



Global Distributed Storage Engineering

EMC Symmetrix DMX

GDSE-03038-SG

Solution Guide

4.0

1. Version Control

Version	Date	Author	Reviewed	Comments
Draft	06-Oct-03	Burt Giron GDSE	-	RFC
1.0	27-Oct-03	Burt Giron	Storage Community	Release
1.1	12-Jul-04	Burt Giron	Storage Community	RFC
1.2	8-Sep-04	Scott Torgerson	Burt Giron	HP-UX Additions
1.3	15-Dec-04	Parul Jain	-	Added new design for vendor remote access: Section 15. Added reference to the SAN Secure Configuration Guide in the "Introduction" section
1.4	8-April-05	Parul Jain	Global Storage Task Force	Updated BIN file standard: Section 10
2.0	11-April-05	Parul Jain	-	Updated Management Section to point to SE SG, updated references section, removed McData fabric section (this is now covered in the McData Switch SG), updated port BIT settings section to include heterogeneous. Section numbering has changed!
2.1	26-May-05	Parul Jain, Eliot Wilson	GDSE	Made changes to BIN file standards, Section 10
2.2	13-Jul-05	Kevin Hobaugh	GDSE	1) Added eTrust Appendix 2) Updated quorum std

2.3	17-Oct-05	Parul Jain	Eliot Wilson, Jesse Reyes	1. Changed name from "Best Practice Guide" to "Solution Guide" on title page 2. Fixed statement about PER bit with DMP in section 12 3. Fixed link to vendor remote access document in section 15 4. Enhanced section 15 with implementation planning guide (15.1) 5. Made section 10 easier to read 6. Replaced or added cylinder count for GB throughout 7. Fixed BIN file drawings to match standard from section 10 8. Removed references to "Secure Configuration Guide". All security information is contained within this document itself.
2.4	22-May-06	Ludwig Derodel		Added DMX-3 Support
2.5	31-May-06	Guy Shimabuku		Added RAID-5 section
3.0	06-Jun-06	Ludwig Derodel	Chris Clayton	RFC
3.01	12-July-06	Ray Young, Ludwig Derodel	Peer Review	1. General reformat of document. 2. Moved benchmark data to release notes. 3. Removed parts list 4. Added DMX3000 specs.

				5. Removed references to SOE 6. Moved security section to main body 7. Moved eTrust audit to main body. 8. Added page numbering. 9. Change document number to 03038 10. Added GDSE DMX-3 guideline table
3.02	09-Aug-06	Ludwig Derodel, Guy Shimabuku, Raymond Young		1. Added High level diagram of DMX-3 models 2. Edited section 8.5.3
3.1	31-Oct-06	Ludwig Derodel, Guy Shimabuku		Updated RAID-1 and RAID-5 cylinder change
3.2	19-Jan-07	Ludwig Derodel	Chris Clayton	Included hardware replication section (sec 10)
3.21	1-Feb-07	Ludwig Derodel	Chris Clayton	Included support statement for RAID-5 SRDF (section 10.2)
3.3	19-Feb-07	Ludwig Derodel	Chris Clayton	Updated VCMDDB access restriction section.
3.4	16-Aug-07	Guy Shimabuku	Ludwig Derodel	1. Included support statement for DMX-3 VCMDDB management (section 9.8) 2. VCMDDB write disable procedure (Appendix B)
3.5	08-Oct-07	Eliot Wilson, Guy Shimabuku		1. Section 7.2.2 – Entry Model DMX-3 950. 2. Appendix A – Symmetrix DMX-3 950 Configuration Summary
4.0	25-Jan-08	Guy Shimabuku	Quang Le Chris Clayton Lugwig Derodel	1. Section 8: DMX-4 & Enginuity 5772 certification. 2. Section 10: DMX-4

				summary information. 3. Appendix D – DMX-4 Backend subsystem 4. Appendix E: MTTDL calculations 5. Section 7.2.1. include DMX-4
--	--	--	--	---

2. Legal Notice

Information contained herein is for Internal Use and may be used only for business purposes authorized by Citigroup.

CITI, CITIBANK, CITICORP, CITIGROUP and the Umbrella Device are trademarks and service marks of Citicorp and are used and registered throughout the world. Citibank, Citicorp, Citigroup and their subsidiaries also claim rights in certain other trademarks and service marks contained in these pages.

3. Contact

Please address any comments or questions regarding the information contained in this document via email to the “*GT Global Distributed Storage Engineering” distribution list or alternatively to gdse@citigroup.com.

4. Latest Versions

It is recommended that the latest version of the document be used from the GDSE website <http://engineering.ti.citigroup.net/gdse>. Newer versions may include expanded scope, or comments on issues discovered after deployment.

5. Table Of Contents

1. VERSION CONTROL	2
2. LEGAL NOTICE.....	6
3. CONTACT.....	6
4. LATEST VERSIONS	6
5. TABLE OF CONTENTS	7
6. INTRODUCTION.....	8
7. EMC SYMMETRIX DMX MODELS	9
7.1. DMX-2 MODELS	9
7.2. DMX-3 MODELS	11
8. EMC SYMMETRIX DMX-4 AND ENGINUITY 5772 MICROCODE	16
8.1. INTRODUCTION	16
8.2. RAID 6 INTRODUCTION	19
8.3. SYMMETRIX PRIORITY CONTROLS	22
8.4. SYMMETRIX DMX-4 AND ENGINUITY 5772 CONFIGURATION GUIDELINES:.....	22
9. PHYSICAL INSTALLATION	26
10. SYMMETRIX DMX STANDARD HARDWARE CONFIGURATION	27
10.1. SYMMETRIX DMX-2.....	27
10.2. SYMMETRIX DMX-3.....	27
10.3. SYMMETRIX DMX-4.....	27
10.4. DISK DRIVE STRATEGY	28
10.5. FIBRE ADAPTER CONFIGURATION.....	29
10.6. BIN FILE CONFIGURATION	32
10.7. FA PORT FLAG SETTINGS	37
10.8. LUN NAMING/NUMBERING	38
10.9. VCMDB MANAGEMENT FOR DMX-3	39
11. HARDWARE REPLICATION	40
11.1. IN-FRAME REPLICATION (BCV/CLONE)	40
11.2. OUT -OF-FRAME REPLICATION (SRDF)	42
12. SYMMETRIX DMX MANAGEMENT	44
13. VENDOR REMOTE ACCESS TO SYMMETRIX	44
13.1. NOTES ON WWDS IMPLEMENTATION.....	45
14. ETRUST AUDIT.....	46

15.	SYMMETRIX OPTIMIZER.....	46
16.	SYMMETRIX DMX SECURITY.....	47
16.1.	SAN INFRASTRUCTURE AND MANAGEMENT COMPONENTS.....	47
16.2.	VOLUME CONTROL MASKING DATABASE (VCMDB) - Host/HBA RESTRICTION	48
16.3.	SYMACL – SYMMETRIX/DMX HOST ACCESS CONTROL MANAGEMENT.....	53
17.	APPENDIX A - EMC SYMMETRIX DMX-3 950 CONFIGURATION SUMMARY:.....	70
18.	APPENDIX B – SEAGATE DRIVE SPECS.....	74
19.	APPENDIX C – VCMDB WRITE DISABLE PROCEDURE:.....	75
	REFERENCES.....	79
20.	APPENDIX D - DMX-4 BACKEND SUBSYSTEM	79
21.	APPENDIX E - RAID 6 MEAN TIME TO FAILURE CALCULATIONS..	80
21.1.	DESCRIPTION	80
21.2.	COMPARISON OF RELIABILITY CALCULATIONS:.....	81
21.3.	CONCLUSIONS:.....	81
21.4.	RESULTS OF RELIABILITY CALCULATIONS:.....	81

6. Introduction

The goal of the DMX solution guide is to provide standards and recommended practices for Citigroup storage teams, creating an efficient support model with consistent configurations that enables common controls and policies to be implemented for the distributed storage environment.

This guide defines the requirements for successful hardware installation and configuration of EMC Symmetrix DMX-2 and the new DMX-3 storage arrays. This is a living document and as such, it will be updated as new functionality or services are certified.

This document is intended for all Citigroup technology groups responsible for the deployment and support of enterprise storage. The information compiled in this document has been tested and certified by Global Distributed Storage Engineering (“GDSE”).

The best practices outlined in this document include Standards and Recommendations. Standards should be adhered to for effective common controls. Recommendations are applicable to most environments and should be employed where possible. There are, however, guidelines to be used by the Storage Manager’s best judgment.

7. EMC Symmetrix DMX Models

EMC currently has several DMX model options available to customers today. However, the only models prevalent at Citigroup today are the DMX-2 and DMX-3 models. The DMX-3 is the newly introduced Symmetrix platform by EMC and is being positioned to succeed the DMX-2 platform at some un-determined timeframe.

Note: The current certification of the DMX-3 will not include BCV or RDF recommendations simply due to resource constraints. The replication portion will be addressed in an updated release of this documentation shortly.

7.1. DMX-2 Models

Symmetrix DMX-2 platform has several model options and has been in deployment within Citigroup for the past few years. The DMX1000, DMX2000 and DMX3000 models were determined to have the capacity, connectivity, and throughput characteristics required to support the enterprise storage requirements of the firm. DMX models 1000P, and 2000P do not provide any required or significantly enhanced functionality and were not included in the certification.

Table 1 - Symmetrix DMX1000/2000/3000 Component Overview

Component	DMX1000	DMX2000	DMX3000
Fibre Channel Directors (Front-End)	2 – 6 (1-3 pairs)	2 – 8 (1 – 4 pairs)	2 – 8 (1 – 4 pairs)
Fibre Channel Disk Directors (Back-End)	2 – 4 (1 - 2 pairs)	4 – 8 (2 – 4 pairs)	8 (4 pairs)
Cache Directors	2 – 4 (1 – 2 pairs) Minimum: 16GB Maximum: 128GB	4 – 8 (2 – 4 pairs) Minimum 16GB Maximum: 256GB	4 – 8 (2 – 4 pairs) Minimum 16GB Maximum: 256GB
Fibre Channel Disk Subsystem	Up to 144 low-profile Fibre Channel disk drives.	Up to 288 low-profile Fibre Channel disk drives.	Up to 576 low-profile Fibre Channel disk drives

Integrated Service Processor

The service processor is used to download the Symmetrix system configuration to the directors and provide diagnostic and maintenance utilities for the Symmetrix system. The service processor connects to the Symmetrix subsystem through an Ethernet network or an RS-232 interface. The service processor uses an external modem for communicating with the EMC Customer Support Center when the Symmetrix system detects an error condition.

7.2. DMX-3 Models

The DMX-3 array is built on the same Direct-Matrix architecture as the DMX-2 but provides better scalability and performance in one package. Similar to the DMX-2, the DMX-3 can also be ordered in several models. The DMX1500, DMX2500, DMX3500 and DMX4500 all provide the necessary redundancy and scalability to be classified as enterprise storage arrays. (see Table 2)

Since the DMX-3 architecture is based on a single expandable entity, the various model mentioned above can easily be identified by the number of DA pairs configured in the array. (See table below).

Compared to the DMX-2, the DMX-3 provides a number of enhancements:

- o Redundant global memory
- o Vaulting – cache data is de-staged in case of extended power outage
- o More and faster processors per director
- o Higher drive scalability

Considering performance and scaling requirements, it is the responsibility of the Storage Manager to determine the appropriate Symmetrix DMX model to deploy. **Table 2** listed below, was created as a high level sizing guide based on several different factors (i.e. drive counts and number of drives per loop).

Table 2: DMX-3 Configuration Guidelines

Component	DMX1500	DMX2500	DMX3500	DMX4500
Number of fibre channel directors (DA) ^a	2	4	6	8
Total fibre channel loops per DA pair	8	16	24	32
Minimum number of disk drives ^b	96	192	360	480
Maximum number of disk drives ^c	240	480	720	960
Minimum raw capacity ^d	14.02 TB	28.03 TB	52.56 TB	70.08 TB
Minimum usable capacity ^e	7.01 TB	14.2TB	26.28 TB	35.04 TB
Maximum raw capacity ^f	72 TB	144 TB	321.6 TB	428.8 TB
Maximum usable capacity ^g	63 TB	126 TB	189 TB	252 TB
Min/Max usable global memory (mirrored)	32/64 GB	64/128 GB	96/192 GB	128/256 GB
Recommended connectivity (FA port pairs) ^h	4	8	12	16

Note: The main decision points for Table 2 were drive count and DA pair configuration. Configurations larger than 960 drives are possible but should be carefully planned and may affect overall performance of the array because some loops will be operating above the recommended configuration of 30 drives per loop.

^a DAs should always be deployed in pairs for redundancy

^b Minimum configuration available for sale by EMC

^c Maximum of 30 drives per loops using all loops

^d Using 146GB drives

^e Using RAID 1 and 146GB drives

^f First 4 drives of each loop are 300 GB drives and remaining are 500 GB drives

^g Using 300 GB drives and RAID 5 (7+1)

^h Using processor port 0 only

7.2.1. Growth Considerations for the DMX-3/4 models

During initial implementation of a DMX-3/4 frame it is important to consider capacity growth for the next several years. This consideration aides in selecting the initial DMX-3/4 model:

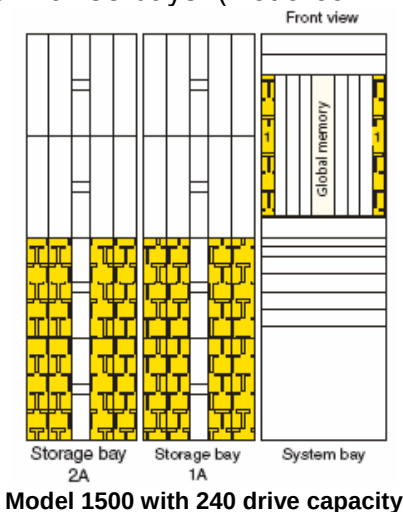
- For example, if an initial deployment has a requirement for 192 drives with a near future requirement of 288 total drives, then a DMX2500 should be selected as the initial DMX-3/4 model.
- Recommended incremental growth should be done in multiples of 96 drives (8 drives per loop) spread across all backend loops.
- Physical space allocation should take into account capacity expansion requiring additional adjacent floor space. In some instances, it may be prudent to deploy a 3-bay configuration initially as opposed to a 2-bay system.
- DMX-2500 and DMX-4500 are the most common DMX-3/4 models purchased and deployed throughout Citigroup, due mainly to our pricing model and configuration flexibility
- The DMX-1500 provides an entry point into the DMX-3/4 family with only a pair of DA adapters. It should be noted that in the rare event of a DA failure, applications may be impacted by performance degradation.

NOTE: If a valid requirement should warrant the use of one of the models not certified or listed above, please contact [GDSE](#) for review.

7.2.1.1. High level configurations of DMX-3/4 models

Symmetrix DMX-3/4 1500

The Symmetrix DMX-3/4 1500 is EMC's entry-level configuration with 1 Disk Adapter (DA) pair, starting at 96 drives. The 1500 model supports array capacities up to 240 drives in three bays. (Not a common configuration)

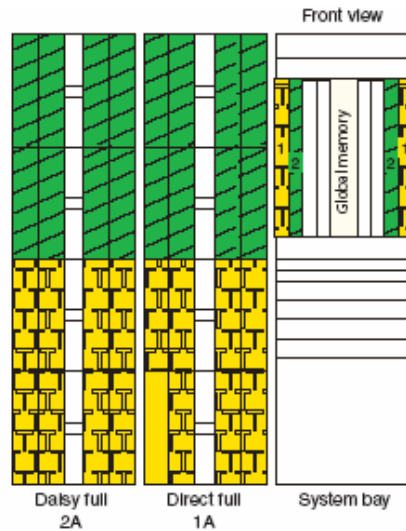


Symmetrix DMX-3/4 2500

The Symmetrix DMX-3/4 2500 contains 2-DA pair supporting 192 - 480 drives in two - three bays. Extending the number of drives beyond 480 requires additional DA pair be installed.



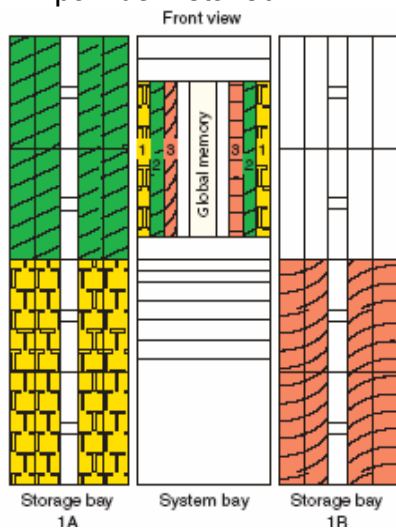
2500 with 240 drive capacity



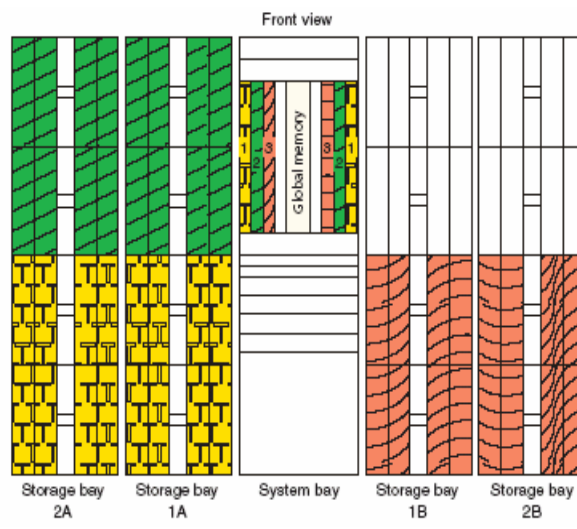
2500 with 480 drive capacity

Symmetrix DMX-3/4 3500

The Symmetrix DMX-3/4 3500 contains 3-DA pair supporting 360-720 drives in three - five bays. Extending the number of drives beyond 720 requires additional DA pair be installed.



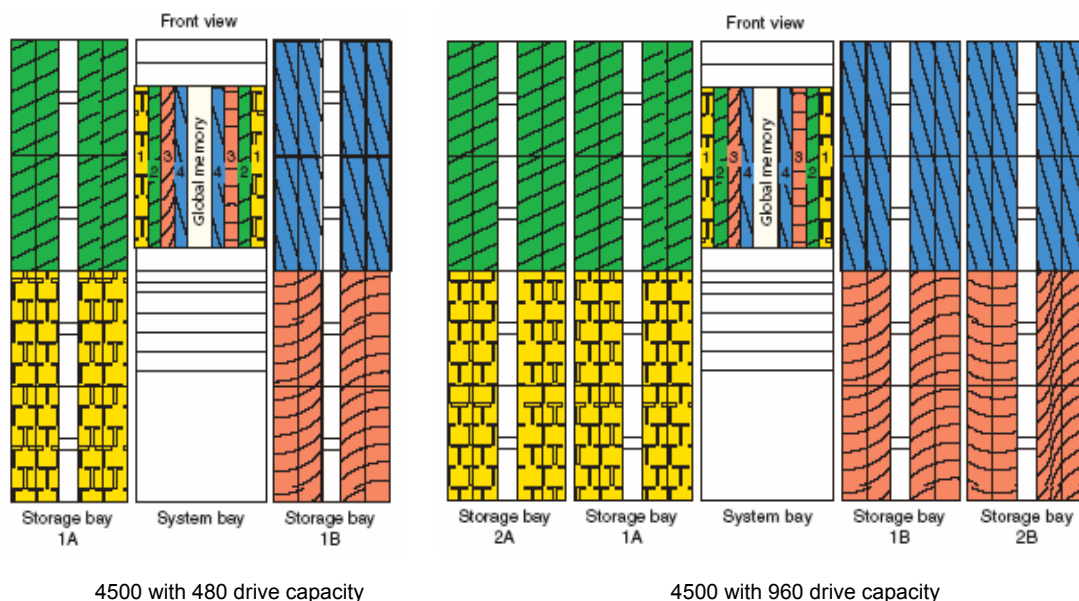
3500 with 360 drive capacity



3500 with 720 drive capacity

Symmetrix DMX-3/4 4500

The Symmetrix DMX-3/4 4500 is EMC's largest array with 4-DA pair supporting 480-960 drives in three - five bays. Extending the number of drives beyond the 960 requires a performance assessment be completed by EMC. The assessment determines the current state of the array, and provides EMC with the information necessary to project capacity expansion capabilities.



