

**Symmetrix
DMX800**

PRODUCT GUIDE

**P/N 300-000-664
REV A04**

EMC Corporation

Corporate Headquarters:
Hopkinton, MA 01748-9103

1-508-435-1000
www.EMC.com

Copyright © 2004 EMC Corporation. All rights reserved.

Printed February, 2004

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.

Regulatory Agency Information

Symmetrix DMX systems have been extensively tested and certified to meet UL60950, CSA 22.2 No 60950, IEC 60950/EN60950; Safety of Information Technology Equipment including Electrical Business Equipment.

Trademark Information

EMC², EMC, Symmetrix, Celerra, CLARiiON, CLARAlert, DG, E-Infostructure, HighRoad, Navisphere, PowerPath, ResourcePak, SnapView/IP, SRDF, VisualSAN, where information lives, and The EMC Effect are registered trademarks and EMC Automated Networked Storage, EMC ControlCenter, EMC Developers Program, EMC Enterprise Storage, EMC Enterprise Storage Network, EMC OnCourse, EMC Proven, EMC Snap, Access Logix, AutoAdvice, Automated Resource Manager, AutoSwap, AVALONidm, C-Clip, CacheStorm, Celerra Replicator, Centera, CentraStar, CLARevent, Connectrix, CopyCross, CopyPoint, CrosStor, Direct Matrix, Direct Matrix Architecture, EDM, E-Lab, Enginuity, FarPoint, FLARE, InfoMover, MirrorView, NetWin, OnAlert, OpenScale, Powerlink, PowerVolume, RepliCare, SafeLine, SAN Architect, SAN Copy, SAN Manager, SDMS, SnapSure, SnapView, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix DMX, TimeFinder, Universal Data Tone, and VisualSRM are trademarks of EMC Corporation. All other trademarks used herein are the property of their respective owners.

Preface	xvii
Warnings and Cautions	xxiii
Regulatory Information	xxvii
Chapter 1 Introducing the Symmetrix DMX800 System	
Symmetrix DMX800 System.....	1-2
Symmetrix DMX800 System Configurations.....	1-2
Symmetrix DMX Performance and Functionality	1-4
Symmetrix Platform and Enginuity Operating Environment...	1-6
Enginuity Operating Environment	1-6
EMC's Solutions Enabler APIs.....	1-7
Storage Capacities and Global Memory Requirements	1-8
Storage Capacities.....	1-8
Global Memory Requirements.....	1-10
Performance Features.....	1-11
Availability and Integrity Features.....	1-12
Serviceability Features	1-13
Supported Software Applications	1-14
Platform Software	1-15
Storage Management Software.....	1-17
Information Management Software.....	1-18
Infrastructure Software	1-20

Chapter 2 **DMX800 System Hardware Components**

Major Components	2-2
DMX800 System Configuration Options.....	2-7
Overview	2-7
DMX800 Rack Configurability Options	2-8
Symmetrix DMX Architecture	2-9
DMX800 System Block Diagrams.....	2-9
Symmetrix DMX Component Scaling Attributes	2-13
DMX800 Card Cage Slot Configuration.....	2-13
Symmetrix Channel Connectivity and Host Integration	2-15
Channel Connectivity	2-15
Symmetrix Channel Configurations.....	2-16
Mainframe Serial Channel Interfaces	2-16
Supported Fibre Channel Interfaces	2-16
Supported Cluster Hosts	2-17
Fibre Channel Disk Devices	2-18
DMX800 System Disk Device Capacities	2-18
Symmetrix DMX800 System Logical Volume Capacities ..	2-19
Symmetrix DMX Disk Device Emulations	2-21
Deleting (and Then Adding) Devices Online.....	2-21
Open System Disk Emulation.....	2-22
IBM DASD Disk Emulation	2-23
Channel Directors, FEBE Boards, and Global Memory	
Directors.....	2-26
Channel Director Connectivity.....	2-26
Channel Director Descriptions	2-26
Fibre Channel Directors (Front-End).....	2-27
Multiprotocol Channel Directors	2-29
Front-End/Back-End Boards	2-32
Fibre Channel Directors for SRDF.....	2-34
Global Memory Directors.....	2-34
Global Memory Director Configuration	2-34
Symmetrix DMX800 Power Subsystem.....	2-35
Storage Processing Enclosure Power Supplies.....	2-35
Disk Array Enclosure Power Supply.....	2-36
Standby Power Supply	2-37
Channel Attachments.....	2-38
FICON Channel Interface Connections.....	2-38
Additional FICON Features.....	2-40
Open Systems Fibre Channel Interface	
Connections	2-41

Chapter 3 Symmetrix DMX800 Input/Output Operations

Symmetrix DMX Disk Subsystem Operation	3-2
Symmetrix Global Memory Management.....	3-3
Elements of a Symmetrix I/O Operation	3-6
I/O Response Time: Mainframe Environment.....	3-6
I/O Response Time: Open Systems Environment	3-7
Symmetrix I/O Operations	3-7
Read Operations.....	3-9
Write Operations	3-10
Write Destaging Operation.....	3-12
Functional Operation.....	3-14
Overview	3-14
MPCD Functional Description.....	3-17
DMX800 Backplane.....	3-19
FEBE Board Functions.....	3-20
Global Memory Director Board Functions	3-21
Host Interface	3-21
Disk Subsystem Interface.....	3-21
DMX800 Communications	3-22
Environmental Control	3-23
Read/Write Operations.....	3-24
Overview	3-24
Read Operations.....	3-24
Write Operations	3-25
I/O Performance Enhancements	3-26

Chapter 4 Performance and Optimization

Overview	4-2
Global Memory Performance Features	4-3
Global Memory ASICs	4-4
Fast Write Capabilities.....	4-5
Dynamic Mirror Service Policy (DMSP) Algorithm	4-5
Disk Rotational Position Ordering (RPO)	4-5
Disk Multiple Priority Queues.....	4-6
PermaCache Option.....	4-6
Quality of Service Control Performance Option	4-7
Symmetrix File System Performance Features	4-9
Dynamically Adjusting Performance Algorithms	4-9
Enhancement of Dynamic Mirror Service Policy	4-9
Enhancement of Symmetrix Optimizer	4-9
More Rapid Recovery from Problems.....	4-10
Enhanced System Audit and Investigation.....	4-10

Multiple Channel Directors	4-11
Channel Speeds and Cable Lengths	4-11
FICON Channels	4-12
iSCSI Channels	4-12
GigE Remote Channels	4-12
Host Connectivity	4-12
Parallel Processing	4-13
RPS Miss Elimination	4-13
Open Systems Hypervolumes	4-14
Hypervolume Extension Feature	4-14
Disk Device Cylinders	4-14
Logical Volume Mapping	4-15
Metavolumes	4-15
Mainframe Systems Hypervolumes	4-18
Hypervolume Extension Options	4-18
Split-Volume Capability	4-18
Extended Cylinder Addressing Option	4-19
Determining Cylinders for Hypervolume User Data	4-20
Optimizing Symmetrix System Performance	4-22
Performance Guidelines for Open System Devices	4-22
Multiport Volume Access for Open Systems Environments ...	4-23

Chapter 5 **Data Integrity, Availability, and Protection**

Overview	5-2
Symmetrix Reliability and Availability Features	5-2
Symmetrix Data Integrity Protection Features	5-2
Data Protection Options	5-3
Reliability and Availability Features	5-7
Reliable Components	5-7
Global Memory Director Data Integrity	5-7
Fibre Channel Front-End Redundancy	5-7
Fibre Channel Back-End Redundancy	5-9
Fibre Channel Arbitrated Loop Design	5-9
Redundant Power Subsystem	5-11
Nondisruptive Component Replacement	5-11
Nondisruptive Engineuity Upgrades	5-12
Nondisruptively Change or Remove Drives	5-13
Deleting (and Then Adding) Devices Online	5-13
Maintaining Data Integrity	5-14
Remote Support	5-14
Error Checking and Correction, and Data Integrity Protection	5-15
Disk Error Correction and Error Verification	5-16
Global Memory Director Data Integrity Logic	5-17

Global Memory Error Correction and Error Verification ..	5-17
Global Memory Chip-Level Redundancy	5-18
Longitude Redundancy Code (LRC)	5-18
Byte-Level Parity Checking	5-18
Global Memory Access Path Protection	5-18
Data Protection Guidelines	5-19
Configuring Data Protection Options	5-20
Disk Mirroring Concepts.....	5-21
Advantages of Mirroring	5-21
Write Operations with Mirroring	5-21
Read Operations with Mirroring	5-22
Error Recovery with Mirroring	5-22
Dynamic Mirror Service Policy	5-23
Business Continuance Volumes	5-23
Symmetrix DMX800 Parity RAID Options.....	5-24
Parity RAID Technology	5-24
Data Protection Flexibility	5-25
Parity RAID Components.....	5-25
Parity RAID Modes of Operation.....	5-29
Writing, Reading, and Rebuilding Data with Parity RAID.....	5-31
Writing Data in a Parity RAID Group.....	5-31
Reading Data in a Parity RAID Group	5-33
Data Recovery with Parity RAID	5-33
Symmetrix DMX800 RAID 5 Options	5-37
RAID 5 Overview	5-37
Temporary Member Sparing	5-38
RAID 5 Configuration Rules	5-38
Symmetrix RAID 10	5-39
Dynamic Sparing.....	5-40
Dynamic Sparing as Additional Protection	5-42
Dynamic Sparing with Locally Mirrored Pairs.....	5-42
Dynamic Sparing with Symmetrix Parity RAID Volumes	5-43
Dynamic Sparing with Remotely Mirrored Pairs (SRDF)	5-45

Chapter 6 Mainframe Features and Support

Introduction	6-2
Supported Mainframe Features	6-3
Dynamic Channel Management	6-3
Dynamic Path Reconnection	6-3
Concurrent Copy	6-4
MultiPath Lock Facility/Concurrent Access	6-4
Multi-Subsystem Imaging	6-4

Sequential Data Striping	6-5
Partitioned Data Set (PDS) Assist	6-5
Multiple Allegiance (MA)	6-5
Compatible Parallel Access Volumes (COM-PAV)	6-5
Dynamic Parallel Access Volumes	6-6
RAID 10 Striping	6-7
Logical Paths per FICON Port and Control Unit	
Image Support	6-7
FICON Cascading and Open Systems Intermix	6-7
PPRC Command Support	6-7
XRC Support	6-8
Configuring CKD Volumes	6-9
Deleting (and then Adding) Devices Online	6-10
Support for 64K Cylinders	6-10
Error Reporting and Recovery	6-11
Types of Errors	6-11
Error Reporting	6-13
Event Messages	6-16
Operator Messages	6-16
EREP Reports	6-18
Error Handling	6-19
Detecting the Error	6-19
Sense Byte Information	6-21
Console Error Messages	6-21
Host Sense Byte Data Formats	6-23

Appendix A DMX800 System Specifications

Storage Control	A-2
Physical Data	A-10
Environmental Data	A-12
Power and Cooling Data	A-13
Power Consumption and Heat Dissipation	A-13
Power Requirements	A-14

Appendix B Power Sequences

Shutting the System Down	B-2
Powering Up the Symmetrix System after a Shutdown	B-4

Appendix C Symmetrix DMX800 System Planning and Installation

Planning Overview	C-2
Physical Specifications	C-2

Transportation and Delivery Guidelines.....	C-3
Power Requirements	C-3
Symmetrix Earth Leakage Current Compliance	C-3
Environmental Specifications.....	C-6
System Cabling Requirements	C-6
Layout and Space Requirements	C-6
Remote Support.....	C-7
Director/Global Memory Board Layout	C-8
Mainframe/Open Systems Installations	C-10
FICON/ FEBE Board Port Designations	C-10
Available EMC FICON Cables	C-12
Open Systems Installations.....	C-13
GigE Remote and iSCSI Director Installations	C-22
GigE and iSCSI Cable Precautions	C-22
Available EMC GigE/iSCSI Channel Cables	C-22
 Glossary	 g-1
 Index	 i-1

1-1	Symmetrix DMX800 System	1-5
1-2	Engenuity and the Storage Platform Relationships	1-6
2-1	DMX800 System Configured with Two FC Directors	2-4
2-2	DMX800 System Rear View	2-5
2-3	Fully Configured DMX800 System with Four FC Directors	2-6
2-4	Block Diagram of a DMX800 Configured with Two Fibre Channel Directors	2-10
2-5	Block Diagram of a DMX800 Configured with Four Fibre Channel Directors	2-11
2-6	Block Diagram of a DMX800 Configured with Two Fibre Channel Directors and Two MPCD	2-12
2-7	Symmetrix DMX800 Two Fibre Channel Director Card Cage Configuration	2-14
2-8	Symmetrix DMX800 Four Fibre Channel Director Card Cage Configuration	2-14
2-9	Track Format for 3390 and 3380 DASD	2-24
2-10	Storage Processing Enclosure Power Supply	2-35
2-11	Disk Array Enclosure (Rear View)	2-36
2-12	Standby Power Supply	2-37
2-13	FICON Channel Attachments	2-39
3-1	Host Global Memory Use	3-2
3-2	Symmetrix Global Memory Management and Data Flow	3-3
3-3	LRU and Age Link Chain Data Flow	3-4
3-4	I/O Response Time (Mainframe Environment)	3-6
3-5	I/O Response Time (Open Systems Environment)	3-7
3-6	Symmetrix I/O Operations	3-8
3-7	Read Operations	3-9
3-8	Read Hit	3-9
3-9	Read Miss	3-10
3-10	Write Operations	3-11

3-11	Fast Write	3-11
3-12	Delayed Fast Write	3-12
3-13	Destaging Operation	3-13
3-14	Functional Block Diagram of a DMX800 Configured with Two FC Directors	3-15
3-15	Functional Block Diagram of a DMX800 Configured with Four FC Directors	3-16
3-16	Functional Block Diagram of a DMX800 Configured with Two FC Directors and Two MPCD	3-18
3-17	DMX800 Two FC Director System Backplane	3-19
3-18	DMX800 Four FC Director System Backplane	3-20
3-19	DAE Fibre Channel Loops	3-22
4-1	Enginuity LRU Resource Pools	4-8
4-2	Logical Volume Mapping (8:1)	4-15
4-3	Concatenated Volumes	4-16
4-4	Striped Data	4-17
5-1	Data Record Format for Conventional DASD	5-15
5-2	Symmetrix Data Record Format	5-16
5-3	Symmetrix Parity RAID Group Disk Director Redundancy	5-27
5-4	Parity RAID Group without Hypervolumes	5-27
5-5	Parity RAID Group with Hypervolumes	5-28
5-6	Parity Protection Logic	5-30
5-7	Symmetrix Parity RAID Write Operation	5-32
5-8	Writing to a Symmetrix Parity RAID Group in Reduced Mode	5-34
5-9	Reading from a Symmetrix Parity RAID Group in Reduced Mode	5-35
5-10	Regenerating Data or Parity after Disk Replacement	5-36
5-11	RAID 5 Data/Parity (3+1)	5-37
5-12	RAID 10 with Dynamic Mirror Service Policy	5-39
5-13	Dynamic Sparing Process	5-40
5-14	Dynamic Sparing with Locally Mirrored Pairs	5-42
5-15	Example of Dynamic Sparing with RAID (3+1) Volumes	5-44
6-1	Dynamic Support of Parallel Access Volumes	6-6
6-2	PPRC and GDPS Support	6-8
6-3	XRC Overview	6-9
6-4	OS/390 or z/OS IEA480E Service Alert Error Message Format (AC Power Failure)	6-16
6-5	OS/390 or z/OS IEA480E Service Alert Error Message Format (Mirror-1 Volume in Not Ready State)	6-17
6-6	OS/390 or z/OS IEA480E Service Alert Error Message Format (Mirror-2 Resynchronization)	6-17
6-7	OS/390 or z/OS IEA480E Service Alert Error Message Format (Mirror-1 Resynchronization)	6-18
6-8	Typical Console Error Message	6-21
B-1	Disk Array Enclosure, Front View	B-2

B-2	Inlines Icon	B-3
B-3	AC Power Cords	B-4
B-4	DAE and SPS Power Switches	B-5
C-1	DMX800 System Service Area	C-7
C-2	DMX800 Two Fibre Channel Director and Global Memory Configuration	C-8
C-3	DMX800 Four Fibre Channel Director and Global Memory Configuration	C-9
C-4	FEBE Board Channel Designations	C-16

1-1	DMX800 Configurations	1-3
1-2	Symmetrix DMX800 System Disk Storage Capacities	1-9
1-3	Performance Features Roadmap	1-11
1-4	Availability and Integrity Features Roadmap	1-12
1-5	Platform Software Options	1-15
1-6	Storage Management Software Options	1-17
1-7	Information Management Software Options	1-18
1-8	Infrastructure Software Options	1-20
2-1	Symmetrix DMX800 Component Overview	2-2
2-2	IBM Controller/DASD Compatibility	2-16
2-3	Symmetrix DMX800 Disk Capacities	2-19
2-4	Logical Volumes Supported on Symmetrix DMX Disk Devices	2-20
2-5	Logical Volumes Supported for Symmetrix DMX Models	2-20
2-6	IBM DASD Emulation Characteristics	2-23
2-7	Symmetrix Devices and Addressing Capabilities for Fibre Channel Directors	2-28
2-8	FICON Director Configurations	2-30
2-9	Fibre Channel Disk Director (Back-End) Configurations	2-33
2-10	Supported Symmetrix Point-to-Point FICON Cable Distances	2-39
2-11	Symmetrix Fibre Channel Cable Distances	2-41
4-1	Symmetrix Open Systems Disk Capacities and Cylinders	4-14
4-2	Symmetrix Mainframe Disk Capacities and Cylinders	4-20
4-3	Device Emulations and Number of Cylinders	4-21
5-1	Number of Supported Channel Directors	5-8
5-2	Data Protection Options	5-19
6-1	Environmental Alert Messages	6-15
6-2	Event Messages	6-16
6-3	Error Handling Steps	6-19
6-4	Unit Status Bits	6-22
6-5	Channel Status Bits	6-22

A-1	Storage Capacities (73 GB Mirrored Disk Devices)	A-3
A-2	Storage Capacities (73 GB SRDF Disk Devices)	A-3
A-3	Storage Capacities (73 GB RAID (3+1) Disk Devices)	A-4
A-4	Storage Capacities (73 GB RAID (7+1) Disk Devices)	A-4
A-5	Storage Capacities (73 GB RAID 5 (3+1) Disk Devices)	A-5
A-6	Storage Capacities (73 GB RAID 5 (7+1) Disk Devices)	A-5
A-7	Storage Capacities (146 GB Mirrored Disk Devices)	A-6
A-8	Storage Capacities (146 GB SRDF Disk Devices)	A-6
A-9	Storage Capacities (146 GB RAID (3+1) Disk Devices)	A-7
A-10	Storage Capacities (146 GB RAID (7+1) Disk Devices)	A-7
A-11	Storage Capacities (146 GB RAID 5 (3+1) Disk Devices)	A-8
A-12	Storage Capacities (146 GB RAID 5 (7+1) Disk Devices)	A-8
A-13	Symmetrix DMX800 Weights	A-10
A-14	Power Consumption/Heat Dissipation	A-13
A-15	DMX800 Electrical Specifications, Single-Phase	A-14
A-16	DMX800 Power Cables and Connectors, Single-Phase	A-14
A-17	DMX800 Extension Line Cords, Single-Phase	A-14
C-1	Preinstallation Responsibility Summary	C-2
C-2	EMC Fibre Cables —FICON 9 Micron Connect	C-12
C-3	Symmetrix DMX800 Checklist for UNIX or PC Server Hosts	C-13
C-4	UNIX or PC Server Host Checklist	C-14
C-5	EMC Fibre Cables —Fibre Channel Connect	C-21
C-6	GigE/iSCSI Channel Cables	C-22

As part of its effort to continuously improve and enhance the performance and capabilities of the Symmetrix product line, EMC periodically releases new versions of the hardware and the EMC Enginuity Operating Environment. Therefore, some functions described in this manual may not be supported by all versions of Enginuity or Symmetrix hardware currently in use.

If your Symmetrix unit does not offer a function described in this guide, please contact your EMC representative for a hardware, software, or Enginuity update.

Audience This manual is part of the Symmetrix DMX series documentation set, and is intended for use by storage administrators, system programmers, or operators who are involved in acquiring, managing, or operating the Symmetrix rackmount system.

Organization Here is an overview of where information is located in this product guide.

Chapter 1, *Introducing the Symmetrix DMX800 System*, provides an overview of the Symmetrix rackmount system highlighting its performance and reliability features, and describes hardware and software options for the unit.

Chapter 2, *DMX800 System Hardware Components*, introduces the hardware components of the Symmetrix rackmount system. It describes its main components, their functions, and the types of host channels and devices to which the Symmetrix system can attach.

Chapter 3, *Symmetrix DMX800 Input/Output Operations*, discusses integrated global memory, disk arrays, I/O operation, and global memory management.

Chapter 4, *Performance and Optimization*, describes the Symmetrix system performance features, how they affect overall performance, and how to use these features to get the best performance from the Symmetrix system.

Chapter 5, *Data Integrity, Availability, and Protection*, discusses the Symmetrix features that affect data availability and reliability.

Chapter 6, *Mainframe Features and Support*, describes the types of errors possible with Symmetrix DMX800 system, mainframe features, error-handling techniques, and provides an error recovery summary.

Appendix A, *DMX800 System Specifications*, lists the operating, physical, and environmental specifications for the rackmount system.

Appendix B, *Power Sequences*, provides step-by-step instructions for powering the Symmetrix system on and off.

Appendix C, *Symmetrix DMX800 System Planning and Installation*, contains several worksheets for product installation.

A *Glossary* and an *Index* are also provided.

Related Documentation

For additional information on all Symmetrix-related publications, contact your EMC Sales Representative or refer to the EMC Powerlink website at:

<http://powerlink.emc.com>

Conventions Used in This Guide

EMC uses the following conventions for notes, cautions, warnings, and danger notices.

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



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.



WARNING

A warning contains information essential to avoid a hazard that can cause severe personal injury, death, or substantial property damage if you ignore the warning.



DANGER

A danger notice contains information essential to avoid a hazard that will cause severe personal injury, death, or substantial property damage if you ignore the message.

Typographical Conventions

EMC uses the following type style conventions in this guide:

AVANT GARDE	Keystrokes
Palatino, bold	◆ Dialog box, button, icon, and menu items in text ◆ Selections you can make from the user interface, including buttons, icons, options, and field names
<i>Palatino, italic</i>	◆ New terms or unique word usage in text ◆ Command line arguments when used in text ◆ Book titles
Courier, italic	Arguments used in examples of command line syntax.
Courier	System prompts and displays and specific filenames or complete paths. For example: working root directory [/user/emc]: c:\Program Files\EMC\Symapi\db
Courier, bold	◆ User entry. For example: sympoll -p ◆ Options in command line syntax

Abbreviations for Units of Measure

Abbreviation	Definition
A or Amp	ampere
b	bit
B	byte
b/s	bits per second
B/s	bytes per second
Btu	British thermal units
cm	centimeters
ft	foot
Gb	gigabits
GB	gigabytes
GHz	gigahertz
Hz	hertz
hr	hours
in.	inches
Kb	kilobits
Kb/s	kilobits per second
KB	kilobytes
KB/s	kilobytes per second
kg	kilograms
KHz	kilohertz
km	kilometers
kV	kilovolt
kW	kilowatt
kWh	Kilowatt-hour
Mb	megabits

Abbreviation	Definition
MB	megabyte
MB/s	megabytes per second
MHz	megahertz
m	meter
mm	millimeter
ms	millisecond
lbs	pounds
ns	nanosecond
s	second
sq.	square
Sq. in.	square inch
Sq. ft.	square foot
V	volt
vol	volume
W	watt

Where to Get Help

For DMX800

Obtain technical support by calling your local sales office.

For service, call:

United States: (800) 782-4362 (SVC-4EMC)

Canada: (800) 543-4782 (543-4SVC)

Worldwide: (508) 497-7901

and ask for Customer Support.

Sales and Customer Service Contacts

For the list of EMC sales locations, please access the EMC home page at:

<http://www.emc.com/contact/>

For additional information on the EMC products and services available to customers and partners, refer to the EMC Powerlink website at:

<http://powerlink.emc.com>

Your Comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Please send a message to **techpub_comments@emc.com** with your opinions of this guide.

Warnings and Cautions

The following warnings and cautions pertain throughout this guide:

WARNING

Trained service personnel only.

This EMC product has more than one power supply cord. To reduce the risk of electric shock, disconnect all power supply cords before servicing.

Ground circuit continuity is vital for safe operation of the machine. Never operate the machine with grounding conductors disconnected. Remember to reconnect any grounding conductors removed for or during any installation procedure.

ATTENTION

Reservé au personnel autorisé.

Cet appareil EMC comporte plus d'un cordon d'alimentation. Afin de prévenir les chocs électriques, débranchez tous les cordons d'alimentation avant de faire le dépannage.

Un circuit de terre continu est essentiel en vue du fonctionnement sécurisé de l'appareil. Ne mettez jamais l'appareil en marche lorsque le conducteur de mise à la terre est débranché.

WARNUNG

Nur für autorisiertes Fachpersonal.

Dieses EMC Produkt verfügt über mehrere elektrische Netzanschlüsse. Zur Vermeidung eines elektrischen Schlages sind vor Servicearbeiten an der Stromversorgung alle Netzanschlüsse zu trennen.

Kontinuierliche Erdung ist notwendig während der gesamten Betriebsdauer des Gerätes. Es ist unzulässig das Gerät ohne Erdung zu betreiben. Gerät muss geerdet werden, bevor es am Stromnetz angeschlossen wird.

Additional Warnings and Cautions

Before attempting to service EMC hardware described in this document, observe the following additional Warnings and Cautions:



WARNING

The hardware enclosure contains no user-serviceable parts, so it should not be moved or opened for any reason by untrained persons. If the hardware needs to be relocated or repaired, only qualified personnel familiar with safety procedures for electrical equipment and EMC hardware should access components inside the unit or move the unit.



WARNING

This product operates at high voltages. To protect against physical harm, power off the system whenever possible while servicing.



WARNING

In case of fire or other emergency involving the EMC product, isolate the product's power and alert appropriate personnel.



CAUTION

Trained personnel are advised to exercise great care at all times when working on the EMC hardware.

Remember to:

- ◆ Remove rings, watches, or other jewelry and neckties before you begin any procedures.
- ◆ Use caution near any moving part and any part that may start unexpectedly such as fans, motors, solenoids, etc.
- ◆ Always use the correct tools for the job.
- ◆ Always use the correct replacement parts.
- ◆ Keep all paperwork, including incident reports, up to date, complete, and accurate.

Static Precautions

EMC incorporates state-of-the-art technology in its designs, including the use of LSI and VLSI components. These chips are very susceptible to damage caused by static discharge and need to be handled accordingly.

**CAUTION**

Before handling printed circuit boards or other parts containing LSI and/or VLSI components, observe the following precautions:

- ◆ Store all printed circuit boards in antistatic bags.
 - ◆ Use a ground strap whenever you handle a printed circuit board.
 - ◆ Unless specifically designed for nondisruptive replacement, never plug or unplug printed circuit boards with the power on. Severe component damage may result.
-

Symmetrix DMX800 System

This section contains the regulatory information pertaining to the Symmetrix DMX800 system and the DMX800 Storage Processing Enclosure (SPE).

Symmetrix DMX800 systems have been extensively tested and certified to meet UL60950, CSA 22.2 No. 60950, IEC 60950/EN60950; Safety of Information Technology Equipment including Electrical Business Equipment.

기기의 명칭(모델명) : Symmetrix DMX800

인증받은 자의 상호 : EMC Corporation

제 조 년 월 일 : 2004

제 조 자/제 조 국가 : EMC Ireland

DMX800 Storage Processing Enclosure

This equipment generates, uses, and may emit radio frequency energy. The equipment has been type tested and found to comply with the limits for a Class A digital device pursuant to Part 15 of FCC rules, which are designed to provide reasonable protection against such radio frequency interference. Operation of this equipment in a residential area may cause interference, in which case the user, at his own expense, will be required to take whatever measures may be required to correct the interference. Any modifications to this device — unless expressly approved by the manufacturer — can void the user's authority to operate this equipment under Part 15 of the FCC rules.

This Class A digital apparatus complies with Canadian ICES-003. Cet appareil numérique de la classe A est conforme à la norme NMB-003 du Canada.

Warning!

This is a Class A product. In a domestic environment this product may cause radio interference in which case the user may be required to take adequate measures.

Achtung!

Dieses ist ein Gerät der Funkstörgrenzwertklasse A. In Wohnbereichen können bei Betrieb dieses Gerätes Rundfunkstörungen auftreten, in welchen Fällen der Benutzer für entsprechende Gegenmaßnahmen verantwortlich ist.

Attention!

Ceci est un produit de Classe A. Dans un environnement domestique, ce produit risque de créer des interférences radioélectriques, il appartiendra alors à l'utilisateur de prendre les mesures spécifiques appropriées.

警告使用者：

這是甲類的資訊產品，在居住的環境中使用時，可能會造成射頻干擾，在這種情況下，使用者會被要求採取某些適當的對策。

この装置は、情報処理装置等電波障害自主規制協議会（V C C I）の基準に基づくクラス A 情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

Introducing the Symmetrix DMX800 System

This chapter provides an overview of the EMC Symmetrix DMX800 system and highlights its performance, availability, and serviceability features, and hardware and software options.

- ◆ Symmetrix DMX800 System.....1-2
- ◆ Symmetrix Platform and Enginuity Operating Environment.....1-6
- ◆ Storage Capacities and Global Memory Requirements.....1-8
- ◆ Performance Features.....1-11
- ◆ Availability and Integrity Features.....1-12
- ◆ Serviceability Features.....1-13
- ◆ Supported Software Applications1-14

Symmetrix DMX800 System

Symmetrix DMX™ is EMC's next generation family of high-end storage solutions targeted to meet the uncompromising requirements of mission-critical applications, environments, and customers. Incorporating a revolutionary and future-proof Direct Matrix Architecture™, the Symmetrix® DMX800 system establishes a new performance trajectory for the high-end, fully leverages EMC's industry-leading storage management functionality, and introduces the economic benefits of modular packaging to the high-end storage marketplace for the first time.

The patented Symmetrix Direct Matrix Architecture (*Fully Configured DMX800 System with Four FC Directors* on page 2-6) and the *Block Diagram of a DMX800 Configured with Two Fibre Channel Directors* on page 2-10 is a fundamentally new storage array technology that employs a matrix of dedicated, point-to-point connections instead of traditional buses or switches. DMX delivers direct access, from the front of the storage array to the back, guaranteeing the highest possible I/O throughput, up to 17.6 GB/s aggregate internal bandwidth. It's faster, less costly, and more reliable, and it avoids contention, latency, and bandwidth issues and potential points of failure associated with bus and switch architectures.

Combined with expanded global memory director technology and the dynamically optimized caching algorithms of the Enginuity™ storage operating environment, systems based on the Symmetrix DMX architecture deliver scalable performance to meet the most demanding information access, protection, and distribution requirements.

Figure 1-1 on page 1-5 provides a front view of the exterior of the Symmetrix DMX800 system (EMC-supplied cabinet).

Symmetrix DMX800 System Configurations

Symmetrix DMX800 system has several configuration options as described in Table 1-1.

Table 1-1 DMX800 Configurations

Component	Two Fibre Channel Directors ^a	Four Fibre Channel Directors
Directors		
Fibre Channel	2	4
Multiprotocol Channel Director (MPCD)^b		
FICON	2	0
GigE/iSCSI	2	0
FICON/GigE/iSCSI	2	0
Disks	8 - 60	61 - 120
Disk Array Assembly (DAE)	2 - 4	5 - 8
RS-232	1 pair	1 pair
Storage Processing Enclosure SPE(s)	1	1
Standby Power Supply Assembly (SPS)	1 for SPE 2 for 4 DAEs	1 for SPE 4 for 8 DAEs
Global Memory	64 GB	64 GB
Message Matrix Board (MMB)^c	1 or 0 if MPCD is installed	NA
Front-End Ports		
Fibre Channel	8	16
FICON/GigE/iSCSI	4	0
Back-End Ports^d		
Fibre Channel	8	16
FICON/GigE/iSCSI	None	None

a. In a two FC director configuration only, Dir 1 and Dir 16 are always Fibre Channel. Dir 2 and Dir 15 can be FICON, GigE, or iSCSI combinations.

b. The MPCD can be added to a DMX800 configured with two FC directors. The MPCD will occupy the slots that the blank panel and the Message Matrix Board (MMB) occupy. The MPCD will become Dir 2 and Dir 15, and will be used for FICON GigE/iSCSI combinations. Refer to *Multiprotocol Channel Directors* on page 2-29.

- c. The MMB is only used in a DMX800 configured with two FC directors. It is removed along with the blank panel when installing the MPCD.
- d. On Dir 2, Dir 15, only FEBE board ports A0, D0 are used for FICON Gige/SCSI configurations.

The Symmetrix DMX800 is the industry's only modular rack-mounted high-end storage array. The two director configuration provides an economic entry for high-end storage, easily upgradeable to a four director DMX800 configuration to meet needs for future capacity, growth, or additional connectivity requirements.

All DMX800 configurations support EMC's industry-leading functional software to meet critical service level requirements. The DMX800 can be deployed inside or outside of the data center to extend the foundation of an Automated Networked Storage infrastructure.

Throughout this product guide, the term *DMX800* is used to describe the common features and operations of the Symmetrix DMX800 system configurations just described in this section.

Symmetrix DMX Performance and Functionality

With full compatibility in integration with EMC's comprehensive EMC Automated Network Storage™ (ANS) storage management suite, the Symmetrix DMX series delivers the ultimate in operational efficiency for both large-scale consolidations and for remote or distributed deployment. Interoperability with existing Symmetrix 8000 series enables the cascading of your current Symmetrix install-base storage that protects your investment even as your storage demands expand.

The performance, functionality, and economic characteristics of the Symmetrix DMX Series establishes a new standard for high-end storage. Combined with EMC's Automated Network Storage solutions, the Symmetrix DMX series allows you to optimize your storage investment to consolidate, protect, and manage ever-growing information assets.



Figure 1-1 Symmetrix DMX800 System

Symmetrix Platform and Enginuity Operating Environment

The Symmetrix DMX800 hardware architecture (refer to *Major Components* on page 2-2) and the Enginuity operating environment are the foundation for the Symmetrix storage platform, which consists of the following:

- ◆ Symmetrix rackmount hardware
- ◆ Enginuity-based operating environment functions
- ◆ EMC Solutions Enabler Applications Program Interfaces (APIs)
- ◆ Symmetrix-based applications
- ◆ Host-based Symmetrix applications
- ◆ Independent Software Vendor (ISV) applications

The relationships among these software layers (and Symmetrix hardware) are illustrated in Figure 1-2.

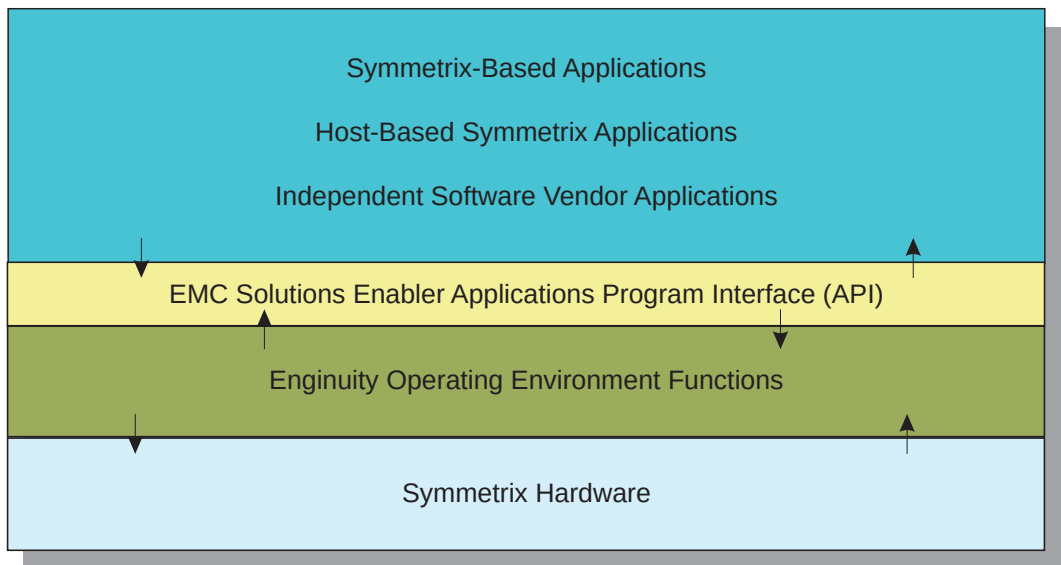


Figure 1-2 Enginuity and the Storage Platform Relationships

Enginuity Operating Environment

Symmetrix Enginuity is the operating environment for the Symmetrix DMX800 systems. Enginuity manages and ensures the optimal flow and integrity of information through the different hardware components of the Symmetrix system.

Enginuity manages all Symmetrix operations from monitoring and optimizing internal data flow, to ensuring the fastest response to the user's requests for information, to protecting and replicating data.

Enginuity Services

Enginuity provides the following services for the Symmetrix DMX800 systems:

- ◆ Manages system resources to intelligently optimize performance across a wide range of I/O requirements.
- ◆ Ensures system availability through advanced fault monitoring, detection, and correction capabilities, and provides concurrent maintenance and serviceability features.
- ◆ Interrupts and prioritizes tasks from microprocessors and, for example, ensures that fencing off failed areas takes precedence over other operations.
- ◆ Offers the foundation for specific software features available through EMC's disaster recovery, business continuance, and storage management software.
- ◆ Provides functional services for both its host Symmetrix systems and for a large suite of EMC storage application software.
- ◆ Defines priority of each task including basic system maintenance, I/O processing, application processing (for example, EMC ControlCenter™, SRDF®, TimeFinder™, and EMC ControlCenter Symmetrix Optimizer).
- ◆ Provides uniform access through APIs for internal calls and provides an external interface to allow integration with other software providers and ISVs.

EMC's Solutions Enabler APIs

EMC's Solutions Enabler APIs, are the storage management programming interfaces that provide an access mechanism for managing the Symmetrix, third-party storage, switches, and host storage resources. They enable the creation of storage management applications that don't have to understand the management details of each piece within the total storage environment.

For more information on EMC's Solutions Enabler APIs and ANS, contact your EMC Sales Representative or refer to the EMC Powerlink™ website at: <http://powerlink.emc.com>

Storage Capacities and Global Memory Requirements

Storage Capacities

A DMX800 configured with two Fibre Channel directors is available with eight to 60, 1-inch 73 GB or 146 GB disk devices, 10K speed. When configured with four FC directors, the DMX800 is available with 61 to 120 disk devices. Table 1-2 on page 1-9 shows the storage capacities for system operation. The capacities are based on storage capacity of each disk drive and the following storage protection options:

- ◆ Mirrored (RAID 1)
- ◆ SRDF
- ◆ Parity RAID (3+1) and Parity RAID (7+1)
- ◆ RAID 5 (3+1) and RAID 5 (7+1)

The GB values shown in Table 1-2 on page 1-9 are based on the following conventions: The GB₂ value is based on 1 GB = 1024 x 1024 x 1024 bytes. The GB₁₀ value is based on 1 GB = 1000 x 1000 x 1000 bytes. Although these values are expressed differently, they are equivalent.

For detailed storage capacities, refer to Appendix A. For Symmetrix DMX systems with 73 GB and 146 GB disk drive capacities when attached to iSeries hosts, contact your EMC Sales Representative.

A Symmetrix system reserves four 3 GB (6,140 cylinders) logical volumes (12 GB total system requirement) of disk capacity for internal Symmetrix file system purposes.

Table 1-2 Symmetrix DMX800 System Disk Storage Capacities

Available Disk Drives	Physical Disk Devices	Usable Open Systems Storage Capacities (GB ₁₀)	Usable Open Systems Storage Capacities (GB ₂)	Usable Mainframe Storage Capacities (GB ₁₀)
73 GB Disk Devices SRDF Capacities				
8 to 120	8 to 120	584 to 8,772	544 to 8,169	577 to 8,660
73 GB Disk Devices RAID 1 Mirrored Capacities				
4 to 60	8 to 120	292 to 4,386	272 to 4,084	288 to 4,330
73 GB Disk Devices Parity RAID (3+1) Capacities				
6 to 90	8 to 120	438 to 6,579	408 to 6,127	433 to 6,495
73 GB Disk Devices Parity RAID (7+1) Capacities				
7 to 105	8 to 120	511 to 7,675	469 to 7,043	505 to 7,577
73 GB Disk Devices RAID 5 (3+1) Capacities				
6 to 90	8 to 120	438 to 6,579	408 to 6,127	433 to 6,495
73 GB Disk Devices RAID 5 (7+1) Capacities				
7 to 105	8 to 120	511 to 7,675	469 to 7,043	505 to 7,577
146 GB Disk Devices SRDF Capacities				
8 to 120	8 to 120	1,168 to 17,520	1,088 to 16,320	1,156 to 17,352
146 GB Disk Devices RAID 1 Mirrored Capacities				
4 to 60	8 to 120	584 to 8,760	544 to 8,160	578 to 8,676
146 GB Disk Devices Parity RAID (3+1) Capacities				
6 to 90	8 to 120	876 to 13,140	816 to 12,240	867 to 13,014
146 GB Disk Devices Parity RAID (7+1) Capacities				
7 to 105	8 to 120	1,022 to 15,330	952 to 14,280	1,012 to 15,183
146 GB Disk Devices RAID 5 (3+1) Capacities				
6 to 90	8 to 120	876 to 13,140	816 to 12,240	867 to 13,014
146 GB Disk Devices RAID 5 (7+1) Capacities				
7 to 105	8 to 120	1,022 to 15,330	952 to 14,280 m	1,012 to 15,183

Global Memory Requirements

The DMX800 system is available with global memory sizes ranging from 4 GB to 64 GB. The total global memory requirement for a Symmetrix system is based upon specific configurations and customer requirements. Besides the customer's applications, other variables that affect the amount of global memory Symmetrix DMX800 systems require include:

- ◆ Number of disk drives
- ◆ Disk capacity (73 GB, 146 GB)
- ◆ Number of logical volumes

Additional information on memory configurations can be found in *Global Memory Directors* on page 2-34.

Performance Features

Table 1-3 identifies many of the Symmetrix DMX and Engenuity supported features that enhance performance and increase throughput.

Table 1-3 Performance Features Roadmap

Feature	Document Sources
Direct Matrix (DMX) Architecture with up to 16 direct nonblocking data paths in the DMX800 two Fibre Channel director configuration, 32 nonblocking data paths in the four Fibre Channel director configuration	• <i>Symmetrix DMX Architecture</i> on page 2-9. • <i>Symmetrix DMX Component Scaling Attributes</i> on page 2-13.
Symmetrix DMX nonvolatile global memory directors for optimized performance	• <i>Global Memory Performance Features</i> on page 4-3. • <i>Memory Striping</i> on page 4-4. • <i>Global Memory Directors</i> on page 2-34. • <i>Global Memory Director Configuration</i> on page 2-34.
One hundred percent global memory fast write capabilities	• <i>Write Operations</i> on page 3-25. • <i>Disk Mirroring Concepts</i> on page 5-21.
PermaCache memory option	• <i>PermaCache Option</i> on page 4-6.
2 Gb/s Fibre Channel drive infrastructure	• <i>Fibre Channel Disk Devices</i> on page 2-18.
• Multiple scalable channel directors and global memory directors • FICON channel speeds up to 2 Gb/s • Fibre Channel speeds up to 2 Gb/s • iSCSI Channel speeds up to 1 Gb/s • GigE Remote Director speeds up to 1 Gb/s • Front-End/Back-end Boards • FICON Cascading and Open Systems Intermix	• <i>Channel Directors, FEBE Boards, and Global Memory Directors</i> on page 2-26. • <i>FICON Mezzanine Cards</i> on page 2-29. • <i>Fibre Channel Directors (Front-End)</i> on page 2-27. • <i>iSCSI Channel Directors</i> on page 2-32. • <i>Gigabit Ethernet Remote Directors</i> on page 2-31. • <i>Front-End/Back-End Boards</i> on page 2-32. • <i>Additional FICON Features</i> on page 2-40.
Hypervolume Extension option	• <i>Open Systems Hypervolumes</i> on page 4-14. • <i>Mainframe Systems Hypervolumes</i> on page 4-18.
Rotational Position Sense (RPS) miss elimination	• <i>RPS Miss Elimination</i> on page 4-13.
3380 and 3390 mixed track geometry	• <i>IBM DASD Disk Emulation</i> on page 2-23.
Compatible Parallel Access Volumes (COM-PAV)	• <i>Compatible Parallel Access Volumes (COM-PAV)</i> on page 6-5.
Dynamic Parallel Access Volumes	• <i>Dynamic Parallel Access Volumes</i> on page 6-6.
PPRC Command Support	• <i>PPRC Command Support</i> on page 6-7.
Quality of Service Control Performance Option	• <i>Quality of Service Control Performance Option</i> on page 4-7.

Availability and Integrity Features

The Symmetrix DMX800 systems include key enhancements that improve reliability, availability, and serviceability. Highlights of the Symmetrix DMX availability include the following features (Table 1-4).

Table 1-4 Availability and Integrity Features Roadmap

Feature	Document Sources
Proactive error detection and remote support	• <i>Maintaining Data Integrity</i> on page 5-14. • <i>Error Checking and Correction, and Data Integrity Protection</i> on page 5-15. • <i>Disk Error Correction and Error Verification</i> on page 5-16. • <i>Global Memory Director Data Integrity Logic</i> on page 5-17.
Support for online Enginuity upgrades and updates	• <i>Nondisruptive Enginuity Upgrades</i> on page 5-12.
Fully fault-tolerant design with redundant critical components and concurrent maintenance support	• <i>Reliability and Availability Features</i> on page 5-7.
Channel director redundancy with end-to-end automatic channel failover and load balancing	• <i>Fibre Channel Front-End Redundancy</i> on page 5-7.
Internal Control Data Path redundancy	• <i>Fibre Channel Front-End Redundancy</i> on page 5-7.
Fibre Channel back-end functionality featuring redundant disk directors, disk channels, FEBC boards, disk ports, and port bypass circuitry	• <i>Fibre Channel Back-End Redundancy</i> on page 5-9. • <i>Fibre Channel Arbitrated Loop Design</i> on page 5-9.
(N+1) Power supply redundancy	• <i>Redundant Power Subsystem</i> on page 5-11.
Symmetrix Mirroring option	• <i>Disk Mirroring Concepts</i> on page 5-21.
Parity RAID (3+1), Parity RAID (7+1), RAID 5 (3+1), RAID 5 (7+1), and RAID 10 data protection options	• <i>Symmetrix DMX800 Parity RAID Options</i> on page 5-24. • <i>Symmetrix DMX800 RAID 5 Options</i> on page 5-37 • <i>Symmetrix RAID 10</i> on page 5-39.
Dynamic sparing option	• <i>Dynamic Sparing</i> on page 5-40.
Nondisruptive component replacement Nondisruptive Enginuity code upgrades Nondisruptive change or remove drives	• <i>Reliability and Availability Features</i> on page 5-7. • <i>Nondisruptive Component Replacement</i> on page 5-11. • <i>Nondisruptive Enginuity Upgrades</i> on page 5-12 • <i>Nondisruptively Change or Remove Drives</i> on page 5-13

Serviceability Features

Each Symmetrix DMX800 unit has an integrated server that continuously monitors the Symmetrix environment. The server communicates with the EMC Customer Support Center through a customer-supplied direct phone line. The service processor automatically dials the Customer Support Center whenever the Symmetrix system detects a component failure or environmental violation. An EMC Product Support Engineer at the Customer Support Center can also run diagnostics remotely through the service processor to determine the source of a problem and potentially resolve it before the problem becomes critical.

Symmetrix DMX800 systems feature a modular design with a low parts count for quick component replacement, should a failure occur. This low parts count minimizes the number of failure points.

The Symmetrix DMX800 systems feature nondisruptive replacement of its major components, including:

- ◆ Channel director boards
- ◆ FEBE boards
- ◆ Global memory director boards
- ◆ Disk devices and DAEs
- ◆ Cooling fan modules
- ◆ Power supplies
- ◆ Batteries (Standby Power Supplies—SPS)
- ◆ Server

For more information on these components, refer to Chapter 2, *DMX800 System Hardware Components* on page 2-1.

Supported Software Applications

Symmetrix Engenuity operating environment is the foundation for many Symmetrix DMX800 software options. Tables 1-5 through 1-8 provide lists and brief descriptions of many of these options. The tables are divided into three categories:

- ◆ *Platform Software* (Table 1-5 on page 1-15)
- ◆ *Storage Management Software* (Table 1-6 on page 1-17)
- ◆ *Information Management Software* (Table 1-7 on page 1-18)
- ◆ *Infrastructure Software* (Table 1-8 on page 1-20)

Documents for these software options are available on the EMC Powerlink website at:

<http://powerlink.emc.com>

All of the software products are furnished under a license. Refer to the copyright page in this manual for the complete licensing statement. For software license, model numbers, prerequisites, and additional information, contact your EMC Sales Representative.

Platform Software

Symmetrix DMX systems support platform software applications for data migration, data mobility, replication, integration, and more.

Table 1-5 describes these software application option.

Table 1-5 Platform Software Options

Option	Description
EMC AutoSwap for Planned Outages	A utility that allows the user to change the contents of a mainframe Unit Control Block (UCB) such that the application does not know that the swap happened. It is useful for planned outages such as maintenance events and the migration of work from one resource to another. This high speed swap function is done at the discretion of the user with either an operator command or a batch job as the triggering event.
EMC AutoSwap for Unplanned Outages	Used in conjunction with EMC consistency groups (part of SRDF for mainframes), which is its prerequisite. This product performs the same operations as AutoSwap for Planned Outages, but automatically. Thus a regional disaster or site-level disaster would result in the contents of a UCB being swapped from the compromised site to the surviving site and the data being operated upon would be the R2 copies instead of the R1 copies. Again, the application would not take an interruption while the swapping was going on.
EMC CopyPoint™ for OS/400	Allows a secondary, offline copy of production data to be accessed by a secondary iSeries system.
EMC Double Checksum	Provides an end-to-end safeguard against data corruption by rechecking Oracle's data validation before data is written to EMC Symmetrix systems.
Symmetrix Enterprise Storage Platform (ESP)	Allows simultaneous storage and access of mainframe data and open systems data on the same Symmetrix system, or between two or more open systems and mainframe Symmetrix models.
EMC Fast Mirror Migrator	A mainframe software product that reads data from one Symmetrix system to another Symmetrix system using the channel connectivity of the mainframe to both control units. It is used for migrating data either on a permanent basis (such as the replacement of an older Symmetrix system) or on a temporary basis (such as changing a workload to a different Symmetrix resource). The user interface to Fast Mirror Migrator is via ISPF panels, a common interface technique in the mainframe community. Both <i>source</i> Symmetrix (the sender) and <i>target</i> Symmetrix (the receiver) can be channel connected with FICON.
EMC InfoMover	Provides high-speed transfer between heterogeneous and similar MVS and UNIX-based host systems through a shared Symmetrix Enterprise Storage system.
SRDF/Asynchronous (SRDF/A)	High performance extended distance asynchronous replication solution using a Delta Set architecture for reduced bandwidth requirements and no host performance impact while also providing a consistent, restartable remote copy at all times.
SRDF/ Automated Availability Manager	Automated high availability through storage based replication and host application clustering via SRDF family and Legato AAM (a division of EMC).

