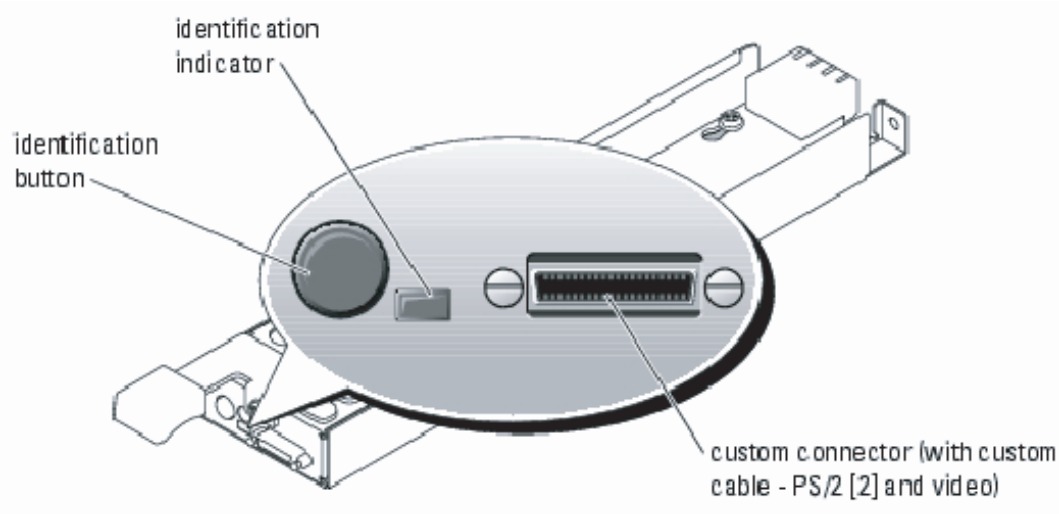


Figure 498 **KVM module example**

Table 93 explains the functionality or indicator codes of the different buttons and indicators which are common to most of the EMC management modules. Please note that the ones that are specific to the example in Figure 498 have not been listed as a part of this publication.

Table 93 **Indicator codes for management modules**

Indicator	Icon	Activity indicator	Indicator code
Identification button	**ID**	None	Press to activate the front and back identification indicators. To turn off the identification feature, press the button again.
Identification indicator		Off	System has not been selected for identification.
		Amber	System is selected for identification.
		Amber blinking	System is being identified.

I/O modules

The main functionality of I/O modules is to provide external port connectivity to the blade servers. I/O modules are always installed in pairs, in specific I/O bays within the chassis.

To provide dual-pathed FC SAN connectivity to servers, the following components must be installed:

- At least one mezannine card / FC Host Bus Adapter card (HBA) in each blade server to be connected must be installed.
- A pair of FC SAN switch modules can be installed. At the time of publication, these embedded FC SAN switch modules are manufactured mainly by Brocade and QLogic. They have different modes of operation depending on the external fabric components they are connected to, and have specific switch firmware revisions that can be supported depending on the manufacturer and blade server vendor. The EMC-supported FC SAN switch modules have been listed with their details in "Blade servers supported by EMC" on page 1309.
- Alternatively, one can also install a pair of supported optical pass-through modules (Figure 499 on page 1305) which provides the external fabric components with direct access to server blades. Basically, the embedded optical pass through modules enable daughter cards to access external switches without the need to place the FC SAN switch modules into the blade server switch bays.

As can be seen, the optical pass-through modules are comparatively much more flexible than the FC SAN switch modules and can be used to hook up the blade server modules to any external standalone switch in a fabric, independent of the firmware and mode running on that switch. The only factors that would affect this connectivity are the type of HBA (manufacturer, driver revisions etc.) or the operating system on the server blade.

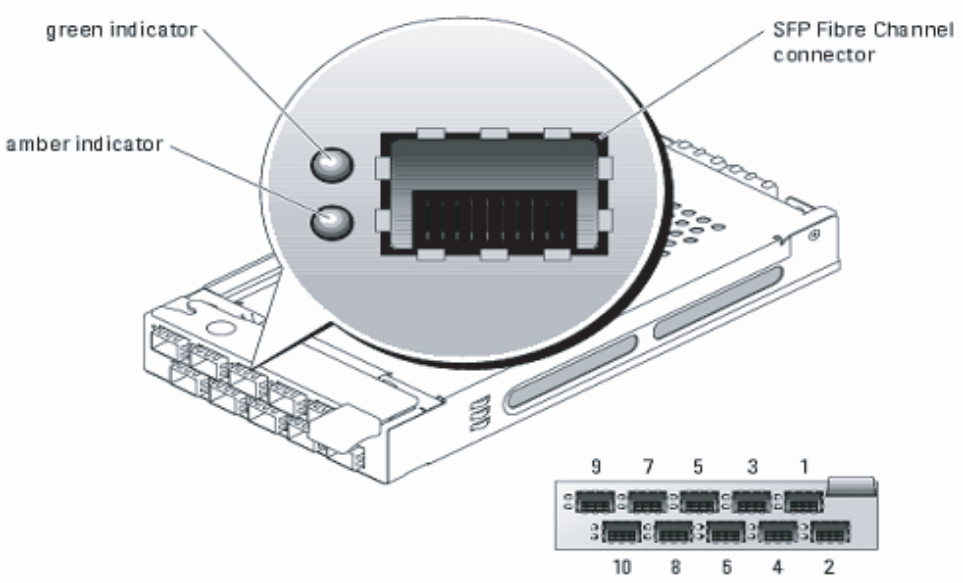


Figure 499 Pass-through example

Table 94 explains the functionality or indicator codes of the different buttons and indicators which are common to most of the EMC supported I/O modules. Please note that the ones that are specific to the example in Figure 499 have not been listed as a part of this publication.

Table 94 Indicator codes for I/O modules

Indicator type	Activity indicator	Indicator code
Fibre Channel indicator (green/amber)	Off	Power is off to the system.
	Green/amber	System has power.
	Green/off	Fibre Channel connection is online.
	Off/amber	The port is connected to a valid link partner on the network.
	Off/flashing (twice per second)	Connection has lost synchronization.

Figure 500 depicts the FC connectivity mechanism. Each mezzanine card provides two I/O ports that connect internally (through buses on the chassis backplane or midplane) to internal ports on the FC switch modules. Correspondingly, each switch module typically provides for one internal port per blade server.

Figure 500 also shows a pair of switch modules with just two external FC ports each. The FC Switch Modules (FCSMs) usually have 2, 4, or 6 external ports depending on the blade server vendor and switch module model.

As shown in Figure 500, a minimum of two fibre-optic cables are needed from the FC switch module on each side of the blade server chassis, to provide a dual-path connectivity to all the server blades in that chassis. Thus, Fabric A and Fabric B can access all the server blades through the respective FC switch modules to which they are connected.

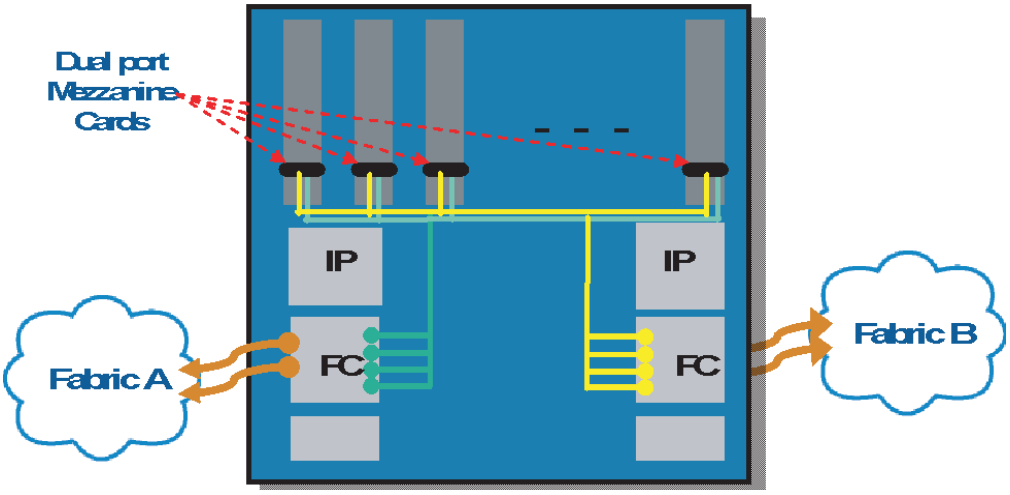


Figure 500 FC blade server – Switch module connectivity mechanism

Ethernet pass-through modules

An Ethernet pass-through module (shown in Figure 501 on page 1307) provides connectivity between the server modules and an external Ethernet device. The Ethernet pass-through modules generally have a certain number of uplinks and downlinks.

- The number of downlinks depends on the maximum number of server blades the chassis can accommodate, and hence the number of embedded Ethernet controllers. The downlinks operate at 1000 Mb only, in most cases. The downlinks are internal ethernet network connections and in most cases operate at 1000 Mb only.
- The uplinks connect to the external Ethernet network and operate at 10/100/1000 Mb.

Since each server blade or server module is equipped with at least one ethernet card, or rather two onboard Gibabit Ethernet ports, it is potentially ready for iSCSI.

The Ethernet modules generally go in I/O bays 1 and 2 in most blade server chasses.

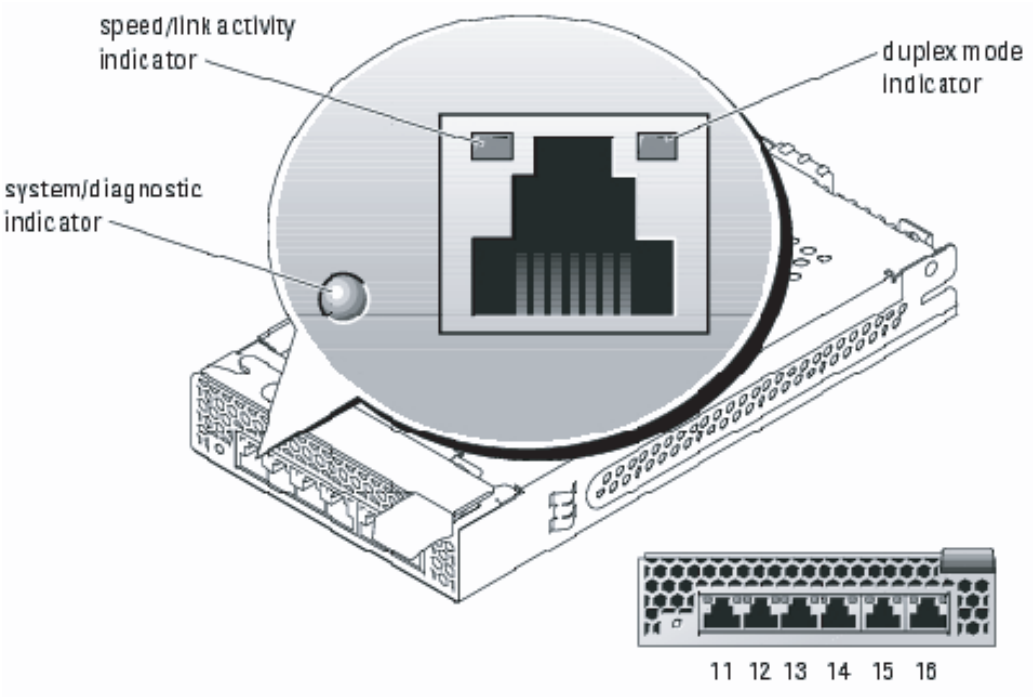


Figure 501 Ethernet switch module example

Power supplies/cooling

There are generally four hot-pluggable, redundant power supplies in a chassis, depending on the power requirement of the blade servers and other operational modules. In most cases, two to three power supplies provide power to the system, while the additional power systems provide redundancy. Figure 502 shows an example of a power module.

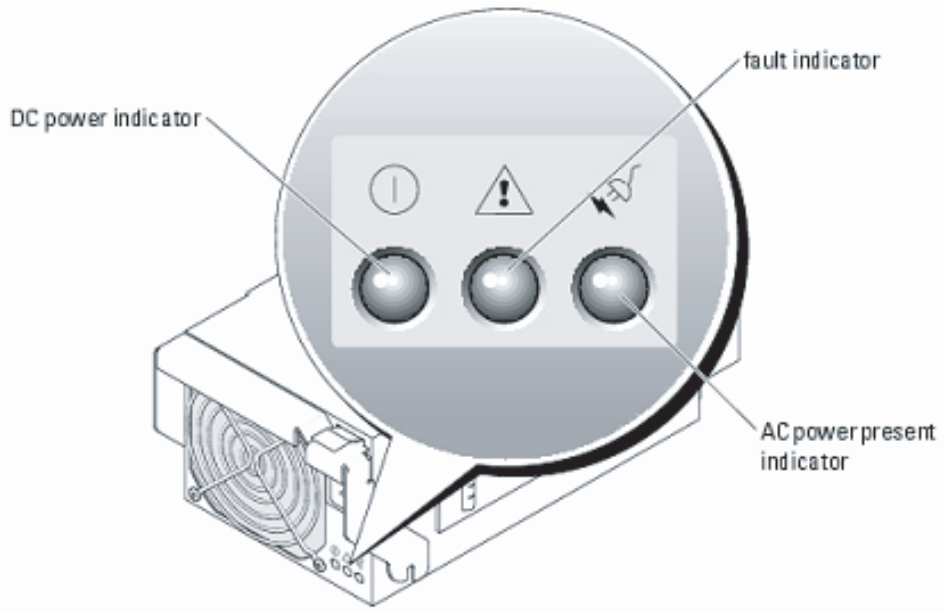


Figure 502 Power module example

In addition to providing power to the system, these power supply modules assist in cooling the system. If a power supply fails, it is advisable to keep the power supply module in the chassis because the fans continue to cool the chassis. Similarly, every chassis usually has two exclusive hot-pluggable system fan modules. Each fan module further has two replacable fans.

Figure 503 shows an example of a fan module.

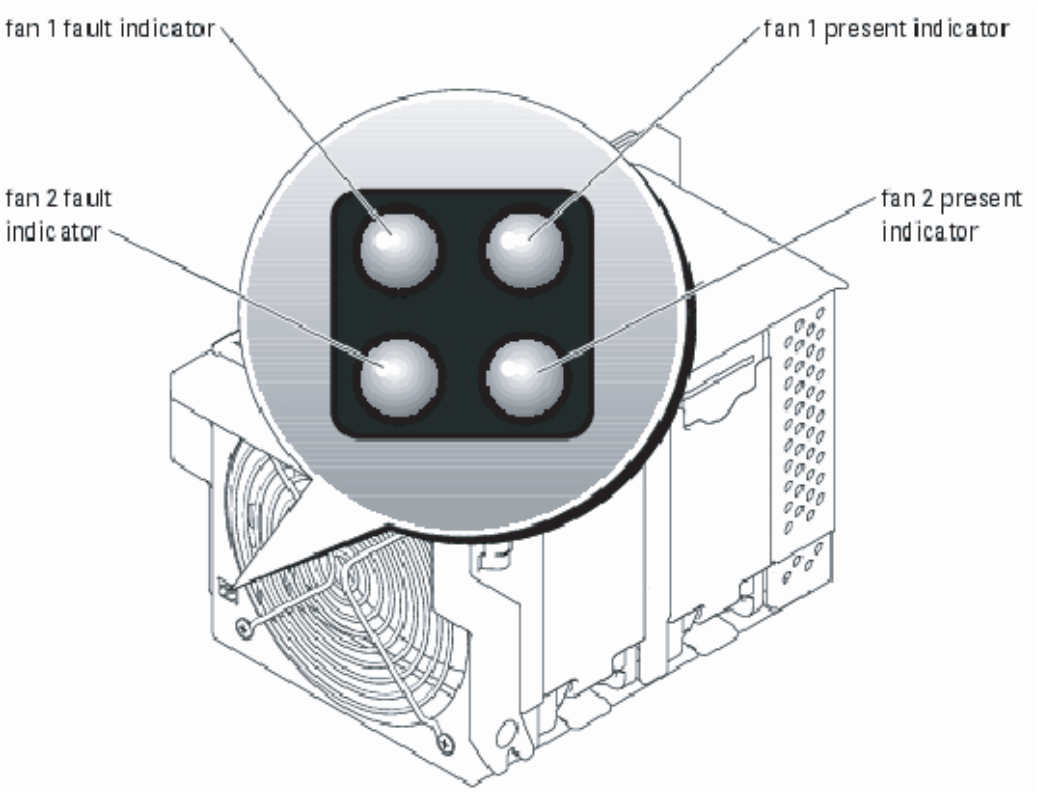


Figure 503 Fan module example

Blade servers supported by EMC

EMC qualifies blade servers manufactured by the following vendors

- Dell
- HP
- IBM
- Intel, and models based on the Intel architecture manufactured by
  - Fujitsu Siemens
  - Digital Henge
  - Lang Chao

Delll, HP, and IBM blade servers

Technical specifications for EMC-supported Dell, HP, and IMBM servers are listed in the Table 95, Table 96, and Table 97.

Table 95 Processor information

Dell blade servers	HP blade servers	IBM blade servers
The Dell blade server (Dell Glacier) or the Poweredge 1855 chassis is a 7U-high cabinet, which supports up to 10 server blades. Each blade server can house up to two Intel Xeon processors.	The HP blade server is a 6U-high cabinet, which supports anywhere from 2 to 16 server blades. There are different types of server blades, which can house 2 to 4 Intel or AMD processors, depending on the type of server blades.	The IBM blade server is a 7U-high cabinet. There are two families of chasses: • The standard BladeCenter can house 14 server blades. • The BladeCenter T can house 8 server blades. Also, there are different type of server blades that can house 2 to 4 Intel or AMD processors ,depending on the type of server blades.

The operating systems that EMC supports on these blade servers are Windows, Linux flavors (RHEL and SuSE), and VMWare.

Table 96 Mezzanine cards

Dell blade servers	HP blade servers	IBM blade servers
All supported mezzanine cards can be installed in a PCI-X slot on the blade server.	The BL20 (G2 and G3), BL25, BL30, BL35 and BL45 blades do not provide a PCI-X slot for mezzanine cards. The dual-ported FC mezzanine cards are installed directly on the motherboard. The cards requiring Qlogic firmware and drivers are the HPQ Dual-port Fibre Channel Adapters for HP Proliant BL20p(G2)/20p(G3)/25p/45p/30p/35p Emulex mezzanine cards are also available and supported for some blades, e.g., the Emulex-based BL25/30/35/45p Fibre Channel Mezzanine HBA, model 394588-B21.	All supported mezzanine cards install in a PCI-X slot on the blade server. The EMC-qualified card models are as follows: • IBM Emulex LP1005DC FC Expansion Card • IBM HS20 4 Gb SFF FC Expansion Card: ***model 26R0890*** • IBM HS20 4 Gb Standard FC Expansion Card: ***model 26R0884*** • IBM HS20 FC Expansion Card: ***model 13N2203*** • IBM HS20 FC Expansion Card: ***model 26K4841*** • IBM HS20 FC Expansion Card: ***model 48P7061***

Check the E-Lab Navigator for currently-qualified blade / mezzanine combinations, host operating system, and card driver versions.

Table 97 FC switch modules

Dell blade servers	HP blade servers	IBM blade servers
EMC supports the following FC switch modules for the Dell Poweredge 1855: • 4-port 2 Gb Brocade FC switch module **(SW3014)** • 4-port 4G b Brocade FC switch module **(SW4016)** • 4-port 2 Gb Brocade M-Series FC switch module **(SW4314)** • Optical Pass-thru module **(J6780)** • 4-port 4 Gb Brocade M-Series FC switch module **(SW4416)**	EMC supports the following FC switch modules for the HP blade server: • 4-port 4 Gb Brocade FC switch module **(Power Pack model A7535A)** • 2-port 4 Gb Brocade FC switch module **(A8001)**	EMC supports the following FC switch modules for the IBM blade server: • 2-port 2 Gb Brocade Entry level FC SAN switch module **(26K5601)** • 2-port 2 Gb Brocade Enterprise level FC SAN switch module **(90P0165)** • 3-port 4 Gb Brocade FC SAN switch module **(32R1813)** • 6-port 4 Gb Brocade FC SAN switch module **(32R1812)** • 6-port 2 Gb Brocade M-Series FC SAN switch module **(32R1790)** • 2-port 4 Gb Qlogic FC SAN switch module **(48P7062)** • 6-port 4Gb Qlogic FC SAN switch module **(26K6477)** • 3-port 4 Gb Brocade M-Series FC SAN switch module (**32R1905**) • 6-port 4 Gb Brocade M-Series FC SAN switch module (**32R1833**)

Important: Each of these switch modules behave like standalone edge switches in a deployed core-edge configuration.

Intel blade server

The architecture of the Intel blade server, is similar to the IBM Blade server. It is a 7U-high cabinet and can house 14 server blades. The Intel FC switch modules SBCEBFCESW and SBCEBFCSW listed in the E-Lab Navigator are the entry and enterprise versions of the Brocade switch modules. These are similar in characteristics to the IBM Brocade SAN switch modules 26K5601 (refer to "IBM PN 26K5601 Brocade 2-port entry-level switch module" on page 1279) and the 90P0165 (refer to "IBM PN 90P0165 Brocade 2-port enterprise-level switch module" on page 1280). The switch modules from other vendors, such as Fujitsu Siemens, Digital Henge, and the Lang Chao are based on the Intel Blade servers.

Modes supported on the switch modules

The Brocade switch modules for blade servers behave and can be managed like any other Brocade standalone edge switch. Hence, they can operate in either of the two Brocade modes: native mode (interopmode 0) and interoperable mode (interopmode 1).

EMC supports all the switches mentioned in Table 97 on page 1311 for native mode support with most of the Brocade switch families and for interoperable mode support with standalone Brocade M-Series and Cisco switches. A detailed listing of the supported firmware revisions and other specifics used to manage the fabric comprising of the switch modules, is in the E-Lab Navigator.

Table 98 gives a brief overview about the various interoperable configurations supported for the Brocade switch modules. The table also highlights the mode in which the switch / switch module operates, and the fabric management application that EMC recommends for the specific configuration.

Table 98 Supported Interoperable configurations

Switch module	Interoperable stand-alone switch	Fabric management application
Vendor: Brocade Mode: interopmode 0	Vendor: Brocade Mode: interopmode 0	Brocade Fabric Manager
Vendor: Brocade Mode: interopmode 1	Vendor: Brocade M-Series Mode: Open Fabric Mode	EMC Connectrix Manager
Vendor: Brocade Mode: interopmode 1	Vendor: Cisco Mode: interop-1	EMC Connectrix Manager or Cisco Fabric Manger
Vendor: Brocade M-Series Mode: Brocade M-Series mode	Vendor: Brocade M-Series Mode: Brocade M-Series mode	SANsurfer switch manager for individual switch management and EFCM for the Fabric Management
Vendor: Brocade M-Series Mode: Standard mode	Vendor: Brocade M-Series Mode: Open Fabric mode	SANsurfer switch manager for individual switch management and EFCM for the Fabric Management

Further reading

The following documentation is available for additional product information:

- IBM blade server related information can be found at:

  www.ibm.com/systems/bladecenter/

- HP blade server related information can be found at:

  h71028.www7.hp.com/enterprise/ cache/80316-0-0-0-121

- A good reference for HP Brocade switch module related information can be found at:

  h18006.www1.hp.com/products/quickspecs/12231_div/12231_div.html

- Dell blade server (Poweredge 1855) related information can be found at:

  http://www1.us.dell.com/content/products/productdetails.aspx/pedge_1855?c=us&cs=555&l=en&s=biz

- http://www.brocade.com
- http://www.Brocade M-Series.com
- http://www.qlogic.com

Brocade Access Gateway

The Access Gateway feature is a software feature that allows a Brocade switch to act as a transparent "port consolidator," connecting multiple Fibre Channel devices (typically host HBA ports) into a single FC fabric port (F_Port). Enabling the Access Gateway mode on a standalone switch or switch module presents all its ports or external ports (in a module) as N_Ports. The Access Gateway connects to the fabric using the FC N_Port protocol rather than the E_Port protocol typically used for switch-to-switch communication. Access Gateway uses the FC-LS N_Port ID Virtualization (NPIV) facility to allow its client devices to connect to the FC fabric.

Although it uses its switching logic to allow the attached devices to send and receive FC traffic to the fabric, it is not recognized by the fabric as a "switch" and therefore does not consume a Domain ID and does not participate in the switch-to-switch traffic normal to an FC fabric.

The Access Gateway provides the following benefits over an embedded switch:

- Since it requires minimal management operations, it reduces management complexities on large blade server fabrics.
- The use of N_Port protocol allows connection to non-Brocade fabrics with reduced interoperability challenges compared to E_Port connections.
- Improved fabric scalability for fabrics with large numbers of embedded blades or small edge switches.

These benefits are further detailed and explained in the following case studies in this section: "Case study #1" on page 1321; "Case study #2" on page 1330; and "Case Study #3" on page 1339).

Hardware and software requirements for Access Gateway

The Access Gateway feature is currently supported for embedded blade server switch modules based on the Goldeneye. For a listing of the latest supported switch modules and firmware revs. please refer to the *EMC Support Matrix*. The Access Gateway feature has been introduced in the v5.2.1 release with a new operational mode, "Access Gateway mode" (also known as "agmode").

Note: Enabling or disabling agmode on a switch can be performed using CLI only. There is no Web Tools option for agmode in this release.

The Access gateway uses the NPIV feature to connect a switch module to the enterprise fabric. The NPIV feature is a standard feature supported on the following fabric switches:

- Connectrix B switches with FOS 5.1.0 and higher
- Connectrix M switches with E/0S 8.0 and higher,
- Connectrix MDS switches with SAN-OS 3.0(1) and higher

Each of Access Gateway's N_Ports will operate correctly when connected to a fabric switch that properly supports the NPIV feature.

Access Gateway theory of operation

The Access Gateway mode of the switch presents standard F_Ports to the hosts, but it connects to the Enterprise fabric as an N_Port (rather than as an E_Port). The purpose of Access Gateway mode is to isolate the enterprise fabric and provide controlled access to it from hosts attached.

Figure 504 shows the “n “server blades attached to the external Fabrics A and B via the blade server embedded Brocade Fibre Channel Switch Modules (FCSM). In this case, all the external ports on the module are E_Ports and their connectivity to the external FC switch is an E_Port to E_Port connection.

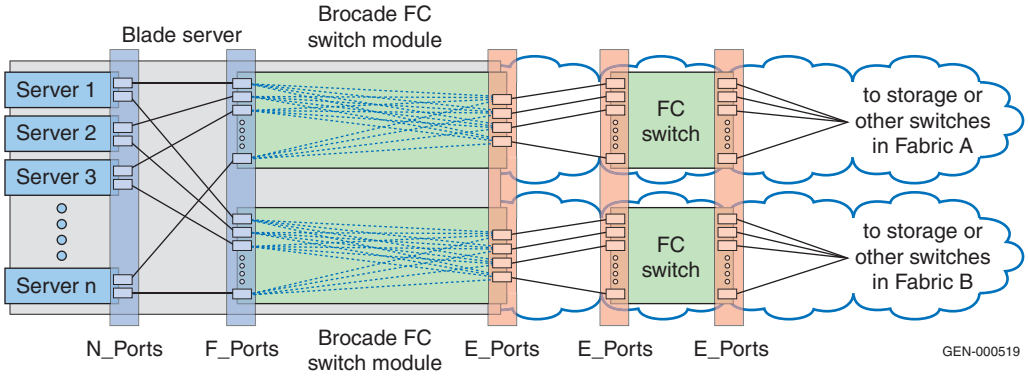


Figure 504 Blade servers using Fibre Channel switch modules

Figure 505 shows the "n" server blades now attached to the external Fabrics A and B via the blade server embedded Brocade Fibre Channel Switch Modules (FCSM) with Access Gateway mode enabled. In this case, all the external ports on the module are N_Ports and their connectivity to the external FC switch is an N_Port to F_Port connection. What we can also see is how one or more F_Ports can be mapped to a single N-port on the Access Gateway module.

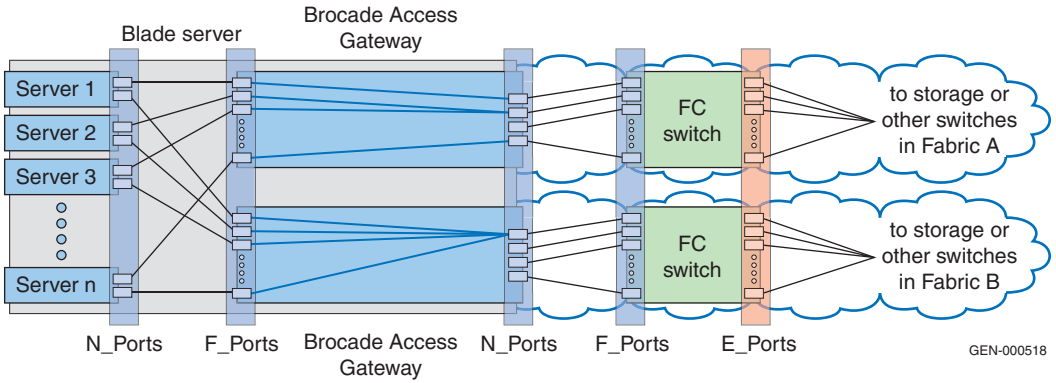


Figure 505 Blade servers using Acess Gateway

Firmware upgrade and downgrade considerations

When Access Gateway mode is enabled for a switch, it will support hot code activation in the same way as base FOS. The code load is non-disruptive.

All firmware downloads are done out-of-band, using CLI. Since v5.2.1 is the first firmware capable of supporting agmode, firmware downgrade is *not* supported when agmode is enabled. Firmware downgrade can be performed only after disabling agmode.

Manageability considerations

Access Gateway management is supported through CLI beginning with the v5.2.1 release. Web Tools and Fabric Manager support are *not* supported in the v5.2.x release. A set of new CLIs have been provided to perform Access Gateway-specific operations (refer to "Access Gateway CLI commands" on page 1318).

N_Port ID Virtualization (NPIV) support

When agmode is enabled, NPIV is automatically enabled and no NPIV license is needed. It cannot be disabled by the user. Enterprise Connectrix-B fabric can support up to 255 PIDs to be

assigned to an N_Port using NPIV. The actual number depends on the edge switch configuration and capability. The order in which hosts come online can be different every time Access Gateway is enabled. To make sure all the hosts get a predictable number of PIDs assigned using NPIV, the default value of the maximum number of logins that will be accepted by an F_Port is set to 20. This value can be changed using the **configure** CLI command.

PID assignment by AG

PIDs assigned by Access Gateway to a host or sever blade (with or without NPIV) can change any time the host or server blade goes offline and comes back online. The host operating system should be capable of handling PID changes. Otherwise, operating systems' specific steps need to be followed to update the host configuration to use new PID(s).

F_Port enable/disable

F_Port enable does not cause any link up or link down of the N_Ports. A link down event, or portdisable on an F_Port, will cause a logout of that F_Port on the enterprise fabric. If NPIV is being used on that F_Port, all the NPIV devices that have logged in to the enterprise fabric from that port will be logged out from the enterprise fabric.

N_Port enable/disable

If N_Port is disabled or the N_Port link goes down, all F_Ports that are mapped to that N_Port will be disabled. If N_Port failover is enabled for this N_Port, then the F_Ports mapped to this N_Port will be mapped to another N_Port connected to the same fabric. When the original N_Port becomes online, the F_Ports mapped to it will not failback to it (if the failover of those F_Ports to another N_Port was successful) because N_Port failover is disruptive for the F_Ports. PIDs assigned to the F_Ports get modified as a result of this failover. If the host drivers can handle the PID change transparently, no operator intervention or reboot will be required for the host.