7.2.2. Entry Model DMX-3 950

The DMX-3 950 model is certified but has numerous limitations and is recommended only when a very small amount of usable storage is required and no expansion is expected for 2-3 years. **GDSE should be engaged prior to any DMX-3 950 deployment to review the requirements and proposed design.**

Considerations for deployment of the DMX-3 950 include:

- The primary benefit of the DMX-3 950 is the ability to accommodate 120 drives in a 6.87sqft single bay footprint.
- Only the 4 director configuration is certified. Data availability risks are too high with the 2 director configuration.
- Single system bay configurations with a maximum of 120 drives are the recommended standard. This is the best use case for the 950 model.
- Dual bay (system and storage bay) configurations will be supported up to a maximum of 240 drives. Initial deployments of dual bay configurations

should strongly consider the other models of DMX-3 which are expandable and may offer better power/cooling requirements.

- Limited number of RDF ports are available on the 950. For SRDF/A, a maximum of 8 RDF groups can be supported by each RDF port pair with 5x71 microcode. This figure will increase with 5772 microcode (see SRDF/A Solutions Guide).

For additional DMX-3 950 information, see section [17 Appendix A – EMC Symmetrix DMX-4 950 Configuration Summary](#).

8. EMC Symmetrix DMX-4 and Enginuity 5772 Microcode

8.1. Introduction

In 2007, EMC introduced the Symmetrix DMX-4 array and Enginuity 5772 microcode to their Direct Matrix line of high-end storage array offerings. Enginuity 5772 is compatible with both the DMX-4 and DMX-3 line of arrays; therefore, sharing much of the management and feature sets. The improvements provided by the DMX-4 and Enginuity 5772 include the following:

- 4Gbps backend fiber loops and point-to-point disk enclosures
- 4Gbps Fibre Channel disk support
- RAID 6 data protection
- Symmetrix Priority Controls

The DMX-4 utilizes the same mid-plane hardware architecture as the DMX-3. The front of the mid-plane contains the channel, disk and global memory directors. The rear of the mid-plane contains the channel adapters and the Communications/Environmental Control Modules. The internal bandwidth, processing capability and memory access are identical between the DMX-4 and DMX-3 arrays. The tables below display the performance and capacity specifications for the DMX-4, which mirrors the DMX-3:

DMX-4 Performance Specifications

Category	Specification	Description
Data Paths	32-128	8 per I/O Director, 16 per global memory Director
Data Bandwidth	32-128GB/s	
Message Bandwidth	4.0-6.4GB/s	
PowerPC Processors	84-130	Four dual 1.3GHz processor complexes per director
Global Memory	16-512GB	Available in 8, 16, 32, 64GB memory Directors
Concurrent Memory Transfers	8-32	4 per global memory Director

DMX-4 Citi Supported Capacity Specifications

Component	DMX1500	DMX2500	DMX3500	DMX4500
Number of fibre channel directors (DA) ⁱ	2	4	6	8
Total fibre channel loops per DA pair	8	16	24	32
Minimum number of disk drives ^j	96	192	360	480
Maximum number of disk drives ^k	240	480	720	960
Minimum raw capacity ^l	35 TB	70 TB	105 TB	140 TB
Minimum usable capacity ^m	17.5 TB	35 TB	52.5 TB	70 TB
Maximum raw capacity ⁿ	72 TB	144 TB	216 TB	288 TB
Maximum usable capacity ^o	63 TB	126 TB	189 TB	252 TB
Min/Max usable global memory (mirrored)	32/64 GB	64/128 GB	96/192 GB	128/256 GB
Recommended connectivity (FA port pairs) ^p	8	16	24	32

EMC supports DMX-4 frame configurations larger than 960 drives or loop populations beyond 30 drives. GDSE artificially limits loop populations to 30 drives to maintain the best balance of performance and capacity. Recommended system and drive cabinet layout for the DMX-4 can be viewed in [section 7.2.1.1](#).

8.1.1. DMX-4 Backend Subsystem

ⁱ DAs are always be deployed in pairs for redundancy

^j Minimum configuration available for sale by EMC

^k Maximum of 30 drives per loops using all loops

^l Using 146GB drives

^m Using RAID 1 and 146GB drives

ⁿ Using 300GB drives

^o Using 300 GB drives and RAID 6 14+2 devices

^p Using DA ports 0 & 1 without RA ports

The primary distinguishing feature of the DMX-4 is the point-to-point Disk Enclosure (DE) that operates a 4Gbps (see Figure 1), as compared to the DMX-3 that utilizes a 2Gbps loop based drive enclosure (see Figure 2). The point-to-point (PtP) architecture enhances the arrays ability to access, diagnose and isolate drives for corrective action. According to EMC, the benefit of the PtP architecture becomes even more evident with highly populated backend end loops. See [Appendix D: DMX-4 Backend Subsystem](#), for additional specifications on the DMX-4 backend subsystem and point-to-point connectivity.

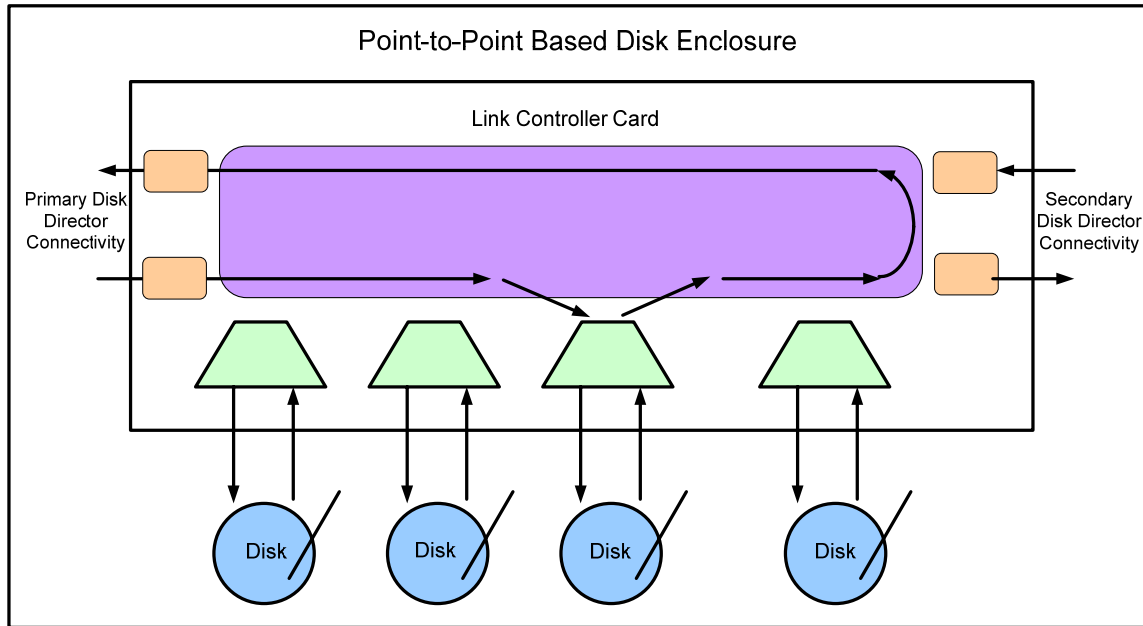


Figure 1

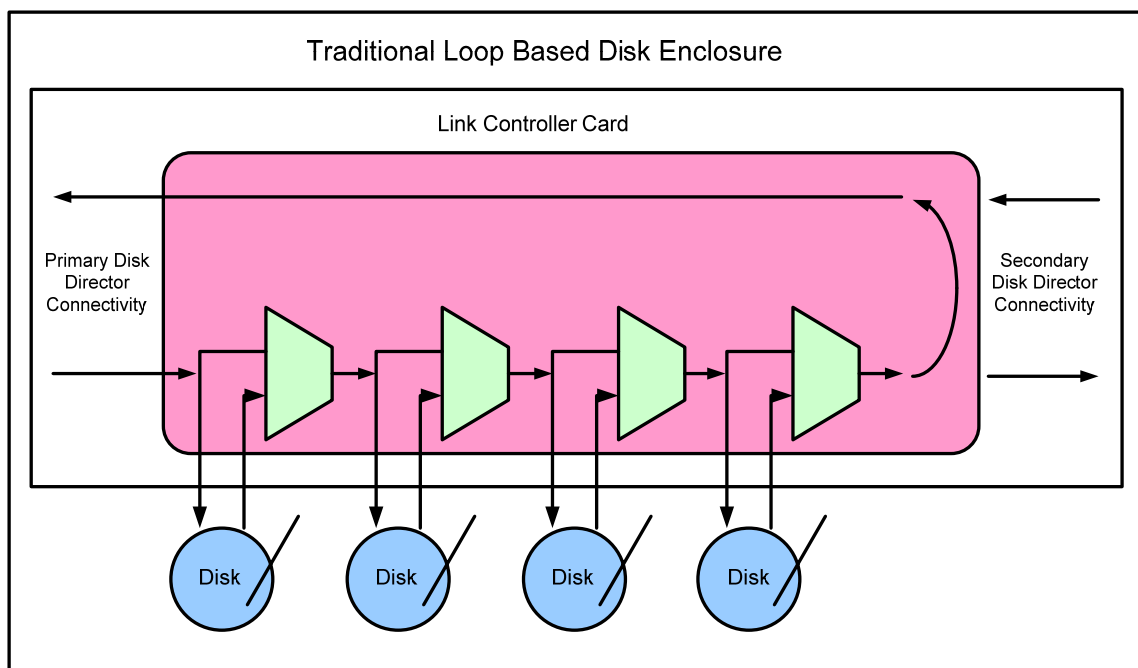


Figure 2

8.2. RAID 6 Introduction

Engenuity 5772 introduces RAID 6 data protection for the Symmetrix DMX-3/4 arrays. Previous data protection schemes such as RAID 1 and RAID 5 protect a system from a single physical drive failure; however, RAID 6 provides data protection if two drives within the RAID group fail. GDSE only supports RAID 6 14+2 (14 data and 2 parity drives), RAID 5 7+1 and mirror data protect schemes. If required, all or any combination of the supported protection schemes may reside within the same array. RAID 6 14+2 and RAID 5 7+1 possess the same data storage efficiency of 87.5%, and utilizes the same management and provisioning command structure.

The primary RAID 6 differentiator is the ability to survive and reconstruct the data in the event of two drive failures within the rank. This ability stems from the XOR Even-Odd algorithm used in calculating multi-dimensional parity. RAID 6 uses both horizontal parity that is calculated from the data across the horizontal rank of the Raid group (identical to RAID 5), and diagonal parity that is calculated on a diagonal vector across a subset of data members. These two independent parity calculations allow RAID 6 to serially reconstruct the data on one disk at a time.

8.2.1. RAID 6 Device Configuration Guidelines

- RAID 6 14+2 configuration provides the best balance of performance and storage efficiency. RAID 6 6+2 is not recommended.

- Support of dynamic disk sparing for array containing RAID 6 devices. Permanent sparing is not recommended due the amount of hardware resources required.
- BCVs can be RAID 6 devices.
- SymmOptimizer Swap supports RAID 6 devices.
- SRDF devices can be RAID 6 devices
- A log device can be a RAID 6 device.
- Concurrent Copy is supported with RAID 6.

8.2.2. RAID 6 Mean Time to Data Loss Probability

RAID 6 provides a level of calculated reliability several orders of magnitude higher than RAID 5 or mirror data structures. There are two methods to calculated data loss: 1) Mean Time to Data Loss due to Disk Failure (MTTDL-DF) determines data loss to physical drive failure, 2) Mean Time to Data Loss due to excessive Bit Error Ratio (MTTDL-BER) that determines the probability of data loss due to excessive bit errors when the Raid group is in a degraded (failed drive(s)) mode. For additional information on MTTDL, see [Appendix E - RAID 6 Mean Time to Data Loss Calculations](#).

RAID 6 vs. RAID 5 MTTDL sample comparison:

A DMX-3 4500 equipped with 300GB 15K.5 drives configured with only RAID 6 14+2 shows a MTTDL-DF of 126,257 years, as compared to 949 years for a RAID 5 7+1 array, amounting to a delta of 133 times. With the same array specifications, MTTDL-BER for RAID 6 is 13,400 years, as compared to 9.99 years for RAID 5 7+1, amounting to a delta of 1341 times.

8.2.3. Throughput Performance

The throughput performance of RAID 6 14+2 as compared to RAID 5 7+1 on 300GB 15K5 disks, for 100% sequential read and write I/O shows a less than 0.5% difference. Sequential I/O usually provides the best performance numbers due to the DMX's prefetch algorithm's high ratio of true positives (cache hit). Random transactions are the most difficult for an array to process due to retrieving the requested data from the physical disk while the request remains in a hold status, thus removing much of the cache I/O enhancement abilities of the array. With 100% Random read and write I/O transactions, RAID 5 7+1 shows 13% throughput advantage over RAID 6 14+2. RAID 5 does hold a slight throughput advantage over RAID 6; however, for most mid-level host applications this performance advantage is practically inconsequential, since rarely is I/O 100% random.

8.2.4. Degradate Mode Operation

RAID 6 provides three levels of online operation: 1) normal, 2) degraded mode (one disk failure) and 3) double-degraded mode (two disk failures). In degraded mode, RAID 6 14+2 on 300GB 15K5 disks, shows 38% lower sequential I/O and 15% lower random I/O versus normal mode. The duration to rebuild a single disk and return the Raid group to normal operation averaged 4.16 hours in lab testing, which corresponds to the 3.51 hours calculated with Mean Time to Repair (recognition/replacement time=zero) formula (see [Appendix E - RAID 6 Mean Time to Data Loss Calculations](#).) Disk rebuild times depend on many factors such as disk type, disk size, bus speed, application load and hot sparing if utilized. A standard rule-of-thumb, the RAID 6 will take 10% longer to rebuild as a RAID 5 device with the same number of data stripes.

8.2.5. Disk Sparing

The DMX-4 supports permanent and dynamic disk sparing, where once a disk has surpassed the soft error threshold the array invokes a physical or logical volume spare. Soft errors are discovered through the error verification process that uses frame-based CRC checks and worst-case data patterns to discover bad blocks. Each bad block increments the soft error counter until the threshold is reached and the disk is designated to be replaced.

RAID 6 only supports permanent disk sparing in single (one drive failure) or double (two drive failure) degraded mode. According to EMC, permanent sparing increases data availability without impacting performance; however, actual testing has shown a 25% (average for sequential and random I/O) loss in throughput that when in degraded mode. Therefore, expect some loss in throughput performance during the sparing operation.

- Permanent sparing averts data loss by replacing a failed physical drive with a spare from the spare drive pool. A permanent spare requires the sparing drive to have same block size, capacity, and speed in a suitable location in the backend subsystem. Therefore, to effectively provide permanent sparing, each backend loop needs to contain at least one spare disk of every type available on the loop. The array will search for the most suitable spare, which includes selecting up to two members on the same disk director, loop, or drive enclosure, which is not a recommended practice when creating the original RAID 6 device.
- Dynamic sparing protects against data loss by relocating logical volumes to available spare drives in the spare drive pool. The dynamic sparing process copies the contents of a failing disk to an available spare disk and/or allows reconstructed data to be placed on the dynamic spare disk. When the original failed drive is replaced, the newly replaced drive is resynchronized with the dynamically spared volumes and the spare drive is returned to the spare pool.

See [Section 8.4 Symmetrix DMX-4 and Enginuity 5772 Configuration Guidelines](#) for supported disk sparing implementation.

Note: Vault drives can be dynamically or permanently spared; however, permanent sparing is only available if a spare of the same speed, block size and capacity is on the same loop as the failing drive. Vault devices can be relocated to different positions on the same loop, but cannot move to a different loop.

8.3. *Symmetrix Priority Controls*

Symmetrix Priority Controls (SPC) is a Quality-of-Service (QoS) parameter that allows the performance metric per device or group of devices to be managed. SPC is supported in Enginuity 5772 on the DMX-3 and DMX-4 arrays. Using the Solutions Enabler symqos command set, SPC provides 16 levels (1 highest – 16 lowest priority) of device prioritization through setting the Priority Service Time (PST). PST prioritizes a device or device group by increasing or decreasing the potential I/O delay of the request.

PST levels are relative to each other; therefore throughput differences between PST levels are dependant upon the load on the array. Lab testing with a low level of background I/O yields a 12% lower sequential read and write with a PST level 16 as compared to level 1. Random read and write shows an 11.5% lower throughput for PST level 16 as compared to level 1. Transactions sizes of 4KB-16KB shows the largest PST difference, with a 22% lower throughput for sequential reads and writes and 21.7% for random reads and writes for PST level 16 as compared to level 1.

EMC defaults the PST levels for all applicable devices to the highest level of 1 for the entire frame; therefore, utilize PST to set the performance metric between 2 and 16. Being that PST levels are relative to each other and dependant upon the load on the array, recommend that PST levels be lowered in small 1 through 3 level increments until an acceptable level of performance is achieved.

PST levels are **not** supported on SFS, VAULT, SAVE, DRV and VDEV.

8.4. *Symmetrix DMX-4 and Enginuity 5772 Configuration Guidelines:*

- The primary drive types supported on the DMX-4 array are the T146155, (146GB 15K) and the T300155 (300GB 15K) 4Gbps Fibre Channel drives. The T300155 drive presents a lower cost per byte as compared to the T146155 drive, and provides sufficient performance for the majority of mid-level server applications.

- RAID 6 14+2 is supported on Enginuity 5772 for the DMX-3/4. RAID 6 14+2 may reside with the current certified data protection schemes of mirror and RAID 5 7+1 in the same array. RAID 6 6+2 is not supported, due to its limited throughput performance and low data to parity capacity ratio. RAID 6 14+2 devices are to be created at 36828 cylinders (34526MB) in size, consistent with Citi's RAID 5 7+1 device profile (see [section 10.6.3.3 for RAID 5 Configuration Settings](#)).

Future Reference:

Larger Fibre Channel (>300GB) and SATA II (>=1TB) disks will be available in the near term, which will require larger 65472 cylinder devices for RAID 6 14+2. The larger devices will reduce the number of individual splits per physical drive, thereby reducing the resources required to manage the backend subsystem. A 65472 cylinder RAID 6 14+2 device has a split size of 4390MB versus a 36828 cylinder device which has a split size of 2470MB. In a 1TB SATA II drive, a 65472 cylinder device creates approximately 220 splits while a 36828 cylinder device creates 400 splits per drive. According to the DMX Product Guides, a maximum of 160 drive splits is recommended for frames supporting RDF and 255 splits for non-RDF frames.

- The VCMDB should use device identifier (devID) of 008, consistent across all DMX-4 models. The VCMDB device is used as an initial gatekeeper device and communication medium for PowerPath. See [section 10.9](#) for additional details on the VCMDB device.

Solutions Enabler command "symdev list -vcm" will display the VCMDB device ID.

- Vault devices are required for maintaining a copy of global memory during planned or unplanned power downs. Each Disk Director pair requires 32 5GB vault devices, therefore a 2500 model will utilize 64 and a 4500 model will use 128 vault devices. EMC reserves device identifiers of 000-007 for vault devices and allows the remaining vault devices to use any available identifier. The bin file should be configured to use vault devID's of 000-007 with the remaining vaults devices using identifiers following the last data device. In the event the DMX-4 array is upgraded with additional Disk Director pair(s), the new vault identifiers should be set contiguously with the previous vault identifiers. For example, upgrading the array from a 2500 to 4500 model requires an additional 64 vault devices. If the last vault devID is FFF, then the new vault devices should use device identifiers of 1000-103F.