Table 1-5 Platform Software Options (continued)

Option	Description
SRDF/Automated Replication (SRDF/AR)	Rapid business restart over any distance with no data exposure through advanced single-hop and multi-hop configurations using combinations of TimeFinder/BCV and SRDF/S and/or SRDF/DM.
SRDF/Cluster Enabler for MSCS and SRDF/Cluster Enabler for VCS	High availability and automated failover through storage based replication and server clustering via SRDF family and MSCS and VCS.
SRDF/Consistency Groups (SRDF/CG)	Maintain data coherency across an SRDF-based configuration by monitoring data propagation from the source volumes to their corresponding target volumes to ensuring data consistency and restartability.
SRDF/Data Mobility	Rapidly transfer data from source volumes to remote volumes quickly, permitting information to be shared, content to be distributed, and access to be local to additional processing environments.
SRDF/Synchronous (Formerly SRDF)	High performance, host-independent, real-time synchronous remote replication from one Symmetrix to one or more Symmetrix systems.
TimeFinder/Consistency Groups	Maintain data coherency across TimeFinder-protected volumes to ensuring data consistency and restartability.
TimeFinder/Exchange and SQL Integration Modules	Tight integration of TimeFinder Family with Microsoft Exchange and SQL applications for automated backup and restore.
TimeFinder/Mirror (Formerly TimeFinder)	Create and use point-in-time business continuance volumes – BCVs, Clones or Mainframe SNAPS - for parallel processing activities and local restore.
TimeFinder/Snap	Pointer-based, space-saving snapshot copies for reduced recovery points and fast local data restore.

Storage Management Software

Symmetrix DMX systems support Storage Management software that improves information management by allowing users to consolidate storage capacity for multiple hosts and servers. EMC offers powerful graphical user interface (GUI)-based tools that simplify and enhance Symmetrix configurations, performance, and status information gathering and management. Table 1-6 describes the storage management software options.

Table 1-6 Storage Management Software Options

Option	Description
EMC ControlCenter (minimum version 5.1.1 required for EMC ControlCenter products)	ControlCenter suite of software enables users to simplify and automate management of multivendor storage infrastructure through a single consistent information-centric strategy. It enables users to: • Discover: Map and display all elements and relationships of multivendor storage topology. • Monitor: View status and performance of those elements in real time; integrate with leading framework applications. • Provision: Configure and reconfigure resources to meet changing business demands. • Report: Obtain real-time and historical information on performance, utilization, and allocation.
EMC ControlCenter Symmetrix Manager	Provides real-time monitoring, configuration, and control of Symmetrix storage systems. SRDF/TimeFinder Manager is now folded into Symmetrix Manager.
EMC ControlCenter Automated Resource Manager	Provides storage resource management capabilities in three key areas: policy-based provisioning, automated alert management, and consolidated management of multivendor hosts, databases, backup systems, and mainframe tape devices.
EMC ControlCenter Symmetrix Optimizer	Automatically monitors, analyzes, and moves highly active logical volumes within the Symmetrix system to maintain optimal performance.
EMC ControlCenter StorageScope	Provides reporting capabilities that enable users to answer questions such as, 'How much storage do I have?', 'Where do I have it?', and 'Who's using it?' to facilitate capacity planning and chargeback operations.
EMC ControlCenter Workload Analyzer	Collects, correlates, and graphically presents performance information for your end-to-end storage infrastructure. It supports EMC CLARiiON® and Symmetrix arrays, multi-vendor switches, open system servers, Oracle databases, and mainframe hosts.
SAN (Storage Area Network) Manager	Provides a central point of access, monitoring, and control for multivendor servers, storage networks, and storage arrays. It provides centralized configuration management (LUN Masking and Zoning) for multivendor storage arrays and network devices.

Information Management Software

Symmetrix DMX Systems support Information Management software that replicate critical business information. The SRDF® and TimeFinder Families enable you to balance your performance, availability, functionality and economic requirements to achieve required service levels for disaster recovery and business continuity.

TimeFinder/Mirror (formerly TimeFinder) enables you to make locally resident business continuance volumes for parallel processing activities like backup, testing and development and local restore, while your business applications proceed uninterrupted, servicing customers and generating revenue. TimeFinder/Snap™ offers instantly available pointer-based snapshot copies for more frequent, cost- and resource-efficient backup, restore, and mirroring tasks.

Symmetrix Remote Data Facility/Synchronous (SRDF/S®, formerly SRDF) is EMC's remote, synchronous, byte-for-byte data-mirroring application for business continuity and disaster recovery needs while SRDF/Asynchronous' (SRDF/A) innovative Delta Set architecture enables unlimited distance replication with no host application and significantly reduced bandwidth requirements.

The SRDF and TimeFinder families can be used in conjunction with each other for even greater flexibility when architecting complete information management solutions.

Table 1-7 describes the Information Management software options.

Table 1-7 Information Management Software Options

Option	Description
SRDF/Asynchronous (SRDF/A)	High performance extended distance asynchronous replication solution using a Delta Set architecture for reduced bandwidth requirements and no host performance impact while also providing a consistent, restartable remote copy at all times.
SRDF/ Automated Availability Manager	Automated high availability through storage based replication and host application clustering via SRDF family and Legato AAM (a division of EMC).
SRDF/Automated Replication (SRDF/AR)	Rapid business restart over any distance with no data exposure through advanced single-hop and multi-hop configurations using combinations of TimeFinder/BCV and SRDF/S and/or SRDF/DM.
SRDF/Cluster Enabler for MSCS and SRDF/Cluster Enabler for VCS	High availability and automated failover through storage based replication and server clustering via SRDF family and MSCS and VCS.
SRDF/Consistency Groups (SRDF/CG)	Maintain data coherency across an SRDF-based configuration by monitoring data propagation from the source volumes to their corresponding target volumes to ensuring data consistency and restartability.

Table 1-7 Information Management Software Options (continued)

Option	Description
SRDF/Data Mobility	Rapidly transfer data from source volumes to remote volumes quickly, permitting information to be shared, content to be distributed, and access to be local to additional processing environments.
SRDF/Synchronous (Formerly SRDF)	High performance, host-independent, real-time synchronous remote replication from one Symmetrix to one or more Symmetrix systems.
TimeFinder/Consistency Groups	Maintain data coherency across TimeFinder-protected volumes to ensuring data consistency and restartability.
TimeFinder/Exchange and SQL Integration Modules	Tight integration of TimeFinder Family with Microsoft Exchange and SQL applications for automated backup and restore.
TimeFinder/Mirror (Formerly TimeFinder)	Create and use point-in-time Business Continuity Volumes (BCVs), Clones or Mainframe SNAPs - for parallel processing activities and local restore.
TimeFinder/Snap	Pointer-based, space-saving snapshot copies for reduced recovery points and fast local data restore.
Catalog Solution	A comprehensive facility for OS/390 and z/OS that enables users to manage, maintain, and recover their catalog environment with great speed, reliability, and flexibility. The Catalog Solution approach includes simple proactive diagnostic tools. Its maintenance and management facilities provide unique DUMP and REBUILD utilities to help you repair many potential logical and structural defects. Catalog Solution's fast and comprehensive validation of catalog, VVDS and VTOC relationships makes it a valuable resource for ensuring system (or data) availability.
EMC ControlCenter Replication Manager	Operates in combination with replication technologies, like EMC TimeFinder, to dramatically simplify and accelerate the management and use of local disk-based replications.
EMC ControlCenter Symmetrix Data Mobility Manager	A centralized management tool that allows you to configure, monitor, simulate, and manage the replication of data between local or remote Symmetrix devices based upon your business applications and/or business rules.
CopyCross	Allows copying mainframe tape data to Symmetrix disk storage for enhanced information availability and protection. Once the tape data has crossed over to Symmetrix storage, the mirroring capabilities of TimeFinder and SRDF can be used to achieve workload compression for increased productivity, easy and frequent local or remote backups, quick disaster recovery for critical business systems, and optimized IT processes.
EMC CopyPoint for OS/400	Allows a secondary, offline copy of production data to be accessed by a secondary iSeries system.
EMC Data Manager	A powerful combination of hardware, software, and services in a centralized, high-performance backup and restore system.

Infrastructure Software

Symmetrix DMX systems support Infrastructure software applications that provide logical presentation and seamless mobility of information. Table 1-8 describes the infrastructure software options.

ResourcePak For Windows is no longer available, however its components can now be found in the following products: Data Relocation Utility, TimeFinder/Exchange Integration Module and TimeFinder/SQL Integration Module.

Table 1-8 Infrastructure Software Options

Option	Description
EMC PowerPath® for Open System	Combines multiple path I/O capabilities, automatic load balancing, path failover, logical volume management, and volume mobility functions into one integrated package—for use on open server platforms connected to EMC Symmetrix, EMC CLARiiON, Hitachi Data Systems (HDS) Lightning, and HP XP (Hitachi OEM) storage systems.
Data Relocation Utility	Enables fast and flexible data mobility and migration operations for Windows environments.

This chapter describes the main hardware components of the Symmetrix DMX800 systems including:

- ◆ Major Components2-2
- ◆ DMX800 System Configuration Options2-7
- ◆ Symmetrix DMX Architecture.....2-9
- ◆ Symmetrix Channel Connectivity and Host Integration2-15
- ◆ Fibre Channel Disk Devices.....2-18
- ◆ Channel Directors, FEBE Boards, and Global Memory Directors...2-26
- ◆ Symmetrix DMX800 Power Subsystem2-35
- ◆ Channel Attachments2-38

Major Components

The Symmetrix DMX800 is a disk subsystem that houses all storage control unit functions and DASD in a rackmount cabinet. This section describes the Symmetrix DMX800 system components.

Table 2-1 on page 2-2 lists and describes the components of the DMX800 system.

Table 2-1 Symmetrix DMX800 Component Overview

Component	Description
Rackmount Cabinet	40 U rack houses all of the DMX800 components.
Storage Processing Enclosure (SPE)	Contains the following components: - Fan module - Two 12V, 1080W power supplies - Director backplane Card cage containing: - Two or four Fibre Channel director boards, or - Two FC directors and two MPCD containing a combination of: - Two FICON, or - One FICON/one GigE, or - One FICON/or one GigE, and/or one iSCSI (combination of two) - Two global memory director boards - Message Matrix Board (MMB) —Two FC director DMX800 system only, no MPCD - Two Front-End/Back-End (FEBE) boards
Disk Array Enclosure (DAE)	Each DAE contains the following components: - Two redundant power and cooling modules for disk devices - Two link control card (LCC) modules - Front panel of each DAE contains EMI shielding - Two to eight DAEs per SPE - Two FC director configuration: - Two to four DAEs per rack - From 4 to 15 FC disk drives per DAE (8 drive system minimum) - Four FC director configuration: - Five to eight DAEs per rack - From 4 to 15 FC disk drives per DAE (61 drive system minimum)
Two AC Power Connectors	Left and right rear of system.
Power Distribution Units (left and right side) Including:	Left and right side AC power strips (rear of cabinet). Two circuit breakers (rear top left and right sides) and the rear door lock.

Table 2-1 Symmetrix DMX800 Component Overview (continued)

Component	Description
Standby Power Supply (SPS) Assembly	One SPS for every two DAEs (up to four SPS assemblies) and one SPS assembly for the SPE (maximum five SPS assemblies per DMX800 system). Provides redundant power (850W).
Symmetrix Server	Contains a CD-ROM drive and a 1.44-inch floppy drive. The DMX800 server supports up to four DMX800 SPEs. Each server contains three Ethernet ports that connect to the following: • One port to the SPE • One port to the customer network • One port used by EMC Global Service
Keyboard, Video Display, Mouse (KVM) (for secure accounts)	1U, connects to server.
RS-232 Switch (Two physical units 1U each)	RS-232 switch is standard for DMX800 systems.

Figure 2-1 on page 2-4 shows the front view and location of the main components of a DMX800 system containing two FC directors and the maximum of four DAEs.

Figure 2-2 on page 2-5 shows the rear view of a DMX800 system containing four DAEs.

Figure 2-3 on page 2-6 shows a four FC director DMX800 system containing the maximum of eight DAEs.



WARNING

To reduce the risk of personal injury, do not open the doors or move the Symmetrix DMX800 systems unless you are qualified and familiar with safety procedures for electrical equipment and the Symmetrix DMX. The Symmetrix DMX800 systems contain no user-serviceable parts. They should not be moved or opened for any reason by untrained persons. If the Symmetrix system is in need of relocation or repair, only qualified personnel should access components inside the unit or move it.

Figure 2-1 shows a fully configured DMX800 with two Fibre Channel directors.

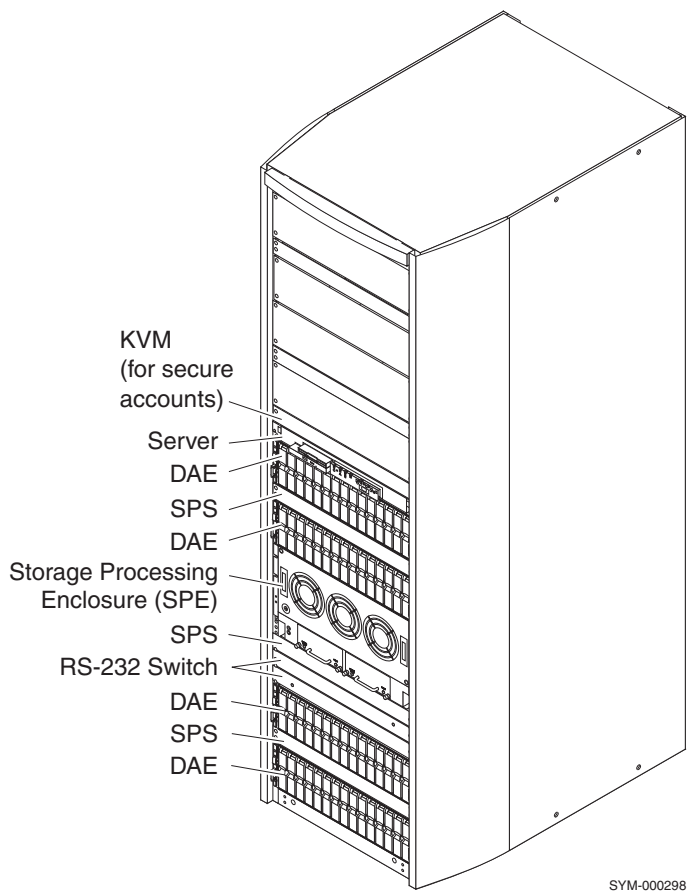
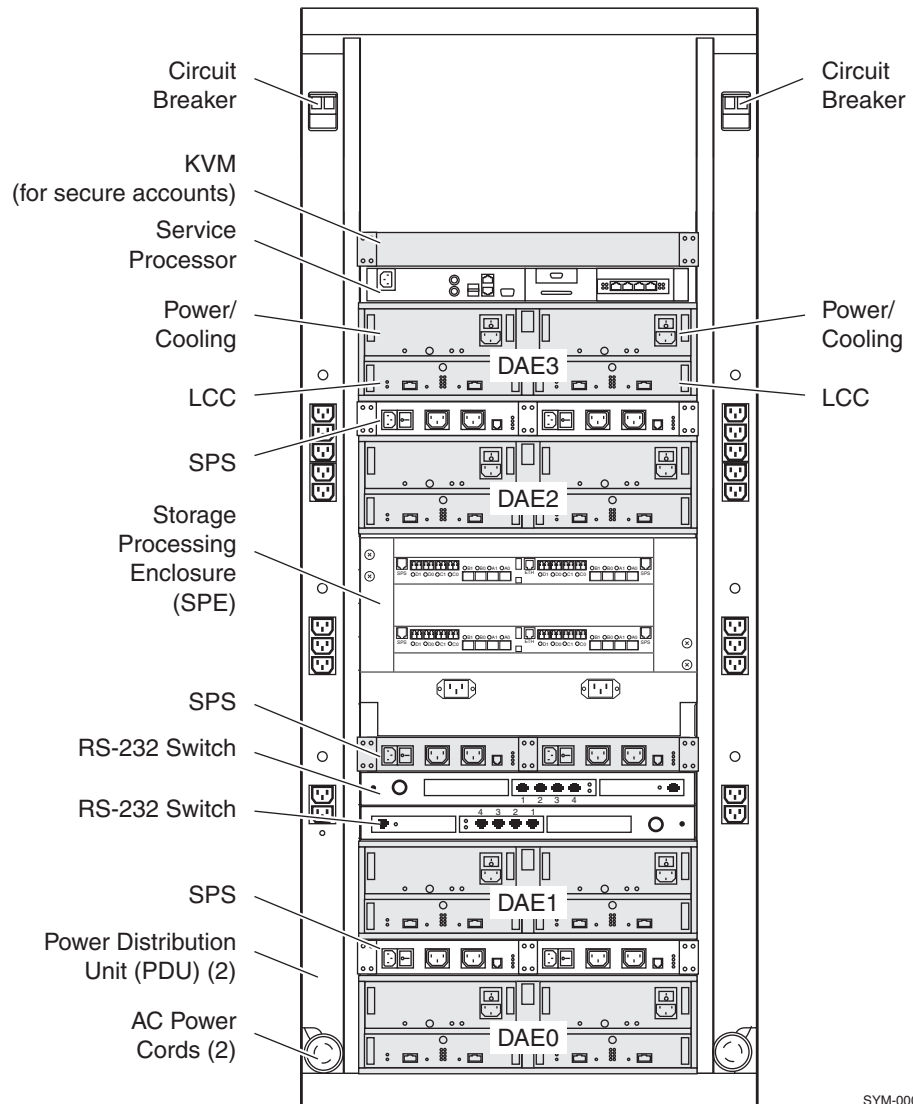


Figure 2-1 DMX800 System Configured with Two FC Directors



SYM-00040

Figure 2-2 DMX800 System Rear View

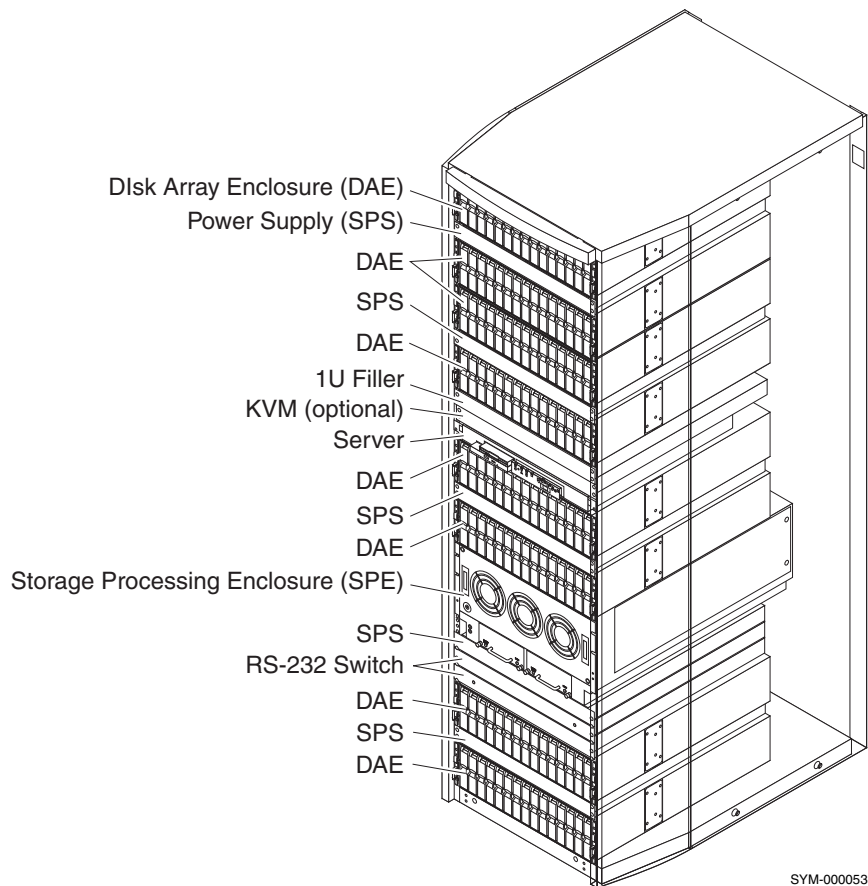


Figure 2-3 Fully Configured DMX800 System with Four FC Directors

DMX800 System Configuration Options

Overview

Each Symmetrix DMX800 with only two FC directors supports two MPCD, two Front-End/Back-End (FEBE) boards, two global memory directors (up to 64 GB of global memory), and from eight to 60, 73 GB or 146 GB disk devices.

Each Symmetrix DMX800 containing four FC directors supports two Front-End/Back-End (FEBE) boards, two global memory directors (up to 64 GB of global memory), and from 61 to 120, 73 GB or 146 GB disk devices.

The Symmetrix Direct Matrix Architecture (DMX) enables flexible configurations for customers to achieve their ideal price/performance configuration.

The DMX800 system is available with the following hardware and software configuration options:

- ◆ DMX800 with two FC directors:
 - From eight to 60 disk drives (four to 15 drives per DAE, eight drive system minimum)
 - From two to four DAEs per DMX800
- ◆ DMX800 with four FC directors:
 - From five to eight DAEs (minimum of 61 drives)
 - From 61 to 120 disk drives (four to 15 drives per DAE)
- ◆ One Standby Power Supply (SPS) assembly for every two Disk Array Enclosures (DAEs)—up to four maximum, plus one SPS assembly per SPE (maximum five SPS assemblies)
- ◆ One Storage Processing Enclosure (SPE) containing:
 - Two FC director boards and two (optional) MPCD containing:
 - FICON (mainframe connectivity)
 - GigE/iSCSI, or
 - FICON/GigE/iSCSI (combination of two)
 - Four FC director boards
 - Two Front-End/Back-End FEBE boards (multimode and single-mode adapters)
 - Message Matrix Board (MMB) when two FC directors are installed and no MPCD

- Two global memory director boards (2 GB, 4 GB, 8 GB, 16 GB, 32 GB), 64 GB maximum global memory
- ◆ Data protection options include:
 - Mirroring RAID 1
 - Mirrored RAID 10 (Striped CKD)
 - Parity RAID (3+1) or (7+1)
 - RAID 5 (3+1) or (7+1)
 - Symmetrix Remote Data Facility (SRDF)
- ◆ Software options (see *Supported Software Applications* on page 1-14)

DMX800 Rack Configurability Options

All DMX800 systems can be upgraded or expanded to meet customer's requirements.

Global memory can be increased in both the two director and four director DMX800 configurations up to 64 GB.

Two Fibre Channel Director Configurability Options

The following configurability options are available in the DMX800 with only two FC directors installed:

- ◆ Two to four DAEs — up to 60 disk drives supported
- ◆ Upgrade from two FC directors to four FC directors
- ◆ Add two MPCD combinations:
 - Two FICON directors
 - Two GigE/iSCSI directors, or
 - Two FICON/GigE/iSCSI combination directors

Four Fibre Channel Director Configurability Options

Presently, a DMX800 system with four FC directors will *only* contain Fibre Channel directors. *No MPCD will be supported.*

The following configurability options are available in the DMX800 with four FC directors:

- From five to eight DAEs (61 to 120 disk drives)
- A DMX800 configured with four FC directors can *only* connect to open systems hosts

Symmetrix DMX Architecture

The Symmetrix DMX800 features a high-performance, Direct Matrix Architecture (DMX) and supports up to 32 point-to-point serial connections in the four Fibre Channel director configuration and 16 point-to-point serial connections in the two Fibre Channel director configuration.

DMX technology is distributed across all directors and global memory boards in Symmetrix DMX800 systems. Enhanced global memory technology supports multiple regions and four connections on each global memory director. In the Direct Matrix Architecture, contention is minimized because control information and commands are transferred across a separate and dedicated message matrix. The major components of DMX800 system are the channel directors, global memory directors, and the FEBE board that functions as a both a front-end adapter and a back-end adapter.

The matrix midplane provides configuration flexibility through the slot configuration. Each director port is hard-wired point-to-point to one port on each global memory director board.

This section describes the following:

- ◆ DMX800 System Block Diagrams
- ◆ Symmetrix DMX Component Scaling Attributes
- ◆ DMX800 Card Cage Slot Configuration

DMX800 System Block Diagrams

Figure 2-4 on page 2-10 illustrates the data path architecture to include the Direct Matrix Architecture (DMX) of the Symmetrix DMX800 configured with two Fibre Channel directors.

Figure 2-5 on page 2-11 illustrates the data path architecture to include the Direct Matrix Architecture (DMX) of the Symmetrix DMX800 configured with four Fibre Channel directors.

Figure 2-6 on page 2-12 illustrates the data path architecture to include the Direct Matrix Architecture (DMX) of the Symmetrix DMX800 configured with two Fibre Channel directors and two multiprotocol channel directors (MPCD).

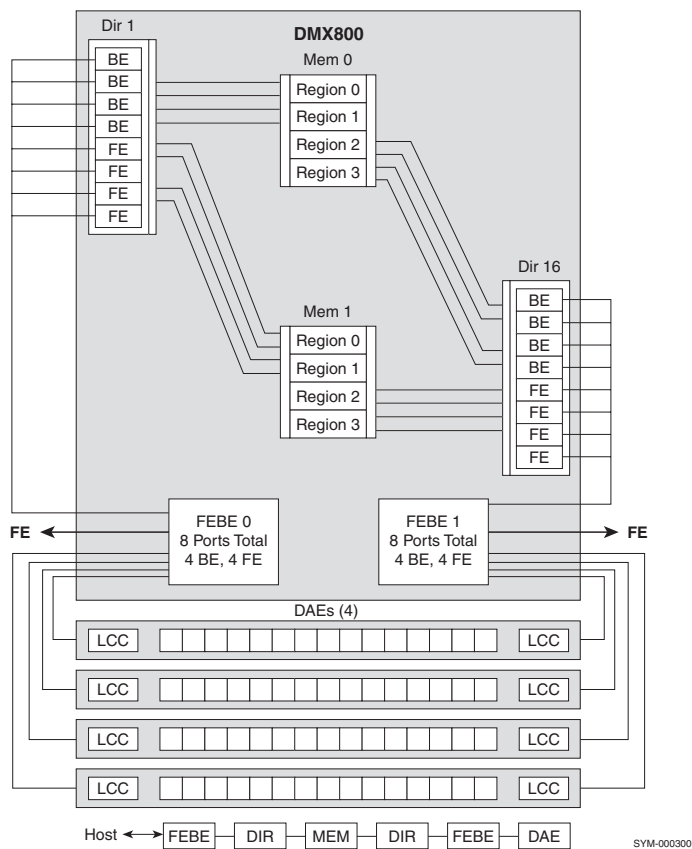


Figure 2-4 Block Diagram of a DMX800 Configured with Two Fibre Channel Directors

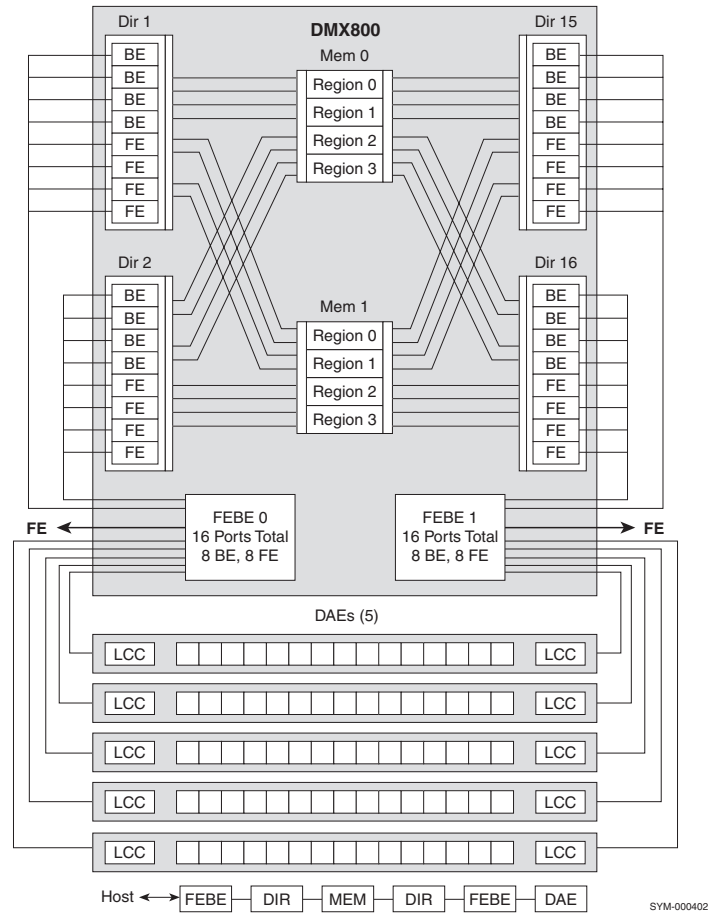


Figure 2-5 Block Diagram of a DMX800 Configured with Four Fibre Channel Directors

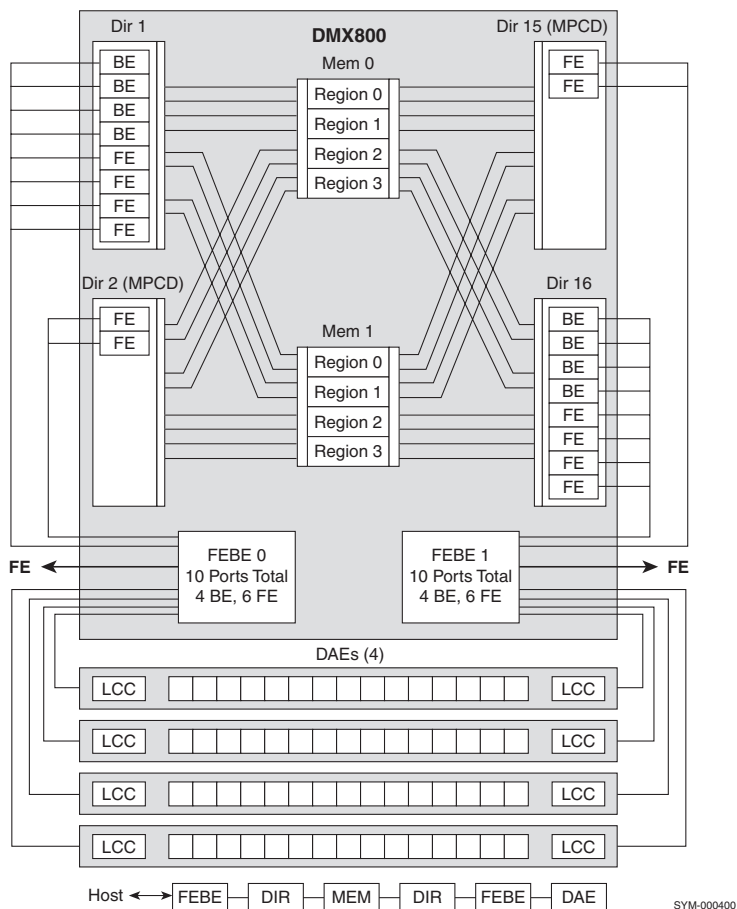


Figure 2-6 Block Diagram of a DMX800 Configured with Two Fibre Channel Directors and Two MPCD

The MPCD are only installed in the two Fibre Channel director configuration. The MPCD are installed as director 2 and director 15. The blank panel and the MMB must be removed to accommodate the MPCD in the DMX800 card cage. The MPCD can be FICON, GigE, iSCSI, or a combination of *any* two.

Symmetrix DMX Component Scaling Attributes

The Symmetrix DMX800 supports multiple configurations of front-end channel directors, global memory directors, and back-end connections to the FEBE boards as described in Table 2-1 on page 2-2. Unlike previous generations of Symmetrix systems, the Direct Matrix Architecture enables flexible configurations to fit your requirements. Two global memory directors (of the same size) are required in the DMX800, however, you can choose the size of the global memory directors (from 2 GB to 32 GB).

Unlike previous Symmetrix generations, you can choose the quantity of back-end connections that you require through the FEBE board. Again, this attribute, along with the quantity of front-end channel directors, has a direct relationship to the performance of the Symmetrix system.

DMX800 Card Cage Slot Configuration

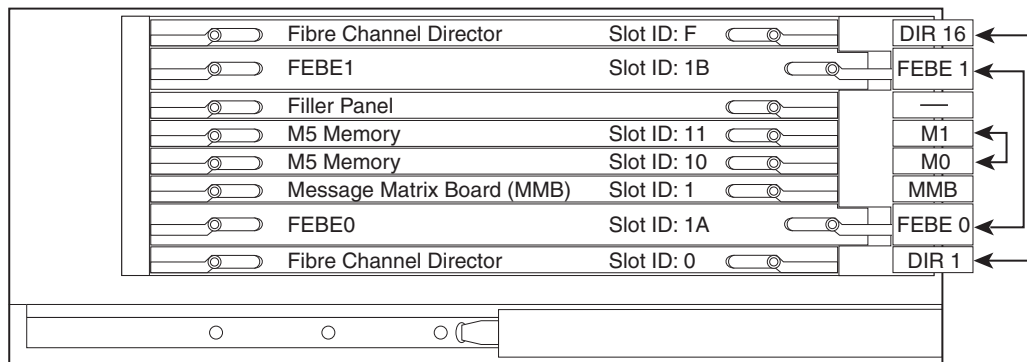
The Symmetrix DMX800 system's card cage (located inside the SPE) contains the channel directors, FEBE boards (front-end, back-end adapters), the global memory directors, and for a DMX800 configured with two FC directors, the Message Matrix Board (MMB).

The Symmetrix DMX systems require certain board types to occupy particular slot locations within the card cage. Adhering to the configurations stated for each model type ensures that the physical routing of the serial signals minimizes contention for memory resources, maximizes bandwidth within the system, and maintains total redundancy to the same memory location through two completely different hardware paths. Termination of TTL signals occurs on the FEBE boards.

Card Cage for DMX800 SPE Configured with Two FC Directors

A DMX800 configured with two Fibre Channel directors has an eight-slot backplane. However, only seven of the slots are utilized. Figure 2-7 shows the card cage layout for the global memory boards, the FEBE boards, the Fibre Channel directors, and the MMB.

When configuring a two FC director system for MPCD, the blank filler panel and the MMB are removed and the two MPCD are installed in their place (Dir 2, Dir 15).

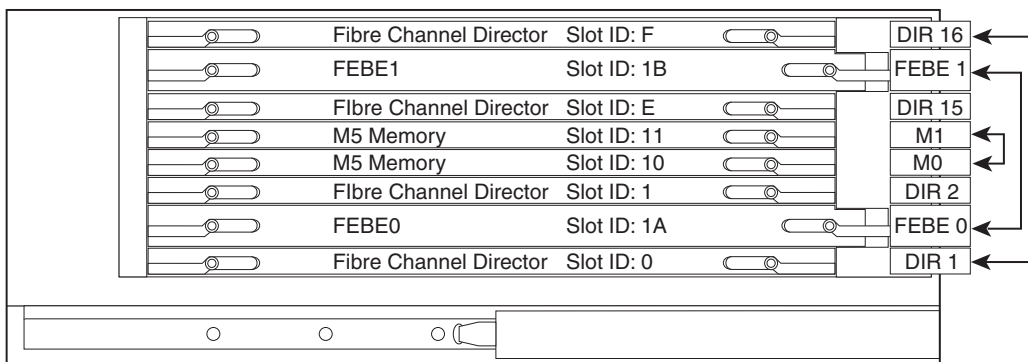


SYM-000663

Figure 2-7 Symmetrix DMX800 Two Fibre Channel Director Card Cage Configuration

Card Cage for a DMX800 SPE Configured with Four FC Directors

A DMX800 configured with four FC directors has an eight slot backplane (Figure 2-8) and is located behind a secure panel on the right side of the DMX800 chassis. It contains four FC director boards, two global memory director boards, and two FEBE boards.



SYM-000664

Figure 2-8 Symmetrix DMX800 Four Fibre Channel Director Card Cage Configuration

Symmetrix Channel Connectivity and Host Integration

The Symmetrix DMX800 systems can be integrated with all major enterprise hosts and servers. This section outlines the emulations and hosts Symmetrix systems support through Fibre Channel, FICON, GigE (SRDF), or iSCSI interfaces.

For the most current information on Symmetrix systems and specific host integration, contact your EMC Sales Representative, or refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

Channel Connectivity

The Symmetrix DMX800 systems support connectivity to mainframe and open systems hosts.

The Symmetrix Enterprise Storage Platform (ESP) software enabler, a software option, is required if you plan to store and access mainframe and open systems data on the same Symmetrix system. For more information on the ESP option, refer to Table 1-5 on page 1-15 or contact your EMC Sales Representative.

Open Systems Connectivity

The Symmetrix DMX800 systems connect to Fibre Channel open system host interfaces such as UNIX, Windows, Linux, and iSeries systems.

Mainframe Systems Connectivity

The Symmetrix DMX800 systems support FICON host connectivity to mainframe systems. Enginuity level 5670 or higher is needed to support connectivity to FICON hosts.

IBM iSeries Fibre Channel Connectivity

With the DMX800 FEBE board, users are able to connect Symmetrix systems through fibre to the iSeries 270 and 8xx models. When using directly connected fibre devices (point-to-point), the maximum distance is 500 meters.

Fibre Channel is currently capable of running up to 2 Gb/s full duplex with iSeries systems. Enginuity level 5670 or higher is needed to support Fibre Channel connectivity to iSeries systems. Symmetrix 73 GB and 146 GB disk drives support iSeries hosts.

Symmetrix Channel Configurations

- EMC offers the following Symmetrix DMX800 channel directors:
- ◆ Eight-port, four-processor (8x4) Fibre Channel directors
 - ◆ Two-port, two-processor (2x2) multiprotocol channel director that supports:
 - FICON host connections
 - iSCSI host connections
 - GigE SRDF connections

Mainframe Serial Channel Interfaces

Symmetrix serial channel interfaces attach to IBM z/9XX, z/8XX, and the plug-compatible manufacturer.

For a more detailed discussion of channel attachment options, refer to *Channel Attachments* on page 2-38.

Table 2-2 shows the IBM DASD models and controllers that Symmetrix systems emulate.

Table 2-2 IBM Controller/DASD Compatibility

IBM DASD	IBM Controller	
	3990-6	2105
3380	X	X
3390-1	X	X
3390-2	X	X
3390-3	X	X
3390-9	X	X
3390-27		X

Supported Fibre Channel Interfaces

The Symmetrix DMX800 system’s Fibre Channel interfaces attach to most open systems and iSeries hosts that have Fibre Channel connectivity.

Consult your EMC Sales Representative for the most current list of supported hosts, models, operating systems, and EMC open systems host support policies, or refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

Supported Cluster Hosts

For the most recent information on supported cluster hosts, refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

Fibre Channel Disk Devices

Symmetrix DMX800 systems use industry-standard 2 GB Fibre Channel disk drive assemblies (DDAs) for physical disks. Each DDA is configured with its own controller consisting of control logic, a microprocessor, and a device-level buffer. The device-level buffer can also be segmented. This allows DMX800 control logic to issue a read command to the device, detach, and then issue a data write to a different segment of the device-level buffer. The on-board microprocessor will manage the read or write, and will notify the DMX800 system when the read is complete.

Each DDA has the following features:

- ◆ Industry standard low-profile form factor
- ◆ Integrated dual-port Fibre Channel Arbitrated Loop (FC-AL) controller
- ◆ 2 Gb/s maximum instantaneous data transfers per port
- ◆ 73 GB disk device is available with 10,000 rpm speed
- ◆ 146 GB disk devices available with 10,000 rpm speed
- ◆ 16 MB device-level data buffer on the 73 GB disk devices
- ◆ 32 MB device-level data buffer on the 146 GB disk devices
- ◆ Fibre Channel interface transports SCSI protocol

DMX800 System Disk Device Capacities

The DMX800 with two FC directors can contain from eight to 60 disk devices, each having a storage capacity of approximately 73 GB or 146 GB. A DMX800 configured with four FC directors can contain from 61 to 120 disk drives. More precise storage capacity values depend on:

- ◆ Whether the user calculates a gigabyte to equal:
 - 1,000 × 1,000 × 1,000 bytes, or
 - 1,024 × 1,024 × 1,024 bytes
- ◆ The number of logical volumes configured on the disk
- ◆ The type of data protection option used: RAID 1 (Mirroring), Parity RAID (3+1), Parity RAID (7+1), RAID 5 (3+1), RAID 5 (7+1), RAID 10, or SRDF

Table 2-3 on page 2-19 shows the available cylinders and storage capacities for 73 GB and 146 GB disk devices based on open systems emulations. The capacities are approximate because the number of logical volumes and data protection options are not factored into the total drive capacity.

The GB values shown in Table 2-3 are based on the following conventions:
 The GB₂ value is based on 1 GB = 1024 x 1024 x 1024 bytes
 The GB₁₀ value is based on 1 GB = 1000 x 1000 x 1000 bytes
 Although the open systems GB values are expressed differently, they are equivalent.

Table 2-3 Symmetrix DMX800 Disk Capacities^a

Disk Device Size	Cylinders for 3380 Emulation	3380 Capacity in GB ₁₀	Cylinders for 3390 Emulation	3390 Capacity in GB ₁₀	Open Systems Cylinders	Open Systems GB ₁₀	Open Systems GB ₂
73 GB	98,204	69.93	84,910	72.17	148,734	73.1	68.08
146 GB	196,393	139.8	170,136	144.6	297,799	146	136

a. These capacities are approximate.

**Symmetrix DMX800
System Logical
Volume Capacities**

The maximum number of logical volumes supported on Symmetrix DMX800 physical disk devices depends on the data protection used. Table 2-4 describes the logical volumes supported on Symmetrix DMX disk devices. The Symmetrix DMX800 systems can support up to 4,840 logical volumes. The logical-to-physical relationship that you choose can automatically apply to all devices in the unit.

Table 2-4 summarizes the requirements for configuring logical volumes for Fibre Channel directors on Symmetrix DMX800 system.

Table 2-4 Logical Volumes Supported on Symmetrix DMX Disk Devices^a

Symmetrix DMX 73 GB and 146 Disk Devices and Data Protection ^b	Maximum Logical Volumes
Logical Volumes per Disk Device with RAID 10 Protection (mainframe data volumes only)	128
Logical Volumes per Disk Device with SRDF Protection (without local data protection)	77
Logical Volumes per Disk Device with RAID 1 Protection	77
Logical Volumes per Disk Device with Parity RAID (3+1) Protection	58
Logical Volumes per Disk Device with Parity RAID (7+1) Protection ^c	39
Logical Volumes per Disk Device with RAID 5 (3+1) Protection	255
Logical Volumes per Disk Device with RAID 5 (7+1) Protection	255

a. The maximum logical volumes represented are the current limits for Enginuity 5670.

b. For information on Symmetrix DMX data protections options, refer to Table 5-2 on page 5-19.

c. Parity RAID (7+1) is not supported on DMX800 systems with less than four DAEs.

For more information on logical volumes, refer to *Open Systems Hypervolumes* on page 4-14. Configuration requirements for Symmetrix systems vary according to the applications used. To configure logical volumes for optimum Symmetrix system performance, consult your EMC Sales Representative.

Table 2-5 describes the maximum logical volumes supported on each Symmetrix DMX800 configuration by data protection option.

Table 2-5 Logical Volumes Supported for Symmetrix DMX Models ^a

Symmetrix DMX800	Data Protection Option	Maximum Number of Disk Drives	Maximum Number of Disk Director Boards	Maximum Logical Volumes per System
DMX800 with Two FC Directors	RAID 10	60	2	529
	RAID 1	60	2	1,210
	Parity RAID (3+1)	60	2	1,815
	Parity RAID (7+1) ^b	60	2	2,048
	RAID 5 (3+1)	60	2	2,048
	RAID 5 (7+1)	60	2	1,024
	SRDF	60	2	2,420
DMX800 with Four FC Directors	RAID 10	120	4	1,059

Table 2-5 Logical Volumes Supported for Symmetrix DMX Models (continued)^a

Symmetrix DMX800	Data Protection Option	Maximum Number of Disk Drives	Maximum Number of Disk Director Boards	Maximum Logical Volumes per System
DMX800 with Four FC Directors	RAID 1	120	4	2,420
	Parity RAID (3+1)	120	4	3,630
	Parity RAID (7+1) ^c	120	4	4,095
	RAID 5 (3+1)	120	4	4,096
	RAID 5 (7+1)	120	4	2,048
	SRDF	120	4	4,840

- a. This table shows the **maximum** logical volumes available for the given Symmetrix models with the maximum number of disk drives and a homogeneous protection scheme on those disk drives. The logical volume limit is a function of the number of disk drives and the number of disk director processors (CPUs). Symmetrix DMX systems that have fewer disk drives or back-end connections may have fewer logical volumes available on a particular configuration.
- b. RAID (7+1) is not supported on a DMX800 with less than four DAEs.
- c. RAID (7+1) is not supported on a DMX800 with less than four DAEs.

Symmetrix DMX Disk Device Emulations

Symmetrix DMX800 systems support connectivity to mainframe systems and open systems hosts. When the Symmetrix DMX800 systems are configured to open systems hosts such as UNIX, Windows, Linux, or iSeries, the Symmetrix disk devices emulate standard disk devices. When the Symmetrix DMX800 systems are configured to IBM z/OS/PCM system hosts, the Symmetrix disk devices emulate IBM CKD DASD.

The Symmetrix Enterprise Storage Platform (ESP) software enabler, a software option, is required if you plan to store and access mainframe and open systems data on the same Symmetrix system.

Deleting (and Then Adding) Devices Online

Enginuity code 5670 supports removing and then adding devices online, which facilitates the following configuration enhancements:

- ◆ Change Device Emulation Online—Remove a CKD volume and add an FBA volume and vice-versa.

Adding a CKD volume to a Symmetrix system requires a global memory configuration change if this is the first CKD volume being added. However, the system can be configured in advance, thus avoiding an offline global memory reformat.

- ◆ Convert between mirrored and RAID protected volumes.

Open System Disk Emulation

On open systems hosts, the Symmetrix DMX logical disk volumes appear as standard SCSI disk devices with data stored in fixed-block architecture (FBA) format. All host logical volume manager software can be used with Symmetrix disk volumes. The following paragraphs describe the FBA disk format and logical volume structure.

FBA Data and Command Format

Fixed block architecture (FBA) disk devices store data in fixed-sized blocks (typically 512 bytes). A disk device using FBA format is viewed as a large array of blocks. The physical position of each block (cylinder and track) is usually not significant to the host. When requesting disk access for read or write, the host addresses a file by the logical block address (LBA) of the starting block and a count of the total blocks used by the file. DMX800 channel directors control access to global memory and disk devices.

Logical Volume Structure (Open Systems)

The channel directors interact with global memory. Therefore, there is no physical meaning to cylinders, tracks, and heads on the DMX800 logical volume from the front-end point of view. However, DMX800 uses a *logical* geometry definition for its logical volume structure. This geometry is reflected in the mode sense data available to the host.

DMX800 uses the following logical volume structure:

- ◆ Each logical volume has n cylinders
- ◆ Each cylinder has 15 tracks (heads)
- ◆ Each track has 64 blocks of 512 bytes

Therefore, a DMX800 logical volume with n cylinders has a usable block capacity of:

$$n \times 15 \times 64$$

n for each volume is defined during DMX800 configuration

To calculate the size of the logical volume:

$$\text{Number of cylinders} \times \text{heads} \times \text{blocks} \times 512 \\ (n \times 15 \times 64 \times 512)$$

For information on configuring open systems logical volumes, refer to *Open Systems Hypervolumes* on page 4-14.

IBM DASD Disk Emulation

The DMX800 supports IBM CKD DASD Disk Emulation. The Symmetrix DMX800 system appears to mainframe operating systems as a 3990-6 or 2105 controller. The physical disk devices can appear to the mainframe operating system as a mix of multiple 3390 and 3380 device types.

All standard models of the 3380 or 3390 volumes can be emulated up to the physical volume sizes installed. A single Symmetrix system can simultaneously support both 3380 and 3390 device emulations. Table 2-6 lists the Symmetrix characteristics for some standard IBM device emulation modes. Symmetrix systems also support non-standard device sizes, as long as the cylinder count does not exceed that of the equivalent IBM device type.

For more IBM DASD emulation characteristics, refer to Appendix A, *DMX800 System Specifications* on page A-1.

Table 2-6 IBM DASD Emulation Characteristics

IBM DASD Model	MB/Volume	Bytes/Track	Bytes/Cylinder	Cylinders/Volume
3390-27	27,844	56,664	849,960	32,760
3390-9	8,514	56,664	849,960	10,017
3390-3	2,838	56,664	849,960	3,339
3390-2	1,892	56,664	849,960	2,226
3390-1	946	56,664	849,960	1,113
3380K	1,891	47,476	712,140	2,655

Mixed Track Geometries

You can configure a Symmetrix DMX800 system with both 3380 and 3390 track geometries on the same disk device. A single disk device can contain up to 128 logical volumes, depending on the data protection used.

For information on configuring mainframe logical volumes, refer to *Mainframe Systems Hypervolumes* on page 4-18.

IBM/PCM Data and Command Formats

All Symmetrix models support the count-key-data (CKD) and extended count-key-data (ECKD) format used by IBM 3390 and 3380 DASD. For a full description of the channel command words (CCW) supported, refer to the *IBM 3990 Storage Control Reference* or the *IBM 3380 Storage Control Model 13* description. Figure 2-9 on page 2-24 shows the CKD track format emulated for 3390 and 3380 DASD.

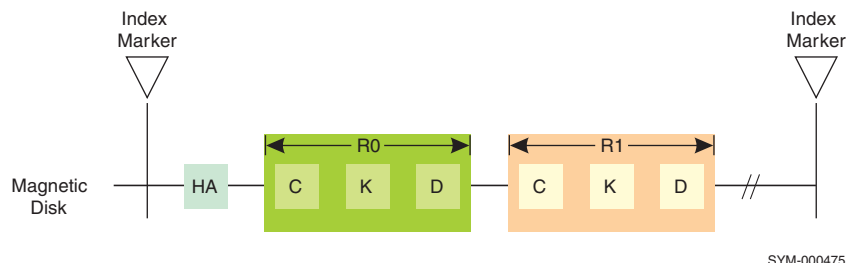


Figure 2-9 Track Format for 3390 and 3380 DASD

Track Format

All tracks are written with formatted records. The start and end of each track are defined by the index marker. Each track has the same basic format as that shown in Figure 2-9.

That is, it has an index marker, home address (HA), record zero (R0), and one or more data records (R1 through Rn). These track formats are discussed in the following sections. Information is recorded on all Symmetrix disk devices in an emulation format chosen at configuration. Each track contains certain *nondata* information (such as the address of the track, the address of each record, the length of each record, and the gaps between each area), and data information.