N_Port Failover policy

N_Port failover will be supported only within N_Ports connected to the same fabric. When a port is configured as an N_Port, failover is enabled automatically. When an N_Port goes offline (cable removal or any other offline event), the F_Ports mapped to that N_Port are disabled. If another N_Port is connected to the same fabric is online, these F_Ports will be failed-over to it and re-enabled. The FLOGI and

FDISC requests will be forwarded from F_Ports through the new N_Port. Failover of F_Ports to next available N_Ports connected to the same fabric will continue until all the N_Ports are exhausted. If multiple N_Ports are available as candidates for failover, Access Gateway will select the port with lowest port number.

N_Port/F_Port bringup latency and error handling

When an N_Port is connected to the fabric, it will take some time to login to the enterprise fabric and get PID assigned. The upper bound on the time it takes to bring up an N_Port is 4 x E_D_TOV (e.g., 4 x 2 = 8 seconds if E_D_TOV is 2 seconds).

The amount of time required to bring up an F_Port depends on the timeout value used by the host between successive FLOGI attempts. Usually, one or more FLOGI attempts from the host are required before the F_Port comes online. The upper bound on the time it takes to bring up an F_Port is the maximum number of FLOGI retries (typically 4) multiplied by the time between consecutive retries (typically R_A_TOV = 10 seconds).

Access Gateway CLI commands

The Access Gateway CLI provides the following functionality:

- Enabling / disabling agmode
- Displaying current configuration and state of AG
- Configure / Display F_Port to N_Port mapping
- Configure N_Port failover / failback policies.

Table 99 on page 1319 lists CLI commands that have been provided to perform Access Gateway-specific operations.

Table 99 New CLI commands

Command	Action	Description
ag <action> [arguments]		
	--modeshow	Displays current agmode setting for the switch.
	--modeeenable	Enabled Access Gateway mode for the switch.
	--modedisable	Disable Access Gateway mode for the switch. Switch is rebooted automatically. Before agmode is disabled the switch needs to be disabled and securemode needs to be disabled (when enabling agmode).
	--show	Displays the current configuration and state of Access Gateway.
	--mapshow [N_Port]	Displays current F_Port to N_Port mapping. N_Port argument optional. If it is not specified, mapping for all the N_Ports is displayed, otherwise mapping for only the specified N_Port is displayed.
	--mapshow [N_Port] --mapset <N_Port> "<F_Port1;F_Port2; ...>"	Set specified list of F_Ports to the mapping for a given N_Port. It will overwrite the current mappings if any for that N_Port. An F_Port cannot mapped to more that one N_Port at the same time.
	--mapadd <N_Port> "<F_Port1;F_Port2; ...>"	Adds specified list of F_Ports to the mapping for a given N_Port. An F_Port cannot mapped to more than one N_Port at the same time.
	--mapdel <N_Port> "<F_Port1;F_Port2; ...>"	Removes specified list of F_Ports from the mapping for a given N_Port.
	--failovershow [N_Port]	If optional N_Port is specified, failover policy for N_Port is displayed. Otherwise failover policy for all the N_Ports is displayed.
	--failoverenable <N_Port>	Enable failover policy for the specified N_Port.
portcfgnport		To disable/enable N_Port capability for a port.
agshow		Display all Access Gateways in the Enterprise fabric.

Advantages of Access Gateway

Access Gateway addresses the following three major issues while maintaining many benefits of using the embedded switch model:

- "Scalability," next
- "Multivendor interoperability" on page 1320
- "Fabric management complexities" on page 1321

Scalability

When switches are added to or removed from a fabric, they tend to precipitate high volumes of interswitch management traffic causing CPUs of the constituent switches to get overloaded, which might result in data disruption due to fabric re-configurations. Fabrics also tend to become unstable while handling large volumes of fabric events. These problems may ultimately result in loss of service (e.g,. HBA logins may timeout) under heavy load conditions in the fabric.

The large number of embedded switch modules can cause fabric instability due to numerous domains, many more places to manage all of the switches in the fabric, and FSPF route calculations due to these events.

Access Gateway acts differently than a switch in the Fibre Channel fabric sense. For example, before we may have had a switch connecting 16 host ports to the fabric through 4 E_Port(s), whereas now we can have Access Gateway connecting 16 host ports to the fabric through 4 N_Ports. By expanding the number of device ports we can connect to a single fabric port, Brocade can now support a much larger fabric (in terms of the number of connected host and storage devices) with a dramatically reduced number of switches in the fabric. Thus, by not presenting an additional domain to the fabric, Access Gateway reduces the domain count when deployed in a fabric. Furthermore, its NPIV-based feature of presenting one N_Port representing multiple server N_Ports reduces the device port count (a concern when resolving scalability issues). Fewer domains increase fabric reliability, stability, and availability.

Multivendor interoperability

Access Gateway connects to the enterprise fabric as N_Port instead of E_Port, therefore interoperability with non-Brocade fabric will be much easier. Whereas E_Port interoperability requires exact compliance with a very large, complicated, and imprecisely defined interswitch protocol suite, the N_Port connects to the enterprise fabric using only the precisely defined FLOGI, FDISC, and LOGO

requests, ensuring compatibility with other Enterprise switch vendors' products.

Fabric management complexities

Access Gateway connects to Enterprise fabric N_Port, making it more of a device management tool instead of a fabric management tool. SAN management activities (e.g., zoning) remain entirely in the external fabric and does not include the switch module within the blade server chassis.

Case studies: How to set up an Access Gateway fabric and its benefits

This section contains the following three case studies for setting up an Access Gateway fabric and its benefits:

- "Case study #1," "Connectrix series switches with NPIV enabled linked through ISLs to Brocade Access Gateway" on page 1321
- "Case study #2," "Migration from Brocade module (native) ISLed to Brocade switches (native) to Brocade module (in AG mode) connected to Cisco switches (interopmode irrelevant)" on page 1330
- "Case Study #3," "Addition of switches or server blades in an existing Connectrix-M fabric without increasing the domain count or port count respectively" on page 1339

Case study #1

Connectrix series switches with NPIV enabled linked through ISLs to Brocade Access Gateway

The objective of this case study is to connect a Brocade switch module in Access Gateway mode to a Connectrix B, Connectrix M and Connectrix MDS switch.

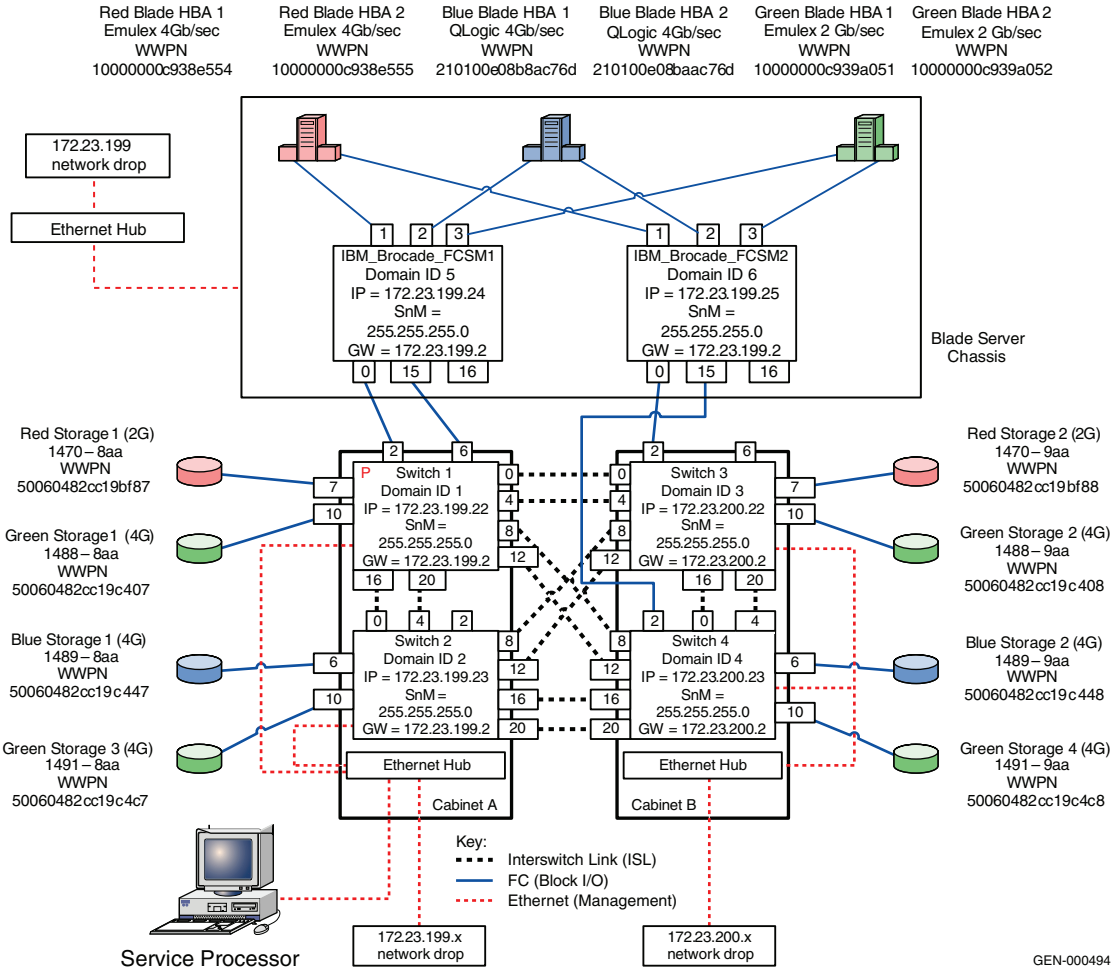


Figure 506 Topology

Figure 506 shows a four switch full mesh fabric topology as discussed in the "Four switch full mesh" on page 147. However, in this case, as can be seen above the red, green and blue hosts are replaced by the red, blue and green server blades that are attached to the fabric via the Brocade switch modules: IBM_Brcd_FCSM1 and IBM_Brcd_FCSM2 as shown above. The connections between the switch modules and the switches are ISLs.

Assumptions specific to this case study:
The following are assumptions specific to this case study.

- An IBM blade server with two 4 G IBM Brocade modules (32R1812).is added to a 4-switch full mesh fabric comprising of Switches A, B, C and D which can be replaced by either Connectrix B, Connectrix M or Connectrix MDS switches which have been discussed in the "Four switch full mesh" on page 147 section.
- The red, blue, and green hosts in the 4-switch full mesh fabrics referred to in Figure 506 on page 1322 now represent red, blue, and green server blades respectively in the blade server chassis.
- The server blades and switch modules in the blade server chassis are configured as explained in "Blade switch with direct attached storage" on page 129. The IPs and Domain IDs on the Brocade switch modules are as follows:
  - For Brocade module 1:

    IP: 172.23.199.24
    Subnet mask: 255.255.255.0
    IP Gateway: 172.23.199.2
    Domain ID: 5

  - For Brocade module 2:

    IP: 172.23.199.25
    Subnet mask: 255.255.255.0
    IP Gateway: 172.23.199.2
    Domain ID: 6

- The IBM Brocade modules and standalone switches in the 4-switch full mesh fabric are running supported firmware versions and operating with the supported interoperable modes.
  - (For the 4-switch full mesh brocade fabric, all the Brocade standalone switches can continue to operate in the **native mode**. The attached IBM Brocade switch modules must also be set to Brocade native mode (**interopmode 0**).
  - For the 4-switch full mesh Connectrix-M fabric, all the Connectrix-M standalone switches must now operate in **McData Open Fabric Mode** to support interoperability with the Brocade switch modules. In order to set the

    Open Fabric Mode, the Connectrix-M switches must be disabled and the Fabric parameters / Operating mode must be changed from McData Fabric Mode to Open Fabric Mode. The

attached IBM Brocade switch modules must be set to Brocade interopmode (**interopmode 1**). The detailed steps to change the interopmode in a Connectrix B/Connectrix M environment have been described in the "Four switch full mesh" on page 147 section.

- For the 4-switch full mesh Connectrix MDS fabric, the VSANs for the Connectrix MDS standalone switches must now be set to **Interop3** mode to support interoperability with the Brocade switch modules. In order to set the VSANs to Interop-3, the Connectrix MDS VSAN containing the switch E_Ports that are to be linked through ISLs to the switch modules and the ports on the path to the storage ports that the server blades are zoned, must be "suspended" and the interopmode must be changed from "default" to "Interop3". The attached IBM Brocade switch modules must be set to Brocade native mode (**interopmode 0**). Steps to change the interopmode in a Connectrix B/ Connectrix MDS environment have been described in the "Connectrix MDS example" on page 159 section.

◆ Two ISLs from one of the Brocade switch modules is going to the same standalone switch. This is referred to as a single-fabric attachment. Please refer to Figure 506 on page 1322 for more details on the physical connections.

◆ The other module has two ISLs going to two different Brocade switches. This is referred to as a multi-fabric attachment. Please refer to Figure 506 on page 1322for more details on the physical connections.

◆ The zoning information that was previously on the 4 switch full mesh Connectrix fabric, as referenced in the "Four switch full mesh" on page 147 case study examples, must now be on the IBM Brocade FCSM1 and FCSM2. This must be validated using the **cfgactvshow** command. The zoning information on the FCSM must appear as follows:

Effective configuration:

```
cfg: Oct_31_06_1140

zone: RedHBA1_1470_8aa
10:00:00:00:c9:38:e5:54
50:06:04:82:cc:19:bf:87

zone: RedHBA2_1470_9aa
10:00:00:00:c9:38:e5:55
```

```
50:06:04:82:cc:19:bf:88

zone: BlueHBA1_1489_8aa
21:01:00:e0:8b:8a:c7:6d
50:06:04:82:cc:19:c4:47

zone: BlueHBA2_1489_9aa
21:01:00:e0:8b:aa:c7:6d
50:06:04:82:cc:19:c4:48

zone: GreenHBA1_AllGreenStorage
10:00:00:00:c9:39:a0:51
50:06:04:82:cc:19:c4:07
50:06:04:82:cc:19:c4:08
50:06:04:82:cc:19:c4:c7
50:06:04:82:cc:19:c4:c8

zone: GreenHBA2_GreenStorage
10:00:00:00:c9:39:a0:52
50:06:04:82:cc:19:c4:07
50:06:04:82:cc:19:c4:08
50:06:04:82:cc:19:c4:c7
50:06:04:82:cc:19:c4:c8
```

- The final goal of this case study is to enable Access Gateway mode on these Brocade switch modules and to re-establish the existing connectivity without causing any disruption to the fabric.

To configure a switch:

In each case, i.e., for the Brocade module added to a Connectrix B, Connectrix M and Connectrix MDS fabric, it is essential to enable NPIV on the switch / switches that is / are linked through ISLs to the Brocade modules. Thus, in our example, Switch 1, Switch 3 ,and Switch 4 must have NPIV enabled. NPIV need not be enabled on Switch 2, although it is a part of the fabric. Refer to the sections below for the steps to enable NPIV on the different vendor switches.

- "Case study A" on page 1326 provides the steps for enabling NPIV on the Connectrix B switches, considering the Brocade modules are added to a Connectrix B fabric.
- "Case study B" on page 1327 provides the steps for enabling NPIV on the Connectrix M switches, considering the Brocade modules are added to a Connectrix M fabric.
- "Case study C" on page 1328 provides the steps for enabling NPIV on the Connectrix MDS switches, considering the Brocade modules are added to a Connectrix MDS fabric.

Case study A

Steps to enable NPIV on the Connectrix B switches

Most Brocade Fibre Channel switches running Fabric OS 5.1.0 or later, support NPIV. All Brocade 4 Gb capable Fibre Channel switches are, by default, configured with NPIV enabled.

Use the following procedure to enable NPIV on the Connectrix B switches when it is not enabled by default:

1. To enable or disable NPIV on a port-by-port basis from Brocade Fabric OS CLI, use the **portCfgNPIVPort** command.

   In our example, NPIV must be enabled on port 2 for switches 1, 3, and 4, and on port 6 for switch 1. The following commands need to be entered at the switch prompts:

   - For switch 1

     switch1:admin> **portCfgNPIVPort 2, 1**
     switch1:admin> **portCfgNPIVPort 6, 1**

   - For switch 3

     switch3:admin> **portCfgNPIVPort 2, 1**

   - For switch 4

     switch4:admin> **portCfgNPIVPort 2, 1**

2. To specify the number of virtual N_Port_IDs per port or per switch, use the **configure** command with either of the following parameters:

   - **`switch.login.perPortMax`**

     Use this parameter to set the number of virtual N_Port_IDs per port to a value between 0 and 255. The default setting is 126. A blade server chassis can house a maximum of 16 server blades, thus the maximum number of virtual N_port_IDs per port may be set to 16.

   - **`switch.login.perSwitchMax`**

     Use this parameter to set the number of virtual N_Port_IDs per switch to a value between 0 and (126 * number of ports). The default setting is (15 * number of ports).

   For our example, the IBM blade server can house a maximum of 14 blades so the Switch.login.perPortMax may be set to 15 = 1 FLOGI from the AG N_Port + 14 FDISCs from the 14 blade server N_Ports. The switch.login.perSwitchMax can be left at the default value.

The following commands need to be run at the switch prompt for switches 1, 3, and 4:

```
switch:admin> switchdisable
switch:admin> configure
Configure ...
Fabric parameters (yes, y, no, n): [no]
Virtual Channel parameters (yes, y, no, n): [no]
F_Port login parameters (yes, y, no, n): [no] y
Maximum logins per switch: (1..4032) [4032] 2048
Maximum logins per port: (1..255) [255] 15
switch:admin> switchenable
```

Case study B

Steps to enable NPIV on the Connectrix M switches

Connectrix M FC switches with E/OS 8.0 or later support NPIV. Connectrix M switches require an optional license to enable this function.

Use the following procedure to apply this license and enable NPIV on switches 1, 3, and 4:

1. From a browser, open the web user interface for all the Connectrix M switches that are to be connected to the IBM Brocade Access Gateway module, i.e., Switch 1, 3, and 4. On the switch EFCM GUI, the **Node List** view details the devices attached to the Connectrix M switch.
2. To install the NPIV license, click **Maintenance** and then select **Options** features.
3. Enter the license key for NPIV in the **Feature Key** field. Select the key as the "**N_Port ID Virtualization (NPIV)**." Link from the window and apply the key by clicking **OK**.

   A checkmark in the left window indicates that the N_Port ID Virtualization key is installed.
4. Click **Configure** and then select **Ports>NPIV**.
5. Click **Enable**.
6. At the prompt, click **OK** if you are sure you want to enable NPIV.
7. In the **Login** column, set the value to 17 or higher for each port connected to the IBM Brocade Access Gateway to ensure proper operation.
8. Click **OK** to save changes.

Case study C

Steps to enable NPIV on the Connectrix MDS switches

Connectrix MDS Fibre Channel switches running SAN-OS 3.0 or later will support NPIV.

To enable NPIV on Connectrix MDS Fibre Channel Switches 1, 3, and 4 running the Cisco Device Manager, use the following procedure:

1. From the **Cisco Device Manager,** click **Admin** and then select **FeatureControl**. The **Feature Control** window appears.
2. Click the row titled **NPIV**.
3. In the **Action** column select **Enable** and then click **Apply**.
4. Click **Close** to return to the **Name Server** screen.
5. Click **Refresh** to display the host ports.

The Cisco CLI may also be used to enable NPIV by running the following commands at the switch prompt for switches 1, 3, and 4:

```
switch# config t
switch(config)# npiv enable
```

To configure the Brocade switch module:

Once NPIV is enabled on the standalone switches, the AG mode can be enabled on the IBM Brocade modules, FCSM1 and FCSM2, by running the following commands in the switch module prompt:

1. Run the **msplmgmtdeactivate** command to deactivate the MS Platform Service before enabling AG mode.
2. Disable the Brocade module by running the **switchdisable** command at the switch prompt.
3. Backup the existing configuration (effective and defined zonesets) using the **configupload** command. Provide details of an ftp server as prompted where the configuration can be saved as a text file.
4. Enable the AG mode by running **ag --modeenable** on the switch prompt. This command also reboots the switch. When the switch comes up, the zoning configuration is erased. However, the zoning configuration still exists on switches 1, 2, 3, and 4.
5. The **ag --mapshow** command displays the factory default N_Port configuration and F_Port to N_Port mappings. The **Configured_F_Ports** display the initial configuration while the

Current_F_Ports display the current mapping which, because of an N_Port failover, is same as the **Configured_F_Ports** until the configuration changes.

In the CLI example below, the external N_Ports on the module specified in Column 1 are mapped to the internal F_Ports specified in Column 3 by default. Column 3, the **Current_F_Ports** column specifies the F_Ports that are currently online which is ports 1, 2 and 3 (the respective red, blue, and green HBA ports connected internally to the Access Gateway module. Their mapping is the same as the default mapping until an N_Port goes down and all the F_Ports mapped to it by default now failover to another N_Port.

FCSM1: **admin> ag --mapshow**

N_Port	Configured F_Ports	Current F_Ports	Failover	Failback
0	1; 2;	1; 2;	1	1
15	3; 4;	3;	1	1
16	5; 6; 7; 8;	None	1	1
17	9; 10; 11;	None	1	1
18	12;	None	1	1
19	13. 14;	None	1	1

If desired, the F_Port to N_Port mapping can be edited by using the following commands:

- **ag --mapadd**: This command adds new F_Ports to the list of currently mapped ports for a N_Port.
- **ag --mapdel**: This command removes F_Ports from the list of currently mapped ports for a N_Port
- **ag --mapset:** This command sets the F_Port mapping for a N_Port to the specified list of F_Ports.

For example, if we want to map F_Port 2 to N_Port 15 instead of N_Port 0, the following set of commands need to be run:

```
FCSM1:admin> ag --mapdel 0 “2”  //F_port2 is not mapped to N_port 0 now

FCSM1:admin> ag --mapset 15 “2” //F_port 2 is now mapped to N_port 15
```

6. Verify that all the F_Ports and N_Ports are now online by running the **switchshow** command.

Thus, the two primary steps that need to be executed to configure the module as an Access Gateway are:

- Enabling the AG mode
- Setting the desired F_Port to N_Port mapping

Case study #2

Migration from Brocade module (native) ISLed to Brocade switches (native) to Brocade module (in AG mode) connected to Cisco switches (interopmode irrelevant)

This example shows how any complexity created by using the interoperability modes for the IBM Brocade module / Connectrix MDS interop is eliminated by using the Access Gateway mode on the Brocade modules.

Figure 507 on page 1331 shows a Brocade-based blade server chassis attached to a 4-switch full mesh Connectrix B fabric. The 4-switch Connectrix B fabric needs to be replaced by a Connectrix MDS fabric as shown in Figure 508 on page 1332.

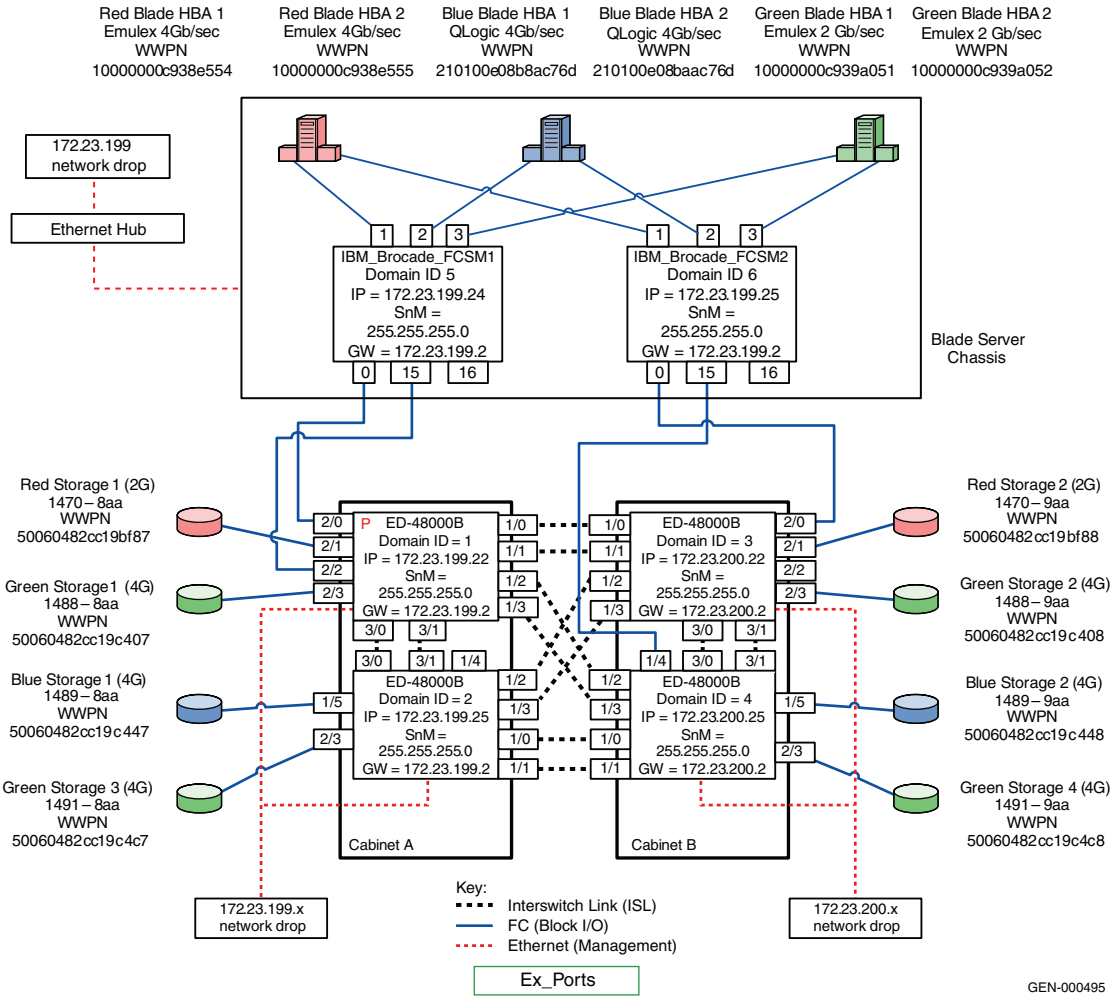


Figure 507 **Brocade-based blade server chassis attached to a 4-switch full mesh Connectrix B fabric**

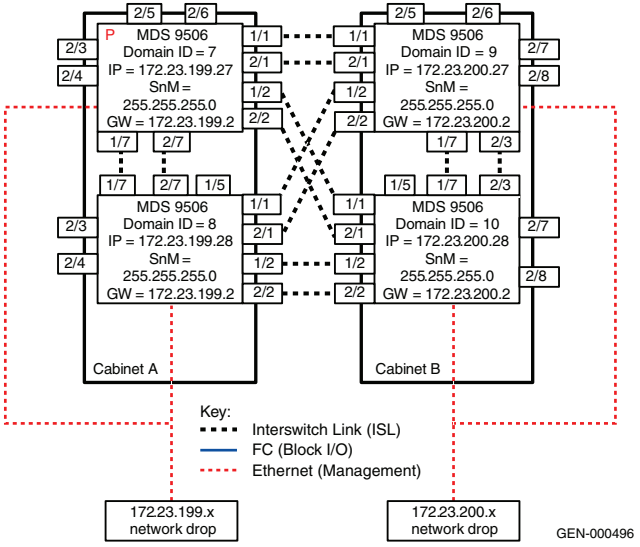


Figure 508 **Connectrix MDS fabric**

The completed migration is shown in Figure 509 on page 1333 with the Brocade based blade server modules now in "Access Gateway" mode and connected to the Connectrix MDS 4-switch full mesh fabric.

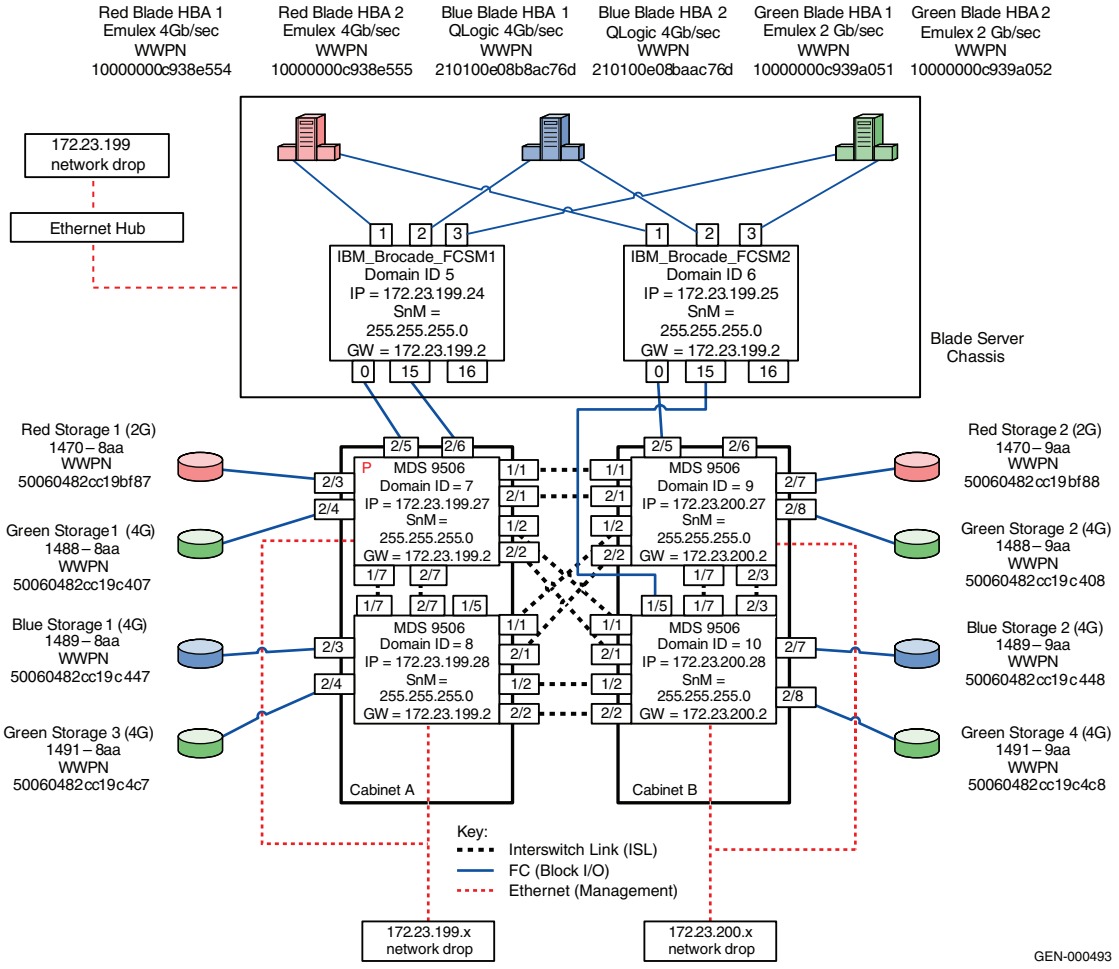


Figure 509 **Completed migration**

Assumptions specific to this case study:

The following are assumptions specific to this case study.

- An IBM blade server with two 4 G IBM brocade modules (32R1812) is added to the 4-switch full mesh fabric comprising of Connectrix B switches that have been discussed in the four switch full mesh "Connectrix B example" on page 148.