Solutions Enabler command "symdev list -all" will list the vault devices.

- Symmetrix disk groups should be utilized to classify physical drives types with specific data protection types. Recommend that Symmetrix disk groups be created in increments of 8 drives when housing mirror and RAID 5 7+1 devices. Disk groups in increments of 16 drives are recommended for RAID 6 14+2 devices. A separate disk group should be created for hot spares. If possible, assign the first four drives in the loop to disk groups with mirror devices, since they contain the vault devices.
- Fibre Adapter (FA) ports 0 and 1 can be utilized simultaneously to provide the widest host distribution through the front-end ports. Each FA contains two 4Gbps Fibre Channel ports, which interconnects with a single Channel Director (CD) processor complex (each CD contains 4 processor complexes). Utilize FA port 0 pairs prior to utilizing FA port 1 pairs. Ports with volume set-addressing (HP-UX servers) should be configured as needed and may use FA port 0 or 1.

Several operating system versions cannot utilize LUN numbers greater than 255 as their target identifier; therefore, distributing host access over ports 0 and 1 increases the total number of host accessible devices on the array.

- The DMX-4 Fiber Adapter (FA) ports should be set with auto-negotiation enabled. By default the bin file will set the FA ports to 1Gbps with auto-negotiate disabled; therefore, GDSE has requested EMC to configure each DMX-4 bin file with auto-negotiate enabled. Once connectivity is establish with the fabric switch verify the port speed for each FA. See [Section 10.7](#) for details on FA port configuration.
- To optimally support dynamic sparing an allotment of 2 spares for every 100 drives per disk type is recommended. A minimum of 8 drives total per array with drive type increments equal to the number of processor pairs is also recommended. For example, a DMX4 4500 equipped with 192qty-146GB 15K, 320qty-300GB 15K and 128qty-1TB 7.2K drives presents an optimal count of 4qty-146GB, 8qty-300GB and 4qty-1TB spare drives.
- Dynamic sparing in Enginuity 5772 supports using two smaller drives to spare one larger drive, where the aggregate capacity of the two smaller spare drives is equal to or greater that the larger drive. Unfortunately for the DMX-4 standard Fibre Channel drives, two T146155 drives with a total capacity of 280028MB, cannot spare one T300155 drive (286103MB actual capacity), since the aggregate capacity is 6074MB less.
- Allocating sufficient disk resources to fully support permanent disk sparing are not recommended, except in cases where access to the array is limited and extreme high availability is required. Full support for

permanent disk sparing requires spare disk(s) to be available on each backend loop, amounting to a high volume of unused capacity within the frame. For example, a DMX-4 4500 with only T300155 disks requires 32 spare drives totaling 9TB of capacity to fully support permanent sparing.

- RAID 6 only supports permanent drive sparing protection. According to EMC engineering, dynamic sparing is never supported on RAID 6 volumes even in double-degraded mode (The DMX-4 Product Guide incorrectly states dynamic sparing will invoke in double-degraded mode). Since allocating resources to fully support permanent sparing is not a recommended practice, the likelihood of sparing RAID 6 members is relatively low. Therefore, the immediate replacement of the failed drive is imperative.
- Permanent sparing may invoke the following configuration anomalies:
 - Locate two RAID 6 device members on the same disk director, loop, or drive enclosure. Having RAID group members share the same backend elements is not a recommended practice.
 - Vault device can only be spared to the same loop. However, a vault device maybe spared to any drive within the loop, which is contrary to the array's policy that all vault devices be located on the first four disks in the loop.
 - Both RAID 5 and 6 members support permanent sparing. However, if the array is unable to locate a suitable permanent spare, RAID 6 members will not be spared and the RAID 5 members will be dynamically spared.
- The DMX-4 backend subsystem supports the industry-standard 4Gbps Fibre Channel (FC) drives. Placing 2Gbps FC drives into the DMX-4 is not supported. A 4Gbps FC drive may be installed on the DMX-3 2Gbps disk subsystem; however, the 4Gbps drive will assume the profile of a 2Gbps FC drive.
- Symmetrix Optimizer (symoptmz) places a lock 15 on the frame when in operation, whereby prohibiting configuration changes on the array until the operation is complete. Since symoptmz operations may take many hours to complete, it is recommended that symoptmz only be initiated at appropriate times when no other configuration changes are eminent.
- Priority Service Time (PST) prioritizes a device or device group by increasing or decreasing the potential I/O delay of the request. PST utilizes a performance metric of 1 through 16, where level 1 is the highest level of throughput and the default value for each device with in the array. Each PST level are relative to each other; therefore throughput differences between PST levels are dependant upon the load on the array.

Recommend that PST only be used to lower a level of performance on a device or device group. PST levels should be lowered in small increments (1-3 levels) until an acceptable level of performance is achieved.

- Dynamic Cache Partitioning is not recommended due to the possibility of misappropriating cache resources. Dynamic Cache Partitioning provides logical partitioning of cache into eight groups (seven new ones with one default). In read operations, slots will be allocated subject to the cache partition size and current slot count. In write operations, each partition will have its own write pending (WP) limits and counters.

9. Physical Installation

Physical and environmental setup of all enterprise storage arrays should conform to vendor specifications and facilities provided by the local data center planning group.

Rack mount	Not Applicable
Power	DMX Environmental Requirements – DC Hardware Planning
Cooling	DMX Environmental Requirements – DC Hardware Planning
Cabling	DMX Environmental Requirements – DC Hardware Planning
SAN	See McData Best Practices.

10. Symmetrix DMX Standard Hardware Configuration

10.1. Symmetrix DMX-2

- Engenuity Code Revision: Consult the GDSE website for the latest certified version <http://engineering.ti.citigroup.net/gdse/>
- RAID Protection: RAID1
- To ensure maximum performance with the DMX-2 family of storage arrays it is recommended to fully populate the array with disk drives of a single capacity and speed during the initial installation. This will allow the logical devices to be spread across the maximum number of drive spindles resulting in the optimal layout. Capacity requirements should be determined upfront and the corresponding DMX-2 model purchased to fit the storage requirements.

10.2. Symmetrix DMX-3

- Engenuity Code Revision: Consult the GDSE website for the latest certified version <http://engineering.ti.citigroup.net/gdse/>
- RAID Protection: RAID1 or RAID-5 7+1. A single RAID type should be used per array whenever possible.
- GDSE has certified up to 30 drives per loop to ensure a good balance between performance and scalability. However, applications that require maximum performance should limit the number of drives to 15 per loop. Configurations with greater than 30 drives per loop are also supported with the pre-requisite of a performance study done by EMC on existing configuration
- The DMX-3 array is scalable array and when fully expanded it can occupy more than 9 contiguous floor tile space. (See [section 7.2.1.1](#) and the *DMX-3 Product Guide* for further details)

10.3. Symmetrix DMX-4

- Engenuity Code Revision: Consult the GDSE website for the latest certified version <http://engineering.citigroup.net/gdse/>
- RAID Protection: mirror, RAID 5 7+1 and RAID 6 14+2. A mixture or any combination of all three RAID types, are supported within the same array.
- GDSE has certified up to 30 drives per loop to ensure a good balance between performance and scalability.
- The DMX-4 array is scalable to occupy more than 9 contiguous floor tile space when fully expanded. (See [section 7.2.1.1](#) and the *DMX-3 Product Guide* for further details).
- See [section 8](#) for details on the DMX-4 and Engenuity 5772.

10.4. Disk Drive Strategy

The certified range of disk drive for the DMX-2 platform compared to the one for DMX-3 is very different. The certification for the DMX-2 was done using mainly 2 disk drive types (73GB and 146GB 10K rpm drives).

The DMX-3 certification on the other hand was done taking into account all the supported fibre channel drive configurations from EMC, ranging from 73GB and 146GB (10K and 15K) to 300GB 10K and including the 500GB 7.2K disk drives.

Throughout the certification both 73GB and 146GB disks were found to perform equally. The sustainable I/O's per second (IOPS) for each disk was equal. Both the 300 GB 10K disk drives and the 500 GB 7.2K drives performed at 25% and 30% degradation respectively when compared to the 73GB and 146GB 10K drives. (See [Appendix A](#) for Seagate drive specs)

The doubling and tripling storage capacity of the 146GB 10K and the 300GB 10K drive respectively while still maintaining the same IOPS as the 73GB 10K must be considered during performance planning.

When considering the consolidation of storage workloads being serviced by two 73GB drives onto a single 146GB drive or 300GB drive, the combined IOPS load of the full 146GB or 300GB of storage must be weighed. If the 73GB drives are performing more than 50% of their sustainable IOPS load, then consolidation onto a 146GB or 300GB drive would not be possible.

Applications requiring high I/O throughput, should strongly consider the use of 73GB or 146GB 15K drives. In other environments where higher storage capacity with lower IOPS per GB of storage is required, the use of 146GB 10K and even the 300GB drives can be considered.

The 500 GB 7.2K drives offer the lowest performance out of all the drives tested. Any application, looking to use these drives should strongly evaluate their requirements and characteristics before production deployment. GDSE recommends using 500 GB drives for applications such as disk-based backup or other tier 3 applications only.

10.4.1. Selecting Drive Types

It is recommended that a single drive type and capacity be selected for an entire array when possible. This maximizes performance and eliminates the manual task of matching drive performance characteristics to application performance requirements. (Note: Mixed drive configurations are also fully supported given that the guidelines in [section 9.5.3](#) are followed)

- i.e. Alleviates possible contention concerns where an application/host could be assigned a LUN from a 73GB drive and another from a 146GB drive or even 300GB drive.

The exception to this rule is when 500 GB drives are used in which case the first 4 drives on each loop must be 10/15K drives.

In SRDF environments the target array must be configured with the same type drives as the source array.

NOTE: One important factor in realizing cost and capacity benefits is using the 146GB or 300GB drives in applications that are best suited for higher capacity environments.

Sequential access applications have a tendency to maximize any caching advantages of the storage system, and take advantage of the *increased internal transfer rate* of the drive. In these applications, the speed at which the drive can transfer data from the platter has the greatest effect on overall performance. Thus, medium-to-large block and sequential I/O applications reap the greatest benefits from both the 146GB and 300GB drive's improved internal transfer rates and higher bit densities.

In summary:

- o To date, the 146GB 15K drives have proven to meet the majority of business requirements and should be the default configuration unless otherwise determined by Storage Management. These drives offer a suitable mix of performance and capacity required for the majority of applications.
- o Applications with extreme high-performance requirements should consider the 15K version of the 73GB drives (C73DX15). For the DMX-3, a minimum purchase for C73DX15 should be 96 for 2500 model and 192 drives for 4500 model for adequate distribution across the backend construct. Other considerations like, 15 drives per loop, DA and proper cache sizing should also be considered.
- o 300 GB drives can also be used when I/O performance is not a primary requirement.
- o 500 GB drives should only be used only in configurations such as disk backups or temporary dump space off loading to tape drives.

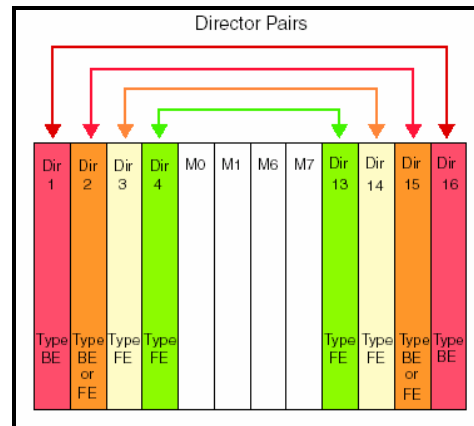
10.5. Fibre Adapter Configuration

Symmetrix DMX arrays are configured with multiple Fibre Channel Adapters (FAs). Each FA has 8 Fibre Channel ports managed by 4 processors (2 ports per processor). Each processor has a direct path to global cache. It is

recommended to only use port 0 of each processor whenever possible. Configurations involving both ports are also supported but should only be used when additional port counts are necessary.

Previous versions of the EMC Symmetrix arrays (Sym 5) were built upon a bus architecture instead of the direct matrix architecture used in the DMX family. To ensure that no disk or redundantly connected host was dependant upon any one bus, EMC developed the Rule of 17. The Rule of 17 stated that all director pairs would be used in a combination that the resulted in a sum of 17. For example if a fibre channel adapter in slot 3 was used it's partner adapter would be the fibre channel director in slot 14.

In the DMX family of arrays the Rule of 17 is no longer a technical requirement but should continue to be followed to retain consistency.



Dir	Port		Dir	Port
FA-3a	0	↔	FA-14b	0
FA-3b	0	↔	FA-14a	0
FA-3c	0	↔	FA-14d	0
FA-3d	0	↔	FA-14c	0
FA-4a	0	↔	FA-13b	0
FA-4b	0	↔	FA-13a	0
FA-4c	0	↔	FA-13d	0
FA-4d	0	↔	FA-13c	0

Table 3: Exampe of "Rule of 17" Director Pairing

10.6. Bin File Configuration

The following are standard configuration requirements for Symmetrix DMX BIN File configurations. The majority of the DMX-2 standards do apply to the DMX-3 except where noted below.

10.6.1. DMX-2 specific standards

	DMX-2 73GB	DMX-2 146GB
Engenuity Code	Current GDSE Certified version	Current GDSE Certified version
Data Protection (Parity)	RAID1	RAID1
Standard Hyper-Volume	18414 cylinders	18414 cylinders
SRDF ^q Replication	Fibre Channel	Fibre Channel
Spare Disks (Single Drive Size)	1 per every 144 drives	1 per every 144 drives

Table 4: DMX-2 standards

All hyper volumes on the array should be of 18414 cylinders (approximately 8GB) in size. Hosts should be provisioned striped meta volumes that are striped across at least 2 hyper volumes and at most 32 hyper volumes. This translates to a minimum LUN size of approximately 16GB and a maximum LUN size of approximately 256GB.

Sites should configure a portion of each storage array into meta volumes at the time the array is being installed through a BIN file request to EMC. GDSE recommends that at least 2/3rd of the total storage on the array should be pre-configured in this fashion. This storage can then be quickly provisioned to hosts. This also ensures well-balanced back end performance and eliminates the need for frequent high-risk BIN file changes^r.

If prior allocation and forecasting do not provide enough information to make a good estimate of meta volume sizes needed, then 32GB (approximately) sized meta volumes are recommended for the 2/3rd portion of the storage array.

The remaining 1/3rd storage should be left as hypervolumes that can be combined into striped metavolumes to fulfill requests that were not anticipated at

^q The use of ESCON for SRDF infrastructure is valid in Symm 5 environments only; DMX installations should employ SRDF over Fibre Channel.

^r GDSE's vision for storage provisioning is to match the requested LUN size as closely as possible, while maintaining the highest backend performance. This can only happen if custom sized meta volumes are created against every provisioning request. Unfortunately, current versions of tools such as Solutions Enabler are very complex to use as day-to-day management tools for these functions.

the time of installation. The minimum and maximum LUN size rules still apply; individual hyper volumes should NOT be provisioned to hosts because of performance concerns.

10.6.2. DMX-3 specific standards

	73GB	146GB	300GB	500GB
Enginuity Code	Current GDSE Certified version	Current GDSE Certified version	Current GDSE Certified version	Current GDSE Certified version
Data Protection (Parity)	RAID1/RAID5	RAID1/RAID5	RAID1/RAID5	RAID1
Standard Hyper-Volume	9207 cylinders for Raid-1 and/or 36828 cylinders for Raid-5	9207 cylinders for Raid-1 and/or 36828 cylinders for Raid-5	9207 cylinders for Raid-1 and/or 36828 cylinders for Raid-5	9207 cylinders for Raid-1 and/or 36828 cylinders for Raid-5
SRDF ^s Replication	Fibre Channel	Fibre Channel	Fibre Channel	Fibre Channel
Spare Disks (Single Drive Size) ^t	2+1/100	2+1/100	2+1/100	2+1/100

Table 5: DMX-3 Standards

The DMX-3 utilizes the same standard block size as the DMX-2 of 512 Bytes. However, the DMX-3 uses a track size of 65.536KB (2^{16}), as compared to the DMX-2, which utilizes a 32.768KB (2^{15}) track size. Therefore, a standard Raid-1 hypervolume extension (HVE) of 18414 cylinders (8630MB) in the DMX-2 is equivalent in size to a 9207 cylinder HVE in the DMX-3. It should also be noted that EMC only supports a stripe size of 960KB when creating metavolumes. (2-cyl for DMX-2 and 1-cyl for DMX-3)

For Raid-5 HVE configuration see section [9.5.3.3 RAID-5 Specific Guidelines](#)

The DMX-3 also introduces a new concept of vaulting for the DMX platform. Vaulting was mainly designed to limit the time needed to power-off the array while insuring the integrity of cache data by de-staging it into the vault drive.

The stripe meta recommendation, as well the $1/3^{\text{rd}}$, $2/3^{\text{rd}}$ rule bin file layout listed above in the DMX-2 section can also apply to the DMX-3. As stated above this rule should be considered when the requirements are not specific enough to yield

^s DMX installations should employ SRDF over Fibre Channel.

^t Each DMX-3 requires a minimum of two global spare drives and one additional global spare drive for every 100 disk drives or a portion thereof. For example, a DMX-3 with 547 disk drives requires $2 + 6 = 8$ global spare drives.

a complete uniform bin file. Furthermore in instances where RAID tiers will be aggregated in a frame, a more accurate forecast maybe needed to limit excess in any particular RAID type.

10.6.3. DMX-3 Disk-Drive Configuration Rules and Guidelines

10.6.3.1. Basic Configuration guidelines

- A minimum of four disk drives must be configured on each drive loop within each Storage Bay.
- Maximum of 30 drives per loop recommended at initial deployment. Configurations needing to surpass that limit should consult EMC to perform a storage frame performance evaluation before increasing the drive count per loop.
- The first four drives on each drive loop in the direct-attached Storage Bay must be either 10,000 or 15,000 rpm drives to support the vault devices. (500 GB not supported)
- Reserved capacities:
 - o 5 GB on the first four drives of every drive loop (160 GB per disk director pair) is reserved for memory vaulting.
 - o 24GB of total disk capacity is reserved for internal Symmetrix file system (SFS) use.
 - o The VCM Database now resides in the SFS volume and is accessible via a 3 cylinder device (Gatekeeper like). For ease of management, the VCMDb device should be assigned Symmetrix device number 008 on the DMX-3, because the first 8 devices, number (0-7) must be occupied by the first set of vault devices. Any remaining vault devices should be assigned at the end of the bin file.
- Mixed-drive protection rules:
 - o RAID 5 3+1 and 7+1 cannot co-exist in the same system.
 - o All other protection types can co-exist in the same system but may require additional management efforts.
 - o A (7+1) RAID 5 configuration requires a minimum of 8 different back-end loops to avoid possible failure cause by a single loop.
- Spare-drive rules:
 - o Dynamic Spare drives will be mandatory for Symmetrix DMX-3 configurations. EMC allows a maximum of 4 mirror positions per device, and sparing occupies a mirror position when invoked. Sparing will not be invoked any Symmetrix array device has all 4 (M1, M2, M3, M4) mirror positions in use. For example, a sparing operation will fail in the following RDF configuration:
 - Source Symmetrix with RDF1 device mirrored (M1+M2), RDF link established M3 and a local BCV (M4) device also established.

- o One additional Global Spare drive will be added to the system for every 100 drives, or portion thereof ($2 + 1/100$). (Example: A Symmetrix DMX-3 system with 547 drives will have: $2 + 6 = 8$ Global Spare drives.)
- o The capacity of all Global Spare drives will be at least the capacity of the largest drive in the system.
- o Sparing with the speed of the fastest largest drive in the system (if available) will be allowed (optional).
- The DMX-3 can be classified into 2 types of tiered storage configuration with different characteristics and considerations. Both configurations listed below are certified and supported by GDSE.
 1. Dedicated tier of storage (different RAID types or different drive types in dedicated storage frames)
 - Provides best overall and predictable performance.
 - Ease of management
 2. “In-the-box tiered storage (different RAID types or different drive types in the same storage frame)
 - Symmetrix disk groups should consist of drives with same rotational speed and size.
 - Drive loops consisting of slower drives can be configured to surpass the recommended maximum of 30 drives. However, performance may be significantly affected. EMC should be consulted to do a performance analysis to determine feasibility of the expansion.
 - 500GB 7200 rpm drives should be evenly distributed across all the backend DA loops for optimal performance. Drive types can easily be represented in the bin file for inspection and allocation.
 - Dedicating physical drives to a RAID type is best in order to predict performance.