Index Marker — An index marker indicates the physical beginning and end of each track for each disk device (Figure 2-9).

Home Address (HA) — There is one home address on each track that defines the physical location of the track by specifying the track address (cylinder and head location) and the condition of the track (flag byte). The flag byte indicates whether the track is usable, defective, or an alternate track.

Record Zero-R0 (Track Descriptor Record) — This is the first record after the home address. The count field indicates its physical location (cylinder and head), record number, key length, and data length. In general, the key length is zero bytes and the data length is eight.

Data Records (R1 through Rn) — All remaining records on a track are data records. The Count field indicates the data record's physical location (cylinder and head), record number, key length, and data length. The key is optional and, when used, contains information used by an application. The data area contains the user data. To determine the number of records a track can hold, refer to the *IBM 3390 Direct Access Storage Introduction* or *IBM 3380 Direct Access Storage Introduction* for the equations for calculating this number.

Track Capacity

Track capacity is the maximum capacity achievable when there is one physical data record per track formatted without a key. Because the track can contain multiple data records, additional address markers, count areas, and gaps reduce the number of bytes available for data. The track capacity is the number of bytes left for data records after subtracting the bytes needed for the home address, record zero, address marker, count area, cyclic check (for error correction), and the gaps for one data record.

For 3390 emulations, the track capacity is 56,664 bytes. For 3380 emulations, the track capacity is 47,476 bytes.

Channel Directors, FEBE Boards, and Global Memory Directors

This section describes the Symmetrix DMX800 directors, the FEBE boards that act as front-end and back-end adapters, and the global memory directors. The channel directors and global memory directors manage the storage control functions. Some FEBE board ports are used as a back-end disk adapter and handle the data storage functions.

Channel Director Connectivity

All Symmetrix DMX800 systems provide channel connectivity through combinations of mainframe systems and open systems channel directors. These include:

- ◆ Fibre Channel directors (also used with SRDF)
- ◆ Multiprotocol channel directors (MPCD) available with these channel connections:
 - FICON
 - iSCSI
 - GigE (Gigabit Ethernet) remote directors used for SRDF

The MultiProtocol channel director and support for FICON, iSCSI and GigE channel connections require Engenuity level 5670 and higher.

Symmetrix DMX systems support mixed FICON, Fibre Channel, and iSCSI interfaces when the required Symmetrix ESP software is installed on the Symmetrix system.

The Symmetrix DMX800 systems support open UNIX systems, Linux systems, Windows, and iSeries systems connectivity through Symmetrix Fibre Channel directors or iSCSI directors. iSeries connectivity is only supported through Fibre Channel directors. The Symmetrix DMX800 systems support mainframe connectivity through FICON directors. Symmetrix systems connect directly to host processors through physical channel attachments.

Channel Director Descriptions

DMX800 systems support Fibre Channel, FICON, GigE, and iSCSI channel directors. Symmetrix channel directors are single cards that occupy one slot on the DMX800 cardcage backplane. Two mezzanine cards (FICON, GigE, or iSCSI combinations) can be installed on the MPCD. All channel directors interface to host channels through the FEBE board interface (front-end, back-end). The front boards are FICON or GigE, or a combination of both.

GigE can be configured as SRDF or iSCSI. There can be up to four of the Fibre Channel directors. Four to eight of the Fibre Channel director ports can be used for disk directors. The rest are available for Open systems host attachment. The Symmetrix DMX800 may support two or four Fibre Channel directors. *Only* the two FC director configurations can support MPCD combinations.

Contact your EMC Sales Representative for specific supported configurations.

All channel directors contain four high-performance microprocessors. The channel directors process data from the host and manage access to global memory over a direct matrix (DMX) technology (Figure 2-5 on page 2-11). Each of the FC directors on the Symmetrix DMX800 system supports four internal links to each global memory director (total of eight internal links to global memory from each FC director), and each MPCD supports eight internal links to global memory. DMX technology is used across the Symmetrix system, and it is also designed on to each global memory director.

Fibre Channel Directors (Front-End)

The Fibre Channel director has four FC ANSI compliant, 2 Gb/s (also configurable to 1 Gb/s) Fibre Channel interfaces for connection to host systems and four high-speed paths to the disks. The Fibre Channel director interfaces to the host channels through the Front-End/Back-End (FEBE) board and is available in the following single-mode and multimode configurations:

- ◆ Fibre Channel director FEBE port designations:
 - DMX-FE-8M0S — Eight multimode ports, zero single-mode ports per board
 - DMX-FE-7M1S — Seven multimode ports and one single-mode port per board
 - DMX-FE-6M2S — Six multimode ports and two single-mode ports per board
 - DMX-FE-4M0S — Four multimode ports and zero single-mode ports per board
 - DMX-FE-3M1S — Three multimode ports and one single-mode port per board
- ◆ Two FC director configurations
 - DMX8-3100 — Three FC multimode ports, one FC single-mode port

- DMX8-4000—Four FC multimode ports, zero FC single-mode ports

Front-end and back-end functionality is split on the director for DMX800 only.

Data transfers between host and global memory can execute concurrently across all Fibre Channel ports on a director. Table 2-7 lists the Fibre Channel front-end support capabilities for directors, their Symmetrix devices, and their addressing.

Refer to *Fibre Channel/FEBE Port Designations* on page C-17 for layouts of the possible Fibre Channel FEBE board port designations.

Table 2-7 Symmetrix Devices and Addressing Capabilities for Fibre Channel Directors

Maximum Symmetrix Devices and Device Addresses	Maximum Symmetrix Devices Per Front-End FC Processor ^a
Symmetrix Devices per Processor	2,048
Symmetrix Devices per Fibre Channel Director Board	8,192
Address Range per Port	0 - 4,095 ^b

a. Each Fibre Channel director has two processors.

b. Address range (0x000 - 0xFFFF).

The numerical values for Symmetrix devices stated in Table 2-7 are the maximum allowed according to the architectural limits of the microcode running on the Fibre Channel front-end director. The actual limits allowed for customer environments will be lower, and are dependent on the host type, HBA and driver type/version, and overall system implementation. Also, note that using metadevices will reduce the number of host-visible volumes for a given number of devices (*Symmetrix Devices*) configured to the Fibre Channel front-end director; that is, each member of the metadevice will be counted to the allowed limit of devices configured to a Fibre Channel front-end director.

For information on Fibre Channel host attachments, refer to the *EMC Interoperability Matrices* located on the EMC Powerlink website at: <http://powerlink.emc.com>.

Fibre Channel Directors for SRDF

The Fibre Channel host director can be Enginuity-configured as the link between Symmetrix units in a Symmetrix Remote Data Facility (SRDF) configuration. The Fibre Channel director interfaces to Symmetrix channels through the Fibre Channel interface adapter.

When FC is used for SRDF, one port is allowed to be utilized, so the other port is disabled.

iSeries Fibre Channel Connectivity

The Symmetrix Fibre Channel Adapters, through fibre, connect to the iSeries 270 and 8xx models. When using directly connected fibre devices (point-to-point,) the maximum distance is 500 meters in 1 Gb/s mode.

Multiprotocol Channel Directors

The Symmetrix DMX800 multiprotocol channel director (MPCD), through mezzanine card technology, supports GigE, iSCSI, and FICON protocols on Symmetrix systems running Enginuity code 5670 and higher. In two-port configurations, mezzanine cards are implemented on the main director board. FICON, Gigabit Ethernet, and iSCSI combinations can be configured on the two-port boards.

FICON Function

FICON mezzanine cards on the multiprotocol director support native FICON channel connectivity to IBM z/OS systems and OS/390 systems. FICON supports multiple concurrent I/O connections, channel program multiplexing, and better link utilization than ESCON path switching.

FICON Mezzanine Cards

The FICON mezzanine cards, on the MPCD, support attachment to IBM 9672 models G5, G6, and z/9XX and z/8XX running z/OS, z/VM. The FICON channel director supports FICON *native mode point-to-point* connections and FICON *native mode switched point-to-point* connections. The FICON card provides two single-mode LC 2 Gb/s (or 1 Gb/s) bidirectional (full duplex) transfer rates.

The DMX FICON design auto-detects 2 Gb/s or 1 Gb/s at switch or channel port login time. The setting will stay at 2 Gb/s unless there is a problem or if the other port is running at 1 Gb/s. Only a switch has settings on each port for 2 Gb/s, 1 Gb/s, or auto-negotiate.

The FICON mezzanine cards on the MPCD support the configurations described in Table 2-8 on page 2-30.

Refer to *FICON/FEBE Board Port Designations* on page C-10 for layouts of the possible FICON FEBE board port designations.

Table 2-8 FICON Director Configurations

Maximum LCUs (Logical control Units) per port	16 (0 to F)
Maximum LCUs per Symmetrix system	64
Maximum devices per LCU	256
Maximum logical paths per port	1,024
Maximum logical paths per LCU per port	64
Maximum DMX800 channel addresses per port	4,096
Maximum DMX channel addresses per Symmetrix system composed of up to 8,000 base devices (physical devices) with up to 7 aliases (logical devices)	16,384
Maximum concurrent I/O host connections per DMX800 system with two FC directors and two FICON MPCD	4

The following are the FICON director FEBE port combinations:

- ◆ Two FICON director FEBE port combinations:
 - DMX8-4002—Four FC multimode ports and two single-mode ports per board
 - DMX8-3102—Three FC multimode ports, one FC single-mode port, two FICON single-mode ports per board
- ◆ One FICON/one GigE director FEBE port combinations:
 - DMX8-3111—Three FC multimode ports, one FC single-mode port, one GigE multimode port, and one FICON single-mode port per board
 - DMX8-4011—Four FC multimode ports, one GigE multimode port, and one FICON single-mode port

For information on FICON Cascading and Open Systems Intermix features, refer to *Additional FICON Features* on page 2-40.

GigE Remote Function

GigE remote mezzanine cards on the MPCD enables remote director functionality based upon Gigabit Ethernet technology that enable direct Symmetrix-to-IP network attachment and eliminate the need for expensive media converter appliances. GigE support for SRDF on Symmetrix DMX systems increases the options for Symmetrix-to-Symmetrix connectivity, and it allows Symmetrix to connect to your existing Ethernet infrastructure and directly access high-speed data transmission conduits by way of Internet Protocol (IP).

Gigabit Ethernet Remote Directors

The Symmetrix GigE remote mezzanine cards on the MPCD provide two 1 Gb/s Ethernet ports and connects by way of LC connectors. The GigE remote mezzanine card supports TCP/IP protocols and function layers. It is fully compatible with the SRDF protocols and other Symmetrix DMX directors, drives, software, and protection schemes.

The GigE remote mezzanine card enables direct native IP connectivity for SRDF deployment. It supports SRDF traffic to one or more remote Symmetrix systems that also have GigE remote directors installed. The Symmetrix DMX800 systems may support up to two, two-port GigE remote mezzanine cards per MPCD.

The following are the GigE director FEBE port combinations:

- ◆ Two GigE director FEBE port combinations:
 - DMX8-4020—Four FC multimode ports, two GigE multimode ports per board
 - DMX8-3120—Three FC multimode ports, one FC single-mode port, two GigE multimode ports per board
- ◆ One GigE/one FICON director FEBE port combinations:
 - DMX8-3111—Three FC multimode ports, one FC single-mode port, one GigE multimode port, and one FICON single-mode port per board
 - DMX8-4011—Four FC multimode ports, one GigE multimode port, and one FICON single-mode port

Refer to *GigE/iSCSI FEBE Port Designations* on page C-19 for layouts of the possible GigE FEBE board port designations.

The iSCSI/FEBE board port designations are the same as the GigE/FEBE board port designations.

iSCSI Functions

The Symmetrix DMX multiprotocol director through mezzanine card technology supports iSCSI channel connectivity by way of Gigabit Ethernet (GigE) protocol for the Symmetrix DMX systems running Engenuity code 5670 and higher.

The GigE 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.

iSCSI Channel Directors

Two FICON mezzanine cards or a mix of one FICON and one GigE/iSCSI mezzanine card can be installed on the MPCD.

The Symmetrix GigE director supports iSCSI connectivity and provides two, 1 Gb/s Ethernet ports and connects by way of LC connectors. It is compatible with Symmetrix DMX and Symmetrix 8000 systems.

The iSCSI director supports TCP/IP protocols and function layers. It is fully compatible with other Symmetrix DMX directors, drives, software, and protection schemes. The MPCD can be configured with two iSCSI ports or one iSCSI port, and one GigE or one FICON port.

Refer to *GigE/iSCSI FEBE Port Designations* on page C-19 for layouts of the possible GigE/iSCSI FEBE board port designations.

Front-End/Back-End Boards

The physical connection to a DMX800 channel interface occurs at the connectors on the FEBE boards. The FEBE board provides an interface between the director and host channels.

In a two-director configuration, FEBE 0 handles all I/O to Dir 1. FEBE 1 handles all I/O to Dir 16. Refer to Figure 2-4 on page 2-10.

In a four-director configuration, FEBE 0 handles all I/O to Dir 1 and Dir 2; FEBE 1 handles all I/O to Dir 15 and Dir 16. Refer to Figure 2-5 on page 2-11.

FEBE Board Features

Each FEBE board provides the following:

- ◆ Up to 16 total FEBE ports per board
- ◆ Four FE and four BE ports per board – two FC director configuration
- ◆ Eight FE and eight BE ports per board – four FC director configuration

These adapters provide the connectivity between the host channels and the Fibre Channel directors (FC-0 layer of the Fibre Channel standard).

Each FEBE board supports a maximum of eight back-end Fibre Channel connections with a maximum of 16 connections per board.

Additionally, the FEBE board contains the following features:

- ◆ Provides system environmental control
- ◆ Two RS-232 ports to SPS and directors
- ◆ One 100BASE-TX Ethernet port to server and directors

FEBE Board Configuration Options

There is a maximum of eight front-end (FE) and eight back-end (BE) ports in a DMX800 with two director (16 FE and 16 BE in a DMX800 with four directors). The number of back-end ports required depends upon the number of DAEs in the DMX800 system. Each DAE requires two back-end ports.

Example

Two DAEs would require four back-end ports with four to 15 drives per DAE. Eight DAEs would require 16 back-end ports (maximum allowable per DMX800 configured with four FC directors) with four to 15 drives per DAE.

SFP Transceivers

The Small Form Pluggable (SFP) optical transceivers provide the interface to the host channels. The FEBE uses an SFP optical transceiver module on each host port for the Fibre Channel interconnect. The SFP multimode transceiver is used for applications where the link distances are less than 300 meters at 2 Gb/s. The SFP single-mode transceiver is capable of longer link distances and supports wavelengths from 1285 to 1335 nm.

Table 2-9 describes the Fibre Channel back-end configurations.

Table 2-9 Fibre Channel Disk Director (Back-End) Configurations

	Two FC Configuration Symmetrix DMX800 15-Drive Loops	Four FC Configuration Symmetrix DMX800 15-Drive Loops
Number of FC directors	2	4
Maximum number of disk drives	60	120
Number of loops per FC director pair	8	8
Range of FC disk drives Per loop	4	4
Maximum Number of FC disk drives per disk director Pair	60	60

For information on maximum logical volumes supported on each Symmetrix DMX physical disk device and the maximum logical volumes supported for each Symmetrix DMX model, refer to *Symmetrix DMX800 System Logical Volume Capacities* on page 2-19.

Fibre Channel Directors for SRDF

The Fibre Channel host director can be Engenuity-configured as the link between Symmetrix units in a Symmetrix Remote Data Facility (SRDF) configuration. The Fibre Channel director interfaces to the remote Symmetrix through the Fibre Channel connections on the FEBE board.

Global Memory Directors

The Symmetrix DMX global memory director technology is one of the most crucial components of a Symmetrix system. All read and write operations transfer data to or from global memory. Any transfers between the host processor, channel directors, and global memory directors are achieved at much greater electronic speeds than transfers involving disks.

Each Symmetrix DMX800 global memory director accommodates four separately addressable, simultaneously accessible memory regions, which greatly reduces the probability of contention for global memory access. Each global memory director is directly connected to each channel director and FEBE board through the direct matrix.

The DMX architecture ensures highest performance due to the following:

- ◆ Requests for global memory are expedited to reduce locking
- ◆ Requests are intelligently arbitrated to optimize available resource usage

Global Memory Director Configuration

The Symmetrix DMX800 system supports two slots in the backplane dedicated to global memory directors and a maximum of 64 GB of global memory. Individual global memory directors are available in 2 GB, 4 GB, 8 GB, 16 GB, and 32 GB sizes. When configuring global memory for the Symmetrix DMX800 systems, follow these guidelines:

- ◆ Two global memory director boards (M0, M1) are required for the DMX800 system configuration.
- ◆ Global memory director configurations require boards of equal size.

Symmetrix DMX800 Power Subsystem

The Symmetrix DMX800 power subsystem includes the following features:

- ◆ SPE power supplies
- ◆ DAE power supplies
- ◆ Standby power supplies (SPS)

Storage Processing Enclosure Power Supplies

The two 1080W redundant power supplies (one power supply is sufficient to power an SPE) located at the front of the Storage Processing Enclosure (Figure 2-10) supply the power to the SPE. Vital product data (VPD) in the power supplies is monitored by the following directors:

- ◆ Power Supply 1 (PS1)—Dir 1
- ◆ Power Supply 2 (PS2)—Dir 16

Each power supply has two LEDs on its front panel (Figure 2-10):

- ◆ Power LED—Illuminates when power is on to the power supply
- ◆ Fault LED—Illuminates when there is a power fault

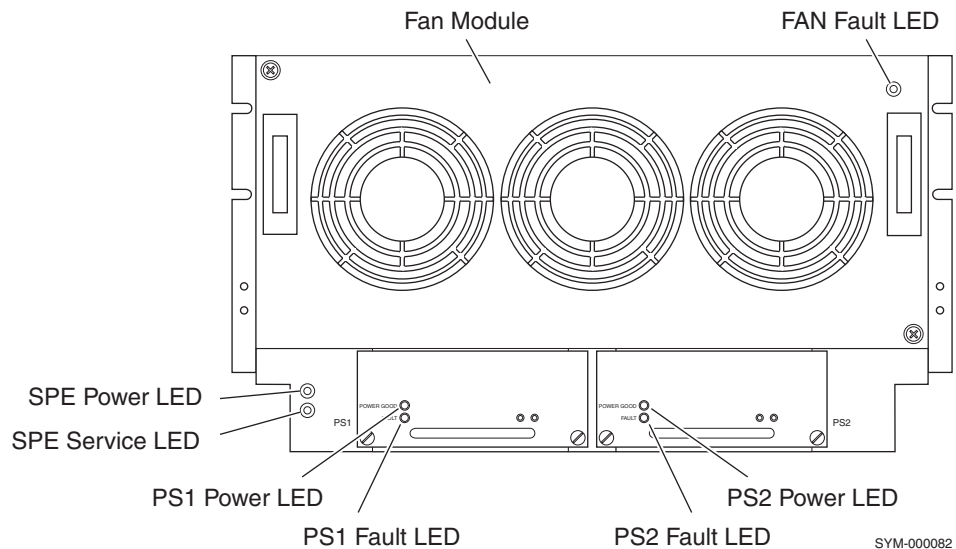


Figure 2-10 Storage Processing Enclosure Power Supply

Disk Array Enclosure Power Supply

The back of each DAE contains two 12 V power supplies (power and cooling modules) and two LCCs (Figure 2-11).

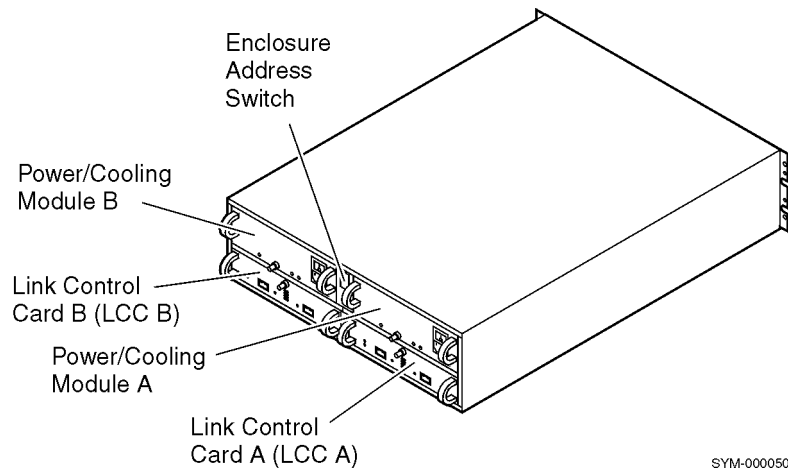


Figure 2-11 Disk Array Enclosure (Rear View)

The two 12 V power supplies (PS A, PS B) provide power and cooling to each DAE.

Each link control card (LCC A, LCC B) supports and controls one Fibre Channel loop and monitors the DAE environment.

The rear of the DAE contains connections and LEDs for fault indication of the DAE, and the AC power inlets for the power supplies and LCC.

Standby Power Supply

There are two SPS units in an SPS tray.

The 1080W SPSs (Figure 2-12) provide the following functions:

- ◆ Two SPS modules (in a 1U unit) provide redundant backup for the SPE
- ◆ Two SPS modules provide redundant backup for every two DAEs

In a DMX800 configuration of eight DAEs, there are five SPS assemblies: four to support the DAEs and one to support the SPE.

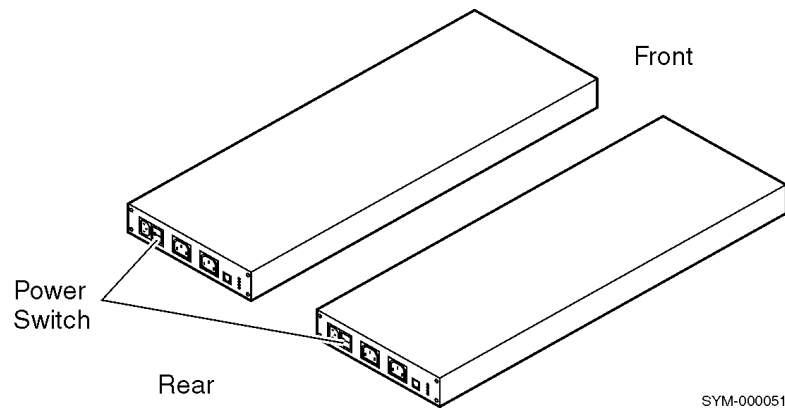


Figure 2-12 Standby Power Supply

Channel Attachments

Symmetrix systems can attach to FICON channels, or Fibre Channels, GigE, iSCSI, or a mix of channel types. The physical connection to a Symmetrix channel interface occurs at the back-end ports of the FEBE board.

The Symmetrix Enterprise Storage Platform (ESP) option is required when both mainframe hosts (FICON channels) and open systems hosts (Fibre Channels) connect to the same Symmetrix system. Consult your EMC Sales Representative for the most current list of supported hosts, models, operating systems, and EMC open systems host support policies, or refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

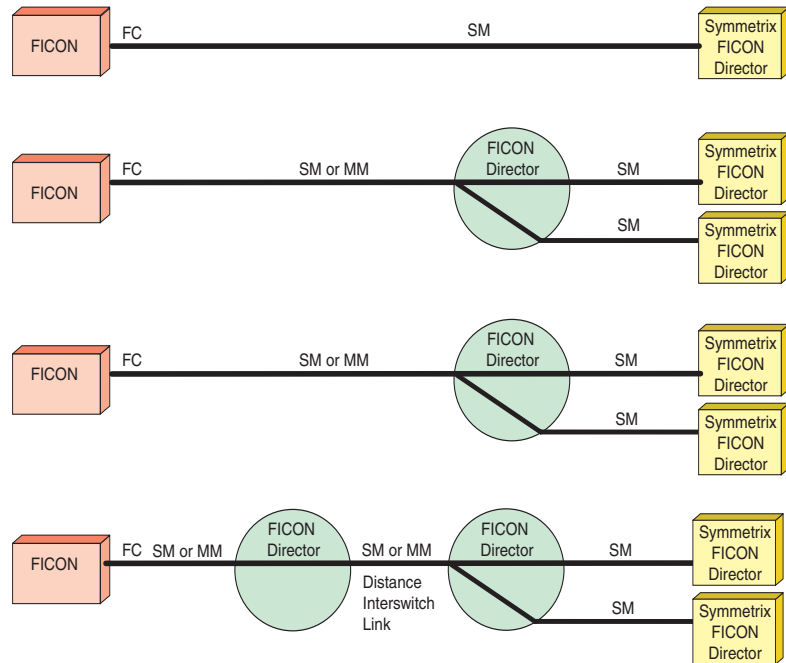
FICON Channel Interface Connections

The FEBE board provides an interface between the FICON director and mainframe host channels. The FEBE boards are located in the DMX800 card cage. These adapters provide the connectivity between the host channels and the FICON channel directors. The two-port, two-processor FICON director provides the capability for 32 concurrent operations per port per LPAR or 128 concurrent operations per port for all LPARs attached through a FICON director to the same port.

FICON Channel directors use fibre-optic cables. The current FICON implementation supports data transfer rates up to 2 Gb/s. There are two types of fiber-optic cables: multimode and single-mode.

Symmetrix DMX800 systems directly connect to FICON single-mode cables. For information on FICON multimode configuration support, contact your EMC Sales Representative. For information on additional configuration options, refer to *Additional FICON Features* on page 2-40.

In the FICON environment, a link connects a host FICON channel with a Symmetrix FICON channel interface. This link can be a direct connection between the processor or LPAR and the FICON channel interface. The link can also have a FICON director that branches off to additional single-mode or multimode links with connections to Symmetrix FICON channel directors. Figure 2-13 on page 2-39 illustrates several types of FICON channel attachments.



SYM-000347

Figure 2-13 FICON Channel Attachments

Table 2-10 describes the maximum distances supported by FICON point-to-point single-mode cables.

Table 2-10 Supported Symmetrix Point-to-Point FICON Cable Distances

Cable Type	Cable Description	Maximum Supported Distance
Single-mode (SM Longwave Laser (1,310 nano meters))	9 micron, 1 Gb/s	10 km (6.2 miles) for each link - Up to 20 km (12.43 miles) for each link with IBM RPQ
	9 micron, 2 Gb/s	10 km (6.2 miles) for each link - Up to 12 km (7.45 miles) for each link with IBM RPQ

Additional FICON Features

All DMX FICON models now support the use of Cascading and Open Systems Intermix. Cascading can be used to reduce the number of FICON adapters and the amount of intersite cabling required by making use of switch to switch communication.

Intermix allows FICON zones to be added to existing Open Systems switches within a site and between sites. These two features help reduce the overall costs while providing greater FICON connectivity, backup, and recovery.

For specific configurations and information, contact your EMC Sales Representative.

Cascading

Cascading provides greatly enhanced FICON connectivity within local and remote sites through the use of switch-to-switch extensions of the CPU to the DMX FICON network. These cascaded switches communicate over long distances using a small number of high speed lines called ISL's (Inter Switch Links). Up to a maximum of two switches max may be connected together within a path between the CPU and the DMX. Same switch vendors are required for a cascaded configuration. The EMC and IBM branded McData and Inrange switches are supported in pairs.

To support Cascading, each vendor requires specific models, hardware and software features, configuration settings, and restrictions. Specific IBM CPU models, MVS release levels, channel hardware, and microcode levels are also required.

Open Systems Intermix

Open Systems Intermix allows separate FICON zones to be defined within new or existing open systems switches. These switches can also be cascaded to further enhance connectivity and remote backup and recovery. The EMC and IBM branded McData and InRange switches are supported. To support Open Systems Intermix, each vendor requires specific models, hardware and software features, configuration settings, and restrictions. Specific IBM CPU models, MVS release levels, channel hardware, and microcode levels are also required.

Open Systems Fibre Channel Interface Connections

The FEBE board provides an interface between the Fibre Channel director and open systems host channels. The FEBE boards are located in the DMX800 card cage. These adapters provide the connectivity between the host channels and the Fibre Channel directors (FC-0 layer of the Fibre Channel standard). The eight-port, four-processor Fibre Channel director provides the capability for four concurrent operations through its four physical interfaces for communicating with the host systems.

For more information on the Fibre Channel director, refer to *Fibre Channel Directors (Front-End)* on page 2-27.

Fibre Channel directors use fiber-optic cables. The channels use Fibre Channel arbitrated loop or switched fabric links. Each link has two physical fibers for transporting data: one for inbound signals and one for outbound signals. The current Fibre Channel implementation supports data transfer rates up to 2 Gb/s.

Table 2-11 Symmetrix Fibre Channel Cable Distances ^a

Cable Type	Cable Description	Maximum Distance
Single-mode	9 micron, 1 Gb/s	10 km (6.2 miles) per cable segment
	9 micron, 2 Gb/s	10 km (6.2 miles) per cable segment
Multimode	50 micron, 1 Gb/s	500 m (1,640 ft) per cable segment
	50 micron, 2 Gb/s	250 m (820 ft) per cable segment
	62.5 micron, 1 Gb/s	300 m (984 ft) per cable segment
	62.5 micron, 2 Gb/s	150 m (492 ft) per cable segment

a. Distances are point-to-point.

For more information on Fibre Channel attachments and cables, refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

Symmetrix DMX800 Input/Output Operations

This chapter describes input and output operations between the Symmetrix DM800 disk subsystem and a host system.

- ◆ Symmetrix DMX Disk Subsystem Operation3-2
- ◆ Elements of a Symmetrix I/O Operation3-6
- ◆ Functional Operation.....3-14
- ◆ Read/Write Operations.....3-24
- ◆ I/O Performance Enhancements3-26

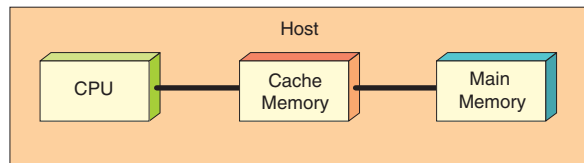
Symmetrix DMX Disk Subsystem Operation

Intelligent global memory configurations allow Symmetrix DMX800 disk subsystems to transfer data at electronic memory speeds that are much faster than physical disk speeds. Symmetrix systems are based on the principle that the working set of data at any given time is relatively small when compared to the total subsystem storage capacity. When this working set of data is in global memory, there is a significant improvement in performance. The performance improvement achieved depends on both of the following principles:

Locality of Reference—If a given piece of information is used, there is a high probability that a nearby piece of information will be used shortly thereafter.

Data Re-Use—If a given piece of information is used, there is a high probability that it will be re-used shortly thereafter.

These global memory principles have been used for years on host systems (CPU and storage devices). Figure 3-1 illustrates this type of host global memory use. The global memory used in this manner is often a high-speed, high-cost storage unit that functions as an intermediary between the CPU and main storage.

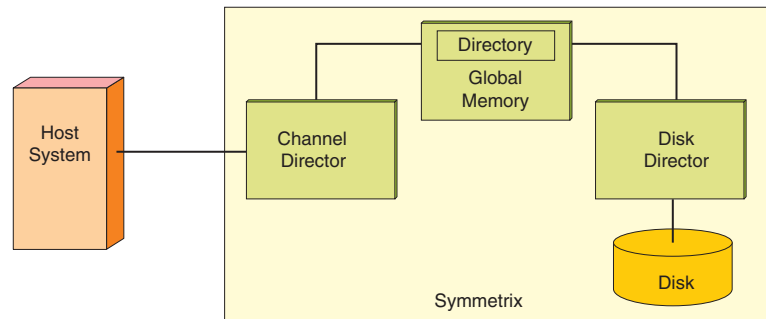


SYM-000420

Figure 3-1 Host Global Memory Use

Symmetrix Global Memory Management

Symmetrix systems use the same global memory principle as host systems, but with enhanced caching techniques. Figure 3-2 illustrates global memory use in Symmetrix systems.



SYM-000450

Figure 3-2 Symmetrix Global Memory Management and Data Flow

In Symmetrix DMX800 systems, the FEBE board's function as a back-end adapter allows it to share global memory. Symmetrix channel directors attach to the CPU channels as well as to global memory. The FEBE ports attach to global memory as well as the disk drives. The Symmetrix directors perform the following functions:

- ◆ The channel director handles I/O requests from the host. It accesses the directory in global memory (Figure 3-2) to determine if the request can be satisfied within the global memory. The directory contains information on each global memory page and blocks within each page.
- ◆ The directors manage global memory using various methods that generally result in the implementation of a Least Recently Used (LRU) algorithm. This usage model ensures that only pages of data that have been recently accessed remain in global memory.
- ◆ Using the prefetch algorithm, the disk director dynamically detects sequential data access patterns to the disk devices. The directors improve the hit ratio of these accesses by promoting blocks from the disk devices to global memory slots before that data has been requested.
- ◆ The disk director manages access to the disk drives. When a read miss occurs, the disk director also stages tracks into global memory. It also performs a background operation that destages written-to blocks to disk.

Least Recently Used (LRU) Algorithm

Figure 3-3 illustrates the data flow with the LRU algorithm. Each time a read hit or write hit occurs, the Symmetrix system marks that global memory slot as most recently used and promotes it to the top of the LRU list. For each write, a written-to flag is set on the initial write to each global memory block and is cleared when the global memory block is destaged. The LRU global memory slot appears at the bottom of the LRU list. Symmetrix global memory can be subdivided into multiple LRU lists. There are sixteen independent LRUs in a Symmetrix system.

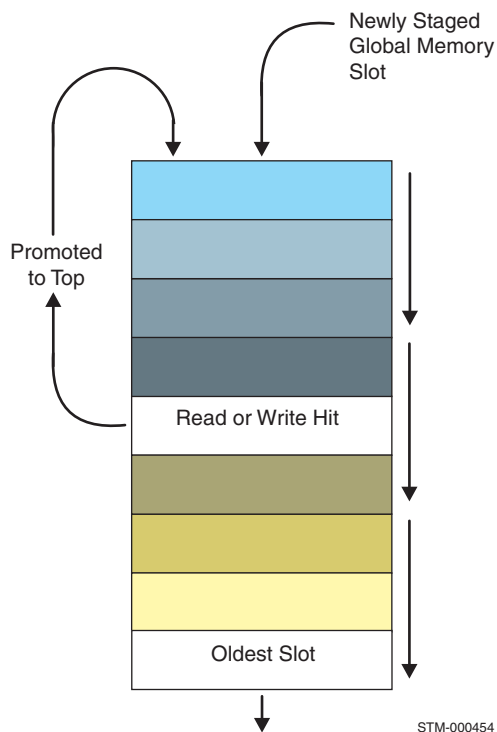


Figure 3-3 LRU and Age Link Chain Data Flow

Global Memory Quality of Service (QoS)

Global memory quality of service (QoS) allows the user to define (and future assign) segments of global memory to be managed by a separate LRU list, to one or more designated logical volumes. This way the user can guarantee that a certain application has a certain number of global memory slots—regardless of the activity levels of the other applications in the system, and, thus, guarantee a certain level of hit ratios and performance.

Prefetch Algorithm

Symmetrix systems continually monitor I/O activity and look for access patterns. When the second sequential I/O to a track occurs, the sequential prefetch process is invoked and the next track of data is read into global memory. The intent of this process is to avoid a read miss. When the host processor returns to a random I/O pattern, the Symmetrix system discontinues the sequential process.

The prefetch algorithm reduces response time and improves the utilization of the disks. The prefetch algorithm maintains, per each logical volume, an array of statistics and parameters based on the latest sequential patterns observed on the logical volume. Prefetch dynamically adjusts, based on workload demand across all resources in the back end of the Symmetrix system. This algorithm also ensures that global memory resources are never overly consumed in order to maintain optimal performance.

Elements of a Symmetrix I/O Operation

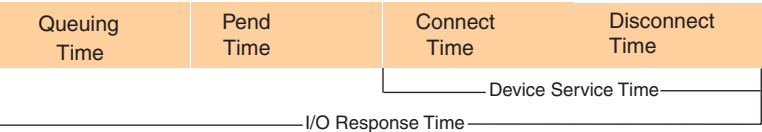
All I/O operations require a certain response time. An I/O request begins when the application issues an I/O command and ends when the data transfer completes. The time interval from I/O request to transfer completion is the I/O response time.

This section describes:

- ◆ I/O Response Time: Mainframe Environment
- ◆ I/O Response Time: Open Systems Environment
- ◆ Symmetrix I/O Operations
- ◆ Read Operations
- ◆ Write Operations
- ◆ Write Destaging Operation

I/O Response Time: Mainframe Environment

In the mainframe environment, I/O response time can be divided into a queuing time, a pend time, a connect time, and a disconnect time, as shown in Figure 3-4.



SYM-000452

Figure 3-4 I/O Response Time (Mainframe Environment)

The **Queuing Time** is the I/O Supervisor (IOS) queue for the next event.

The **Pend Time** consists of:

- ◆ Control Unit Busy (CUB)
- ◆ Device Busy (DB)
- ◆ Director Port Busy

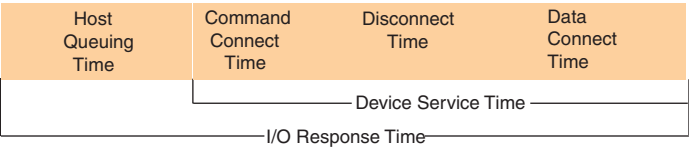
The **Connect Time** is the length of time the channel processes commands and transfers data.

The **Disconnect Time** is:

- ◆ The length of time it takes to retrieve data from the physical disk (device seek and latency)
- ◆ The length of time it takes to reconnect to the host
- ◆ SRDF write overhead (protocol, line latency, and so on)

I/O Response Time: Open Systems Environment

In the open systems environment, I/O response time can be divided into a host queuing time, a command connect time, a disconnect time, and a data connect time, as shown in Figure 3-5.



SYM-000453

Figure 3-5 I/O Response Time (Open Systems Environment)

The **Host Queuing Time** is the length of time the request is in the host queue before it is dispatched on the Fibre Channel path.

The **Command Connect Time** is the length of time the channel is transferring a Fibre Channel command.

The **Disconnect Time** is the length of time involving device seek and latency. During this time, the Fibre Channel path can be used by other devices.

In case of a global memory hit in an I/O request, the Disconnect Time requirement is eliminated.

The **Data Connect Time** is the length of time the channel is transferring data.

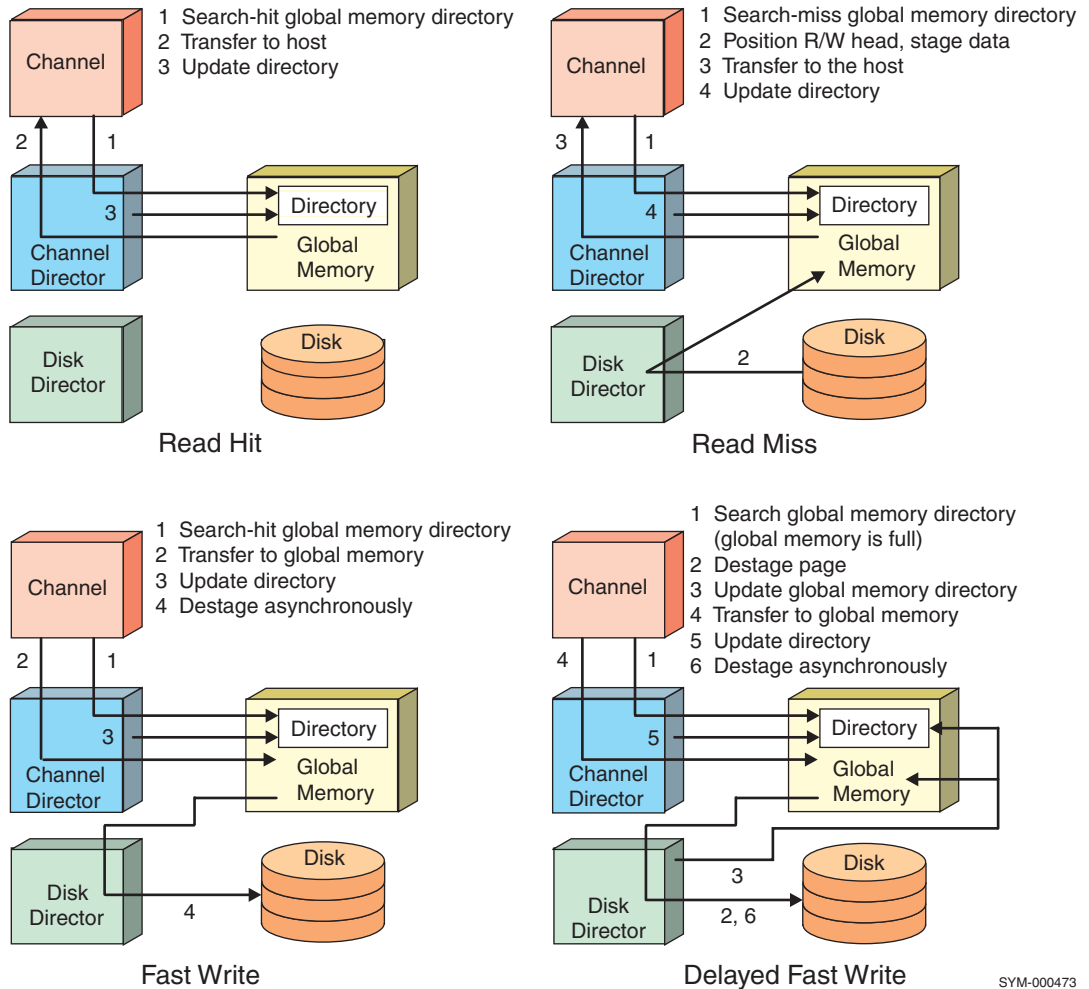
Symmetrix I/O Operations

There are four basic types of Symmetrix I/O operations (Figure 3-6 on page 3-8):

- ◆ Read hit
- ◆ Read miss
- ◆ Fast write
- ◆ Delayed fast write

The Symmetrix system performs read operations from global memory and always the global memory's write operations. This global memory operation is transparent to the host operating system. A read operation causes the channel director to scan the global memory directory for the requested data. If the requested data is in global memory, the channel director transfers this data immediately to the channel with a channel end and device end (or a SCSI good ending status).

If the requested data is not in global memory, the disk director transfers the data from the disk device to the global memory, and the channel director transfers the requested data from the global memory to the channel.



SYM-000473

Figure 3-6 Symmetrix I/O Operations

Read Operations

There are two types of read operations: Read hit and read miss. Figure 3-7 illustrates the data flow for read operations.

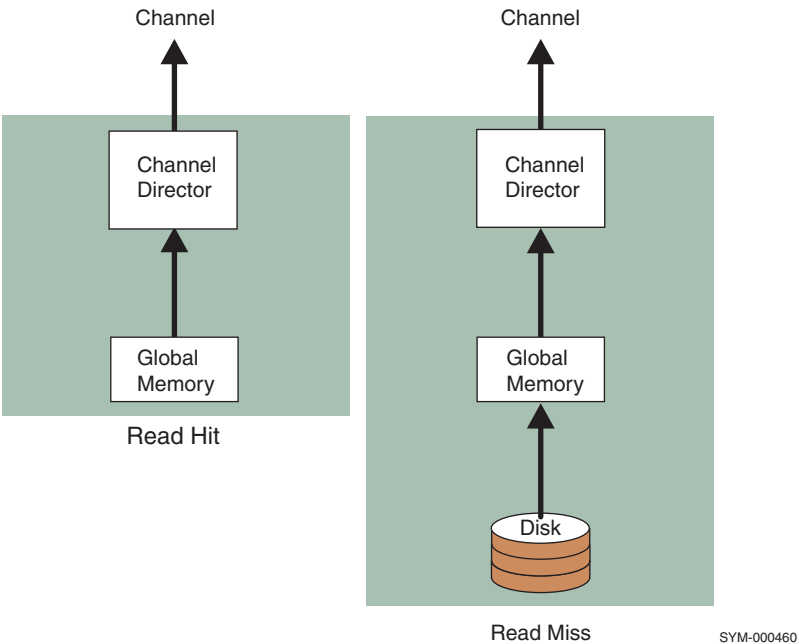


Figure 3-7 Read Operations

Read Hit In a read hit operation (Figure 3-7), the requested data resides in global memory. The channel director transfers the requested data through the channel interface to the host and updates the global memory directory. Since the data is in global memory, there are no mechanical delays due to seek, latency, and Rotational Position Sensing (RPS) miss (Figure 3-8).

For more information on RPS, refer to *Fibre Channel Disk Devices* on page 2-18.

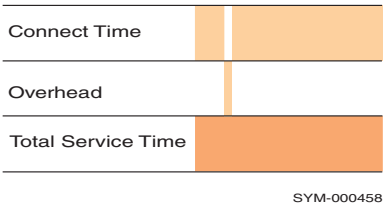
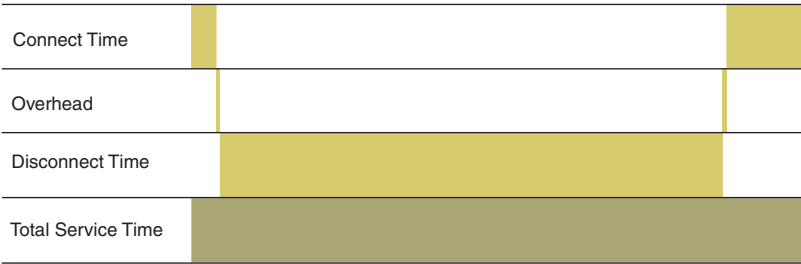


Figure 3-8 Read Hit

Read Miss In a read miss operation (Figure 3-7 on page 3-9), the requested data is not in global memory and must be retrieved from a disk device. While the channel director creates space in the global memory, the disk director reads the data from the disk device. The disk director stores the data in global memory and updates the directory table. The channel director then reconnects with the host and transfers the data. If the requested data is in the process of being prefetched (sequential read ahead), the miss is considered to be a short miss. If the requested data is not in the process of being read into global memory, the disk director requests the data from the drive. This miss is considered to be a long miss. Because the data is not in global memory, the Symmetrix system must search for the data on disk and then transfer it to the channel. This adds seek and latency times to the operation (Figure 3-9). During the disconnect time, other commands can be executed on other devices on the bus, or commands can queue to the same device.



SYM-000459

Figure 3-9 Read Miss

Write Operations

Symmetrix systems use large global memory configurations, and 100 percent global memory fast writes to ensure the highest possible performance when writing data. Symmetrix systems write operations occur as either fast write or delayed fast write operations (Figure 3-6 on page 3-8).

Fast Write A fast write occurs when the percentage of modified data in global memory is less than the fast write threshold. On a host write command, the channel director places the incoming block(s) directly into global memory. For fast write operations (Figure 3-10), the channel director stores the data in global memory and sends a channel end and device end (or a SCSI good ending status) to the host computer. The disk director then asynchronously destages the data from global memory to the disk device.

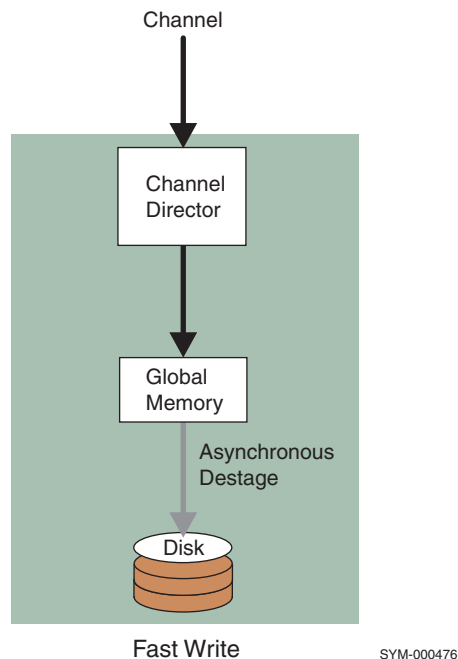
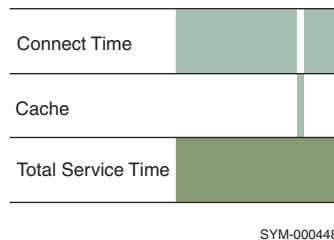


Figure 3-10 Write Operations

Fast write also allows Symmetrix systems to accommodate bursts of writes at a speed above and beyond the speed that the hard drives allow. The negative effects of *bursty* writes are minimized through this algorithm.

Because Symmetrix systems write the data directly to global memory and not to disk, there are no mechanical delays due to seek, latency, and RPS miss (Figure 3-11).



SYM-000448

Figure 3-11 Fast Write

Delayed Fast Write

A delayed fast write occurs only when the fast write threshold has been exceeded. That is, the percentage of global memory containing modified data is higher than the fast write threshold. If this situation occurs, the Symmetrix system disconnects the channel directors from the channels.

The disk directors then destage the least recently used(LRU) data to disk. When sufficient global memory space is available, the channel directors reconnect to their channels and process the host I/O request as a fast write (Figure 3-12). The Symmetrix system continues to process read operations during delayed fast writes. With sufficient global memory present, this type of global memory operation rarely occurs.

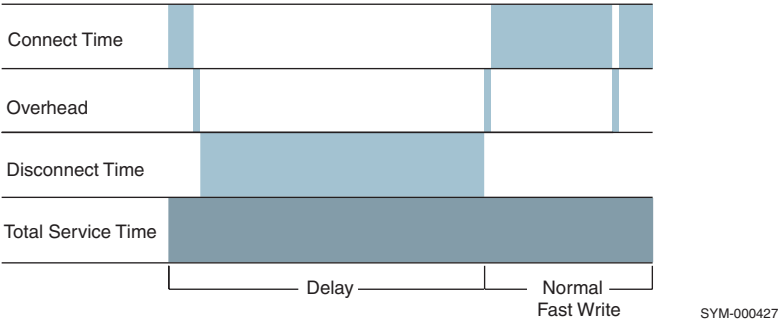


Figure 3-12 Delayed Fast Write

Write Destaging Operation

In addition to the four types of I/O operations explained earlier, Symmetrix systems perform a background operation that destages blocks back to disk. This allows any written-to or changed data to be maintained in two locations: in global memory for performance in the occurrence of re-use of that data and on disk to maintain the highest levels of data integrity. Any pending writes are assured of arrival to the intended disk, even in the event of power failure. Figure 3-13 illustrates this destaging operation.

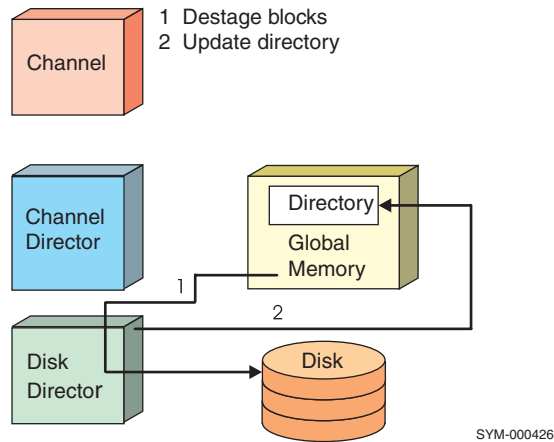


Figure 3-13 Destaging Operation

For each logical volume, the Symmetrix system maintains a special data structure that points to the data that needs to be destaged in the global memory Write Pending Indicators (WPI) already discussed above. This dynamically adjusting algorithm saves disk seek and latency time by destaging data in groups of up to four tracks concurrently per logical volume.