- The red, blue and green hosts in the 4-switch full mesh fabrics referred to in the four switch full mesh "Connectrix B example" on page 148 , now represent red, blue, and green server blades respectively in the blade server chassis.
- The server blades and switch modules in the blade server chassis are configured as explained in chapter xx section yy. The IPs and domain IDs on the brocade switch modules are as follows:
  - For Brocade module 1:

    IP: 172.23.199.24
    Subnet mask: 255.255.255.0
    IP Gateway: 172.23.199.2
    Domain ID: 5
  - For Brocade module 2:

    IP: 172.23.199.25
    Subnet mask: 255.255.255.0
    IP Gateway: 172.23.199.2
    Domain ID: 6
- The IBM Brocade modules and standalone switches in the 4-switch full mesh fabric are running supported firmware versions and operating with the supported interoperable modes.For the 4-switch full mesh brocade fabric, all the brocade standalone switches can continue to operate in the **native mode**. The attached IBM brocade switch modules must be set to brocade native mode (**interopmode 0**) also.
- 2 ISLs from one of the brocade switch modules is going to the same standalone switch. This is referred to as a single-fabric attachment. Please refer to Fig xx for more details on the physical connections.
- The other module has 2 ISLs going to two different brocade switches. This is referred to as a multi-fabric attachment. Please refer to Figure 507 on page 1331 for more details on the physical connections.
- The zoning information that was previously on the 4 switch full mesh Connectrix B fabric, per the "Connectrix B example" on page 148, must now be on the IBM Brocade FCSM1and FCSM2. This must be validated using the **cfgactvshow** command. The zoning information on the FCSM must appear as follows:

  Effective configuration:

  ```
  cfg: Oct_31_06_1140
  ```

```
zone: RedHBA1_1470_8aa
10:00:00:00:c9:38:e5:54
50:06:04:82:cc:19:bf:87

zone: RedHBA2_1470_9aa
10:00:00:00:c9:38:e5:55
50:06:04:82:cc:19:bf:88

zone: BlueHBA1_1489_8aa
21:01:00:e0:8b:8a:c7:6d
50:06:04:82:cc:19:c4:47

zone: BlueHBA2_1489_9aa
21:01:00:e0:8b:aa:c7:6d
50:06:04:82:cc:19:c4:48

zone: GreenHBA1_AllGreenStorage
10:00:00:00:c9:39:a0:51
50:06:04:82:cc:19:c4:07
50:06:04:82:cc:19:c4:08
50:06:04:82:cc:19:c4:c7
50:06:04:82:cc:19:c4:c8

zone: GreenHBA2_AllGreenStorage
10:00:00:00:c9:39:a0:52
50:06:04:82:cc:19:c4:07
50:06:04:82:cc:19:c4:08
50:06:04:82:cc:19:c4:c7
50:06:04:82:cc:19:c4:c8
```

- The final goal of this case study is to replace all the standalone Connectrix B switches by Connectrix MDS switches. The blade server switch modules in Brocade native mode are then switched to Brocade Access Gateway mode with all external N_Ports connected to the NPIV enabled Connectrix MDS switches. The operating mode on the Conectrix MDS switch is irrelevant.

To migrate from a Connectrix B 4-switch fabric to a Connectrix MDS 4-switch fabric with brocade based blade servers, without changing the interoperability modes

1. While the Brocade modules are linked through ISLs to the Connectrix B standalone switches, enable the AG mode by executing the following steps:

   a. Run the **msplmgmtdeactivate** command to deactivate the MS Platform Service before enabling AG mode.

b. Disable the Brocade module by running the **switchdisable** command at the switch prompt.

c. Backup the existing configuration (effective and defined zonesets) using the **configupload** command. Provide details of an ftp server as prompted where the configuration can be saved as a text file.

d. Enable the AG mode by running **ag --modeenable** on the switch prompt. This command also reboots the switch. When the switch comes up, the zoning configuration is erased.

e. The **ag --mapshow** command displays the factory default N_Port configuration and F_Port to N_Port mappings. The **Configured_F_Ports** display the initial configuration while the **Current_F_Ports** display the current mapping which, because of an N_Port failover, is same as the **Configured_F_Ports** until the configuration changes.

f. The F_Port to N_Port mapping can be edited as is desired by using any of the following commands:

– **ag --mapadd**: This command adds new F_Ports to the list of currently mapped ports for a N_Port.

– **ag --mapdel**: This command removes F_Ports from the list of currently mapped ports for a N_Port

– **ag --mapset**: This command sets the F_Port mapping for a N_Port to the specified list of F_Ports.

Note: For more detailed information on **ag –mapshow** and changing the N_Port to F_Port mapping please refer to Step 5 on page 1328.

---

g. Verify that all the F_Ports and N_Ports are now online by running the **switchshow** command.

2. Configure the IP for the individual Connectrix MDS standalone switches. Install the Fabric Manager and Device Manager as explained in the "Connectrix MDS example" on page 159, and link through ISLs the target Connectrix-MDS fabric with only E_Ports as shown in Figure 508 on page 1332.

3. Create an interop-3 (Brocade native mode) VSAN on the Connectrix MDS switches 7, 8, 9, and 10.

a. Click the **Create VSAN** icon

b. The **Create VSAN** dialog box displays.

c. Check the switches that you want in this VSAN: Switch 7, 8, 9, and 10.

d. Fill in the **VSAN Name** as “Brcd_module_interop_VSAN” and **VSAN ID** fields with an unused VSAN ID, for example “4”.

e. Set the **LoadBalancing** value to default and the **InterOperValue** to **Interop-3**.

f. Set the **Admin State** to **active**.

g. Check the **Static Domain Ids** checkbox to assign an unused static Domain IDs to the switches in the VSAN. For example, Switch 7, 8, 9, and 10 have to be assigned Domain IDs 7, 8, 9, and 10.

h. Leave the other fields in this dialog box at default and click **Create** to add the VSAN.

4. Assign and enable all the E_Ports in Figure 508 on page 1332 to VSAN ID 4 using the Device Manager for the respective switches. From the **Device Manager Menu** bar, go to **FC>VSANs** and click the **Membership** tab to add the ports to the desired VSAN.

5. Validate E_Port connectivity by comparing the topology in the Figure 507 on page 1331 with the topology view of the Connectrix MDS fabric comprising of Switches 7, 8, 9, and 10 obtained using Fabric Manager.

6. Physically connect an ISL between the Connectrix MDS Switch 8 port FC 2/8 (or any other unused port) and Connectrix B Switch 2 port 1/5 (or any other unused port) so the existing active zoneset from the Connectrix B switches gets pulled onto the Connectrix MDS fabric.

   This can be validated by opening the **Zoning** menu in the Fabric Manager for VSAN ID 4.

   The **Active Zoneset** on the Connectrix MDS switch 8 will appear as follows:

```
Zone set name = Oct_31_06_1140
Zone name = "RedHBA1_1470_8aa"
Zone Member = "10000000c938e554"
Zone Member = "50060482cc19bf87"

Zone name = "RedHBA2_1470_9aa"
Zone Member = "10000000c938e555"
Zone Member = "50060482cc19bf88"
```

```
Zone name = "BlueHBA1_1489_8aa"
Zone Member = "210100e08b8ac76d"
Zone Member = "50060482cc19c447"

Zone name = "BlueHBA2_1489_9aa"
Zone Member = "210100e08baac76d"
Zone Member = "50060482cc19c448"

Zone name = "GreenHBA1_AllGreenStorage"
Zone Member = "10000000c939a051"
Zone Member = "50060482cc19c407"
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"

Zone name = "GreenHBA2_AllGreenStorage"
Zone Member = "10000000c939a052"
Zone Member = "50060482cc19c407"
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"
```

7. All switches in the Connectrix MDS fabric distribute active zone sets when new E_Port links come up or when a new zone set is activated in a VSAN if the full zone set propagation is set.

   To propagate the full zoneset across all switches in a VSAN, execute the following steps:

   a. Select **VSAN ID 4 > Default Zone** from the **Logical Domains** pane.The zone set configuration displays in the **Information** pane.

   b. Select the **Policies** tab.

   c. Set the **Propagation** column to **fullZoneset** from the drop-down menu.

   d. Click **Apply Changes** to propagate the full zone set, or click **Undo Changes** to discard any changes you made.

8. Enable NPIV on the Connectrix MDS Switches 7, 9, and 10 which are to be connected to the AG module by executing the following steps:

   a. From the **Cisco Device Manager**, click **Admin**, and then select **FeatureControl**. The **Feature Control** window appears.

   b. Click the row titled **NPIV**.

   c. In the **Action** column select **Enable**, and then click **Apply**.

d. Click **Close** to return to the **Name Server** screen.

e. Click **Refresh** to display the host ports.

9. Move half of the access gateway N_Ports and the storage ports from the Connectrix B fabric to the Connectrix MDS fabric.

   This requires connections to be physically removed from the Connectrix B director and connected to a Connectrix MDS 9506 director. The active zonesets propagates across the Connectrix MDS switches ensuring that connectivity is maintained even if the server blade connection is in one fabric and the storage connection is in the other fabric.

   a. Move one device at a time. Start by moving the cable for the "red host HBA 1".

   b. After it has been connected to the appropriate port in the Connectrix MDS fabric (Domain ID 7, 2/5), use the copa inquiry (**inq** command) on the red server blade to verify the connectivity.

   c. Review the "Checkpoints" on page 305 (in "Case study #3" on page 290) and verify that this action did not affect the connectivity and functioning of the fabric.

   d. Repeat this procedure as you move each connection for the red host and red storage, and the blue host and storage, from the Connectrix B to the Connectrix MDS fabric.

10. This step is an extension of Step 9. After validating that the stable fabric exists after the completion of Step 9, execute this step by pulling the remaining host and storage connections in the Connectrix B fabric. Review the "Checkpoints" on page 305 (in "Case study #3" on page 290) to validate the connectivity.

11. Disable the ISLs between the Connectrix B and Connectrix MDS switches and validate stability of the Connectrix MDS fabric attached to the blade servers via the Access Gateway modules.

    After the migration is completed the fabric must appear as shown in Figure 509 on page 1333.

Case Study #3 Addition of switches or server blades in an existing Connectrix-M fabric without increasing the domain count or port count respectively

This case study discusses the advantage of using the brocade Access Gateway modules for scalability purposes.

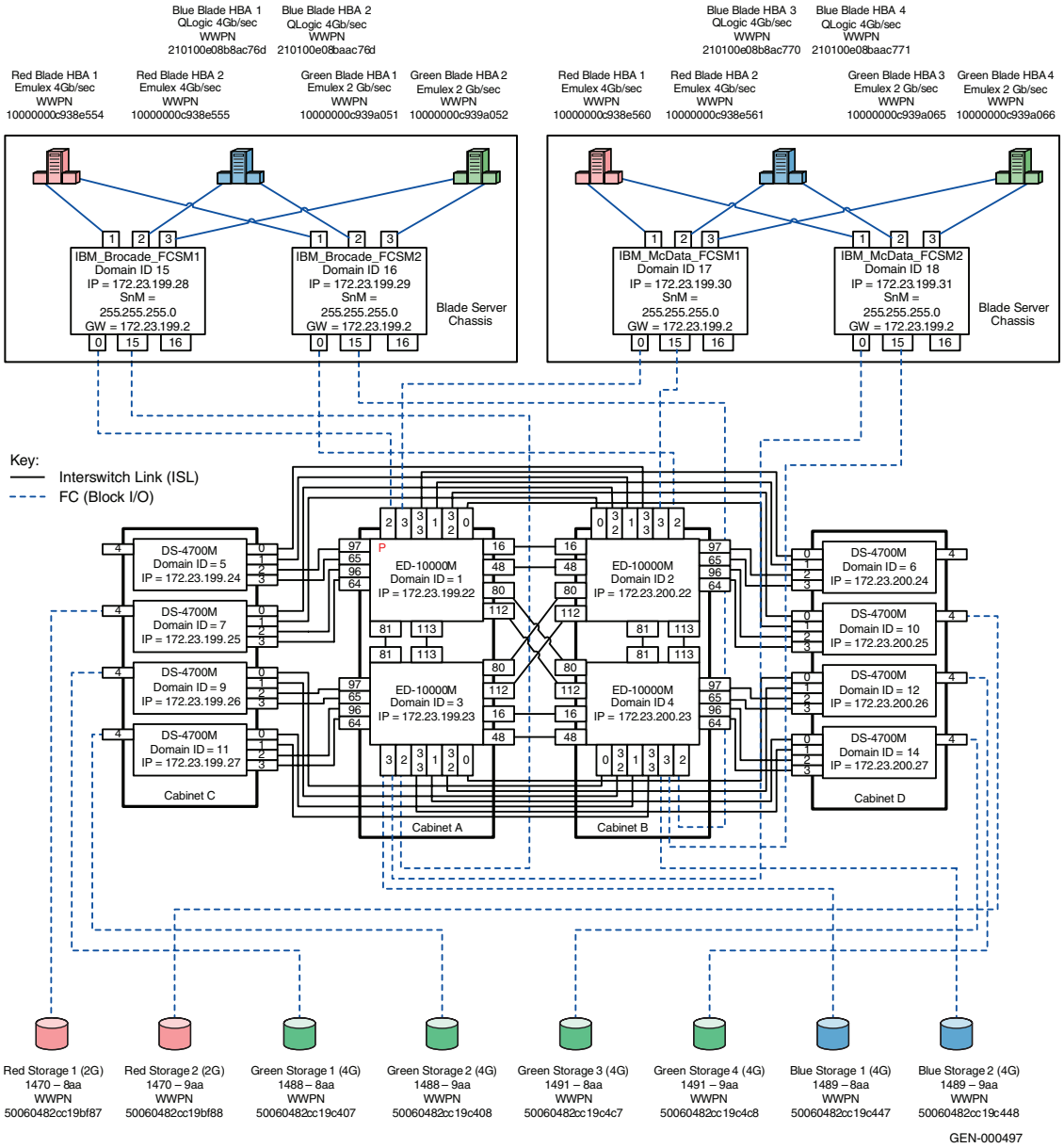


Figure 510 Topology

Figure 510 on page 1340 shows the compound-core edge Connectrix M fabric discussed in the "Connectrix M example" on page 234 with the red, green, and blue hosts being replaced by red, blue, and green blade server chasses connecting via the Brocade and McDATA modules with ISLs, as described in the assumptions and steps to follow in this section.

Assumptions specific to this case study:

The following are assumptions specific to this case study.

- Consider the compound core-edge Connectrix-M topology discussed in the "Connectrix M example" on page 234 with a mix of Brocade and McDATA-based blade servers attached to the fabric as shown in the Figure 510 on page 1340. The Connectrix M switches are in Open Fabric 1.0 mode while the Brocade modules are operating in interopmode, and the McDATA modules are in standard mode. Setting up this kind of a configuration has not been explicitly explained in any previous section, and thus is explained in this case study.
- This example assumes that the customer already has a Connectrix M fabric with blade server modules and needs to increase the fabric size by adding more switch modules and server blades. If the customer is to start designing a scalable fabric with switch modules, then it is recommended to add switch modules with Access Gateway enabled.
- The final goal of this case study is to increase the fabric size by adding more switches to the existing setup, yet maintain the switch domain count, keeping it to a minimum. This is done by enabling Access Gateway on the existing Brocade modules and swapping the McDATA modules if required by Brocades in Access Gateway mode.

  If more blade servers are to be added into the fabric without significantly increasing the N_Port count, the existing Brocade modules can be set as Access Gateway and the newly added blade servers must have AG modules as the I/O modules. In that way, although we are expanding the number of device ports, we can connect them through a single AG fabric port, thus enabling support for a much larger fabric in terms of number of connected host and storage devices.

To set up a Connectrix-M core edge fabric with blade servers:

1. Set up a Connectrix-M core edge fabric with E_Ports or ISLs only (i.e., without the server and storage ports) as explained in section Figure 510 on page 1340.

2. Configure the Connectrix M switches as explained in the "Connectrix M example" on page 234 with Interop Mode set to Open Fabric 1.0.

3. Configure the IBM Brocade modules as explained in "IBM Brocade example" on page 136**.** Telnet into the FCSM1 and FCSM2 and run the following commands at the switch prompt to set the interopmode on.

   Fcsm1> **switchdisable**
   Fcsm1> **interopmode 1**

   Press **Yes** when prompted to continue.

   Fcsm1>**reboot**

   Note: A switch reboot is recommended after switching the interop mode.

4. ISL the blade server McDATA modules in standard mode to the Connectrix-M fabric as shown in the Figure 510 on page 1340.

   a. Using EFS 2007, (the McDATA switch module Fabric Manager application), add the IPs of the McDATA modules. These switches display in a column on the left pane of the GUI.

   b. Click on one of the switch modules from the left pane..

   c. When the selected switch shows up as a faceplate display:

      – Open the **Switch** menu on the top of the screen.
      – Select **Advanced Switch Properties**.
      – When the **Advanced Switch Properties** dialog displays, set the **Interop Mode** to **Standard**.

      The switch will automatically be taken offline and restored once the changes (if any) are completed.

5. Attach the storage ports as shown in Figure 510 on page 1340**.**

6. Create and activate the zoneset on the Connectrix M switches, discussed in the "Connectrix M example" on page 234. However, please note the addition of the new zones to this zoneset, which are created as a result of the addition of three more server blades

(6 more server ports in the fabric): Red HBA3,4, Blue HBA3,4, and Greeen HBA 3,4. The zoneset on the Connectrix M switches must appear as follows:

```
Zone set name = "Oct_31_06_1140"

Zone name = "RedHBA1_1470_8aa"
Zone Member = "10000000c938e554"
Zone Member = "50060482cc19bf87"

Zone name = "RedHBA2_1470_9aa"
Zone Member = "10000000c938e555"
Zone Member = "50060482cc19bf88"

Zone name = "RedHBA3_1470_8aa"
Zone Member = "10000000c938e560"
Zone Member = "50060482cc19bf87"

Zone name = "RedHBA4_1470_9aa"
Zone Member = "10000000c938e561"
Zone Member = "50060482cc19bf88"

Zone name = "BlueHBA1_1489_8aa"
Zone Member = "210100e08b8ac76d"
Zone Member = "50060482cc19c447"

Zone name = "BlueHBA2_1489_9aa"
Zone Member = "210100e08baac76d"
Zone Member = "50060482cc19c448"

Zone name = "BlueHBA3_1489_8aa"
Zone Member = "210100e08b8ac770"
Zone Member = "50060482cc19c447"

Zone name = "BlueHBA4_1489_9aa"
Zone Member = "210100e08baac771"
Zone Member = "50060482cc19c448"

Zone name = "GreenHBA1_AllGreenStorage"
Zone Member = "10000000c939a051"
Zone Member = "50060482cc19c407"
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"

Zone name = "GreenHBA2_AllGreenStorage"
Zone Member = "10000000c939a052"
Zone Member = "50060482cc19c407"
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"
```

```
Zone name = "GreenHBA3_AllGreenStorage"
Zone Member = "10000000c939a065"
Zone Member = "50060482cc19c407"
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"

Zone name = "GreenHBA4_AllGreenStorage"
Zone Member = "10000000c939a066"
Zone Member = "50060482cc19c407"
Zone Member = "50060482cc19c408"
Zone Member = "50060482cc19c4c7"
Zone Member = "50060482cc19c4c8"
```

7. At this step, the connectivity and fabric stability can be verified by checking that all the ports are up, the switches are in the appropriate modes of operation , and that the zones have been propagated across all the Connectrix M switches and Brocade, McDATA switch modules.

To increase the fabric size by adding more switches or blade servers:

To keep the switch domain count to the minimum possible, we will use the scalability aspect of Access Gateway and enable Access Gateway on all the Brocade modules in the fabric. At the same time, if more blade servers are to be added, the total server N_Port count can be kept low by using the NPIV functionality of Access Gateway modules. Multiple server blade N_Ports can be mapped to a single N_Port on the Access Gateway module reducing the total number of N_Ports seen by the Connectrix-M fabric.

The mode of operation (i.e., Open Fabric 1.0 in this case) on the Connectrix M switches is irrelevant for Access Gateway connectivity. Therefore, we can leave the interop mode as it is on the Connectrix M switches. The first step is to enable NPIV on the Connectrix M switches that are currentlylinked through ISLs to the blade server Brocade modules.

Note: Connectrix M switches require an optional license to enable this function.

Use the following procedure to apply this license and enable NPIV:

1. From a browser, open the web user interface for the Connectrix M switch that is to be connected to the Brocade AG module. The **Node List** view details the devices attached to the Connectrix M switch.

2. To install the NPIV license, click **Maintenance**, and then select the **Options** features.
3. Enter the license key for NPIV in the **Feature Key** field. Select the key as the "**N_Port ID Virtualization (NPIV)**." Link from the window and apply the key by clicking **OK**.

   A checkmark in the left window indicates that the N_Port ID Virtualization key is installed.
4. Click **Configure** and then select **Ports>NPIV**.
5. Click **Enable**.
6. At the prompt, click **OK** if you are sure you want to enable NPIV.
7. In the **Login** column, set the value to 17 or higher for each port connected to the IBM Brocade Access Gateway to ensure proper operation.
8. Click **OK** to save changes.

The AG mode must be enabled on the Brocade modules by executing the following steps:

1. Run the **msplmgmtdeactivate** command to deactivate the MS Platform Service before enabling AG mode.
2. Disable the Brocade module by running the **switchdisable** command at the switch prompt.
3. Backup the existing configuration (effective and defined zonesets) using the **configupload** command. Provide details of an ftp server as prompted where the configuration can be saved as a text file.
4. Enable the AG mode by running **ag --modeenable** on the switch prompt. This command also reboots the switch. When the switch comes up, the zoning configuration is erased. However, the zoning configuration still exists on switches 1, 2, 3, and 4.
5. The **ag --mapshow** command displays the factory default N_Port configuration and F_Port to N_Port mappings. The **Configured_F_Ports** display the initial configuration while the **Current_F_Ports** display the current mapping which, because of an N_Port failover, is same as the **Configured_F_Ports** until the configuration changes.
6. The F_Port to N_Port mapping can be edited as is desired by using any of the following commands:

- **ag --mapadd**: This command adds new F_Ports to the list of currently mapped ports for a N_Port.
- **ag --mapdel**: This command removes F_Ports from the list of currently mapped ports for a N_Port
- **ag --mapset:** This command sets the F_Port mapping for a N_Port to the specified list of F_Ports.

Note: For more detailed information on ag –mapshow and changing the N_Port to F_Port mapping please refer to Step 5 on page 1328.

7. Verify that all the F_Ports and N_Ports are now online by running the **switchshow** command.

To replace McDATA modules with Brocade modules in Access Gateway mode:

If required, the McDATA modules can also be replaced with Brocade modules in Access Gateway mode. This will further reduce the switch domain count. The following steps must be executed for the same:

1. Plan a downtime for the McDATA-based blade server.
2. Disable the McDATA switch module.
3. Physically disconnect the ISLs running between the McDATA modules and the Connectrix-M series switches.
4. Replace the McDATA modules with Brocade modules.
5. Configure the Brocade modules with an IP address.
6. Enable NPIV on the Connectrix switches that were originally connected through ISLs to the McDATA modules and which will now beconnected through ISLs to the AG modules.
7. Telnet into the Brocade module, and enable Access Gateway mode using Step 1 on page 1345 through Step 7 on page 1346 under "The AG mode must be enabled on the Brocade modules by executing the following steps:".
8. Reconnect the Connectrix-M switches to the Brocade Access Gateway module and validate the end-to-end connectivity and device discovery.

Access Gateway qualification plan

This section lists the “areas of concern” that are qualified as a part of the E-Lab Switch Interoperability qualification plan for Access Gateway features and functionality. The test plan has been designed to qualify both the F_Port to N_Port connectivity within the AG module, and the N_Port to F_Port connectivity between the AG modules and external NPIV switches.

Each of the Access Gateway features / functions have been tested for a specific 'purpose' as listed in this sub section. The 'background' provides a reasoning behind designing a test case around that particular area of concern.

AG manageability

Purpose: To verify the functionality of essential Access Gateway configuration commands.

Background: On enabling Access Gateway on the Brocade switch module, it loses its FC fabric switch functionality and introduces a new command set for management. As a part of this testing, E-Lab validates the custom configuration and management of Access Gateway using the supported interfaces, such as Web Tools and CLI commands used to enable AG mode, revert back to standard switch mode, configure / display F_Port to N_Port mapping, configure N_Port failover / failback policies and disable / enable N_Port capability must be validated. Verifying that NPIV logins are properly displayed in the Connectrix Manager fabric tree is also essential.

AG login

Purpose: To characterize the login behavior of an AG switch module which has been enabled for NPIV and the behavior after an N_Port failover.

Background: A clean and efficient login is essential for proper N_Port operation. By ensuring a similar login process can be accomplished across all platforms, we ensure that the users’ experience will be similar across all the NPIV edge switch vendors tested by EMC. This will reduce support costs.

AG fabric connectivity

Purpose: To ensure that end-to-end (physical and logical) connectivity is possible across multiple platform types in a multivendor switch fabric.

Background: Seamless end-to-end connectivity is the whole purpose of an FC SAN. Ensuring that devices can communicate with each

other across all different platform types in an interop fabric with AG is critical to achieving seamless connectivity.

AG stress test

Purpose: To validate that adding and removing NPIV logins does not impact other NPIV login sessions.

Background: Since NPIV virtual logins will be used to give each virtual host access to storage, it is important that each NPIV implementation insulates the actions of one virtual host from another.

AG firmware download

Purpose: To verify the hot code load feature for AG.

Background: The hot code load feature is supported on the Brocade switch with AG mode enabled. It would be interesting to verify whether the configuration details, AG port mapping, nameserver, and zoning details on the edge switches, etc., are maintained after a firmware upgrade / downgrade without a reboot.

AG N_Port failover/failback

Purpose: To validate the N_Port failover feature on the AG module.

Background: With the failover policy enabled on the AG modules, all F_Ports will be mapped to other N_Ports connected to the same fabric and re-enabled. It will first try the N_Ports connected to the same switch and then the same fabric. The delay between the disable and re-enable of F_Ports is of the order of 5-10 ms. The PIDs assigned to the F_Ports get modified in this case. When the original N_Port becomes active, the F_Ports configured to it will failback if failback policy is enabled for the N_Port. This needs to be validated.

Qualification results

The following are qualification results.

Supported configurations

The following are supported configurations:

- All the 4 G IBM, Dell ,and HP Blade Server Brocade modules currently listed in the *EMC Support Matrix* are supported for Access Gateway mode. No standalone switches running Brocade v5.2.1 and up are supported for Access Gateway at this time.
- For a list of the Connectrix-B, Connectrix-M, and Connectrix-MDS edge switches that can be connected to the Access gateway with NPIV enabled, please refer to the most current *EMC Support Matrix*.
- The Connectrix-M and Connectrix-MDS connectivity will be supported starting from the following firmware versions only.

- Connectrix M EOS v9.x
- Connectrix MDS firmware v3.x

◆ Only FCP initiator ports can be connected to Access Gateway as F_Ports (ports from other Access gateway, FCP Target ports, loop device, and FICON channels / control unit connectivity are not supported).

◆ Dynamic POD will continue to be supported when agmode is enabled.

Limitations

The following are limitations:

◆ Access gateway does not support loop devices.

◆ Cascading an Access Gateway to another Access Gateway will not be supported.

◆ N_Port failover is supported only within N_Ports connected to the same switch for Brocade FOS 5.2.x (i.e. , v5.2.1b, v5.2.2) and will be supported for N_Ports connected to the same fabric from Brocade FOS 5.3.

◆ The Access Gateway feature is supported only through CLI and *not* Web Tools.

Some current FOS testing-related limitations that affectAccess Gateway are:

◆ Maximum number of Access Gateways that can be connected to an edge switch is 30 (due to testing limitation).

◆ Maximum number of devices that can be connected to a FOS switch through Access Gateway depends on the maximum number of local devices that are supported in a fabric.

◆ Table 100 shows the scalability bounds for an Access Gateway based fabric due to the testing limitations in EMC's and Brocade's scalability testing lab.

Table 100 Fibre Channel Access Gateway support limitations

Max# of Access Gateways in a fabric	39
Max# of Access Gateways that can be connected to a switch	30
Max# of F-ports that can be mapped to the N-ports on an Access Gateway module	16
Max# of N-ports supported per Access Gateway module	8

D

Legacy Information

This appendix contains information on legacy technologies.

- Fibre Channel arbitrated loop (FC-AL) 1352
- QuickLoop 1391
- Bridges 1396
- Interfacing arbitrated loop to switched fabric 1408
- Storage area network management 1417
- CNT (Inrange) 1421

Fibre Channel arbitrated loop (FC-AL)

The following subjects are discussed in this section:

- "Overview"
- "Symmetrix and Fibre Channel connectivity" on page 1359
- "Arbitrated loop addressing" on page 1361
- "Primitive signals and sequences" on page 1365
- "Loop initialization" on page 1369
- "Arbitration process" on page 1379
- "Alternate Buffer-to-Buffer Credit (BB_Credit)" on page 1388

Overview

When Fibre Channel was first introduced, it was a new technology and everything was expensive. Switches, hubs, and node transceivers proved to be costly. Arbitrated Loop topology lies between point-to-point and switched fabric in that it provides more connectivity than point-to-point with up to 126 NL_Ports in a loop, but less than switched fabric which has the ability in theory to support up to 16 million ports. It was a cost-effective way of connecting a limited number of ports in a loop single network.

Fibre Channel arbitrated loop (FC-AL) is a *daisy-chain* connecting up to 126 devices in a loop configuration over attachment points called *L_Ports* (loop ports). FC-AL is a low-cost connectivity solution because it does not require switches. FC-AL is a good choice for small to medium-sized configurations, and provides a growth path by allowing connection of a loop to a switched fabric.

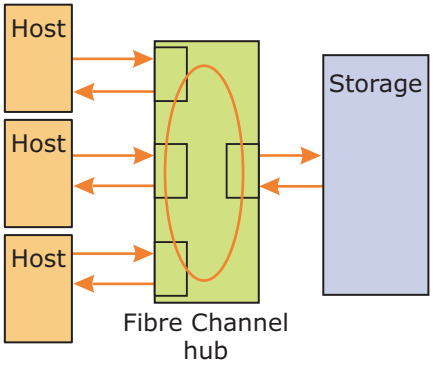


Figure 511 Arbitrated loop example

Efficiency and connectivity is enhanced by incorporating one or more hubs into the loop. Routing traffic through a hub on each leg of a loop eliminates the loss of the entire loop, as happens in a hubless loop. ("Loops with hubs" on page 1356 provides more information.)

The arbitrated loop topology promoted the introduction of Fibre Channel by removing the cost of a fabric switch and, depending on the number of nodes in the loop, the amount of transceivers could also be reduced.

Arbitrated loop provides more connectivity than point-to-point in that it can support 126 NL_Ports and 1 FL_Port on a single loop and is a middle ground between point-to-point and switched fabric. In arbitrated loop, the transmit output of one port is connected to the receive input of the next and these connections are made between all the nodes until a closed loop is formed (refer to Figure 512 on page 1354). This type of configuration is usually made using a Fibre Channel hub which eliminates the need to form the logical loop using cabling.

In arbitrated loop each port sees all messages on the loop and ignores or passes those messages which are not addressed to that particular port.

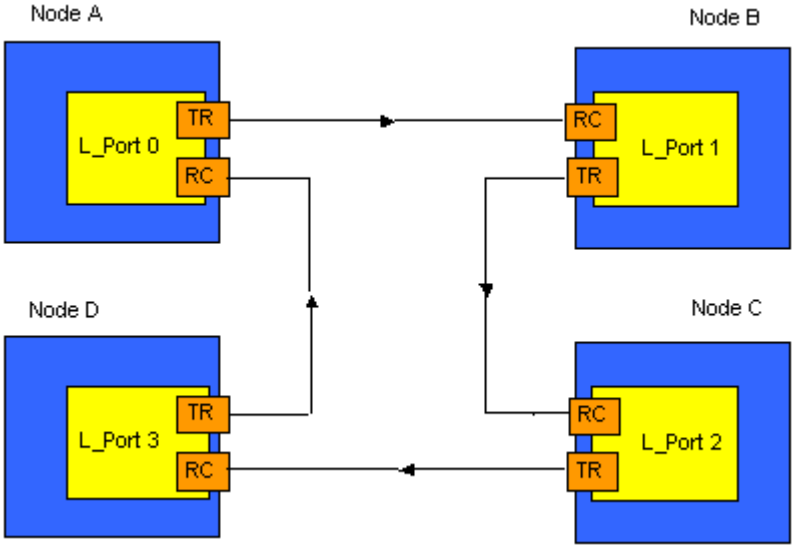


Figure 512 **Arbitrated loop**

Loop construction

An arbitrated loop is constructed by connecting nodes together in a single loop. Loops can be constructed by physically connecting each node in the loop or through the use of a Fibre Channel hub. The transmit of one port is attached to the receive of the next. This is continued until the loop has been formed. A loop can contain 126 N_L Ports and one F_L Port which is used for connection to a Fabric environment. Information that is passed around the loop is repeated by each port and passed on if it is not the required destination.