10.6.3.2. General RAID protection best practices

- Data Availability requirements should drive the solution whenever available.
- RAID 1 protection is the highest level of protection available and provides the best performance overall in most applications; mission-critical applications (high availability and low tolerance for performance impact due to drive failure) should use RAID 1 to minimize the potential of data loss.
- RAID 5 recommendations are described in the next section.
- Recommended Drive Spare rules
 - o Each Symmetrix DMX-3 configuration ordered will have at least a base global sparing of two Global Spares.
 - o 500 GB drives should not be used as global spares. They should only be used as spares when the data drives are 500 GB drives.

10.6.3.3. RAID-5 Specific Guidelines

New to Citigroup is the support of RAID-5 with EMC DMX arrays. RAID-5 is supported only on the DMX-3 arrays only and not on previous DMX models. RAID-5 devices are utilized and provisioned as standard, BCV and SRDF^u devices similar to 2-way mirrored devices. The DMX-3 supports two types of RAID-5 variants:

- RAID-5 3+1 - configuration consists of four Symmetrix devices with effective data storage of 75% for the Raid device.
- RAID-5 7+1 - configuration consists of eight Symmetrix devices with effective data storage of 87.5% for the Raid device.

Based on cost, utilization and performance analysis, the EMC RAID-5 7+1 configuration is best for Citigroup overall usage.

- RAID-5 7+1 configuration is EMC's highest effective storage ratio, providing 43% greater effective storage utilization as compared to mirrored protection and 14% greater than RAID-5 3+1 protection.
- RAID-5 7+1 configurations can only tolerate a 1:8 drive failure ratio within the RAID group, as opposed to a 1:2 ratio for mirror and 1:4 ratio for RAID-5 3+1.

The DMX-3 Raid-5 hypervolume is 34.53GB in size as compared to the Raid-1 hypervolume of 8.63GB. The RAID-5 hypervolume encompasses 36,828 cylinders created from 552,420-65,536B tracks across the entire RAID group. RAID-5 Metavolumes can be created by striping multiple RAID-5 hypervolumes together, similar in principle to creating RAID-1+0 metavolumes.

The larger Raid-5 hypervolume reduces the number of disk volumes on each physical drive. For optimum performance, the number of disk volumes on a single physical drive should be limited to no more than 56 disk volumes, per EMC Symmetrix Performance Engineering. Extending beyond these disk volume limits increases the probability of performance issues with prefetch tasks, internal copies (Clone/TimeFinder/SRDF), and disk rebuilds. Expanding the hypervolume size for Raid-5 from 8.63GB to 34.53GB reduces the number of disk volumes on a physical drive by a factor of 4. In a 146GB drive, the 34.53GB hypervolume is reduces the number of disk volumes from 112-1.235GB to 28-4.95GB disk volumes. In a 300GB drive, the 34.53GB hypervolume reduces the number of disk volumes from 224-1.235GB to 56-4.95GB.

RAID-5 guideline summary

- RAID 5 should not be used for mission critical applications and should be on dedicated physical spindles other than RAID 1.

^u DMX-3 phase II certification includes SRDF Raid-5 testing.

- o RAID 5 configurations only have one copy of the data (although the RAID-5 implementation occupies 2 mirror position)
- o RAID 5 (7+1) is seven times more susceptible to a second drive failure than RAID 1.
- o RAID 5 provides best performance in sequential type applications and does poorly in random write-intensive workloads.
- o Always use smaller drive types for RAID 5 whenever possible to avoid longer rebuilt time and increase data availability.
- o 500 GB drives are not supported by GDSE in RAID 5 configuration.
- o RAID-5 holds the advantage of lower cost, and represents a sweet spot in second tier business critical data storage applications. Please see the Appendix [DMX release notes](#) for performance benchmark results

10.7. FA Port Flag Settings

The following bits need to be set on all FA ports of any Symmetrix to provide full interoperability between operating systems: (see Table 10 below)

EMC Bits	HP-UX	Shared OS ^v	Notes
T ^w		X	Tagged Commands
D		X	Disable Queue Reset
C	X	X	Common Serial Number
V ^x	X		Enable Volume Set Addressing
SC3		X	SCSI 3 Interface
PP	X	X	Enable Point to Point Mode
UWN	X	X	Unique World Wide Name
VCM	X	X	Enable Fibre Path on this port

Table 6 - Symmetrix BIN File Default BIT Settings

- Sun Cluster 2.2 requires an additional bit “SCL” and should only be set on FA ports that will be used for Sun Cluster 2.2. Sun Cluster 2.2 does not support FA port sharing.
- Sun Cluster 3.0 does not require the “SCL” bit. The PER volume flag is required to allow sharing FA’s between Sun Cluster 3.0 with other operating systems. Currently Sun Cluster 3.0 cannot be used with a BCV LUN located on the R2 target array. In this configuration even though the

^v Sun Solaris, IBM AIX, Microsoft Windows 2000 and Windows 2003, VMware, RedHat Linux

^w T bit not required on microcode 5568 or later.

^x Setting the V bit will require a bin file change to the LUN addressing for the FA port. Once the V bit is on, the addressing of the LUN's on the FA only range from 0-7 for the LUN value and the TID and VBUS values get incremented.

R1 or source system was a Sun Cluster, the DR/COB system would have to be a single server without Sun Cluster.

NOTE: Although PER flag setting is incompatible with Veritas DMP, it does not impact Citigroup environment because Citigroup uses EMC PowerPath with DMP.

- Veritas Cluster Server v3.5^y is supported. However, certain VCS cluster configurations (i.e. Oracle RAC) do require the PER flag to be set on the coordinator disks. The common configuration for a VCS cluster is to allocate 3 quorum devices to a pair of cluster server with the PER flag enabled on all 3 quorum devices.
- Although, AIX, Linux and HP-UX were not part of the GDSE certification testing, they are inherent in Citigroup's technology infrastructure and are considered certified for DMX use based on testing and implementations from storage groups in other regions.
- Setting the "V" bit on an FA port for HP-UX servers will make the port dedicated for HP servers only since other Operating Systems use Peripheral Set Addressing. GDSE recommends using the FA "B" ports which are reserved for expansion, to hold HP-UX servers also.
- Heterogeneous masking capability should be used to share an FA port between hosts requiring different bit settings. Please see the GDSE solution guide for Solutions Enabler for more information. Please see the references section for the location of this document.

Please reference the EMC Support Matrix for latest support configurations - https://powerlink.emc.com/LowFreq/P_esm20031003.pdf

10.8. LUN Naming/Numbering

GDSE Best Practices for LUN assignments are as follows:

- LUN 0 (VCMDB) should be assigned to all FA ports. (Symmetrix device number greater than 007 on the DMX-3 but mapped to hosts/FA ports as LUN 0)
- LUNs are assigned to FA ports as volumes are provisioned, in ascending order.
- Each FA segment is assigned a corresponding host SCSI target.

^y Please reference the Global Distributed Unix Engineering website for configuration details. <http://ue.ny.ssmc.com/>

- o e.g., hosts assigned to FA Segment 1, composed of ports FA3a, port 0 and FA14B port 0, would use SCSI Target 1 in their sd.conf (Solaris) configuration file.
- Quorum devices, used with clustering software applications, are created as approximately 500MB devices, usually made from remaining space on each physical disk after all the hyper volumes are assigned. These devices start with LUN assignment number 5A Hex → 90 Decimal. It is recommended to assign at least 6 quorum devices behind each FA segment. It is not necessary to setup RDF relationships for those devices in a R1 -> R2 Symmetrix configuration.
- Gatekeepers devices sized 6 cylinders (approximately 2.88MB) are also created from remaining space on each physical disk after all the hyper volumes are assigned. These devices start with LUN assignment number F0 Hex → 240 Decimal. It is recommended to assign at least 6 gatekeeper devices behind each FA segment.

10.9. VCMDB Management for DMX-3

The Volume Configuration Management Database (VCMDB) contains the masking and access records for a specific array. Prior to Enginuity 5771, the VCMDB resided on a specific VCM device, and by default accessible to any HBA logged in to the director port where the database is configured. The VCM device is represented as volume ID 000 at 23MB in size and set as **write disabled**.

In Enginuity 5771 and later, the VCMDB is used as the initial gatekeeper device and communication medium for PowerPath. The masking and access record database has been moved from the VCM device to the internal Symmetrix Filesystem. The VCM device is represented by volume ID 008 (vault devices are 000-007) at 2.88MB in size (3 DMX-3 cylinders) and is **write enabled**.

As with all Enginuity releases the VCMDB should **never** be use in creating a host volume. Therefore, for Enginuity 5771 and later, GDSE recommends that VCMDB device be write disabled at the earliest possible opportunity. The VCMDB is required to support PowerPath masking and discovery operations; therefore, it should continue to be mapped to each FA port in a write disabled state. The VCMDB write disabling procedure is located in [APPENDIX C – VCMDB WRITE DISABLE PROCEDURE](#).

- Enginuity 5771 and 5772 VCMDB Specifications:
 1. VCMDB device are read/write enabled by default.
 2. VCMDB is mapped to all FA ports on the DMX-3 array by default.

3. The DMX-3 utilizes a type 6 VCMDB that resides in the Symmetrix File System, and supports 64,000 devices and 256 fibre/512 iSCSI initiator records.
4. Devices 000-007 are vault devices and are not depicted on FA ports.
5. The VCMDB should be mapped to all ports on the DMX-3 to support PowerPath masking and discovery operations.

VCMDB_restrict_access metric has been obsolete since Enginuity 5770.

11. Hardware Replication

EMC offers two types of replication on the Symmetrix DMX architecture: SRDF (Symmetrix Remote Data Facility) which enables remote replication and TimeFinder which handles replication within a frame. Both replication methods have implications on the total number of mirror positions allowed on a Symmetrix. The maximum number of mirror positions available on the EMC Symmetrix platform has always been 4 (M1, M2, M3 and M4).

11.1. In-Frame Replication (BCV/Clone)

In the past any reference to the EMC TimeFinder was immediately associated to the business continuance solution that allows you to use special Symmetrix devices called business continuance volumes (BCV) which can provide a point-in-time copy of data from standard volumes. Each BCV device has its own host address, and is configured as a stand-alone Symmetrix device. With recent enhancements to the Symmetrix enginuity code and Solutions Enabler software, EMC Timefinder has been expanded beyond the simple BCV model and now provides both snapshot and clone technology.

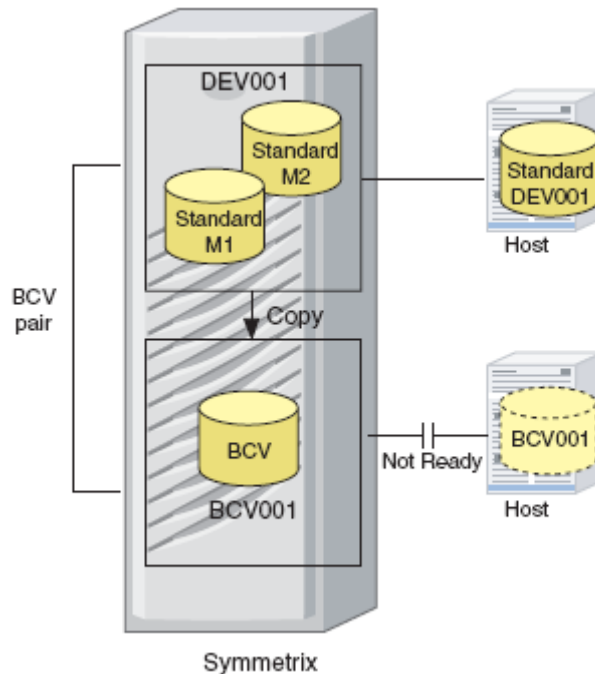
The current TimeFinder product set can be broken down into three main different operations.

TimeFinder/Mirror (symmir) – Is the only local replication product that creates true mirror images. It also provides a full-volume mirror of the source device. The mirror BCV requires a mirror position of the standard device and can be either a protected or unprotected but not supported in RAID 5 protection. Since TimeFinder/Mirror (BCVs) are full physical copies and appear as a mirror of the standard device, they can also be used to increase availability if the source volumes are destroyed.

TimeFinder/Clone – Provides full-volume point-in-time copies of the source data. BCVs formed from RAID-5 devices always use clone technology and hence do not require mirror positions on the standard device. Clone operations also support TimeFinder/Mirror commands via emulation mode. This will allow existing BCV (shell/perl) scripts to perform clone operations. Default clone operations do not copy the entire data to the clone volume. It is copied on

access. Although not the default setting, it is a recommended best practice to always use the `-precopy` option to enable a full copy of the standard device upon activation. (Latest version of *Solutions Enabler Symmetrix TimeFinder Family CLI Product Guide* from <http://powerlink.emc.com>)

TimeFinder/Snap – Provides a pointer-based snapshot copy of the source. A TimeFinder/Snap copy relies on the availability of the source and can be made from any device type. It is also different from TimeFinder/Clone of TimeFinder/Mirror as it does not actually create a full copy of the source data.



11.1.1. TimeFinder/Mirror (symmir) Configuration Guidelines

- BCV devices should use the same device types as standard volumes. Creating BCVs from a slower device than the standard volume may have some performance impact on the application while the BCV is established.
- Where possible BCV devices should be implemented on the DR (R2) side of the replicated environment. Having BCV devices on the source Symmetrix increases the risk where the storage operator or systems administrator may in error issue the wrong BCV command and overwrite the data on the standard volumes. This scenario is unlikely on the remote side where a BCV restore command automatically suspends the RDF links to prevent an accidental restore to the source volumes.
- BCV devices should be spread across all the physical devices on the Symmetrix to provide best overall performance. Segregation of devices from standard devices may cause performance degradations. This can be done using the following BCV flag options (`-opt` and `-opt_rag`).

- In a SRDF/S configuration where the R2 devices are not mirrored, BCV devices are usually configured and used as the R2 mirror to provide a test DR environment when needed.
- SRDF RAID-5 configuration with RAID-5 BCV volumes will not use any additional mirror position because RAID-5 BCVs make use of clone technology ([Please reference the Appendix D of the SRDF/A Solution Guide for further details](#))

11.2. Out -of-frame Replication (SRDF)

The Symmetrix Remote Data Facility (SRDF) is a business continuance solution that maintains a mirror image of data at the device level in Symmetrix arrays located in physically separate sites.

The main goal of SRDF is to provide a recovery point solution in the event of a site or component failure between remotely mirrored devices. On older DMX-2 and Symm5 arrays SRDF was only supported in a RAID-1 configuration. Starting with the DMX-3, SRDF using RAID-5 protection is being offered/supported as an alternative (lower price point) option to RAID-1 configuration. As usual, business applications owners will be required to clearly convey application requirements to storage management in order to deploy the optimal storage solution.

There are 3 basic protocols (modes) for SRDF of which only 2 of them are prevalent at Citigroup.

Synchronous (SRDF/S) – Provides disaster recovery within the customer's campus or metropolitan network through real-time synchronous remote replication from one Symmetrix to one or more Symmetrix systems. Synchronous SRDF *writes* from the host to the *primary* (sometimes referred to as the *source*) do not finish until it is acknowledged at the primary that the complete write is resident in the *secondary* (sometimes referred to as the *target*) Symmetrix system's cache. It should be noted that in a SRDF/S configuration, there is always a performance penalty on writes because of the data has to be written to both local and target Symmetrix arrays.

Asynchronous (SRDF/A) - mode is a high-performance extended distance asynchronous replication feature that uses a Delta Set architecture resulting in reduced bandwidth requirements with typically no performance impact to hosts. There is no write performance penalty in SRDF/A configuration because the host gets the acknowledgement once the write is committed to the cache of the local Symmetrix array. ([Please consult SRDF/A Solutions Guide for further details](#))

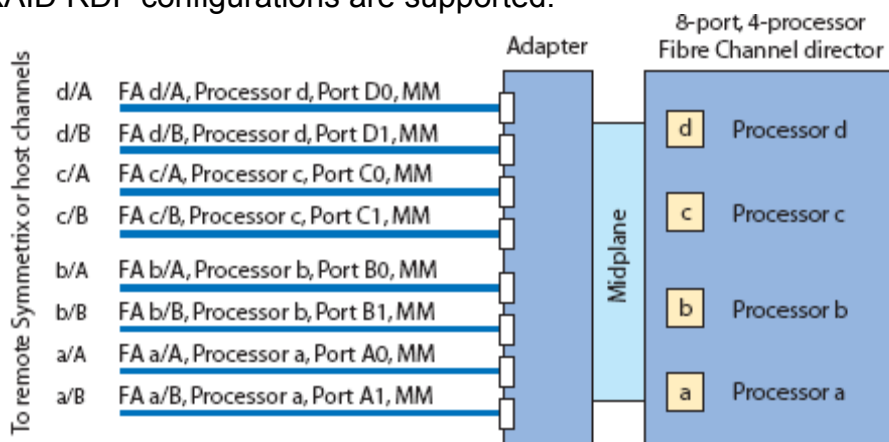
Adaptive Copy - Provides long-distance bulk data transfers for data center relocations and content replication. This option is well suited for array migrations.

The two prevalent SRDF modes at Citigroup are SRDF/S and SRDF/A. Adaptive Copy is not widely used at Citigroup and therefore not considered standard. With the introduction of the DMX architecture Citigroup has changed the SRDF Adapters from ESCON to Fibre Channel (RF).

11.2.1. SRDF Configuration guidelines

Fibre Channel director (RF)

- All Fibre Channel director ports have only seven (7) buffer-to-buffer credits. To support remote replication, the use of a Fibre Channel switch is necessary to increase or alleviate buffer-buffer credit starvation).
- A standard Eight-Port/Four-Processor RF in the Symmetrix DMX Fibre Channel directors support up to four SRDF/Fibre Channel links (one per processor) at 2Gb/s. If one port is used the other port on that processor cannot be used.
 - o Each processor provides 2 host connections, If not configured for SRDF.
 - o SRDF must be configured using Port 0 on each processor slice, starting with D0 downwards. However, if multiple RF directors are installed, port 0 of processor “d” should be used on all cards before moving downwards.
- SRDF/A is only supported today with McData IPS routers ([Please consult SRDF/A Solutions Guide for further details](#))
- SRDF/S configuration should always be set to “switched” on the DMX environment at all times. This configuration enables multiple SRDF device groups to share the same “switched” RF ports, either by static definition in the Symmetrix Configuration file or by dynamic SRDF device groups.
- The DMX architecture supports up to 64 RDF groups.
- Older configuration of Symm 5s using ESCON SRDF can only support one SRDF device group.
- Mixing RAID-1 and RAID-5 SRDF is not recommended. Only like-to-like RAID RDF configurations are supported.



12. Symmetrix DMX Management

Symmetrix DMX Array is managed with the Solutions Enabler application. Please see the GDSE solutions guide for Solutions Enabler at http://engineering.ti.citigroup.net/gdse/Products/ESS_Symmetrix.htm for information on the latest certified version of Solutions Enabler.

13. Vendor Remote Access to Symmetrix

According to a new standard released on December 9, 2004, vendor remote access to the Symmetrix will now go through Citigroup standard dial-in infrastructure called “WWDS” as illustrated in the following figure. Please see the “Vendor Remote Access to All Third Parties” document from the WWDS engineering web site at <http://www.citigroup.net/tie/neteng/external/wwds> for details.