Functional Operation

The Symmetrix DMX800 is a rackmount system containing channel directors, global memory directors, disk drives, and associated software. A DMX800 two FC director configuration contains eight front-end ports for host connections and eight back-end ports for data transfer (16 ports each for the four FC director configuration).

Refer to Figure 3-14 on page 3-15, Figure 3-15 on page 3-16, and Figure 3-16 on page 3-18.

Overview

The DMX800 two FC director configuration contains two FC channel directors and up to 64 GB global memory residing on two global memory director boards. The DMX four FC director configuration contains four FC directors and up to 64 GB global memory on two global memory director boards. Each channel director has eight direct paths to each global memory director. Each of the two FEBE boards has eight front-end ports and eight back-end ports for connection to host systems and the Fibre Channel directors.

The channel directors are paired for failover protection: Dir 1 and Dir 16 for the two-director configuration and Dir 1 and Dir 16; Dir 2 and Dir 15 for the four-director configuration.

Refer to *Fibre Channel Front-End Redundancy* on page 5-7 for information on failover protection.

In the DMX four FC director configuration, FEBE 0 handles all I/O to Dir 1 and Dir 2; FEBE 1 handles all I/O to Dir 15 and Dir 16.

The Link Control Cards (LCC) — Two cards per each DAE supports and controls one Fibre Channel loop (up to 15 disk drives in one DAE) and monitors the DAE environment. The LCC sends and receives signals from the FEBE board to the FC directors.

A server within the DMX800 system monitors status and sends a call-home message if an error occurs.

The Ethernet interface provides connection into an Ethernet switch fabric between the FEBEs and directors in the DMX800 system. It is the primary communication link between the server and the DMX800. Communication with the DMX800 is initiated by the server. The server also contains an Ethernet link to the customer's network.

The following sections describe the DMX800 functionality.

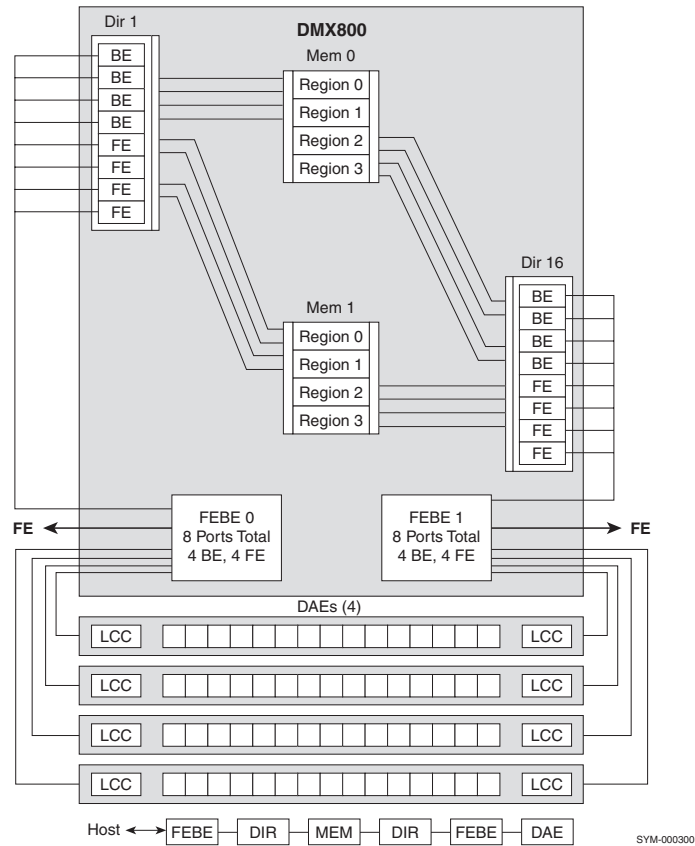


Figure 3-14 Functional Block Diagram of a DMX800 Configured with Two FC Directors

Up to eight DAEs can be installed in the DMX800 four FC director system.

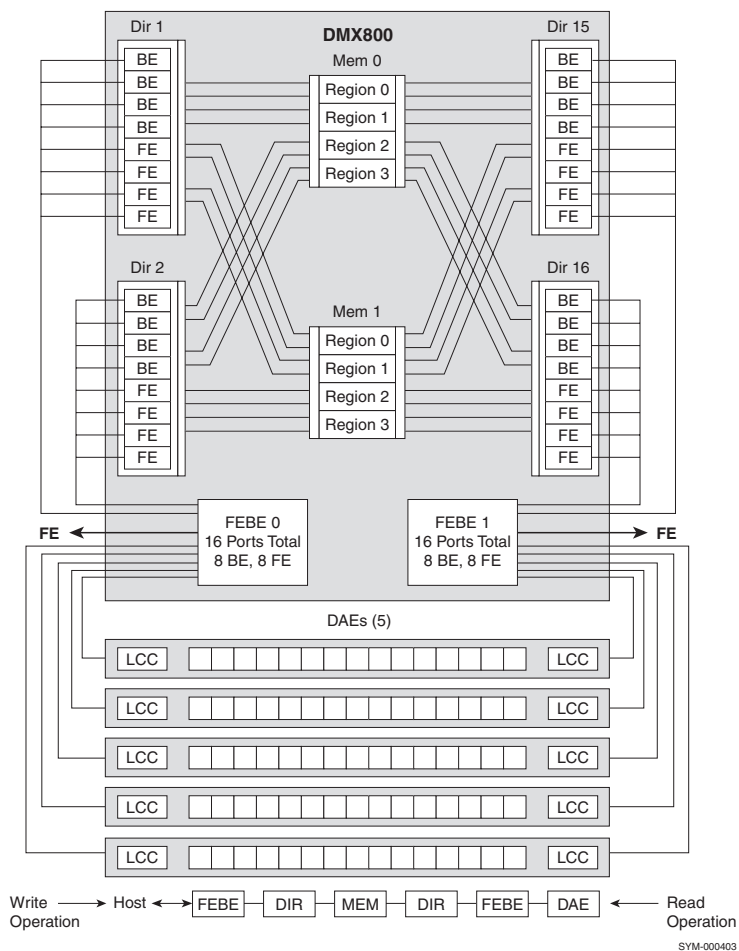


Figure 3-15 Functional Block Diagram of a DMX800 Configured with Four FC Directors

MPCD Functional Description

The Symmetrix DMX multiprotocol channel directors (MPCD), through mezzanine card technology, supports GigE, iSCSI, and FICON protocols on Symmetrix systems running Engenuity code 5670 and higher. Two-port configurations (all DMX800 configurations), and two mezzanine cards are located on the director. FICON, Gigabit Ethernet (GigE), and iSCSI combinations can be configured on the MPCD. Refer to Figure 3-16 on page 3-18.

The iSCSI capabilities of the MPCD only support host connections and cannot be used for Symmetrix-to-Symmetrix SRDF connectivity.

Mezzanine card configuration is one port per mezzanine card, two mezzanine cards per MPCD. The type of mezzanine card (FICON, GigE, iSCSI, or combination of two) determines the host protocol support on the channel. The director can have two channels; each mezzanine card provides one connection per Fibre Channel loop. Each MPCD has eight paths to global memory.

In the DMX800 two FC director card cage, both the blank panel and the MMB are removed to accommodate the MPCD which then become Dir 2 and Dir 15 (Figure 3-16 on page 3-18).

Refer to *Multiprotocol Channel Directors* on page 2-29 for additional information. Refer to *Mainframe/Open Systems Installations* on page C-10 and *GigE Remote and iSCSI Director Installations* on page C-22 for MPCD/FEBE port designations for the MPCD combinations.

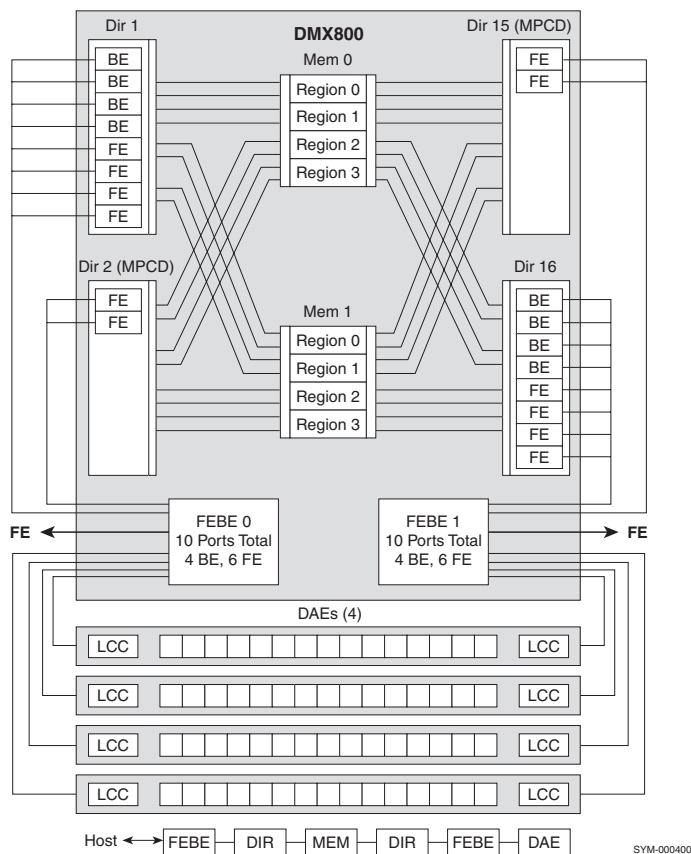
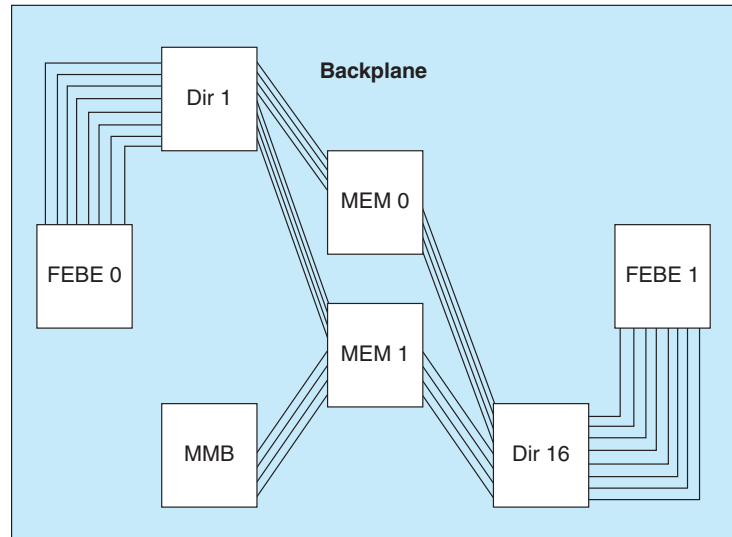


Figure 3-16 Functional Block Diagram of a DMX800 Configured with Two FC Directors and Two MPCD

DMX800 Backplane

On the DMX800 two FC director configuration, the two channel directors, the two FEBEs, the two global memory directors, and the Message Matrix Board (MMB) plug into the eight-slot DMX800 backplane (Figure 3-17).

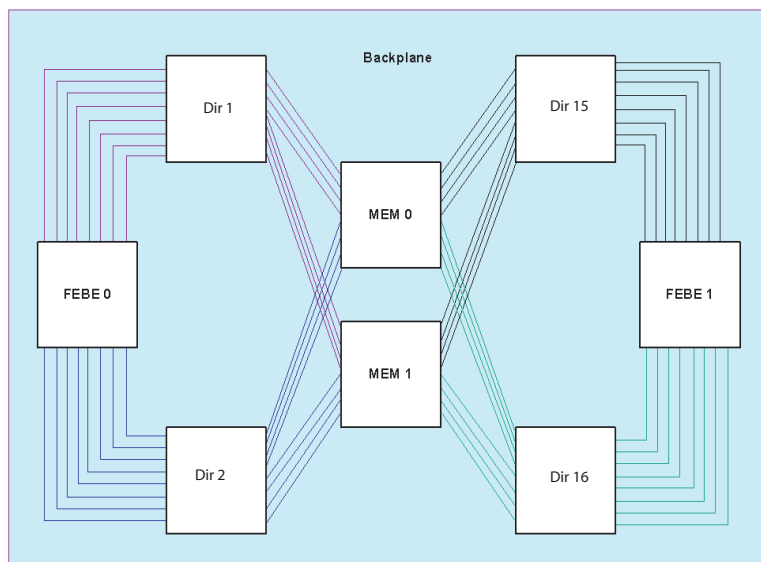
Only seven slots are utilized in a two FC director configuration.



SYM-000301

Figure 3-17 DMX800 Two FC Director System Backplane

The four FC channel directors, two FEBEs, and the two global memory directors plug into the eight-slot DMX800 backplane (Figure 3-18 on page 3-20).



SYM-000419

Figure 3-18 DMX800 Four FC Director System Backplane

FEBE Board Functions

The DMX800 uses the FEBE adapter to provide both the interface between a host and the channel director, and the interface with the disk subsystem (DAE). There are two FEBE boards (FEBE 0, FEBE 1) in the DMX800. FEBE 0 interfaces with Dir 1, Dir 2; FEBE 1 interfaces with Dir 15, Dir 16.

The transmission of control and status between the FEBE board and the disk subsystem is accomplished through an RS-232 interface. Diplexing circuitry provides the transmission of the port's channel director data and RS-232 control/status signals to and from the disk subsystem via the LCC and the FEBE board. The FEBE board has 16 total ports. Each FEBE board is configured to have up to six front-end (FE) host port connections and four back-end (BE) ports in a system configured with two FC directors, and up to eight front-end and eight back-end ports in a DMX800 configured with four FC directors.

Front-End Ports

The main function of the front-end portion of the 16-port 16-link FEBE is to provide an interface between the hosts and the Symmetrix FC director. One FEBE board can provide an interface for two directors in the system. Each adapter provides one connection per FC loop for two directors. Each FEBE board can support from two to eight host ports for two directors and eight host ports for four director configurations.

Back-End Ports

The FEBE board supports a maximum of eight back-end FC connections to disks.

Independent Power Zones

The FEBE contains two independent power zones —VCC_A and VCC_B. Zone A supplies power to all top power blades. Zone B supplies power to all bottom power blades. All boards remain powered on if one or the other power zones fails.

Global Memory Director Board Functions

The DMX800 system contains two global memory director boards (M0, M1). Each board contains 16 full-duplex serial ports; there are four port connections to each of the four FC directors (Figure 3-15 on page 3-16). All paths in the directors and memory are fully independent for I/O operations.

Memory Striping

Memory striping ensures that contiguous memory is spread over the four global memory director board regions in 64 memory word blocks (512 bytes). The stripe block size is 64 memory words because this is the maximum block size that can be transferred in one read or write frame. Faster memory access time is achieved because the access to memory is spread out over the four memory regions. Refer to *Memory Striping* on page 4-4 for additional information.

Dual-Path Feature

The dual-path feature provides redundant access to data in the event of a failure of any of the components in the primary path to that data.

Refer to *Fibre Channel Front-End Redundancy* on page 5-7 for an explanation of this feature.

Host Interface

The DMX800 connects to host systems through the front-end adapter ports of the FEBE board. The back-end FEBE ports connect to the channel director. There are 16 front-end ports and 16 back-end ports in each DMX800 system. The host sends and receives data through the FEBE board (acting as a front-end or back-end adapter) to the directors.

Disk Subsystem Interface

Each DAE contains from four to 15 dual-ported fibre drives (73 GB, 146 GB) that are configured on a redundant pair of 2 Gb/s FC loops. For the loop to operate reliably, switch elements (port bypass circuits — PBC) are inserted between each loop node. Each 15-drive DAE is an FC loop (Figure 3-19 on page 3-22).

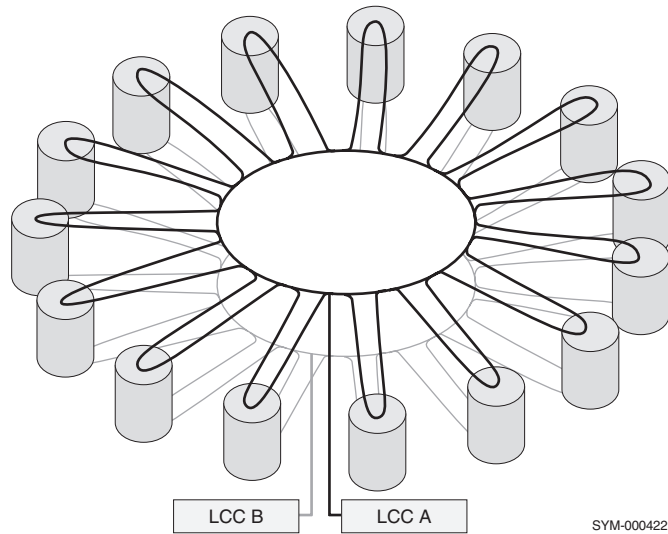


Figure 3-19 DAE Fibre Channel Loops

The LCC independently receives and electronically terminates the incoming FC-AL signal. The transmission of control and status between the FEBE board and the disk subsystem is accomplished through diplexing circuitry to the LCC. Diplexing circuitry provides the transmission of the port's Fibre Channel data and RS-232 control/status signals to and from the disk subsystem and the FEBE board.

DMX800 Communications

The directors in the DMX800 communicate with each other. The SPS status information will be sent through an RS-232 interface to the director boards. The RS-232 interface from the SPS connects to the director through the FEBE board. (Refer to Figure 3-15 on page 3-16, and Figure 3-18 on page 3-20.) The RS-232 port on the system backplane will be sent to the directors.

SPS Communications

Each SPS communicates through an RS-232 interface with a specific director:

- ◆ FEBE 0 connects SPS 0 to Dir 1
- ◆ FEBE 1 connects SPS 1 to Dir 16

Dir 1 and Dir 16 monitor the SPS power to the RMS enclosure and power to the DAEs.

FEBE Communications

Each FEBE board has a connection to each director and the outside world (RJ-45).

**Server
Communications**

The server contains six Ethernet ports to support the following:

- ◆ One port for the SPE (up to four)
- ◆ One port for customer service connection
- ◆ One port for connection to the customer's network

A server communicates to FEBE 0 through an Ethernet port. FEBE 0 communicates to each of the four directors through an Ethernet port.

**Environmental
Control**

The FEBE board has an environmental control capability. It monitors the SPE chassis environmental — PS1 and PS2 and the fan module. DAEs perform their own monitoring of an AC power failure. The SPS monitors AC power failure and reports the information to the directors through an RS-232 interface.

In case of a SPS power failure, the power will remain on for 90 seconds. The SPS can support up to two consecutive 90-second discharge events.

Read/Write Operations

The most critical part of any Symmetrix DMX800 system is its ability to process read/write operations in the most efficient and cost-effective manner possible. The DMX800 system has the following characteristics:

- ◆ 17.6 Gb/s total system aggregate bandwidth
- ◆ Four back-end ports to disks; four front-end ports to host(s) (eight ports each in a four FC director system)

The following sections describe the read/write operations in a Symmetrix DMX800 system.

Overview

Read/write operations in the DMX800 system are unique because the FEBE board acts as both a front-end adapter and a back-end adapter for the Fibre Channel directors.

For *write* operations, the FEBE receives data from the host and passes it on as a front-end adapter to the director.

In a *read* operation, the FEBE receives data from the DAE (disk drives) and passes it on as a back-end adapter to the director.

Refer to Figure 3-15 on page 3-16 for additional information.

Read Operations

The basic flow of a *read* operation is as follows:

DAE -> FEBE (BE) -> Dir -> Memory -> Dir -> FEBE (FE) -> Host

Read operations begin with the data stored on the disk drives within each DAE. The data is sent from the disks to the LCCs by way of the port bypass circuitry, which then sends the data to the back-end (BE) FEBE ports to the Fibre Channel directors (Dir, Dir 16). The directors then transmit the data to global memory. From memory, the data is transmitted to the Fibre Channel director to the front-end (FE) ports of the FEBE, which transmits the data to the host(s) system.

Write Operations

The basic flow of a *write* operation is as follows:

Host -> FEBE (FE) -> Dir-> Memory -> Dir -> FEBE (BE) -> DAE

Write operations begin when the host(s) system sends data to the back-end adapter ports of the FEBE board to the Fibre Channel director. The channel director (Dir 1, Dir 16) transmits the data to memory, then on to the back-end (BE) adapter ports of the FEBE board to the Fibre Channel director. The data is then sent to the DAE through the LCCs, where it is then written to the disk drives.

I/O Performance Enhancements

Symmetrix systems use these techniques to enhance performance:

- ◆ Split director functions—The FEBE board (acting as the front-end, back-end adapter) services requests between global memory and disk. Splitting the director functions eliminates the processing overhead and global memory locking associated with control units that perform both functions.
- ◆ Multiple back-end connectivity—The FEBE board's function as a back-end adapter supports 15-drive Fibre Channel drive loops. For more information on director and disk combinations, refer to Figure 3-19 on page 3-22.
- ◆ High speed global memory—Global memory speed is greater than the total speed of all components (for example, the directors) that access it.
- ◆ Disk microprocessor and buffer—Each disk device has its own microprocessor and buffer that respond to the actuator level, providing parallel processing of data. These features add another level of caching and improve overall performance.
- ◆ RPS miss elimination (mainframe devices)—The dedicated microprocessor and segmented data buffer on each disk device can temporarily store data until the disk director is ready to read or write data. This eliminates rotational positional sensing (RPS) misses that occur in conventional mainframe DASD when the heads are positioned over the desired sector, but the channel path is not ready for read or write operations. The segmented data buffer of the disk device allows multiple operations to occur to the head/disk assemblies.
- ◆ Sequential access patterns—Access patterns can be sequential, random, or a combination of both. When a miss occurs on a sequential access pattern, the number of blocks brought into global memory increases, which improves the hit rate because the requested data is in global memory.

Performance and Optimization

This chapter describes the performance features of the DMX800 system and explains how to optimize these features.

- ◆ Overview4-2
- ◆ Global Memory Performance Features4-3
- ◆ Symmetrix File System Performance Features4-9
- ◆ Multiple Channel Directors4-11
- ◆ Open Systems Hypervolumes.....4-14
- ◆ Mainframe Systems Hypervolumes.....4-18
- ◆ Optimizing Symmetrix System Performance4-22
- ◆ Multiport Volume Access for Open Systems Environments4-23

Overview

Real-world workloads consist of many different types of I/O activity. They can be read or write requests, have different data block sizes, or be skewed (some disks or host channels doing more work than others). They can be highly random, sequential, or mixed, and they are often *bursty* (peak reads or writes can come at unexpected times). The workloads used for envelope measurements are normally static, simple, and designed to always yield certain levels of hit ratio (access of read/write data directly out of global memory), regardless of the global memory size and algorithms. In reality, the actual application behavior is greatly influenced by the performance optimization algorithms.

Symmetrix Enginuity contains extensive algorithmic intelligence that is designed to deliver fast application response time. These algorithms optimize use of internal resources and allow end-user definition of priorities for Symmetrix operations. These unique algorithms balance the load evenly among Symmetrix components, save valuable resources, and optimize data layout based on detection of long-term workload patterns.

The Symmetrix DMX800 systems offer improved performance over conventional storage control units and DASD designs. The Symmetrix features described in this section allow high global memory hit ratios and less processing overhead, thus reducing response time and improving throughput.

The Symmetrix Enginuity performance optimization features are summarized in the following sections.

Global Memory Performance Features

The Symmetrix DMX global memory director design provides point-to-point, nonblocking connectivity between the front-end channel directors and the FEBE board back-end ports. Each global memory director is physically partitioned into four separately addressable, simultaneously accessible regions. These offload engines, utilizing Symmetrix DMX technology, implement advanced techniques while securing data integrity and optimizing available resource usage. This EMC exclusive global memory design managed by EMC's Enginuity Storage Operating Environment delivers a high-performance, fault-tolerant, ultra-reliable system.

The global memory directors deliver consistently high levels of system performance, improve responsiveness and consistency of responsiveness, and manage peak I/O requests through a series of techniques that essentially eliminate contention for shared global memory and optimize utilization of system resources. The underlying principles are:

- ◆ Global memory is partitioned into four separately addressable regions on each memory board. Two global memory director boards provide from four to 64 GBs of total global memory in the DMX800.
- ◆ Requests for global memory are expedited to reduce locking.
- ◆ Requests are intelligently arbitrated to optimize available resource usage.

This section describes the following Enginuity-supported global memory performance features:

- ◆ Global Memory ASICs
- ◆ Fast Write Capabilities
- ◆ Dynamic Mirror Service Policy (DMSP) Algorithm
- ◆ Disk Rotational Position Ordering (RPO)
- ◆ Disk Multiple Priority Queues
- ◆ PermaCache Option
- ◆ Quality of Service Control Performance Option

Global Memory ASICs

The global memory directors expedite transactions between process requests and global memory. EMC's global memory directors use application-specific integrated circuits (ASICs) technology that acts as intelligent offload engines to perform repetitive system-critical functions. Global memory ASICs consist of the following components:

- ◆ Parallel Global Memory Regions
- ◆ Buffering
- ◆ Access Arbitration
- ◆ Memory Striping

Parallel Global Memory Regions

Symmetrix DMX global memory architecture drastically reduces global memory contention by partitioning the global memory into four separately addressable, simultaneously accessible regions. In a Symmetrix DMX800 system with two global memory directors, there are eight separately addressable and accessible global memory regions.

Buffering

Global memory ASICs buffer incoming requests from front-end channel directors and, as soon as possible, free up the global memory region being accessed. The result is a truly nonblocking architecture that is capable of massive performance scaling.

Access Arbitration

Global memory directors also arbitrate incoming requests for global memory resources in such a way as to optimally allocate global memory regions to incoming requests by appropriately timing and intelligently prefetching required information from global memory into the buffer. This results in optimal utilization of available resources.

Memory Striping

To eliminate any contention that may occur while working with large chunks of data, global memory directors stripe any data that resides on global memory across all the four regions within each global memory director. Striping is carried out in units of 64 memory words¹ across the four regions of each memory director.

1. David E. Culler, Jaswinder Pal Singh, 1999: *Parallel Computer Architecture*, Morgan Kaufmann

For example, for an instruction requesting to write 256 memory words of data to a given address, the global memory director will break out this request into four bursts:

- ◆ The first 64-word chunk (from 1 to 64) is written to region 1
- ◆ The second chunk (from 65 to 128) is written to region 2
- ◆ The third chunk (from 129 to 192) is written to region 3
- ◆ The last chunk (from 193 to 256) is written to region 4

Striping of data across the four global memory regions on the global memory directors mitigates any contention for I/O through global memory.

Fast Write Capabilities

Symmetrix systems process all write operations in global memory, eliminating the need to write data to the disk immediately. This capability results in faster response times and improved overall subsystem performance. Channel directors and the FEBE board back-end-function dynamically allocate global memory space between reads and writes, depending on I/O activity.

Dynamic Mirror Service Policy (DMSP) Algorithm

When implementing mirrored (RAID 1) data protection, DMSP determines which of the mirrors of a given logical volume should service a read miss (when the data is *not* found in global memory) operation. BCVs can also participate in the DMSP. DMSP provides a dynamic approach to setting the optimal mirror service policy. The DMSP algorithm monitors the access patterns to the different logical volumes in the Symmetrix back end and, based on this, determines a policy for the next few minutes. The DMSP tries to achieve two goals:

- ◆ Balance the load among all of Symmetrix back-end components
- ◆ Minimize the seek-time spent on the physical drives

DMSP also works with the Symmetrix File System (SFS) as described in *Enhancement of Dynamic Mirror Service Policy* on page 4-9.

Disk Rotational Position Ordering (RPO)

Disk Rotational Position Ordering (RPO) optimization can more than double the number of random I/Os a disk can do. Whenever multiple I/O requests are queued on the disk, the EMC Enginuity Operating Environment can optimize the order in which the I/Os are executed. The RPO optimization reorders the I/Os based on their physical location on the drive.

RPO optimization significantly reduces the effect of seeks and latency times on the overall performance of the disk. To take full advantage of the RPO optimization, the Symmetrix system needs to queue enough I/Os on the physical drives. The more I/O demand the Symmetrix system encounters, the better it will perform.

Disk Multiple Priority Queues

Disk multiple priority queues enables the Symmetrix system to give better response times to I/Os the hosts are waiting for, without sacrificing the disk RPO optimization. The multiple-priority queues algorithm handles starvation situations, so even the low priority I/Os are serviced within a time-optimized period.

PermaCache Option

The PermaCache memory option allows you to permanently assign mission-critical data requiring very high performance to global memory. A variable number of contiguous cylinders on the disk devices can be reserved for PermaCache memory backup.

PermaCache memory is best used for infrequently accessed data that needs instantaneous response because this data is normally not in global memory when it is requested. The Symmetrix large global memory and intelligent caching algorithms try to keep frequently accessed data in global memory, making its assignment to PermaCache memory unnecessary.

Other Symmetrix options that require global memory resources affect the amount of global memory available for PermaCache memory. For example, enabling a parity RAID option RAID (3+1) or RAID (7+1) may require you to adjust memory assignments by reducing the existing PermaCache memory area, lowering the write-pending ceiling of the system, or reducing the number of Symmetrix Parity RAID groups defined.

Global Memory Requirements

Your system must be configured with more than the minimum (base) amount of memory for it to use part of that as PermaCache memory. To determine the minimum memory requirement for your Symmetrix configuration, consult your EMC Sales Representative. If a power failure occurs, records that have been updated in PermaCache memory will be destaged to disk.

Power Failures

If a power failure occurs, records that have been updated in PermaCache memory will be destaged to disk.

**Global Memory
Director Replacement**

If a global memory board requires online replacement, PermaCache memory is unaffected (PermaCache is turned off during the removal procedure, then restarted manually after replacement), provided that the replacement board contains the same amount of global memory. Replacement with a smaller capacity global memory requires that the system be brought offline and re-initialized before PermaCache memory will function.

**Online PermaCache
Configuration
Changes**

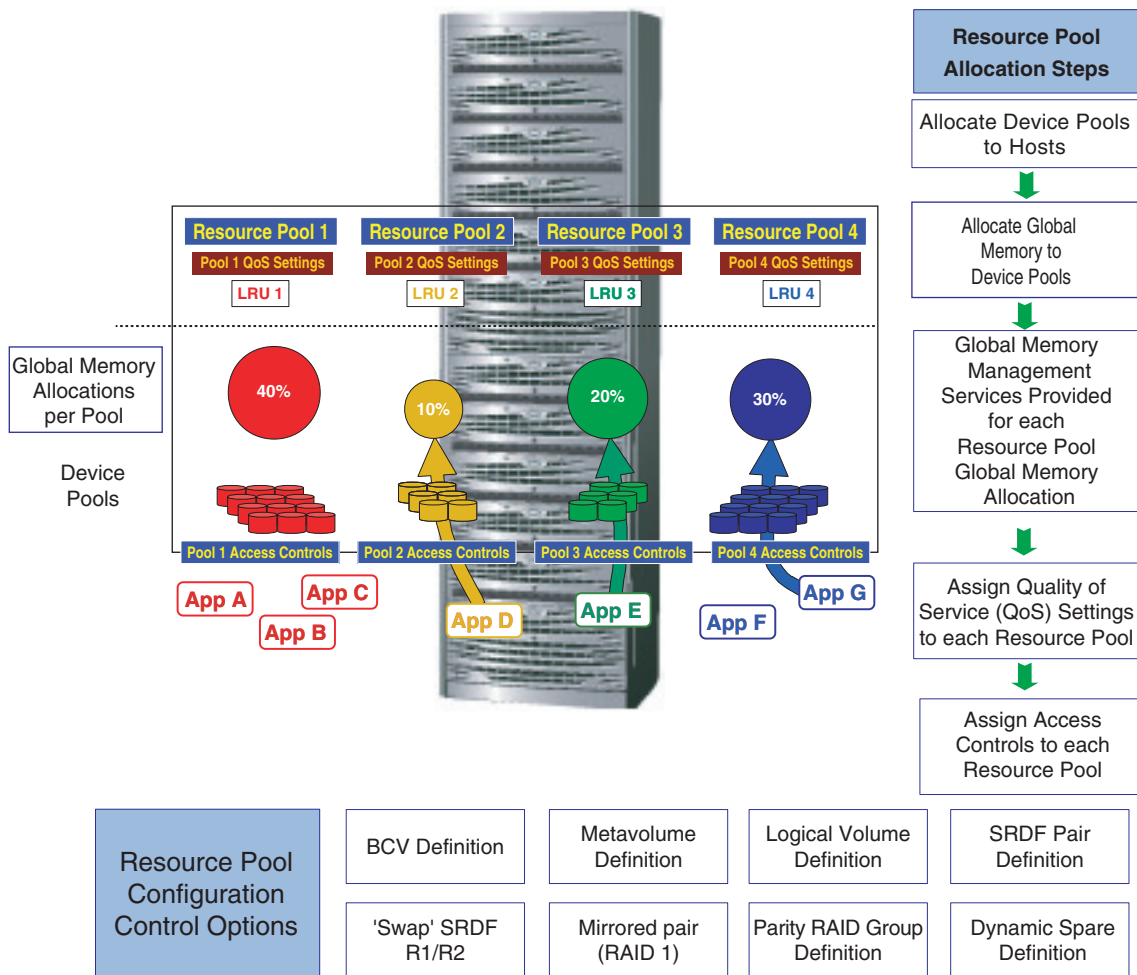
The Symmetrix DMX800 Enginuity permits PermaCache configurations to be changed while the Symmetrix system is online to the host. Enabling PermaCache modifications to take place via an online configuration change allows users to set a configuration with no disruptions and avoid requirements for future maintenance.

**Quality of Service
Control
Performance Option**

Enginuity provides Symmetrix users with the ability to assign higher or lower priority (quality of service [QoS] level) to various applications in their production environment by creating multiple global memory-slot rings. Symmetrix users decide which applications need to have priority over others. Logical device pools connected with these applications are mapped on a one-to-one basis with global memory slot rings. Each global memory-slot ring is associated with a global memory resource pool. Global memory resource pools are carved out of the Symmetrix global memory. A larger share of global memory leads to faster performance for the associated applications.

Enginuity gauges effectiveness of global memory resource pool allocation to each ring by collecting and storing hit ratios pertaining to each resource pool (Figure 4-1 on page 4-8). Enginuity running on each front-end director automatically stores collected statistics in their local memory. Periodically, a background task within Enginuity collects all these statistics and combines them in an area in global memory to generate a true hit ratio for each ring.

This true hit ratio helps deliver true performance on demand to the various applications and automatically fine-tunes global memory allocation based on the customer-defined priority of applications associated with resource pools.



SYM-000311

Figure 4-1 Engenuity LRU Resource Pools

Symmetrix File System Performance Features

Enginuity automatically reserves 12 GB (raw) (four 3 GB 6,140 cylinders) logical volumes for internal use as a Symmetrix File System (SFS). This space is automatically allocated while initially loading the Enginuity Operating Environment on Symmetrix systems and is not visible to the host environment.

The SFS generates several benefits for Symmetrix users, namely:

- ◆ Dynamically adjusting performance algorithms
- ◆ Enhancement of dynamic mirror service policy
- ◆ Enhancement of Symmetrix Optimizer
- ◆ More rapid recovery from problems
- ◆ Enhanced system audit and investigation

Dynamically Adjusting Performance Algorithms

State-of-the-art information management algorithms implemented within Enginuity need to store and use statistical data gathered over time. SFS provides a place to store this statistical database that Enginuity creates and uses to dynamically optimize algorithms and optimize performance.

Enhancement of Dynamic Mirror Service Policy

Enginuity uses SFS to store the data from sampling for Symmetrix Dynamic Mirror Service Policy (DMSP). DMSP algorithms within Enginuity collect logical volume activity samples, store this information in SFS, and utilize this information to make mirror service policy decisions that further enhance Symmetrix RAID 1 performance.

Enhancement of Symmetrix Optimizer

Enginuity uses SFS to store information on physical spindle activity and other mirror activity. This data is used to plot a thermograph of hot (busy) and cold (idle) spindles. This thermographical information gathered over a long time period and preserved in SFS, is used to determine the hottest (busiest) and the coldest (idlest) spindles and recommend swaps.

More Rapid Recovery from Problems

SFS capacity allows the collection and analysis of longer error and event traces than previously possible. Enginuity can now record all historical errors and event trace information in the extra capacity provided to Enginuity by SFS. These capabilities enhance EMC's ability to analyze and precisely resolve customer problems in a timely manner and provide better proactive customer service.

Enhanced System Audit and Investigation

Enginuity stores an audit log in the SFS. This enables improved investigation, both at system level and customer environment level. The Symmetrix audit log collects and presents a chronological list of host-initiated Symmetrix actions and activities. SFS tracks and records manual activities. (For example, physically removing/replacing a component as well as automatically initiated scripts and API activities (TimeFinder or SRDF routines). This provides a means to oversee and historically recall how and when a Symmetrix device is being used.

Enginuity also expands capabilities for host applications built on SymmAPI™. Earlier, host applications were allowed to make entries to a Symmetrix write-only buffer. Audit log features now enable Symmetrix ISVs and other service providers to use logged information for their own reporting purposes. The built-in system security in SymmAPI does not allow host applications to change audit logs. This reduces IOSQ (I/O Supervisor Queue Time) exposure and increases system performance.

Multiple Channel Directors

The Symmetrix DMX800 Fibre Channel directors process data from the host and manage access to global memory over a direct matrix (DMX) technology (Figure 2-4 on page 2-10 and Figure 2-5 on page 2-11). Each of the FC directors on the Symmetrix DMX800 system supports four internal links to each global memory director (total of eight internal links to global memory from each FC director), and each MPCD supports eight internal links to global memory.

Contact your EMC Sales Representative for specific supported channel director configurations.

The Symmetrix DMX800 two FC director system with MPCD installed, supports *simultaneous* connections to mainframe systems hosts and open systems hosts when the required Symmetrix ESP software is installed on the Symmetrix system.

For detailed information on each Symmetrix channel director, refer to *Channel Directors, FEBE Boards, and Global Memory Directors* on page 2-26.

The DMX800 systems support these channel directors:

- ◆ Fibre Channel directors (open systems hosts)
- ◆ MPCD (open systems and mainframe hosts)
 - GigE remote directors for SRDF connectivity
 - iSCSI (open systems hosts) (GigE supports iSCSI connectivity)
 - FICON directors (mainframe hosts)

Channel Speeds and Cable Lengths

Fibre Channels

Symmetrix system channel speeds (data transfer rate) and cable lengths vary according to the type of channel director. This section describes the data transfer rates and supported cable lengths for the different channel directors.

Fibre Channels transfer data at speeds to 2 Gb/s. Symmetrix systems support multimode cable lengths from 10 feet (three meters) to 984.25 feet (300 meters). For specific cable distances refer to Table 2-11 on page 2-41.

For information on Fibre Channel host adapters and cable requirements for your Symmetrix DMX800, consult your EMC Sales Representative.

FICON Channels

Symmetrix FICON channels transfer data at speeds up to 2 Gb/s. The DMX FICON design auto-detects 2 Gb/s or 1 Gb/s at switch or channel port login time. The setting will stay at 2 Gb/s unless there is a problem or if the other port is running at 1 Gb/s. Only a switch has settings on each port for 2 Gb/s, 1 Gb/s, or auto-negotiate.

Symmetrix systems support FICON fiber cable connections up to 7.45 miles (10 km) with single-mode 2Gb/s connections only. For specific cable distances, refer to Table 2-10 on page 2-39.

Symmetrix systems directly connect to FICON single-mode cables only. For information on FICON multimode configurations, consult your EMC Sales Representative. For information on additional configuration options, refer to *Additional FICON Features* on page 2-40.

iSCSI Channels

Symmetrix iSCSI channels transfer data at speeds up to 1.25 Gb/s. Symmetrix systems support iSCSI cable connections up to 524.93 feet (160 meters) with multimode connections.

GigE Remote Channels

Symmetrix GigE Remote channels transfer data at speeds up to 1.25 Gb/s. Symmetrix systems support GigE remote cable connections up to 524.93 feet (160 meters) with multimode connections.

Host Connectivity

Symmetrix DMX800 systems support open systems host connectivity to open UNIX, Linux, Windows, and iSeries systems through Fibre Channels. Symmetrix DMX800 systems support mainframe systems host connectivity through FICON channels.

Symmetrix systems only supports Fibre Channel connectivity for iSeries hosts.

Symmetrix DMX800 systems support *simultaneous* connections to mainframe systems hosts and open systems hosts when the required Symmetrix ESP software is installed on the Symmetrix system.

Parallel Processing

Each Fibre Channel director contains four resident microprocessors, and each disk device contains one resident microprocessor. These microprocessors use advanced parallel processing to reduce processing time and improve throughput. The MPCD contains two microprocessors on each mezzanine card.

RPS Miss Elimination

In Symmetrix systems, each disk device contains a dedicated microprocessor and segmented data buffer that can temporarily store data until the disk director is ready to read or write data. This eliminates rotational position sensing (RPS) misses that occur in conventional DASD when the heads are positioned over the desired sector, but the channel path is not ready for read or write operations. The segmented data buffer of the disk device allows multiple operations to occur to the head/disk assemblies.

Open Systems Hypervolumes

This section describes the hypervolume extension feature on Symmetrix open systems units. It includes the following information:

- ◆ Hypervolume Extension Feature
- ◆ Disk Device Cylinders
- ◆ Logical Volume Mapping
- ◆ Metavolumes

To use the hypervolume extension feature in Symmetrix units connected to mainframe hosts, refer to *Mainframe Systems Hypervolumes* on page 4-18.

Hypervolume Extension Feature

The hypervolume extension feature provides configuration flexibility by allowing one physical device to be split into two or more logical volumes. This capability is particularly useful for some 32-bit implementations of UNIX that allow only 2 GB file systems per single logical disk. The Symmetrix DMX800 system allows up to 77 logical volumes on each Symmetrix DMX 73 GB or 146 GB open systems physical disk device and up to 4,840 logical volumes per Symmetrix DMX800 system depending on the hardware configuration and the data protection options.

For more information on logical volumes supported on Symmetrix DMX disk devices and systems, refer to *Symmetrix DMX800 System Logical Volume Capacities* on page 2-19. Configuration requirements for Symmetrix systems vary according to the applications used. To configure logical volumes for optimum Symmetrix system performance, consult your EMC Sales Representative.

Disk Device Cylinders

The Symmetrix DMX800 system use 3.5-inch low profile disk devices that are available in the following storage capacities and maximum Symmetrix cylinders (Table 4-1).

Table 4-1 Symmetrix Open Systems Disk Capacities and Cylinders

Drive Type	Maximum Usable Symmetrix Cylinders ^a
73 GB	148,734
146 GB	297,799

a. The maximum number of usable cylinders for one logical volume is 65,520.

UNIX and PC Server Hosts

For Symmetrix disk devices attached to UNIX or PC server hosts, each cylinder contains 15 tracks, and each track contains 64 512-byte blocks.

Logical Volume Mapping

The logical-to-physical relationship that you choose can automatically apply to all devices in the unit. You can customize the logical-to-physical relationship on each device as well as the size of each logical volume. For example, if the logical-to-physical ratio chosen is 8:1, the logical volume mapping is similar to that shown in Figure 4-2.

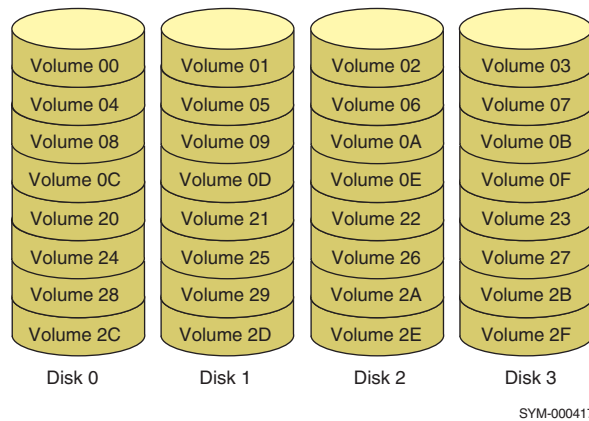


Figure 4-2 Logical Volume Mapping (8:1)

Metavolumes

Several operating systems, such as Windows NT or Windows 2000, some applications software, and some open systems environments require larger volumes than are provided by standard Symmetrix physical disk devices. A metavolume is a logical volume set created from individual physical disks to define volumes larger than the current Symmetrix maximum hypervolume size of approximately 32 GB. Metavolumes are functionally the same as logical volume sets implemented with host volume manager software. Physically, a metavolume is two or more Symmetrix hypervolumes presented to a host as a single addressable device. The metavolume consists of a *head device*, some number of *member devices* (optional), and a *tail device*.

Metavolume Size Requirements

Symmetrix metavolumes can contain up to 255 devices and be up to 7.65 terabytes in size. Metavolumes can be composed of nonsequential and nonadjacent volumes.

When configuring a metavolume, each metavolume device is counted as a single logical volume. Using metadevices will reduce the lot of host-visible devices. Each member of the metadevice must be counted toward the maximum number of host-supported logical volumes. For information on Fibre Channel host attachments, refer to the *EMC Interoperability Support Matrix* located on the EMC website at: <http://www.emc.com>.

Metavolume Performance

By allowing individual physical disk devices to be grouped together into a metavolume, and the use of metavolume addressing, Symmetrix systems enhance disk system functionality. To increase throughput and further improve performance, Symmetrix systems provides multiple I/O drive queues for metavolumes.

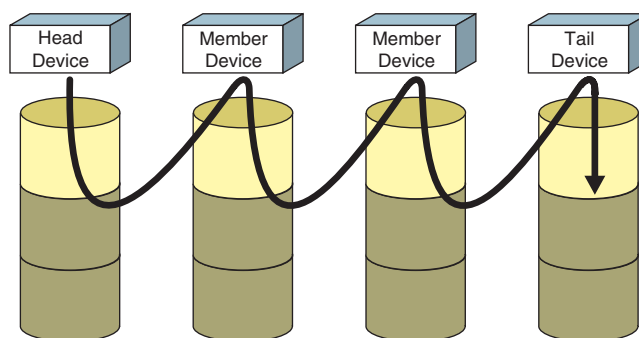
Accessing Data in a Metavolume

You can access data contained in a metavolume in two different ways:

- ◆ Concatenated volumes
- ◆ Striped data

Concatenated Volumes

Concatenated volumes are volume sets that are organized with the first byte of data at the beginning of the first volume (Figure 4-3). Addressing continues to the end of the first volume before any data on the next volume is referenced. When writing to a concatenated volume, the first slice of a physical disk device is filled, then the second, and so on, to subsequent physical disk devices.

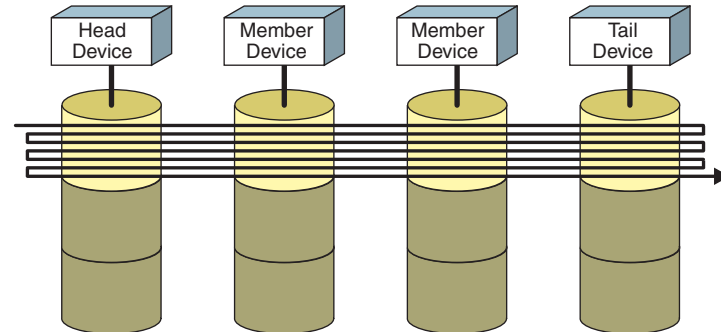


SYM-000421

Figure 4-3 Concatenated Volumes

Striped Data

Metavolume addressing by striping also joins multiple slices to form a single volume. However, instead of using sequential address space, striped volumes use addresses that are interleaved between slices (Figure 4-4). In data striping, equal size stripes of data from each participating drive are written alternately to each member of the set.



SYM-000455

Figure 4-4 Striped Data

Configuring Metavolumes

Although members of a striped set do not have to be the same size, the effective size of each member is the actual size of the smallest member rounded down to an even cylinder count. Therefore, when configuring such a metavolume, be careful to minimize wasted space.

A customer can reconfigure metavolumes. While the Symmetrix system is online to the host, the customer can:

- ◆ Expand both concatenated and striped metavolumes
- ◆ Convert an unused volume to a concatenated or striped metavolume
- ◆ Convert a populated volume to a concatenated or striped metavolume

Striping data across the multiple drives is designed to benefit *random reads* by avoiding stacking multiple reads to a single spindle and disk director. This scheme creates a large volume, but also balances the I/O activity between the disk devices and the Symmetrix disk directors.

DMX systems support two cylinder stripe size and 1,920 stripe capacity (512 byte) blocks.

Mainframe Systems Hypervolumes

To use this feature in Symmetrix units connected to open systems hosts, refer to *Open Systems Hypervolumes* on page 4-14.

This section describes how to use the hypervolume extension feature on Symmetrix mainframe volumes. It contains the following topics:

- ◆ Hypervolume Extension Options
- ◆ Split-Volume Capability
- ◆ Extended Cylinder Addressing Option
- ◆ Determining Cylinders for Hypervolume User Data

Hypervolume Extension Options

The Symmetrix system enhances disk system functionality by supporting multiple logical volumes on each physical device.

The hypervolume extension feature has two usage options:

- ◆ **Split-Volume Capability**—Allows up to 128 logical volumes on each Symmetrix DMX800 physical disk device depending on the data protection option used.
- ◆ **Extended Cylinder Addressing**—Establishes a small logical volume at the end of physical disk device for data requiring high performance on a small volume.

Split-Volume Capability

Using the split-volume option of the hypervolume extension (HVE) feature, Symmetrix systems allow multiple logical volumes to reside on a single physical drive. This split-volume option provides for the consolidation of many physical DASD devices into far fewer physical high-capacity high-performance disks.

Support is provided for native IBM 3390 and 3380 track emulation with all 3390 and 3380 disk volumes being supported. No modifications are required to the operating system, application, or program software to take advantage of HVE.

The split-volume option can override the one-to-one logical-to-physical relationship on all devices in the Symmetrix unit. The logical-to-physical relationship can automatically apply to all devices in the unit. You can also customize the logical-to-physical relationship on each device, as well as the size of each logical volume.

Symmetrix DMX800 systems support up to 128 logical volumes per physical disk device and up to 4,840 logical volumes per unit, depending on the Symmetrix DMX hardware configuration and the data protection options.

For more information on logical volumes supported on Symmetrix DMX disk devices and systems, refer to *Symmetrix DMX800 System Logical Volume Capacities* on page 2-19. Configuration requirements for Symmetrix systems vary according to the applications used. To configure logical volumes for optimum Symmetrix system performance, consult your EMC Sales Representative.

Extended Cylinder Addressing Option

Extended Cylinder Addressing places a small logical volume at the end of a disk device. Each small logical volume can occupy a variable number of contiguous cylinders. This small logical volume can be as small as one cylinder in size. The maximum number of cylinders for this volume depends on the capacity of the disk device and the emulation mode selected.

This flexibility in logical volume configuration increases system performance for datasets requiring extremely high performance such as Multi-Image Manager files, JES Checkpoint, RACF Control files, catalogs, and so on.

Because the data of interest resides on a small volume, the Unit Control Block (UCB) busy conditions that arise when this data is placed on larger capacity volumes with high activity are eliminated. This reduces IOSQ (I/O Supervisor Queue Time) exposure and increases system performance.