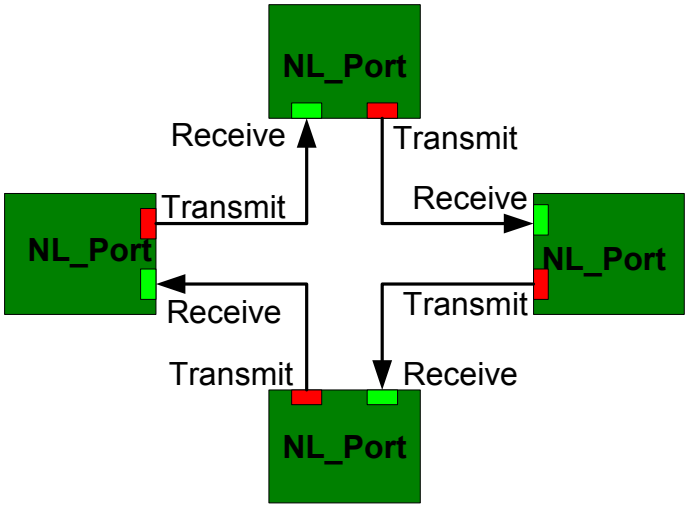


Figure 513 Arbitrated loop with 4 nodes

Figure 513 shows an example of how an arbitrated loop might be constructed. In this loop, there are four nodes and four transceivers, keeping the Fibre Channel circuitry to a minimum. In this configuration, the bandwidth is shared between all nodes on the loop and if any new nodes were added it would be further reduced. In this configuration if any of the ports failed then the complete loop would be lost. With the absence of a hard failure, this loop could be thrown into turmoil if any intermittent type failure should arise. Any *blip* whatsoever would force the whole loop into a re-initialization state and halt all I/O operations until the loop is back to a stable state. This is not desirable in a highly-available storage subsystem which is why today most fibre connectivity to a Symmetrix is through a fabric switch (refer to "Symmetrix and Fibre Channel connectivity" on page 1359). However, with the emergence of dual port fibre devices, the arbitrated loop is a viable option, especially as a back-end solution within a storage array.

Arbitrated loop performance is dependant on a number of factors. Some of these are obvious, such as the number of nodes on the loop. Clearly, in a shared bandwidth medium, the population fighting for that medium will have a direct impact on performance. However, this is not the only way the number of nodes can affect performance.

With the introduction of more nodes, the roundtrip time of the loop is also increased. This is because each frame or sequence may have additional ports to pass through to reach the final destination, and with each additional port adding some latency, overall throughput can be affected. Likewise, the time it takes to win arbitration can increase.

Loops with hubs

Another way of constructing a loop is to use a hub. As the cost of transceivers and hubs have dropped dramatically from when they were first introduced, the benefits obtained with this configuration can outweigh the extra expense. As shown in Figure 514, the loop is constructed within the hub.

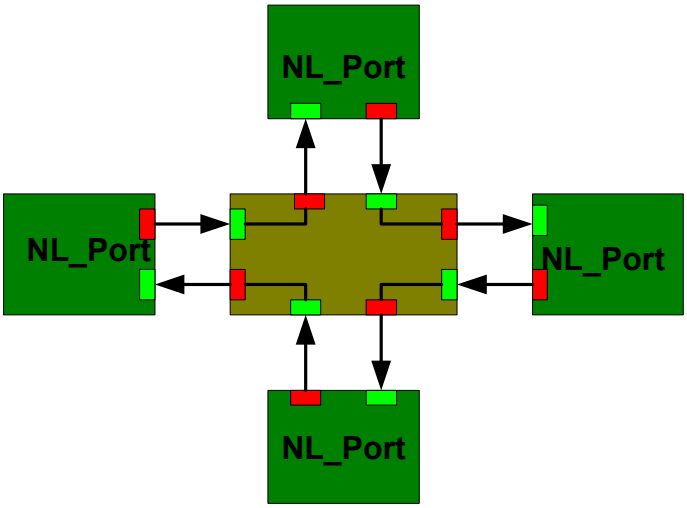


Figure 514 **Arbitrated loop with a hub**

The first improvement is the ease of cabling. In Figure 514, each transmitter had to be routed to the receiver of the next node and so on. As shown in Figure 514, all that is needed to construct the loop is to plug each cable directly to the hub. This example shows that for the same number of nodes in the loop, the number of transceivers will double, adding to the cost.

The hub brings another useful feature to the table. A failure would cause the whole loop to collapse. However, the hub can bypass a faulty port and allow the remaining ports to operate as normal, as illustrated in Figure 515.

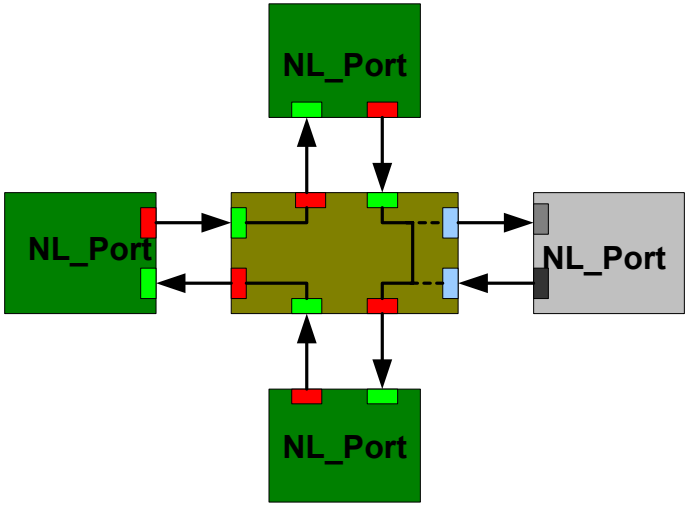


Figure 515 **Hub port bypass**

Abitration

The process through which a fibre port gains sole control of the loop is called *arbitration*. Once a port has control of the loop, and it has opened another port, it is free to transmit frames to that port. Figure 516 on page 1358 shows a loop which is available; in other words devices are free to arbitrate to gain access to this loop.

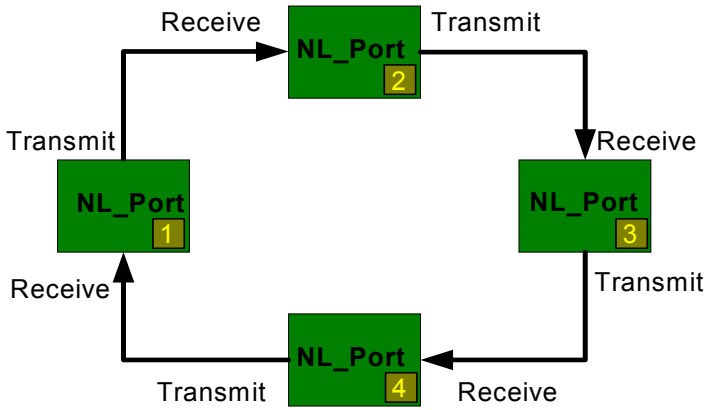


Figure 516 Arbitrated loop (FC-AL)

Arbitration prevents multiple ports from sending frames at the same time on the loop. After a device has won arbitration and opened another device then, in effect, a point-to-point connection has been established. All remaining ports cannot participate in any way other than passing along whatever comes their way until the loop has been closed and is once again available for arbitration to all.

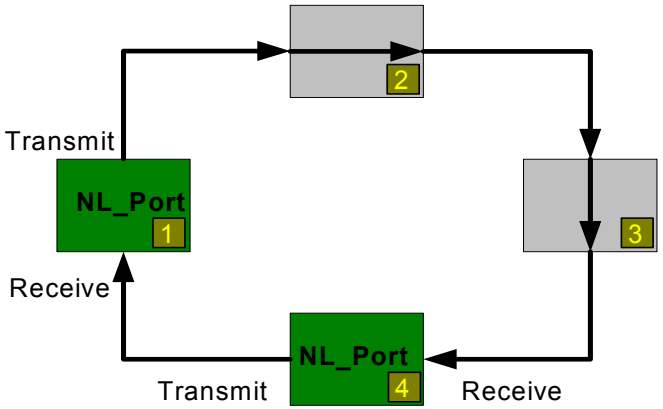


Figure 517 Loop after arbitration won and ports opened

Figure 517 illustrates the state of the loop after a successful arbitration by port one or four. The arbitration winner opened the other port and now ports two and three have been logically removed from the loop forming a virtual point-to-point connection between port one and four.

Symmetrix and Fibre Channel connectivity

EMC first introduced arbitrated loop in a direct connect from HBA to Symmetrix FA configuration (2 Node Arbitrated Loop). (Refer to Figure 518). This was the beginning of Fibre Channel connectivity to the Symmetrix and the configuration was limited to a minimum until all of the teething problems associated with a new technology had been identified and corrected.

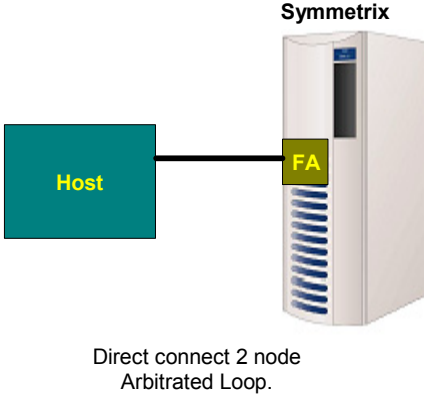


Figure 518 **First implementation on Symmetrix**

Very quickly, hubs were introduced to the configuration and the connectivity was expanded (Figure 519 on page 1360). The hubs provided resiliency in that they provided a bypass circuit on each port which is used to bypass a particular port if there is not a valid signal present, allowing the rest of the loop to continue to operate. Basically, they ensure that a bad port or a broken fibre will not take down the entire loop.

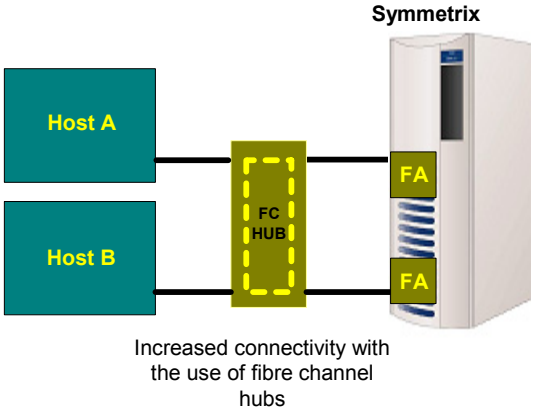


Figure 519 Expanded connectivity

Dual port devices

In dual loop drive systems each disk drive has two ports and thus the drive can be connected on two separate loops. As shown in Figure 520, if a failure occurs on one loop, for example Initiator 1 fails or a port fails on one of the disks, the disks will remain accessible through the second initiator. This could be compared to PowerPath at the disk level.

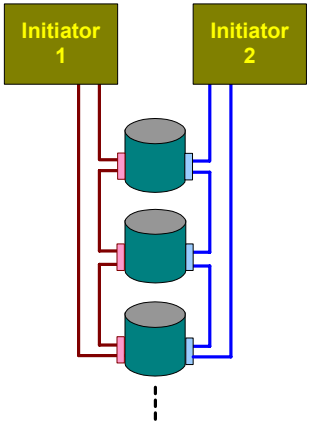


Figure 520 Dual loop disk drives

This is not the only solution available with dual port disks. Hubs could also be included which would provide further connectivity and resiliency options (see Figure 521).

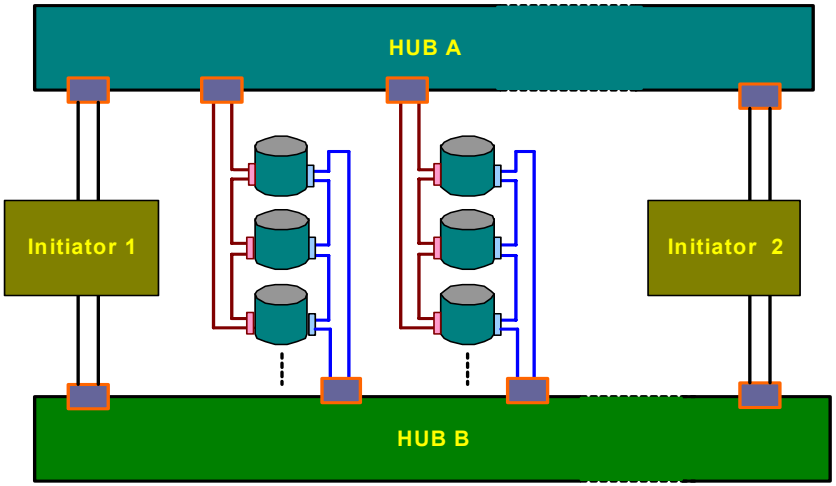


Figure 521 **Highly available dual port disk solution**

Arbitrated loop addressing

Fibre Channel specifies a three-byte field for the address used in routing frames. In arbitrated loop, only one of these three bytes (least significant 8 bits) is used for the address which is known as the Arbitrated Loop Physical Address (AL_PA). This address is used in the Source ID (S_ID) and Destination ID (D_ID) of frames transmitted in the loop.

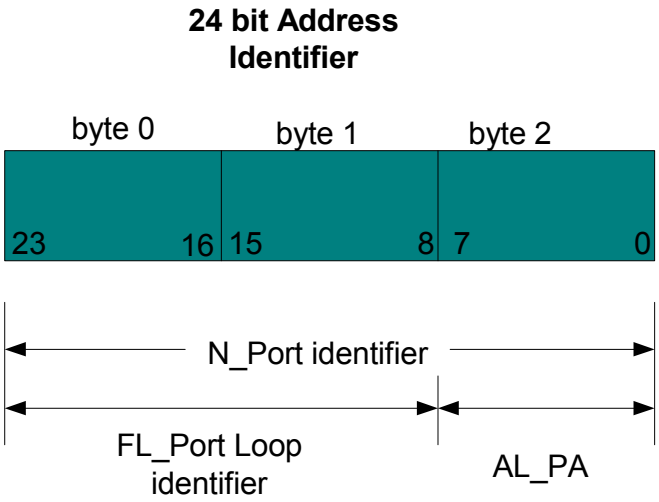
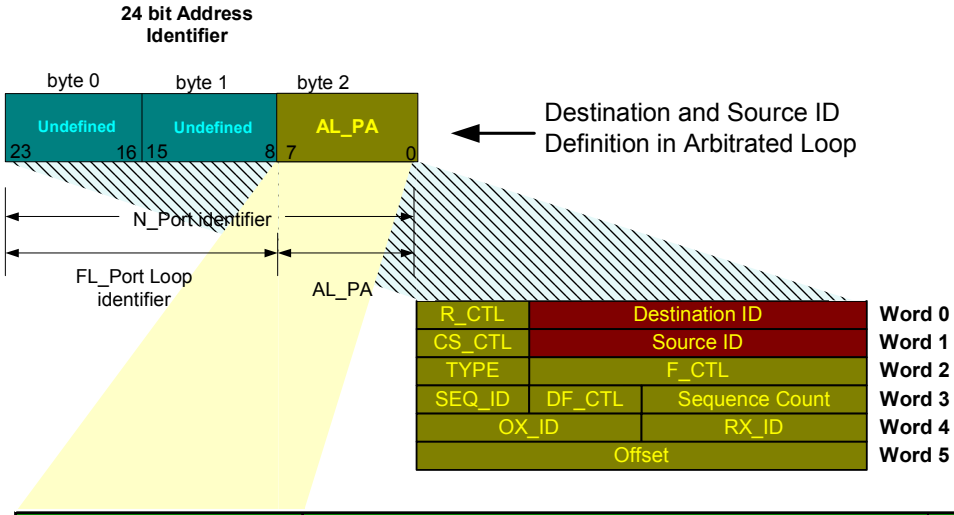


Figure 522 Loop addressing

Figure 522 shows the full 24-bit address defined by the Fibre Channel standard. Eight (8) bits used by the AL_PA. Bits 8 to 23 are used for the FL_port identifier and the full 24 bits are used by an N_Port in a fabric switch environment.

The AL_PA values used are limited to characters that result in neutral disparity after encoding. AL_PA 00 is reserved for FL_Port and the remaining 126 AL_PA are distributed irregularly between 01 and EF. If two ports arbitrate to get access to the loop at the same time then only one port can be given access. This is decided by assigning priorities to the AL_PA addresses with 00 having the highest and 01 through to EF having decreasing priority.

Figure 523 shows AL_PA priorities.



AL_PA	Assigned to	Priority
00	FL_Port (1 per Loop)	Highest
01 to EF	Available for NL_Ports (126 values)	Lowest
F0-FF	Reserved for Primitive Signals and Sequences	

Figure 523 AL_PA priority

Loop ID

The usable AL_PA values are made up of the neutral disparity characters in the range 0x00 to 0xFF and are not distributed regularly. The used AL_PA values range from 0x00 to 0xEF as is shown in Figure 524 on page 1364. When you configure a Symmetrix FA to have a certain AL_PA you do not set an AL_PA value directly but instead specify a Loop ID value in the bin file and this Loop ID then corresponds to a certain AL_PA value. The Loop ID values are sequential between 0x00 and 0x7E and thus can be easier to manage. The correlation of Loop ID to AL_PA is shown in Figure 524.

	Loop ID			Loop ID	
AL_PA	HEX	Decimal	AL_PA	HEX	Decimal
EF	0	0	72	40	64
E8	1	1	71	41	65
E4	2	2	6E	42	66
E2	3	3	6D	43	67
E1	4	4	6C	44	68
E0	5	5	6B	45	69
DC	6	6	6A	46	70
DA	7	7	69	47	71
D9	8	8	67	48	72
D6	9	9	66	49	73
D5	0A	10	65	4A	74
D4	0B	11	63	4B	75
D3	0C	12	5C	4C	76
D2	0D	13	5A	4D	77
D1	0E	14	59	4E	78
CE	0F	15	56	4F	79
CD	10	16	55	50	80
CC	11	17	54	51	81
CB	12	18	53	52	82
Ca	13	19	52	53	83
C9	14	20	51	54	84
C7	15	21	4E	55	85
C6	16	22	4D	56	86
C5	17	23	4C	57	87
C3	18	24	4B	58	88
BC	19	25	4A	59	89
BA	1A	26	49	5A	90
B9	1B	27	47	5B	91
B6	1C	28	46	5C	92
B5	1D	29	45	5D	93
B4	1E	30	43	5E	94
B3	1F	31	3C	5F	95
B2	20	32	3A	60	96
B1	21	33	39	61	97
AE	22	34	36	62	98
AD	23	35	35	63	99
AC	24	36	34	64	100
AB	25	37	33	65	101
AA	26	38	32	66	102
A9	27	39	31	67	103
A7	28	40	2E	68	104
A6	29	41	2D	69	105
A5	2A	42	2C	6A	106
A3	2B	43	2B	6B	107
9F	2C	44	2A	6C	108
9E	2D	45	29	6D	109
9D	2E	46	27	6E	110
9B	2F	47	26	6F	111
98	30	48	25	70	112
97	31	49	23	71	113
90	32	50	1F	72	114
8F	33	51	1E	73	115
88	34	52	1D	74	116
84	35	53	1B	75	117
82	36	54	18	76	118
81	37	55	17	77	119
80	37	56	10	78	120
7C	39	57	0F	79	121
7A	3A	58	8	7A	122
79	3B	59	4	7B	123
76	3C	60	2	7C	124
75	3D	61	1	7D	125
74	3E	62	0	7E	126
73	3F	63			

Figure 524 AL_PA to loop ID chart

Primitive signals and sequences

Arbitrated loop has several *ordered sets* used in loop arbitration and opening and closing of loop circuits (refer to "Opening and Closing the loop" on page 1385). An ordered set is a group of four transmission characters, the first being the K28. Five special characters and the remaining three (data) characters define the meaning of the ordered set. These ordered sets can exist either at the start or end of the frame (in the case of frame delimiters), or can exist on their own (in the case of primitive signals and sequences). For more information on ordered sets, refer to "Ordered sets" on page 769.

Primitive signals are normally used to indicate events or actions at the sending port. A table of these primitive signals is shown in Figure 525. For example, the ARB(x) primitive signal is used by a port in the arbitrating state to indicate that it is arbitrating for access to the loop. The x indicates the AL_PA assigned to that specific port.

Primitive Signal	Abbr	Ordered Set
Arbitrate	ARB(x)	K28.5 D20.4 AL_PA AL_PA
Arbitrate Fairness	ARB(F0)	K28.5 D20.4 D16.7 D16.7
Arbitrate (No AL_PA)	ARB(F7)	K28.5 D20.4 D23.7 D23.7
Clock Synchronization X	SYN x	K28.5 D31.3 CS_x CS_x
Clock Synchronization Y	SYN y	K28.5 D31.5 CS_y CS_y
Clock Synchronization Z	SYN z	K28.5 D31.6 CS_z CS_z
Close	CLS	K28.5 D05.4 D21.5 D21.5
Dynamic Half Duplex	DHD	K28.5 D10.4 D21.5 D21.5
Idle	IDLE	K28.5 D21.4 D21.5 D21.5
Mark	MRK(x)	K28.5 D31.2 MK_TP AL_PS
Open Full-Duplex (Point-to-Point)	OPN(yx)	K28.5 D17.4 AL_PD AL_PS
Open Half-Duplex (Point-to-Point)	OPN(yy)	K28.5 D17.4 AL_PD AL_PD
Open Broadcast Replicate	OPN(fr)	K28.5 D17.4 D31.7 D31.7
Open Selective Replicate	OPN(yr)	K28.5 D21.4 AL_PD D31.7
Receiver Ready	R_RDY	K28.5 D21.4 D10.2 D10.2

Figure 525 Primitive signals

Primitive sequences are used to indicate states or conditions and are normally transmitted continuously until something causes the state to change. A minimum of three consecutive occurrences of the same

ordered set is required before a primitive sequence is recognized and action taken. Figure 526 on page 1366 shows Arbitrated Loop primitive sequences used in link initialization and loop port bypass. Switched fabric uses a different method of link initialization than arbitrated loop, but it does still use primitive sequences.

Primitive Sequence	Abbr	Ordered set
Loop Initialization - F7,F7	LIP	K28.5 D21.0 D23.7 D23.7
Loop Initialization - F8,F7	LIP	K28.5 D21.0 D24.7 D23.7
Loop Initialization - F7,x	LIP	K28.5 D21.0 D23.7 AL_PS
Loop Initialisation - F8,x	LIP	K28.5 D21.0 D24.7 AL_PS
Loop Initialization - reset	LIPyx	K28.5 D21.0 AL_PD AL_PS
Loop Port Enable	LPEyx	K28.5 D5.0 AL_PD AL-PS
Loop Port Enable All	LPEfx	K28.5 D5.0 D31.7 AL_PS
Loop Port Bypass	LPByx	K28.5 D9.0 AL_PD AL_PS

Figure 526 Primitive sequences

An AL_PA identifies either a source or destination port in arbitrated loop. In some cases it is necessary to identify whether it is the source or destination AL_PA that is being referred to. When it is necessary to identify a destination port, the term **AL_PD** is used. In the case of the source port, the term **AL_PS** is used. Figure 526 shows different types of LIP (Loop Initialization Primitive) sequence used in loop initialization. Each is discussed below.

Loop Initialization LIP(F7, F7). A port transmitting LIP(F7, F7) indicates that the port in the initializing state is requesting loop initialization but does not have a valid AL_PA. This mainly occurs when a device is hot-plugged into a Loop or when a port that was non-participating wants to become participating and requires an AL_PA to do so.

Loop Failure LIP(F8, F7). A port transmitting LIP(F8, F7) indicates that the port in the initializing state is requesting a loop initialization due to a loop failure. The port does not have an AL_PA and uses F7. This could occur if a non-participating loop port without an AL_PA detects a loop failure or a node in the process of getting an AL_PA detects a loop failure.

Loop Initialization LIP(F7, AL_PS). This LIP indicates that the loop port identified in the AL_PS value is requesting loop initialization. This can be used if the port detects a performance degradation, arbitration wait time-out, or for another unspecified reason.

Loop Failure LIP(F8, AL_PS). This LIP indicates that the loop port identified in the AL_PS value has detected a loop failure. This may occur when a loop interconnection has failed, a loop port has failed, a loop port has been powered off or removed from the loop when no bypass circuit is present, or if the bypass circuit fails.

Selective Reset LIP(AL_PD, AL_PS). The selective reset LIP is used to perform a vendor specific reset at the loop port specified in the AL_PD value. The AL_PS value indicates the port that originated the request. This LIP could be used during error recovery to reset a port that is in a hung state.

Along with the loop initialization primitive sequences, there are a number of sequences to set and reset the LP_Bypass variable in the Loop Port State Machine (LPSM). With this variable set, the LPSM retransmits frames and does not attempt to arbitrate or participate in the loop. The state of this variable is also used to control an optional port bypass circuit to electrically bypass the loop port.

Loop Port Enable LPE(yx), LPE(fx). These primitive sequences cause either a designated port (yx) or all ports (fx) to reset the LP_Bypass variable and deactivate a control line to an external port bypass circuit if present.

Loop Port Bypass LPB(yx), LPB(fx). These primitive sequences, when received, cause the designated port (yx) or all the ports (fx) to set the LP_Bypass variable and optionally activate a control line to an external port bypass circuit if present.

Loop Port State Machine (LPSM)

A port on an arbitrated loop is required to behave in a certain manner in order to operate correctly on the Loop. The LPSM defines this behavior (refer to Figure 527 on page 1368). The particular states that can occur on a loop include: initialization, arbitration, open circuit, close circuit, and implement fairness. These various states are implemented using specific ordered sets which are processed by the LPSM of the ports on the loop.

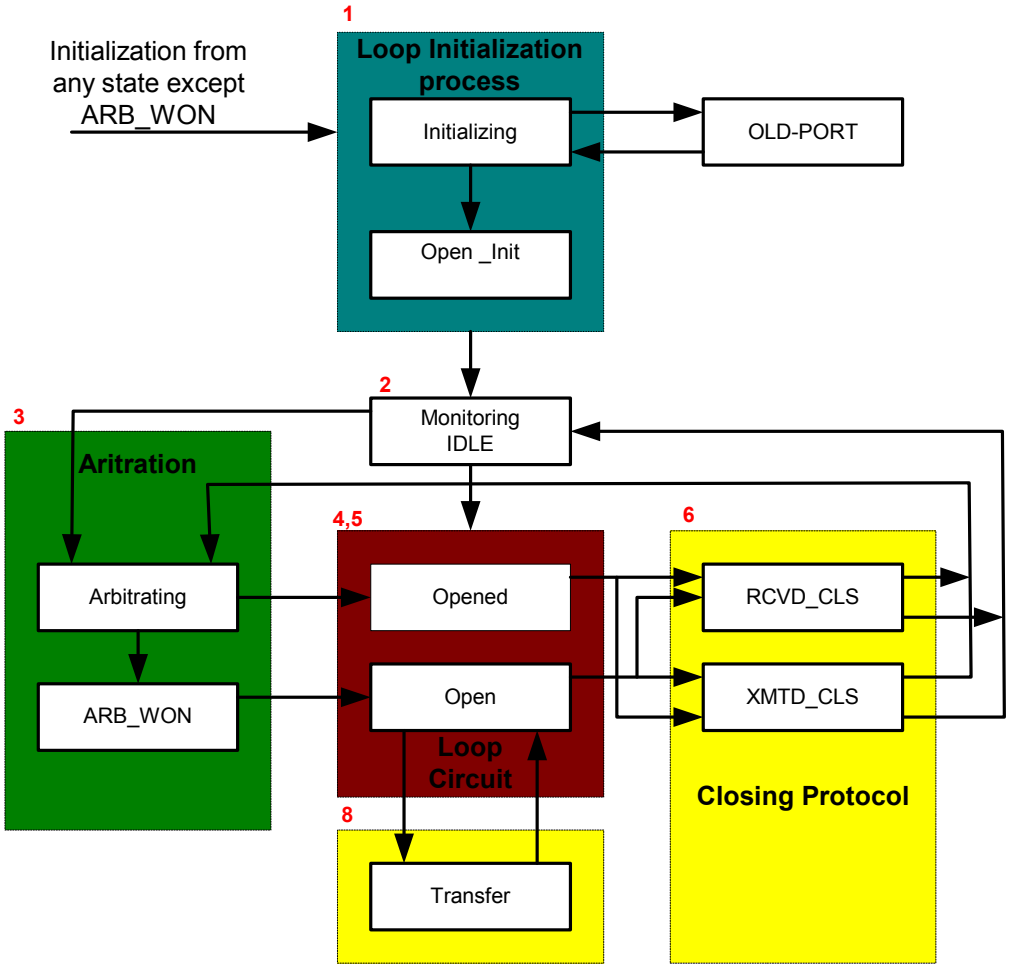


Figure 527 Loop Port State Machine (LPSM)

The basic operations of a Loop Port State Machine are:

- When a port is first attached to the loop it enters the initializing state and starts loop initialization where each port transmits a LIP continuously and monitors for a LIP returning. On receipt of a LIP the ports go into a Open-Init state where most of the Loop initialization steps take place. If the loop is not operational the port may go to into the Old_Port state disabling Arbitrated Loop functionality and begin operating as an N_Port.

- If the ports have no work to do they are in the monitoring state transmitting and receiving Idles.
- When the loop port requires access to other ports in the loop it enters the arbitrating state and begins arbitrating for access. When the port wins arbitration it enters the ARB-WON state.
- When the port needs to access another port on the loop it sends an OPN to select the destination port. The source port goes into the OPEN state and the destination port goes into the OPENED state.
- Once both ports are in these states transfer of frame can begin.
- When either of the ports has completed its transfer and wants to close the connection it sends a close (CLS). The port that sends the CLS enters the transmitted close state and the when the other port receives the CLS it enters the received close state. It transmits its remaining frames, if necessary, as long as it has available credit and then transmits a CLS back to the originator. The two ports are now logically disconnected from the loop and enter the monitoring state again.
- If a loop port is in the monitoring state or arbitrating for the loop and receives an OPEN from another port it enters the OPENED state.
- There is another state that a port can enter called the *transfer state*. If a port has frames to send to multiple ports and is finished communicating with the first of these ports it can send a CLS, move to the transfer state, and once it receives the CLS from the first port it can open the second port without going through an arbitration cycle.

Loop initialization

Before discussing all the steps in detail that take place during a loop initialization, it is important to understand why this step must be done and exactly what is accomplished during this step.

Overview

The loop initialization step performs a number of functions in arbitrated loop including the assignment of addresses (AL_PAs) to loop ports, notification that the configuration may have changed, and notification of a loop failure. Events that can cause a loop initialization include: a) if a port was powered on it may need to acquire an AL_PA and notify other ports on the loop that the

configuration has changed; or b) a port detects a physical connection problem and begins the loop initialization process to notify other ports and check if the loop is still operational. The main steps involved in loop initialization are listed in Figure 528.