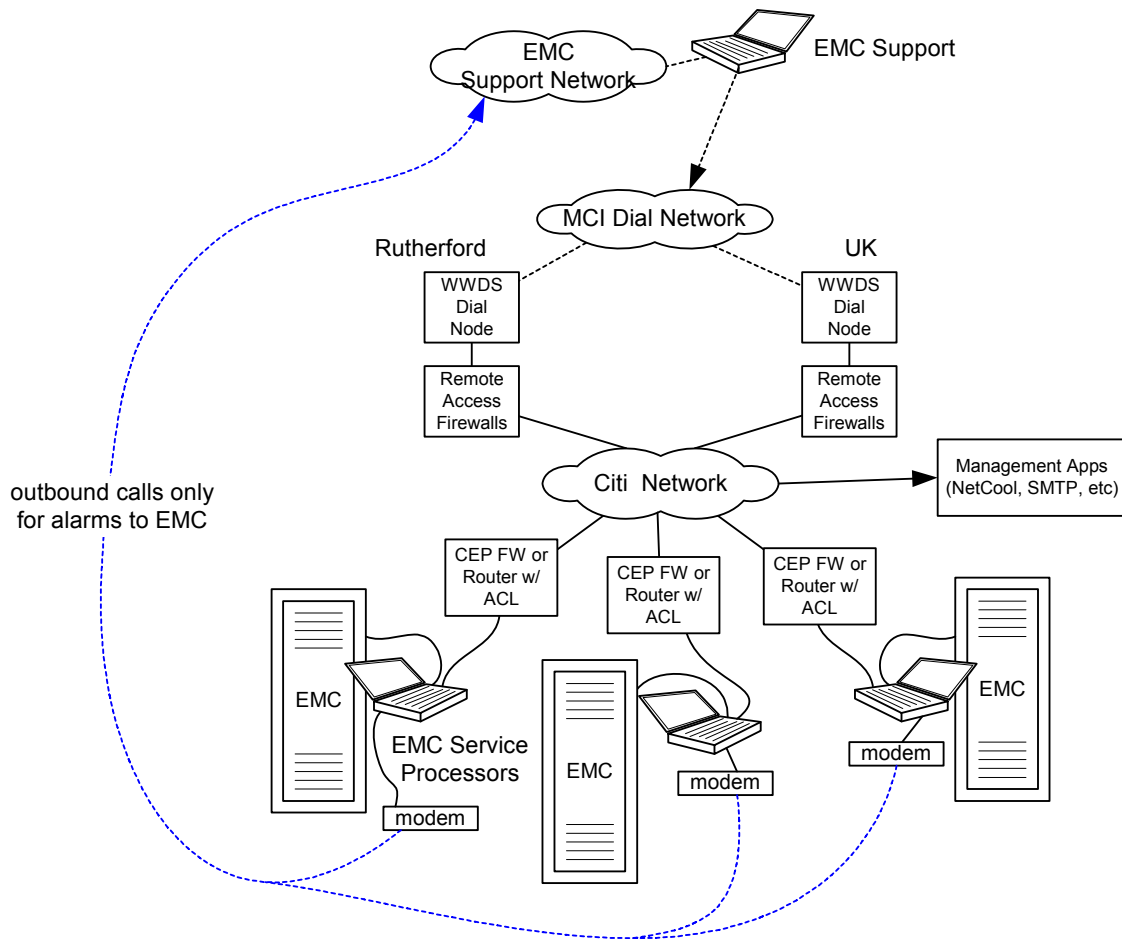


Figure 1. WWDS based Vendor Remote Access

13.1. Notes on WWDS Implementation

The implementation of WWDS for vendor remote access is a major change in the environment and requires detailed planning with several teams including local and corporate EMC customer support, Citigroup Network Operations, Firewall Operations and WWDS Operations. The following is a high level summary of tasks that should give the reader an idea of the work involved:

- Currently the Symmetrix Service Processor is not connected to a Citigroup IP network. In some data centers, Ethernet/IP infrastructure that includes Ethernet switches and a UTP cable plant must be built to service these processors.
- Routable IP addressing is needed for each Service Processor
- The Service Processors must be reconfigured to connect to the Citigroup IP network and accept remote connections from the IP network instead of the modem. Further, the PrivateWire firewall must be reconfigured to allow such

access. Some Service Processors cannot be configured in this fashion and must be replaced. Local EMC personnel performing the assessment, replacement and configuration should read EMC internal and confidential document "Customer Service Procedure: Connectivity for Citigroup", released on 7/25/05 (EMC reference #CSP2068)

- Certain data centers do not have the Common Extranet Platform ("CEP"). In such cases, one must either be built, or a firewall/router based ACL must be used to isolate the Service Processors from the Citigroup Corporate Network
- The nearest WWDS dial node must be identified and checked for capacity: additional node capacity may be required to be built.
- Hardware tokens must be obtained and issued to personnel who will own the access control on Citigroup side
- One or more IP address pools must be assigned to EMC callers on WWDS and associated with the hardware tokens
- The WWDS and CEP firewalls must be configured to allow remote access protocol traffic based on WWDS IP pool and IP addresses assigned to Service Processors
- A process would have to be created for verifying need for remote access and subsequently granting and logging the access
- Citigroup as well as EMC local and support center personnel would need to be trained on the new remote access process

14. ETrust Audit

Etrust Audit is being used to collect SymmWin/SymmRemote/EMCRemote log events and store them in a centralized repository. Refer to the [eTrust Solutions Guide](#) for details on establishing and configuration of an eTrust environment.

Etrust Audit is supported on the DMX-3 architecture with the help of a new installer script created by EMC, which allows Citigroup to install the Etrust software. The main reason for the modified script was the change in the symwin installation directory which is now drive "o:\\" as opposed to the older installation directory "c:\\".

15. Symmetrix Optimizer

Symmetrix Optimizer which runs in the background of the service processor is fully tested and certified across the DMX storage platform. (Please consult the GDSE Website for further details on Symmetrix Optimizer

http://engineering.ti.citigroup.net/gdse/GDSE_docs/GDSE-06008-SG.pdf)

- Given that the DMX-3 is more versatile than the DMX-2 architecture and yet similar in architecture, optimizer should be performed as well on the DMX-3. However, the use of optimizer is not yet recommended on the DMX-3 due to some outstanding best practices around in-the-box tiers of storage. Please stay tuned for the updated optimizer documentation.

16. Symmetrix DMX Security

16.1. SAN Infrastructure and Management Components

In lieu of the current CITMP SAN IS Product TRAM analysis, GDSE has taken action to certify improved security options for the EMC Symmetrix/DMX storage arrays.

The following SAN infrastructure was used for the testing and certification of the “VCMDB_restricted_access” and SYMACL:

Network Infrastructure:

McData ED6140 – Core Fabric Switch
McData ED3216 – Host Edge Switch
McData E/OS: v07.01.03, 09.00.00 76
McData Connectrix Manager: v8.7.1
Symmetrix DMX1000: Enginuity code v5670.73.69
Symmetrix DMX-3: Enginuity code v5771.83.92, v5771.87.95

Management Hosts:

UNIX Hardware:	Sun-Fire V480R Sun-Fire V240 Sun-Fire V440 Emulex LP-9002L, LP10000L
UNIX Software:	Solaris v2.8 EMC Solutions Enabler v6.0.3 (required for DMX-3) PowerPath v4.5 LPFC Revision: 3.92A3, Driver: 6.0e2
WinTel Software:	Windows 2000, SP3 PowerPath: v4.5 LPFC Revision: 3.92A2, Driver: v5-2.40a2

16.2. Volume Control Masking Database (VCMDB) - Host/HBA Restriction

The Symmetrix has a defined set of Symmetrix-Wide configuration metrics that include the following:

max_hypers_per_disk
dynamic_rdf
FBA_multi_access_cache
raid_s_support
VCMDB_restricted_access (no longer used)
raid_s_members
raid_5_support

In this section, we will focus mainly on the restricting unauthorized access on the EMC Symmetrix architecture (i.e. Sym5, Sym6 or Sym7 (DMX-3) arrays).

Please note that “VCMDB_restricted access” feature is no longer supported on 5771 microcode (i.e. DMX-3). The same functionality has been ported into symacl. For further information on why VCMDB_restrict is no longer supported, please visit the following primus case([emc12841](#)).

Although the “VCMDB_restricted_access” parameter, which enables you to restrict access to the device-masking database, is still supported on older Symmetrix architecture, it will no longer be the recommended method to secure the VCMDB. The new method for securing the VCMDB will be via “symacl” using a dedicated device pool for the vcm device. This method stated above, applies to all Symmetrix array platforms ranging from Symm5 to Symm7.

16.2.1. Restricting *symmask* Access on EMC Symmetrix

As stated above, “VCMDB_restricted access” is no longer supported on the 5771 microcode or DMX-3 due to a known issue. The new method to restrict SYMMASK/VCM access across all Symmetrix array is detailed below.

Once symacl access has been configured on the admin host(s) ([Refer to sec 15.3](#)), the storage administrator would need to create a device pool (i.e. VCMdb) and add the vcm device into that pool. Once the vcm device has been added to a device pool, in this case VCMdb pool, all other hosts with the exception of hosts with admin rights will NOT be able to access this device for change functionality. For hosts/servers running symagent for ECC use, a special ECC access would need to be granted to those hosts for that functionality ([see section 15.2.2](#))

This can be done by following the steps below:

Step 1: Create the command file on the Management Server

```
# cat vcmdb_pool.cmd
  create accpool VCMdb;
  add dev 0008 to accpool VCMdb;
```

In the example the VCM device is symmetrix device id **0008**. Once added to the symacl database, the output will be as listed below.

```
server1# symacl list -accpool
```

```
Symmetrix ID: 000184502xxx
```

Pool Name	Number of Devices	Number of ACLs
-----	-----	-----
VCMdb	1	0

```
server1# symacl show accpool VCMdb
```

```
Access Pool: VCMdb
```

```
Symmetrix ID           : 000184502xxx
Number of Member Devices : 1
```

```
Member Devices (1):
```

```
{
  Device Name
  -----
  Sym Physical
  -----
  0008 /dev/vx/rdmp/c2t6d0s2
  }
```

Device	Config	Attribute	Sts	Cap (MB)
-----	-----	-----	-----	-----
0008 /dev/vx/rdmp/c2t6d0s2	2-Way Mir	N/Grp'd	WD	8

```
server1# symacl list -acl
```

```
Symmetrix ID: 000184502xxx
```

Group Name	Pool Name	Access Type
-----	-----	-----
AdminGrp	ALL_DEVS	ADMIN
AdminGrp	ALL_DEVS	ALL

Please note that the default access for *UnknwGrp* has been removed from the symacl database and should **only** exist on systems configured for FDR backup only. Once the above rights have been setup, any client host will have very limited access to the vcm device even if it is visible to that host. (see examples below)

client1# **syminq**

Device		Product				Device	
Name	Cap (KB)	Type	Vendor	ID	Rev	Ser Num	
/dev/rdisk/c1t0d0s2		FUJITSU	MAP3367N SU*	0401 0322N026	35566479		
/dev/rdisk/c1t1d0s2		FUJITSU	MAP3367N SU*	0401 0322N025	35566479		
/dev/rdisk/c1t2d0s2		FUJITSU	MAP3367N SU*	0401 0322N027	35566479		
/dev/rdisk/c1t3d0s2		FUJITSU	MAP3367N SU*	0401 0322N026	35566479		
/dev/rdisk/c3t9d0s2		EMC	SYMMETRIX	5568 2100000000	7680		
/dev/rdisk/c3t9d240s2 GK		EMC	SYMMETRIX	5568 21000FD000	2880		
/dev/rdisk/c3t9d241s2 GK		EMC	SYMMETRIX	5568 21000FE000	2880		
/dev/rdisk/c3t9d242s2 GK		EMC	SYMMETRIX	5568 21000FF000	2880		

Note that the vcm lun /dev/rdisk/c3t9d0s2 is visible to the client host along with 3 gatekeepers. However, access to manipulate the VCMDB or any other symconfigure parameters is removed as listed in the example below.

client1# **symmaskdb -sid 2421 list database**

Symmetrix ID : 000184502xxx

Database Type : Type2

Last updated at : 04:32:07 PM on Mon Jan 29, 2007

Director Identification : FA-3A

Director Port : 0

Identifier		Type	User-generated		Devices
-----			Node Name	Port Name	-----
10000000c92e5939		Fibre	client1	port0	00FD:00FF

Director Identification : FA-14A

Director Port : 0

Identifier		Type	User-generated		Devices
-----			Node Name	Port Name	-----
10000000c9324e9a		Fibre	client1	port0	00FD:00FF

client1# **symmaskdb -sid 2421 init -f test**

Symmetrix access control denied the request

client1# **symconfigure -sid 2421 -f vcmdb_restrict preview**

Execute a symconfigure operation for symmetrix '000184502xxx' (y/[n]) ? y

A Configuration Change operation is in progress. Please wait...

Establishing a configuration change session.....Established.
Processing symmetrix 000184502xxx

Performing Access checks.....Access denied.

Error occurred while Submitting
Symmetrix access control denied the request

Terminating the configuration change session.....Done.

The configuration change session has failed.

client1#

Based on the above example, if symacl access is properly implemented, the vcm device can adequately be secure against unwanted user access.

16.2.2. Enabling ECC Support on DMX-3

Use the following procedure to set up each Symmetrix system to work with ControlCenter:

Each Symmetrix system that has Access Control enabled must set up an acc group for the ControlCenter application with the ControlCenter PIN added as a user ID. The ECC setup can be done in 3 steps

Step 1:

Create a group for the ControlCenter application:

Step 2:

Add an access PIN for ECC applications:

Where acid is **ECCPIN** and name is **ECC_PIN**

Step 3:

Grant **ALL** access to the ECC application group

Next, each host running a Symmetrix agent monitoring this Symmetrix system must have ECC access rights. Set up an access group that contains all of the Symmetrix agent hosts and give that group the Base and ECC rights.

Step 1

Create an access group for hosts running ECC Agents

Step 2

Add an access ID for an ECC agent-running host to the group where <UNIQUE ID> is the unique ID of the host returned by running symacl -unique on that host and <HOST NAME> is the name of the host:

When you finish, the entries in the access database display as follows:

symacl list -acl

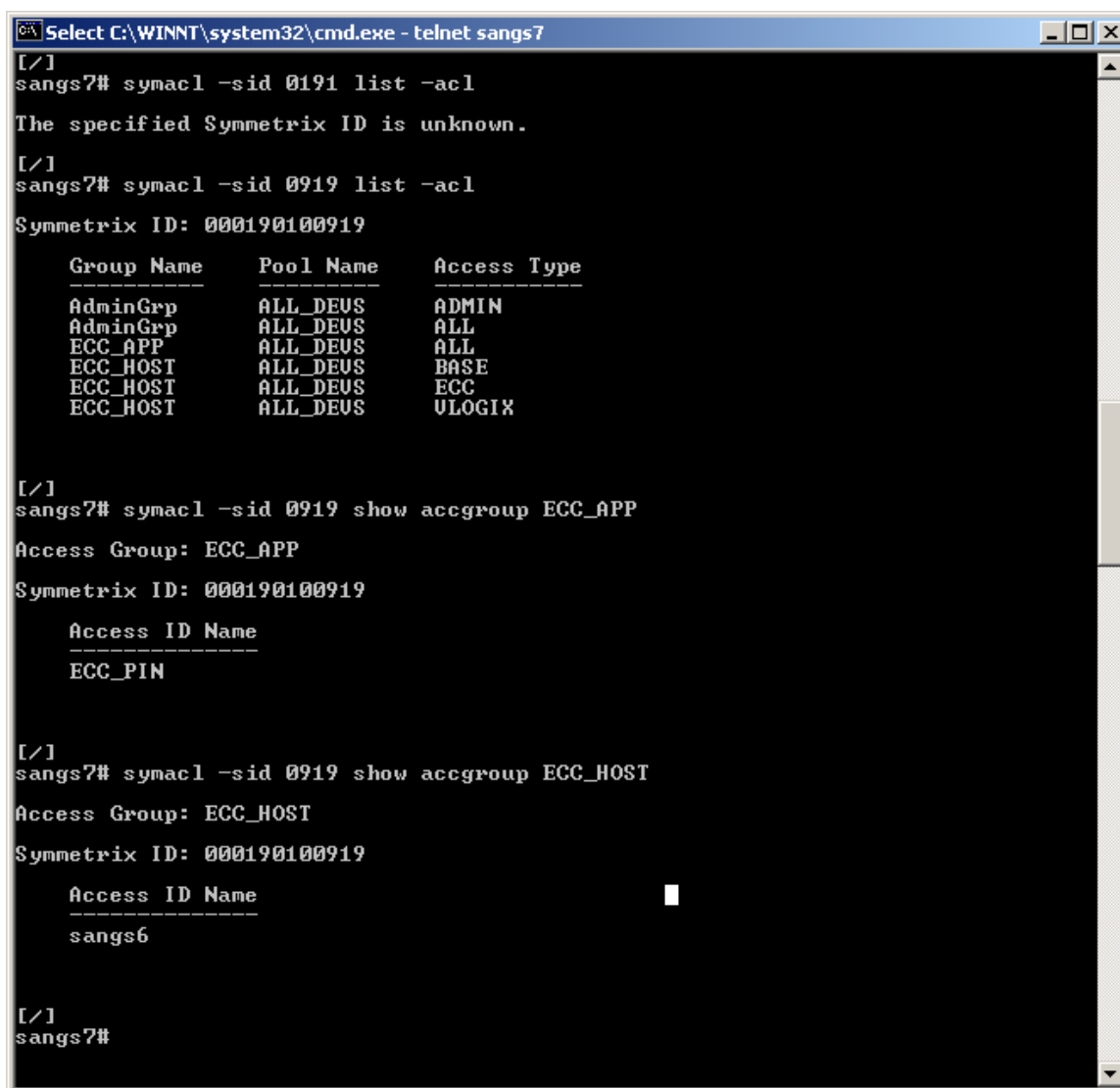
```
Symmetrix ID: 000184502xxx
Group Name  Pool Name  Access Type
-----
ECC_HOST    ALL_DEVS   BASE
ECC_HOST    ALL_DEVS   ECC
ECC_APP     ALL_DEVS   ALL
```

Symmetrix Access Control configuration is complete. Repeat this procedure for each Symmetrix system requiring ControlCenter access.

The figure below shows an example of the setup on a SAN management workstation with the necessary ECC access rights. The VLOGIX Access line is to grant other ECC_HOST (i.e. symagent hosts) access to symmask commands.

The following script combine all the steps listed above into one command file.

```
# cat sample_ecc.cmd
create accgroup ECC_APP;
add user accid ECCPIN name ECC_PIN to accgroup ECC_APP;
grant access=ALL to accgroup ECC_APP for ALL devs;
create accgroup ECC_HOST;
add host accid XXXXXX-XXXXXX-XXXXXX name sangs6 to accgroup
ECC_HOST;
grant access=BASE, ECC to accgroup ECC_HOST for ALL devs;
grant access=VLOGIX to accgroup ECC_HOST for ALL devs;
```



```

Select C:\WINNT\system32\cmd.exe - telnet sangs7
[/]
sangs7# symacl -sid 0191 list -acl
The specified Symmetrix ID is unknown.
[/]
sangs7# symacl -sid 0919 list -acl
Symmetrix ID: 000190100919

  Group Name      Pool Name      Access Type
  -----
AdminGrp         ALL_DEVS      ADMIN
AdminGrp         ALL_DEVS      ALL
ECC_APP          ALL_DEVS      ALL
ECC_HOST         ALL_DEVS      BASE
ECC_HOST         ALL_DEVS      ECC
ECC_HOST         ALL_DEVS      ULOGIX

[/]
sangs7# symacl -sid 0919 show accgroup ECC_APP
Access Group: ECC_APP
Symmetrix ID: 000190100919

  Access ID Name
  -----
ECC_PIN

[/]
sangs7# symacl -sid 0919 show accgroup ECC_HOST
Access Group: ECC_HOST
Symmetrix ID: 000190100919

  Access ID Name
  -----
sangs6

```

Figure 22: Symacl with ECC access

16.3. SYMACL – Symmetrix/DMX Host Access Control Management

Today, anyone with access to Symmetrix-based management software can execute any function on any Symmetrix device. EMC Solutions Enabler (SYMCLI) supports the “symacl” command component that allows you to set up and maintain an access-controlled environment over the Symmetrix/DMX resources (device pools).