Consult your EMC Systems Engineer regarding dataset placement on these small logical volumes.

This logical volume can be added to Symmetrix units that are configured and running without affecting existing data on the disk device by giving the volume an address beyond any currently used on the unit. This option may also be used with the PermaCache memory option, in which infrequently accessed datasets requiring instantaneous access reside permanently in global memory.

Determining Cylinders for Hypervolume User Data

Before you can use either HVE feature — the split-volume option or extended cylinder addressing—you must determine the *total user cylinders* on the physical disk device available for user data.

The number of cylinders available for user data depends on the:

- ◆ Symmetrix disk device cylinder capacity
- ◆ Selected disk emulation type
- ◆ Alternate, diagnostic, and device support (ADDS) cylinders for the emulation type
- ◆ Internal device support cylinders
- ◆ Desired number of logical volumes for the Symmetrix disk device

Table 4-2 lists the 73 GB and 146 GB disk device capacities for 3380 and 3390 emulations.

Table 4-2 Symmetrix Mainframe Disk Capacities and Cylinders

Drive Size 3.5-inch, Low Profile	Available Symmetrix Cylinders	
	3380 Emulation	3390 Emulation
73 GB	98,204	84,910
146 GB	196,393	170,136

Table 4-3 outlines the cylinders required for full emulation per logical volume for several emulation types. It also lists the number of alternate, diagnostic, and device support cylinders (ADDS), and the Symmetrix internal device support cylinders required for each logical volume by that particular emulation type.

Table 4-3 Device Emulations and Number of Cylinders

Emulation Type	Number of Emulated Cylinders	Alt, Diag, Device Support Cylinders (ADDs)	Symmetrix Internal Device Support Cylinders	Total Support Cylinders
3380D	885	3	2	5
3380E	1,770	4	2	6
3380K	2,655	5	2	7
3390-1	1,113	4	2	6
3390-2	2,226	4	2	6
3390-3	3,339	4	2	6
3390-9	10,017	7	2	9
3390-27	32,760	0	2	2

Symmetrix DMX systems running Enginuity level 5670 and higher can support 64K cylinder CKD devices for operating systems that can exploit it.

Optimizing Symmetrix System Performance

Achieving optimal performance requires careful and detailed planning of the Symmetrix configuration according to the requirements of the host(s) you are connecting to the Symmetrix system and your performance needs.

Carefully review the issues listed below with your EMC Sales Representative before the EMC Customer Engineer installs the Symmetrix system. Appendix C, *Symmetrix DMX800 System Planning and Installation*, includes sample worksheets and checklists to assist you in this process.

Performance Guidelines for Open System Devices

Resolve the following open system device issues before installing your Symmetrix system:

- ◆ Rank distributed workloads from the busiest to the least busy (when configuring multiple hosts to a Symmetrix unit)
- ◆ Data storage capacity required for each host connected to the Symmetrix system
- ◆ Number of channels available from each host
- ◆ Nature of the applications executed on the host connected to the Symmetrix system
- ◆ Availability of a logical volume manager (LVM) on the host and the use of data striping
- ◆ Use of hypervolumes on the Symmetrix system, hypervolume size, and the allocation of hypervolumes between different hosts, different channels, and different applications
- ◆ Maximum drive and file system sizes supported by each host connected to the Symmetrix system
- ◆ Requirements for device sharing
- ◆ Number of Fibre Channel directors used and the number of ports used on each director
- ◆ Special considerations for host-level mirroring for device distribution in the Symmetrix system
- ◆ Possibility of upgrading the Symmetrix system with additional drives (DAEs) in the future, and its effect on the configuration if the model installed is not at maximum capacity

Multiport Volume Access for Open Systems Environments

Unlike most disk arrays, the Symmetrix unit can present any logical volume through any number of Fibre Channel channels to the hosts connected to the unit. Usually, a Symmetrix volume will be presented through one host channel. However, in different failover or cluster scenarios, the volume will be *visible* to different hosts on two or more Fibre Channel channels.

For example:

- ◆ A volume may be configured to be visible on two channels for host channel failover (such as EMC PowerPath or HP-UX PV links).
- ◆ A volume can be accessed by up to 32 separate paths using EMC PowerPath.
- ◆ A volume can be configured to be visible on one channel for a hot standby scenario where one host can assume the devices of another host in case of a controller or host failure.
- ◆ A volume can be visible through all channels in cluster environments that can take advantage of such multiport volume access (such as NCR Teradata).

For more information on multiport volume access, and the most current information on Symmetrix systems and specific host integration, contact your EMC Sales Representative, or refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

This chapter discusses the DMX800 features and options that affect data availability and reliability.

◆ Overview	5-2
◆ Reliability and Availability Features	5-7
◆ Maintaining Data Integrity	5-14
◆ Data Protection Guidelines	5-19
◆ Disk Mirroring Concepts	5-21
◆ Symmetrix DMX800 Parity RAID Options	5-24
◆ Symmetrix DMX800 RAID 5 Options	5-37
◆ Symmetrix RAID 10	5-39
◆ Dynamic Sparing	5-40

Overview

DMX800 has many features and options to ensure a high degree of system and data availability. Many of these features and options are built into the Enginuity operating environment. Other availability options may be purchased separately and implemented into the DMX800 system operation.

Symmetrix Reliability and Availability Features

The DMX800 design offers the following reliability and availability features:

- ◆ High-reliability components
- ◆ Global memory director redundancy
- ◆ FEBE board redundancy
- ◆ Redundant internal data paths
- ◆ Internal control data path redundancy
- ◆ Fibre Channel back-end redundancy
- ◆ Redundant power subsystem
- ◆ Nondisruptive maintenance and microcode upgrades and loads

These basic Symmetrix features provide protection against loss of system and data availability due to a power loss or failed component. A redundant design allows the Symmetrix system to remain online and operational during component repair. For example, if a power supply fails, the remaining power supplies share the load until the failed component is replaced.

Symmetrix Data Integrity Protection Features

The Symmetrix system is designed with these data integrity features:

- ◆ Error checking, correction, and data integrity protection
- ◆ Disk error correction and error verification
- ◆ Global memory access path protection
- ◆ Global memory error correction and error verification
- ◆ Periodic system checks

Error verification prevents temporary errors from accumulating and resulting in permanent data loss. The Symmetrix system also looks at the error verification frequency as a signal of a potentially failing component. The periodic system check tests all components as well as microcode integrity. The Symmetrix system reports errors and environmental conditions to the host system as well as the EMC Customer Support Center.

Data Protection Options

Although the Symmetrix system has standard features that provide a higher level of data availability than conventional DASD, the following data protection options ensure an even greater level of data recoverability and availability. These data protection options are configurable at the physical volume level so that different levels of protection can be applied to different datasets within the same Symmetrix system.

You can choose from the following Symmetrix data protection options to match your critical data requirements:

- ◆ Mirroring (RAID 1)
- ◆ Parity RAID (3+1) or Parity RAID (7+1)
- ◆ RAID 5 (3+1) or RAID 5 (7+1)
- ◆ Symmetrix Remote Data Facility (SRDF)
- ◆ Symmetrix Remote Data Facility/Synchronous (SRDF/S, formerly SRDF)
- ◆ Symmetrix Remote Data Facility/Asynchronous (SRDF/A)
- ◆ Symmetrix Remote Data Facility /Data Mobility (SRDF/DM)
- ◆ Dynamic Sparring

Mirroring (RAID 1) Option

Mirroring (RAID 1) provides the highest performance, availability, and functionality for all mission-critical and business-critical applications. With the mirroring option, DMX800 maintains two identical copies of a logical volume on separate disk devices. Should the DMX800 system be unable to read data from one volume of a mirrored pair, it immediately retrieves the data from the other logical volume.

For more information on the Symmetrix Mirroring option, refer to *Disk Mirroring Concepts* on page 5-21.

Parity RAID Option

The Symmetrix Parity RAID option provides high-availability data with high performance, and higher usable storage capacity than Symmetrix mirroring. Parity RAID improves data availability on drives in Symmetrix DMX800 disk subsystems by using a portion of the array to store redundancy information. The data protection feature is based on a Parity RAID (3+1) volume configuration (three data volumes to one parity volume) or Parity RAID (7+1) volume configuration (seven data volumes to one parity volume).

For more information on Symmetrix Parity RAID options, refer to *Symmetrix DMX800 Parity RAID Options* on page 5-24.

RAID 5 Option

Enginuity code 5670 includes RAID 5 support. A RAID 5 device has the following attributes:

- ◆ Data blocks are striped/interleaved horizontally across the members of a RAID 5 group, similar to a metavolume.
- ◆ There is no separate parity drive in a RAID 5 group. Instead, parity blocks rotate amongst the group members.

Symmetrix Remote Data Facility (SRDF)

EMC's SRDF family of software products help companies manage planned data center events such as scheduled maintenance, daily backups, migrations, and application testing, as well as recovery from unplanned outages such as disasters. The following sections briefly describe the SRDF product offerings.

Each of the SRDF product offerings is sold under a separate license by EMC. For more in depth information about SRDF, refer to the *Symmetrix Remote Data Facility Product Guide*.

SRDF/Synchronous (SRDF/S formerly SRDF)

The Symmetrix Remote Data Facility/Synchronous offers high performance, host-independent, real-time synchronous remote replication from one Symmetrix to one or more Symmetrix systems. SRDF/S duplicates production (source) site data at a logical volume level to a recovery (target) site transparently to users, applications, databases, and host processors.

When a primary (source) device is down, SRDF/S enables fast switchover to the secondary (target) copy data so that critical information is again available in minutes. SRDF/S provides complete business continuance capability during the unlikely event of a data center disaster or during planned events such as daily backups, scheduled maintenance, and data center migrations or consolidations.

In either case, the same information protection capabilities are provided. After a system event, SRDF/S can resynchronize data to the source or to the target system at the user's discretion, thereby ensuring information and data consistency.

SRDF/Asynchronous (SRDF/A)

SRDF/Asynchronous (SRDF/A) is an SRDF application used for remote replication that allows you to maintain an asynchronous *dependent write consistent* copy of Symmetrix stored data at a remote site. In SRDF/A implementations:

- ◆ The recovery data from the target side lags behind the source by seconds or minutes.

- ◆ There is minimal host impact since the host adapter does not wait for an acknowledgement for background writes. Global memory is used as a data buffer.
- ◆ It has the potential to more efficiently use network bandwidth due to *locality of reference* (for example, applications tend to write data in proximity of time and space). SRDF/A sends data to a remote site as cycles of dependent write consistent data. The write I/Os coming in are managed in global memory buffers through complex Enginuity code algorithms and, regardless of how many times a specific block is written to, only the last write is sent over to the target box with its associated predefined timed cycles called *delta sets* (there are two alternating cycles). Therefore, less data needs to be transferred to the target system.

SRDF/ Data Mobility (SRDF/DM)

Symmetrix Remote Data Facility/Data Mobility (SRDF/DM) is a special form of SRDF that permits operation in SRDF *Adaptive Copy* mode only and is designed for data replication and/or migration between two or more Symmetrix systems. SRDF/DM transfers data from source volumes to remote volumes, quickly permitting information to be shared, content to be distributed, and access to be local to additional processing environments.

Adaptive copying modes facilitate data sharing and migration. These modes allow the primary and secondary volumes to be more than one I/O out of synchronization. The maximum number of I/Os that can be out of synchronization is known as the *maximum skew value*. The default value is equal to the entire logical volume. The maximum skew value for a volume can be set using the SRDF monitoring and control software.

There are two adaptive copying modes: adaptive copy write-pending (AW) mode and adaptive copy disk (AD) mode. Both modes allow write tasks to accumulate on the local system before being sent to the remote system.

Adaptive Copy mode enables applications using that volume to avoid propagation delays while data is transferred to the remote site. SRDF/DM supports all Symmetrix systems and all Enginuity levels that support SRDF, and can be used for local or remote transfers. Unlike full-function SRDF, SRDF/Data Mobility is not intended for disaster recovery.

SRDF/Automated Replication (SRDF/AR)

SRDF/Automated Replication provides rapid business restart over any distance with no data exposure through advanced single-hop and multi-hop configurations using combinations of TimeFinder/BCV and SRDF/S and/or SRDF/DM.

SRDF/Consistency Groups (SRDF/CG)

SRDF/Consistency Groups maintains data coherency across an SRDF-based configuration by monitoring data propagation from the source volumes to their corresponding target volumes to ensuring data consistency and restartability.

Dynamic Sparing

Dynamic sparing is another data protection option you can use in conjunction with mirroring, RAID, or SRDF. It limits the exposure after drive failure and before drive replacement. With dynamic sparing, the Symmetrix system maintains a small pool of spare volumes that function only when the Symmetrix system detects a potentially failing device. The dynamic sparing feature performs a fast copy of the data and uses the spare until the original device can be replaced.

For more information on the Symmetrix Dynamic Sparing option, refer to *Dynamic Sparing* on page 5-40.

Reliability and Availability Features

The Symmetrix system has several features that allow it to maintain data integrity and maximize system availability. This section discusses several features in detail.

Reliable Components

Symmetrix systems use components that have a mean time between failure (MTBF) of several hundred thousand to millions of hours for a minimal component failure rate. A redundant design allows Symmetrix systems to remain online and operational during component repair.

A periodic system check tests all components as well as microcode integrity. Symmetrix systems report errors and environmental conditions to the host system as well as the EMC Customer Support Center.

Global Memory Director Data Integrity

Every DMX800 system is configured with two global memory directors to allow for online nondisruptive replacement of a failing global memory director. When a hard error is detected, or temporary errors reach a predetermined threshold, the Symmetrix service processor calls home to request an immediate maintenance action.

The memory design and global memory management algorithms allow a failing area of memory to be fenced off to prevent a loss of data. When board replacement is required, global memory usage is redirected to the remaining good boards in the system, and the suspect board is removed and replaced while the system remains online.

For information on global memory director data integrity protection, refer to *Maintaining Data Integrity* on page 5-14.

Fibre Channel Front-End Redundancy

Fibre Channel front-end redundancy is provided by configuring multiple connections from the host servers (direct connect) or Fibre Channel switch (SAN Connect) to the Symmetrix system. With SAN connectivity through Fibre Channel switches, each Symmetrix DMX port can support multiple host attachments, enabling storage consolidation across a large number of host platforms.

The multiple connections are distributed across separate Fibre Channel front-end ports to ensure uninterrupted access in the event of a channel failure. A minimum of two connections per server or SAN is required to provide full redundancy.

The following is a summary of the number of host/SAN connections available across the Symmetrix DMX series:

Table 5-1 Number of Supported Channel Directors

Model	Number Channel Directors ^a	Total Ports
Symmetrix DMX800 (Two FC directors)	2	16
Symmetrix DMX800 (Four FC directors)	4	32

a. For specific channel director configurations, consult your EMC Sales Representative.

Channel failover functionality is required to automate the failover and failback process to avoid an interruption to data access.

Without this functionality, a path failure due to a problem with the host bus adapter, Fibre Channel switch, fibre cable, or channel director would create the potential for the application to go down. EMC PowerPath (refer to *Infrastructure Software Options* on page 1-20 for a description of PowerPath for open systems) provides this functionality and supports end-to-end (or switch-to-storage) channel failover functionality. PowerPath provides automatic, nondisruptive failover capabilities by redirecting I/Os from a failed front-end port to all remaining ports for open systems devices. PowerPath can automatically detect when a path has failed and sends an error log to notify the host if there is an inactive path. PowerPath then fails over the existing I/O request to another active path to maintain data access and application availability, offering intelligent load balancing capabilities to optimize performance and minimize bottlenecks. All this occurs transparently to the host so the application is not stopped and data is continuously available.

Once the failed path is fixed or repaired, PowerPath automatically detects activity, brings the path back into operation, and automatically sends I/O requests down to the now active path.

The major benefits of the path failover capability is the ability to automatically detect a front-end port failure, transparently redirect an I/O to other available ports, and easily service and restore the failed port without interrupting applications.

Fibre Channel Back-End Redundancy

The Symmetrix DMX800 architecture incorporates a 2 Gb/s Fibre Channel back-end design to ensure high performance and full redundancy. The Fibre Channel back-end disk subsystem provides redundant paths to the data stored on the disk drives to provide non-stop access to information, even in the event of a component failure and/or replacement. The following is an overview of the redundant components of the Symmetrix DMX back-end disk subsystem.

Redundant Disk Directors

A pair of directors (Dir 1 and Dir 16; Dir 2 and Dir 15) is used to access each disk drive. One director is connected to one physical path to the drive, and the other director is connected to a second physical path to the drive. The directors are responsible for moving data between the central global memory and the disks and, as such, are each connected to the global memory through redundant internal paths, to eliminate any possible single points of failure.

Redundant Cable Paths

Each director port is attached to its associated set of Fibre Channel back-end FEBE ports by a separate, independent cable assembly that carries the Fibre Channel loop, control bus, and other protocol signals.

Redundant Disk Ports

Each Fibre Channel disk drive has two fully independent Fibre Channel ports that are designed to connect to two separate loops. In the Symmetrix DMX architecture, these two ports are each connected to port bypass circuitry on the FEBE boards (FEBE 0, FEBE1), which in turn is connected by way of separate cables to two different disk directors. These two pathways are completely independent of one another.

Redundant Port Bypass Circuitry

Each port in a redundant pair connects to port bypass circuitry on the FEBE boards (FEBE0, FEBE1).

Fibre Channel Arbitrated Loop Design

Symmetrix DMX systems employ an arbitrated loop design that contains monitoring and control hardware and software to maximize the performance and availability of each loop. The loop is connected in a Star-Hub topology, with the hub ports gated with a bypass switch that allows individual loop segments (that connect to a Fibre Channel disk drive) to be dynamically inserted or removed.

Data Integrity for Fibre Channel Arbitrated Loops

The loop initiator is a Symmetrix Fibre Channel disk director, which feeds data into and controls the hub. All of the monitoring and control logic is contained on the port bypass circuitry, which also contains the bypass/hub logic for the loop.

To ensure the highest level of availability, the communications link to the monitoring and control logic, is implemented through a separate path. If the Fibre Channel loop is not operating for some reason, the director can access the port bypass card through the control bus and reconfigure the loop into a working state.

In order to insure that the data transmitted by the disk director is the same as the data received by the disk device (and vice-versa) a multi-layer protection strategy is employed. At the highest level, as data is transmitted between the global memory and the disk, a connection referred to as an exchange is established. Logical parameters that identify this connection are initialized, and then all information transferred as part of this exchange is identified with these parameters.

In addition, as the data in global memory is encapsulated into Fibre Channel frames, it is marked with a sequence number that identifies the order of the frame within the exchange and, in addition, a cyclic redundancy checksum is computed over the data and appended to the data frame. The receiving port (after validating the exchange parameters) will re-compute the checksum for the data received and compare it to the transmitted checksum to determine if there are any bytes-in-error. It will also check the sequence number to insure that the received frame is the next one expected for that exchange.

The next level of protection is designed in at the transmission layer of the Fibre Channel protocol. The data frame, with the additional information appended, is translated into special ten-bit data transmission symbols. These symbols are chosen for their ability to increase the reliability of the transmitted data, whether it is through a fiber-optic or copper cable. The transmission codes also help to assist in detecting errors in transmission.

Finally, once the data has been received at the disk drive (for writes) and decoded from the transmission codes, it will then be further re-coded using a very powerful linear block coding algorithm before being stored on the disk's magnetic surface. This code has the ability to not only detect errors, but to also correct them on the fly, insuring that the data read from the drive at a later time will be read perfectly.

Redundant Power Subsystem

The Symmetrix DMX800 systems have a modular power subsystem featuring a redundant architecture that facilitates field replacement of any of its components without any interruption in processing.

The Symmetrix DMX800 power subsystem supports an (N+1) power zone for redundancy and drive expansion. The DMX800 system connects to two dedicated or isolated AC power lines. If AC power fails on one line, the power subsystem automatically switches to the other line. If any Symmetrix DMX800 power supply fails, the remaining power supplies continue to share the load. The Symmetrix DMX800 senses the fault and reports it as an environmental error.

Symmetrix Power Failure on Mainframe Channels

When a power failure occurs, all directors send a unit check status with the environmental data-present bit set to all channels (error code 047A). Power switches immediately to the backup power supply, and the Symmetrix system continues to operate normally. When the timer window elapses, the Symmetrix system presents SCU busy to prevent the host system from writing or reading any data at the unit. The Symmetrix system destages any fast write data still in global memory to disk, spins down the disk devices and retracts the heads, and powers down, turning off the battery at that time. The Symmetrix system will respond `not operational` to the host after it powers down.

Symmetrix Power Failure on Open Systems Channels

When a power failure occurs, power switches immediately to the backup power supply and the Symmetrix system continues to operate normally. When the timer window elapses, the Symmetrix system presents a busy status to prevent the host system from writing or reading any data at the unit. The Symmetrix system destages any Fast Write data still in global memory to disk, spins down the disk devices and retracts the heads, and powers down, turning off the battery at that time. The Symmetrix system will not respond to SCSI commands after it powers down.

Nondisruptive Component Replacement

Symmetrix DMX systems implement a modular design with a low parts count that improves serviceability by allowing nondisruptive component replacement, should a failure occur. The low parts count minimizes the number of failure points. The Symmetrix DMX features concurrent maintenance of all major components, including:

- ◆ Global memory director boards
- ◆ Channel director boards
- ◆ FEBE boards
- ◆ Disk devices (DAEs)

- ◆ Cooling fan modules
- ◆ Link Control Cards
- ◆ Power supplies (SPS)
- ◆ Server

The Symmetrix DMX800 systems provide full component-level redundancy to protect against a component failure and ensure continuous and uninterrupted access to information. A *Disable Interface* feature ensures that at power up the link interfaces between any new channel director, its FEBE board, and its corresponding global memory directors are disabled. Only after safeguards have been satisfied to ensure that a director has been properly replaced will a new board be activated and brought online.

This nondisruptive replacement capability allows the EMC Customer Engineer to install a new component, initialize it if necessary, and bring it online without:

- ◆ Disrupting access to unaffected volumes
- ◆ Powering down the Symmetrix unit
- ◆ Stopping the operating system
- ◆ Taking unaffected channel paths offline
- ◆ Taking devices offline (other than the affected device)

Nondisruptive Enginuity Upgrades

Interim updates of the Enginuity Operating Environment can be performed at the customer site by the EMC Customer Engineer (CE). The nondisruptive upgrade will complete in 15 seconds or less. These updates provide enhancements to performance algorithms, error recovery and reporting techniques, diagnostics, and Enginuity fixes. They also provide new features and functionality for Enginuity.

During an online Enginuity code load, the EMC Customer Engineer downloads the new Enginuity code to the service processor. The new Enginuity code loads into the EEPROM areas within the channel and disk directors, and remains idle until requested for a hot load in control store. The Symmetrix system does not require customer action during the performance of this function. All channel and disk directors remain in an online state to the host processor, thus maintaining application access. The Symmetrix system will load executable Enginuity code within each director hardware resource until all directors have been loaded.

Once the executable Enginuity code is loaded, internal processing is synchronized and the new code becomes operational.

Dynamic Reconfigurations

Symmetrix systems support dynamic reconfiguration activity without disruption to online applications such as:

- ◆ Establish/deestablish mirrored pairs
- ◆ Add/reallocate hypervolumes
- ◆ Modify channel assignments or change emulation modes

Modification to channel assignments without disruption to online applications can occur only if the application has an alternate path to the data.

Nondisruptively Change or Remove Drives

This feature nondisruptively allows an FBA open systems device to be changed or deleted online regardless of protection type. This feature uses the same mechanisms as other online configuration changes to ensure no impacts.

Deleting (and Then Adding) Devices Online

Enginuity code 5670 supports removing and then adding devices online, which facilitates the following configuration enhancements:

- ◆ Change Device Emulation Online—Remove a CKD volume and add an FBA volume and vice-versa.

Adding a CKD volume to a Symmetrix system requires a global memory configuration change if this is the first CKD volume being added. However, the system can be configured in advance, thus avoiding an offline global memory reformat.

- ◆ Convert between mirrored and RAID protected volumes.

Maintaining Data Integrity

The Symmetrix DMX800 system preserves data integrity by performing extensive error checking and correction on all data and addresses it passes internally. The Symmetrix DMX architecture provides the most reliable hardware design in the industry. However, all hardware is subject to the effects of aging and occasional failures. The unique methods used by Symmetrix DMX systems for detecting and preventing these failures in a proactive way set it apart from all other storage solutions in providing continuous data integrity and high availability. The Symmetrix system is designed with these data integrity features:

- ◆ Remote support
- ◆ Error checking, correction, and data integrity protection
- ◆ Disk error correction and error verification
- ◆ Global memory error correction and error verification
- ◆ Periodic system checks

Remote Support

Remote support is an important and integral part of EMC Customer Service and Support. Every Symmetrix unit has an integrated service processor that continuously monitors the Symmetrix environment. The service processor can communicate with the EMC Customer Support Center through a customer-supplied direct phone line. Through the service processor, the Symmetrix system actively monitors all I/O operations for errors and faults. By tracking these errors during normal operation, the Symmetrix system can recognize patterns of error activity and predict a potential hard failure before it occurs. This proactive error tracking capability can often prevent component failures by fencing off, or removing from service, a suspect component before a failure occurs. The *call home* capabilities allow the Symmetrix system to automatically notify EMC Customer Support of potential issues before a failure actually occurs. An EMC Product Support Engineer handles these calls and can dispatch a local Customer Engineer to install a new component without disrupting access to data.

To provide remote support capabilities, the Symmetrix system is configured to phone home and alerts EMC Customer Support of a failure or potential failure. The appropriate authorized EMC Product Support Engineer is able to run system diagnostics remotely for further troubleshooting and resolution.

Configuring the EMC products to allow inbound dial also enables EMC Customer Support to proactively dial in to the Symmetrix system to gather needed diagnostic data or to attend to identified issues. The current dial-in support program for the Symmetrix system uses the latest digital key exchange technology for strong authentication, layered application security, and a centralized support infrastructure that places calls through an encrypted tunnel between EMC Customer Support and the service processor located inside the Symmetrix system.

Before any individual can initiate a call to a customer site, that person must be individually authenticated and determined to be an appropriate member of the EMC support team. Field-based personnel who might be known to the customer must still be properly associated with the specific customer’s account. An essential part of the design of the dial-in support program is that the call to the customer’s Symmetrix service processor must originate from one of several specifically designed Remote Support Networks at EMC. Within each of those EMC Support Centers, the necessary networking and security infrastructure has been built to enable both the call-EMC and call-device functions.

Error Checking and Correction, and Data Integrity Protection

In conventional DASD, the subsystem adds Error Checking and Correction (ECC) bytes to each data record field, as shown in Figure 5-1. The system uses these error checking and correction bytes to check the data and correct it, if possible. If it detects an uncorrectable error, the DASD subsystem informs the host that it has encountered bad data to avoid affecting data integrity.

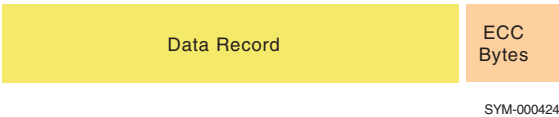


Figure 5-1 Data Record Format for Conventional DASD

Symmetrix systems, like conventional DASD, perform this level of error checking and correction when they pass data and addresses. Symmetrix systems, however, go further to ensure that the information passed belongs to the record specified. The system does this by including additional bytes with the data field of each record. These bytes contain the record ID and a double longitude redundancy code (LRC) check byte, as shown in Figure 5-2 on page 5-16.

The Symmetrix system uses these bytes to check that the data is from the specified record and alarms the host if it is not. This second level of protection further ensures data integrity by preventing incorrect data from being transferred.



SYM-0004

Figure 5-2 Symmetrix Data Record Format

Symmetrix systems assure the highest level of data integrity by checking data validity through the various levels of the data transfer in and out of global memory. Should an error be undetected at one level, it will be detected at one of the other levels.

Disk Error Correction and Error Verification

The disk directors use idle time to read data and check the polynomial correction bits for validity. All data and command words passed between the disk directors and the disk devices include frame-based CRC used to check integrity at each data transfer. If a disk read error occurs, the disk director reads all data on that track to Symmetrix global memory. The disk director writes several worst-case patterns to that track, searching for media errors.

When the test completes, the disk director rewrites the data from global memory to the disk device, verifying the write operation. The disk microprocessor maps around any bad block (or blocks) detected during the worst-case write operation, thus skipping defects in the media. If necessary, the disk microprocessor can reallocate up to 32 blocks of data on that track. To further safeguard the data, each disk device has several spare cylinders available. If the number of bad blocks per track exceeds 32 blocks, the disk director rewrites the data to an available spare cylinder. This entire process is called *error verification*.

The disk director increments a soft error counter with each bad block detected. When the internal soft error threshold is reached, the Symmetrix service processor automatically dials the EMC Customer Support Center and notifies the host system of errors through sense data. The Symmetrix system also invokes dynamic sparing (if the dynamic sparing option is enabled). This feature maximizes data availability by diagnosing marginal media errors before data becomes unreadable.

Global Memory Director Data Integrity Logic

The global memory director technology implements the following data integrity logic on the Symmetrix DMX global memory directors:

- ◆ Global Memory Error Correction and Error Verification
- ◆ Global Memory Chip-Level Redundancy
- ◆ Longitude Redundancy Code (LRC)
- ◆ Byte-Level Parity Checking
- ◆ Global Memory Access Path Protection

Global Memory Error Correction and Error Verification

In addition to monitoring recoverable conditions during normal access, all locations in global memory are periodically read and rewritten to detect any increase in single-bit errors. This global memory verification technique maintains a record of errors for each global memory segment. If the predetermined (single-bit) error threshold is reached in a certain segment, the service processor generates a call-home for immediate attention.

Constant global memory verification to detect and correct single-bit and nonconsecutive double-bit errors dramatically reduces the potential for multibit or hard errors. Should a multibit error be detected during the scrubbing process, it is considered a permanent error. This error is corrected and the segment is immediately fenced (removed from service) and the segment's contents are moved to another area in global memory.

A server call-home alerts EMC Global Service Call Centers whenever an unacceptable level of errors has been detected and a nondisruptive global memory replacement is ordered. Customer service is immediately notified of all call-home alerts, and a customer service engineer can be dispatched with the appropriate parts for speedy repair. Even in cases where errors are occurring and are easily corrected, if they exceed a pre-set level, the call home is executed. This represents the EMC philosophy of not accepting any errors.

Error verification maximizes data availability by significantly reducing the probability of encountering an uncorrectable error by preventing bit errors from accumulating in global memory.

In the mainframe host environment, the Symmetrix system reports uncorrectable bit errors as Equipment Checks to the CPU. These errors appear in the IBM EREP file.

Global Memory Chip-Level Redundancy

Traditional global memory systems usually provide for eight bits of parity information to support bit error correction and detection in a 64-bit long word. The global memory directors incorporate single-nibble correction double-nibble detection capabilities. (A nibble is four consecutive bits of information.) Global memory chip-level redundancy is achieved by internally generating 16 bits of ECC parity information and replacing the incoming parity information. This enables the system to correct up to four bit errors associated with a 64-bit-long word and can detect up to eight bit errors. It also interleaves 64 bits of information plus 16 global memory director parity information (total 80 bits) across 20 memory chips on the global memory board. This results in each memory chip storing only a nibble of information corresponding to a word. So, a chip-level error will disable access only to the nibble stored on that faulty chip. However, the global memory director enables regeneration of data from the faulty chip. This leads to chip-level redundancy making every chip on the global memory board redundant.

Longitude Redundancy Code (LRC)

Symmetrix global memory directors also incorporate sector-level Longitudinal Redundancy Checks (LRC), which further assure data integrity. The check bytes are the XOR (exclusive OR) value of the accumulated bytes in a 4 KB sector. LRC checking can detect both data errors and incorrect block access problems.

Byte-Level Parity Checking

All data paths and control paths have parity generating and checking circuitry that verify data integrity at the byte or word level. All data and command I/Os passed through the direct matrix interconnect, and within each channel/disk director and global memory director, include parity bits used to check integrity at each stage of the data transfer. This provides a system wide error checking capability.

Global Memory Access Path Protection

Before Symmetrix global memory can accept data from a host connection, it must ensure that the area to which the data is to be written is without error. Symmetrix systems assure the highest level of data integrity by checking data validity through the various levels of the data transfer in and out of global memory.

Data Protection Guidelines

The Symmetrix data protection options ensure a higher level of data protection, recoverability, and availability than the standard Symmetrix availability and reliability features. The options in Table 5-2 can be purchased separately and implemented into the Symmetrix operation.



CAUTION

EMC strongly recommends that you use one or more of the data protection schemes described here for your Symmetrix volumes.

Table 5-2 Data Protection Options ^a

Data Protection Option	Description
Mirroring (RAID Level 1)	Provides the highest level of performance and availability for all mission-critical and business-critical applications by maintaining a duplicate copy of a volume on two disk devices. For more information, refer to <i>Disk Mirroring Concepts</i> on page 5-21.
Parity RAID	Provides: • High performance dependent upon volume layout and external striping • High availability - data from lost volume is regenerated from remaining members A Parity RAID (3+1) group consists of three data volumes to one parity volume. A Parity RAID (7+1) group consists of seven data volumes to one parity volume. For more information, refer to <i>Symmetrix DMX800 Parity RAID Options</i> on page 5-24.
RAID 5	Provides: • High performance with automatic striping across hypervolumes • High availability - lost hypervolumes regenerated from remaining members RAID 5 is configured in (3+1) and (7+1) groups. RAID 5 technology stripes data and distributes parity blocks across all the disk drives in the disk array. For more information, refer to <i>Symmetrix DMX800 RAID 5 Options</i> on page 5-37.
Symmetrix Remote Data Facility (SRDF)	Provides an information protection/business continuance solution by maintaining a mirror image of data in multiple Symmetrix systems that can be in physically separate locations. For more information on SRDF, refer to the <i>Symmetrix Remote Data Facility Product Guide</i> .
Dynamic Sparing	Increases data availability by copying the data on a failing volume to a spare volume until the original device is replaced. Dynamic sparing is used as additional protection for mirrored volumes. For more information, refer to <i>Dynamic Sparing</i> on page 5-40 and <i>Dynamic Sparing as Additional Protection</i> on page 5-42.
Symmetrix RAID 10	RAID 10 (one-zero) is a mirroring feature with striping used for mainframe environments. For more information, refer to <i>Symmetrix RAID 10</i> on page 5-39.

a. When configuring multiple data protection options, consult your EMC Sales Representative for specific configuration rules.

Configuring Data Protection Options

Follow these guidelines when configuring data protection options:

- ◆ Parity RAID (3+1) and (7+1) cannot co-exist in the same frame.
- ◆ RAID 5 (3+1) and (7+1) cannot co-exist in the same frame.
- ◆ Any combination of (3+1) and (7+1)—Parity RAID and/or RAID 5, cannot co-exist in the same frame. (For example, Parity RAID (3+1) is not supported with RAID 5 (7+1).)
- ◆ A single parity RAID protection scheme can be configured within a frame with any combination of SRDF, business continuance volumes (BCVs), and mirroring protection.
- ◆ A single RAID 5 protection scheme can be configured within a frame with any combination of SRDF, BCV, and mirroring protection.
- ◆ A single parity RAID protection scheme and RAID 5 of the same scheme can be configured within a frame with any combination of SRDF, BCV, and mirroring protection. (For example, Parity RAID (3+1) is supported with RAID 5 (3+1).)
- ◆ Parity RAID (7+1) is not supported on DMX800 systems with less than four DAEs.

Disk Mirroring Concepts

Mirroring provides the highest level of performance and availability for all mission-critical and business-critical applications. Mirroring maintains a duplicate copy of a logical volume on two physical disk devices. The Symmetrix system maintains these copies internally by writing all modified data to both devices. The mirroring operation is transparent to the host.

The mirroring feature designates from two-to-four logical volumes residing on different physical devices as a *mirrored volume*, one volume being *mirror-1* and the other volumes being *mirror-2*, *mirror-3*, and *mirror-4*. The host views the mirrored volumes as the same logical volume because each has the same unit address.

Advantages of Mirroring

Mirroring offers the following advantages:

- ◆ Improved performance over traditional mirrored (RAID 1) by supporting 100 percent fast write and two simultaneous internal data transfer paths
- ◆ Protection of mission-critical data from any single point of failure
- ◆ Assurance that the second copy of data is identical to the first copy
- ◆ Superior read performance since both mirrors are used for read operations
- ◆ Continuous business operation in a situation where there is a device failure
- ◆ Automatic resynchronization of the mirrored pair after repairing the defective volume

Write Operations with Mirroring

The Symmetrix system handles a write operation to a mirrored logical volume as a normal write operation. The channel director presents channel end and device end (or a good ending status) to the channel after data is written to and verified in global memory. The disk directors then destage the data to each drive of the mirrored pair of drives asynchronously. As such, mirroring on Symmetrix systems exploits the 100-percent fast write capability, and the application does not see additional time associated with physically performing two disk write I/Os (one to each drive of the mirrored pair) as is normally associated with mirroring.

Read Operations with Mirroring

During read operations, if the data is not available in global memory, the Symmetrix system reads the data from the disk pointed to by its performance algorithms for best system performance. The performance algorithms track path-busy information as well as actuator location and which sector is currently under the disk head in each device. If a data check occurs on the device being read, Symmetrix systems automatically read the data from the other device.

The Symmetrix performance algorithms for read operations offer three service policies to best balance the use of the Symmetrix architecture:

- ◆ **Interleave Service Policy**—Shares the read operations of the mirrored pair by reading tracks from both disk devices in an alternating method, a number of tracks from M1, and a number of tracks from M2. Interleave is designed to achieve maximum throughput.
- ◆ **Split Service Policy**—Differs from Interleave because read operations are assigned to either the M1 or the M2 logical volume, but not both. Split is designed to minimize head movement.
- ◆ **Dynamic Mirror Service Policy (DMSP)**—Utilizes both Interleave and Split for maximum throughput and minimal head movement. DMSP adjusts each logical volume dynamically based on access patterns detected. This is the default mode within the Enginuity operating system.

Error Recovery with Mirroring

In the unlikely event that one volume in the mirrored pair fails, the Symmetrix system automatically uses the other volume without interruption of data availability. The Symmetrix system notifies the host operating system of the error and the EMC Customer Support Center through the call-home feature. The EMC Customer Support Center Product Support Engineer (PSE) then begins the diagnostic process and, if necessary, dispatches a Customer Engineer (CE) to the customer site. Once the suspect disk device is nondisruptively replaced, the Symmetrix system reestablishes the mirrored pair and automatically resynchronizes the data with the new disk. During the data resynchronization process, the Symmetrix system gives priority to host I/O requests over the copy I/O to minimize the impact on performance. All new writes take place to both devices. The time it takes to resynchronize the mirrored pair depends on the I/O activity to the volume, the disk device, and the disk capacity.

Dynamic Mirror Service Policy

The Symmetrix Dynamic Mirror Service Policy (DMSP) is an enhancement to the adaptive algorithms in the Symmetrix architecture that improves the performance of read operations in mirrored and business continuance volume (BCV) environments. The improved system performance is a result of the Symmetrix system balancing the load between physical disk drives and disk directors, and minimizing actuator movement.

To achieve this improved performance, the Symmetrix system measures and tracks I/O activities of logical volumes (including business continuous volumes), physical volumes, and disk directors (FEBE board back-end function). Then, based on these measurements, the Symmetrix system directs read operations for mirrored data to the appropriate mirror resulting in the best overall performance of the Symmetrix system. As the access patterns and workloads change, the dynamic algorithm analyzes the new workloads and adjusts the service policies as needed.

Business Continuance Volumes

A business continuance volume (BCV) is a standard device (logical volume) used for dynamic mirroring. It has additional attributes that allow it to independently support host applications and processes. It may be configured as a single mirror, a locally mirrored device, or an SRDF source (R1) device. A BCV device can be RAID 1 or RDF protected, but it *cannot* be Parity RAID (3+1) or (7+1) protected.

A business continuance sequence first involves setting, or establishing, the BCV device as an additional mirror of a standard device. Once the BCV is established as a mirror of the standard device, it is not accessible through its original device address. The BCV device may later be separated, or split, from the standard device with which it was previously paired. After a split, the BCV device has valid data and is available for backup or other host processes through its original device address. Once host processes on the BCV device are complete, the BCV may again be mirrored to a standard device (either the same device to which it was previously attached or a different device). It can then acquire new data for other business continuance processes or update the standard device with any new data from the completed business continuance processes.

Symmetrix DMX800 Parity RAID Options

RAID (redundant array of independent disks) is a combination of hardware and software functionality that improves data availability on disk drives in Symmetrix DMX800 systems by using a portion of the array to store redundancy information. This redundancy information, called parity, can be used to regenerate data if the data on a disk drive becomes unavailable.

The following sections describe how RAID functions in Symmetrix systems and how it differs from conventional RAID:

- ◆ *Parity RAID Technology* on page 5-24
- ◆ *Data Protection Flexibility* on page 5-25
- ◆ *Parity RAID Components* on page 5-25
- ◆ *Parity RAID Modes of Operation* on page 5-29
- ◆ *Writing, Reading, and Rebuilding Data with Parity RAID* on page 5-31
- ◆ *Writing Data in a Parity RAID Group* on page 5-31
- ◆ *Reading Data in a Parity RAID Group* on page 5-33
- ◆ *Data Recovery with Parity RAID* on page 5-33

Parity RAID Technology

Parity RAID technology offers the following advantages:

- ◆ Protects a volume requiring high availability from being a single point of failure
- ◆ High performance, even in the event of a disk failure within a Parity RAID group
- ◆ In the case of a single disk failure, all logical volumes that were not physically stored on the failed disk device perform at the level typical of standard Symmetrix devices
- ◆ In the event of a multiple disk failure within a Parity RAID group, data on all remaining devices within the group remains accessible
- ◆ Automatically restores parity protection on the global memory level to the Parity RAID group after repair of a defective device

Parity RAID employs the same technique for generating parity information as many other commercially available RAID solutions, that is, the Boolean operation EXCLUSIVE OR (XOR).

However, EMC's Parity RAID implementation reduces the overhead associated with parity computation by moving the operation from controller microcode to the hardware on the XOR-capable disk drive. Additional XOR hardware-assist built into the Symmetrix global memory boards further distributes the XOR function throughout the system to improve performance in the regeneration mode of operation.

Data Protection Flexibility

Compared to a mirrored Symmetrix system, Parity RAID offers more usable capacity than a mirrored system containing the same number of disk drives. Like the mirroring or Dynamic Sparing options, Symmetrix RAID parity protection can be dynamically added or removed. For example, for higher performance requirements and high availability, a Parity RAID group of volumes can be reconfigured as multiple mirrored pairs. Within the same Symmetrix system, data can be protected through Parity RAID, mirroring, and SRDF. Dynamic sparing can be added to any of these data protection options.

Parity RAID Components

A Parity RAID group consists of the physical disk devices within the Symmetrix unit that are related to one another for common parity protection. The Parity RAID group is defined by the EMC Customer Engineer at the time Symmetrix system is installed, and includes disk volumes that are designated as either data volumes or parity volumes.

The following sections describe these Parity RAID components:

- ◆ Logical Volume
- ◆ Parity RAID—Data Volume
- ◆ Parity RAID—Parity Volume
- ◆ Parity RAID Group
- ◆ Parity RAID Rank
- ◆ Parity RAID (3+1)
- ◆ Parity RAID (7+1)
- ◆ Hypervolume Extension (HVE) with Parity RAID

Logical Volume

A logical volume is a unit of storage implemented on a single Symmetrix disk drive. When hypervolume extension (HVE) is not used, the size of a logical volume is usually the same as a physical volume. With HVE, up to 58 logical volumes for Parity RAID (3+1) configurations and up to 39 logical volumes for Parity RAID (7+1) configurations of varying sizes can exist on a physical volume.

Parity RAID—Data Volume

A data volume is similar to a traditional logical volume. It is the *virtual volume* image presented to the host operating system and defined as a separate unit address to the host. All data volumes within a rank must be the same size. Parity RAID does not stripe data across members of a rank as in many other RAID implementations. Instead, each data volume emulates a complete 3380 or 3390 device mapped to a mainframe host, or a complete FBA logical volume mapped to an open systems host. This methodology allows the group to sustain the loss of more than one disk drive and still service requests from all the surviving disk drives. In conventional parity-based RAID implementations that stripe data, the loss of more than one member would result in loss of all data in the entire group.

Parity RAID—Parity Volume

A parity volume is a logical volume that holds the parity information for the rank. It must be the same size as the data volumes it supports. Parity volumes do not have unit addresses and are transparent to the host software. As with the secondary (M2) volumes in a mirrored Symmetrix system, parity volumes are not included in the total device limit within a single Symmetrix system.

Parity RAID Group

A Parity RAID group consists of the physical disk devices within the Symmetrix unit that are related to one another for common parity protection. The Parity RAID group is defined within the Symmetrix system, and includes disk volumes that are designated as either data volumes or parity volumes.

Symmetrix systems employ four-processor Fibre Channel disk directors. These disk director boards are divided into an *a* side, *b* side, *c* side, and a *d* side, with each side supporting two Fibre Channel paths (labeled *A* and *B*) (Figure 5-3 on page 5-27). For optimum availability and redundancy, members of a RAID group are configured to span across multiple disk directors. The overriding configuration requirement is that each member of a RAID group must be on a different Fibre Channel path on the back end of the system.

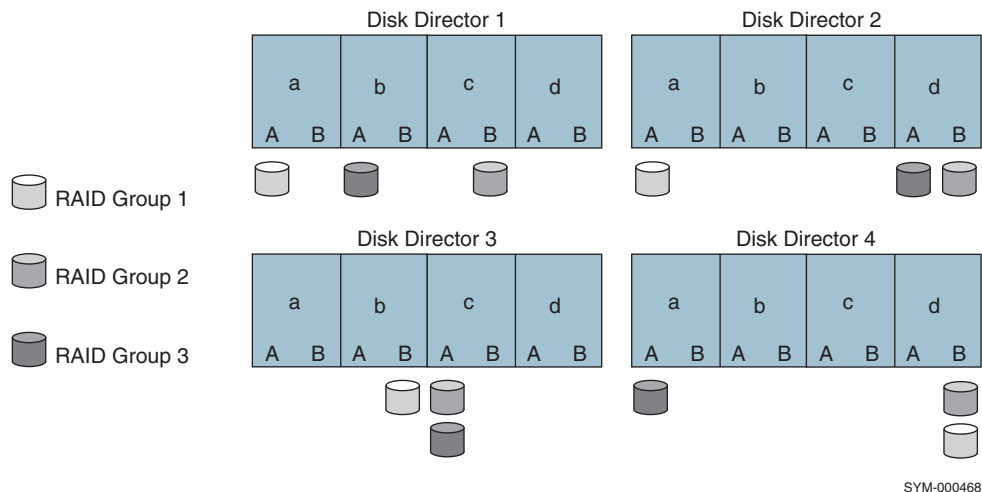


Figure 5-3 Symmetrix Parity RAID Group Disk Director Redundancy

Parity RAID Rank

A Parity RAID rank is the set of logical volumes related to each other for parity protection. Each Parity RAID group supports a minimum of one rank and, with hypervolume extension (HVE) enabled, a maximum of 58 ranks for Parity RAID (3+1) configurations and a maximum of 39 ranks for Parity RAID (7+1) configurations. Figure 5-4 shows a Parity RAID group consisting of four Symmetrix drives with one Parity RAID (3+1) rank defined across the group.

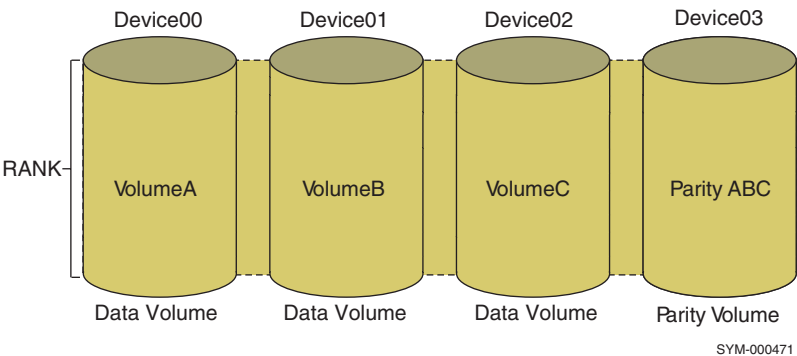


Figure 5-4 Parity RAID Group without Hypervolumes

Parity RAID (3+1)

A Parity RAID (3+1) configuration consists of Parity RAID ranks containing three data volumes and one parity volume. With this approach, effectively 75 percent of the total storage capacity of each Parity RAID group is available for storing data.

Parity RAID (7+1)

Parity Raid (7+1) with less than four DAEs will not be supported in Enginuity 5670.

A Parity RAID (7+1) configuration consists of Parity RAID ranks containing seven data volumes and one parity volume. With this approach, effectively 87.5 percent of the total storage capacity of each Parity RAID group is available for storing data.

Multiple Symmetrix RAID groups with one configuration can coexist on the same Symmetrix unit. Symmetrix RAID (3+1) groups and RAID (7+1) groups cannot be configured in the same Symmetrix system, although each can be configured together with mirrored volumes. All logical volumes participating in a Symmetrix RAID group must have identical storage capacity. A Symmetrix system allows intermixing of different capacity physical disk devices within a single Symmetrix unit.

Hypervolume Extension (HVE) with Parity RAID

Symmetrix HVE is supported with Parity RAID. When using HVE, parity and data volumes are distributed among the members of a Parity RAID group, as shown in Figure 5-5.

For more information on the Symmetrix HVE feature, refer to the hypervolume sections for mainframe and open systems environments sections in Chapter 4.

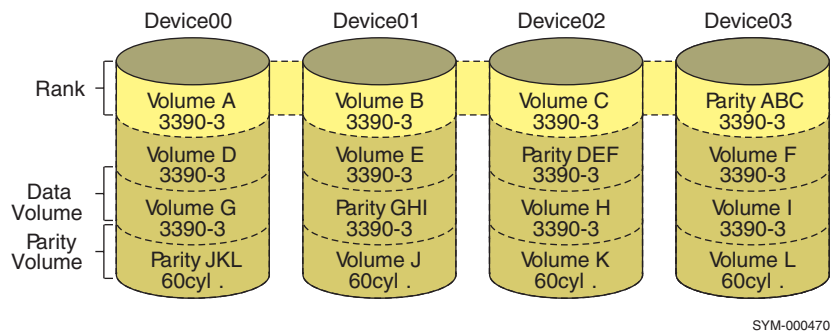


Figure 5-5 Parity RAID Group with Hypervolumes

HVE not only allows multiple logical volumes to be allocated to one physical volume, but those logical volumes that are members of a Parity RAID rank can be distributed across the multiple physical drives. The parity volume for each rank can reside on any physical volume within the Parity RAID group, as long as it is a different physical volume from those that compose the data volumes of that rank.

This distributed parity provides improved performance over a single physical volume, which could become a performance bottleneck in a heavy write workload. It also provides the performance advantages normally associated with RAID 5 *striping*, without the disadvantages of data volume striping.