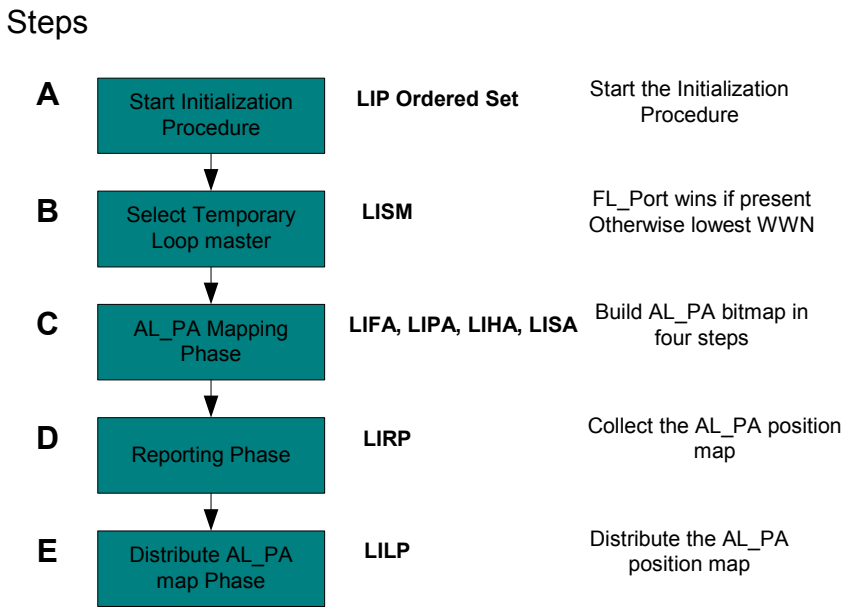


Figure 528 Initialization procedure

Each node on a loop is capable of initiating loop initialization by transmitting one of the LIP sequences shown in Figure 526 on page 1366. When a node transmits this LIP sequence all the other ports on the loop recognize this and enter the **OPEN_ INIT** state and retransmit the LIP along the loop. This occurs until the LIP arrives back at the port that is initiating the initialization and it too enters the 'OPEN_ INIT' state.

With all ports on the loop now in the OPEN_ INIT state, the next step is a process to select a port to become the temporary loop master. This step is initiated by each node on the loop that has entere 'OPEN_ INIT' state continuously sending out LISM (Loop Initialization Select Master) frames. LISM frames have a certain format, detailed Figure 529 on page 1371, and serve the purpose of allowing each node on the loop a chance at becoming the temporary loop master.

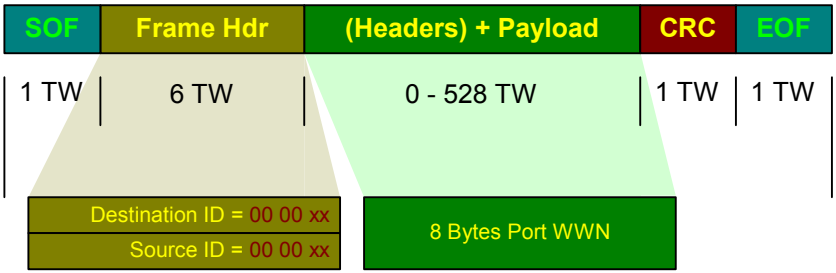


Figure 529 LISM frame format

The structure of the LISM frame is shown in Figure 529 with the S_ID and D_ID fields in the frame header showing a value of '0000xx'. The least significant byte is a value of 00 if the port is an FL_Port, or EF if the port is an NL_Port which identifies the port type. The payload of the LISM frame contains the 8 byte World Wide Name of the port.

If there is a single FL_Port on the loop it becomes the loop master. If there are more than one FL_Port then the one with the lowest port worldwide number becomes the master. The reason the FL_Port becomes the master is based on the fact that the FL_Port is part of a fabric and therefore has more knowledge of the configuration. If there are no FL_Ports present then the NL_Port with the lowest port worldwide number is selected as temporary loop master. Normally, it is an HBA that becomes the loop master as its worldwide number is lower than that of a Symmetrix FA.

The process involved at the LISM stage involves each port transmitting LISM frames and each port checking the LISM frames it is receiving for the port type field which is denoted in the least significant byte of the S_ID and D_ID fields. If an NL_Port receives a LISM frame from an FL_Port it stops transmitting its own LISM frames and begins retransmitting the LISM of the FL_Port. If an FL_Port receives a LISM frame from an NL_Port it discards the received frame and begins transmitting its own frame. If the port type in the S_ID and D_ID fields of the frame header is the same as that of the receiving port, the port worldwide number in the payload of the LISM frame is compared with that of the receiving port. If the port worldwide number in the received LISM frame is higher than that of the receiving port, the port discards the frame and the port continues to transmit its own LISM. If the port worldwide number in the received LISM frame is not higher, it stops transmitting its own LISM frame and begins transmitting the received frame. Eventually one of

the loop ports will receive back around the loop its own LISM frame and when this happens this port becomes the temporary loop master. This port then begins transmitting the ARB(F0) primitive signal to inform the other ports that the LISM procedure has been completed and a loop master selected.

The next step in the process is *AL_PA assignment*. This process involves the assignment of Arbitrated Loop Physical Addresses (AL_PA) to each port on the loop. A port's AL_PA is its Fibre Channel address on the loop and is used to identify it on the loop and is put in the S_ID of the frame header of all data frames that are transmitted by the port and is in the D_ID of all data frames that are to be received by the port. This is a multistep process where addresses are assigned using four distinct steps depending on the particular ports operation. These four steps are discussed beginning on page 1373 and all use the concept of populating an AL_PA bitmap with a value depending on whether that corresponding AL_PA is assigned or not.

In arbitrated loop there are 127 possible addresses on a loop (126 for NL_Ports and one for the FL_Port). To identify which AL_PA values have been taken in a loop a 128 bit (four word) map is used where each bit corresponds to a certain AL_PA. If a bit is set to 1 then the corresponding AL_PA is assigned and that address is in use. If the bit is 0 then the corresponding AL_PA is available to be acquired. Word 0 bit 31 of the AL_PA bit map is the login-required bit (L-bit) which is set by an FL_Port to indicate that the configuration has changed and that all ports are logged out.

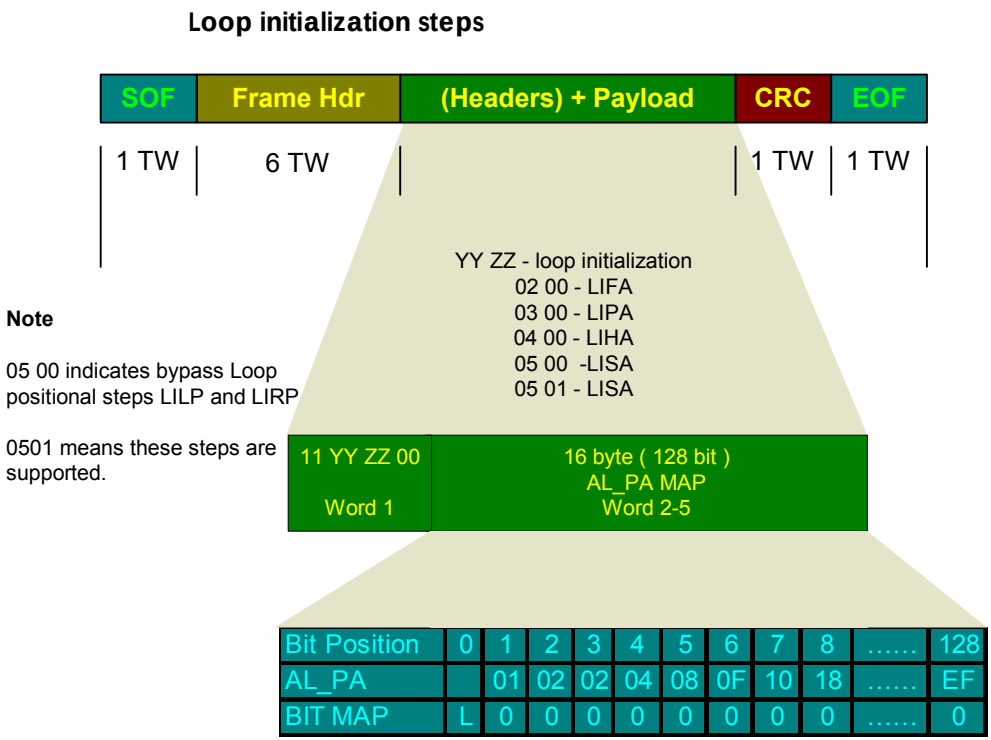


Figure 530 **LIFA / LIPA / LIHA / LISA frames**

Step 1. The first step in the AL_PA assignment is the **LIFA** (Loop Initialization Fabric Address) which basically allows public ports that had previously been logged in with the fabric (FL_Ports) to reclaim the AL_PA they had been using prior to loop initialization. The temporary loop master first initializes the AL_PA bitmap to all ‘0’s and builds the LIFA frame with this bitmap in payload words two through five and the loop initialization identifier in word 1, as is shown in Figure 530.

If the loop master had a fabric assigned address prior to the loop initialization it sets the bit for that AL_PA in the bitmap and then transmits this frame. If the next port on the loop had a fabric assigned AL_PA it also sets the corresponding bit in the AL_PA bitmap (if not already set) and retransmits the frame. If the next loop port did not previously have a fabric assigned AL_PA then it would leave the AL_PA map unchanged and retransmits the frame. This continues until every port on the loop processes the LIFA frame and is received back by the temporary loop master.

Step 2. The next step of the process is **LIPA** (Loop Initialization Previous Address) where private ports (NL_Ports) that had an AL_PA prior to loop initialization can reclaim the same AL_PA. This step is initiated by the temporary loop master by changing the loop initialization identifier in word 1 of the LIFA frame it received to LIPA identifier and then transmitting this around the loop. Each port starting with the loop master checks to see if it had a nonfabric assigned AL-PA prior to loop initialization. If it had, it first checks to see if that bit is set in the AL_PA bit map and if it is not set then sets it to 1 and reclaims the AL_PA. If the bit is set then the port would have to wait for a soft-assigned address. This continues until every port on the loop has processed the frame and it is received back by the loop master.

Step 3. The next step of the initialization process is for ports that did not have an AL-PA prior to loop initialization but do have a preferred AL_PA that is set by either jumper settings or some other configuration method. This step is known as the **LIHA** (Loop Initialization Hard Address)/ The Symmetrix FA would be an example where the AL_PA can be configured in the bin file by setting a corresponding Loop ID value. The temporary loop master changes the received LIPA frame loop initialization identifier in word 1 of the payload to LIHA which indicates this is now a LIHA frame which will be transmitted around the loop. Each port that did not have an AL_PA prior to link initialization but does have hard assigned AL_PA checks the bit corresponding to that AL_PA in the LIHA frame payload words two to five and then sets the bit and claims that AL_PA. If the bit is already set then the port would again need to wait for a soft-assigned address. This continues until every port has processed the LIHA frame and it is received back around the loop by the temporary loop master.

Step 4. The final step in acquiring an AL_PA is the **LISA** (Loop Initialization Soft Address) where the port may select the first available AL_PA in the bitmap. Once the loop master has received back the LIHA frame it changes the identifier in word 1 from LIHA to LISA and leaves the AL_PA map unchanged. It then transmits the frame and every port that does not have an AL_PA scans the bitmap in the LISA frame to find the first available AL_PA and then claims that soft-assigned address by setting that bit in the AL_PA bitmap. The AL_PA values are normally assigned starting at the most significant bit of the AL_PA map and proceeding to the least significant bit. This provides the most efficient ordering of AL_PA values around the loop. For the best performance the AL_PA values

should be arranged in descending priority in the direction of information flow around the loop. AL_PA 00 is the highest priority.

If after these steps a port has not acquired an address, for example if all the available AL_PA s have been taken or if a ports hard address is already in use, then the port will enter the non- participating mode.

If the port does not support loop positional mapping steps which follow AL_PA assignment it sets the third byte of the loop initialization identifier (Word 1) to 00.

A method assigning addresses during the LISA step is to have initiators acquire AL_PA values in the higher priority end of the AL_PA bitmap starting at the most significant bit and proceeding to the least significant bit and targets do the opposite. However, this does not lead to the most efficient ordering of the AL_PAs on a loop for target devices as they would get AL_PAs assigned starting with the lower priority.

A preferred method of assigning AL_PA values during the LISA process is to have a range of AL_PA values at the higher order end of the AL_PA bit map reserved for initiators and have targets assign soft addresses beginning at the end of this range. This would avoid the AL_PA ordering problem in that the initiators would get the higher priority AL_PA values in descending order and the targets would get the lower priority AL_PA values also in descending order around the loop.

There are two additional steps in the initialization stage that are needed to provide information on the positioning of AL_PAs in the loop for managing the configuration and problem analysis. These two steps are Loop Initialization Report Position (LIRP**)** and Loop Initialization Loop Position (LILP):

- The **LIRP** step involves building a map of the AL_PA values according to their position on the loop relative to the temporary loop master. The temporary loop master begins this process by building a LIRP frame with the structure shown in Figure 531 on page 1376.

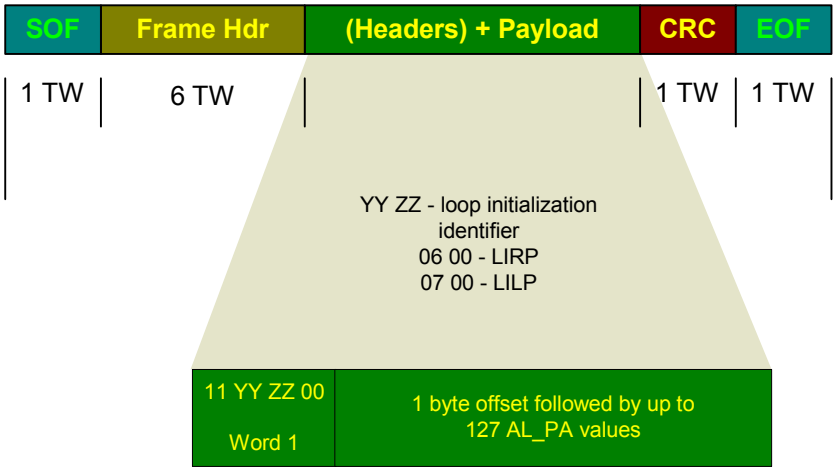


Figure 531 **LIRP/LILP frame format**

The basic structure of the frame payload is a one word loop initialization identifier, a one byte offset and up to 127 AL_PA entries. The temporary loop master first of all initializes the structure by setting the offset to 01 and storing its AL_PA at offset 01 in the AL_PA map and puts FF in all the other remaining positions. This frame is then passed to the next port on the loop, the offset is incremented by 1 and the next port stores its AL_PA at that location in the map. This process continues until the LILP frame comes back around the loop to the temporary loop master.

- When the loop master receives the LIRP frame back it changes the identifier in the payload of the frame to a **LILP** and retransmits the entire frame so each port can have a copy of the AL_PA positional map. Once this frame comes back around the loop the loop initialization is complete.

Login process

Once the loop is initialized each port has acquired an address but each initiator (HBA in server) does not know what target devices are on the loop. For each HBA to discover what targets are on the loop it needs to perform some extra steps. These steps provide ports with a means of exchanging information about each other that is used to control any communication that is initiated by these ports. Fibre Channel provides three different types of login that can occur between ports by the use of the following extended link services:

- Fabric Login (FLOGI)
- N_Port Login (PLOGI)
- Process Login (PRLI)

The **Fabric login** process is used in a switched fabric environment to allow an N_Port establish a session with the fabric. During this step both the N_Port and the fabric exchange parameters with each other and identify themselves to each other. This step also assigns an address to the attached N_Port.

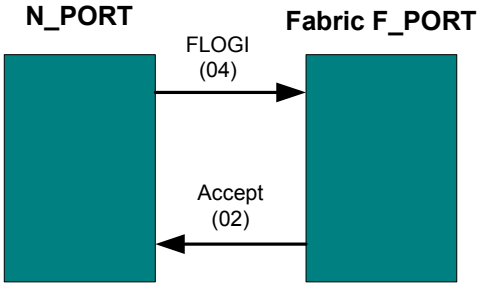


Figure 532 FLOGI and Accept

Figure 532 shows the exchange of information that takes place with the N_Port sending the initial FLOGI frame (Command code 04) with all its information and the Fabric returning the Accept (Command code 02) with its associated parameters.

In arbitrated loop the Fabric login does not take place and the ports exchange service parameters with **N_Port login** , which is also known as PLOGI. Service parameters are basically information regarding the FC-2 capabilities of a port such as maximum frame size that can be received or end-to-end credit values. Again the N_Port login is performed by the initiator sending a PLOGI request frame and the target returning an accept.

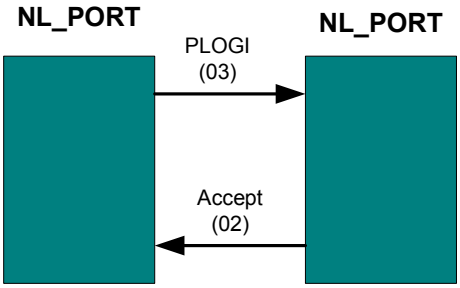


Figure 533 PLOGI and Accept

Figure 533 shows the PLOGI and the Accept back from the target with their respective command codes. The payload of both the request and the accept PLOGI frames contain the service parameters of the initiator and the target.

Process Login (PRLI) is the process to allow two ports exchange service parameters relating to the FC-4 type they are using. Specifically what you would see here is SCSI type information being exchanged by both ports in a request and accept fashion as is shown in Figure 534.

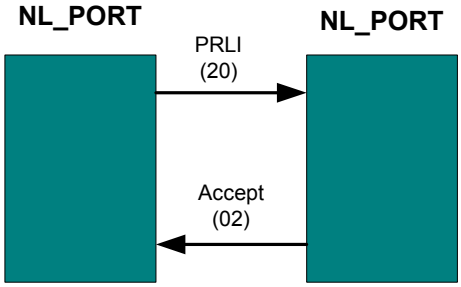


Figure 534 PRLI and Accept

Some of the parameters exchanged during process login are whether the port is an initiator or target, transfer ready being used for reads or writes, and the FC4 type being used (0x08 for SCSI in these implementations).

Arbitration process

Overview

Arbitration is the process that allows a loop port to gain access to the loop and ensure that only one port at a time is sending information. If this was not the case then ports could send data at the same time and interfere with each other. If two ports arbitrate at the same time to get on to the loop then there has to be a mechanism to handle multiple simultaneous requests. This is done by allowing the port with the lowest value AL_PA to have priority over ports with higher value AL_PAs. This could cause a situation where higher priority loop ports monopolize the loop, but to handle this a fairness mechanism is incorporated into the arbitration protocol (refer to "Access fairness" on page 1383).

This section discusses the following topics:

- "Arbitration process steps" on page 1380
- "Access fairness" on page 1383
- "Opening and Closing the loop" on page 1385

Before going into the arbitration process the concept of *Fill Words* needs to be understood. In Fibre Channel, even if there are no frames being transmitted, idle words are continuously transferred around the loop. When frames are being sent around the loop a certain amount of idles are also transmitted between these frames. These idles are a form of fill word and must exist between frames in a Fibre Channel environment. During arbitration it is necessary for a port to remove a fill word it has received and transmit a different fill word in its place. This is known as *fill word substitution*. When a port needs to transmit a fill word it will use the value contained in the current fill word. For example, if a port wants to start arbitrating on a loop which is not being used it needs to change the current fill word from ARB(x) to IDLE and once this is done the port can transmit ARB(x) instead of the received IDLEs.

Once the loop initialization phase has completed the loop is filled with IDLEs as each port is in the monitoring state. Once a port needs access to the loop it has to arbitrate to get this access. The six steps involved are discussed next, using an example of a loop containing four ports with one port arbitrating as shown in Figure 535.

Arbitration process steps

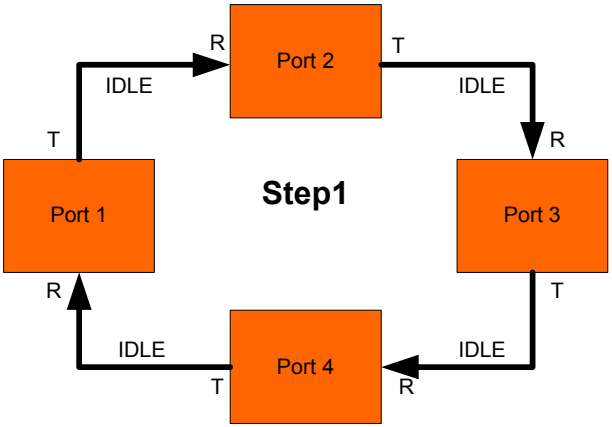


Figure 535 **Arbitration Step 1**

Step 1 of the Arbitration process, shown in Figure 535, shows a loop which is filled with idles and all ports in the monitoring status. The current fill word on all the ports is IDLE and thus received IDLEs are substituted with the current fill word which is IDLE on all the ports.

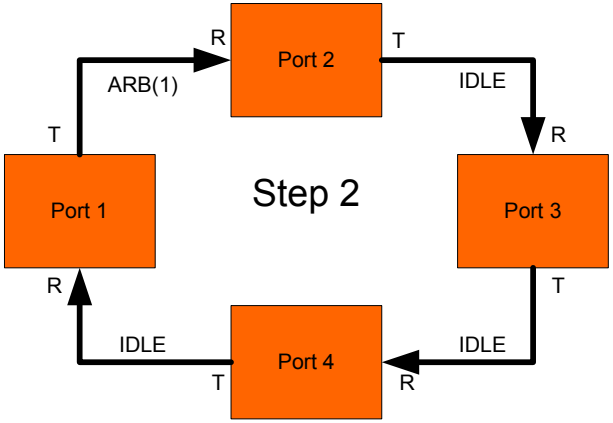


Figure 536 **Arbitration Step 2**

Step 2 (Figure 536) shows Port 1 arbitrating for access to the loop by changing its current fill word to ARB(1) and substituting ARB(1) for every received IDLE. If Port 1 receives any ARBs, it needs to check the

priority of these ARBs and then either discard if they are a lower priority than ARB(1) or substitute if they are a higher priority.

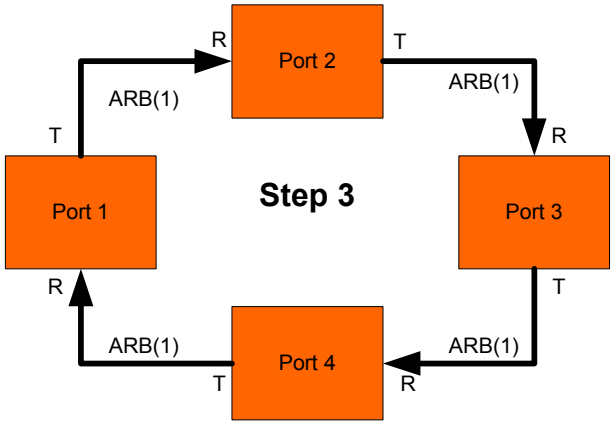


Figure 537 **Arbitration Step 3**

In **Step 3** (Figure 537), once the ARB(1) is received by port 2 it changes its current fill word to ARB(1) and the ARB(1) is transmitted to the next port on the loop. The current fill word on the remaining ports is also changed in this fashion allowing the ARB(1) to propagate around the loop.

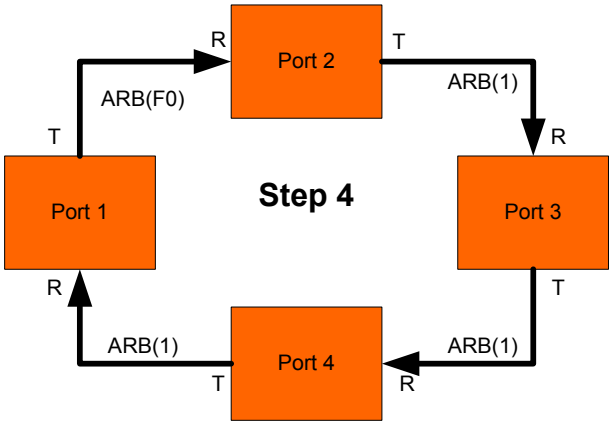


Figure 538 **Arbitration Step 4**

In **Step 4** (Figure 538 on page 1381), once the ARB(1) is received by Port 1 it has won arbitration and immediately changes its current fill word to ARB(F0). Port 1 now will discard any received ARB(x) and

thus this prevents any other loop port from winning arbitration as its ARB(x) could not make it around the loop.

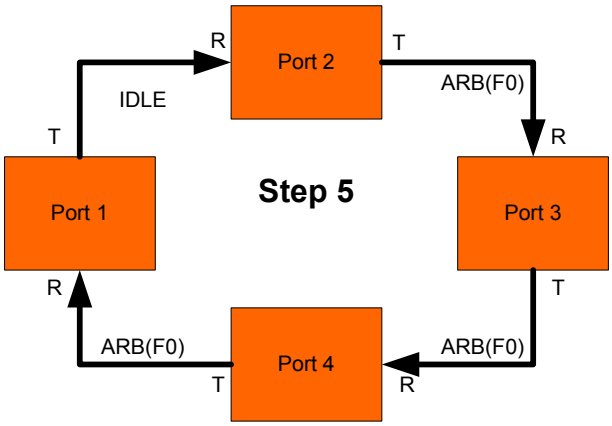


Figure 539 **Arbitration Step 5**

In **Step 5** (Figure 539) each port that receives the ARB(F0) changes its current fill word to ARB(F0) and transmits the ARB(F0) whenever required. This is transmitted by the winner of the arbitration to determine if any other ports are arbitrating. If there are other ports arbitrating then they substitute their ARB(x) for the ARB(F0).

As long as no other port attempts to start arbitrating, the ARB(F0) makes it around the loop to port 1 and the current fill word on port 1 is changed to IDLE.

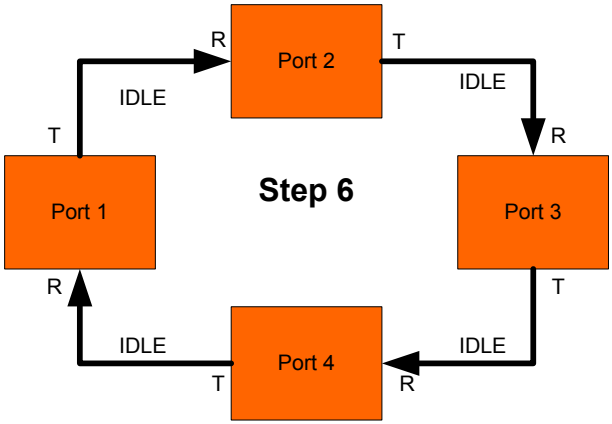


Figure 540 **Arbitration Step 6**

In **Step 6** (Figure 540) each port that receives the IDLE changes its current fill word to IDLE and as long as no other port is arbitrating the IDLE's travel around the loop. As long as Port 1 owns the loop it discards any received ARB(x) or IDLE and continues to send IDLE as its fill word. As it is discarding any received ARB(x) it prevents any other loop port from winning arbitration.

Figure 540 details how a single port arbitrating win ownership of the loop but one must also consider what happens if multiple ports begin arbitrating at the same time. This is handled by the fact that AL_PAs have different prioritiesand in this case the higher priority AL_PA (lower numerical value) will win the arbitration.

Access fairness

Arbitrated loop assigns different priorities to different AL_PA values and thus the situation could occur on a loop where the higher priority AL_PA ports could dominate the loop and prevent access from the lower AL_PA ports. To prevent this, **access fairness** is used to counteract AL_PA priority and ensure every loop port has an equal chance to access the loop. Access fairness does not mean how long a port that has won arbitration can use the loop, but rather ensures that once a loop port has won arbitration and given up control of the loop it will not arbitrate again until all other loop ports that are arbitrating have had an opportunity to access the loop. Ports that are not arbitrating do not affect the access fairness protocol. A loop port that

follows the access fairness protocol is called a *fair* port and a port that does not is called an *unfair* port.

Access fairness applies in a certain time interval from when the first port on a loop wins arbitration to when no other port is arbitrating for ownership of the loop. This is controlled by each port on the loop setting a control variable called the *Access variable*. When a port wins arbitration it sets the access variable to 0 to ensure it cannot arbitrate again in the current fairness window until all other ports arbitrating have had a chance to win arbitration. The winning port knows what other ports are arbitrating as it is sending ARB(F0) and the ports that are trying to arbitrate are substituting their ARB(x) for this.

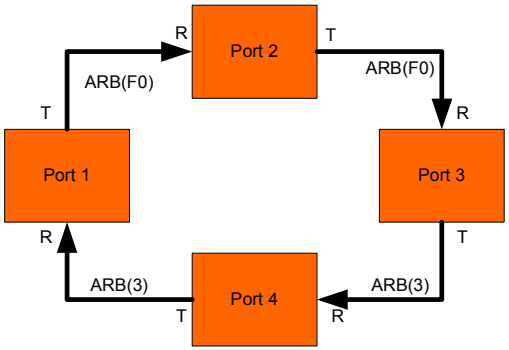


Figure 541 Access fairness window

If the winning port receives around an ARB(F0) it knows that no other port was arbitrating during the time that ARB(F0) was being sent around the loop which means the end of the current access fairness window. At the end of a fairness window all fair ports that have their access variable set to 0 to prevent them arbitrating in that fairness window have this bit set again allowing them to arbitrate once more.

If the winning port does not receive back an ARB(F0), but instead an ARB(x) from a port trying to arbitrate, then it knows that another port is also arbitrating for access to the loop. Once the current winner relinquishes ownership of the loop it cannot arbitrate in the same fairness window so the other port arbitrating has a chance to win arbitration. This is shown in Figure 541 where Port 1 has won arbitration but Port 3 is also arbitrating for ownership of the loop. Access fairness will ensure that Port 3 will win arbitration once Port 1 has relinquished ownership.

Opening and Closing the loop

After a port has won arbitration it needs to select a destination port before it can send frames to it. It does this by opening a connection with the destination port by sending an open (OPN) ordered set with the AL_PA of the destination port. Once the destination port has received the OPN ordered set the loop circuit between both ports is open. At this stage all other ports on the loop behave as repeaters and allow frames to pass through on the way to the destination port.

Opening the loop

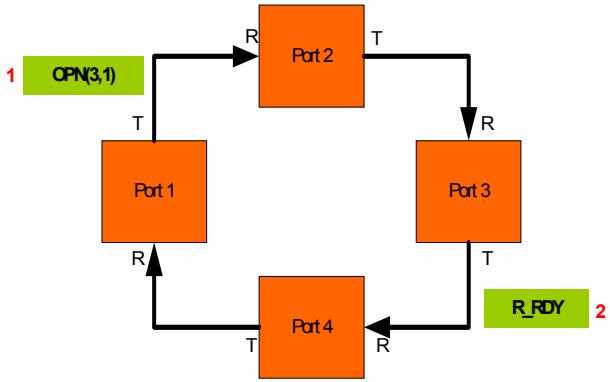


Figure 542 **Opening a loop circuit**

Figure 542 shows Port 1 on a loop opening a circuit with Port 3 by sending an OPN ordered set with the destination port of 3 and once this is received Port 3 responds saying it was ready to receive a frame by sending back the R_RDY.

As can be seen in Figure 543, there are different open ordered sets to open a circuit between ports.

Open Full-Duplex (Point-to-Point)	OPN(yx)	K28.5 D17.4 AL_PD AL_PS
Open Half-Duplex (Point-to-Point)	OPN(yy)	K28.5 D17.4 AL_PD AL_PD
Open Broadcast Replicate	OPN(fr)	K28.5 D17.4 D31.7 D31.7
Open Selective Replicate	OPN(yr)	K28.5 D21.4 AL_PD D31.7

Figure 543 **Open primitive signals**

- **OPN(yx)** is used to open a connection between two ports where the frame transmission will be in both directions. This operation allows both ports transmit at the same time and gives maximum utilization of the bandwidth since frame flow can be in both directions simultaneously.
- **OPN(yy)** is used when the originator of the open only wants to send frames to the destination port and wants to prevent the destination from sending frames to it other than link control frames such as Acknowledge, Busy, or Reject.
- **Open Replicate** is a mode ordered set used to implement multicast and broadcast capabilities in an arbitrated loop environment. The open ordered sets set the mode on the destination port to replicate which basically causes it to retransmit all received information and process the frames that have its address in the frame header.
  - **OPN(yr)** is used to selectively set replicate mode on a port specified by the AL_PD field of the ordered set.
  - **OPN(fr)** causes all ports on the loop to set replicate mode except the port transmitting the OPN(fr) ordered set.