The symacl database uses the following access control mechanisms to govern access to the Symmetrix/DMX device pools:

- *Access control groups* - Unique *access IDs* and *names* are assigned (together) to hosts and users and then sorted into *access control groups* according to similar needs (determined by an Administrator). Access groups are allowed to act on device pools based on *rights (access types)* granted by the Administrator.

The unique host ID for open systems can be viewed by running “symacl – unique”

- *Device pools* - Access rights (types), such as BCV, SRDF, ADMIN, give permission to a host or user to perform certain Solutions Enabler functionality on a specified device pool. Device pools are sometimes referred to as access pools or accpool.
- *Access control entries (ACES)*- Once the group and device pool mechanisms are established, the *access control entries (ACEs)* are created, which grant access rights to these pools. The ACEs of the various access groups (along with groups and pools) are managed and stored in the access control database.
- *Access control list (ACL)* - is a group of ACEs that are associated with the same “accgroup”. For example, the ACEs that are associated to a control group comprise an ACL in the database.

The following table illustrates a Symmetrix/DMX that is managing the access rights of various types of host connections to the Symmetrix devices. Each Symmetrix has its own access control database, which are separately managed by “symacl”.

Groups		Device Pools		ACE/ACLs		
Name	AccessID (hostname)	Name	Devices	accgroup	rights	dev/accpool
AdminGrp	SunHost1	ProdDevs	002 - 011 01C - 01F	UnknwGrp _{ACE}	ALL	ALL_DEVS
ProdB	WinHost	PoolA	012 - 019 022 - 02A	AdminGrp _{ACE}	BASE	ALL_DEVS
HR	WinHR1	PoolHR	081 - 0AA	AdminGrp _{ACE}	ADMIN	ALL_DEVS
Sales	WinSale2	SalePool	03A - 03F	ProdB _{ACE}	BASE	PRODdevs
GroupA	UXNodeB			ProdB _{ACE}	RDF	PRODdevs
UnknwGrp	Default			HR _{ACE}	BASE	PoolHR
				HR _{ACE}	BCV	PoolHR
				Sales _{ACE}	BASE	Salepool
				Sales _{ACE}	SDR	Salepool
				GroupA _{ACE}	ALL	PoolA

Table 6 - Sample "symacl" database entries

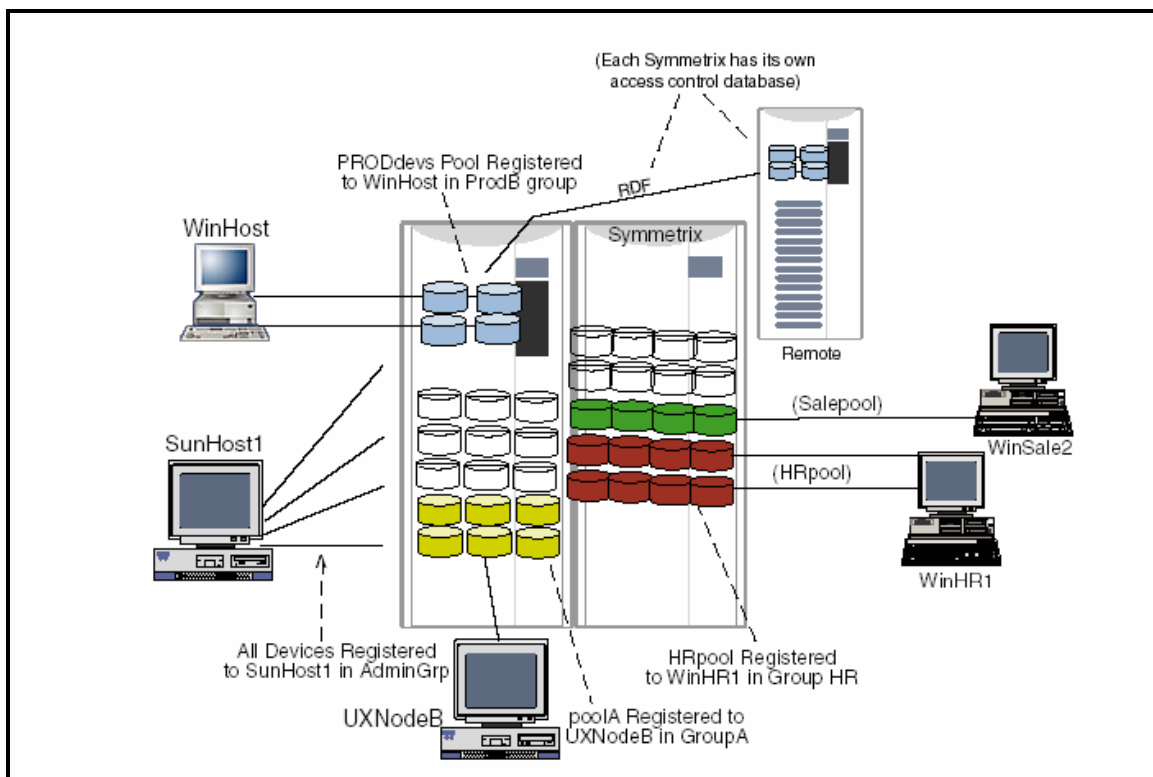


Figure 33 - Sample Symmetrix/DMX "symacl" configuration

Like "symconfigure", "syclacl" uses a command file to create or change the access control information in the database. The following are the various types of rights creation and changes possible in the command file:

- Create new access groups
- Add and remove access IDs to access groups
- Move an access ID from one group to another
- Create new device pools
- Add and remove devices to device pools
- Delete device pools and access groups
- Add ACEs to grant access
- Remove ACEs to deny access

16.3.1. SYMACL Symmetrix/DMX Setup

Initial setup of “symacl” requires the EMC CE to configure a SAN Management host for minimal access control parameters to enable you to begin administrative actions. From the SAN Management host the following is required:

- Symmetrix Serial number: *000XXXXXXXXX*
- Host ID ("symacl -unique"): *2F5802AF-524389CF-XXXXXXXX*
- Host Name (up to 8 characters): *sangs7*
- User Name (up to 8 characters): *gdseacIs*
- User ID: *0120* (*Note this is also the SYMACL_ACCESS_PIN*).

The Symmetrix serial number and symacl Host Id are specific to the Symmetrix/DMX and SAN Management host respectively. All other variables are user defined and need only follow restrictions for character length.

These variables can be retrieved using the following symcli commands:

- Symmetrix/DMX S/N
 - o **symcfg list**
- Host ID (The symacl -unique option returns the encrypted access ID in the following segmented, 24-digit numeric form:
xxxxxxxx-yyyyyyyy-zzzzzzzz)
 - o **symacl -unique**

```

C:\WINNT\System32\cmd.exe
C:\>symacl -unique
The unique id for this host is: 2F5A05A6-54408EC9-99367578
C:\>symcfg list

      S Y M M E T R I X

SymnID      Attachment  Model      Mode      Cache      Num Phys      Num Symn
              Version      Size <MB>    Devices    Devices
000184502421 Local      8530      5568      16384      3      411
000387430120 Local      1000S-M2  5670      16384      8      1075
000184700743 Remote     8730      5568      32768      0      411
000185503289 Remote     8530      5568      16384      0      791
000387430127 Remote     DMX1000S  5670      24576      0      563

C:\>_

```

Figure 44 – Host and Symmetrix/DMX variables required for initial setup of symacl

16.3.1.1. Default Configuration

Typically following initial setup of “symacl” on a Symmetrix/DMX the database contains the following default ACLs:

Groups		Pools		ACE/ACLs		
Name	Access/ID (hostname)	Name	Devices	accgroup	rights	dev/accpool
AdminGrp	SANADx	!!INPOOLS	001 - FFF	AdminGrp _{ACE}	ALL ADMIN	ALL_DEVS
UnknwGrp	Default			UnknwGrp _{ACE}	ALL	!!INPOOLS
				UnknwGrp _{ACE}	BASE	ALL_DEVS

Table 7 - Default "symacl" ACLs setup

As shown in the above table for an initial setup, access group AdminGrp has access to all devices (ALL_DEVS) and this group was granted both ADMIN and ALL rights. In addition, a group called UnknwGrp was established for unknown users, making all devices in the Symmetrix accessible to this unknown group. By default, this gives all unknown hosts BASE rights to all devices (ALL_DEVS) until such time it becomes clear as to what restrictions should be established. The initial UnknwGrp setup also grants ALL rights to (!!INPOOLS), where all devices not already assigned to an “accpool” are pooled by default.

16.3.2.2. Adding Host Access Ids to a Group

To add a host access ID to an access group, use the following syntax in the command file:

add host accid *Id* name *IdName* to accgroup *GroupName*;

- 1) Edit (addnewgroups.cmd), add the following line:
 - a) "add host accid 2F5A05A6-54408EC9-99367578 name emcsrv07 to accgroup AdminGrp;"

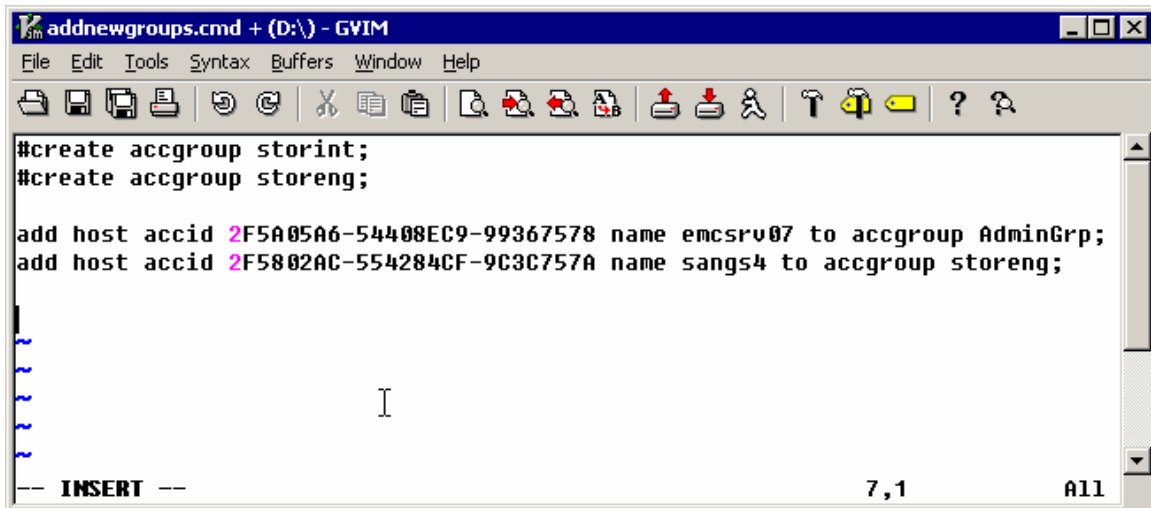


Figure 66 - symacl adding host Access ID command file syntax

16.3.2.3. Add User Access Ids to a Group

User access IDs are 4 - 12 character passwords or PINs defined by the users. When a user attempts a commit, prepare, or release operation, symacl prompts the user for a PIN. To add a user access ID to an access group, use the following syntax in the command file:

add user accid *Id* name *IdName* to accgroup *GroupName*;

- 1) Edit (addnewgroups.cmd), add the following line:
 - a) "add user accid 0120 name gdseac1s to accgroup engineering;"

```
#create accgroup storint;
#create accgroup storeng;

#add host accid 2F5A05A6-54408EC9-99367578 name emcsrv07 to accgroup AdminGrp;
#add host accid 2F5802AC-5542846F-9C3C757A name sangs4 to accgroup storeng;

add user accid sym0120 name gdseac11 to accgroup storeng;
```

Figure 77 - symacl adding User Access ID command file syntax

16.3.2.4. Access Group Maintenance

Once access groups are established, you can remove access IDs or ACEs from a group, move IDs from one group to another, or delete an entire group.

- Remove a host accid from a Group - To remove an access ID from an access group, use the following syntax in the command file:
 - o remove accid name *IdName* from accgroup *GroupName*;
- Move an Access ID to Another Group - To move an access ID from one access group to another, use the following syntax in the command file:
 - o move accid name *IdName* to accgroup *GroupName*;
- Remove all ACEs from a Group - To remove all ACEs from an access group, use the following syntax in the command file:
 - o remove aces from accgroup *GroupName*;
- Delete a Group - To delete an access group from the database, use the following syntax in the command file:
 - o delete accgroup *GroupName* [remove_aces=true];

Note: that all ACEs must be removed before you can delete a group, or you can use the “remove_aces=true” option in the command to remove all ACEs at the same time.

```
#create accgroup storint;
#create accgroup storeng;

#add host accid 2F5A05A6-54408EC9-99367578 name emcsrv07 to accgroup AdminGrp;
#add host accid 2F5802AC-554284CF-9C3C757A name sangs4 to accgroup storeng;

#add user accid sym0120 name gdseac11 to accgroup storeng;

delete accgroup storint remove_aces = true;
```

Figure 88 - symacl Delete Access Group command file syntax

16.3.3. Device Pool Management

When the various Symmetrix devices used by a host application must function as non-shareable resources, the target devices must be identified and assigned into a device pool for protection. Once a device pool is created, the pool can be a target to create access control entries (ACEs). More than one access group can access a pool with different rights.

16.3.3.1. Create a Device Pool

To create a Device Pool you must first create the pool and later add the specific Symmetrix/DMX devices that will be controlled by Access Groups.

From the SAN Management host (SANADx):

- 1) Using your favorite editor, create a command file and add the appropriate Pool name:
 - a) Edit addnewpool.cmd, add the following command:
 - i) "create accpool engBCV;"

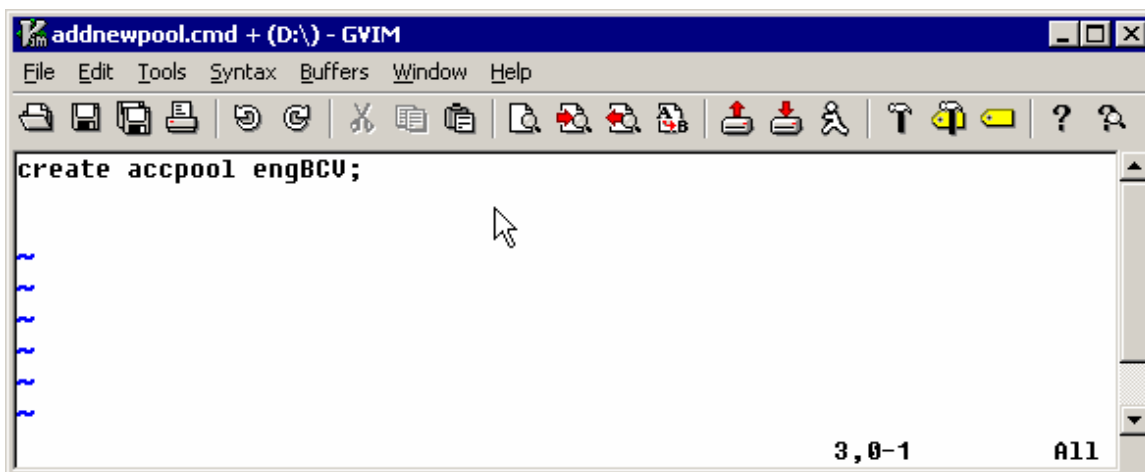


Figure 19 - symacl create accpool command file syntax

16.3.3.2. Add Devices to a Pool

- 1) Edit the command file addnewpool.cmd, use the following syntax ;
 - a) add dev *StartDevName[:EndDevName]* to accpool *PoolName*;
 - i) "add dev 0001:0200 to accpool engBCV;"

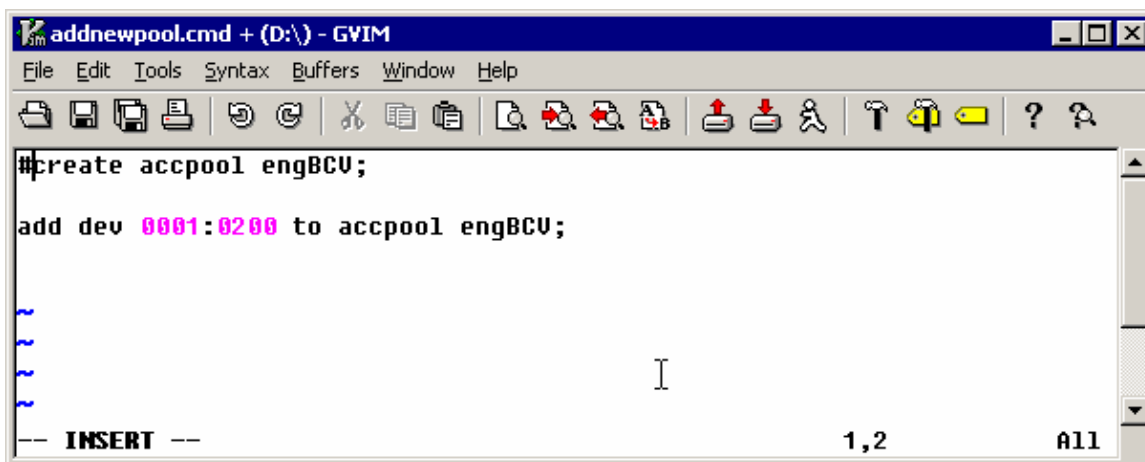


Figure 90 - symacl add dev to accpool command file syntax

16.3.3.3. Device Pool Maintenance

You can edit and manage the existing device pool information in the access control database.

- Remove Devices from a Pool - To remove specific Symmetrix devices from a pool, use the following syntax in the command file:
 - o remove dev *StartDevName[:EndDevName]* from accpool *PoolName*;

- Delete a Device Pool - To delete a device pool and remove all ACEs, use the following syntax in the command file:
 - o delete accpool *PoolName* [remove_aces=true];
Note: that all ACEs must be removed before you can delete a pool, or you can use the “remove_aces=true” option in the command to remove all ACEs at the same time.

16.3.4. ACE/ACLs Management

Once you have created groups and device pools, you can now create access control entries (ACEs) that grant access rights to the groups and pools. A group can have multiple access rights and pools, which requires an ACE for each.

The rights (*AccessType*) granted to a group for a specified device pool uses the following command file syntax:

```
“grant access=<AccessType,,,> to accgroup GroupName for <accpool PoolName> | <ALL|NON-POOLED devs>;”
```

The *AccessType* parameter specifies the rights to the SYMCLI or SYMAPI features or functionality that is to be granted to selected devices.