When using HVE and Parity RAID, all volumes that compose the Parity RAID ranks of a physical Parity RAID group must be identical in format (all 3390-1, 3390-3, 3390-9, 2105, or all FBA.) However, separate Parity RAID physical groups on the same Symmetrix system can support different hypervolume configurations.

Parity RAID Modes of Operation

This section describes the following Parity RAID modes of operation:

- ◆ Normal Mode/Parity Generation
- ◆ Reduced Mode/Regeneration
- ◆ Non-RAID Mode
- ◆ Parity Rebuild

Normal Mode/Parity Generation

When a Parity RAID rank operates with all data and parity volumes functioning, it is operating in normal mode. In normal mode, the Symmetrix system accomplishes data redundancy by using the standard Parity RAID EXCLUSIVE OR (XOR) logic to generate and store XOR parity data that can then be used to reconstruct the data of a failed drive. In parity generation, a parity volume is initially formed by performing an XOR calculation on the contents of all member data volumes and writing the resulting parity to the parity volume. This bit-by-bit parity generation is illustrated in the first example in Figure 5-6 on page 5-30.

Reduced Mode/Regeneration

When a Parity RAID rank operates with one failed data volume, it is running in reduced mode. Referring to Figure 5-6 on page 5-30, the failure of Device00 would force the rank to operate in reduced mode. In Figure 5-5 on page 5-28, the failure of Device00 would cause the first three ranks to operate in reduced mode.

The data on the failed volume is reconstructed by XORing the parity volume with the remaining data volumes in the same rank. This process is called *regeneration*. Regeneration is shown in the bottom two examples of Figure 5-6 on page 5-30.

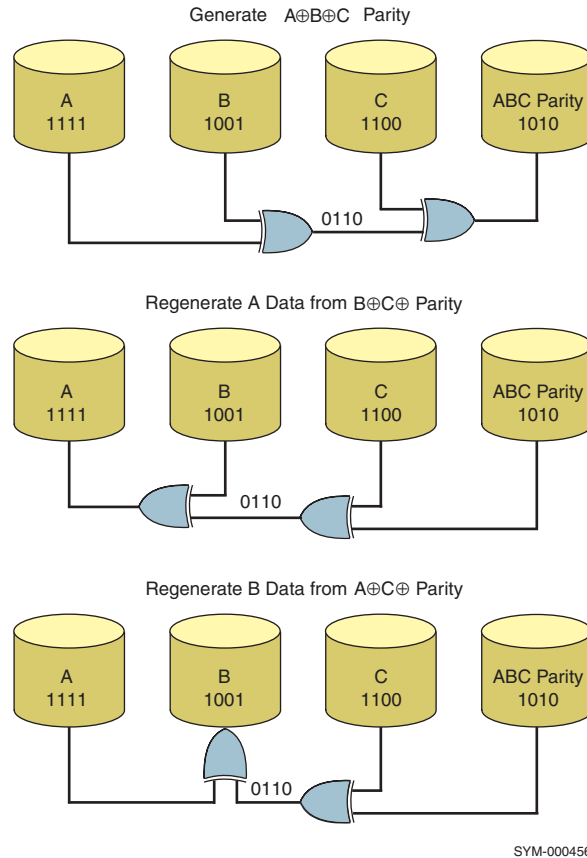


Figure 5-6 Parity Protection Logic

Non-RAID Mode

When a Symmetrix Parity RAID rank is operating with one failed *parity* volume, it is said to be running in non-RAID mode. As in reduced mode, parity protection is suspended for the rank. Again referring to Figure 5-5 on page 5-28, the failure of Device00 would cause the fourth rank to operate in non-RAID mode.

Parity Rebuild

When a parity volume fails, Symmetrix Parity RAID protection is suspended for the rank. When the failed device is replaced as part of a service action, the parity volume is reconstructed. This process is called *parity rebuild*.

Writing, Reading, and Rebuilding Data with Parity RAID

The Symmetrix Parity RAID write process is significantly different from other parity-based RAID implementations. The typical RAID implementation requires that the controller execute four discrete sequential I/O operations:

- ◆ Read old data
- ◆ Read old parity
- ◆ Write new data
- ◆ Write new parity

In addition, two processing steps must be executed by the controller microcode:

- ◆ XOR old data with new data (creating difference data)
- ◆ XOR difference data with old parity (creating new parity)

In contrast, Symmetrix Parity RAID requires that the controller execute only two discrete sequential I/O operations:

- ◆ Write new data
- ◆ Write difference data

The design of Symmetrix Parity RAID distributes the work of computing parity between the disk director and the disk drives, using the XOR chip located on the disk drive and the disk level buffer. The disk containing the data volume reads the old data, XORs it with the new data to compute difference data, and sends the difference data to the disk director, which then places the data in global memory. The disk containing the parity volume reads the old parity, XORs it with the difference data received from the global memory, and writes the new parity to the disk.

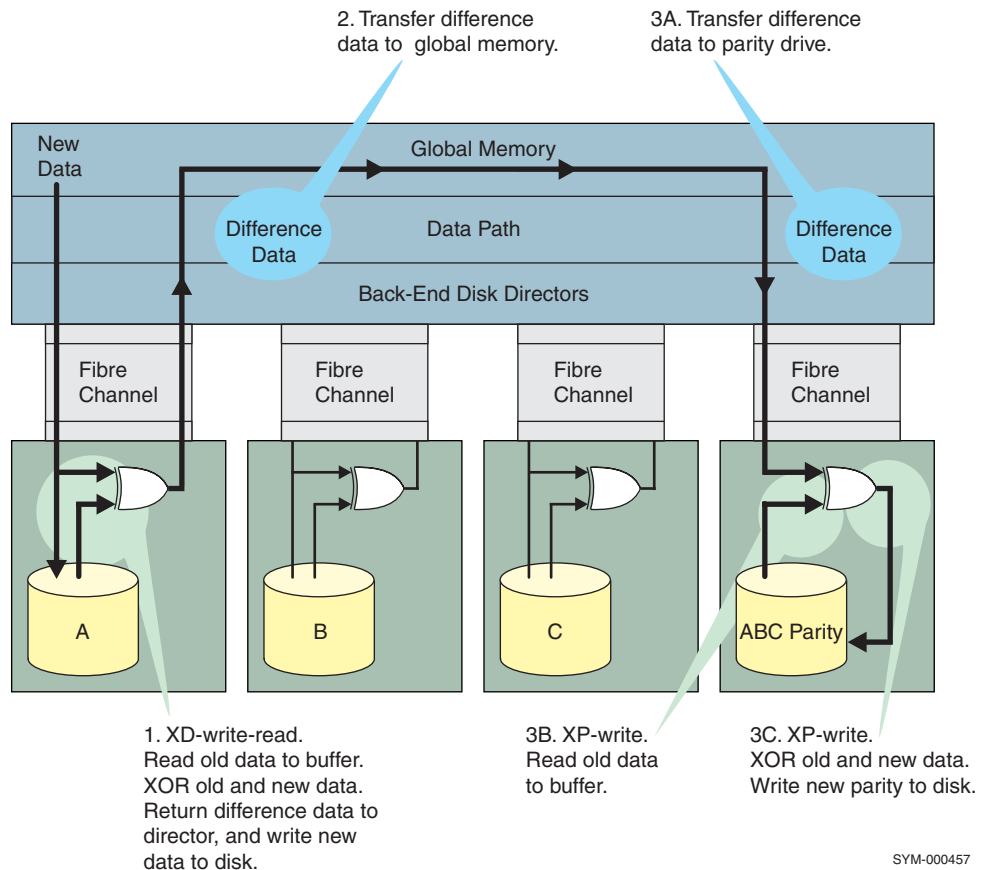
The parallelism introduced into the parity computation process through the use of XOR drives allows the disk director to do only half the number of back-end I/Os as do competitive RAID solutions. This reduces the impact of the write penalty significantly and improves the overall performance of Symmetrix Parity RAID compared to competitive implementations.

Writing Data in a Parity RAID Group

As with all Symmetrix operating modes, 100 percent of writes are fast writes and are satisfied in global memory. In a destaging write operation, the channel director presents a channel end/device end (or a good ending status) message to the host after data is written and verified in global memory.

In the Parity RAID write process, performing the *read old data* and *XOR* functions at the disk device level reduces the disk director's operations to a single read (*difference data*) and two writes (new data to the data volume and difference data to the parity volume). Figure 5-7 illustrates how data is destaged to disk through the following sequence of Fibre Channel commands used in the Parity RAID write process:

- ◆ XD-write-read
- ◆ XP-write



SYM-000457

Figure 5-7 Symmetrix Parity RAID Write Operation

Reading Data in a Parity RAID Group

Read hit operations in a Symmetrix Parity RAID group are processed through global memory as in normal Symmetrix system processing. Read misses are directed to the disk drive and processed as normal Symmetrix system read misses. There is no XORing of the data, and only one disk drive is involved in servicing the request.

This process offers an advantage over other implementations that stripe data across multiple disk drives. In these implementations, more than one disk drive may be required to service the request, resulting in possible contention when multiple I/Os are serviced simultaneously.

Data Recovery with Parity RAID

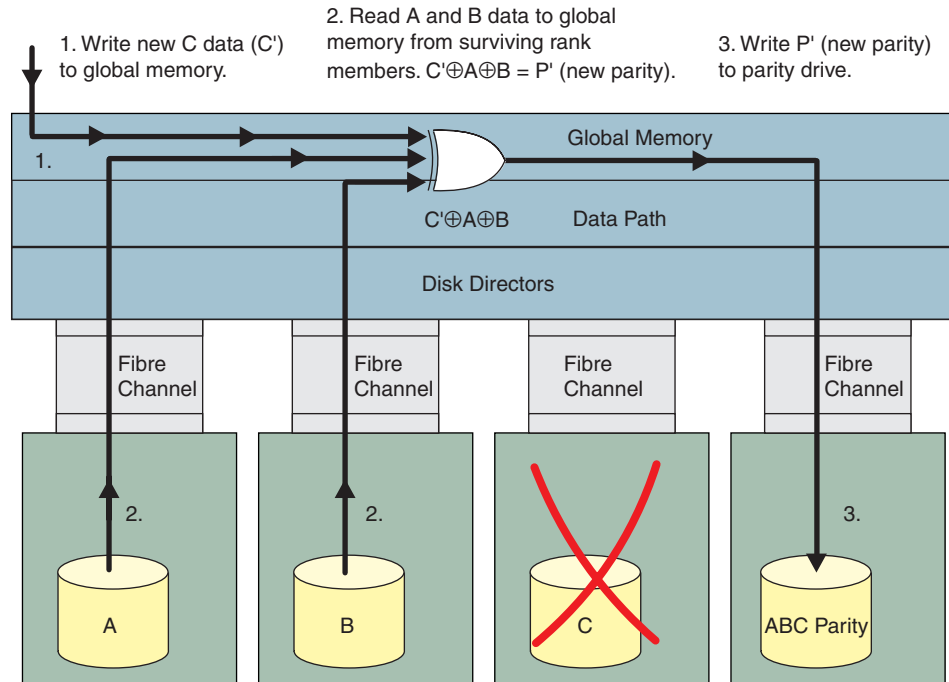
Symmetrix Parity RAID provides continuous availability for all data in a Symmetrix Parity RAID group if any single physical device in the group that contains the rank fails or becomes unavailable. If a media error occurs, the affected tracks will be regenerated. When a device within the Symmetrix Parity RAID group fails or becomes unavailable, the Symmetrix Parity RAID group is put in reduced mode (Figures 5-8 and 5-9). These volumes will now serve all their I/O requests as standard Symmetrix Parity RAID data volumes. All data is still available to the host, but is unprotected against additional failures unless protected by dynamic sparing.

If a physical device reports too many errors, or fails completely, that volume will be taken offline by the Symmetrix subsystem. This condition causes the Symmetrix system to place a remote service call to the Customer Support Center. The Product Support Engineer at the Support Center determines if a disk drive has been identified for replacement and dispatches a Customer Engineer (CE) who starts the Symmetrix interactive drive replacement process. The logical volumes on the disk being replaced are placed in a not-ready state and the associated ranks will begin operating in either reduced mode or non-RAID (depending on whether the logical volume that was made not-ready contained data or parity information). After the new drive is in place, the rebuild process begins.

If the Symmetrix system is configured with dynamic sparing, it copies the data from the failing volume to the spare, reconstructing the data, if necessary, from parity. The Symmetrix system also invokes available spares for the remaining volumes in the Symmetrix Parity RAID group, if they are available. For more information on dynamic sparing used with Symmetrix Parity RAID, refer to *Dynamic Sparing with Remotely Mirrored Pairs (SRDF)* on page 5-45.

Write Requests on a Failing Volume

As write requests are made to the failing data volume, the new data is XORed with the old data and the new parity is written to the parity drive. The Symmetrix system follows the process shown in Figure 5-8.

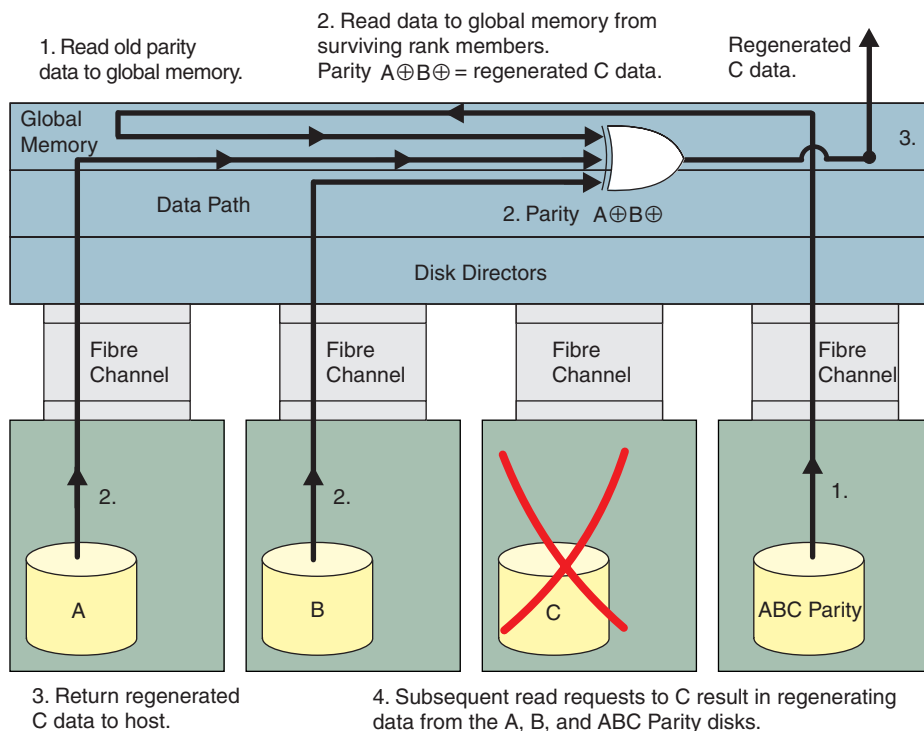


SYM-000463

Figure 5-8 Writing to a Symmetrix Parity RAID Group in Reduced Mode

Read Requests on a Failing Volume

Read requests for data in global memory (read hit) are serviced immediately from global memory. For read requests not found in global memory (read misses), as a read request to access data from the failing volume is received, the data is automatically calculated in global memory from the XOR of the parity volume and the remaining active data volumes (Figure 5-9 on page 5-35). Then, from global memory, the regenerated data is delivered to applications that request it. If the regenerated data received frequent read requests, it will remain in global memory according to the most recently used (MRU) global memory tables. If the data is not in global memory, it is again recalculated from the remaining active drives and delivered to the applications that request it.



SYM-000462

Figure 5-9 Reading from a Symmetrix Parity RAID Group in Reduced Mode

Data Volume Function when a Parity Volume Fails

If the parity volume fails, the active data volumes within the Symmetrix Parity RAID group stop creating parity and function as standard data volumes in non-RAID mode. Because data associated with individual logical volumes is not striped across multiple volumes, no single drive depends on the availability of any other drive in the Symmetrix Parity RAID group to provide access to its own data. All volumes on the remaining drives are available.

Data Recovery with Symmetrix Parity RAID and HVE

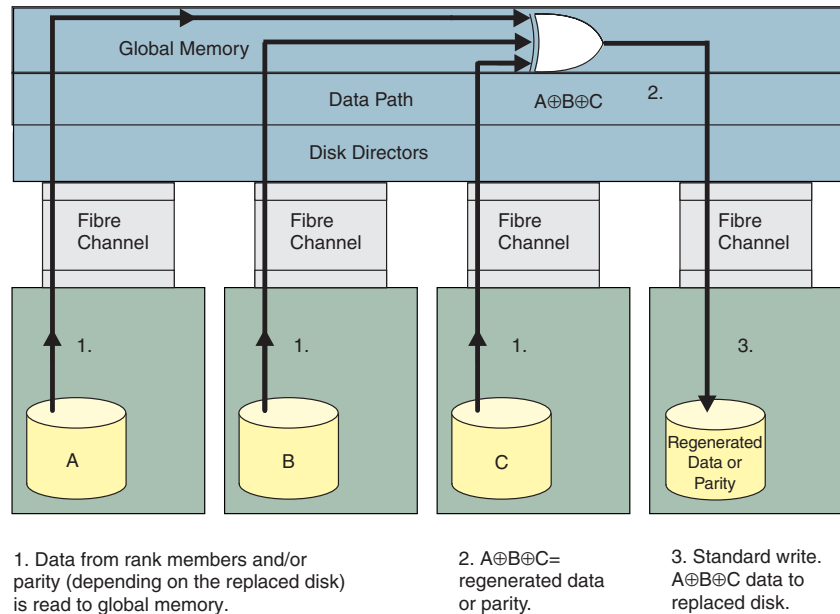
Data recovery in Symmetrix Parity RAID when using HVE is logically identical to Symmetrix Parity RAID recovery without HVE. For example, in Figure 5-5 on page 5-28, assume that volume B starts exceeding error thresholds and is about to fail. The Symmetrix data protection mechanisms then decide to remove volume B. Parity is immediately stopped within the entire Symmetrix Parity RAID group, and all subsequent requests to read or write to volume B will be served by the parity ABC volume, as described in *Data Recovery with Parity RAID* on page 5-33.

Regenerating Data or Parity after Replacing a Disk

As read or write requests are made to volumes E and J, data is regenerated from the data and parity volumes in their respective ranks. No action needs to be taken on parity volume GHI, because it is now ignored. After the physical drive on which volumes B, E, and J resided is replaced, the data in these volumes is regenerated, and then the parity volumes rebuilt. Parity protection is now available.

After the failed disk device is replaced, data is regenerated for the data volumes or parity is rebuilt for the parity volume. For the data volume(s) on the replaced disk device, data is regenerated from the parity volume and the remaining data volumes in the Symmetrix Parity RAID group. For the parity volume on the replaced disk device, parity is generated from the Symmetrix Parity RAID group data volumes (Figure 5-10). After this process is completed, the Symmetrix system reestablishes parity protection for all volumes of the Symmetrix Parity RAID group and resumes normal mode operation.

Replace failed drive. For a replaced parity drive, rebuild parity from the data volumes. For a replaced data drive, regenerate data from the surviving data drives and the parity drive.



SYM-000461

Figure 5-10 Regenerating Data or Parity after Disk Replacement

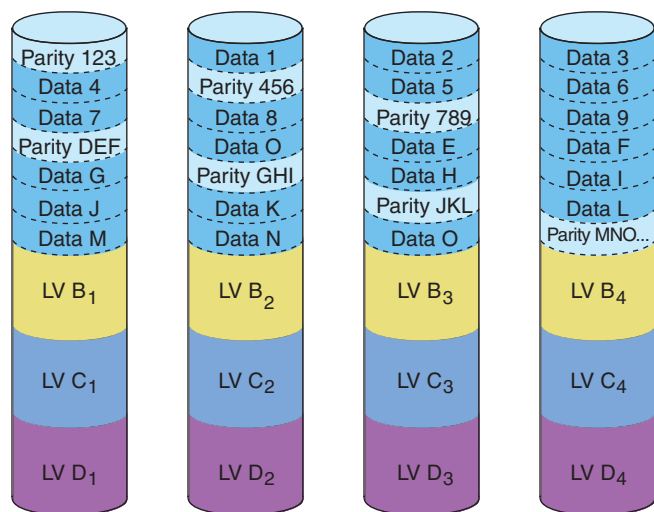
Symmetrix DMX800 RAID 5 Options

RAID 5 Overview

Enginuity level 5670 supports RAID 5 data protection. A RAID 5 device has the following attributes:

- ◆ Data blocks are striped (or interleaved) horizontally across the members of a RAID 5 group, similar to a striped metavolume. Each member owns some data tracks and some parity tracks.
- ◆ There is no separate parity drive in a RAID 5 group. Instead, parity blocks rotate amongst the group members.
- ◆ RAID 5 groups can be:
 - Four members per logical device, RAID 5 (3+1)
 - Eight members per logical device, RAID 5 (7+1)

Figure 5-11 shows a RAID 5 group composed of four hypervolume physical drives.



SYM-000547

Figure 5-11 RAID 5 Data/Parity (3+1)

Temporary Member Sparing

Unlike Dynamic Sparing in the case of a Parity RAID volumes (refer to *Dynamic Sparing with Symmetrix Parity RAID Volumes* on page 5-43), the RAID 5 spare is a part of the RAID group and *replaces* the failing member. In particular if you look at the tracks on the spare you would see some data tracks and some parity tracks; the entire layout would be identical to that of the failing member. Enginuity supports a *direct copy* mode between the spare and the failing member. For example, once the failing drive is replaced, the Enginuity code will avoid rebuilding the data to the new drive. Instead, it will copy the tracks (both data and parity) directly from the spare to the new drive. This is a much faster way of validating the new drive. The Symmetrix system has a sense of *copy direction*. When a spare is invoked, the default copy direction is from the failing member to the spare. The spare has a mirror number of its own that never changes.

Example

In an example configuration, the spare will be invoked as an M3 or as an M4. The invalid bits describing the state of the tracks on the spare will always be in the spare's mirror position, regardless of whether they are data or parity tracks.

RAID 5 Configuration Rules

When configuring RAID 5 devices, follow these rules:

- ◆ Only one spare can be invoked per RAID group
- ◆ When a spare is invoked against a member, the same track can be valid on either the member or the hot spare, not both at the same time
- ◆ Striped CKD volumes cannot be RAID 5 devices
- ◆ BCVs cannot be RAID 5 devices
- ◆ A virtual device cannot be a striped or a RAID 5 device
- ◆ While Parity RAID and RAID 5 are allowed in the same Symmetrix system, they are not allowed on the same device
- ◆ A Symmetrix RAID 5 group can contain four or eight members, as in Parity RAID (3+1) and Parity RAID (7+1)
- ◆ SymmOptimizer Swap support for RAID 5 devices
- ◆ SRDF devices are allowed to be RAID 5 devices
- ◆ eSNAP source and destination devices are allowed to be RAID 5 devices
- ◆ A log device is allowed to be a RAID 5 device
- ◆ Concurrent Copy is supported with RAID 5

Symmetrix RAID 10

To improve mainframe volume performance, Symmetrix RAID 10 stripes data of a logical device across multiple Symmetrix logical devices. (The feature is analogous to metavolume stripes on open systems.) Four Symmetrix devices (each one-fourth the size of the original mainframe device) appear as one mainframe device to the host. Any four Symmetrix logical devices can be chosen to define a RAID 10 group provided they are the same type (for example, IBM 3390) and have the same mirror configuration. Striping occurs across this group of four devices with a striping unit of one cylinder, as shown in Figure 5-12. Since each member of the stripe group is mirrored, the entire set is protected. Dynamic Mirror Service Policy (DMSP) can then be applied to the mirrored devices. The combination of DMSP with mirrored striping and concatenation to create a mainframe volume, as illustrated below, enables greatly improved performance in mainframe systems. RAID 10 uses four pairs of disks in its Symmetrix implementation, as illustrated below.

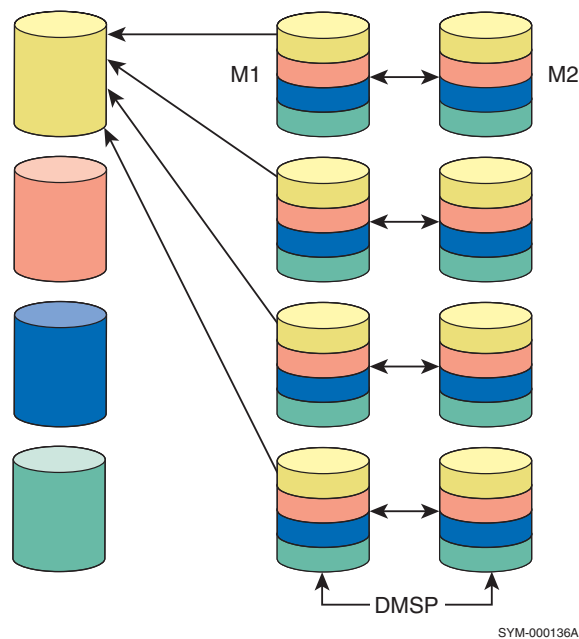


Figure 5-12 RAID 10 with Dynamic Mirror Service Policy

Dynamic Sparing

Symmetrix systems have a dynamic sparing option that reserves volumes as standby spares. These volumes are not user-addressable. The dynamic sparing function determines when a logical volume is about to fail and copies the contents of the disk device on which that volume resides to an available spare without any interruption in processing (Figure 5-13). The Symmetrix system notifies the EMC Customer Support Center of this event with an Environmental-Data Present error, and then uses the spare until the device on which the original logical volume resides can be replaced. The dynamic sparing option increases data availability without impacting performance. It is used in combination with mirroring, Parity RAID, or SRDF.

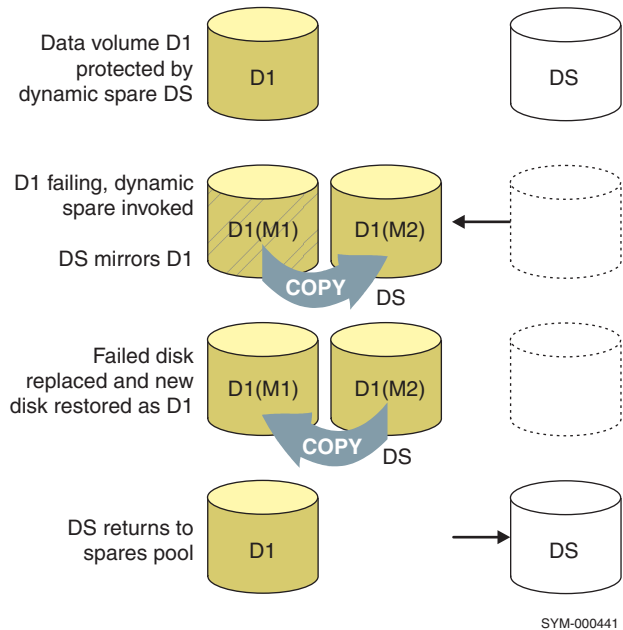


Figure 5-13 Dynamic Sparing Process

EMC Customer Service is usually able to replace a failed drive within four hours. Dynamic sparing is usually used in remote locations where EMC personnel cannot guarantee this service-level agreement. The time it takes to replace and resynchronize a failed drive within a Parity RAID group depends on the I/O activity to the logical volumes, the disk device, and the disk capacity.

When dynamic sparing is enabled, the Symmetrix system makes its copying decision based on error statistics maintained by its directors, the intelligent disk microprocessor self-testing information, and its active error checking system. If the Symmetrix dynamic sparing algorithms determine that the number of errors occurring on a volume is excessive and that a hard failure is probable, it looks for an available spare in its small pool of spares.

The Symmetrix system dynamically copies all data from the device containing the failing volume to the available spare. The Symmetrix system continues to process host I/O requests at the highest priority while this copy operation takes place to minimize the effect on performance. When the copy operation completes, the Symmetrix system notifies the EMC Customer Support Center of the event. The spare and the original device work as a mirrored pair until the defective unit is replaced.

After the disk is replaced, the EMC Customer Engineer (CE) then issues commands for the Symmetrix system to dynamically copy the contents of the spare to the new device. The spare remains in use until the copy completes. When the copy has completed, the CE issues commands for the Symmetrix system to return the spare to its pool, making it available should another volume fail in the future.

Dynamic sparing has the added benefit of allowing an EMC Customer Engineer operating on site or from the EMC Customer Support Center to force the Symmetrix system to copy the contents of a device to an available spare. This forced copy allows the EMC Customer Engineer to test or service a disk device nondisruptively while access to data continues.

In summary, dynamic sparing offers the following advantages:

- ◆ Increases protection of all volumes from loss of data
- ◆ Automatically activates the spare volume without interruption prior to loss of access of a potentially failing volume
- ◆ Ensures that the spare copy is identical to the original copy
- ◆ Resynchronizes a new disk device with the dynamic spare after repair of the defective device is complete
- ◆ Increases data availability of all volumes in use without loss of any data capacity
- ◆ Dynamic sparing is transparent to the host and requires no user intervention

Dynamic Sparing as Additional Protection

Dynamic sparing is used as additional protection for volumes already protected by the Symmetrix mirroring, RAID, or SRDF options. Dynamic sparing provides incremental protection against failure of a second disk during the time between when a disk is taken offline and when it is ultimately replaced and resynchronized.

Dynamic Sparing with Locally Mirrored Pairs

When a dynamic spare is invoked for a locally mirrored pair, the Symmetrix system automatically augments the original mirrored pair with a dynamic spare volume that joins the mirror pair as an additional or third mirror (Figure 5-14). Data is copied to the dynamic spare volume from the failing volume. If any data cannot be copied from the failing volume, it is copied from the other mirror. The Symmetrix system continues processing I/Os with the spare functioning as a mirror with no interruption in operation.

The failing disk can then be replaced and resynchronized with the mirror group. Then the dynamic spare can be returned to the spare pool.

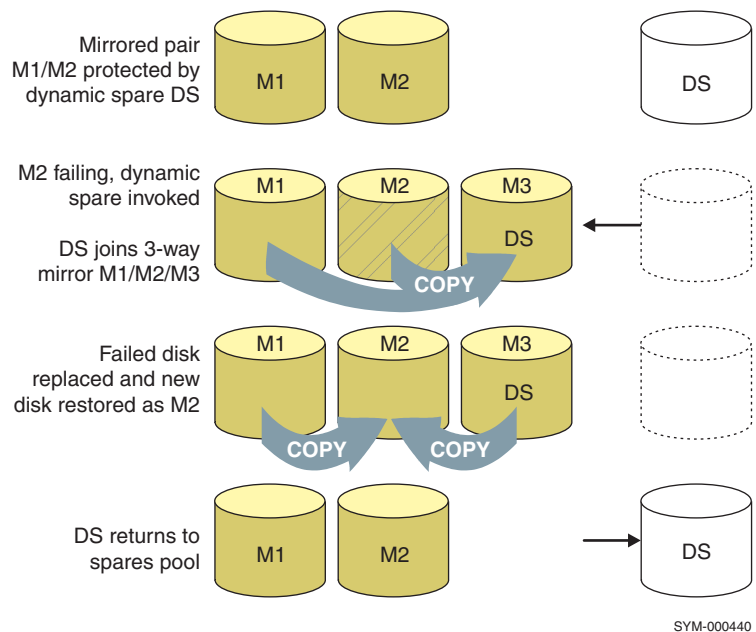


Figure 5-14 Dynamic Sparing with Locally Mirrored Pairs

**Dynamic Sparing
with Symmetrix
Parity RAID Volumes**

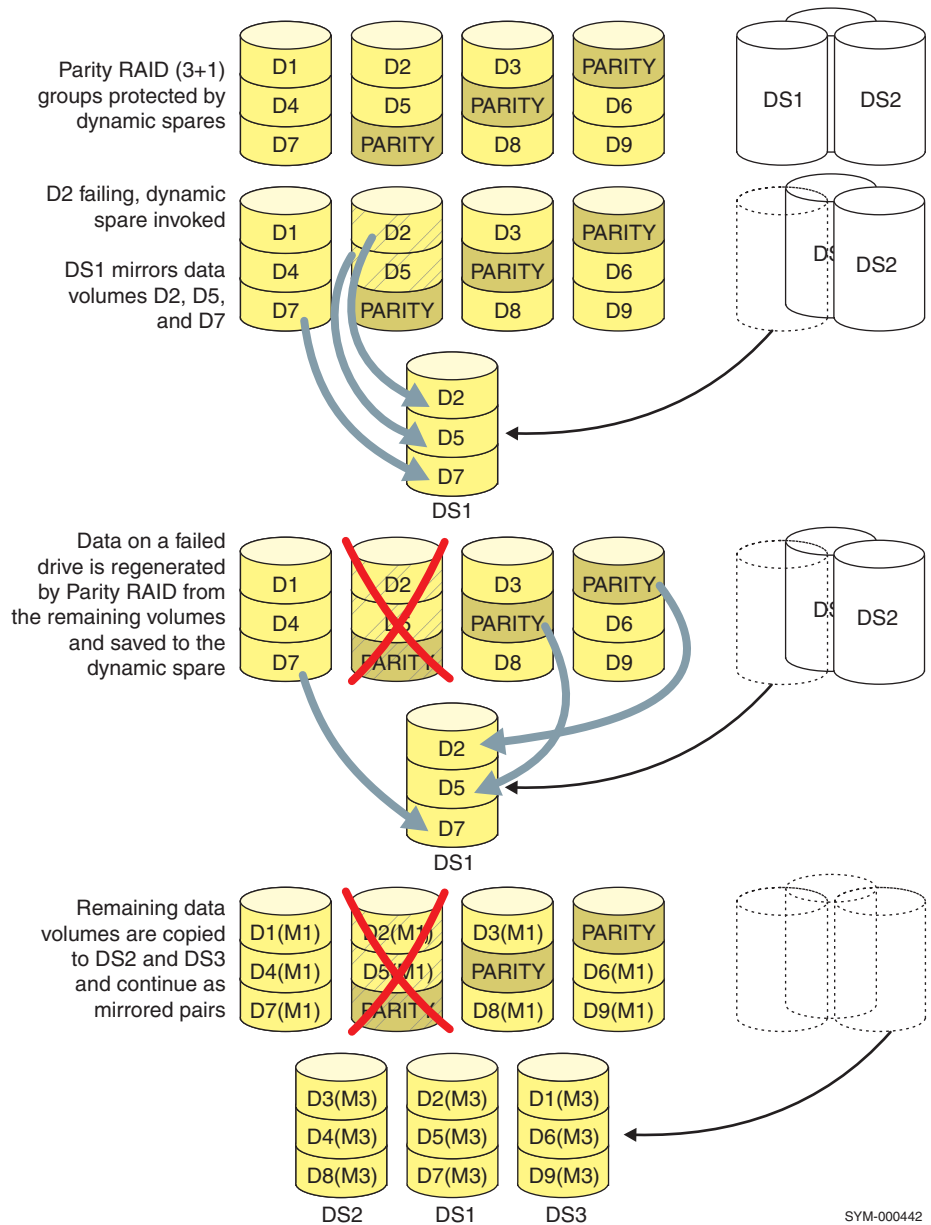
In a Symmetrix Parity RAID system, all data volumes of the Symmetrix Parity RAID group will be spared if there are enough dynamic spares available for all the data volumes. If a device in a Symmetrix Parity RAID group fails, the Symmetrix system will attempt to copy data from the failing device to the first spare (Figure 5-15 on page 5-44). If the system cannot copy from the failing device to the spare, the Symmetrix system then uses the parity algorithms to reconstruct the data for the failing device and then copies the data to the spare. As the failing volume receives read or write requests, the data is regenerated. Regenerated read data is sent to the requesting application. Write data is regenerated as described in Figure 5-8 on page 5-34. The Symmetrix system also invokes available spares for the remaining data devices in the Symmetrix Parity RAID group, allowing all volumes in the Parity RAID group to function as mirrored pairs.

**Data Device Failure
Within a Parity RAID
Group**

If the failing data volume becomes not ready before it can be replaced, the Symmetrix system recalculates the data for the failed device from the remaining data devices and parity volume. The dynamic spare functions as a mirrored pair for that data volume. Parity RAID protection is not available for the remaining volumes of the Parity RAID rank until the failing device is replaced. All volumes of the Parity RAID group are spared if there are enough dynamic spares available. If there are not enough spares for the remaining data volumes in the group, the Symmetrix system invokes as many dynamic spares as possible for the Parity RAID group.

**Parity Device Failure
Within a Parity RAID
Group**

Upon detecting that a parity volume is failing, the Symmetrix system turns off parity protection, and mirrors all data volumes to available spares. Parity RAID protection is not available until the failing device is replaced. If there are not enough spares for the data volumes in the group, the Symmetrix system will invoke as many dynamic spares as are available for the Parity RAID group.



SYM-000442

Figure 5-15 Example of Dynamic Sparing with RAID (3+1) Volumes

Dynamic Sparing with Remotely Mirrored Pairs (SRDF)

When the dynamic sparing option is invoked for a remotely mirrored SRDF pair, the Symmetrix system automatically activates an available spare in the Symmetrix unit containing the failing device and copies data from the failing device to the spare. The Symmetrix system continues processing I/Os (with the spare functioning as one of a mirrored pair with the failing device), and its remote mirror continues with no interruption in operation. If the Symmetrix system cannot copy all data from the failing device to the spare, it retrieves the unavailable data from the *good* member of the remote pair.

For examples of dynamic sparing in SRDF configurations, refer to the *Symmetrix Remote Data Facility Product Guide*.

Mainframe Features and Support

The information presented in this chapter is only applicable to Symmetrix disk devices connected to mainframe hosts. The following information is described:

- ◆ Introduction6-2
- ◆ Supported Mainframe Features6-3
- ◆ Error Reporting and Recovery6-11
- ◆ Sense Byte Information6-21

Introduction

This chapter provides an overview of the Symmetrix mainframe storage solutions supported by the Enginuity operating environment. Also included are descriptions of the types of errors possible when the Symmetrix system is connected to a mainframe host, error handling techniques, and an error recovery summary.

Supported Mainframe Features

The following IBM/PCM functions and features are supported on Symmetrix DMX800 systems that have connections to mainframe hosts:

- ◆ Dynamic Channel Management
- ◆ Dynamic Path Reconnection
- ◆ Concurrent Copy
- ◆ MultiPath Lock Facility/Concurrent Access
- ◆ Multi-Subsystem Imaging
- ◆ Sequential Data Striping
- ◆ Partitioned Data Set (PDS) Assist
- ◆ Multiple Allegiance (MA)
- ◆ Compatible Parallel Access Volumes (COM-PAV)
- ◆ Dynamic Parallel Access Volumes
- ◆ RAID 10 Striping
- ◆ Logical Paths per FICON Port and Control Unit Image Support
- ◆ FICON Cascading and Open Systems Intermix
- ◆ PPRC Command Support
- ◆ XRC Support (level 3)
- ◆ Configuring CKD Volumes
- ◆ Deleting (and then Adding) Devices Online
- ◆ Support for 64K Cylinders

Dynamic Channel Management

The Symmetrix DMX800 supports the Dynamic Channel Management (DCM) feature of the Intelligent Resource Director function inherent in IBM zSeries mainframe processors. DCM allows dynamic reconfiguration of channels between control units under the direction of the Workload Manager component of z/OS. Support for DCM requires the installation of a software module on the z/OS operating system that describes a control unit's internal structure. This module is called IOSTEMC and is available at the EMC FTP site:
<ftp://ftp.emc.com/pub/MVSsoft/DCM-IOSTEMC>

Dynamic Path Reconnection

Dynamic Path Reconnection (DPR) permits the storage control unit (SCU) to reconnect to the host on any available channel path between the device and the host system if the original channel is busy with other operations. Without DPR, the SCU waits for the original channel path to become available again. Refer to the *IBM 3390 Direct Access Storage Introduction* or the *IBM 3380 Direct Access Storage Introduction* for more information on this function.

The DPR option must be invoked in a FICON environment to facilitate reduction of director port busy conditions. DPR must also be enabled when using extended platform functions, such as IBM's Concurrent Copy.

DPR support is enabled by the EMC Customer Engineer at installation or service time. Consult your EMC Systems Engineer to determine if DPR is appropriate for your operating environment.

Concurrent Copy

Symmetrix systems support the IBM Concurrent Copy facility. Concurrent Copy can significantly reduce the time that data on the Symmetrix system volumes is unavailable during backup operations.

To use Concurrent Copy, the Symmetrix system must be emulating 3990-6, 3990-3, or 2105 storage control, and 3390 DASD with Dynamic Path Reconnect (DPR) and Symmetrix Differential Data Facility (SDDF) must be enabled.

MultiPath Lock Facility/Concurrent Access

Symmetrix systems support the IBM MultiPath Lock Facility/Concurrent Access (MPLF/CA) feature for use with the ultra-high performance Airline Control Program (ACP) and Transaction Processing Facility (TPF) host operating system environments. MPLF/CA allows multiple concurrent I/O requests to the same logical device from multiple TPF mainframes.

The Symmetrix system maintains the names and status of logical locks currently in use and responds to requests to obtain or release a lock. This allows multiple hosts to share DASD through multiple paths in an active OnLine Transaction Processing (OLTP) environment while maintaining data integrity. MPLF/CA is an enhancement and replacement for the Extended Limited Lock Facility (ELLF) and the Limited Lock Facility (LLF). The Symmetrix system must be emulating the 3990-6, 3990-3, or 2105 storage control and running Enginuity supporting the MPLF/CA feature.

Multi-Subsystem Imaging

The Symmetrix DMX800 systems support multiple z/OS environments by providing maximum connectivity through the use of its 3990-3, 3990-6, and 2105 emulation modes and the hypervolume extension feature. The Symmetrix systems support up to 16 SubSystem IDentifiers (SSIDs) with up to 256 devices per SSID.

Sequential Data Striping

The Symmetrix DMX800 systems are fully compatible with the IBM Sequential Data Striping function for 2105, 3990-6, and 3990-3 storage control with Extended Platform. Sequential Data Striping provides faster batch execution on large I/O-bound sequential processing requests by allowing I/O operations to be managed in parallel across as many as 16 devices.

Sequential Data Striping is available only in z/OS with DFSMS environments. The Symmetrix system must be emulating 2105 or a 3990-6 storage control and running Enginuity supporting this feature. Additionally, the Symmetrix system must have SMS-managed volumes.

Partitioned Data Set (PDS) Assist

The Symmetrix DMX800 systems support the IBM Partitioned Data Set (PDS) Search Assist feature for 2105, 3990-6, and z/OS storage control with Extended Platform. PDS Assist improves performance on large, heavily used partitioned datasets by modifying the directory search process. PDS Assist is automatically invoked with the appropriate level of DFSMS, and the Enginuity revision supporting this feature.

Multiple Allegiance (MA)

The Symmetrix DMX800 systems support Multiple Allegiance (MA), an IBM feature that improves throughput across a shared storage environment. MA allows different hosts to concurrently access the same device (have concurrent implicit allegiances), as long as I/Os do not conflict with each other.

The Symmetrix DMX800 systems support MA as a 2105 control unit if the COM-PAV feature is enabled on the Symmetrix system. The Symmetrix system can also support MA when the Symmetrix system is defined to the host as a 3990-6 by enabling this feature on the Symmetrix service processor.

Compatible Parallel Access Volumes (COM-PAV)

The Symmetrix DMX800 systems support Compatible Parallel Access Volumes (COM-PAV), an IBM feature that improves response time by reducing device contention, resulting in higher performance and throughput.

The Symmetrix DMX systems must be defined to the host as a 2105 control unit to support COM-PAV.

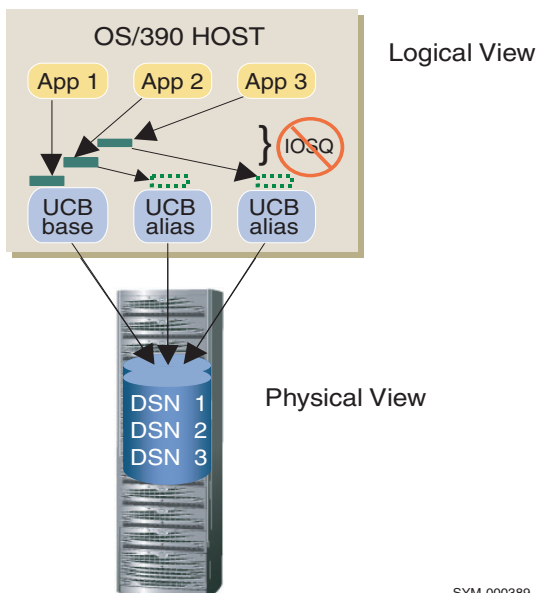
Dynamic Parallel Access Volumes

Parallel Access Volumes/Multiple Allegiance (PAV/MA) is a mainframe-exclusive feature that resolves the z/OS limitation allowing only one outstanding I/O operation to a device (Figure 6-1). The PAV/MA feature was introduced with the IBM ESS 2105 storage subsystem. EMC's COM-PAV/MA feature provides significant performance improvements for Symmetrix system users who are experiencing high levels of device queuing (high IOSQ time).

Enginuity adds dynamic support to the existing EMC PAV implementation. This enables the MVS Workload Manager (in goal mode) or the EMC PAVManage utility to reassign alias UCBs to a base UCB *on the fly*. In addition to this support for dynamic PAV, COM-PAV/MA now allows up to seven aliases to be associated with one base device, improving the opportunity for parallel I/O operations. This further reduces, and can even eliminate, IOSQ time and enables EMC's PAV support for larger mainframe logical devices.

Enginuity level 5670 supports adding and deleting PAV base and alias features online for FICON channels.

The EMC implementation is designed to be 100 percent compatible with IBM's implementation. The aliases can also be managed with the PAVManage utility in the ResourcePak Extended for OS/390.



SYM-000389

Figure 6-1 Dynamic Support of Parallel Access Volumes

RAID 10 Striping

RAID 10 (one-zero) is a mirroring feature with striping used for mainframe environments. Four Symmetrix devices (each one-fourth the size of the original IBM device) appear as one IBM device to the host, accessible by way of one channel address. Any four devices can be chosen to define a group provided they are equally sized, same type (for example, all 3390), and have the same mirror configuration.

For more information on RAID 10, refer to *Symmetrix RAID 10* on page 5-39.

Logical Paths per FICON Port and Control Unit Image Support

The Symmetrix DMX systems support up to the maximum of 4,096 channel addresses per FICON port. The two-port, two-processor FICON adapter supports up to 1,024 logical paths per port, per processor.

For more FICON director configuration information, refer to *FICON Mezzanine Cards* on page 2-29.

FICON Cascading and Open Systems Intermix

For information on FICON Cascading and Open Systems Intermix features, refer to *Additional FICON Features* on page 2-40.

PPRC Command Support

Enginuity enables the Symmetrix system to support native IBM Peer-to-Peer Remote Copy (PPRC) commands. PPRC is the remote copying solution available with IBM Storage Systems (Figure 6-2). PPRC is implemented as a subset of the dynamic SRDF feature. As a result, Symmetrix DMX800 systems running Enginuity level 5669 (or higher) will support PPRC commands and IBM's Geographically Dispersed Parallel Sysplex (GDPS) solution.

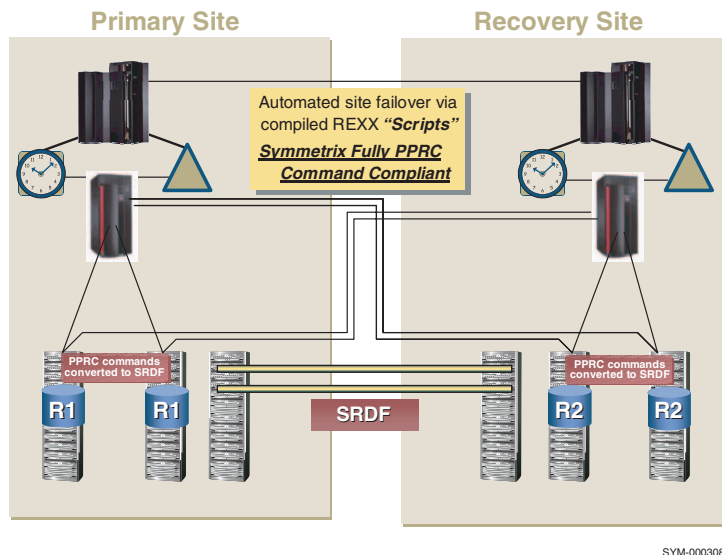


Figure 6-2 PPRC and GDPS Support

XRC Support

Engenuity code 5670 supports IBM's XRC level 3 functionality. IBM designed XRC as a host-based disaster recovery and workload migration solution that provides asynchronous copying of OS/390 data to a remote data center located virtually anywhere in the world.

With XRC level 3, a remote site has a point-in-time copy of the source/primary data with the following attributes:

- ◆ The copy at the remote site is not the most current one; there is a lag between the local and remote copies
- ◆ The remote copy's data is re-startable
- ◆ Since this is an asynchronous solution, write operations do not suffer from performance degradation while transferring the data across the link

XRC provides update sequence integrity through its mechanism of consistency groups, and journal, control, and state datasets. The basic concept is that the System Data Mover (SDM) polls, or receives attention interrupts from, the primary control units it is servicing and receives from them all new or updated records.

The SDM then determines the timestamp of the most recent I/O from each control unit in the session. Once these timestamps are determined, SDM selects the smallest of these timestamps as a starting point and forms a group which includes all I/Os received at that time or prior. This set of I/Os constitutes a *consistency group* which is written to the journal data sets and then written to the volumes at the recovery site. This is depicted in Figure 6-3.

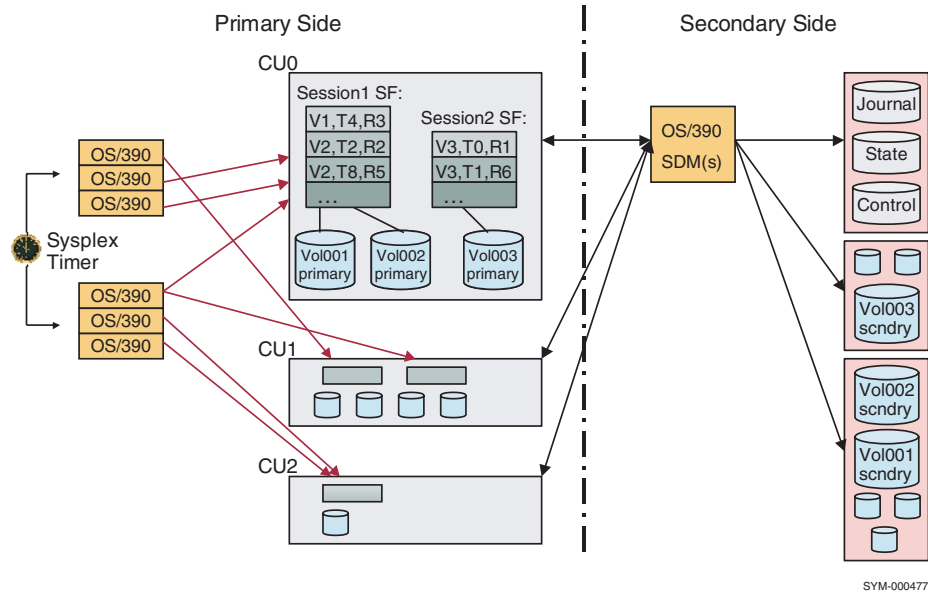


Figure 6-3 XRC Overview

Configuring CKD Volumes

Enginuity code 5670 supports mapping and unmapping CKD volumes. This capability is necessary because deleting mainframe capacity cannot take place if the volumes cannot be unmapped from a front-end port.