Closing the loop

When all the frames are transmitted from one port to another then the ports may close the loop by sending a **Close** ordered set from one port to the other. When a port is finished transmitting frames it sends a CLS to the destination port and it enters the transmitted close state and waits for the other port to finish. When a port receives a CLS ordered set it enters the Received Close state finishes its frame transmission and sends a CLS. Once the port that initiated the close receives back a CLS from the other port it either enters the monitoring state to allow another port win arbitration or it may send an open to establish a connection with another port.

The following three steps explain how to close a loop.

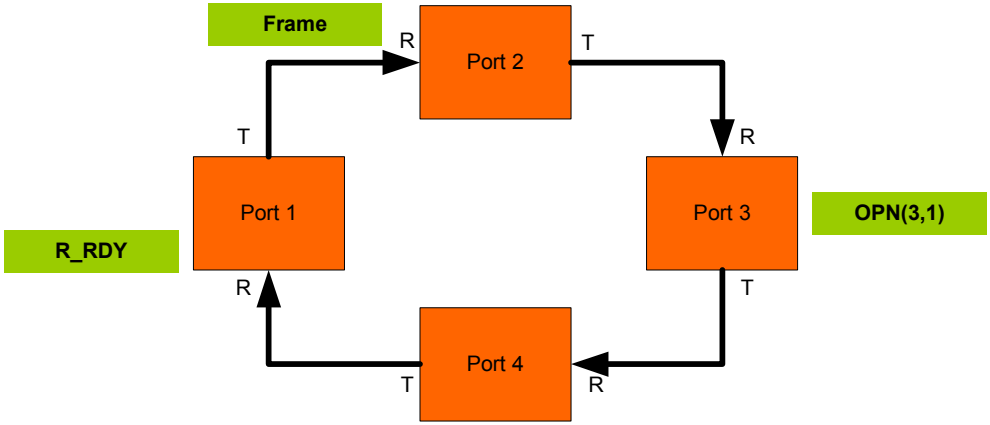


Figure 544 **Close loop: Step 1**

Step 1. Figure 544 shows an existing open loop between Port 1 and Port 3 Port 3 has received the OPN, responded by sending the R_RDY, Port 1 has received the R_RDY and is transmitting a frame to Port 3.

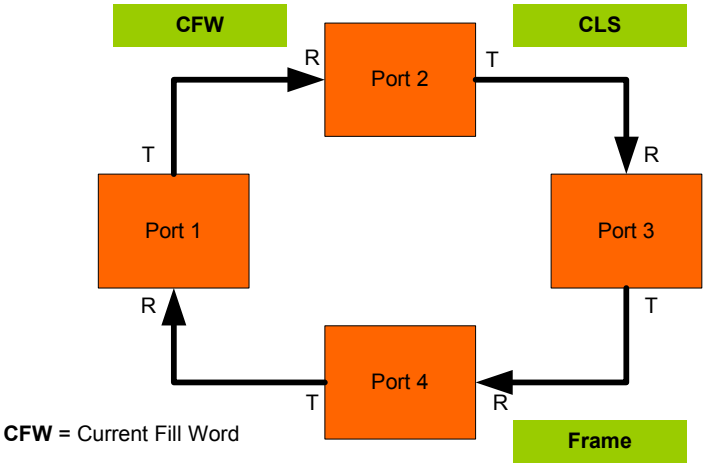


Figure 545 **Close loop: Step 2**

Step 2. Figure 545 shows Port 1 sending a Close to Port 3 and Port 1 enters the transmitted close state. Port 3 can continue to transfer frames as long as it has credits to do so.

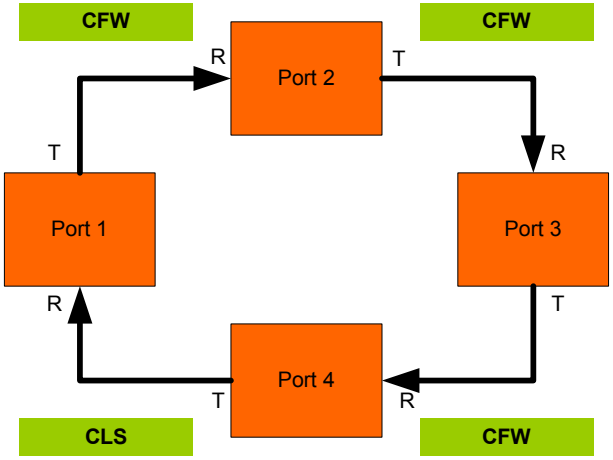


Figure 546 **Close loop: Step 3**

Step 3. When Port 3 receives the Close it enters the received close state, finishes its frame transmission and sends a CLS of its own (Figure 546). Port 3 also enters the monitoring state. When Port 1 receives the close it can enter the monitoring state or open a loop circuit with another port.

Alternate Buffer-to-Buffer Credit (BB_Credit)

In arbitrated loop the flow of frames between two ports is handled using a different credit model than the ones mentioned previously. Buffer-to-buffer credit and end-to-end credit flow control are both credit models that are negotiated between ports during the login process. For more information on buffer-to-buffer credit, refer to "Buffer-to-buffer credit (BB_Credit)" on page 928. For more information on end-to-end credit, refer to "End-to-End Credit" on page 902.

Alternate BB_Credit is used in arbitrated loop environments instead of login BB_Credit as each port on a loop may open up to 126 other ports on the loop and thus would have to keep track of the credit it has with each of the other ports on the loop from N_Port login time.

This would be a significant complication. The Alternate BB_Credit model is an enhancement of the login BB_Credit model in that it begins with an initial login credit value which can be zero and can then be increased dynamically once a loop circuit between two ports is opened.

Before the ports log in, the available credit is assumed to be zero and even the login BB_Credit granted to another port may be zero. This login BB_Credit is used to set the available BB_Credit value when the loop circuit is opened, although ports may or may not use this login BB_ credit value.

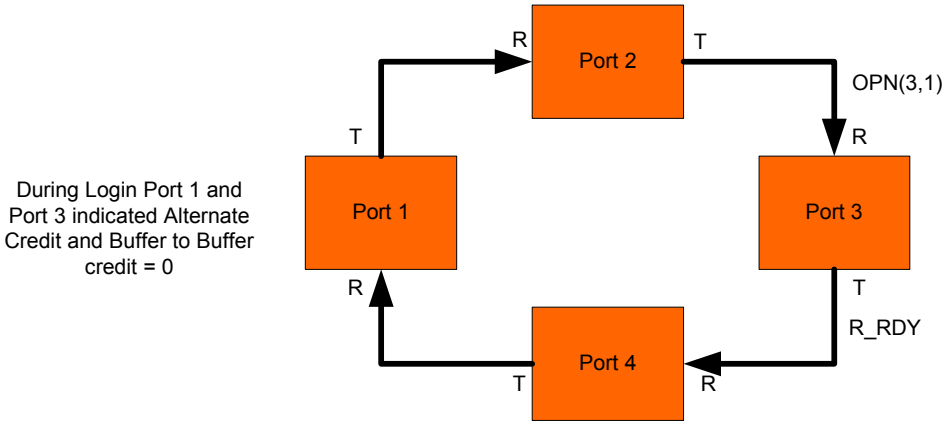


Figure 547 **Alternate credit**

BB_Credit is then signalled dynamically by a port by sending one R_RDY for each available receive buffer. This allows ports that were opened with an available BB_Credit value of zero to begin frame transmission as soon as they receive an R_RDY from the receiving port. A loop port must indicate whether it supports login BB_Credit or Alternate Credit model. This is done during the N_Port login where the BB_Credit management bit (BBM bit) is either **on** for support of Alternate credit or **off** for support of the login BB_Credit model.

Figure 547 shows the situation where Port 1 and Port 3 during N_Port login indicated Alternate credit support and BB_Credit of zero. Thus when port 1 opens a loop circuit with port 3 by sending the OPN(3, 1) then port 1 has to wait until the R_RDY is received back from port 3 before a receive it can send a frame. The number of R_RDYs received

by port 1 indicates the number of receive buffers available on port 3. Thus, when the loop is opened with an initial available BB_Credit value of zero there is a time delay in waiting for the first R_RDY to come back around where no data transfer can take place.

When the loop is closed the available BB_Credits are set to zero and any outstanding R_RDYs are not sent. The available BB_Credit value is then re-established when a loop circuit is opened between two ports.

QuickLoop

With data centers supporting HP-UX FC-AL and a variety of FC-SW types, EMC-supported solutions include the following:

- Single DS-16B operating in QuickLoop mode
- Two DS-16Bs participating in a QuickLoop *Partnership* (described under "Concepts, terms, and rules" on page 1392)
- Single DS-16B operating in mixed mode with QuickLoop and FC-SW nodes

Note: RPQs may be accepted for mixed-mode switches participating in a multi-switch fabric.

Need for QuickLoop

QuickLoop offers customers with HP-UX FC-AL hosts a migration path to full fabric environments. Using a Fibre Channel switch, FC-AL nodes can communicate with other private nodes over the fabric infrastructure.

EMC also supports QuickLoop with OpenVME hosts from ICL.

Defining QuickLoop

In a normal FC-AL hub environment, all traffic (for example: arbitrations, data requests, reads, and writes) is passed around the loop from node to node to node. The limitation is that only one node may send out traffic at a time. This solution does not scale well; all traffic on the loop is constrained to 100 MB/s total bandwidth.

QuickLoop allows hosts and storage to connect to a DS-16B using FC-AL drivers and HBAs.

Each port enabled for QuickLoop mode on the DS-16B is considered a *looplet*. If a host is on Port 0, one node is considered to be in that looplet. If an eight-port hub is attached to Port 0, there are seven nodes on that looplet. Data transfers between looplets are supported at 100 MB/s. With QuickLoop, there may be multiple data transfers capable of 100 MB/s each.

Concepts, terms, and rules

Note the following:

- A *Private Loop* device does not attempt to log in with a fabric and communicates only with other devices on the same loop.
- A *Public Loop* device (for example: HP-UX L-Class with A5158A HBA) logs in with a fabric and may communicate with both public and switched nodes.

  Note: This is not a QuickLoop function, nor is it supported by EMC.

- QuickLoop nodes do not interact with fabric nodes.

  FC-AL public nodes added to a QuickLoop effectively become private nodes. As a result, they cannot initiate communication with fabric nodes.
- EMC supports one QuickLoop per switch.
- QuickLoop may span a maximum of two switches. (This is known as a QuickLoop Partnership.) However, both switches may be part of only one QuickLoop.

  A QuickLoop Partnership is supported in a standalone configuration. RPQs may be accepted for multiswitch fabrics with two of the switches participating in a QuickLoop Partnership.
- EMC supports the zoning of QuickLoop-enabled ports zoned by port number.
  - Since QuickLoop ports are hardware enforced, port zoning must be in effect.
  - You may have up to 16 of these zones in a single QuickLoop.
  - Port-zoning of QuickLoop enables ports isolates LIPs (loop initialization primitives).
  - For zoning requirements, refer to the following documents, available on Powerlink:
    - *Topology Guide for DS-16B2 in a Symmetrix Environment*
    - *Departmental Switch Models DS-16B and DS-8B Zoning Reference Manual*
    - *DS-16B and DS-8B Topology Guide For Departmental Switches in a Symmetrix Environment*
- EMC supports a maximum QuickLoop of two DS-16Bs and two FC-AL hubs.

- EMC supports a maximum of 16 HBAs in a single QuickLoop.
- No hosts may be attached to a hub.
- There is a maximum of eight Symmetrix Fibre Channel director ports on a hub.

◆ Supported firmware revisions for QuickLoop include 2.1.4a, 2.2.1a, 2.3.0, and 2.5.0d.

◆ Supported firmware revisions for port zoning of QuickLoop-enabled ports are 2.2.1a, 2.3.0, and 2.5.0d.

◆ Supported firmware revisions for single-switch mixed mode are 2.3.0 and 2.5.0d.

◆ When connecting to a QuickLoop enabled port, the Symmetrix Fibre Channel director port must be set to FC-AL.

◆ When set to FC-AL mode, a Symmetrix Fibre Channel director port will support private loop only.

Configuration examples

Single QuickLoop configuration

Example: Single DS-16B operating in QuickLoop; HP K-Class host connected to a Symmetrix 8430, as shown in Figure 548:

Figure 548 **Single QuickLoop configuration**

In an environment with only HP-UX FC-AL hosts, the switch may be 100% dedicated to a QuickLoop. The switch may only be part of single QuickLoop. Zoning of QuickLoop-enabled ports is not required because the entire switch is QuickLoop-enabled. With multi-initiators, however, zoning is recommended to restrict LIPs to that particular zone.

Dual QuickLoop configuration

Example: Two DS-16Bs operating in a QuickLoop Partnership, as shown in Figure 549.

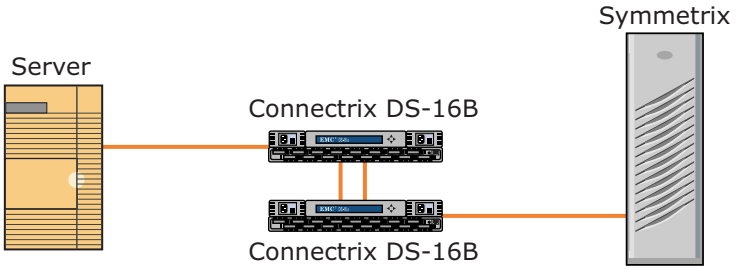


Figure 549 **Dual QuickLoop configuration**

Hosts and Symmetrix Fibre Channel director ports may be located anywhere on either switch. In a QuickLoop Partnership, there are two switches with one QuickLoop. However, there may be multiple-port zones of QuickLoop-enabled ports within a single QuickLoop.

Mixed-mode configuration: Fabric and QuickLoop

A mixed-mode switch environment refers to a DS-xxB switch that has both fabric nodes and QuickLoop ports, as shown in Figure 550.

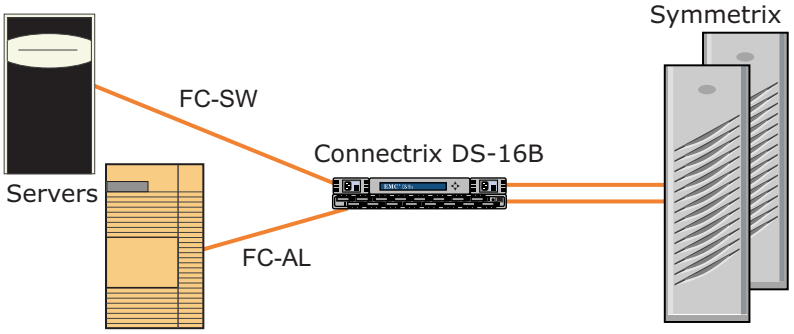


Figure 550 **Fabric/QuickLoop mixed-mode configuration**

Note: EMC supports only mixed mode with a single DS-xxB.

EMC requires that zoning of QuickLoop enabled ports be implemented whenever FC-SW and QuickLoop nodes are running on the same DS-xxB. The zone prevents LIPs from any particular zone from propagating into the fabric or into other zones.

In mixed mode, the switch is set in fabric mode with specific ports in QuickLoop mode.

EMC supports fabric zoning by World Wide Port Name (WWPN). Zoning with QuickLoop is hardware-enforced and therefore requires port zoning. Both WWPN and port zoning may be mixed in the same configuration.

References

For more information on latest qualified drivers and configurations, refer to the following:

- E-Lab Navigator
- *EMC Departmental Switches Model DS-16B2 Topology Guide for DS-16B2 in a Symmetrix Environment*, available on Powerlink.

The fabric OS user's manual contains additional information on DS-xxB Telnet commands.

Bridges

A bridge is a collection of both hardware and software whose primary function is facilitate communication between two dissimilar communication protocols. The bridge is the physical point were the dissimilar networks meet and the translation of protocols takes place. Bridges may also control traffic and security, filtering where necessary to boost SAN performance and contain sensitive data to particular areas of the SAN.

This guide discusses the following bridge-related topics:

- "SCSI-to-Fibre Channel bridges"
- "Crosspoint 4200 SCSI-to-fabric configuration" on page 1399
- "ADIC SAN Gateway SCSI-to-fabric configuration" on page 1402
- "ADIC SAN Gateway loop-to-fabric configuration" on page 1404

SCSI-to-Fibre Channel bridges

SCSI-to-Fibre Channel bridges allow an administrator to place SCSI and Fibre Channel devices into the same SAN environment. These devices communicate by using the bridges to translate commands between the two dissimilar protocols. Commercial bridges are available that support the connection of both target and initiator devices to the bridge's SCSI or Fibre Channel ports.

SCSI-to-Fibre Channel bridges can also be used to extend the distance between SCSI targets and initiator devices. For example, Fast-Wide Differential SCSI devices must be no farther than 25 meters from the initiator. A SCSI-to-Fibre Channel bridge allows extending the distance between the target and initiator to a total of 525 meters.

Operating modes

Storage mode

SCSI-attached tape storage devices, tape library robotics and disk storage have been, and continue to be, a viable solution in the data storage environment. Whether these devices were legacy hardware, new purchases or advanced SCSI technology, they will also have to communicate with the Fibre Channel portion of the SAN environment.

In this mode of operation, such SCSI target devices as tape drives, tape libraries and disk drives are connected to the bridge's SCSI bus connections, and Fibre Channel initiators on the Fibre Channel link.

Figure 551 displays a simple storage mode configuration. This configuration, with the additional Fibre Channel switch support, can allow the incorporation of SCSI devices into the fabric.

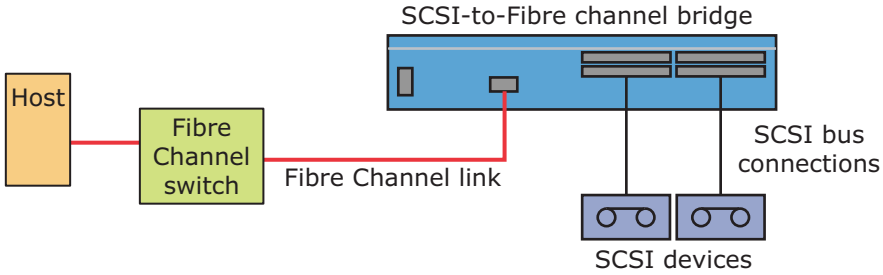


Figure 551 Simple storage mode configuration

Host mode

You may find that some servers and operating systems will not support a Fibre Channel HBA. This topology is not qualified, and will be handled as required on a case-by-case basis.

SCSI-attached hosts, though supported by many bridge vendors, have not yet been validated in the EMC Quality Assurance process. The differences between Host and Storage mode should not be taken lightly and therefore should not be implemented without a thorough review of the customer's configuration and SAN's operational requirements.

Supported SCSI-to-Fibre Channel fabric bridges

EMC evaluates the reliability, functionality of performance of all products prior to their inclusion into our SAN solution sets. Table 101 lists the products and their current feature support, as well as the EMC validation status. E-Lab Navigator lists the EMC-qualified features.

Table 101 Supported SCSI-to-Fibre Channel fabric bridges

Feature	Crossroads CrossPoint 4200	ADIC SAN Gateway
Maximum Fiber Ports	1	6
Maximum SCSI Busses	4	4
Fibre Channel Single-Mode Ports	Yes	No

Table 101 Supported SCSI-to-Fibre Channel fabric bridges (continued)

Feature	Crossroads CrossPoint 4200	ADIC SAN Gateway
Fibre Channel Multimode Ports	Yes	Yes
Fibre Channel Switched Fabric	Yes	Yes
Fibre Channel Arbitrated Loop	Yes	Yes
FibreAlliance MIB Compliance	Yes	Yes
SCSI Host Mode	Yes	Yes
SCSI Storage Mode	Yes	Yes
User-Initiated Diagnostics	Yes	Yes
Serial Communications Port	Yes	Yes
Ethernet Configuration Management	Yes	Yes
Web-Based Configuration Management	Yes	Yes
LAN-Free Backup	Yes	Yes
External LED Status Interface	Yes	Yes
Storage-Side LUN Masking	No	Yes
Storage-Side Fibre Channel/SCSI Channel Masking	No	Yes
Field-Upgradeable Fimware	Yes	Yes

About ADIC Technology, Inc.

ADIC Technology, Inc. manufactures the SAN Gateway, which can operate as a SCSI-to-Fibre Channel fabric bridge.

For more information on the ADIC SAN Gateway, SAN Router and value-added features unique to the ADIC SAN Gateway, visit www.adic.com.

About Crossroads Systems, Inc.

Crossroads Systems Inc. manufactures the Crosspoint 4200, which is validated to operate as SCSI-to FC-SW bridge. Crossroads also manufactures a new+er series of SCSI-to-Fibre Channel bridges using the 4x50 family name.

It should be noted that may vendors may OEM the Crosspoint 4200 under a variety of different names (such as Brocade M-Series EB1200, STK 3200, or Compaq FCTC II). The configuration guidelines below may be applied to these components as well as with the Crossroads 4x50 class of SAN appliances.

For more information on the Crossroads 4200, the 4x50 family of products and other value added features unique to the Crossroads 4200 and 4x50 family, visit www.crossroads.com.

SCSI-to-fabric configuration envelope

The following are guidelines and limitations for using the SCSI-to-Fibre Channel bridge:

- Only FC-SW configurations are supported for SCSI-to-Fibre Channel bridges.
- Only non-high-availability bridge configurations are supported.
- Bridges are currently supported only in a storage mode.
- Single-HBA zoning rules apply for HBA-to-bridge port zoning.
- ANSI-standard 100 [MBytes/sec] Fibre Channel.
- Fibre Channel Distance up to 500 meters between Bridge and Switch port.
- Index (or Auto) Addressing scheme for SCSI-to-Fibre Channel device mapping.
- E-Lab Navigator lists the supported tape libraries and tape drives.
- For performance, a maximum of two high-bandwidth SCSI devices per bridge SCSI bus is recommended.

Crosspoint 4200 SCSI-to-fabric configuration

It should be noted that may vendors may OEM the Crosspoint 4200 under a variety of different names (such as Brocade M-Series EB1200, STK 3200, or Compaq FCTC II). The configuration guidelines below may be applied to these components as well as with the Crossroads 4x50 class of SAN appliances.

To configure a SCSI tape drive or tape library device into the SAN environment using the Crossroads Crosspoint 4200 as a SCSI-to-Fabric Bridge, follow the steps outlined in the following paragraphs.

Initial setup

Follow the Crossroads procedures for the initial setup and power-on of the bridge. If you intend to configure the Crosspoint 4200 over the Web interface, make sure that you have properly configured the unit's IP address, netmask and TCP/IP gateway address. This must be initially done through the serial interface, using a PC and a terminal server program. You may also use the PC and Telnet into the bridge to complete the configuration.

1. Set SCSI port characteristics — Prior to the connection of a SCSI tape devices (targets), use the Crossroads procedures to set the mode of operation. Using Fibre Channel hosts to communicate with SCSI devices is referred to as Initiator Mode. Initiator Mode is the default operation mode for the Crosspoint 4200.
2. Set fibre port characteristics — The Crosspoint 4200's fiber connection can automatically detect if it is connected to a fabric and configure the port for fabric operation without intervention. If this is not the initial configuration for the bridge, Telnet into the bridge to verify the status of the configuration before proceeding.
3. Employ value-added features — At this point you may also decide to start to employ some of the value added features supplied by Crossroads. One of interest would be Fibre Channel-to-SCSI mapping. This allows you to manually configure how the physical target device LUNs will be seen on the HBAs connected to the bridge.

   Refer to your current SAN/ESN management suite prior to enabling and configuring any additional management options available on the Crosspoint 4200. Any redundancy in this area may result in unexplained outages and difficulty in troubleshooting SAN issues in the future.

Target connections

It is recommended when installing new SCSI devices into the Crosspoint 4200 that you start with both the bridge and the SCSI target devices (Tapes or Disks) powered off. The Crosspoint 4200 will assign new LUN numbers and present them to the Hosts based on how they are discovered. With this in mind, connect the SCSI devices into the Crosspoint 4200's SCSI busses as you would like them to be numbered.

After the SCSI connections have been made, you may power on SCSI target devices; then power on the Crosspoint 4200.

At this point you may also decide to start to employ some of the value-added features supplied by Crossroads. One of interest would be Fibre Channel to SCSI Mapping. This allows you to manually configure how the physical target device LUNs will be seen on the HBAs connected to the bridge.

Refer to your current SAN/ESN management suite prior to enabling and configuring any additional management options available on the Crosspoint 4200. Any redundancy in this area may result in unexplained outages and difficulty in troubleshooting SAN issues in the future.

If SCSI devices are moved, added, or removed, you may need to reboot the bridges for the changes to take effect.

Remember to record the Crossroads, assigned LUN numbers, because they will be needed later when you perform any specific HBA-to-LUN discovery, mapping and persistent binding.

Bridge-to-SAN connections

The fiber port on the Crosspoint 4200 is provided a unique World Wide Port Name (WWPN) by the Crosspoint 4200. You should verify the current WWPN currently configured on the bridge and record this future use during the fabric zoning stage.

Host connections and configuration

As in any other SAN configuration, you will connect the host HBAs directly into the fabric backbone. Once both the host and Crossroads, Crosspoint 4200 are visible to the fabric you may start the procedures for zoning and the distribution of devices across the SAN.

At this point you may also want to perform any host specific procedures necessary to discover or map the new devices LUNs that will be discovered at the completion of the configuration.

Note: Refer to the appropriate host-specific configuration information in the EMC Host Connectivity Guides.

Fabric zoning

For the Crossroads, Crosspoint 4200 to communicate with the host's HBA, follow the current supported practices for HBA and bridge WWPN zoning on your fabric switch.

Final setup

Once you have completed the fabric zoning, you should return to the Crosspoint 4200 to ensure that the server's HBA can communicate with the Crosspoint 4200. After communication has been validated you should reboot your host machines so that it can discover the list of new SCSI devices.

ADIC SAN Gateway SCSI-to-fabric configuration

To configure a SCSI tape drive or tape library device into the SAN environment using the ADIC SAN Gateway as a SCSI-to-fabric bridge, follow the steps outlined in this section.

Initial setup

Follow ADIC's documented procedures for the initial setup and power-on of the Gateway. If you intend to configure the SAN Gateway through the Web interface, make sure that you have properly configured the unit's IP address, Netmask and TCP/IP Gateway address. You must also install the client/server portion of the Web software somewhere on your IP network.

Initial configuration of the IP interface must be done through the serial interface, using a PC and a terminal server program. You may also use a PC and telnet into the SAN Gateway to complete the configuration.

1. Set SCSI port characteristics — Prior to the connection of a SCSI tape devices (targets), use the ADIC procedures to set the Channel Mode of the SCSI port on the SAN Gateway in use to **Initiator**. The correct port designation can easily be remembered by always configuring the port to act as the opposite of the devices that you are installing. Since you are adding a target you need to configure the SAN Gateway port to act as an initiator.

2. Set fibre port characteristics — Before you connect the SAN Gateway into the fabric you will need to set the characteristics of the fiber port. Use the ADIC procedures for setting the Port Mode of the fiber connection to **Target**. This can be remembered in the same fashion as discussed above by always configuring the port to act as the opposite of the devices that you are installing.

   Since you are going to be communicating with a host initiator device the SAN Gateway port must be seen as a target device. Lastly, since you are connecting to a fabric switch, you will want to set the Connection Options to **Point-to-Point** mode.

3. Employ value-added features — At this point you may also decide to start to employ some of the value-added features supplied by ADIC. EMC recommends that you review and possibly employ Channel Zoning and Virtual Private SAN.

The Channel Zoning option is currently a free offering that allows you to designate which SCSI channels can communicate with which Fibre Channels. You may choose to allow total communication between the different channels or select any combination of individual target-to-initiator paths.

The Virtual Private SAN (VPS) option requires a license and is used to designate which device LUNS can be seen on which individual HBAs. Once enabled you will have to either install a supplied software package on your host machines to register the HBA information with the SAN Gateway or manually configure the information into the Gateway. Refer to the ADIC User Manual for additional features and procedures associated with the Virtual Private SAN option.

It should be noted that these features may already be supplied in your current SAN/ESN management tools. Refer to your current SAN/ESN management suite prior to enabling and configuring any additional management options. Any redundancy in this area may result in unexplained outages and difficulty in troubleshooting SAN issues in the future.

Target connections

It is recommended when installing new SCSI devices into the SAN Gateway that you start with both the Gateway and the SCSI target devices powered off. The SAN Gateway will assign new LUN numbers and present them to the hosts based on how they are discovered. With this in mind, connect the SCSI devices into the SAN Gateway's SCSI busses as you would like them to be numbered.

After the SCSI connections have been made, power on the SCSI target devices. You may then power on the SAN Gateway.

If a SCSI device is moved, added, or removed, you can run the SCSI **rescan** command on the Gateway to get the correct device picture.

Remember to record the ADIC-assigned LUN numbers, because they will be needed later when you perform any specific HBA-to-LUN discovery, mapping and persistent binding.

Bridge-to-SAN connections

Each fiber port on the SAN Gateway is provided a unique World Wide Port Name (WWPN) by the SAN Gateway. The Gateway does this by:

1. Changing the first pair of digits in the WWN from **10** to **20**.
2. Substituting the next two digits in the WWN into a two-digit representation of the fiber port number being used.

For example, if you had connected fiber port 2 on the SAN Gateway (WWN 10:00:00:60:45:16:0A:0D) into the fabric switch, the WWPN that would appear on the fabric switch would be 20:02:00:60:45:16:0A:0D.

Host connections and configuration

As in any other SAN configuration you will connect the host HBAs directly into the fabric backbone. Once both the host and ADIC SAN Gateway are visible to the fabric you can start the procedures for zoning and the distribution of devices across the SAN.

At this point you may also want to perform any host-specific procedures necessary to discover or map the new devices LUNs that will be discovered at the completion of the configuration.

Note: Refer to the appropriate host-specific configuration information in the EMC Host Connectivity Guides.

Fabric zoning

For the ADIC SAN Gateway to communicate with the host's HBA, follow the current supported practices for HBA and Gateway WWPN zoning on your fabric switch.

Final setup

Once you have completed the fabric zoning, you should return to the SAN Gateway to ensure that the server's HBA can communicate with the SAN Gateway. After communication has been validated you should reboot your host machines so that it can discover the list of new SCSI devices.

ADIC SAN Gateway loop-to-fabric configuration

To configure an FC-AL tape drive into the SAN environment using the ADIC SAN Gateways as a loop-to-fabric bridge, follow the steps outlined below:

Initial setup

Follow ADIC's documented procedures for the initial setup and power-on of the Gateway. If you intend to configure the SAN Gateway through the Web interface, make sure that you have properly configured the unit's IP address, Netmask, and TCP/IP Gateway address. You must also install the client/server portion of the Web software somewhere on your IP network.

Initial configuration of the IP interface must be done through the serial interface, using a PC and a terminal server program. You may also use a PC and Telnet into the SAN Gateway to complete the configuration.

1. Set fibre port characteristics as a target — Prior to the connection of an FC-AL tape device (Target), use the ADIC procedures to set the **Fibre Channel Mode** of the Fibre Channel port on the SAN Gateway in use to **Public Initiator** and the **Connection Options** to **Loop only**. The correct port designation can easily be remembered by always configuring the port to act as the opposite of the devices you are installing. Since you are adding a target you need to configure the SAN Gateway port to act as an initiator.

2. Set fibre port characteristics as link connection — Before you connect the SAN Gateway into the fabric you will need to set the characteristics of the fiber port. Use the ADIC procedures for setting the **Port Mode** of the fiber connection to **Target**. This can be remembered in the same fashion as discussed above by always configuring the port to act as the opposite of the devices that you are installing.

   Since you are going to be communicating with a host initiator device the SAN Gateway port must be seen as a target device. Lastly, since you are connecting to a fabric switch, you will want to set the **Connection Options** to **Point to Point** mode.

3. Employ value-added features — At this point you may also decide to start to employ some of the value-added features supplied by ADIC. EMC recommends that you review and possibly employ Channel Zoning and Virtual Private SAN.

   The Channel Zoning option is currently a free offering that allows you to designate which SCSI channels and Fibre Channels can communicate with which Fibre Channels. You may choose to allow total communication between the different channels or select any combination of individual target-to-initiator paths.