The following table describes the possible rights (*AccessType*) available to symacl-managed hosts on the Symmetrix DMX:

Rights (<i>AccessType</i>)	Rights Description	Commands Affected
ADMIN ^a	Grants administrator privilege to grant/deny access control entries to hosts and users.	symacl <prepare, commit, release, list, show>
ADMINRD ^a	Grants read access only to all access control information.	symacl <list, show>
ALL ^c	All possible access types granted except ADMIN and ADMINRD. Must be directed to ALL devices.	All
BASE ^a	Allows the discovery of devices and to obtain states and statistics from the Symmetrix array (directors and devices). Must be directed to ALL devices.	Base component commands
BASECTRL	Allows base control operations on devices and device groups.	symld <controls>
BCV	Allows TimeFinder (BCV) control and status operations.	symmir <controls>
BDF	Allows FasTrax functionality.	Not implemented
CACHCTRL	Allows Cache control operations concerning LRU partition management.	symqos <set LRU >

CFGDEV	Allows powerful configuration control operations that manage various types of configuration changes on devices in the Symmetrix.	symconfigure
CFGSYM ^c	Allows access to set Symmetrix attributes, set port flags, and swap RA groups with symconfigure. Must be directed to ALL devices.	symconfigure
CHECKSUM	Allows Symmetrix device Oracle checksum operations.	symchksum <controls>
CREATEDV ^c	Allows the creation of Symmetrix devices (part of symconfigure). Must be directed to ALL devices.	symconfigure
DIRCTRL ^c	Allows you to take directors and their ports offline and online. Must be directed to ALL devices.	symcfg <online/offline>
ECC ^{ab}	Allows the ECC symmetrix agent to run on the requested host.	Not applicable
ERASE	EMC Symmetrix Secure Erase allows the user to permanently and securely erase Symmetrix devices.	symerase
OPTMZR	Allows user configurable attributes that may affect the Optimizer behavior.	symoptmz <controls>
POWRPATH ^{bc}	Access to PowerPath-directed devices in an RDF consistency group. Must be directed to ALL devices.	Not applicable
QOS	Allows the execution of Quality of Service (QOS) performance control operations. Excludes LRU cache control functionality.	symqos <set pace>
RDF	Allows SRDF control and set operations.	symrdf <control>
SDDF	Allows the DeltaMark™ (Change Tracker) functionality that monitors track changes.	symchg
SDR	Allows mapping/unmapping of devices to directors/ports for the Symmetrix Disk Reallocation (SDR) feature.	symconfigure
SNAP	TimeFinder/Snapshot allows the creation and management of virtual copy sessions between a source device and multiple virtual (VDEV) target devices.	symsnap
VLOGIX ^{bc}	Enables access to Device Masking or Volume Logix devices.	symmask, symmaskdb

a. Note that this access type can only be specified for all the devices in a Symmetrix and cannot be specified for device pools.

b. See the appropriate product documentation for use of these access types.

c. These access types must be granted to either ALL devices or all NON-POOLED devices in a Symmetrix.

To effectively manage access controls in the Citigroup SOE SAN infrastructure the following symacl rights (AccessTypes) must be adhered to:

Green Rights – SAN Management Hosts (SANADx) only.

Yellow Rights – SAN Application Hosts requiring BCV and or RDF functionality.

Red Rights – These rights are inherent in the ADMIN rights and should not be required of any Application hosts, unless business justification implies otherwise.

16.3.4.1. Grant Rights (AccessType) to an Access Group

To grant “ADMIN” rights to all devices “ALL DEVS” for group “storeng” the following sequence must be completed:

- 1) From the SAN Management Host (SANADx), create a command file:
 - a) Edit “addnewrights.cmd”
 - b) Enter the following command syntax:
 - i) Grant access=BASE to accgroup storing for accpool ALL DEVS;
 - ii) grant access=ADMIN to accgroup storeng for accpool ALL DEVS;

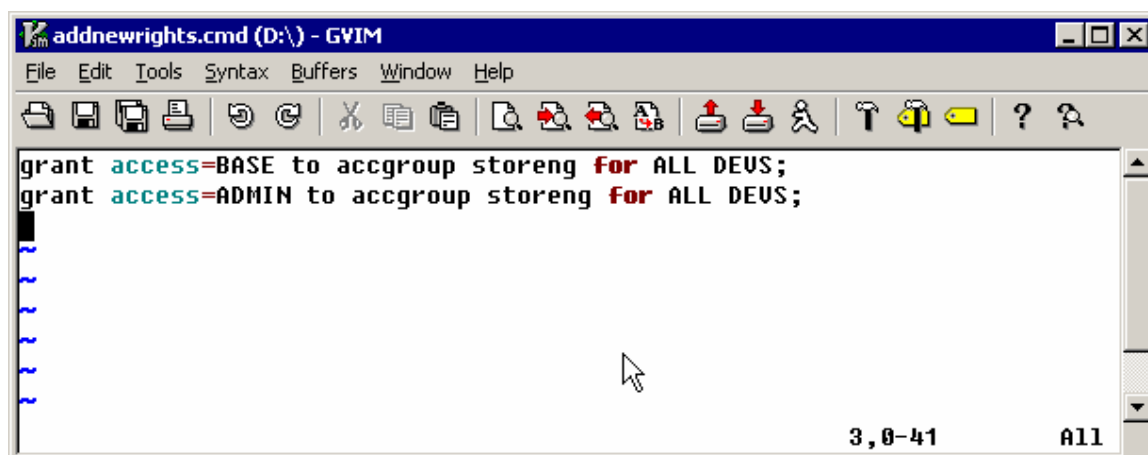


Figure 101 - granting symacl rights command file syntax

- c) commit command file:
 - i) symacl –sid xxxx –f addnewrights.cmd commit

NOTE: Before you can grant a group the rights to any value-added non-base feature, you must first grant the group BASE rights to all devices.

16.3.4.2. Remove Rights (Access Type) from an Access Group

To remove rights for an access group, use the following command in the command file:

remove access=<AccessType,,> from accgroup *GroupName* for <accpool *PoolName*> | <ALL|NON-POOLED devs>;

NOTE: At some point, you must commit the command file to execute the changes with the entries. In addition, you should execute the “symcfg configure” command after every configuration change in order to update the SYMAPI database.

16.3.5. Access Control Information

You can check the current Symmetrix/DMX access control database to view the various groups, pools, and also the ACLs. Only if you have ADMIN or ADMINRD rights, can you view all access objects in the access control database.

16.3.5.1. Listing Access Control information

Using the list argument in the following command, you can list information about groups, pools, and ACLs with the following syntax:

- `symacl [-sid SymmID|ALL] list -accgroup | -accpool | -acl | -v`

16.3.5.2. Showing Access Control Information

Using the show actions when you don't have administrative rights, you only see access objects that are associated with the access group to which your host belongs. To show detail information about a specified group or pool, use the following syntax:

- `symacl -sid SymmID|ALL show accgroup GroupName | show accpool PoolName`

16.3.6. SYMACL Backup and Restore

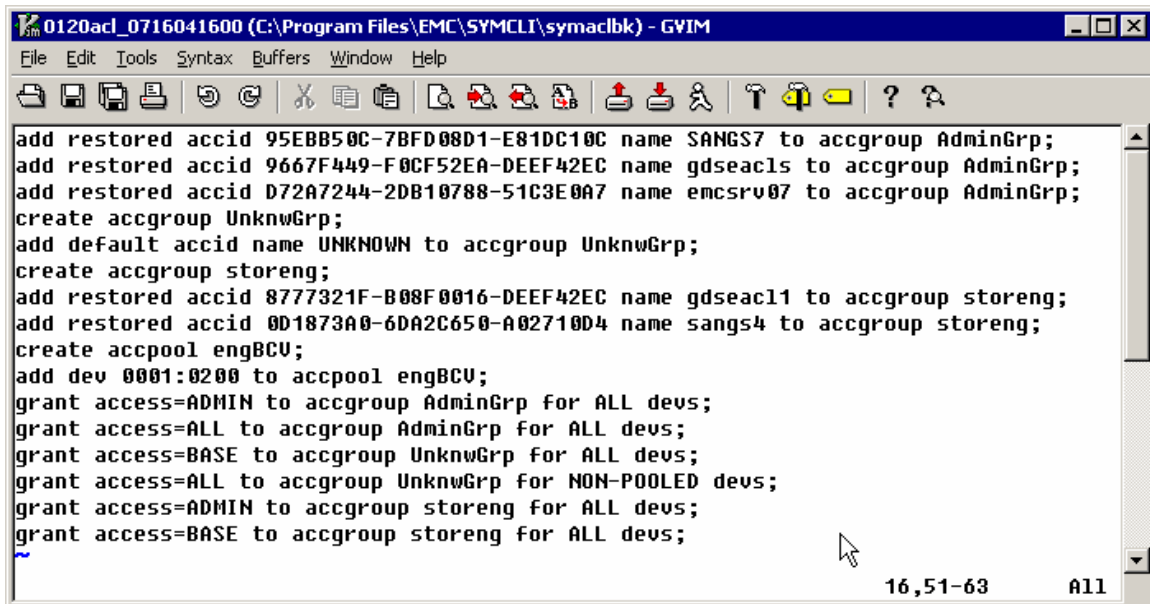
EMC has recently included the ability to backup and restore the symacl database. As a best practice, GDSE recommends, a backup copy of the symacl database be created before and after any symacl configurations are made

16.3.6.1. Symacl Backup

The backup operation saves the contents of the access control database in the file specified by the file option. The file must not previously exist. The backup file created is a text command file compatible for use with the symacl utility.

To create a backup of the symacl database, use the following command:
`symacl -sid SymmID backup -file CommandFile`

Note: The backup file contains encrypted versions of the unique IDs, therefore if you compare the values in the backup file to the original file used to create the database, they will be different.



```

add restored accid 95EBB50C-7BFD08D1-E81DC10C name SANGS7 to accgroup AdminGrp;
add restored accid 9667F449-F0CF52EA-DEEF42EC name gdseac1s to accgroup AdminGrp;
add restored accid D72A7244-2DB10788-51C3E0A7 name emcsrv07 to accgroup AdminGrp;
create accgroup UnknwGrp;
add default accid name UNKNOWN to accgroup UnknwGrp;
create accgroup storeng;
add restored accid 8777321F-B08F0016-DEEF42EC name gdseac11 to accgroup storeng;
add restored accid 0D1873A0-6DA2C650-A02710D4 name sangs4 to accgroup storeng;
create accpool engBCU;
add dev 0001:0200 to accpool engBCU;
grant access=ADMIN to accgroup AdminGrp for ALL devs;
grant access=ALL to accgroup AdminGrp for ALL devs;
grant access=BASE to accgroup UnknwGrp for ALL devs;
grant access=ALL to accgroup UnknwGrp for NON-POOLED devs;
grant access=ADMIN to accgroup storeng for ALL devs;
grant access=BASE to accgroup storeng for ALL devs;

```

Figure 112 - symacl backup file command file syntax

16.3.6.2. Symacl Restore

The restore operation replaces the contents of the access control database with the contents of the file specified by the file option. It eliminates whatever currently exists in the Symmetrix/DMX configuration database and replaces it with the contents of the backup file.

To restore a backup using symacl you use the `-restore` option as shown below:
`syclbk -sid SymmID commit [-v|-noecho] -restore -file CommandFile`

```

C:\WINNT\System32\cmd.exe
C:\Program Files\EMC\SYMCLI\symaclbk>symacl -sid 0120 commit -restore -file 0120acl_0716041600
Enter Access PIN:
Command file: <0120acl_0716041600>
PREVIEW.....Started.
PREVIEW.....Done.
PREPARE.....Started.
Adding Restored access id SANGS7 to group AdminGrp.....Done.
Adding Restored access id gdseac1s to group AdminGrp.....Done.
Adding Restored access id emcsrv07 to group AdminGrp.....Done.
Creating group UnknwGrp.....Done.
Adding the DEFAULT access id UNKNOWN to group UnknwGrp.....Done.
Creating group storeng.....Done.
Adding Restored access id gdseac1l to group storeng.....Done.
Adding Restored access id sangs4 to group storeng.....Done.
Creating pool engBCU.....Done.
Adding 512 devs, dev 0001 through 0200 to pool engBCU.....Done.
Granting access ADMIN to group AdminGrp for ALL devs.....Done.
Granting access ALL to group AdminGrp for ALL devs.....Done.
Granting access BASE to group UnknwGrp for ALL devs.....Done.
Granting access ALL to group UnknwGrp for NON-POOLED devs.....Done.
Granting access ADMIN to group storeng for ALL devs.....Done.
Granting access BASE to group storeng for ALL devs.....Done.
PREPARE.....Done.
Starting COMMIT.....Done.
C:\Program Files\EMC\SYMCLI\symaclbk>
C:\Program Files\EMC\SYMCLI\symaclbk>symacl -sid 0120 commit -restore -file 0120acl_0716041600_

```

Figure 123 - symacl database -restore command syntax

16.3.7. Strategy for Hosts requiring BCV/Symmir Controls

To set up an access-controlled environment for BCV/Symmir operations, you need set up both the standard and BCV devices as follows:

- 1) Define the working device pool to contain both the STD and BCV Symmetrix device names.
- 2) For the group name, grant BASE rights to access all devices.
- 3) For the group name, grant BCV rights to the device pool holding the pairs.

16.3.8. Strategy for Hosts requiring SRDF Controls

To set up an access-controlled environment for SRDF operations, you need set up both the local and remote Symmetrix/DMX. Because both Symmetrix arrays have their own access-controlled database, you must:

- 1) From the SAN Management Host (SANADx), create an access control group for the local Symmetrix/DMX. Then with a local or remote SANADx host, define the same access control group name for the remote Symmetrix.

The policy for complete control of the Symmetrix/DMX, is to only register access to certain devices on an as-needed basis. To accomplish this, the UNKNOWN user ACEs and userID are removed from the UnknwGrp. Therefore, the Symmetrix/DMX must know the application host/node, and only specific users/hosts in defined groups with limited or unlimited rights have access to the devices defined in their working pools.

[illegible]

17. Appendix A - EMC Symmetrix DMX-3 950 Configuration Summary:

DMX-3 950 Introduction:

The role of the EMC Symmetrix DMX-3 950 is to provide top tier array features and functionality in a compact package, allowing for use in datacenters where space and power are severely limited. The DMX-3 950 is the smallest model in DMX-3 product line, with 25% of the drive capacity and 17% of the throughput as compared to the DMX-3 4500 model. The DMX-3 950 achieves this compact size through the use of multi-functional directors and adapters, providing a combination of both front-end and back-end I/O processing on the same hardware module.

DMX-3 950 Specifications:

1. The primary architectural difference between the DMX-3 950 and the standard DMX-3 is the combination of functions provided by the MultiProtocol Channel Directors (MPCD) and Front-End Back-End (FEBE) adapters. Figure 1, depicts the connectivity relationship between the MPCD director and FEBE adapter:
 - a. The MPCD simultaneously functions as a channel director (front-end) and disk director (back-end), where a typical DMX-3 handles these functions on separate physical director units. Functionally, the MPCD is similar to a standard director with the appropriate mezzanine boards and firmware. The MPCD is not field upgradeable to a standard director.
 - b. The FEBE simultaneously functions as a channel adapter (front-end) interfacing the Fibre Channel SAN and a disk adapter (back-end) accessing the physical disk drives. Standard DMX-3 models divide these functions on separate channel and disk adapter modules. The FEBE is not field upgradeable to a standard adapter.

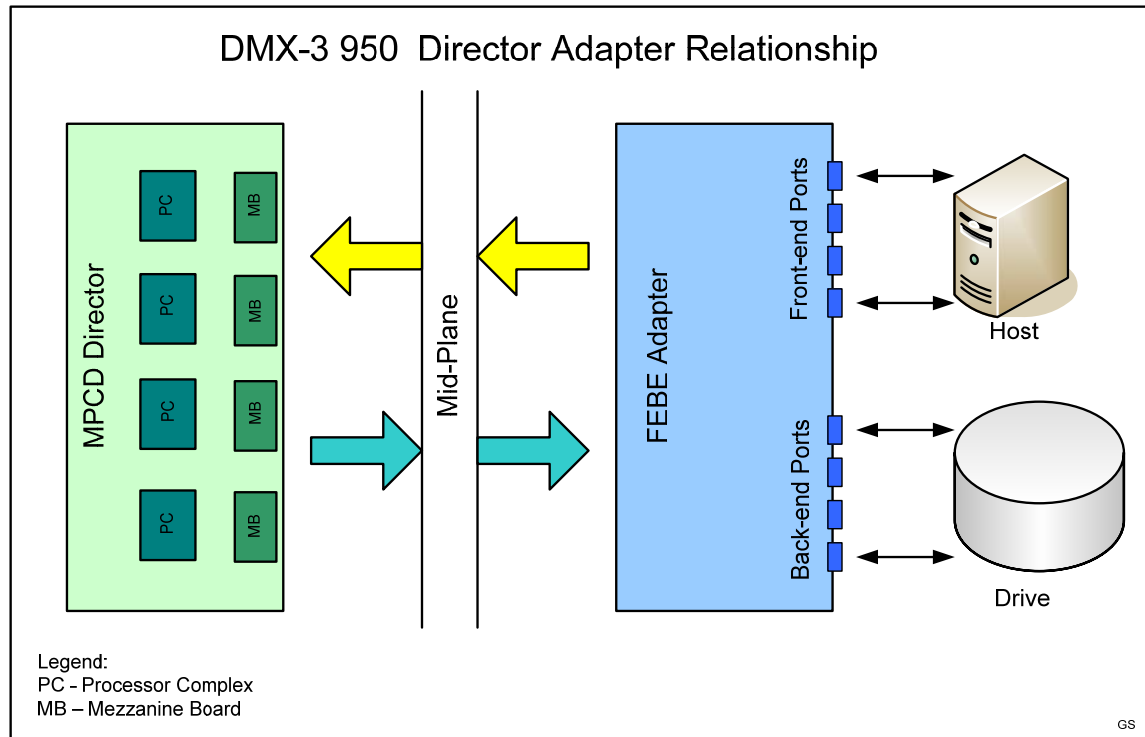


Figure 1:

2. The DMX-3 950 system bay use a mid-plane design that houses a 6 slot horizontal layout card cage, as compared in the typical DMX-3 that uses a 24 vertical slot layout. The horizontal layout allows 8 disk enclosures (DE) with a total capacity of 120 drives to be incorporated into the system bay.
3. The DMX-3 950 storage bay accommodates 120 drives in 8.73sqft footprint. The physical disk drives are partitioned into DE's of 15 drives each. Figure 2, depicts the DE layout for a system and storage bay.

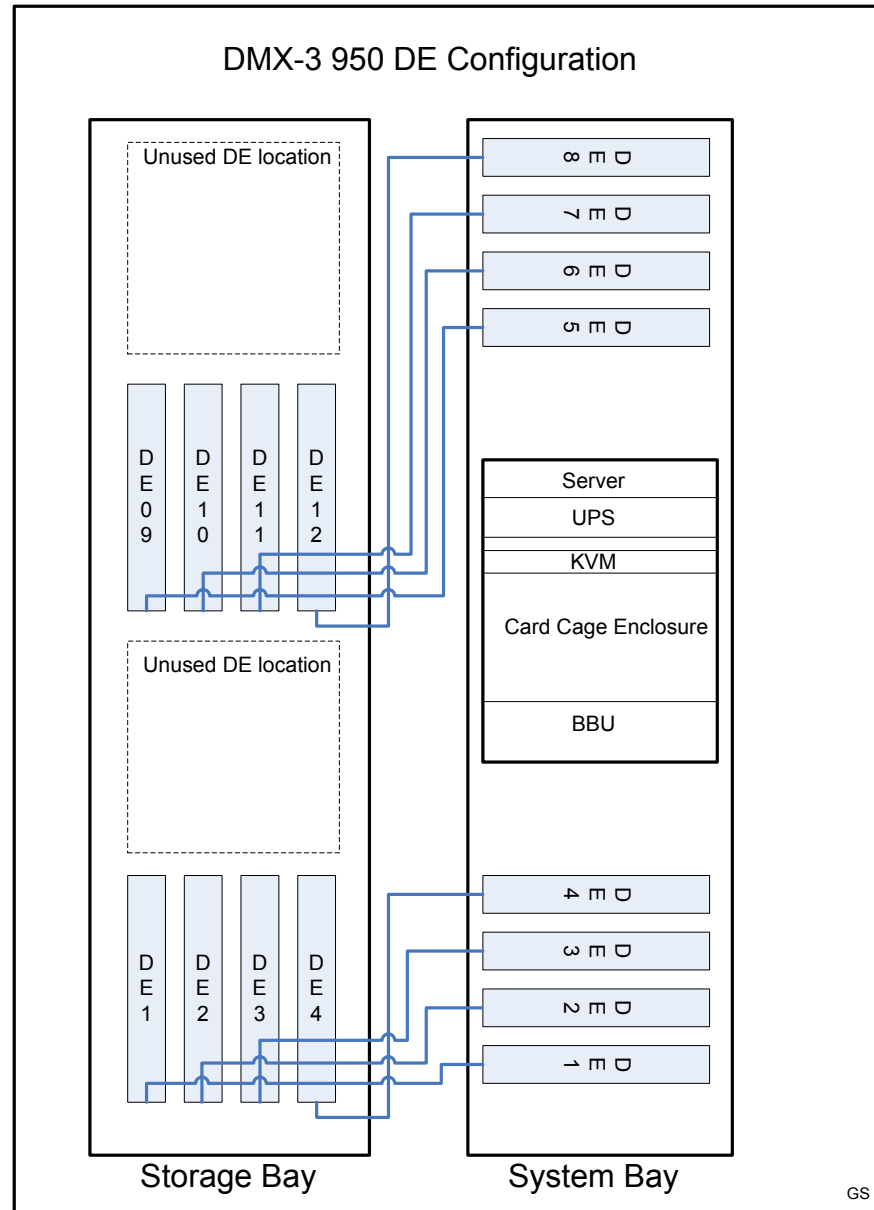


Figure 2:

4. The four director DMX-3 950 accommodates a maximum of 31,140 mirror and 8192 RAID5 (7+1) logical volumes.
5. Fan In/Out ratio for the DMX-3 950 is 128 maximum initiators per port, same as a standard DMX3.
6. DMX-3 950 supports the SRDF family of software (e.g. synchronous and asynchronous) as a means to provide high data availability and disaster recovery.