Deleting (and then Adding) Devices Online

Enginuity code 5670 supports removing and then adding devices online, which facilitates the following configuration enhancements:

- ◆ Change Device Emulation Online—Remove a CKD volume and add an FBA volume and vice-versa.

Adding a CKD volume to a Symmetrix system requires a global memory configuration change if this is the first CKD volume being added. However, the system can be configured in advance, thus avoiding an offline global memory reformat.

- ◆ Convert between mirrored and RAID protected volumes.

Support for 64K Cylinders

Symmetrix DMX systems running Enginuity code 5670 and higher can support 64K cylinder CKD devices for operating systems that can exploit it.

Error Reporting and Recovery

This section discusses the types of errors possible when the Symmetrix system is connected to a mainframe host. It also discusses error handling techniques and an error recovery summary.

Types of Errors

There are several types of errors associated with data storage. The error type is indicated in the sense information, and in error messages and reports. You should be aware, however, that the error type doesn't necessarily identify the source or the cause of the problem.

The error types detected by the Symmetrix system include the following:

- ◆ **Data Check**—The Symmetrix system has detected an error in the bit-pattern read from the disk. Data checks are due to hardware problems when writing or reading data, media defects, or random events.
- ◆ **System or Program Check**—The Symmetrix system has rejected the command. This condition is attributed to the construction of the channel program. This type of error is indicated to the processor and is always returned to the requesting program.
- ◆ **Overrun**—The Symmetrix system cannot receive data at the rate it is transmitted from the host system. This error indicates a timing problem. Resubmitting the I/O operation usually corrects this error.
- ◆ **Equipment Check** —The Symmetrix system has detected an error in hardware operation.

Temporary and Permanent Errors

Whenever the Symmetrix system detects a data or equipment error, either the Symmetrix system or the host operating system will attempt to recover from the error, depending on the situation and the type of hardware involved. Error recovery can be temporary or permanent.

An error is temporary if the Symmetrix system or the host operating system can recover from the error successfully. The application is not notified of a temporary error. Temporary errors recovered within the Symmetrix system are not listed in the error report unless its internal soft error threshold has been exceeded. Temporary errors recovered by the operating system are logged by the host Error Recovery Procedures (ERP).

When requested, the Symmetrix system generates sense bytes and sends them to the operating system. These sense bytes define the error. An error is permanent from a system view if neither the host operating system nor Symmetrix system can recover from the error condition. For example, a channel director may cause a data check error. That data check is permanent to the system and is recorded in the error recording dataset (ERDS) as a permanent path error.

However, if the host system retries the read from an alternate path, and the data is read successfully from the Symmetrix system, the system does not notify the application of the error, so the error is not permanent to the application. If the read on the alternate path was not successful, the data check is permanent from both the system and application points of view.

Recoverability by Error Type

An error is recoverable if the application does not see it as a permanent error.

Data Check

When data is written, the Symmetrix system records check bytes with the data to enable data check detection. These are the error checking and correction (ECC) bytes. These bytes often provide sufficient information for the Symmetrix system to reconstruct the data should an error occur. When the Symmetrix system reconstructs the data using the ECC bytes, the data check is ECC-correctable.

ECC-correctable data checks are always recoverable. The Symmetrix system relocates the block as necessary.

When the Symmetrix system cannot reconstruct the data using the ECC bytes, the data check is *ECC-uncorrectable*. In this case, the Symmetrix system retries the I/O operation. If the retry is unsuccessful, the data check is uncorrectable. This appears as a permanent error to the Symmetrix system.

For Symmetrix systems protected by mirroring, the data check is not reported to the application. The data is provided to the host through the protection mechanism (mirroring) and the Symmetrix system places a remote service call.

ICKDSF Control Statement

For an uncorrectable data check, on an unprotected Symmetrix system, use the following device support facilities (ICKDSF) control statement to re-format the home address and record zero of the track.

Do *not* use ICKDSF to assign alternate tracks in the classic IBM manner since no alternate will be assigned.

This control statement may require an operator response to permit purging of datasets on the volume. Refer to the *IBM ICKDSF User's Guide and Reference* for more information.



CAUTION

All data remaining on the track(s) operated on by this control statement will be *DESTROYED*. Restore the data from your most recent backup. Using any other means to reallocate data will affect data integrity.

Enter the following ICKDSF control statement to reassign the defective track(s):

```
INSPECT UNIT(ccuu) | SYSNAME(sysxxx) |
DDNAME (ddname)
```

```
DEVTYPE(3380/3390) VERIFY (serial) NOPRESERVE NOASSIGN
NOMAP -
```

```
TRACKS((cyl,head),...)
```

Specify the desired (cyl,head) in decimal notation, or as (X'cyl',X'head') if hex notation is desired.

Overrun Errors and Equipment Checks

When an overrun or equipment check is detected, the host system retries the operation. If the operation is successful on the retry, the error is recorded as recoverable. If the retry is unsuccessful, the error is recorded as unrecoverable.

Error Reporting

The Symmetrix system reports error conditions to the host and to the EMC Customer Support Center through its Autocall feature. The Symmetrix system presents a unit check status in the status byte to the channel whenever it detects an error condition such as a data check, command reject, overrun, or equipment check.

For error reporting specific to the Symmetrix Remote Data Facility (SRDF) environment, refer to the *Symmetrix Remote Data Facility Product Guide*.

Unit Check Status

The Symmetrix system also presents a unit check status (environmental-data present) whenever it detects an environmental violation. The Symmetrix system runs a series of internal tests on its components at least once every 24 hours and monitors critical components continuously. It also runs these tests when initially powered up or when a software re-set has occurred. These tests check for a low battery charge or AC power failure, or redundant component failure such as the failure of one device in a mirrored pair or the activation of dynamic sparing.

**Environmental Tests
Error Flag**

If, while running its environmental tests, the Symmetrix system detects an error condition, it sets a flag to indicate a pending error and presents a unit check status to the host on the next I/O operation. The Symmetrix system then schedules the test that detected the error condition to be rerun more frequently. (Each test has specific deltas to regulate its execution.) If a device-level problem is detected and reported, subsequent failures of that device are not reported until the failure is fixed. If a second type of failure is detected for a device while there is a pending error-reporting condition in effect, the Symmetrix system reports the pending error on the next SIO and then the second error.

Host Sense Data

When the Symmetrix system presents a unit check status, the host retrieves the sense data from the Symmetrix system and, if logging action has been requested, places it in the Error Recording Data Set (ERDS). The EREP (Environment Recording, Editing, and Printing) program prints the error information. The sense data identifies the condition that caused the interruption and indicates the type of error and its origin. The format of the sense data may be found in the appropriate IBM reference manuals. For interpretation of the EREP reports, the appropriate IBM manual should be consulted.

Service Alerts

The Symmetrix system reports the exception conditions listed in Table 6-1 to the host as Service Alerts. In 3990 controller emulation, the returned sense data is in SIM format. In 3880 emulation, environmental data is returned. In both cases, the code listed will be contained in sense bytes 22 and 23. These messages are reported across all logical paths to the device experiencing the error.

Table 6-1 Environmental Alert Messages

Sense Bytes 22-23	Meaning
E460	Dynamic sparing was invoked by a disk director.
1463	A disk director dual-initiator failed to IMPL monitor ('DD').
E466	Hot sparing was automatically invoked by a disk director for an RDF device in the other system.
E467	RDF initiated a SIM message.
E46D	All RDF links are not operational.
E46E	All RDF links are operational now.
1470	An over-temperature condition exists.
1471	A low battery/high charge state exists.
1472	A Symmetrix power subsystem alarm was detected.
E473	A locally mirrored drive (source volume) is in a <i>not ready</i> state.
E474	A locally mirrored drive (source volume) is write disabled.
E475	A remotely mirrored drive is in a <i>not ready</i> state.
1476	The service processor is not responding.
1477	The AutoCall feature failed to complete a call to the EMC Customer Support Center.
1478	There is a 12V On error.
1479	The environment cable is missing.
147A	There is an AC line failure or interruption.
147B	A high charge state was not detected within two minutes of power up; or, a clock inconsistency was found between the Symmetrix system and the service processor; or, the director was inserted without system re-set.
147C	Latched alarms were detected.
147D	There is an SRDF adapter link problem.
147E	An RDF link is now operational.
147F	The service processor called home to report an error condition (customer option).

Event Messages

Table 6-2 describes the events that appear at the host console. These events are reported at the time they occur and also appear at the Symmetrix service processor. Not all events are reported to the EMC Customer Support Center. The next section shows the format of these messages.

Table 6-2 Event Messages

Sense Bytes 22-23	Meaning
0464	Migration has completed for all migration devices.
E451	Event trace trigger. PC should call home.
E452	Event trace trigger on remote RDF Symmetrix system. PC should call home.
E453	SYSCALL EMC transfer poll from host. PC should call home.
E454	Too many SUSPEND/HALT chains. Switching to Adaptive Copy Write Pending Mode. PC should call home.
E456	Automatic memory disable invoked. PC should call home.
E461	M2 is resynchronized with the M1 device after M2 is brought back to ready.
E462	M1 is resynchronized with the M2 device after M1 is brought back to ready.
E465	Device resynchronization process has begun.
E469	Fibre Channel optical module adapter reported an error. PC should call home.
E46F	Arbiter Problem. PC should call home.

Operator Messages

On OS/390 or z/OS, SIM messages will be displayed as IEA480E Service Alert Error messages. They have the format shown in Figure 6-4 and Figure 6-5.

```
*IEA480E 0D03,SCU,SERVICE ALERT,MT=3990-3,SER=,
REFCODE=147A-0000-0000
```

147A = AC Line Failure or Interruption

SYM-000418

Figure 6-4 OS/390 or z/OS IEA480E Service Alert Error Message Format (AC Power Failure)

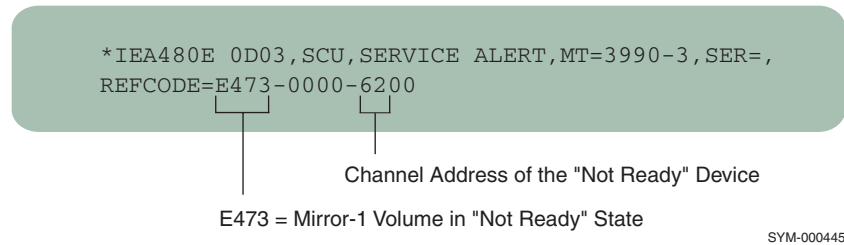


Figure 6-5 OS/390 or z/OS IEA480E Service Alert Error Message Format (Mirror-1 Volume in Not Ready State)

All host channel paths to that device (target volume) will report this error message. Therefore, this same message may appear several times.

The Symmetrix system also reports events to the host and to the service processor. These events are:

- ◆ The mirror-2 volume has synchronized with the source volume
- ◆ The mirror-1 volume has synchronized with the target volume
- ◆ Device resynchronization process has begun

On OS/390 or z/OS, these events are displayed as IEA480E Service Alert Error messages. They have the format shown in Figure 6-6 and Figure 6-7.

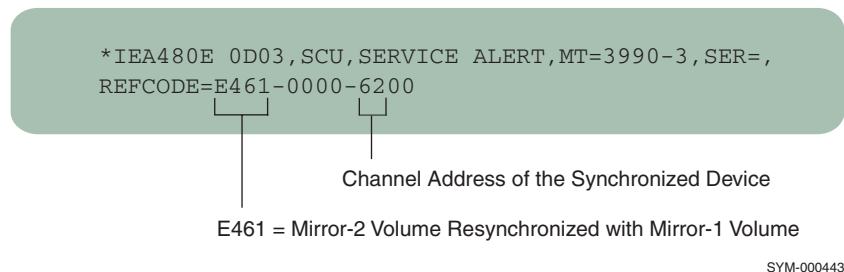


Figure 6-6 OS/390 or z/OS IEA480E Service Alert Error Message Format (Mirror-2 Resynchronization)

```
*IEA480E 0D03, SCU, SERVICE ALERT, MT=3990-3, SER=,
REFCODE=E462-0000-6200
```

Channel Address of the Synchronized Device

E462 = Mirror-1 Volume Resynchronized with Mirror-2 Volume

SYM-000444

Figure 6-7 OS/390 or z/OS IEA480E Service Alert Error Message Format (Mirror-1 Resynchronization)

EREP Reports

The Environmental Record Editing and Printing (EREP) program is an applications program that runs under the z/OS operating systems. EREP helps you monitor the functioning of various units in your system such as the processor, I/O devices, controller, and channels by supplying information on errors that have occurred in these components. When an error occurs, the operating system creates a record from the data captured by the hardware or software and writes it in the Error Recording Dataset (ERDS). EREP reads records directly from the ERDS and processes them to produce the report or reports you request.

EREP Error Records

Error records processed by EREP are described in detail in the IBM EREP manual.

Refer to the *IBM EREP User's Guide and Reference* for the formats of these records.

EREP uses the information in these records to produce many types of reports. There are three System Exception Reports applicable to disk media errors with the Symmetrix system. These reports are listed below in the order you use them to identify and handle a media error situation.

- ◆ **System Error Summary (Part 2)**—This report lists permanent I/O errors (data or equipment checks) and identifies each error by job name and time.
- ◆ **Subsystem Exception DASD**—This report lists accumulated permanent and temporary I/O errors.
- ◆ **DASD Data Transfer Summary**—This report presents details on data checks.

See the *IBM EREP User's Guide and Reference* for information on how to define report requests to meet your needs.

Error Handling

As part of your routine maintenance procedure, you should generate System Exception Reports daily. From these reports, you can determine whether errors are permanent or temporary, the number of errors that have occurred, and their frequency.

If your system reports permanent errors occurring on the Symmetrix system, save the Subsystem Exception DASD report output and contact your EMC Customer Engineer.

If your system reports that temporary data check errors are occurring on the Symmetrix system and you consider the number of temporary errors for a volume to be excessive, notify your EMC Customer Engineer.

Table 6-3 describes the error handling process you should follow for the Symmetrix system.

Table 6-3 Error Handling Steps

Step	Task	Tool(s)	Actions
1	Detect error occurrence	System Error Summary (Part 2) report (permanent errors) Subsystem Exception DASD report (temporary errors)	If permanent error, or if temporary error requires investigation, perform Step 2
2	Determine source of errors using EREP	Subsystem Exception DASD report	If the source is hardware, call the EMC Customer Engineer

Detecting the Error

Review the System Error Summary (Part 2) to determine if permanent errors are occurring on the Symmetrix system. This report lists permanent I/O errors in sequence according to the time they occurred. Also review the Subsystem Exception DASD report. This report highlights problems related to data storage that may need further investigation. It gives the number and frequency of both permanent and temporary errors.

You can use this information to defer further handling of a particular error based on its recoverability status, location, and frequency, or take immediate action to correct the error source.

Determining the Error Source



CAUTION

Permanent data errors with a probable failing unit of VOLUME require your IMMEDIATE attention!

If the System Error Summary (Part 2) and Subsystem Exception DASD reports indicate problems with a Symmetrix device, review the Subsystem Exception DASD report to determine the failing component. This report lists error information by volume. It identifies the failing volume, the track location of the failure, the sense information received from the Symmetrix system at the time of the last failure, and the permanent and temporary error counts on that device.

Sense Byte Information

The information presented in this chapter is only applicable to Symmetrix disk devices connected to mainframe hosts.

Symmetrix systems present a unit check status in the status byte to the channel whenever it detects an error condition. The error condition can be one of the error types associated with data such as a data check, command reject, overrun, or equipment check.

The channel issues a sense command to retrieve the sense data from the Symmetrix system. The host places the sense data in the Error Recording Dataset (ERDS) where the EREP subsequently uses it for its reports.

Sense byte error data may also appear on the system operating console. The format of the error data reported to the operator console depends on your operating system. Sense byte data at the host is presented in 24-byte mode for 3880 controller emulation and 24-byte compatibility mode or 32-byte mode for 3990 emulation.

Console Error Messages

Figure 6-8 is an example of a typical OS/390 console error message for 24-byte sense data.

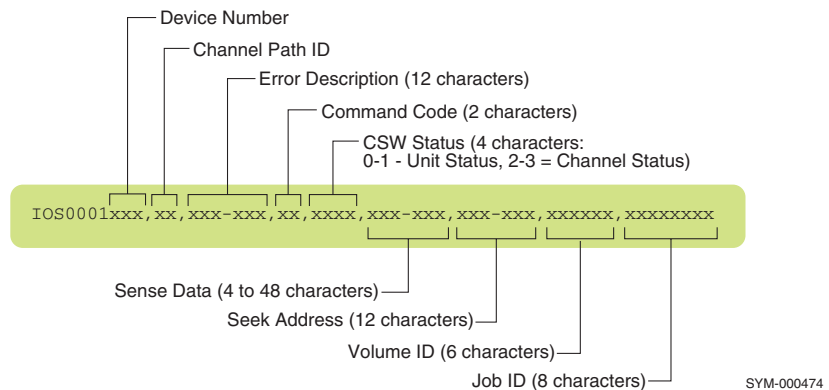


Figure 6-8 Typical Console Error Message

The unit status and channel status (CSW characters) indicate why the operation terminated. The unit status is bits 32-39 in the CSW (370 mode) and bits 0-7 of Word 2 of the SCSW (XA and ESA mode). The channel status is bits 40-47 in the CSW (370 mode) and bits 8-15 of Word 2 of the SCSW (XA and ESA mode).

Unit Status Bits Table 6-4 lists and describes the unit status bits.

Table 6-4 Unit Status Bits

370	XA, ESA	Meaning
Bit 32	Bit 0	Attention—READY status when set in conjunction with Device End and Unit Exception.
Bit 33	Bit 1	Status modifier—When set with Unit Check, indicates an usual condition, command retry; when set with Busy, indicates the Symmetrix system is busy; when set with Device End, next CCW has been skipped.
Bit 34	Bit 2	Control-unit end—Symmetrix system no longer busy.
Bit 35	Bit 3	Busy—If Status Modifier clear, the drive is busy; if Status Modifier set, the channel director is busy.
Bit 36	Bit 4	Channel end—Data or command transfer to/from channel is complete.
Bit 37	Bit 5	Device end—Device operation complete.
Bit 38	Bit 6	Unit check—Symmetrix system detected an error condition.
Bit 39	Bit 7	Unit exception—EOF on addressed track during a Read R0, Read IPL, or Read CKD, or Write Key Data or Write Data operation. If Attention and Device End set, indicates a READY status.

Channel Status Bits Table 6-5 lists and describes the channel status bits.

Table 6-5 Channel Status Bits

370	XA, ESA	Meaning
Bit 40	Bit 8	Program-controlled interruption (PCI)—CCW specified interrupt occurred.
Bit 41	Bit 9	Incorrect length—Number of bytes transferred were not equal to number of bytes specified by CCW.
Bit 42	Bit 10	Program check—CCW programming error.
Bit 43	Bit 11	Protection check—Channel attempted to address a storage area prohibited by a protection key.
Bit 44	Bit 12	Channel-data check—Incorrect parity.
Bit 45	Bit 13	Channel-control check—Channel hardware error.
Bit 46	Bit 14	Interface-control check—Invalid signal from the Symmetrix system (invalid signal sequence, overly slow response, or parity error on Bus In).
Bit 47	Bit 15	Chaining check—Channel overrun during input operation.

You can find a detailed description of these status bits in the appropriate IBM *Principles of Operation*.

Host Sense Byte Data Formats

Sense byte formats are found in the IBM reference manual appropriate for the specific Symmetrix system emulation.

**DMX800 System
Specifications**

This appendix contains the specifications of the DMX800 system.

◆ Storage Control.....	A-2
◆ Physical Data	A-10
◆ Environmental Data.....	A-12
◆ Power and Cooling Data.....	A-13
◆ Power Requirements	A-14

Storage Control

Emulation 3990-6, 2105, and FBA

Channel Speeds Up to 2 Gb/s FICON channel
Up to 2 Gb/s Fibre Channel
1 Gb/s iSCSI channel
1 Gb/s GigE remote channel (SRDF)

Storage Capacities The following tables outline capacities for the Symmetrix DMX800 systems for both the two FC director configurations and the four FC director configurations. The capacities are presented based the disk device capacities of 73 GB and 146 GB, and these methods of data protection:

- ◆ SRDF (Symmetrix Remote Data Facility)
- ◆ Mirroring RAID 1
- ◆ Parity RAID (3+1) and Parity RAID (7+1)
- ◆ RAID 5 (3+1) and RAID 5 (7+1)

The GB values shown in the following tables are based on the following conventions:

The GB₂ value is based on 1 GB = 1024 x 1024 x 1024 bytes.

The GB₁₀ value is based on 1 GB = 1000 x 1000 x 1000 bytes.

Although the open systems values are expressed differently, they are equivalent.

A Symmetrix DMX800 system reserves four 3 GB (6,140 cylinders) logical volumes (12 GB total system requirement) of usable disk capacity for internal Symmetrix file system (SFS) purposes.

Table A-1 Storage Capacities (73 GB Mirrored Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	4	272 ₂ (292 ₁₀)	288 ₁₀
16	8	544 ₂ (584 ₁₀)	577 ₁₀
24	12	816 ₂ (877 ₁₀)	866 ₁₀
32	16	1,089 ₂ (1,169 ₁₀)	1,154 ₁₀
48	24	1,633 ₂ (1,754 ₁₀)	1,732 ₁₀
64	32	2,178 ₂ (2,339 ₁₀)	2,309 ₁₀
96	48	3,267 ₂ (3,508 ₁₀)	3,464 ₁₀
120	60	4,085 ₂ (4,386 ₁₀)	4,330 ₁₀

Table A-2 Storage Capacities (73 GB SRDF Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	8	544 ₂ (584 ₁₀)	577 ₁₀
16	16	1,089 ₂ (1,169 ₁₀)	1,154 ₁₀
24	24	1,633 ₂ (1,754 ₁₀)	1,732 ₁₀
32	32	2,178 ₂ (2,339 ₁₀)	2,309 ₁₀
48	48	3,267 ₂ (3,508 ₁₀)	3,464 ₁₀
64	64	4,357 ₂ (4,678 ₁₀)	4,618 ₁₀
96	96	6,535 ₂ (7,017 ₁₀)	6,928 ₁₀
120	120	8,169 ₂ (8,772 ₁₀)	8,660 ₁₀

Table A-3 Storage Capacities (73 GB RAID (3+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	6	408 ₂ (438 ₁₀)	433 ₁₀
16	12	816 ₂ (877 ₁₀)	866 ₁₀
24	18	1,225 ₂ (1,315 ₁₀)	1,299 ₁₀
32	24	1,633 ₂ (1,754 ₁₀)	1,732 ₁₀
48	36	2,450 ₂ (2,631 ₁₀)	2,598 ₁₀
64	48	3,267 ₂ (3,508 ₁₀)	3,464 ₁₀
96	72	4,901 ₂ (5,263 ₁₀)	5,196 ₁₀
120	90	6,127 ₂ (6,579 ₁₀)	6,495 ₁₀

Table A-4 Storage Capacities (73 GB RAID (7+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	7	469 ₂ (511 ₁₀)	505 ₁₀
16	14	953 ₂ (1,023 ₁₀)	1,010 ₁₀
24	21	1,429 ₂ (1,535 ₁₀)	1,515 ₁₀
32	28	1,902 ₆ (2,046 ₁₀)	2,020 ₁₀
48	42	2,859 ₂ (3,070 ₁₀)	3,031 ₁₀
64	56	3,812 ₂ (4,093 ₁₀)	4,041 ₁₀
96	84	5,718 ₂ (6,140 ₁₀)	6,062 ₁₀
120	105	7,148 ₂ (7,675 ₁₀)	7,577 ₁₀

Table A-5 Storage Capacities (73 GB RAID 5 (3+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	6	408 ₂ (438 ₁₀)	433 ₁₀
16	12	816 ₂ (877 ₁₀)	866 ₁₀
24	18	1,225 ₂ (1,315 ₁₀)	1,299 ₁₀
32	24	1,633 (1,754 ₁₀)	1,732 ₁₀
48	36	2,450 ₂ (2,631 ₁₀)	2,598 ₁₀
64	48	3,267 ₂ (3,508 ₁₀)	3,464 ₁₀
96	72	4,901 ₂ (5,263 ₁₀)	5,196 ₁₀
120	90	6,127 ₂ (6,579 ₁₀)	6,495 ₁₀

Table A-6 Storage Capacities (73 GB RAID 5 (7+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	7	469 ₂ (511 ₁₀)	505 ₁₀
16	14	939 ₂ (1,023 ₁₀)	1,010 ₁₀
24	21	1,408 ₂ (1,535 ₁₀)	1,515 ₁₀
32	28	1,878 (2,046 ₁₀)	2,020 ₁₀
48	42	2,817 ₂ (3,070 ₁₀)	3,031 ₁₀
64	56	3,756 ₂ (4,093 ₁₀)	4,041 ₁₀
96	84	5,634 ₂ (6,140 ₁₀)	6,062 ₁₀
120	105	7,043 ₂ (7,675 ₁₀)	7,577 ₁₀

Table A-7 Storage Capacities (146 GB Mirrored Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	4	544 ₂ (584 ₁₀)	578 ₁₀
16	8	1,088 ₂ (1,168 ₁₀)	1,156 ₁₀
24	12	1,632 ₂ (1,752 ₁₀)	1,735 ₁₀
32	16	2,176 ₂ (2,336 ₁₀)	2,313 ₁₀
48	24	3,264 ₂ (3,504 ₁₀)	3,470 ₁₀
64	32	4,352 ₂ (4,672 ₁₀)	4,627 ₁₀
96	48	6,528 ₂ (7,008 ₁₀)	6,940 ₁₀
120	60	8,160 ₂ (8,760 ₁₀)	8,676 ₁₀

Table A-8 Storage Capacities (146 GB SRDF Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	8	1,088 ₂ (1,168 ₁₀)	1,156 ₁₀
16	16	2,176 ₂ (2,336 ₁₀)	2,313 ₁₀
24	24	3,264 ₂ (3,504 ₁₀)	3,470 ₁₀
32	32	4,352 ₂ (4,672 ₁₀)	4,627 ₁₀
48	48	6,528 ₂ (7,008 ₁₀)	6,940 ₁₀
64	64	8,704 ₂ (9,344 ₁₀)	9,254 ₁₀
96	96	13,056 ₂ (14,016 ₁₀)	13,881 ₁₀
120	120	16,320 ₂ (17,520 ₁₀)	17,352 ₁₀

Table A-9 Storage Capacities (146 GB RAID (3+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	6	816 ₂ (876 ₁₀)	867 ₁₀
16	12	1,632 ₂ (1,752 ₁₀)	1,735 ₁₀
24	18	2,448 ₂ (2,628 ₁₀)	2,602 ₁₀
32	24	3,264 ₂ (3,504 ₁₀)	3,470 ₁₀
48	36	4,896 ₂ (5,256 ₁₀)	5,205 ₁₀
64	48	6,528 ₂ (7,008 ₁₀)	6,940 ₁₀
96	72	9,792 ₂ (10,512 ₁₀)	10,411 ₁₀
120	90	12,240 ₂ (13,140 ₁₀)	13,014 ₁₀

Table A-10 Storage Capacities (146 GB RAID (7+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	7	952 ₂ (1,022 ₁₀)	1,012 ₁₀
16	14	1,904 ₂ (2,044 ₁₀)	2,024 ₁₀
24	21	2,856 ₂ (3,066 ₁₀)	3,036 ₁₀
32	28	3,808 ₂ (4,088 ₁₀)	4,048 ₁₀
48	42	5,712 ₂ (6,132 ₁₀)	6,073 ₁₀
64	56	7,616 ₂ (8,176 ₁₀)	8,097 ₁₀
96	84	11,424 ₂ (12,264 ₁₀)	12,146 ₁₀
120	105	14,280 ₂ (15,330 ₁₀)	15,183 ₁₀

Table A-11 Storage Capacities (146 GB RAID 5 (3+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	6	816 ₂ (876 ₁₀)	867 ₁₀
16	12	1,632 ₂ (1,752 ₁₀)	1,735 ₁₀
24	18	2,448 ₂ (2,628 ₁₀)	2,602 ₁₀
32	24	3,264 ₂ (3,504 ₁₀)	3,470 ₁₀
48	36	4,896 ₂ (5,256 ₁₀)	5,205 ₁₀
64	48	6,528 ₂ (7,008 ₁₀)	6,940 ₁₀
96	72	9,792 ₂ (10,512 ₁₀)	10,411 ₁₀
120	90	12,240 ₂ (13,140 ₁₀)	13,014 ₁₀

Table A-12 Storage Capacities (146 GB RAID 5 (7+1) Disk Devices)

Total Disk Devices	Usable Disk Devices	Open Systems Capacity (GB)	Mainframe Capacity (GB)
8	7	952 ₂ (1,022 ₁₀)	1,012 ₁₀
16	14	1,904 ₂ (2,044 ₁₀)	2,024 ₁₀
24	21	2,856 ₂ (3,066 ₁₀)	3,036 ₁₀
32	28	3,808 ₂ (4,088 ₁₀)	4,048 ₁₀
48	42	5,712 ₂ (6,132 ₁₀)	6,073 ₁₀
64	56	7,616 ₂ (8,176 ₁₀)	8,097 ₁₀
96	84	11,424 ₂ (12,264 ₁₀)	12,146 ₁₀
120	105	14,280 ₂ (15,330 ₁₀)	15,183 ₁₀

MB/Volume by Emulation Type	3380D emulation	630
	3380E emulation	1,260
	3380K emulation	1,891
	3380K(+) emulation	2,378
	3380K(++) emulation	2,843
	3390-1 emulation	946
	3390-2 emulation	1,892
	3390-3 emulation	2,838
	3390-9 emulation	8,514
	3390-27 emulation	27,844
Bytes/Track	Fixed Block Architecture (FBA) capacities are based on 512 bytes per block, 64 bytes per track, and 15 tracks per cylinder.	
	3380 emulations	47,476
	3390 emulations	56,664
	FBA	32,768
Bytes/Cylinder	Fixed Block Architecture (FBA) capacities are based on 512 bytes per block, 64 bytes per track, and 15 tracks per cylinder.	
	3380 emulations	712,140
	3390 emulations	849,960
	FBA	491,520
Cylinders/Volume	3380D emulation	885
	3380E emulation	1,770
	3380K emulation	2,655
	3380K(+) emulation	3,339
	3380K(++) emulation	3,993
	3390-1 emulation	1,113
	3390-2 emulation	2,226
	3390-3 emulation	3,339
	3390-9 emulation	10,017
	3390-27 emulation	32,760
	FBA (Maximum)	65,520
Disk Device Form Factor	1-inch	Height (Low-Profile)
	3.5-inch	Width

Physical Data

Depth

36 in. (92 cm)

Width

24 in. (61 cm)

Height

75 in. (190 cm)

Weights

Table A-13 Symmetrix DMX800 Weights

Single SPE in Rack	Number of Drives	lbs (kgs)	Dual SPE in one Rack (Full Config.) lbs (kgs)
Rack		353 (160)	353 (160)
SPE (including global memory, directors, FEBE)		127 (58)	254 (115)
Server		25 (11)	25 (11)
KVM		32 (15)	32 (15)
Total Component Wt. w/o DAEs, Drives, SPS, and RS232		509 (231)	636 (288)
Drives		2.5 (1.1)	2.5 (1.1)
SPS Assembly		57 (26)	57 (26)
RS232		4 (2)	4 (2)
DAE (w/o Drives)		58 (26)	58 (26)
2 DAE	8	788 (357)	
2 DAE	15	805 (365)	
2 DAE	30	843 (382)	

Table A-13 Symmetrix DMX800 Weights (continued)

Single SPE in Rack	Number of Drives	lbs (kgs)	Dual SPE in one Rack (Full Config.) lbs (kgs)
3 DAE	45	1,003 (455)	
4 DAE	60	1,099 (498)	
5 DAE ^a	75	1,195 (542)	
6 DAE ^a	90	1,348 (611)	1,475 (669)
7 DAE ^a	105	1,501 (680)	
8 DAE ^a	120	1,596 (724)	

a. Requires a four FC director system.

Access (Raised) Floor
Tile Requirements

Not required for DMX800.

Service Area

Front 48 inches (122.0 cm)
Rear 48 inches (122.0 cm)
Right 28 inches (71.0 cm) (access to SPE card cage)

Floor Space (with
Service Area)

57.4 ft² 5.3 m²

Environmental Data

Operating Temperature	50° to 90° F (10° to 32° C)
Operating Altitude (maximum)	8,000 ft (2500 m)
Operating Humidity	20% to 80% non-condensing

Power and Cooling Data

Power Consumption and Heat Dissipation

Table A-14 shows the power consumption and heat dissipation for the DMX800 system with optional DAE configurations.

Table A-14 Power Consumption/Heat Dissipation

DAE Count	2 DAEs	3 DAEs	4 DAEs	5 DAEs	6 DAEs	7 DAEs	8 DAEs
Drive Count	30	45	60	75	90	105	120
Power Consumption (VA)	1,971	2,406	2,816	3,222	3,636	4,046	4,456
Heat Dissipation (Btu/hr)	6,725	8,209	9,608	11,007	12,406	13,804	15,203

Under normal N+1 operation with two fully charged SPSs, the total input load on an SPE is 492 VA. Under N+1 operation with both SPEs charging, the total input load on the SPE is 1062 VA.

Power Requirements

Tables A-15 through A-17 describe the North American and International electrical specifications, power cables, connectors, and extension cords for the Symmetrix DMX800 systems.

Table A-15 DMX800 Electrical Specifications, Single-Phase

Specification	North America	International
Input Voltage	208 / 240 VAC L-L nom (240 VAC optimal)	220 / 230 / 240 VAC L-N nom (240 VAC optimal)
Range	200–240 VAC	200–240 VAC
Frequency	47–63 Hz	47–63 Hz
Circuit Breakers	30 A	32 A
Power Zones	One	One
Power Requirements at Customer Site (minimum)	Two 30 A, single-phase drops	Two 32 A, single-phase drops

Table A-16 DMX800 Power Cables and Connectors, Single-Phase

Power Cables and Connectors Included with Each System		Mating Connectors		
Power Connection	Connector Type	Vendor Part Number	EMC Part Number	EMC Model Number
Two 12-inch (approximate) power cables affixed to each system	Two Hubbell 30A L6-30 twist-lock connectors	NEMA L6-30 IEC-309-332P6, IP-x7 Clipsal 56PA332 (or equivalent) (North American and International connectors are industry standards)	118-032-023 118-032-025 118-032-024	DMX-PWR40U-US DMX-PWR40UIEC3 DMX-PWR40UASTL (Connectors are not sold separately. Line cord is sold with connectors on both ends. Customer supplies mating connector.)

Table A-17 DMX800 Extension Line Cords, Single-Phase

EMC Part Number	EMC Model Number	Description
118-032-023	DMX-PWR40U-US	North American 15-ft single-phase line cord with connectors on each end (one L6-30R and one L6-30P)
118-032-025	DMX-PWR40UIEC3	International 15-ft single-phase line cord with connectors on each end (one L6-30R and one IEC-309-332P6)
118-032-024	DMX-PWR40UASTL	Australian 15-ft single-phase line cord with connectors on each end (one L6-30R and one Clipsal 56PA332)

Power Sequences

This appendix provides step-by-step instructions for powering the DMX800 on and off. The unit is powered on and off through the power switches on the rear door of the unit.

- ◆ Shutting the System Down B-2
- ◆ Powering Up the Symmetrix System after a Shutdown..... B-4

Shutting the System Down

Follow these steps when powering down the Symmetrix system.

Powering down the entire system is not necessary for some disruptive replacement procedures. Review the specific procedure before starting.

This procedure requires an attached KVM or a PC running VNC.

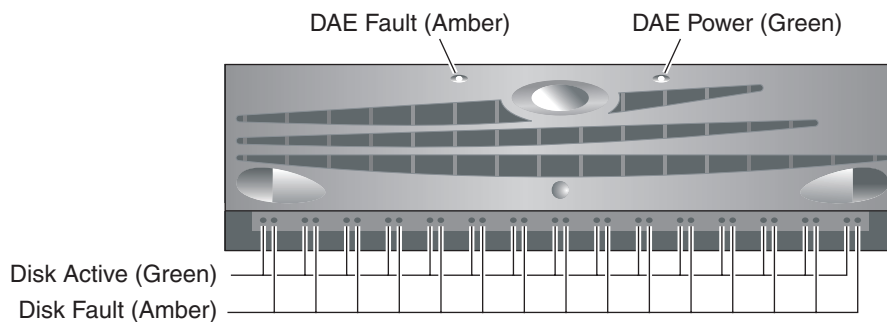


CAUTION

Do not power down the Symmetrix system to replace a component if it can be replaced nondisruptively.

1. Have the customer stop all processes to the Symmetrix system.
2. Wait for all pending write operations to complete to the disk devices.

The green *Disk Active* LEDs (shown in Figure B-1) on all disk modules flash on and off during write operations, and remain constantly on when all writes are completed.



SYM-000343

Figure B-1 Disk Array Enclosure, Front View

All pending write operations should complete within five minutes. If a device fails to complete its write operations within this time period, it could indicate a problem with that device. Call the EMC Service Support Center for assistance if necessary.

3. From the SymmWin main menu, go to Inlines. (Refer to Figure B-2.)

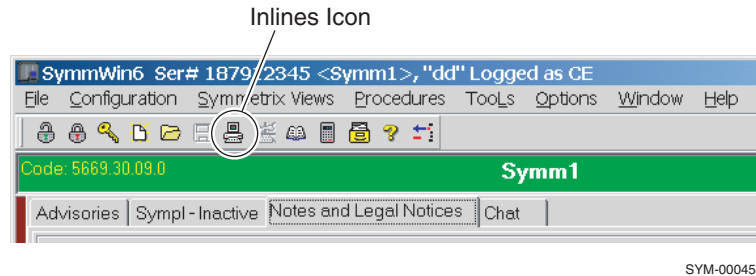


Figure B-2 Inlines Icon

4. Verify that all pending write operations have stopped, as follows:
 - a. Type **A4** and press ENTER.
 - b. Type **A7** and press ENTER and verify that the **WRT PEND** column is all zeroes.
 If any column is not zero, contact the EMC Customer Support Center for assistance.
5. Take all channels and devices offline, as follows:
 - a. Issue **\$BC,CA,F0,CE,0F** to take all host channels offline.
 - b. If there are any RDF connections, issue **\$BC,F3,F0,CE,0F** to take all RDF channels offline.
 - c. Issue **\$BC,DA,F0,CE,0F** to take all disk devices offline.
6. Switch the PDU's two main AC circuit breakers (one on each power strip, shown in Figure B-3 on page B-4) OFF (down).

When the system detects the loss of AC power, the SPS modules power on automatically and keep the system running for up to 90 seconds, during which time the system destages write operations in an orderly manner. Once the write pendings are destaged, the DMX800 system powers off.

If there are no write pendings detected and AC power is lost, the DMX800 will immediately start power off.

The maximum allowable number of write pendings is determined by the number of DAEs in the system and the number of drives per DAE. This limit ensures that all write pendings will be destaged within 90 seconds.

Powering Up the Symmetrix System after a Shutdown

To power up after the system has been shut down:

1. If not already done, plug each of the two 240-volt AC power cords (shown in Figure B-3) into a different branch circuit.

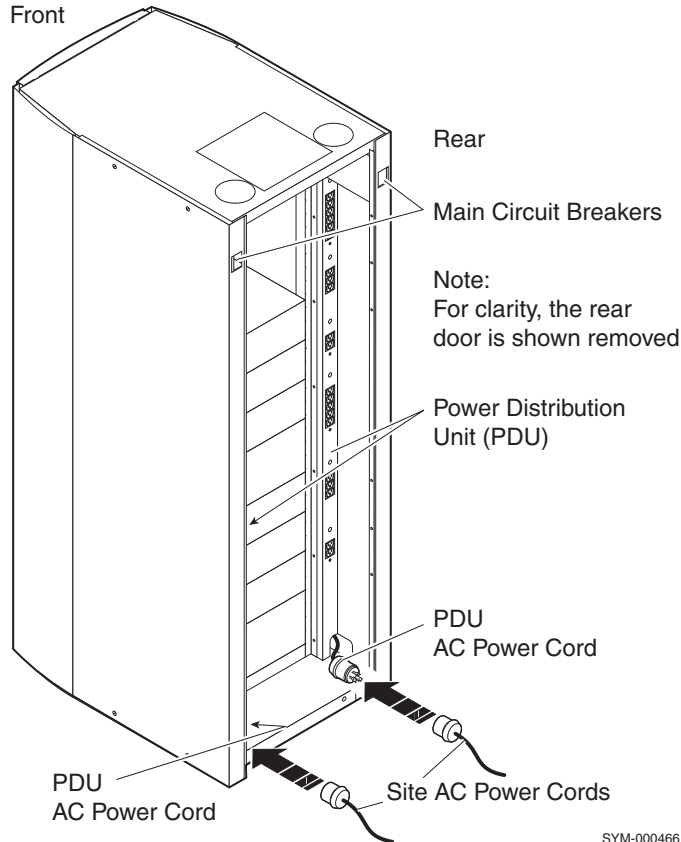


Figure B-3 AC Power Cords

2. Ensure that all DAE and SPS power switches (shown in Figure B-4) are in the ON (1) position.

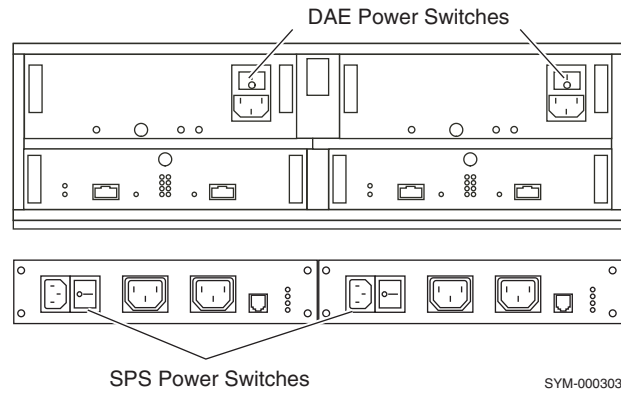


Figure B-4 DAE and SPS Power Switches

3. Switch the PDU's two main AC circuit breakers (shown in Figure B-3) ON (up).
4. Wait at least 30 minutes for the IML procedure to complete.
5. Ensure that the system is online and error free before connecting to the host(s).
6. Put all channels and devices online, as follows:
 - a. Issue **\$BC,DA,F0,CE,0** to put all disk devices online.
 - b. If there are any RDF connections, issue **\$BC,F3,F0,CE,0** to put all RDF channels online.
 - a. Issue **\$BC,CA,F0,CE,0** to put all host channels online.

Symmetrix DMX800 System Planning and Installation

This appendix covers the tasks you need to complete when planning or verifying the physical configuration of a Symmetrix DMX800 system or creating I/O addressing schemes.

- ◆ Planning Overview C-2
- ◆ Director/Global Memory Board Layout C-8
- ◆ Mainframe/Open Systems Installations C-10
- ◆ GigE Remote and iSCSI Director Installations..... C-22

Planning Overview

This section guides you through the physical details related to delivery and installation of the Symmetrix DMX800 system. It reviews all the necessary installation details performed by EMC personnel and outlines customer responsibilities.

Inform EMC of any labor union-based restrictions or security clearance requirements prior to delivery.

The Symmetrix system is designed for installation in a properly equipped computer room with controlled temperature and humidity, proper airflow and ventilation, proper power and grounding, system cable routing facilities, fire equipment, and so on. One or more planning sessions with your EMC Systems Engineer and Customer Engineer may be necessary to finalize all the details related to installation. Table C-1 lists the responsibility summary at the first planning session.

Table C-1 Preinstallation Responsibility Summary

EMC Responsibilities	Customer Responsibilities
Provide all details necessary for site planning and preparation.	Provide an environment that supports the safe installation of the rackmount Symmetrix system and promotes its reliable long-term operation.
Complete and process the Installation Planning Task Sheet and Pre-site Survey.	Provide appropriate power, cooling and ventilation, humidity control, floor load capability, and service clearances as required.
Arrange for shipment and delivery through the appropriate method.	Participate in planning sessions as required to ensure a smooth and uncomplicated installation.
Install a properly working system as promised and on schedule.	

Physical Specifications

Physical Data on page A-10, provides overall dimensions and weights for the Symmetrix system.

Transportation and Delivery Guidelines

Symmetrix systems delivered within the United States or Canada travel by air-ride truck or van. The system is shrouded by custom-designed shipping material, crated, and palletted. Integrated shock-absorbing casters on which the Symmetrix system rests facilitates its movement during shipping and installation.

Systems delivered internationally normally involve air freight and are, therefore, crated for shipment. Unless otherwise instructed, the EMC Traffic Department arranges for delivery directly to your computer room. To ensure successful delivery of the system, EMC has formed partnerships with specially selected moving companies. These companies have moving professionals trained in the proper handling of large and very sensitive systems. These companies provide the appropriate personnel, floor layments, and any ancillary moving equipment required to facilitate delivery.

Power Requirements

The Symmetrix system operates on 208 or 240 VAC single-phase input power at frequencies of 47 Hz to 63 Hz. Domestic units are cable-ready to plug directly into your prewired receptacles. International units arrive ready for hard-wiring at your facility. You are responsible for supplying and installing two receptacles on *separate* circuits before delivery. EMC also recommends that these circuits reside on separate circuit panels.

Power Requirements on page A-14 lists the AC power requirements.

Symmetrix Earth Leakage Current Compliance

The Earth Leakage Compliance statement *only* applies to an entire Symmetrix rackmount configuration consisting of the rack and its associated components. It does *not apply* to a DMX800 stand-alone SPE chassis.

The Symmetrix rackmount system is designated as fixed (stationary) electrical equipment with high earth leakage markings. The intention is to connect the cabinets to the customer AC supply with the recommended rated current breakers (refer to *Power Requirements* on page A-14 for recommended ampere breakers) and the attached cables. These recommended breaker ratings are based on the maximum kW loading of the cabinet and not on the kVA figures of a particular configuration. The product manual's kVA figures (refer to *Power and Cooling Data* on page A-13) are intended for air conditioning and utility loading purposes only.

Because the DMX800 system is a high earth leakage device, differential trip devices (typically 30 mA) are *not* recommended due to random activations during utility feed distortions and powerline transients interacting with the cabinet noise filters. These differential devices can be called GFCI (Ground Fault Circuit Interrupter, GFI (Ground Fault Interrupter), ELCB (Earth Leakage Circuit Breaker), or other names but typically have trip ratings of 5 mA to 500 mA and are mostly intended for consumer goods rather than fixed devices.

If the Symmetrix cabinet is correctly grounded, leakage current should not produce a voltage leading to electrical shock. Serious insulation breakdowns should trip the breakers feeding the cabinet or, preferably, should first clear fuses in individual internal modules closest to the fault.

The Symmetrix system meets regulatory requirements as referenced by the following IEC 60950 code (Third Edition, 1999):

5.1.7 Equipment with touch (leakage) current exceeding 3.5mA

For Stationary Permanently Connected Equipment, or Stationary Pluggable Equipment Type B, with a main protective earthing terminal, and if the Touch (Leakage) Current measurements exceed 3.5mA r.m.s., all of the following conditions apply:

- The r.m.s. Protective Conductor Current shall not exceed five percent of the input current per line under normal operating conditions. If the load is unbalanced, the largest of the three line (phase) currents shall be used for this calculation.
- To measure the Protective Conductor Current, the procedure for measuring Touch Current is used, but the measuring instrument is replaced by an ammeter of negligible impedance; and
 - The cross-sectional area of the Protective Bonding Conductor shall not be less than 1.0 mm squared in the path of high Protective Conductor Current
 - A label with similar wording shall be affixed adjacent to the equipment AC MAIN SUPPLY connection:

HIGH LEAKAGE CURRENT

Earth connection essential
before connecting supply.

The Symmetrix DMX800 system is a Stationary Pluggable Type B system, and has been extensively tested and certified to meet the above standard, including the application of this IEC 60950 warning label, EMC P/N 046-000-309, in English and French. The label location for the Symmetrix DMX800 is on the inside of the rear door.

Choosing a UPS

When the system detects the loss of AC power, the SPS modules power on automatically and keep the system running for up to 90 seconds, during which time the system destages write operations in an orderly manner. Once the write pendings are destaged, the DMX800 system powers off.

If you need to extend this time period, you will need to purchase a UPS from a qualified vendor. When you are planning the UPS solution for the Symmetrix system, and the host system is presently (or is going to be) protected with a UPS, the battery backup time you propose for the Symmetrix UPS solution should match that of the host system.

EMC neither offers nor recommends any specific UPS suppliers or product type for its customers. This said, EMC uses preferred suppliers for UPS systems in their facilities. Therefore, if you, the customer, are implementing a UPS, EMC recommends the following:

- ◆ When you are planning the UPS solution for the Symmetrix system, and the host system is presently (or is going to be) protected with a UPS, the battery backup time you propose for the Symmetrix UPS solution should match that of the host system.
- ◆ The Symmetrix system requires independent main and auxiliary power feeds.
- ◆ The UPS should be equipped with an internal output isolation transformer.
- ◆ The UPS should be installed as a separately derived AC source using neutral and ground wiring to preserve the (N+1) or 2 (N+1) fault tolerance specification of the Symmetrix power system.
- ◆ Depending on the power requirements for your Symmetrix system operation, an isolation transformer/stabilizer installed in front of the UPS could further buffer the AC utility environmental factors from reaching the Symmetrix system. To determine if an isolation transformer/stabilizer is needed, consult a licensed electrician and your EMC Customer Engineer.

Environmental Specifications

The DMX800 system requires the environmental specifications outlined in *Environmental Data* on page A-12. You must make sure that the site meets or exceeds the specifications listed.

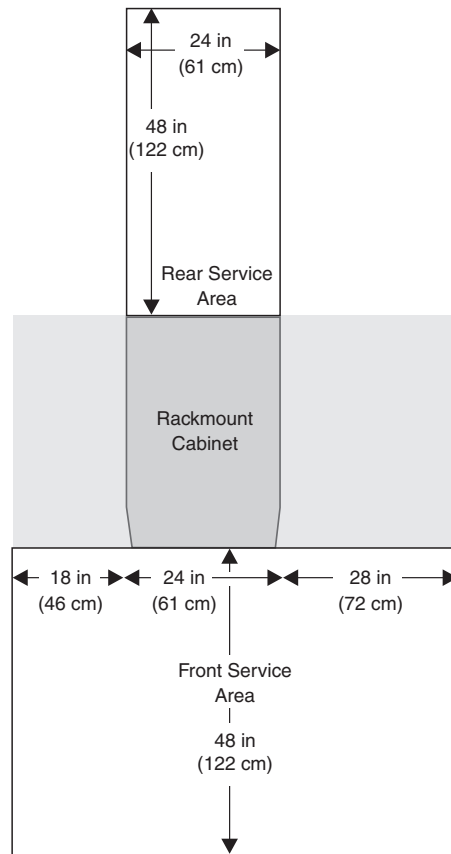
System Cabling Requirements

The Presite Survey completed with the EMC Systems Engineer reports the cable lengths (Fibre Channel) required for each host connection to the Symmetrix system. From a physical planning perspective, review the routing path(s) from the host(s) to the Symmetrix system. Resolve any physical access issues *before* installation.