   The Virtual Private SAN (VPS) option requires a license and is used to designate which device LUNS can be seen on which individual HBAs. Once enabled you will have to either install a supplied software package on your host machines to register the HBA information with the SAN Gateway or manually configure the information into the Gateway. Refer to the ADIC User Manual for additional features and procedures associated with the Virtual Private SAN option.

   Note that these features may already be supplied in your current SAN/ESN management tools. Refer to your SAN/ESN management suite prior to enabling and configuring any

additional management options. Any redundancy in this area may result in unexplained outages and difficulty in troubleshooting SAN issues in the future.

Target connections

It is recommended when installing a new FC-AL devices into the SAN Gateway that you start with both the Gateway and the FC-AL target devices powered off. The SAN Gateway will assign new LUN numbers and present them to the hosts based on how they are discovered. With this in mind, connect the FC-AL device into the SAN Gateway's Fibre Channel port as you would like it to be numbered.

After the loop connections have been made, power on the FC-AL target devices. You may then power on the SAN Gateway.

Remember to record the ADIC-assigned LUN numbers, because they will be needed later when you perform any specific HBA-to-LUN discovery, mapping and persistent binding.

Bridge-to-SAN connections

Each fiber port on the SAN Gateway is provided a unique World Wide Port Name (WWPN) by the SAN Gateway. The Gateway does this by:

1. Changing the first pair of digits in the WWN from **10** to **20**.
2. Substituting the next two digits in the WWN into a two-digit representation of the fiber port number being used.

For example, if you had connected fiber port 2 on the SAN Gateway (WWN 10:00:00:60:45:16:0A:0D) into the fabric switch, the WWPN that would appear on the fabric switch would be 20:02:00:60:45:16:0A:0D.

Host connections and configuration

As in any other SAN configuration you will connect the host HBAs directly into the fabric backbone. Once both the host and ADIC SAN Gateway are visible to the fabric you can start the procedures for zoning and the distribution of devices across the SAN.

At this point you may also want to perform any host specific procedures necessary to discover or map the new devices LUNs that will be discovered at the completion of the configuration.

Note: Refer to the appropriate host-specific configuration information in the EMC Host Connectivity Guides.

Fabric zoning

For the ADIC SAN Gateway to communicate with the host's HBA, follow the current supported practices for HBA and Gateway WWPN zoning on your fabric switch.

Final setup

Once you have completed the fabric zoning, you should return to the SAN Gateway to ensure that the server's HBA can communicate with the SAN Gateway. After communication has been validated you should reboot your host machines so that it can discover the list of new FC-AL devices.

Reference

"Interfacing arbitrated loop to switched fabric" on page 1408 contains more information on connection FC-AL tapes into a switched fabric.

Interfacing arbitrated loop to switched fabric

Loop-to-fabric Fibre Channel bridges allow an administrator to place FC-AL and FC-SW devices into the same SAN environment. These devices communicate by using the bridges to translate commands between the dissimilar protocols. Fibre Channel switches and bridges are available that support FC-AL to FC-SW protocol translation. E-Lab Navigator lists the supported connectivity devices that allow the connection of FC-AL initiators.

Operating modes

Storage mode

In *storage mode* operation, such target devices as tape drives and disk drives are connected to the bridge's Fibre Channel ports. Other Fibre Channel ports on the bridge are used to connect the bridge into the fabric. The number of fabric connections from bridge to fabric depends on the actual bandwidth needed to consistently stream data to the target devices on the bridge.

Host mode

In *host mode* operation, initiator devices are attached to either a SCSI port or Fibre Channel port on the bridging device. The bridge is then connected through another Fibre Channel port to the fabric.

You may find that some servers and operating systems will not support Fibre Channel in the chassis or FC-SW settings on an HBA. This topology is not qualified, and will be handled as required on a case-by-case basis.

Connectivity devices that support FC-AL

The EMC Fibre Channel implementation supports these Fibre Channel connectivity devices for tape:

- EMC Products:
  - Connectrix DS-16B (Brocade Silkworm 2xxx)
  - Connectrix DS-8B (Brocade Silkworm 2xxx)
  - Connectrix DS-16B2 (Brocade Silkworm 3800)
  - Connectrix DS-24M2 (Brocade M-Series 4500)

- Brocade Products:
  - Silkworm 2xxx series (EMC Connectrix DS-16B / 8B)
  - Silkworm 3800 / 3200 (EMC Connectrix DS-16B2 / DS-8B2)
- Brocade M-Series Products
  - Brocade M-Series 4500 (EMC Connectrix DS-24M2)
- ADIC Products: SAN Gateway

Connectrix DS-16B, DS-16B2 (Brocade SilkWorm series)

These switches all provide the same functionality. Each port on these switches is capable of auto-negotiating and auto-initializing the port for public loop or fabric operation. Auto-configuration occurs immediately after the devices are connected to the switch port. At the completion of the auto-configuration phase (almost instantaneously), the port should appear as an FL_Port in the name server list on the switch. No additional software or licensing is required for this functionality.

Note: Do not configure these ports as QuickLoop / Fabric Assist ports. The QuickLoop / Fabric Assist mechanisms are not required for public loop device support.

When the negotiation and initialization is complete, you will also be able to view the device's WWPN in the switch's name server list.

Port count

Each DS-16B and DS-16B2 (Brocade SilkWorm) has 16, and each DS-8B has 8, Fibre Channel ports that individually and automatically negotiate whether to communicate using Fibre Channel arbitrated loop or switched fabric protocols.

Features

The Connectrix DS-16B / 8B and DS-16B2 are Departmental Switches that can provide all the Fibre Channel protocol features associated with a fabric switch. ISLs may be connected to either another Brocade switch (in native Brocade Fabric mode) or to a Brocade M-Series switch (using Open Fabric mode). Using Connectrix DS-16B / B2 in Open Fabric mode allows you to easily connect FC-AL devices into a Brocade M-Series fabric.

The ability of the Connectrix DS-16B / 8B / 16B2 to automatically negotiate the correct protocol on an individual port basis allows you to mix E_Ports, N_Ports, and NL_Ports on the same switch.

Fabric

Since the Connectrix DS-16B/8B/16B2 is a switch, it requires a unique Domain ID before connecting into an existing fabric. EMC provides recommendations for fabric topology configurations and sizes that must be followed when using a switch to incorporate FC-AL tapes into a SAN environment.

The Connectrix DS-16B/8B/16B2 can be used in either a homogeneous Brocade fabric or a heterogeneous fabric that includes Brocade M-Series switches. When incorporating a Brocade switch into a fabric that includes Brocade M-Series switches, the fabric mode on all switches must be set to Open Fabric.

Maintenance

Each tape drive is individually attached to the DS-16B/8B/16B2, and forms a single-device loop with the switch port. This allows the user to work individually on any tape device without experiencing an interruption associated with a Loop Initialization Protocol (LIP) exchange across the other devices.

Note: The tape library robot may attached over either Fibre Channel or SCSI. Be sure to understand your specific connectivity requirements for your library. ("SCSI-to-Fibre Channel bridges" on page 1396 provides more information.)

Zoning

Public loop and switched tapes attached to a switch register their unique WWNs with the switch's name server. This allows the user to directly zone the tape's WWN with the WWN of the initiator device (HBA).

"Sharing tape and disk on the same HBA" on page 593 provides more information.

SAN configuration recommendations

EMC recommends that you first identify your actual data throughput requirements prior to locking down any configuration. As a best practice, start by taking the theoretical throughput of the of the switch's Fibre Channel connectivity capabilities (for example: 1 Gb/s, 2 Gb/s, and so on) and divide this by the tape drive's theoretical throughput capability (not its Fibre Channel connector specifications) to identify how many tape drives can be streaming simultaneously on a single ISL.

Example:

- Assume a two switch fabric (A and B).
- Each tape drive in the library can support a throughput of 25 MB/s.
- Both switches in the fabric have 100 MB (half-duplex) switch ports.
- MB per ISL / MB per tape drive = tapes per ISL (rounded down).
- 100 MB / 25 MB = 4 tape drives.

For particular environments that do not require the same simultaneous tape access, you may be able to reduce the required ISLs and increase the number of tape drives per switch. You may also be able to use the backup application's drive pooling facility to distribute your simultaneous drive load across multiple ISLs.

Figure 552 shows two tape pools, each on its own core switch. Placing the pools on separate switches guarantees that different ISLs are used when each pool is accessed. You can also increase the overall availability of tape resources by spreading them across the fabric.

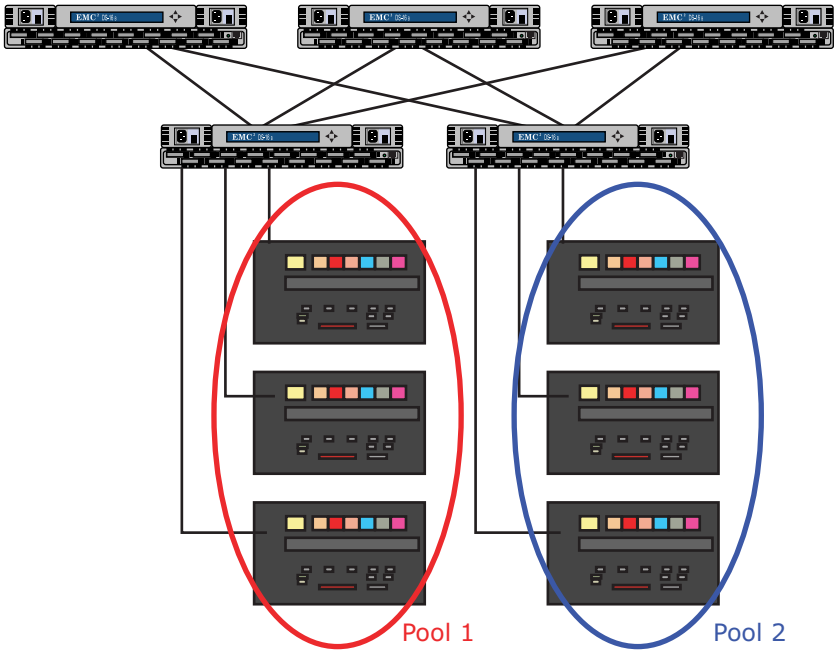


Figure 552 Examples of tape pools

Brocade M-Series ES-1000

Port count

Each Brocade M-Series ES-1000 has eight loop-only Fibre Channel ports acting as an 8-port hub, as well as a single Fibre Channel B_Port (Bridge Port) used for attaching the ES-1000 to other Brocade M-Series Fibre Channel switches.

Features

- The ES-1000 product is fully manageable by EMC Connectrix Manager.
- Each device attached to the ES-1000 registers its WWN with the switch's name server, allowing the individual devices to be zoned independently.
- All devices attached to the 8-port hub will share the bandwidth of the 1 Gb B_Port ISL to the remainder of the fabric.
- Not all industry switches support the B_Port protocol; for this reason the ES-1000 may be connected only to ED-64M, DS-16M, DS-32M, ED-1032.
- Current versions of the ES-1000 firmware offer support for Open Fabric mode operation in a heterogeneous environment. It should be noted that the ES-1000 must still be directly connected to the Brocade M-Series family of switches.

Fabric

Since the ES-1000 is a workgroup switch and supports B_Ports (specialized, limited functionality E_Ports), it requires a unique Domain ID before connecting into an existing fabric. EMC provides recommendations for fabric topology configurations and sizes that must be followed when using a switch to incorporate tapes into a SAN environment.

It should be noted that since the ES-1000 provides only one B_Port, the loss of this connection would constitute a single point of failure for the fabric.

Users may use the ES-1000 in either a homogeneous Brocade M-Series fabric or a heterogeneous fabric, as long as the ES-1000 is attached to a Brocade M-Series switch. Users should also familiarize themselves with the requirements and operation of an heterogeneous fabric before making the decision to include a Brocade switch in a Brocade M-Series switch fabric.

Maintenance

The ES-1000 forms a Fibre Channel loop consisting of the entire set of devices that are connected to the FC-AL ports. Since all devices are in the same loop, maintenance on any one device may cause a LIP (loop initialization protocol) message to be set to all devices on the loop. LIPs may cause interruptions in backup jobs.

Zoning

Since devices attached to the switch register their own unique World Wide Names (WWN) with the switch's name server, the user will be able to directly zone the individual tape drives or other target devices with the WWN of the initiator device (HBA).

SAN configuration recommendations

EMC recommends that you first identify your actual data throughput requirements prior to locking down any configuration. As a best practice, start by taking the theoretical throughput of the of the switch's Fibre Channel connectivity capabilities (for example, 1G) and divide this by the tape drive's theoretical throughput capability (not its Fibre Channel connector specifications) to identify how many tape drives can be streaming simultaneously on a single ISL.

Since the Brocade M-Series ES-1000 has only a single ISL, the simultaneous access requirements to the tape drives may need to be relaxed when fully populating the switch with tape drives. Users may also decide to distribute their tape devices across multiple ES-1000 devices to limit the impact of an ISL or ES-1000 link failure.

ADIC SAN Gateway

Port count

Each ADIC SAN Gateway has up to six Fibre Channel ports. Each Fibre Channel port can be manually configured for either FC-AL or FC-SW, or can be configured to detect and configure itself based on the current attachment.

The SAN Gateway may also be configured with up to four additional SCSI adapters.

Features

The ADIC SAN Gateway comes with its own client/server management application that may be used to configure, assign host-to-device access, or assign devices to specific channels. A single-client/server setup may be used to manage multiple gateways simultaneously.

Fabric

Since the ADIC SAN Gateway is not a switch, it may be added to any size fabric without the need for a unique fabric Domain ID. Depending on the individual Fibre Channel port setting, the ADIC SAN Gateway connects to the switch and logs in to the switch using the FC-SW protocols. When the ADIC SAN Gateway logs in to the switch it registers the WWN of the port of the SAN Gateway that is being used. All Fibre Channel ports on the SAN Gateway have the same lower WWN. Only the second pair of the upper WWN is changed to indicate the local Fibre Channel port of the gateway.

EMC provides recommendations for fabric topology configurations and sizes that must be followed when using a SAN Data Gateway to incorporate tapes into a SAN environment.

Maintenance

Each FC-AL tape drive is individually attached to a Fibre Channel port on the SAN Gateway, and forms a single-device loop with the gateway. This allows the user to work individually on any tape device without experiencing an interruption associated with a Loop Initialization Protocol exchange across the other devices.

Zoning

Devices attached to the gateway are automatically assigned a unique LUN by the gateway. The user may use the management application to change this number, but EMC recommends that the configuration list a tape library first with the lowest LUN, followed by its tape drives listed in sequential order. This ordering can be repeated for multiple tape libraries on the same gateway.

A feature on the gateway called *channel zoning* allows the user to associate target devices with specific gateway Fibre Channel ports. Once the gateway's Fibre Channel port(s) are connected to a switch in the fabric, the gateway registers its WWN(s) as stated above.

Zoning is performed using the fabric's management environment (for example, ESN Manager). A zone will consist of the server's HBA WWN and the gateway Fibre Channel port WWN. Devices will appear to the host in a similar fashion to the way disk devices are displayed behind a SCSI controller (that is, each tape device will be referenced under a target and LUN combination). Unique LUNs will be assigned for each tape driver or library robot. The target number may differ depending on how the devices were spread across the Fibre Channel ports on the SAN Gateway.)

SAN configuration recommendations

EMC recommends that you first identify your actual data throughput requirements prior to locking down any configuration. As a best practice, start by taking the theoretical throughput of the of the gateway's Fibre Channel connectivity capabilities (for example, 1G) and divide this by the tape drive's theoretical throughput capability (not its Fibre Channel connector specifications) to identify how many tape drives can be streaming simultaneously on a single Fibre Channel connection from the gateway to the switch.

For particular environments that do not require the same simultaneous tape access, you may be able to reduce the required ISLs and increase the number of tape drives per gateway.

Note: Multiple links between the SAN Gateway and the fabric are possible. To further increase the availability you may also attach links to separate fabrics; however, this lowers the number of tape drives that can be attached.

Configuration guidelines

"ADIC SAN Gateway loop-to-fabric configuration" on page 1404 provides some configuration guidelines.

Summary

E-Lab Navigator lists the EMC-qualified features.

Table 102 Interfacing arbitrated loop to switched fabric: Summary

Feature	Connectrix DS-16B/8B, DS-16B2	Brocade SilkWorm	ADIC SAN Gateway	Brocade M-Series ES-1000
Max Fiber Ports	16 FC-AL/FC-SW (DS-8B has 8)	8 or 16 FC-AL/FC-SW	6 FC-AL/FC-SW	8 FC-AL, 1 FC-SW
Max SCSI Busses	None	None	4	None
Fibre Channel Single-Mode Ports	Yes	Yes	Yes	Yes
Fibre Channel Multimode Ports	Yes	Yes	Yes	Yes
Fibre Channel Switched Fabric	Yes	Yes	Yes	Yes
Fibre Channel Arbitrated Loop	Yes	Yes	Yes	Yes
FibreAlliance MIB Compliance	Yes	Yes	Yes	Yes
SCSI Host Mode	No	No	Yes	No

Table 102 Interfacing arbitrated loop to switched fabric: Summary (continued)

Feature	Connectrix DS-16B/8B, DS-16B2	Brocade SilkWorm	ADIC SAN Gateway	Brocade M-Series ES-1000
SCSI Storage Mode	No	No	Yes	No
User-Initiated Diagnostics	Yes	Yes	Yes	Yes
Serial Communications Port	Yes, DS-8B only	Yes, 8-port switch only	Yes	Yes
Configuration Management	• Web Tools • ESN Manager • Telnet • Fabric Manager	• Web Tools • ESN Manager • Telnet • Fabric Manager	• SAN Director • HyperTerminal	• EFCM • ESNM • Telnet
LAN-Free Backup	No; however, you can back up the switch/zoning configuration	No; however, you can back up the switch/zoning configuration	Yes	No
External LED Status Interface	Yes	Yes	Yes	Yes
Storage-Side LUN Masking	No	No	Yes	No
Storage-Side Fibre Channel/SCSI Channel Masking	N/A	N/A	Yes	N/A
Field-Upgradeable Fimware	Yes (disruptive)	Yes (disruptive)	Yes	Yes

Loop-to-fabric configuration envelope

The following are guidelines and limitations for using the loop-to-fabric bridge:

- Only non-high-availability bridge configurations are supported.
- Single-HBA zoning rules apply for HBA-to-bridge port zoning.
- ANSI-standard 100 (MB/s) Fibre Channel.
- Fibre Channel Distance up to 500 meters between bridge and switch port.

E-Lab Navigator lists the supported tape drives.

Storage area network management

Distance topology

In the early years of ESCON, both multimode for short distances and single mode for long distances were offered. However, it was found that most data centers did not want their long distance links to terminate at their mainframes or storage equipment. For both availability and resource management reasons, they preferred to terminate the long distance links on an ESCON director, and then implement a connection to another ESCON director.

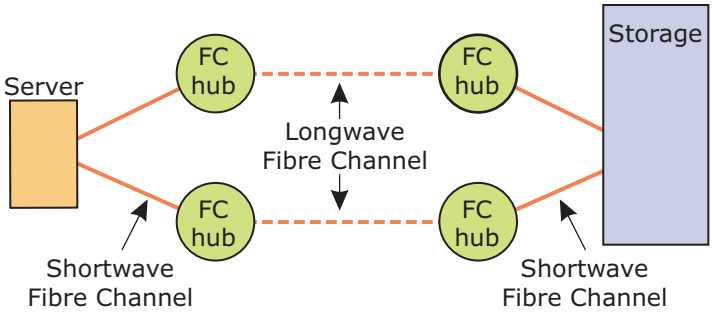


Figure 553 **FC-AL high-availability distance topology example**

By replicating the hubs (as shown in Figure 553), as well as the server and storage connections at each site, a highly fault-resistant topology ensures maximum availability for the site-to-site interconnection.

Capacity topology in the loop environment

By using a hub, the host system includes a fault-resilient connection to a large pool of Symmetrix storage, reducing server I/O slot count.

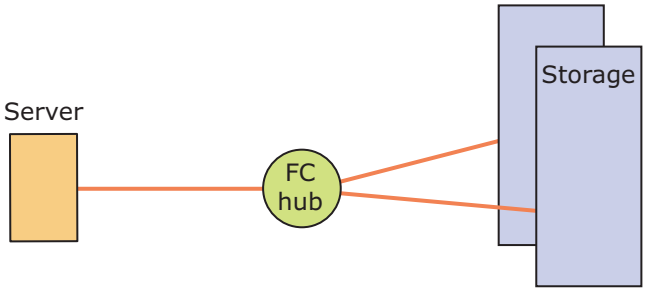


Figure 554 FC-AL capacity expansion topology example

HP-UX PV Links play an essential role in providing high availability in many Fibre topologies. EMC PowerPath® does the same when deployed with Fibre Channel. These host-based software products provide availability during the failure and repair of one of the hubs. In addition, PowerPath provides load balancing services, which facilitates performance management of the system.

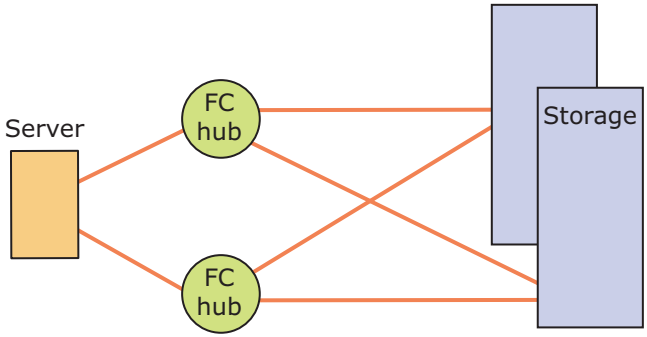


Figure 555 FC-AL high-availability capacity expansion topology example

Consolidation topology in the arbitrated loop environment

Figure 556 and Figure 557 are examples of the consolidation topology in a loop environment.

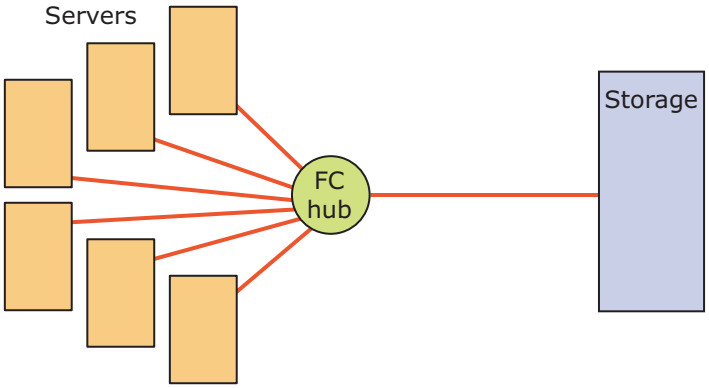


Figure 556 FC-AL consolidation topology example

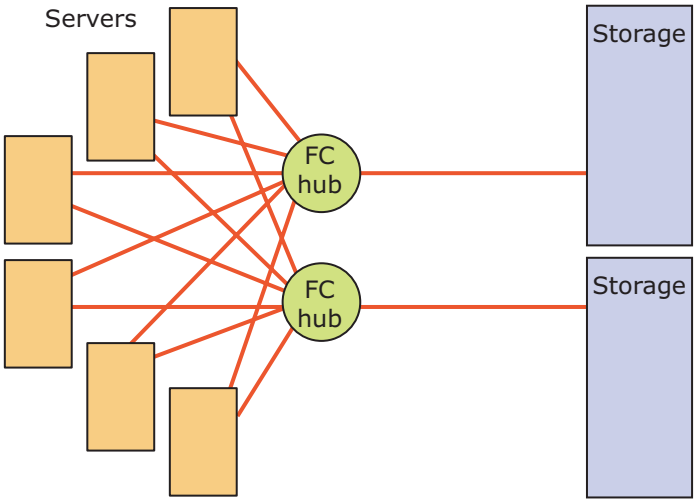


Figure 557 FC-AL high-availability consolidation topology example

Figure 557 shows six hosts, each with two FC-AL ports, which require a total of twelve storage connections. By using a pair of Fibre Channel hubs, all twelve server connections can be consolidated into two storage connections. Servers with low I/O demands (for example, they may be bottlenecked on CPU or networking resources) can share the storage bandwidth of two Symmetrix connections.

Combined topologies

Topologies can be combined for maximum efficiency, achieving large Symmetrix storage capacity for many servers while minimizing the necessary number of host bus adapters and Symmetrix Fibre Channel ports.

The three basic topologies can be combined to take advantage of the benefits of each.

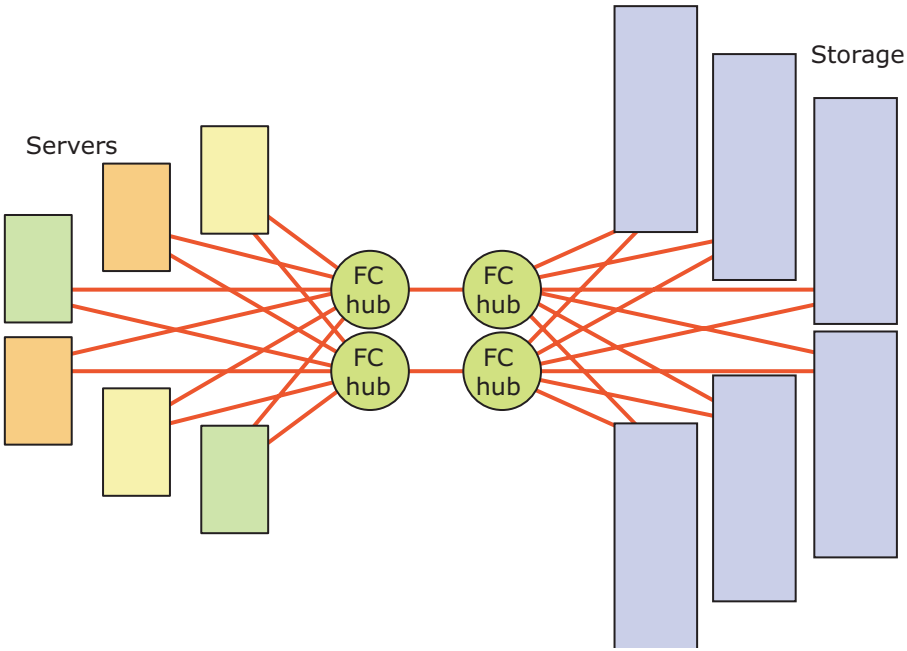


Figure 558 FC-AL combined hub topologies example

CNT (Inrange)

Supported products

FC-9000

Configuring

Topology support

Note the following:

- Up to two domains and a single hop are supported.
- Intermixing FICON and Symmetrix-FCP is not supported.
- Intermixing FICON and SRDF is not supported.
- CUP is supported.
- The switch address range is **x'01'** to **x'EF** (1-239).

Note: You can find more information in the FC-9000 Redbook, at http://www.ibm.com/redbooks.

IOCP considerations

CNT (Inrange FC-9000) Switch ID Definition — No offsets on the switch ID or port address, but every value must be in hex for the mainframe.

Note: Domain ID 32 (the default) would be x'20'.

Glossary

This glossary contains terms related to EMC products and EMC networked storage concepts.

A

access control

A service that allows or prohibits access to a resource. Storage management products implement access control to allow or prohibit specific users. Storage platform products implement access control, often called LUN Masking, to allow or prohibit access to volumes by Initiators (HBAs). *See also* “persistent binding” and “zoning”.

active domain ID

The domain ID actively being used by a switch. It is assigned to a switch by the principal switch.

active zone set

The Active Zone Set is the Zone Set Definition currently in effect and enforced by the Fabric or other entity (for example, the Name Server). Only one zone set at a time can be active.

agent

An autonomous agent is a system situated within (and is part of) an environment that senses that environment, and acts on it over time in pursuit of its own agenda. Storage management software centralizes the control and monitoring of highly distributed storage infrastructure. The centralizing part of the software management system can depend on agents that are installed on the distributed parts of the infrastructure. For example, an agent (software component) can be installed on each of the hosts (servers) in an environment to allow the centralizing software to control and monitor the hosts.

alarm An SNMP message notifying an operator of a network problem.

any-to-any port connectivity A characteristic of a Fibre Channel switch that allows any port on the switch to communicate with any other port on the same switch.

application Application software is a defined subclass of computer software that employs the capabilities of a computer directly to a task that users want to perform. This is in contrast to system software that participates with integration of various capabilities of a computer, and typically does not directly apply these capabilities to performing tasks that benefit users. The term application refers to both the application software and its implementation which often refers to the use of an information processing system. (For example, a payroll application, an airline reservation application, or a network application.) Typically an application is installed “on top of” an operating system like Windows or LINUX, and contains a user interface.

application-specific integrated circuit (ASIC) A circuit designed for a specific purpose, such as implementing lower-layer Fibre Channel protocols (FC-1 and FC-0). ASICs contrast with general-purpose devices such as memory chips or microprocessors, which can be used in many different applications.

ASIC family Different switch hardware platforms that utilize the same port ASIC can be grouped into collections known as an ASIC family. For example, the Fuji ASIC family which consists of the ED-64M and ED-140M run different microprocessors, but both utilize the same port ASIC to provide Fibre Channel connectivity, and are therefore in the same ASIC family. For inter operability concerns, it is useful to understand to which ASIC family a switch belongs.

arbitration The process of selecting one respondent from a collection of several candidates that request service concurrently.

ASCII ASCII (American Standard Code for Information Interchange), generally pronounced [aeski], is a character encoding based on the English alphabet. ASCII codes represent text in computers, communications equipment, and other devices that work with text. Most modern character encodings, which support many more characters, have a historical basis in ASCII.

audit log A log containing summaries of actions taken by a Connectrix Management software user that creates an audit trail of changes. Adding, modifying, or deleting user or product administration

values, creates a record in the audit log that includes the date and time.

authentication Verification of the identity of a process or person.

B

backpressure The effect on the environment leading up to the point of restriction. See "congestion."

BB_Credit *See* "buffer-to-buffer credit".

beaconing Repeated transmission of a beacon light and message until an error is corrected or bypassed. Typically used by a piece of equipment when an individual Field Replaceable Unit (FRU) needs replacement. Beaconing helps the field engineer locate the specific defective component. Some equipment management software systems such as Connectrix Manager offer beaconing capability.

BER *See* "bit error rate".

bidirectional In Fibre Channel, the capability to simultaneously communicate at maximum speeds in both directions over a link.

bit error rate Ratio of received bits that contain errors to total of all bits transmitted.

blade server A consolidation of independent servers and switch technology in the same chassis.

blocked port Devices communicating with a blocked port are prevented from logging in to the Fibre Channel switch containing the port or communicating with other devices attached to the switch. A blocked port continuously transmits the off-line sequence (OLS).

bridge A device that provides a translation service between two network segments utilizing different communication protocols. EMC supports and sells bridges that convert iSCSI storage commands from a NIC-attached server to Fibre Channel commands for a storage platform.

broadcast Sends a transmission to all ports in a network. Typically used in IP networks. Not typically used in Fibre Channel networks.

broadcast frames Data packet, also known as a broadcast packet, whose destination address specifies all computers on a network. *See also* "multicast".

buffer Storage area for data in transit. Buffers compensate for differences in link speeds and link congestion between devices.

buffer-to-buffer credit The number of receive buffers allocated by a receiving FC_Port to a transmitting FC_Port. The value is negotiated between Fibre Channel ports during link initialization. Each time a port transmits a frame it decrements this credit value. Each time a port receives an R_Rdy frame it increments this credit value. If the credit value is decremented to zero, the transmitter stops sending any new frames until the receiver has transmitted an R_Rdy frame. Buffer-to-buffer credit is particularly important in SRDF and Mirror View distance extension solutions.