DMX-3 950 Configuration Guidelines:

1. **The four director DMX-3 950 is the only director configuration supported by Citi.** The 4 director configuration supports 8 loops at 15 drives pre loop in a single bay and 30 drives per loop in a dual bay configuration.
2. **The DMX-3 950 two director variant that will not be supported by Citi.** EMC does not support Raid-5 7+1 (Citigroup certified) on a 2 director (4 loops) system, due to the potential of data loss when distributing 8 stripes across four loops.
3. **The DMX-3 950 system bay with an associated storage bay will be supported by Citi.** Extending the DMX-3 950 capacity beyond 120 drives requires the addition of a storage bay for an additional 120 drives. The storage bay has a maximum capacity of 240 drives; however, for performance concerns the storage bay will be limited to 120 drives. A system and storage bay will utilize 15.6sqft floor space and approximately 8.3KVA of power.
4. Host Fibre Channel connectivity and RDF groupings with SRDF replication. The DMX-3 950 contains 16 mezzanine boards, of which 8 is dedicated to support front-end processing. Each mezzanine board dedicated to the front-end accommodates 1 RDF port or 2 Fibre Channel ports. A maximum of 4 RDF ports, or one per each director module, will be supported to maintain a proper balance of remote and local connectivity:
 - With 2 RDF ports (8 dual path SRDF/A applications), 12 ports are available for host connections.
 - With 4 RDF ports (16 dual path SRDF/A applications), 8 ports are available for host connections.

Note: For SRDF/A, a maximum of 8 RDF groups can be supported by each RDF port pair with 5x71 microcode. This figure will increase with 5772 microcode (see SRDF/A Solutions Guide).

5. Data protection schemes and volume sizes remains consistent across all DMX-3 models including the 950. The supported logical host volumes of 9208 cylinders (8633MB) for mirror and 36828 cylinders (34526MB) for Raid-5 7+1 volumes
6. The role of the DMX-3 950 array is to provide the highest level of storage performance and reliability in a compact package. As with standard DMX-3 models, the DMX-3 950 supports both SRDF and host cluster functionality. Selecting between the DMX-3 950 and the larger DMX-3

models should be based on the initial and projected capacity requirements and available datacenter resources (e.g. floor space, power, cooling). If the 24-36 month projected storage requirement is less than **7TB for mission critical** data (mirror, 146GB drives) or **25TB for business critical** data (Raid-5 7+1, 300GB drives) for the entire array, the DMX-3 950 system should be selected. If the projected storage requirements are greater than the previous stated values, the DMX-3 2500 model or larger should be selected. Upgrading of the DMX-3 950 with a storage bay (120 maximum) is supported in the event the initial system bay installation provides insufficient capacity. The table below compares several of the distinguishing variables between the DMX-3 950 and 2500 models:

Features/Specification	DMX-3 950	DMX-3 2500
Drive loops	8	16
Maximum number of disks	120	480
Maximum raw capacity (300GB)	36TB	144TB
Maximum Logical Volumes R5 7+1	8192	15,300
Global Memory	64GB	144GB
Maximum Data Paths	32	64
Maximum FE Connectivity	16 Fibre Channel	32 Fibre Channel
Foot Print	6.87-15.6sqft	15.4sqft (240 drv)
Power consumption	4.7-8.7kVA	11.5kVA
Heat dissipation	16,036-29,684BTU/hr	39,250BTU/hr

18. Appendix B – Seagate Drive Specs

The Seagate chart below depicts the profiles of the 73GB and 146GB drives:

Model Number	ST3146855LW/LC/FC/SS	ST373455LW/LC/FC/SS
Capacity		
Formatted 512 Kbytes/Sector (GB)	146.8	73.4
External Transfer Rate (Mbytes/sec)		
Ultra320 SCSI	320	320
4-Gbit/sec FibreChannel	400	400
3-Gbit/sec Serial Attached SCSI	300	300
Performance		
Spindle Speed (RPM)	15K	15K
Average Latency (msec)	2	2
Seek Time	3.5/4.0	3.5/4.0
Average Read/Write (msec)	0.2/0.4	0.2/0.4
Track-to-Track Read/Write (msec)	960 to 1607	960 to 1607
Transfer Rate	89 to 150	89 to 150
Internal (Mbits/sec)	73 to 125	73 to 125
Internal Formatted (Mbytes/sec)	16	16
Configuration/Organization		

Disks	2	1
Heads	4	2

19. Appendix C – VCMDB Write Disable Procedure:

VCMDB should **never** be used in creating a host volume. Therefore, in Engenuity 5771 and later, GDSE recommends that VCMDB device be write disabled at the earliest opportunity. See [section 10.9](#) for additional information:

1. From a Solution Enabler 6.0.3 or 6.3.1 server, write disable the VCMDB device: (note – device 008 will represent the VCMDB throughout this example)

```
“symdev –sid XXX write_disable 008”
```

2. If one of the following error messages appears during the attempt to write disable the VCMDB device, proceed to step 3:

The gatekeeper device (while using the Base Daemon) has an error (Please see the Log file)

There is no gatekeeper device that can be used to communicate to the Symmetrix

(assume additional gatekeeper devices are available, if not add gatekeepers to the host and rerun step 1)

3. List and shutdown all stordemons:

```
“stordemons –list”
```

```
“stordemons shutdown all”
```

4. With a text editor, open the “/usr/emc/API/symapi/config/daemon_options” file and verify that the “storapid:client_use=” option is set to the default entry of “require” or is commented out. In certain versions of Solution Enabler 6.3, the “storapid:client_use” option is not present in the “daemon_options” file, which is acceptable for write disabling the VCMDB.
5. Create a file “/usr/emc/API/symapi/config/symapi_gkmgt” with the contents of “symapi_gkmgt=disable”. (do not include the quotes in the file)
6. Restart each stordemon listed in step 3:

Sample command:

```
“stordemons start storapid”
```

7. Write disable and verify the status of VCMDB device:

```

“symdev –sid XXX write_disable 008”
“symdev –sid XXX show 008”

```

If disabling the VCMDB is successful (see highlighted test in the box below) this procedure is complete. If disabling the VCMDB fails, proceed to step 8.

Device Physical Name	: /dev/rdsk/c2t16d0s2	
Device Symmetrix Name	: 0008	(VCM)
Device Serial ID	: 1900008000	
Symmetrix ID	: 000190100919	
Attached BCV Device	: N/A	
Attached VDEV TGT Device	: N/A	
Vendor ID	: EMC	
Product ID	: SYMMETRIX	
Product Revision	: 5772	
Device WWN	: 60060480000190100919533030303038	
Device Emulation Type	: FBA	
Device Defined Label Type	: N/A	
Device Defined Label	: N/A	
Device Sub System Id	: 0x0001	
Device Block Size	: 512	
Device Capacity	{	
Cylinders	:	3
Tracks	:	45
512-byte Blocks	:	5760
MegaBytes	:	3
KiloBytes	:	2880
	}	
Device Configuration	: 2-Way Mir	(Non-Exclusive Access)
Device is WORM Enabled	: No	
Device is WORM Protected	: No	
SCSI-3 Persistent Reserve	: Disabled	
Dynamic Spare Invoked	: No	

Dynamic RDF Capability : None

STAR Mode : No

STAR Recovery Capability : None

STAR Recovery State : NA

Device Service State : Normal

Device Status : Write Disabled (WD)

Device SA Status : Write Disabled (WD)

8. Reset all 0x45 semaphores on the Solution Enabler server.

A. Example of Resetting the Semaphores:

1) List semaphores:

```
# ipcs -s
IPC status from <running system>
T      ID   KEY      MODE      OWNER   GROUP
Semaphores:
s       0  0x45000cd3 --ra-ra-ra-  root   root
s       1  0x45000dae --ra-ra-ra-  root   root
s       2  0xcace      --ra-ra-ra-  root   root
s       3  0x450001ae --ra-ra-ra-  root   root
s       4  0x4500027e --ra-ra-ra-  root   root
s       5  0x4500028e --ra-ra-ra-  root   root
```

2) Remove semaphores beginning with 0x45:

```
# ipcrm -s 0
# ipcrm -s 1
# ipcrm -s 3
# ipcrm -s 4
# ipcrm -s 5
```

3) Rediscover available frames:

```
#symcfg discover
```

This operation may take up to a few minutes. Please be patient...

4) Verify the restart of 0x45 semaphores:

```
# ipcs -s
```

IPC status from <running system>						
T	ID	KEY	MODE	OWNER	GROUP	
Semaphores:						
s	65536	0x450001ae	--ra-ra-ra-	root	other	
s	65537	0x45000cd3	--ra-ra-ra-	root	other	
s	2	0xcace	--ra-ra-ra-	root	root	
s	65539	0x45000dae	--ra-ra-ra-	root	other	
s	65540	0x4500027e	--ra-ra-ra-	root	other	
s	65541	0x45000161	--ra-ra-ra-	root	other	

9. List and shutdown all stordaemons:

```
"stordaemons -list"
"stordaemons shutdown all"
```

10. Verify that the "/usr/emc/API/symapi/config/daemon_options" file has the "storapid:client_use=" option is set to the default entry of "require" or is commented out.

11. Verify the contents of file "/usr/emc/API/symapi/config/symapi_gkmgt" contains the contents of "symapi_gkmgt=disable". (do not include the quotes in the file)

12. Restart each stordaemons listed in step 3:

```
Sample command:
"stordaemons start storapid"
```

13. Write disable and verify the status of VCMDDB device:

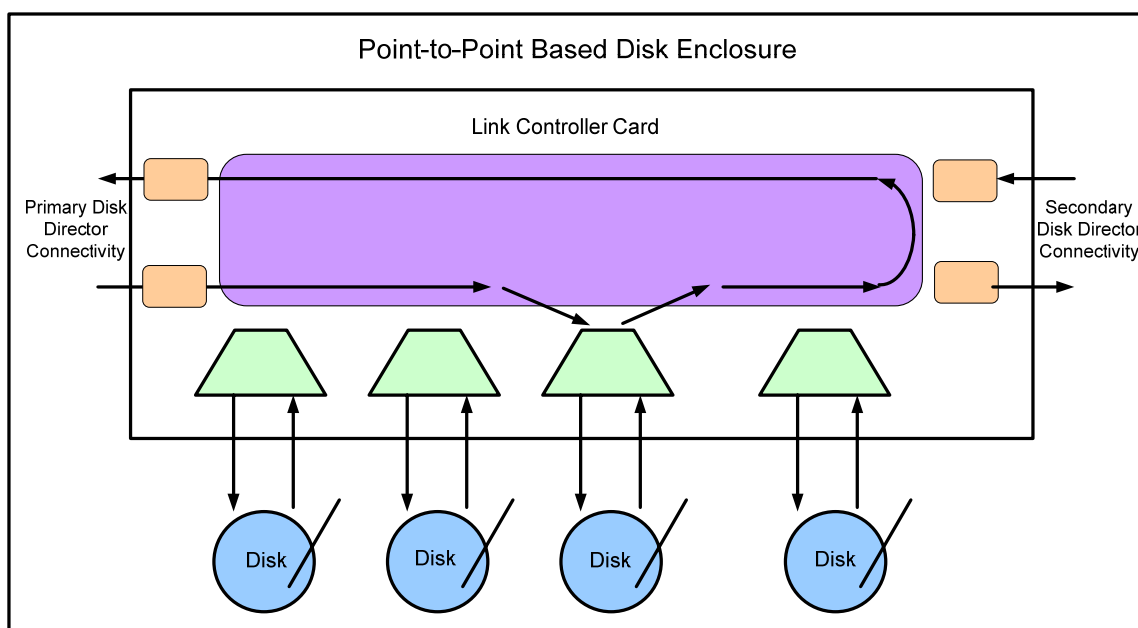
```
"symdev -sid XXX write_disable 008"
"symdev -sid XXX show 008"
```

References

Document	Location
Global Unix Engineering website	http://unixeng.ssmb.com
GDSE Website	http://engineering.ti.citigroup.net/gdse/
Solutions Enabler Solution Guide	http://engineering.ti.citigroup.net/gdse/products/platforms/symmetrix/stdsdocs/
Vendor Remote Access to All Third Parties	http://www.citigroup.net/tie/neteng/external/wwds
SRDF/A Solutions Guide	http://engineering.ti.citigroup.net/gdse/products/platforms/symmetrix/stdsdocs/SRDF/A

20. Appendix D - DMX-4 Backend Subsystem

The DMX-4 backend subsystem employs an arbitrated loop design that logically interconnects the Disk Director (DD) with the Disk Enclosure (DE). The DD, which contains monitoring and control programs, operates as the loop initiator that feeds data into specific DE's through the Link Controller Cards (LCC). The LCC provides the point-to-point FC connectivity to disk drives and monitors environmental status of the DE (see figure below). This architecture creates an independent relationship between the drive and DD, thus simplifying problem isolation and improving serviceability.



Each DE contains two LCC's for redundancy and accommodates from 4-15 disks. The LCC uses Fibre Channel Arbitrated Loop (FC_AL) protocols to

logically interconnect a disk with the respective DD. The LCC monitors the loop for Fibre Channel layer 0/1 signal connection and protocol errors. If errors are detected, the controlling software program is alerted and failing segment is isolated. The LCC also maintains basic configuration information, in the event the loop configuration is changed (i.e. new disk assembly), it will alert the control program to take the appropriate action.

A multilayer protection strategy is employed to ensure content accuracy between the DD and disk element. The connection to which data is transferred between the global memory and the disk is referred to as an exchange. Each exchange creates a logical set of parameters for the connection that is transferred as part of this exchange. 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 a Cyclic Redundancy Checksum (CRC) is computed and appended to the frame. The receiving port validates the exchange parameters and re-computes the CRC for the data received and compares it with the CRC embedded within the exchange. The sequence number is verified to ensure that the received frame is properly ordered. The FC-1 transmission layer provides another level of data integrity through 8b/10b encoding. This encoding technique takes an 8-bit word and converts it into a 10-bit transmission character, where a form of parity and checksum algorithms provides error checking and correction. Once the data has been received at the disk the overhead symbols are decoded then recoded before being stored on the disk, ensuring that future request for this data are accurate.

21. Appendix E - RAID 6 Mean Time to Failure Calculations

21.1. Description

Predicting data storage reliability can often be confusing due to the large numbers used by vendors to describe their products reliability. For many storage vendors the Mean Time Between Failure (MTBF) value of an individual drive represents the reliability of the entire frame. The MTBF value provided by vendors represents a statistical approximation of how long a set of devices should last before failure. This approximation is derived by estimating the MTBF of individual components on the device mixed with reliability experiences of similar products. Therefore, MTBF values can be useful by the marketing department to compare individual devices; however it is not valuable at predicting the reliability of a series of devices on the backend of a storage array.

A practical method to predict the overall reliability of the disks in an array is the Mean Time To Data Loss (MTTDL) calculation. MTTDL calculates the probability that data will be lost by the frame using a number of metrics: 1) MTBF, 2) Mean Time To Replace (MTTR), 3) total number of disks in the array, 4)

number of data/parity volumes in the Raid group and 5) sector error rates. The higher the MTDDL value the less likely a failure will occur.

In RAID 5 and RAID 6 structures, losing one drive over the total number of parity drive(s) will cause the Raid group to fail. Therefore, losing 2 drives in a RAID 5 group or 3 drives in a RAID 6 group will cause a failure of the Raid group an in-turn loss of data. This type of failure can be calculated using the Mean Time to Data Loss due to disk failure (MTDDL-DF) formula. A second method where data loss occurs is caused by bit error(s) when the Raid group is in a degraded (failed drive(s)) or in a rebuilding mode (replacement of failed drive(s)). This type of calculation uses the Mean Time to Data Loss due to bit errors (MTDDL-BER) formula.

21.2. Comparison of Reliability Calculations:

1. For a DMX-4 4500 equipped with 300GB 10K RPM drives, RAID 6 14+2 shows a MTDDL-DF that is 145 times greater than RAID 5 7+1 and 20 times greater than mirror devices. (note: higher MTDDL value equates to a lower probability of failure)
2. For a DMX-4 4500 equipped with 300GB 15K RPM drives, RAID 6 14+2 shows a MTDDL-DF that is 228 times greater than RAID 5 7+1 and 33 times greater than mirror devices.
3. For a DMX-4 4500 equipped with 300GB 10K RPM drives, RAID 6 14+2 shows a MTDDL-BER that is 1465 times greater than RAID 5 7+1 and 211 times greater than mirror devices.
4. For a DMX-4 4500 equipped with 300GB 15K RPM drives, RAID 6 14+2 shows a MTDDL-BER that is 2297 times greater than RAID 5 7+1 and 330 times greater than mirror devices.
5. For a DMX-3 4500 equipped with 500GB 7200 RPM drives, RAID 6 14+2 shows a MTDDL-BER that is 1128 times greater than RAID 5 7+1.

21.3. Conclusions:

- RAID 6 shows higher MTDDL-DF and MTDDL-BER values than RAID 5 or mirror devices regardless of drive type.
- RAID 6 14+2 with 300GB 15K drives provides the best balance of reliability and capacity.
- RAID 5 on SATA drives presents MTDDL-BER of approximately half a year.

21.4. Results of Reliability Calculations:

Reliability: DMX-3 4500 frame with 960 300GB 10K.7 RPM drives	MTTR High I/O (hr)	MTTR Low I/O (hr)	MTTDL disk fail High I/O (yr)	MTTDL disk fail Low I/O (yr)	MTTDL bit error High I/O (yr)	MTTDL bit error Low I/O (yr)
RAID 5 7+1	3.2	6.56	1040	507	9.99	9.99
RAID 6 14+2	3.2	6.56	151,632	36,036	14,632	7159
Mirror	3.2	6.65	7,278	3,548	69.4	69.4

Reliability: DMX-3 4500 frame with 960 300GB 15K.5 RPM drives	MTTR High I/O (hr)	MTTR Low I/O (hr)	MTTDL disk fail High I/O (yr)	MTTDL disk fail Low I/O (yr)	MTTDL bit error High I/O (yr)	MTTDL bit error Low I/O (yr)
RAID 5 7+1	2.05	3.51	1624	949	9.99	9.99
RAID 6 14+2	2.05	3.51	370,196	126,257	22,945	13,400
Mirror	2.05	3.51	11,372	6641	69.4	69.4

Reliability: DMX-3 4500 frame with 960 500GB 7.2K RPM SATA drives	MTTR High I/O (hr)	MTTDL disk fail (yr)	MTTDL bit error (yr)
RAID 5 7+1	4.06	601	0.58
RAID 6 14+2	4.06	59,217	654

References:

P. Chen, E. Lee, G. Gibson, R. Katz, D. Petterson: RAID: High-Performance, Reliable Secondary Storage. ACM Computing Surveys