Layout and Space Requirements

The DMX800 system requires the following space to accommodate Fibre Channel cabling and service access (Figure C-1 on page C-7).

It is necessary to support all four corners of the floor tiles, and the overall floor load capacity must be sufficient to safely support the particular model to be installed. For the access (raised) floor tile load requirements and service clearance requirements, refer to *Physical Data* on page A-10.



SYM-000472

Figure C-1 DMX800 System Service Area

**CAUTION**

Before the Symmetrix system is rolled into position in the computer room, note how close the wheels are to the edge of the cutout. If the DMX800 unit needs to be relocated, contact the EMC Customer Support Center.

Remote Support

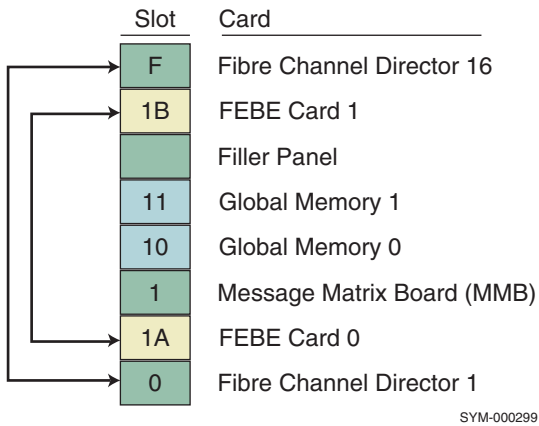
Remote support is an important and integral part of the EMC Customer Service and support strategy. Communication between the EMC Customer Support Center and the Symmetrix unit occurs through the serial port external modem connected to the server.

Director/Global Memory Board Layout

Figure C-2 shows the board layout in the card cage for the DMX800 with two FC directors installed.

Figure C-3 shows the board layout in the card cage for the DMX800 with four FC directors installed.

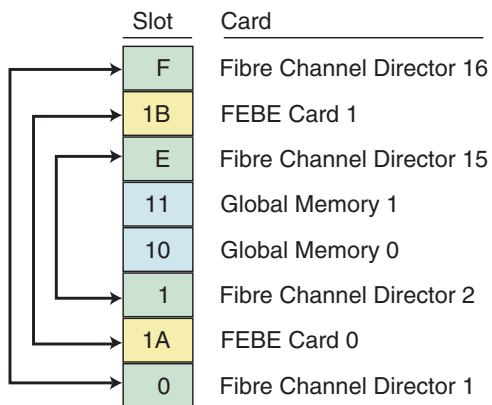
For SRDF Remote Link director configuration, refer to the *Symmetrix Remote Data Facility Product Guide*.



SYM-000299

Note:
The filler panel and the MMB are removed
in order to accommodate two MPCD
(FICON, GigE/iSCSI, or FICON/GigE/iSCSI).

Figure C-2 DMX800 Two Fibre Channel Director and Global Memory Configuration



SYM-000464

Figure C-3 DMX800 Four Fibre Channel Director and Global Memory Configuration

Mainframe/Open Systems Installations

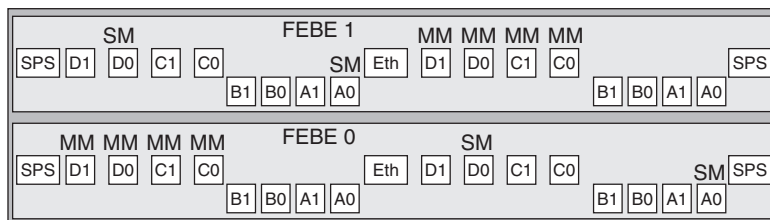
For Open Systems installation requirements, refer to *Open Systems Installations* on page C-13.

FICON/ FEBE Board Port Designations

The following models show the possible FEBE board port configurations with mainframe FICON (MPCD) directors and a mix of one FICON director and one GigE/iSCSI (MPCD) director. For card cage slot assignments, refer to Figure C-2 on page C-8 and Figure C-3 on page C-9.

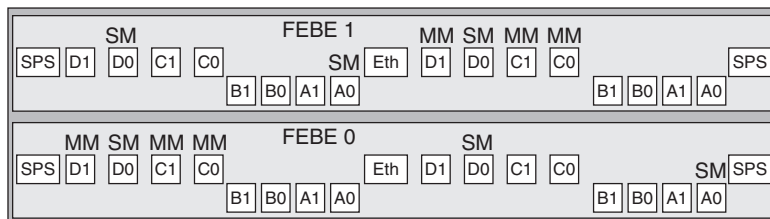
FICON Directors

- ◆ **DMX8-4002**—Four Fibre Channel multimode ports, two FICON single-mode ports:



SYM-000447

- ◆ **DMX8-3102**—Three Fibre Channel multimode ports, one Fibre Channel single-mode port, and two FICON single-mode ports:

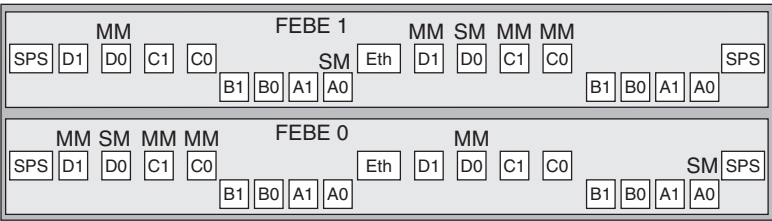


SYM-000429

FICON/GigE/iSCSI Directors

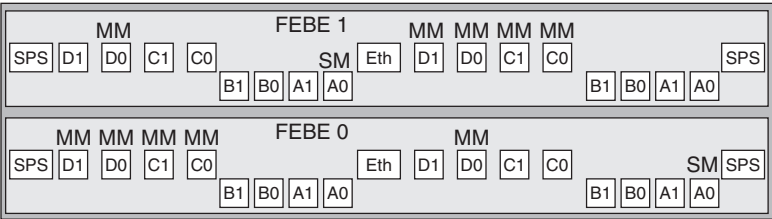
The following FEBE models contain one FICON mezzanine card and one GigE mezzanine card on the MPCD:

- ◆ **DMX8-3111**—Three Fibre Channel multimode ports, one Fibre Channel single-mode port, one GigE multimode port, and one FICON single-mode port:



SYM-000430

- ◆ **DMX8-4011**—Four Fibre Channel multimode ports, one GigE multimode port, one FICON single-mode port:



SYM-000433

Available EMC FICON Cables

Table C-2 lists the Symmetrix DMX800 FICON cables currently available from EMC. To obtain these cables, contact your EMC Sales Representative.

Table C-2 EMC Fibre Cables —FICON 9 Micron Connect

Model Number	Description
FC1M-9MSLC	SC/LC-1 meter w/coupler
FC3M-9MSLC	SC/LC-3 meter w/coupler
FC5M-9MSLC	SC/LC-5 meter
FC10M-9MSLC	SC/LC-10 meter
FC30M-9MSLC	SC/LC-30 meter
FC50M-9MSLC	SC/LC-50 meter
FC100M-9MSLC	SC/LC-100 meter
FC5M-9MLC	LC/LC-5 meter
FC10M-9MLC	LC/LC-10 meter
FC30M-9MLC	LC/LC-30 meter
FC50M-9MLC	LC/LC-50 meter
FC100M-9MLC	LC/LC-100 meter

Open Systems Installations

For mainframe installation requirements, refer to *Open Systems Installations* on page C-13.

This section contains a hardware and host checklist that you can use when connecting the DMX800 system to open systems hosts. For additional information on open systems installations, refer to the Powerlink website at <https://powerlink.emc.com>. The *EMC Support Matrix* is available from emc.com and the Powerlink website.

DMX800 System Hardware Checklist

Make sure you discuss with and obtain the following site profile information from the Customer Engineer and/or Systems Engineer. This information is necessary for each Symmetrix system you are installing (Table C-3).

Table C-3 Symmetrix DMX800 Checklist for UNIX or PC Server Hosts

Total number of physical drives to be configured on the Symmetrix DMX800 system (8 to 120 disk drives)	
Physical drive type (size) — 73 GB, 146 GB	
Total amount of Symmetrix global memory (4 GB to 64 GBs)	
Number of Fibre Channel directors	
Number of MPCD (FICON, GigE, iSCSI combinations)	
Host channel director models used	
Number of channels used per director ^a	
Remote link director (SRDF) type (RA-2, RA-4) and quantity	

a. From a performance perspective, it is better to spread Fibre Channels across as many Fibre Channel directors as possible.

Host Checklists

Make sure you discuss with and provide the following host information to the Customer Engineer and/or Systems Engineer. This information is necessary for each host you are attaching to the Symmetrix system. Make a copy of this form (Table C-4 on page C-14) for each additional host you will attach to the Symmetrix system. When you are done you should have a checklist for each host.

Table C-4 UNIX or PC Server Host Checklist

Host Configuration Requirements	Host 1	Host 2
Host (CPU) vendor and model number		
Host controller type and model number		
Memory capacity of host		
O/S revision level of host		
I/O rate per second expected per host		
Is this a clustered environment? Which one?		
Will devices be shared? Which ones?		
Total number of Fibre Channels per host, to which FA ports will they attach?		
Specify if any narrow channels used		
Number of logical devices needed per Fibre Channel path.		
Size of volumes required to be visible to host		
Total customer-usable data storage required		
Will host-level mirroring be used? Which volumes?		
Will Symmetrix Dynamic Sparing (specify number of spares) be used, which volumes?		
Will Symmetrix Mirrored (RAID 1) be used? Which volumes?		
Will Symmetrix RAID (3+1) or RAID (7+1) be used? Which volumes?		
Will Symmetrix RAID 5 (3+1) or RAID (5 7+1) be used? Which volumes?		
Will SRDF be used? Which volumes?		

Table C-4 UNIX or PC Server Host Checklist (continued)

Host Configuration Requirements	Host 1	Host 2
Data storage utilization per host ^a		
Average transfer size of data		
Using raw devices or file systems?		
Size of file system		
Will data striping be used? What type of data striping package?		
Partitioning?		
Partition sizes?		
LVM used ^b		
What major applications are to be run?		
Database used: Oracle/Sysbase/Informix/other?		
Size of database		
Database release version		
Supply typical high-level database schema and queries		
Any patches or modifications related to I/O?		
Additional comments		

a. Percentage of available Symmetrix data capacity used by that host.

b. Special attention is required when using a Logical Volume Manager (LVM) and/or data striping when using hypervolumes. *In general, EMC recommends using data striping on the Symmetrix system. Keep in mind that the larger the granularity of the striping, the less effective it becomes.*

Fibre Channel Cabling

The physical connection to a Fibre Channel interface occurs on the FEBE board located in the SPE card cage. Refer to Figure C-4 on page C-16, and *Fibre Channel/FEBE Port Designations* on page C-17.

When connecting to hosts with differential Fibre Channel interfaces, refer to the *EMC Host Connectivity Guides* on the Powerlink website.

To obtain adapter part numbers for Fibre Channel configurations, consult your EMC Sales Representative, or refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink homepage, select the menu options: Services, Document Library, Host Connectivity.

FEBE Board Channel Designations

The FEBE board (FEBE0, FEBE1) is a 16-port board with up to eight front-end ports and eight back-end ports (Figure C-4 on page C-16). It serves as the interface between the MPCD (FICON, GigE, or iSCSI) or Fibre Channel directors and the host bus adapter (HBA) or disk devices. The following sections show the FEBE board port designations for the following channel directors:

- ◆ Fibre Channel
- ◆ GigE
- ◆ iSCSI

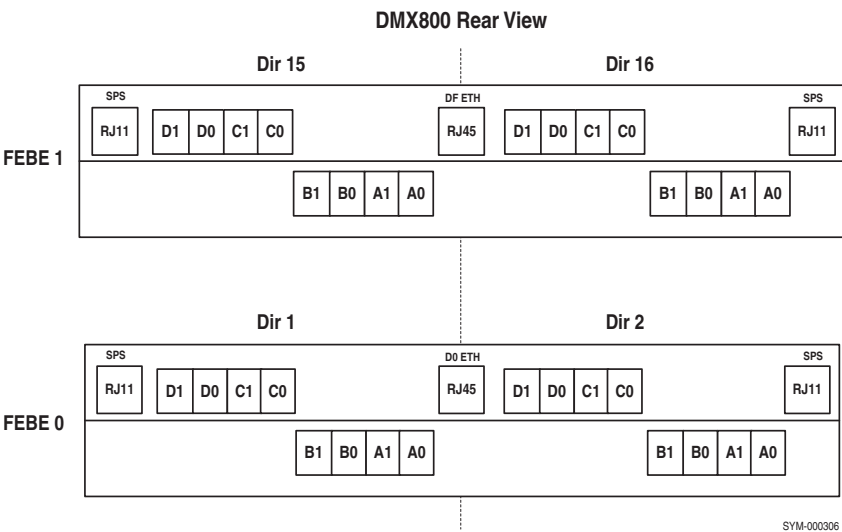


Figure C-4 FEBE Board Channel Designations

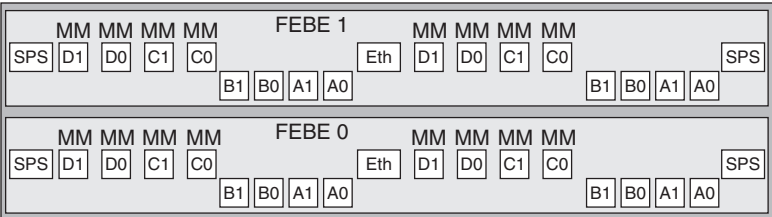
Fibre Channel/FEBE Port Designations

EMC offers several Fibre Channel/FEBE port configurations.

The number of back-end ports used depends upon the number of DAEs in the system. Each DAE requires two back-end ports.

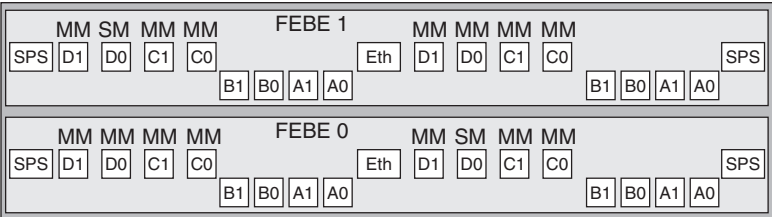
All configurations use the port pairing shown in Figure 2-4 on page 2-10.

- ◆ **DMX-FE-8M0S**—Eight Fibre Channel multimode ports, no Fibre Channel single-mode ports:



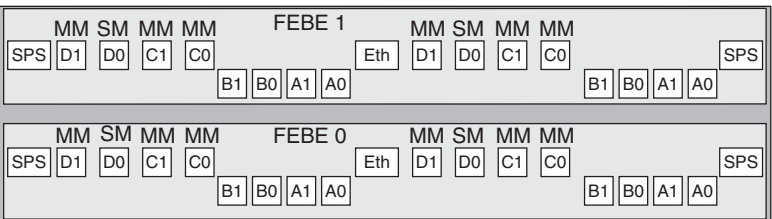
SYM-000439

- ◆ **DMX-FE-7M1S**—Seven Fibre Channel multimode ports , one Fibre Channel single-mode port:



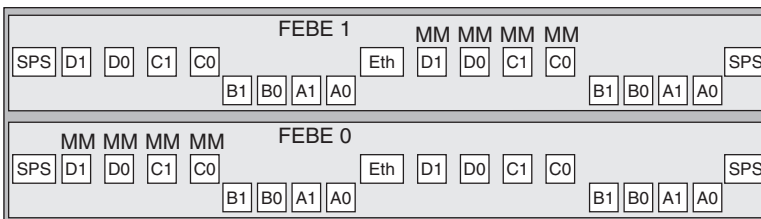
SYM-000438

- ◆ **DMX-FE-6M2S**—Six Fibre Channel multimode ports, two Fibre Channel single-mode ports:



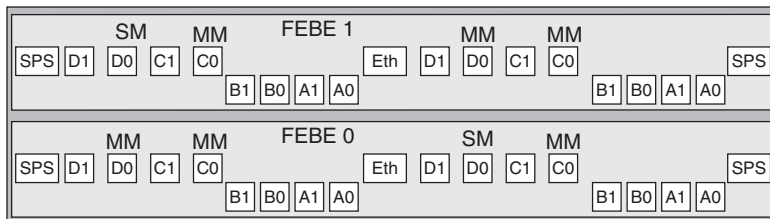
SYM-000437

- ◆ **DMX-FE-4M0S**—Four Fibre Channel multimode ports and zero Fibre Channel single-mode ports:



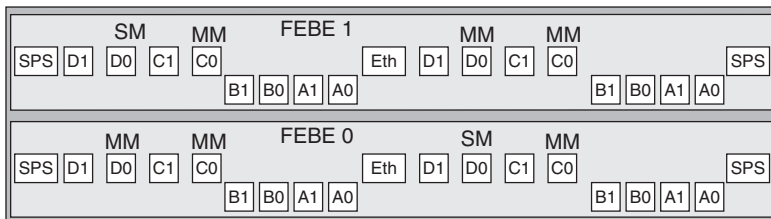
SYM-000435

- ◆ **DMX-FE-3M1S**—Three Fibre Channel multimode ports and one Fibre Channel single-mode port:



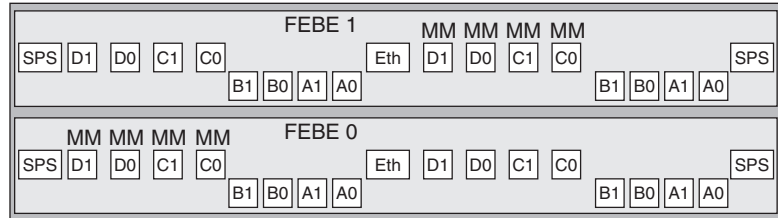
SYM-000435

- ◆ **DMX8-3100**—Three Fibre Channel multimode ports and one Fibre Channel single-mode port:



SYM-000428

- ◆ **DMX8-4000**—Four Fibre Channel multimode ports and zero Fibre Channel single-mode ports:



SYM-000436

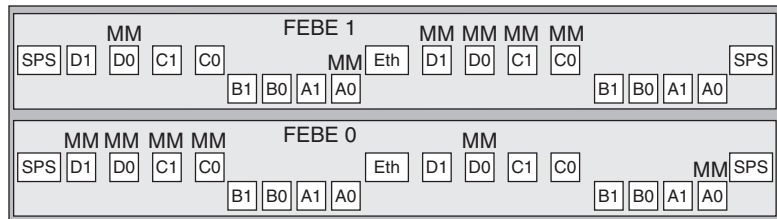
GigE/iSCSI FEBE Port Designations

EMC offers several GigE/iSCSI/FEBE port configurations.

GigE/iSCSI Directors

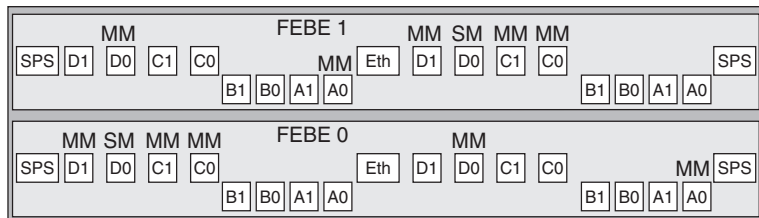
The following GigE/iSCSI FEBE port configurations are offered:

- ◆ **DMX8-4020**—Four Fibre Channel multimode ports, two GigE/iSCSI multimode ports:



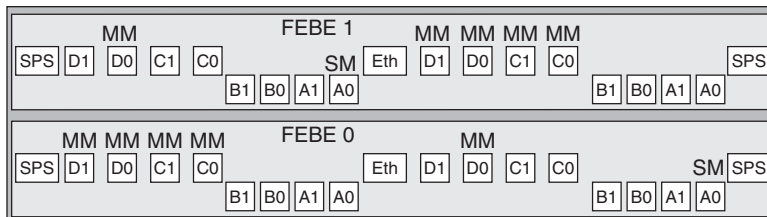
SYM-000434

- ◆ **DMX8-3120**—Three Fibre Channel multimode ports, one Fibre Channel single-mode port, two GigE/iSCSI multimode ports:



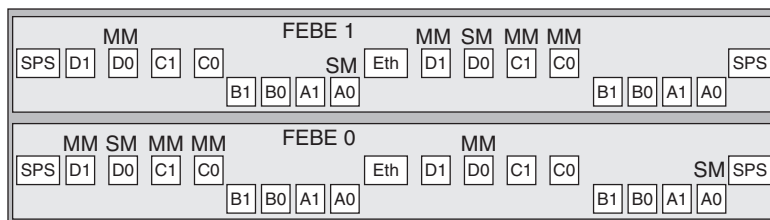
SYM-000431

- ◆ **DMX8-4011**—Four Fibre Channel multimode ports, one GigE/iSCSI multimode port, and one FICON single-mode port:



SYM-000433

- ◆ **DMX8-3111**—Three Fibre Channel multimode ports, one Fibre Channel single-mode port, one Gige/iSCSI multimode port, and one FICON single-mode port:



SYM-000430

Available EMC Fibre Channel Cables

Table C-5 lists the Symmetrix Fibre Channel cables currently available from EMC. To obtain these cables, contact your EMC Sales Representative.

Table C-5 EMC Fibre Cables —Fibre Channel Connect

Multimode Fibre Cables - 50 Micron		Single-mode Fibre Cables - 9 Micron	
Model Number	Description	Model Number	Description
FC1M-50MSLC	SC/LC-1 meter w/SC coupler	FC1M-9MSLC	SC/LC-1 meter w/coupler
FC3M-50MSLC	SC/LC-3 meter w/SC coupler	FC3M-9MSLC	SC/LC-3 meter w/coupler
FC5M-50MSLC	SC/LC-5 meter	FC5M-9MSLC	SC/LC-5 meter
FC10M-50MSLC	SC/LC-10 meter	FC10M-9MSLC	SC/LC-10 meter
FC30M-50MSLC	SC/LC-30 meter	FC30M-9MSLC	SC/LC-30 meter
FC50M-50MSLC	SC/LC-50 meter	FC50M-9MSLC	SC/LC-50 meter
FC100M-50MSLC	SC/LC-100 meter	FC100M-9MSLC	SC/LC-100 meter
FC1M-50MLC	LC/LC-1 meter	N/A	N/A
FC3M-50MLC	LC/LC-3 meter	N/A	N/A
FC5M-50MLC	LC/LC-5 meter	FC5M-9MLC	LC/LC-5 meter
FC10M-50MLC	LC/LC-10 meter	FC10M-9MLC	LC/LC-10 meter
FC30M-50MLC	LC/LC-30 meter	FC30M-9MLC	LC/LC-30 meter
FC50M-50MLC	LC/LC-50 meter	FC50M-9MLC	LC/LC-50 meter
FC100M-50MLC	LC/LC-100 meter	FC100M-9MLC	LC/LC-100 meter

Fibre Channel Cable Precautions

When connecting Fibre Channels channels to host channels, it is important to familiarize yourself with the DMX800 Fibre Channel hardware components and their channel designations.

Each Symmetrix Fibre Channel director occupies one slot on the Symmetrix DMX800 backplane. Each director interfaces to the host channels through a FEBE board in the backplane card cage. The Symmetrix Fibre Channel directors contain four advanced microprocessors. Refer to Figure C-1 on page C-7 for the FEBE board port designations, and *Fibre Channel/FEBE Port Designations* on page C-17.

GigE Remote and iSCSI Director Installations

GigE and iSCSI Cable Precautions

The physical connection to a GigE/iSCSI interface occurs on the FEBE board in the DMX800 backplane. Refer to Figure C-4 on page C-16, *GigE/iSCSI FEBE Port Designations* on page C-19.

When connecting to hosts with differential GigE/iSCSI interfaces, consult your EMC Sales Representative for the most current list of supported hosts, models, operating systems, and EMC open systems host support policies, or refer to the EMC Powerlink website at: <http://powerlink.emc.com>. From the Powerlink home page, select the menu options: Services, Document Library, Host Connectivity.

Available EMC GigE/iSCSI Channel Cables

Table C-6 lists the Symmetrix GigE/iSCSI channel cables currently available from EMC. To obtain these cables, contact your EMC Sales Representative..

Table C-6 GigE/iSCSI Channel Cables

Model	Description	Where Used
FC5M-62MSLC	SC/LC-5 meter	For Symmetrix DMX GigE/iSCSI connection to patch panels or Ethernet switches using LC connectors.
FC10M-62MSLC	SC/LC-10 meter	
FC30M-62MSLC	SC/LC-30 meter	
FC50M-62MSLC	SC/LC-50 meter	
FC100M-62MSLC	SC/LC-100 meter	
FC1M-62MLC	LC/LC-1 meter	For Symmetrix DMX GigE/iSCSI connection to patch panels or Ethernet switches using LC connectors.
FC3M-62MLC	LC/LC-3 meter	
FC5M-62MLC	LC/LC-5 meter	
FC10M-62MLC	LC/LC-10 meter	
FC30M-62MLC	LC/LC-30 meter	
FC50M-62MLC	LC/LC-50 meter	
FC100M-62MLC	LC/LC-100 meter	

This glossary contains terms related to disk storage subsystems. Many of these terms are used in this manual.

A

Alternate Track	A track designated to contain data in place of a defective primary track. See also <i>Primary Track</i> .
Actuator	A set of access arms and their attached read/write heads, which move as an independent component within a head and disk assembly (HDA).
Adapter	Board that provides the physical interface between the Fibre Channel director and disk devices.
ADT	Automatic Diagnostic Test.
ANSI	American National Standards Institute. A standards-setting, nongovernment organization which develops and publishes standards for voluntary use in the U.S.
Asynchronous Transmission	Transmission in which synchronization is done on a per-byte basis. The synchronizing handshake is done using REQuest and ACKnowledge signals.

B

Backplane	Card that accommodates the director, global memory, and adapter cards.
------------------	--

Bit	The smallest unit of computer memory. A bit can hold a value of zero or one.
Business Continuance Volumes (BCVs)	A standard Symmetrix device with special attributes that allow it to independently support applications and processes, such as backup operations, restore operations, decision support operations, and application testing. BCV devices are available through the <i>EMC TimeFinder</i> software.
Byte	Any eight-bit unit of data storage.
C	
Cache	<i>See Global Memory.</i>
Channel Director	The component in the DMX800 subsystem that interfaces between the host channels and data storage. It transfers data between the channel and global memory.
CCOPY	Concurrent Copy Facility used on IBM DASD.
Command Descriptor Block (CDB)	The structure used to communicate commands from an initiator to a target. This structure may be six bytes, 10 bytes, or 12 bytes in size.
Compatible Parallel Access Volumes (COM-PAV)	Compatible Parallel Access Volumes is an IBM feature, supported by the Symmetrix unit, that improves response time, resulting in greater resource availability to the host.
Controller ID	Controller identification number of the director the disks are channeled to for EREP usage. There is only one controller ID for DMX800.
Count-Key-Data (CKD)	A data recording format employing self-defining record formats in which each record is represented by a count area that identifies the record and specifies its format, an optional key area that may be used to identify the data area contents, and a data area that contains the user data for the record. CKD can also refer to a set of channel commands that are accepted by a device that employs the CKD recording format.

D

DASD	Direct access storage device.
Data Availability	Access to any and all user data by the application.
Delayed Fast Write	An operation used when there is no room in global memory for the data presented by the write operation.
Destage	The asynchronous write of new or updated data from global memory to disk device.
Device	A uniquely addressable part of the DMX800 subsystem that consists of a set of access arms, the associated disk surfaces, and the electronic circuitry required to locate, read, and write data. See also <i>Volume</i> .
Device Address	The hexadecimal value that uniquely defines a physical I/O device on a channel path in an MVS environment. See also <i>Unit Address</i> .
Device Form Factor	Refers to the physical size and shape of a device. Example: 36 GB disk device is a one-inch disk device in a one-inch form factor.
Device Number	The value that logically identifies a disk device in a string.
Device Support Facilities Program (ICKDSF)	A program used to initialize DMX800 at installation and provide media maintenance.
Diagnostics	System-level tests or firmware designed to inspect, detect, and correct failing components. These tests are comprehensive and self-invoking.
Direct Matrix Architecture (DMX)	New line of Symmetrix systems encompassing the DMX800 (rackmount), DMX1000, DMX2000, and DMX3000 systems.
Director	The component in the DMX800 subsystem that allows DMX800 to transfer data between the host channels and disk devices. See also <i>Channel Director</i> and <i>Disk Director</i> .
Disk Director	The component in the Symmetrix subsystem that interfaces between global memory and the disk devices.
Disk Array Enclosure (DAE)	This enclosure houses from four to 15 disk drives per DAE (73 GB, 146 GB). There is a minimum of two DAEs and a maximum of eight DAEs per each DMX800 cabinet.

Dynamic Sparing A DMX800 feature that automatically transfers data from a failing disk device to an available spare disk device without affecting data availability. This feature supports all non-mirrored devices in the DMX800 subsystem.

Dynamic Path Reconnect (DPR) A function that allows disconnected I/O operations with the Symmetrix system to reconnect over any available channel path rather than be limited to the one on which the I/O operation was started. This function is available only on System 370/XA, System 370/ESA, and System 390/ESA systems.

E

EMIF ESCON Multiple Image Facility. Allows sharing of ESCON channels among logical partitions (LPAR).

Enginuity Enginuity is the operating environment for the EMC Symmetrix Enterprise Storage Platforms. Enginuity provides functional services for both its host Symmetrix systems as well as for a large suite of EMC storage application software.

EPO Emergency Power Off.

EREP Program The program that formats and prepares reports from the data contained in the Error Recording Data Set (ERDS).

Error Verification The process of reading, checking the error correction bits, and writing corrected data back to the source.

ESA Enterprise System Architecture (mainframe systems only).

ESP Enterprise Storage Platform. Symmetrix ESP is a functional enhancement that allows simultaneous storage and access of mainframe data and open systems data on the same Symmetrix system.

F

Fast Write In DMX800, a write operation at global memory speed that does not require immediate transfer of data to disk. The data is written directly to global memory and is available for later destaging.

FBA Fixed Block Architecture. Disk device data storage format using fixed size data blocks.

Fibre Channel A nominally 2 Gb/s serial data transfer interface technology, although the specification allows data transfer rates from 133 Mb/s up to 4.25 Gb/s. Data can be transmitted and received simultaneously.

Common transport protocols, such as Internet Protocol (IP) and Small Computer Systems Interface (SCSI), run over Fibre Channel. Consequently, a single connectivity technology can support high-speed I/O and networking.

FC-AL Fibre Channel Arbitrated Loop. A standard for a shared access loop, in which a number of Fibre Channel devices are connected (as opposed to point-to-point transmissions). Requires a port to successfully negotiate to establish a circuit between itself and another port on the loop.

FC-SW Fibre Channel Switch Fabric. A Fibre Channel network standard where nodes are connected to a fabric topology implemented by one or more switches. Each node's N_Port connects to an F_Port on a switch. Pairs of nodes connected to an FC-SW network can communicate concurrently.

Field Replaceable Unit (FRU) A component that is replaced or added by service personnel as a single entity.

Frame Data packet format in an ESCON environment.

Front End Back End (FEBE) Board The FEBE board, unique to the DMX800 system, contains both front-end and back-end connections.

G

Gigabyte (GB) 10^9 bytes.

Global Memory Random access electronic storage used to retain frequently used data for faster access by the channel.

Global Memory Slot Unit of global memory equivalent to one track.

Group The physical disks related to each other for common parity protection when implementing the parity RAID option.

H

HACMP	High Availability Clustered Multiprocessing. IBM's cluster system developed by CLAM Associates.
Head and Disk Assembly (HDA)	A field replaceable unit in the DMX800 subsystem containing the disk and actuator.
Home Address (HA)	The first field on a CKD track that identifies the track and defines its operational status. The home address is written after the index point on each track.
Hypervolume Extension	The ability to define more than one logical volume on a single physical disk device making use of its full formatted capacity. These logical volumes are user-selectable in size. The minimum volume size is one cylinder and the maximum size depends on the disk device capacity and the emulation mode selected.

I

ICKDSF	See <i>Device Support Facilities program</i> .
Identifier (ID)	A sequence of bits or characters that identifies a program, device, controller, or system.
IML	Initial microcode program loading.
Index Marker	Indicates the physical beginning and end of a track.
Index Point	The reference point on a disk surface that determines the start of a track.
INLINES	An EMC-provided host-based global memory Reporter utility for viewing short and long-term global memory statistics at the system console.
IOCP	Input/Output Configuration Program.
I/O Device	An addressable input/output unit, such as a disk device.
IPL	Initial Program Loading.

K

Keyboard, Video Display, and Mouse (KVM)

The KVM is an optional item in a 1U rack that connects to the DMX800 server to monitor system operation.

Kilobyte (K)

1024 bytes.

KVM

See *Keyboard, Video Display, and Mouse (KVM)*.

L

Least Recently Used Algorithm (LRU)

The algorithm used to identify and make available the global memory space by removing the least-recently-used data.

Link Control Card (LCC)

These cards are located at the rear of each DAE. Each link control card supports and controls one FC loop and monitors the DAE environment. There are two link control cards per DAE.

A physical or virtual device addressable through a target. A physical device can have more than one logical unit.

Logical Unit Number (LUN)

An encoded three-bit identifier for the logical unit of a SCSI device.

Logical Volume

A user-defined storage device.

Long Miss

Requested data is not in global memory and is not in the process of being fetched.

Longitude Redundancy Code (LRC)

Exclusive OR (XOR) of the accumulated bytes in the data record.

LPAR

Logical Partition or Logical Partitioning Mode. Mainframe system function that allows different operating systems to run concurrently in separate logical partitions. See also EMIF.

M

MA	Multiple Allegiance is an IBM feature, supported by the Symmetrix system, that improves throughput across a shared storage environment. MA allows more than one I/O from different hosts to access the same device as long as the I/Os do not conflict with each other.
Media	The disk surface on which data is stored.
Megabyte (MB)	10 ⁶ bytes.
Metadevice	A group of components (physical partitions) accessed as a single logical device through concatenating, striping, mirroring, logging the physical devices, or setting up RAID devices.
MII	Machine Initiated Interrupt.
Mirroring	The DMX800 maintains two identical copies of a designated volume on separate disks. Each volume automatically updates during a write operation. If one disk device fails, DMX800 automatically uses the other disk device.
Mirrored Pair	A logical volume with all data recorded twice, once on each of two different physical devices.
MMB	Message Matrix Board. Only used in two Fibre Channel director DMX800 systems.
MPCD	Multiprotocol channel director. This director board can contain the following configurations: • Two FICON cards • Two FICON/GigE cards • Two GigE/iSCSI cards
MVS	Multiple Virtual Storage (mainframe systems only).

P

Page	Several commands use regular parameter structures that are referred to as pages. These pages are identified with a value known as a page code.
-------------	--

Partitioned Data Set (PDS) Assist An IBM feature for 3990 Model 6 and 3990 Model 3 with Extended Platform units. PDS Assist improves performance on large, heavily used partitioned datasets by modifying the directory search process.

BX (Private Branch Exchange) Short for private branch exchange, a private telephone network used within an enterprise. Users of the PBX share a certain number of outside lines for making telephone calls external to the PBX.

Physical Partition (PP) A physical partition is the smallest unit of disk space that can be allocated in a volume group in an AIX environment. Any disk space allocated is an integral number of physical partitions. By default, a PP is 4 MB in size.

Physical ID Physical identification number of the DMX800 director for EREP usage. This value automatically increments by one for each director installed in DMX800.

Primary Track The original track on which data is stored. See also *Alternate Track*.

Promotion The process of moving data from a track on the disk device to global memory slot.

R

RAID (Parity RAID) A Symmetrix feature that provides parity data protection on the disk device level using physical parity volumes. A parity RAID (3+1) group consists of three data drives and one parity drive. A parity RAID (7+1) group consists of seven data drives and one parity drive. The parity RAID option can be also used with Hypervolume Extension to establish distributed parity.

Read Hit Data requested by the read operation is in global memory.

Read Miss Data requested by the read operation is not in global memory.

Reconnect The function that occurs when a target selects an initiator to continue an operation after a disconnect.

Reconnection A reconnection exists from the assertion of the BSY signal in a RESELECTION phase until the next BUS FREE phase occurs. A reconnection can only occur between a target and an initiator.

Record Zero The first record after the home address.

Reselect The function that occurs when a target disconnects from an initiator in order to perform a time-consuming function and then, after performing that function, reestablishes the connection.

Reserved The term used for bits, bytes, fields, and code values that are set aside for future standardization.

S

SCU_ID For 3880 storage control emulations, this value uniquely identifies the storage director without respect to its selection address. It identifies to the host system, through the EREP, the director detecting the failing subsystem component. This value automatically increments by one for each director installed. The SCU_ID must be a unique number in the host system. It should be an even number and start on a zero boundary.

Scrubbing The process of reading, checking the error correction bits, and writing corrected data back to the source.

SDMS Symmetrix Data Migration Service allows continuous business operations and continuous data availability while data is automatically migrated from existing DASD to Symmetrix 5000 systems.

Short Miss Requested data is not in global memory, but is in the process of being fetched.

SIO Start I/O.

SRDF Symmetrix Remote Data Facility. SRDF consists of the microcode and hardware required to support Symmetrix remote mirroring.

SSID For the Symmetrix DMX800 system storage control emulations, this value identifies the physical components of a logical DASD subsystem. The SSID must be a unique number in the host system. It should be an even number and start on a zero boundary.

Stage The process of writing data from a disk device to global memory.

Standby Power Supply (SPS) There are two SPS modules in each SPS assembly housed in a 1U rack. One SPS assembly within the storage processing enclosure (SPE) supplies redundant power to the SPE. Each DAE has an SPS module that provides redundant power.

Storage Control Unit The component in the DMX800 subsystem that connects DMX800 to the host channels. It performs channel commands and communicates with the disk directors and global memory. See also *Channel Director*.

Storage Processing Enclosure (SPE) The SPE contains: one SPS assembly (2 SPS modules in a 1U rack), DMX800 card cage, a fan assembly, and two power supplies.

String A series of connected disk devices sharing the same disk director.

Striping The process of segmenting logically sequential data and writing the segments to multiple physical disk devices.

Symmetrix DMX Symmetrix Direct Matrix is the architecture of nonblocking direct matrix interconnect design that supports up to 32 direct paths through the Symmetrix DMX global memory directors for the Symmetrix DMX800 systems.

T

Target A SCSI device that performs an operation requested by an initiator.

Terabyte (TB) The TB₂ value is based on the convention:

1 TB = 1024x1024x1024x1024 bytes.

The TB₁₀ value is based on the convention:

1 TB = 1000x1000x1000x1000 bytes.

Although these values are expressed differently, their values are equivalent.

U

Unit Address The hexadecimal value that uniquely defines a physical I/O device on a channel path in an MVS environment. See also *Device Address*.

UNIX UNIX is an interactive, multitasking, multiuser operating system. UNIX is written in "C" language.

There are three types of UNIX files: Directories, data files, and special files. A directory is a file containing certain information about another file. A directory contained within another directory is a sub-directory.

The two most common types of UNIX are BSD (Berkeley Software Distribution) and System VR4 (developed by AT&T). Most UNIX systems are a *mix* of both types.

V

Volume A general term referring to a storage device. In the DMX800 subsystem, a volume corresponds to single disk device.

W

Write Hit There is room in global memory for the data presented by the write operation.

Write Miss There is no room in global memory for the data presented by the write operation.

Numerics

3390 DASD
track format 2-24

A

Adaptive copying mode 5-5
address mark 2-24
APIs 1-7
Automated Network Storage 1-4
availability features
disk error correction and verification 5-16
dynamic sparing 5-40 to 5-45
error checking and correction 5-15 to 5-16
mirroring 5-21, 5-22
multiport volume access (SCSI) 4-23
nondisruptive component replacement 5-11
redundant power subsystem 5-11
Availability guidelines
SRDF/S option 5-4

B

battery backup
power failure 5-11
BCV pairs 5-23
block capacity 2-22, 4-15
block format sizes 2-22
business continuance volumes (BCVs) 5-23

C

cables
fiber optic
multimode 2-38

single-mode 2-38
cabling requirements C-6
card cage 2-14
Catalog Solution 1-19
channel
attachments 2-38 to 2-41
cable length 4-12
interfaces C-21
channel director boards 1-13
channel status bits 6-22
Compatible Parallel Access Volumes (COM-PAV)
1-11, 6-5
COM-PAV 1-11
Concurrent Copy 6-4
connect time 3-6, 3-7
cooling fan modules 1-13
CopyCross 1-19
Count-key-data (CKD) format 2-23
cylinders
maximum per logical volume 4-14

D

DASD Data Transfer Summary 6-18
data
delayed fast write 3-12
fast write 3-10
read hit 3-9
read miss 3-10
data and command formats 2-23 to 2-25
data records 2-24
home address (HA) 2-24
index marker 2-24
record zero (RO) 2-24

- track capacity 2-25
- track format 2-24
- data check error 6-11, 6-12
- data integrity protection 5-14 to 5-18
 - error verification 5-2
 - periodic system check tests 5-2
- data protection options 5-3, 5-6
 - dynamic sparing 5-6
 - mirroring 5-3, 5-21 to 5-22
- data recovery
 - dynamic sparing with RAID 5-43
 - RAID 5-33 to 5-36
- data storage
 - disk emulation 2-23
- defective tracks 6-13
- delayed fast write operation 3-12
- detecting the error 6-19
- determining error source 6-20
- differential adapter C-21
- director/global memory configuration C-8
- directors 2-26 to 2-34
- disconnect time 3-6, 3-7
- disk capacities 2-19
- disk devices 1-13, 2-18
 - capacity 4-20
 - emulation types 2-23
 - error correction and verification 5-16
 - FBA data and command format 2-22
 - form factor A-9
 - RPS miss elimination 3-9
- disk emulation 2-18, 2-23, 2-25
- disk error correction 5-16
- disk error verification 5-16
- disk subsystem operation 3-2 to 3-5
- distributed parity 5-29
- DMX800
 - backplane 3-19
 - capacities 1-8
 - configuration options 2-7
 - service area C-7
- domestic shipments C-3
- Dynamic Mirror Service Policy 4-5, 4-9
- Dynamic Path Reconnection (DPR) 6-3
- dynamic sparing 5-6, 5-40 to 5-45
 - advantages 5-41
 - with SRDF 5-45

E

- Electrical specifications A-14
- EMC APIs 1-7
- EMC ControlCenter 1-17
- EMC ControlCenter Automated Resource Manager 1-17
- EMC ControlCenter Replication Manager 1-19
- EMC ControlCenter StorageScope 1-17
- EMC ControlCenter Symmetrix Data Mobility Manager 1-19
- EMC ControlCenter Symmetrix Manager 1-17
- EMC ControlCenter Symmetrix Optimizer 1-17
- EMC ControlCenter Workload Analyzer for Symmetrix 1-17
- EMC CopyPoint for OS/400 1-15, 1-19
- EMC Data Manager 1-19
- EMC Double Checksum 1-15
- EMC Service Support Center 1-13
- Enginuity 1-2
- Enginuity operating environment 1-6
- environmental alert messages 6-15
- environmental specifications C-6
- equipment checks 6-11, 6-13
- EREP file 6-11
- EREP reports 6-18 to 6-19
 - DASD Data Transfer Summary 6-18
 - subsystem exception DASD 6-18
 - system error summary (part 2) 6-18
- error checking and correction
 - disk 5-16
- error handling
 - detecting the error 6-19
 - determining error source 6-20
- error recoverability 6-12 to 6-13
- error reporting 6-13 to 6-19
 - EREP reports 6-18 to 6-19
- error verification 5-2
- errors 6-11 to 6-13
 - permanent 6-11
 - temporary 6-11
- evaluating performance 4-22
- event messages 6-16
- Extended Count-Key-Data (ECKD) 2-23
- extended cylinder addressing 4-18, 4-19

F

fast write ceiling 4-5
 fast write operation 3-10, 4-5
 FEBE boards 1-13
 configuration options C-17, C-19
 Fibre Channel front-end director 2-27
 Fibre optic cables 2-41
 FICON 2-29
 FICON cascading 2-40
 FICON Open Systems Intermix 2-40
 front end back end (FEBE) 2-32, 3-20

G

global memory
 basic operation 3-9
 host cache use 3-2
 management 3-3 to 3-5
 prefetch algorithm 3-5
 requirements for PermaCache 4-6
 sizes 2-34
 global memory director 2-34, 4-3
 global memory director boards 1-13
 global memory techniques 3-26
 disk microprocessor and buffer 3-26
 high speed global memory 3-26
 sequential access patterns 3-26
 split director functions 3-26

H

high availability
 multiport volume access 4-23
 high performance files 4-19
 Home Address (HA) 2-24
 host integration 2-15 to 2-17
 hosts
 channels 2-15 to 2-16
 PC servers 4-15
 supported Fibre Channel hosts 2-16
 UNIX 4-15
 hypervolumes 4-14
 mainframe device emulations 4-21
 mainframe extended cylinder addressing 4-19
 mainframe split-volume capability 4-18
 open systems 4-14 to 4-17

I

I/O operations
 delayed fast write 3-12
 fast write 3-10
 overview 3-6 to 3-13
 read hit 3-9
 read miss 3-10
 I/O response time
 connect time 3-7
 connect time (mainframe) 3-6
 disconnect time 3-6, 3-7
 pend time (mainframe) 3-6
 queuing time 3-7
 queuing time (mainframe) 3-6
 IBM DASD 2-21
 index marker 2-24
 Information Management
 EMC ControlCenter 1-17
 Infrastructure Services Software Options
 EMC PowerPath 1-20
 installation planning C-2 to C-7
 mainframe systems C-10 to C-12
 open systems C-13 to C-22
 international shipments C-3
 iSeries Fibre Channel connectivity 2-15

L

logical volume capacities 2-19
 logical volumes 2-20
 capacity calculation 2-22
 LPAR 2-38

M

main components 2-2 to 2-3
 mainframe
 operating systems supported 2-16
 mainframe split-volume 4-18
 mirroring 5-3
 advantages 5-21
 mixed track geometries 2-23
 Multi-Path Lock Facility (MPLF) 6-4
 Multiple Allegiance (MA) 6-5
 multiple LRU lists 3-4
 multi-subsystem imaging 6-4
 MVS error records 6-18

N

nondisruptive component replacement 5-11
 non-RAID Mode 5-30
 normal mode/parity generation 5-29

O

operating systems supported
 mainframe hosts 2-16
 open systems hosts 2-15
 operator messages 6-16
 overrun error 6-11, 6-13

P

parallel processing 4-13
 Parity RAID 1-8, 5-29
 XOR logic 5-29
 Parity RAID (3+1) 1-8
 parity rebuild 5-30
 Partitioned Dataset (PDS) Search Assist 6-5
 PC server hosts 4-15
 PDS Assist 6-5
 pend time 3-6
 performance features
 fast write capabilities 4-5
 hypervolume extension 4-14 to 4-15, 4-18
 mainframe systems hypervolume extension
 4-18 to 4-21
 multiple channel directors 4-11
 multiport volume access 4-23
 parallel processing 4-13
 PermaCache 4-6
 permanent errors 6-11
 planning overview C-2 to C-16
 Platform Software Options
 EMC InfoMover 1-15
 Symmetrix Enterprise Storage Platform 1-15
 power requirements C-3
 power subsystem 2-35, 5-11
 power supplies 1-13
 powering down
 for service B-2
 powering the system on
 after a shutdown B-4
 prefetch algorithm 3-5
 program check error 6-11

Q

queuing time 3-6, 3-7

R

RAID 5-24 to 5-36
 data recovery 5-31 to 5-36
 data recovery with hyper-volumes 5-35
 data volumes 5-35
 device-level parity protection 5-32
 disk devices 5-25
 Hypervolume Extension 5-28
 modes of operation 5-29
 non-Raid mode 5-29
 normal mode 5-29
 parity rebuild 5-29
 reduced mode 5-29
 parity rebuild 5-36
 parity volumes 5-35
 read/write operations 5-34
 with hypervolume extension 5-28
 writing data 5-34
 RAID 10 6-7
 RAID 5 5-37, 5-38
 read miss 3-10
 read operations 3-9 to 3-10
 long miss 3-10
 read hit 3-9
 read miss 3-10
 short miss 3-10
 read/write operations 3-24
 reassigning defective tracks 6-12
 Record Zero (R0) 2-24
 reduced mode/regeneration 5-29
 reliability features
 disk error correction and verification 5-16
 dynamic sparing 5-40 to 5-45
 error checking and correction 5-15 to 5-16
 mirroring 5-21
 nondisruptive component replacement 5-11
 redundant power subsystem 5-11
 remote Fibre Channel directors 2-29, 2-34
 remote support C-7
 RPS Miss Elimination 4-13

S

- SAN (Storage Area Network) Manager 1-17
- segmented data buffer(ing) 4-13
- sequential data access patterns 3-5
- sequential data striping 6-5
- server 1-13
- service alert messages 6-16
 - events 6-16
- service alerts 6-15
- serviceability features 1-13
- short miss 3-10
- shutdown
 - for service B-2
- soft error thresholds 5-16
- software options 1-15 to 1-20
- space requirements C-6
- split-volume option
 - mainframe 4-18 to 4-21
- SRDF 1-8
- SRDF/Automated Replication (SRDF/AR) 5-5
- SRDF/Consistency Groups (SRDF/CG) 5-6
- SRDF/Synchronous (SRDF/S formerly SRDF) 5-4
- SSID 6-4
- standby power supplies 1-13
- standby power supplies (SPS) 2-37
- storage directors 4-11
- Storage Management
 - software 1-17
- Subsystem Exception DASD 6-18
- Symmetrix
 - cache 2-34
 - channel interfaces C-21
 - dynamic sparing 5-40 to 5-45
 - hardware components 1-13
 - I/O operations 3-6 to 3-12
 - mirroring 5-21
 - PermaCache 4-6
 - physical specifications C-2
 - RAID option 5-24 to 5-36
 - serviceability features 1-13
 - split-volume option 4-20 to 4-21

- Symmetrix DMX800 slot configuration 2-14
- Symmetrix file system (SFS) 4-9
- Symmetrix Optimizer 4-9
- Symmetrix RAID option
 - regenerating parity 5-36
 - write operations 5-30
 - write requests 5-34
- Symmetrix Remote Data Facility (SRDF) 5-4
- system check error 6-11
- system check tests 5-2

T

- temporary errors 6-11
- track capacity 2-25
- track format 2-24
- transportation and delivery guidelines C-2
- types of errors 6-11 to 6-13
 - data check 6-11
 - equipment checks 6-11
 - overrun 6-11
 - system or program check 6-11

U

- unit status bits 6-22
- UNIX 2-21
- UNIX hosts 4-15

V

- VM error records 6-18

W

- windows 2-21, 2-26
- write operations
 - delayed fast write 3-12
 - fast write 3-10