C

Call Home A product feature that allows the Connectrix service processor to automatically dial out to a support center and report system problems. The support center server accepts calls from the Connectrix service processor, logs reported events, and can notify one or more support center representatives. Telephone numbers and other information are configured through the Windows NT dial-up networking application. The Call Home function can be enabled and disabled through the Connectrix Product Manager.

channel With Open Systems, a channel is a point-to-point link that transports data from one point to another on the communication path, typically with high throughput and low latency that is generally required by storage systems. With Mainframe environments, a channel refers to the server-side of the server-storage communication path, analogous to the HBA in Open Systems.

Class 2 Fibre Channel class of service In Class 2 service, the fabric and destination N_Ports provide connectionless service with notification of delivery or nondelivery between the two N_Ports. Historically Class 2 service is not widely used in Fibre Channel system.

Class 3 Fibre Channel class of service Class 3 service provides a connectionless service without notification of delivery between N_Ports. (This is also known as datagram service.) The transmission and routing of Class 3 frames is the same

as for Class 2 frames. Class 3 is the dominant class of communication used in Fibre Channel for moving data between servers and storage and may be referred to as “Ship and pray.”

Class F Fibre Channel class of service

Class F service is used for all switch-to-switch communication in a multiswitch fabric environment. It is nearly identical to class 2 from a flow control point of view.

community

A relationship between an SNMP agent and a set of SNMP managers that defines authentication, access control, and proxy characteristics.

community name

A name that represents an SNMP community that the agent software recognizes as a valid source for SNMP requests. An SNMP management program that sends an SNMP request to an agent program must identify the request with a community name that the agent recognizes or the agent discards the message as an authentication failure. The agent counts these failures and reports the count to the manager program upon request, or sends an authentication failure trap message to the manager program.

community profile

Information that specifies which management objects are available to what management domain or SNMP community name.

congestion

Occurs at the point of restriction. See “backpressure.”

connectionless

Non dedicated link. Typically used to describe a link between nodes that allows the switch to forward Class 2 or Class 3 frames as resources (ports) allow. *Contrast with* the dedicated bandwidth that is required in a Class 1 Fibre Channel Service point-to-point link.

Connectivity Unit

A hardware component that contains hardware (and possibly software) that provides Fibre Channel connectivity across a fabric. Connectrix switches are example of Connectivity Units. This is a term popularized by the Fibre Alliance MIB, sometimes abbreviated to connunit.

Connectrix management software

The software application that implements the management user interface for all managed Fibre Channel products, typically the Connectrix -M product line. Connectrix Management software is a client/server application with the server running on the Connectrix service processor, and clients running remotely or on the service processor.

Connectrix service processor An optional 1U server shipped with the Connectrix -M product line to run the Connectrix Management server software and EMC remote support application software.

Control Unit In mainframe environments, a Control Unit controls access to storage. It is analogous to a Target in Open Systems environments.

core switch Occupies central locations within the interconnections of a fabric. Generally provides the primary data paths across the fabric and the direct connections to storage devices. Connectrix directors are typically installed as core switches, but may be located anywhere in the fabric.

credit A numeric value that relates to the number of available BB_Credits on a Fibre Channel port. *See*"buffer-to-buffer credit".

D

DASD Direct Access Storage Device.

default Pertaining to an attribute, value, or option that is assumed when none is explicitly specified.

default zone A zone containing all attached devices that are not members of any active zone. Typically the default zone is disabled in a Connectrix-M environment which prevents newly installed servers and storage from communicating until they have been provisioned.

Dense Wavelength Division Multiplexing (DWDM) A process that carries different data channels at different wavelengths over one pair of fiber optic links. A conventional fiber-optic system carries only one channel over a single wavelength traveling through a single fiber.

destination ID A field in a Fibre Channel header that specifies the destination address for a frame. The Fibre Channel header also contains a Source ID (SID). The FCID for a port contains both the SID and the DID.

DID An acronym used to refer to either Domain ID or Destination ID. This ambiguity can create confusion. As a result E-Lab recommends this acronym be used to apply to Domain ID. Destination ID can be abbreviated to FCID.

device A piece of equipment, such as a server, switch or storage system.

dialog box A user interface element of a software product typically implemented as a pop-up window containing informational messages and fields for modification. Facilitates a dialog between the user and the application. Dialog box is often used interchangeably with window.

director An enterprise-class Fibre Channel switch, such as the Connectrix ED-140M, MDS 9509, or ED-48000B. Directors deliver high availability, failure ride-through, and repair under power to insure maximum uptime for business critical applications. Major assemblies, such as power supplies, fan modules, switch controller cards, switching elements, and port modules, are all hot-swappable.

The term director may also refer to a board-level module in the Symmetrix that provides the interface between host channels (through an associated adapter module in the Symmetrix) and Symmetrix disk devices. (This description is presented here only to clarify a term used in other EMC documents.)

DNS *See* "domain name service name".

domain ID A byte-wide field in the three byte Fibre Channel address that uniquely identifies a switch in a fabric. The three fields in a FCID are domain, area, and port. A distinct Domain ID is requested from the principal switch. The principal switch allocates one Domain ID to each switch in the fabric. A user may be able to set a Preferred ID which can be requested of the Principal switch, or set an Insistent Domain ID. If two switches insist on the same DID one or both switches will segment from the fabric.

domain name service name Host or node name for a system that is translated to an IP address through a name server. All DNS names have a host name component and, if fully qualified, a domain component, such as *host1.abcd.com*. In this example, *host1* is the host name.

dual-attached host A host that has two (or more) connections to a set of devices.

E

E_D_TOV A time-out period within which each data frame in a Fibre Channel sequence transmits. This avoids time-out errors at the destination Nx_Port. This function facilitates high speed recovery from dropped frames. Typically this value is 2 seconds.

E_Port
Expansion Port, a port type in a Fibre Channel switch that attaches to another E_Port on a second Fibre Channel switch forming an Interswitch Link (ISL). This link typically conforms to the FC-SW standards developed by the T11 committee, but might not support heterogeneous inter operability.

edge switch
Occupies the periphery of the fabric, generally providing the direct connections to host servers and management workstations. No two edge switches can be connected by interswitch links (ISLs). Connectrix departmental switches are typically installed as edge switches in a multiswitch fabric, but may be located anywhere in the fabric

Embedded Web Server
A management interface embedded on the switch's code that offers features similar to (but not as robust as) the Connectrix Manager and Product Manager.

error detect time out value
Defines the time the switch waits for an expected response before declaring an error condition. The error detect time out value (E_D_TOV) can be set within a range of two-tenths of a second to one second using the Connectrix switch Product Manager.

error message
An indication that an error has been detected. *See also* "information message" *and* "warning message".

Ethernet
A baseband LAN that allows multiple station access to the transmission medium at will without prior coordination and which avoids or resolves contention.

event log
A record of significant events that have occurred on a Connectrix switch, such as FRU failures, degraded operation, and port problems.

expansionport
See "E_Port".

explicit fabric login
In order to join a fabric, an Nport must login to the fabric (an operation referred to as an FLOGI). Typically this is an explicit operation performed by the Nport communicating with the F_port of the switch, and is called an explicit fabric login. Some legacy Fibre Channel ports do not perform explicit login, and switch vendors perform login for ports creating an implicit login. Typically logins are explicit.

F

FA Fibre Adapter, another name for a Symmetrix Fibre Channel director.

F_Port Fabric Port, a port type on a Fibre Channel switch. An F_Port attaches to an N_Port through a point-to-point full-duplex link connection. A G_Port automatically becomes an F_port or an E-Port depending on the port initialization process.

fabric One or more switching devices that interconnect Fibre Channel N_Ports, and route Fibre Channel frames based on destination IDs in the frame headers. A fabric provides discovery, path provisioning, and state change management services for a Fibre Channel environment.

fabric element Any active switch or director in the fabric.

fabric login Process used by N_Ports to establish their operating parameters including class of service, speed, and buffer-to-buffer credit value.

fabric port A port type (F_Port) on a Fibre Channel switch that attaches to an N_Port through a point-to-point full-duplex link connection. An N_Port is typically a host (HBA) or a storage device like Symmetrix or CLARiiON.

fabric shortest path first (FSPF) A routing algorithm implemented by Fibre Channel switches in a fabric. The algorithm seeks to minimize the number of hops traversed as a Fibre Channel frame travels from its source to its destination.

fabric tree A hierarchical list in Connectrix Manager of all fabrics currently known to the Connectrix service processor. The tree includes all members of the fabrics, listed by WWN or nickname.

failover The process of detecting a failure on an active Connectrix switch FRU and the automatic transition of functions to a backup FRU.

fan-in/fan-out Term used to describe the server:storage ratio, where a graphic representation of a 1:n (fan-in) or n:1 (fan-out) logical topology looks like a hand-held fan, with the wide end toward n. By convention fan-out refers to the number of server ports that share a single storage port. Fan-out consolidates a large number of server ports on a fewer number of storage ports. Fan-in refers to the number of storage ports that a single server port uses. Fan-in enlarges the storage capacity used by a server. A fan-in or fan-out rate is often referred to as just the

n part of the ratio; For example, a 16:1 fan-out is also called a fan-out rate of 16, in this case 16 server ports are sharing a single storage port.

FCP *See* "Fibre Channel Protocol".

FC-SW The Fibre Channel fabric standard. The standard is developed by the T11 organization whose documentation can be found at T11.org. EMC actively participates in T11. T11 is a committee within the InterNational Committee for Information Technology (INCITS).

fibre A general term used to cover all physical media types supported by the Fibre Channel specification, such as optical fiber, twisted pair, and coaxial cable.

fiber optics The branch of optical technology concerned with the transmission of radiant power through fibers made of transparent materials such as glass, fused silica, and plastic.

Either a single discrete fiber or a non spatially aligned fiber bundle can be used for each information channel. Such fibers are often called optical fibers to differentiate them from fibers used in non-communication applications.

Fibre Channel The general name of an integrated set of ANSI standards that define new protocols for flexible information transfer. Logically, Fibre Channel is a high-performance serial data channel.

Fibre Channel Protocol A standard Fibre Channel FC-4 level protocol used to run SCSI over Fibre Channel.

Fibre Channel switch modules The embedded switch modules in the back plane of the blade server. See "blade server" on page 1425.

firmware The program code (embedded software) that resides and executes on a connectivity device, such as a Connectrix switch, a Symmetrix Fibre Channel director, or a host bus adapter (HBA).

F_Port Fabric Port, a physical interface within the fabric. An F_Port attaches to an N_Port through a point-to-point full-duplex link connection.

frame A set of fields making up a unit of transmission. Each field is made of bytes. The typical Fibre Channel frame consists of fields: Start-of-frame, header, data-field, CRC, end-of-frame. The maximum frame size is 2148 bytes.

frame header Control information placed before the data-field when encapsulating data for network transmission. The header provides the source and destination IDs of the frame.

FRU Field-replaceable unit, a hardware component that can be replaced as an entire unit. The Connectrix switch Product Manager can display status for the FRUs installed in the unit.

FSPF Fabric Shortest Path First, an algorithm used for routing traffic. This means that, between the source and destination, only the paths that have the least amount of physical hops will be used for frame delivery.

G

gateway address In TCP/IP, a device that connects two systems that use the same or different protocols.

gigabyte (GB) A unit of measure for storage size, loosely one billion (10^9) bytes. One gigabyte actually equals 1,073,741,824 bytes.

G_Port A port type on a Fibre Channel switch capable of acting either as an F_Port or an E_Port, depending on the port type at the other end of the link.

GUI Graphical user interface.

H

HBA *See* "host bus adapter".

hexadecimal Pertaining to a numbering system with base of 16; valid numbers use the digits 0 through 9 and characters A through F (which represent the numbers 10 through 15).

high availability A performance feature characterized by hardware component redundancy and hot-swappability (enabling non-disruptive maintenance). High-availability systems maximize system uptime while providing superior reliability, availability, and serviceability.

hop A hop refers to the number of InterSwitch Links (ISLs) a Fibre Channel frame must traverse to go from its source to its destination.

Good design practice encourages three hops or less to minimize congestion and performance management complexities.

host bus adapter

A bus card in a host system that allows the host system to connect to the storage system. Typically the HBA communicates with the host over a PCI or PCI Express bus and has a single Fibre Channel link to the fabric. The HBA contains an embedded microprocessor with on board firmware, one or more ASICs, and a Small Form Factor Pluggable module (SFP) to connect to the Fibre Channel link.

I

I/O

See "input/output".

in-band management

Transmission of monitoring and control functions over the Fibre Channel interface. You can also perform these functions out-of-band typically by use of the ethernet to manage Fibre Channel devices.

information message

A message telling a user that a function is performing normally or has completed normally. User acknowledgement might or might not be required, depending on the message. *See also* "error message" *and* "warning message".

input/output

(1) Pertaining to a device whose parts can perform an input process and an output process at the same time. (2) Pertaining to a functional unit or channel involved in an input process, output process, or both (concurrently or not), and to the data involved in such a process. (3) Pertaining to input, output, or both.

interface

(1) A shared boundary between two functional units, defined by functional characteristics, signal characteristics, or other characteristics as appropriate. The concept includes the specification of the connection of two devices having different functions. (2) Hardware, software, or both, that links systems, programs, or devices.

Internet Protocol

See "IP".

interoperability

The ability to communicate, execute programs, or transfer data between various functional units over a network. Also refers to a Fibre Channel fabric that contains switches from more than one vendor.

interswitch link (ISL) Interswitch link, a physical E_Port connection between any two switches in a Fibre Channel fabric. An ISL forms a hop in a fabric.

IP Internet Protocol, the TCP/IP standard protocol that defines the datagram as the unit of information passed across an internet and provides the basis for connectionless, best-effort packet delivery service. IP includes the ICMP control and error message protocol as an integral part.

IP address A unique string of numbers that identifies a device on a network. The address consists of four groups (quadrants) of numbers delimited by periods. (This is called *dotted-decimal* notation.) All resources on the network must have an IP address. A valid IP address is in the form *nnn.nnn.nnn.nnn*, where each *nnn* is a decimal in the range 0 to 255.

ISL Interswitch link, a physical E_Port connection between any two switches in a Fibre Channel fabric.

K

kilobyte (K) A unit of measure for storage size, loosely one thousand bytes. One kilobyte actually equals 1,024 bytes.

L

laser A device that produces optical radiation using a population inversion to provide light amplification by stimulated emission of radiation and (generally) an optical resonant cavity to provide positive feedback. Laser radiation can be highly coherent temporally, spatially, or both.

LED Light-emitting diode.

link The physical connection between two devices on a switched fabric.

link incident A problem detected on a fiber-optic link; for example, loss of light, or invalid sequences.

load balancing The ability to distribute traffic over all network ports that are the same distance from the destination address by assigning different paths to different messages. Increases effective network bandwidth. EMC PowerPath software provides load-balancing services for server IO.

logical volume A named unit of storage consisting of a logically contiguous set of disk sectors.

Logical Unit Number (LUN) A number, assigned to a storage volume, that (in combination with the storage device node's World Wide Port Name (WWPN)) represents a unique identifier for a logical volume on a storage area network.

M

MAC address Media Access Control address, the hardware address of a device connected to a shared network.

managed product A hardware product that can be managed using the Connectrix Product Manager. For example, a Connectrix switch is a managed product.

management session Exists when a user logs in to the Connectrix Management software and successfully connects to the product server. The user must specify the network address of the product server at login time.

media The disk surface on which data is stored.

media access control *See* "MAC address".

megabyte (MB) A unit of measure for storage size, loosely one million (10^6) bytes. One megabyte actually equals 1,048,576 bytes.

MIB Management Information Base, a related set of objects (variables) containing information about a managed device and accessed through SNMP from a network management station.

multicast Multicast is used when multiple copies of data are to be sent to designated, multiple, destinations.

multiswitch fabric Fibre Channel fabric created by linking more than one switch or director together to allow communication. *See also* "ISL".

multiswitch linking Port-to-port connections between two switches.

N

name server (dNS) A service known as the distributed Name Server provided by a Fibre Channel fabric that provides device discovery, path provisioning, and

state change notification services to the N_Ports in the fabric. The service is implemented in a distributed fashion, for example, each switch in a fabric participates in providing the service. The service is addressed by the N_Ports through a Well Known Address.

network address A name or address that identifies a managed product, such as a Connectrix switch, or a Connectrix service processor on a TCP/IP network. The network address can be either an IP address in dotted decimal notation, or a Domain Name Service (DNS) name as administered on a customer network. All DNS names have a host name component and (if fully qualified) a domain component, such as *host1.emc.com*. In this example, *host1* is the host name and *EMC.com* is the domain component.

nickname A user-defined name representing a specific WWxN, typically used in a Connectrix -M management environment. The analog in the Connectrix -B and MDS environments is alias.

node The point at which one or more functional units connect to the network.

N_Port Node Port, a Fibre Channel port implemented by an end device (node) that can attach to an F_Port or directly to another N_Port through a point-to-point link connection. HBAs and storage systems implement N_Ports that connect to the fabric.

NVRAM Nonvolatile random access memory.

O

offline sequence (OLS) The OLS Primitive Sequence is transmitted to indicate that the FC_Port transmitting the Sequence is:

a. initiating the Link Initialization Protocol

b. receiving and recognizing NOS

c. or entering the offline state

OLS *See* “offline sequence (OLS)”.

operating system Software that controls the execution of programs and that may provide such services as resource allocation, scheduling, input/output control, and data management. Although operating

systems are predominantly software, partial hardware implementations are possible.

operating mode Regulates what other types of switches can share a multiswitch fabric with the switch under consideration.

optical cable A fiber, multiple fibers, or a fiber bundle in a structure built to meet optical, mechanical, and environmental specifications.

OS *See* "operating system".

out-of-band management Transmission of monitoring/control functions outside of the Fibre Channel interface, typically over ethernet.

oversubscription The ratio of bandwidth required to bandwidth available. When all ports, associated pair-wise, in any random fashion, cannot sustain full duplex at full line-rate, the switch is oversubscribed.

P

parameter A characteristic element with a variable value that is given a constant value for a specified application. Also, a user-specified value for an item in a menu; a value that the system provides when a menu is interpreted; data passed between programs or procedures.

password (1) A value used in authentication or a value used to establish membership in a group having specific privileges. (2) A unique string of characters known to the computer system and to a user who must specify it to gain full or limited access to a system and to the information stored within it.

path In a network, any route between any two nodes.

persistent binding Use of server-level access control configuration information to persistently bind a server device name to a specific Fibre Channel storage volume or logical unit number, through a specific HBA and storage port WWN. The address of a persistently bound device does not shift if a storage target fails to recover during a power cycle. This function is the responsibility of the HBA device driver.

port (1) An access point for data entry or exit. (2) A receptacle on a device to which a cable for another device is attached.

port card Field replaceable hardware component that provides the connection for fiber cables and performs specific device-dependent logic functions.

port name A symbolic name that the user defines for a particular port through the Product Manager.

preferred domain ID An ID configured by the fabric administrator. During the fabric build process a switch requests permission from the principal switch to use its preferred domain ID. The principal switch can deny this request by providing an alternate domain ID only if there is a conflict for the requested Domain ID. Typically a principal switch grants the non-principal switch its requested Preferred Domain ID.

principle downsteam ISL The ISL to which each switch will forward frames originating from the principal switch.

principle ISL The principal ISL is the ISL that frames destined to, or coming from, the principal switch in the fabric will use. An example is an RDI frame.

principal switch In a multiswitch fabric, the switch that allocates domain IDs to itself and to all other switches in the fabric. There is always one principal switch in a fabric. If a switch is not connected to any other switches, it acts as its own principal switch.

principle upsteam ISL The ISL to which each switch will forward frames destined for the principal switch. The principal switch does not have any upstream ISLs.

product (1) Connectivity Product, a generic name for a switch, director, or any other Fibre Channel product. (2) Managed Product, a generic hardware product that can be managed by the Product Manager (a Connectrix switch is a managed product). Note distinction from the definition for "device".

Product Manager A software component of Connectrix Manager software such as a Connectrix switch product manager, that implements the management user interface for a specific product. When a product instance is opened from the Connectrix Manager software products view, the corresponding product manager is invoked. The product manager is also known as an Element Manager.

product name A user configurable identifier assigned to a Managed Product. Typically, this name is stored on the product itself. For a Connectrix switch, the Product Name can also be accessed by an SNMP Manager as the System Name. The Product Name should align with the host name component of a Network Address.

products view The top-level display in the Connectrix Management software user interface that displays icons of Managed Products.

protocol (1) A set of semantic and syntactic rules that determines the behavior of functional units in achieving communication. (2) A specification for the format and relative timing of information exchanged between communicating parties.

R

R_A_TOV *See* "resource allocation time out value".

resource allocation time out value A value used to time-out operations that depend on a maximum time that an exchange can be delayed in a fabric and still be delivered. The resource allocation time-out value of (R_A_TOV) can be set within a range of two-tenths of a second to 120 seconds using the Connectrix switch product manager. The typical value is 10 seconds.

remote access link The ability to communicate with a data processing facility through a remote data link.

remote notification The system can be programmed to notify remote sites of certain classes of events.

remote user workstation A workstation, such as a PC, using Connectrix Management software and Product Manager software that can access the Connectrix service processor over a LAN connection. A user at a remote workstation can perform all of the management and monitoring tasks available to a local user on the Connectrix service processor.

S

SAN *See* "storage area network (SAN)".

segmentation A non-connection between two switches. Numerous reasons exist for an operational ISL to segment, including interop mode incompatibility, zoning conflicts, and domain overlaps.

segmented E_Port E_Port that has ceased to function as an E_Port within a multiswitch fabric due to an incompatibility between the fabrics that it joins.

service processor *See* "Connectrix service processor".

session *See* "management session".

single attached host A host that only has a single connection to a set of devices.

small form factor pluggable (SFP) An optical module implementing a shortwave or long wave optical transceiver.

SMTP Simple Mail Transfer Protocol, a TCP/IP protocol that allows users to create, send, and receive text messages. SMTP protocols specify how messages are passed across a link from one system to another. They do not specify how the mail application accepts, presents or stores the mail.

SNMP Simple Network Management Protocol, a TCP/IP protocol that generally uses the User Datagram Protocol (UDP) to exchange messages between a management information base (MIB) and a management client residing on a network.

storage area network (SAN) A network linking servers or workstations to disk arrays, tape backup systems, and other devices, typically over Fibre Channel and consisting of multiple fabrics.

subnet mask Used by a computer to determine whether another computer with which it needs to communicate is located on a local or remote network. The network mask depends upon the class of networks to which the computer is connecting. The mask indicates which digits to look at in a longer network address and allows the router to avoid handling the entire address. Subnet masking allows routers to move the packets more quickly. Typically, a subnet may represent all the machines at one geographic location, in one building, or on the same local area network.

switch priority Value configured into each switch in a fabric that determines its relative likelihood of becoming the fabric's principal switch.

T

TCP/IP Transmission Control Protocol / Internet Protocol. TCP / IP refers to the protocols that are used on the Internet and most computer networks. TCP refers to the Transport layer that provides flow control and connection services. IP refers to the Internet Protocol level where addressing and routing are implemented.

topology Logical and / or physical arrangement of switches on a network.

trap An asynchronous (unsolicited) notification of an event originating on an SNMP-managed device and directed to a centralized SNMP Network Management Station.

toggle To change the state of a feature / function that has only two states. For example, if a feature / function is *enabled*, toggling changes the state to *disabled*.

U

unblocked port Devices communicating with an unblocked port can log in to a Connectrix switch or a similar product and communicate with devices attached to any other unblocked port if the devices are in the same zone.

Unicast Unicast routing provides one or more optimal path(s) between any of two switches that make up the fabric. (This is used to send a single copy of the data to designated destinations.)

upper layer protocol (ULP) The protocol user of FC-4 including IPI, SCSI, IP, and SBCCS. In a device driver ULP typically refers to the operations that are managed by the class level of the driver, not the port level.

URL Uniform Resource Locater, the addressing system used by the World Wide Web. It describes the location of a file or server anywhere on the Internet.

V

virtual switch A Fibre Channel switch function that allows users to subdivide a physical switch into multiple virtual switches. Each virtual switch

consists of a subset of ports on the physical switch, and has all the properties of a Fibre Channel switch. Multiple virtual switches can be connected through ISL to form a virtual fabric or VSAN.

virtual storage area network (VSAN) An allocation of switch ports that can span multiple physical switches, and forms a virtual fabric. A single physical switch can sometimes host more than one VSAN.

volume A general term referring to an addressable logically contiguous storage space providing block IO services.

VSAN Virtual Storage Area Network.

W

warning message An indication that a possible error has been detected. *See also* "error message" *and* "information message".

World Wide Name (WWN) A unique identifier, even on global networks. The WWN is a 64-bit number (XX:XX:XX:XX:XX:XX:XX:XX). The WWN contains an OUI which uniquely determines the equipment manufacturer. OUIs are administered by the Institute of Electronic and Electrical Engineers (IEEE). The Fibre Channel environment uses two types of WWNs; a World Wide Node Name (WWNN) and a World Wide Port Name (WWPN). Typically the WWPN is used for zoning (path provisioning function).

Z

zone An information object implemented by the distributed Nameserver(dNS) of a Fibre Channel switch. A zone contains a set of members which are permitted to discover and communicate with one another. The members can be identified by a WWPN or port ID. EMC recommends the use of WWPNs in zone management.

zoning Zoning allows an administrator to group several devices by function or by location. All devices connected to a connectivity product, such as a Connectrix switch, may be configured into one or more zones.

zone set An information object implemented by the distributed Nameserver(dNS) of a Fibre Channel switch. A Zone Set contains a set of Zones. A Zone Set is activated against a fabric, and only one Zone Set can be active in a fabric.

zonie A storage administrator who spends a large percentage of his workday zoning a Fibre Channel network and provisioning storage.

Index

Symbols

(IP over Fibre Channel (FC-IP) 1041

A

ADIC SAN Gateway
    loop to fabric 1404
    SCSI to fabric 1402
arbitrated loop 1352

B

B Series examples 1156
backup
    direct-attached 559
    LAN-based 561
    LAN-free 566
    NAS 582
    serverless 571
balanced fabric 713
Blade Server switch modules 1273
bridge
    described 1396
    loop-to-fabric 1408
    SCSI-to-Fibre Channel 1396
Brocade AP7420 474
Brocade applications supported 1036
Brocade ISL Trunking 1035
Brocade Silkworm 2xxx 1065
Brocade Silkworm 3800 1060
business unit fabric 718

C

Celerra 582
channels, described 695
CIENA CN 2000 1266
Ciena ONLINE7000 1185
Cisco MDS 9000 482
Cisco MDS 9216 1091
Cisco Metro 1500 1210
Cisco ONS 15540 1214
class
    Fibre Channel classes supported 1037
Class of service 985, 996
CN 2000 1266
CNT/Inrange FC/9000 1174
connectivity tier fabrics 737
Connectrix
    DS-16B 1065
    DS-16B2 1060
    DS-16B3 1051
    DS-16M/DS-16M2 1145
    DS-24M2 1141
    DS-32B2 1058
    DS-32M/DS-32M2 1143
    DS-4400M/DS-4400M 1134
    DS-4700M/DS-4700M 1127
    DS-8B 1065
    DS-8B3 1054
    EC-1200 1166
    EC-1500 1158
    ED-10000M 1096
    ED-1032 1124
    ED-12000B 1015
    ED-140M 1108
    ED-24000B 1009
    ED-64M 1122
    MDS 9216 1091

Connectrix DS-4900B 1042
Connectrix MDS 9513 1069
Control Unit Port 371
Crosspoint 4200 1399
CUP 371

D

default
    maintenance port password 239
direct-attached backup 559
domain ID
    described 808
    negotiation 809
    recommendations 750
DS-16B 1065
DS-16B2 1060
DS-16B3 1051
DS-16M / DS-16M2 1145
DS-24M2 1141
DS-32B2 1058
DS-32M / DS-32M2 1143
DS-8B 1065
DS-8B3 1054

E

E_Port 804
EC-1200 1166
EC-1500 1158
EC-1700 cabinet component placement 1155
EC-1700-B cabinet 1154
ED-10000M 1096
ED-1032 1124
ED-12000B
    described 1015
ED-140M 1108
ED-24000B
    described 1009
ED-64M 1122
Ethernet Port 1034

F

F_Port 804
fabric
    balanced 713
    business unit 718
    design practices 720
    mirrored 714
fan-in / fan-out 707
Fbre Channel classes, supported 1037
FC / 9000 1174
FC-AL 1352
FCIP
    described 699
    with Brocade AP7420 474
    with Cisco MDS 9000 family 482
FCIP Tunneling Service 995
Fibre Channel
    arbitrated loop 1352
    connecticity equipment 806
    levels 696
    logical topologies 702
    physical topologies 702
Fibre Channel port specifications 1023
Fibre Channel Routing Services 994
Fibre Channel switch support 994
FICON
    and EMC ControlCenter 370
    cascading 367
    zoning 367
FICON connectivity
    for CNT (Inrange) FC-9000 1421
    for Connectrix B series 371
    for Connectrix M series 374
    for Connectrix MDS series 379
    general information 366
Finisar 40 km link extender 1261

G

G_Port 804
GbE port specifications 982, 1024
gigabit Ethernet. *See* GigE
GigE remote director (Symmetrix) 968

H

hardware components
    nonport side 993
    port side 990
HBA configuration 783
hop count 746
host bus adapter configuration 783
hubs

compared to switches 806

I

iFCP
EMC-McDATA Distance Extension Solution 484
immunity, electromagnetic 1041
internetwork 962
IOCP considerations 369
IP
addressing 965
IPS 3300, IPS 4300 484
iSCSI
via Symmetrix MPCD 971
ISL guidelines 751
ISL Trunking, 994

L

LAN-based backup 561
LAN-free backup
described 566
link extenders 1261
link speeds 1033
Linux with Oracle 10g 10.2 RAC 604
logical fabric segregation 717
Lucent OptiStar EdgeSwitch 1271

M

maintenance port
configure director network addresses 237
default password 239
management
ports 992
McDATA Eclipse 1620
for iFCP 484
McDATA IPS 3300, IPS 4300 484
MDS 9000 482
MDS 9216 1091
MDS series examples 1156
Metro 1500 1210
mirrored fabric 714
MPCD
for iSCSI 971
for SRDF 968
MSRS (see Multiprotocol SAN Routing Services) 994
Multiprotocol Ports 992
Multiprotocol SAN Routing Services 994
multisite fabrics 740

N

N_Port 804
Name Server support 994
NAS backup 582
networks, described 695
nonport side view 1023
Nortel Optera Metro 1180

O

ONI. *See* Ciena
ONLINE7000 1185
Open Trunking (McDATA / Connectrix M series) 1147
Optera Metro 1180
OptiStar EdgeSwitch 1271
Oracle 10g 10.2 RAC installation for Linux 604
Oracle 10g RAC with CLARiiON or Symmetrix 598
OUSP 1266
overview 978

P

password
default maintenance port 239
persistent binding 778
port
Ethernet port 1034
serial port 1033
trunking 1035
port side view 1021
ports
Fibre Channel, description of 992
management 992
serial 992
power supply
description of 991
general information 1034
principal switch placement 755
principal switch placement and negotiation 809

protocol, ANSI 1041

Q

QPM Quad Port Module, Connectrix M 1111
QuickLoop 1391

R

RJ-45 connector 1034
routing (Fibre Channel fabric) 811
RS-232 connector 1033

S

SAN Security 496
security for SAN 496
Serial Port 1035
serial port 992
serial port specifications 981, 1024
serverless backup
    overview 571
Silkworm 2xxx 1065
Silkworm 3800 1060
SLES 9 Oracle 607
SRDF
    via Symmetrix MPCD 968
support
    Name Server 994
    Zone Server 994
switch, principal 755
switches
    compared to hubs 806
    maximum per fabric 747
Symmetrix iSCSI multiprotocol channel director 971
Symmetrix MPCD
    for iSCSI 971
    for SRDF 968

T

tape connectivity 743
tapes and fabrics 591
topologies, Fibre Channel
    logical and physical 702
trunking
    about 1035
trunking, open (McDATA / Connectrix M series) 1147

U

UDP 963
User Datagram Protocol 963
user interfaces 995

X

XPM port card 1119

Z

Zone Server support 994