

Symmetrix Foundations

Upon completion of this module, you will be able to:

- Identify the front-end directors, back-end directors, cache and disk location in the Symmetrix DMX series
- Explain the relationship between Symmetrix physical disk and Symmetrix logical volumes
- Identify volume protection options available on the Symmetrix
- Explain the I/O path through Symmetrix cache
- List the Symmetrix DMX series connectivity options
- Describe Symmetrix DMX3 Vaulting

The objectives for this module are shown here. Please take a moment to read them.



Symmetrix Foundations

EMC Symmetrix DMX Offerings

© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 2 

Now we will take a look at EMC Symmetrix DMX offerings.

Symmetrix DMX2

DMX800



DMX1000



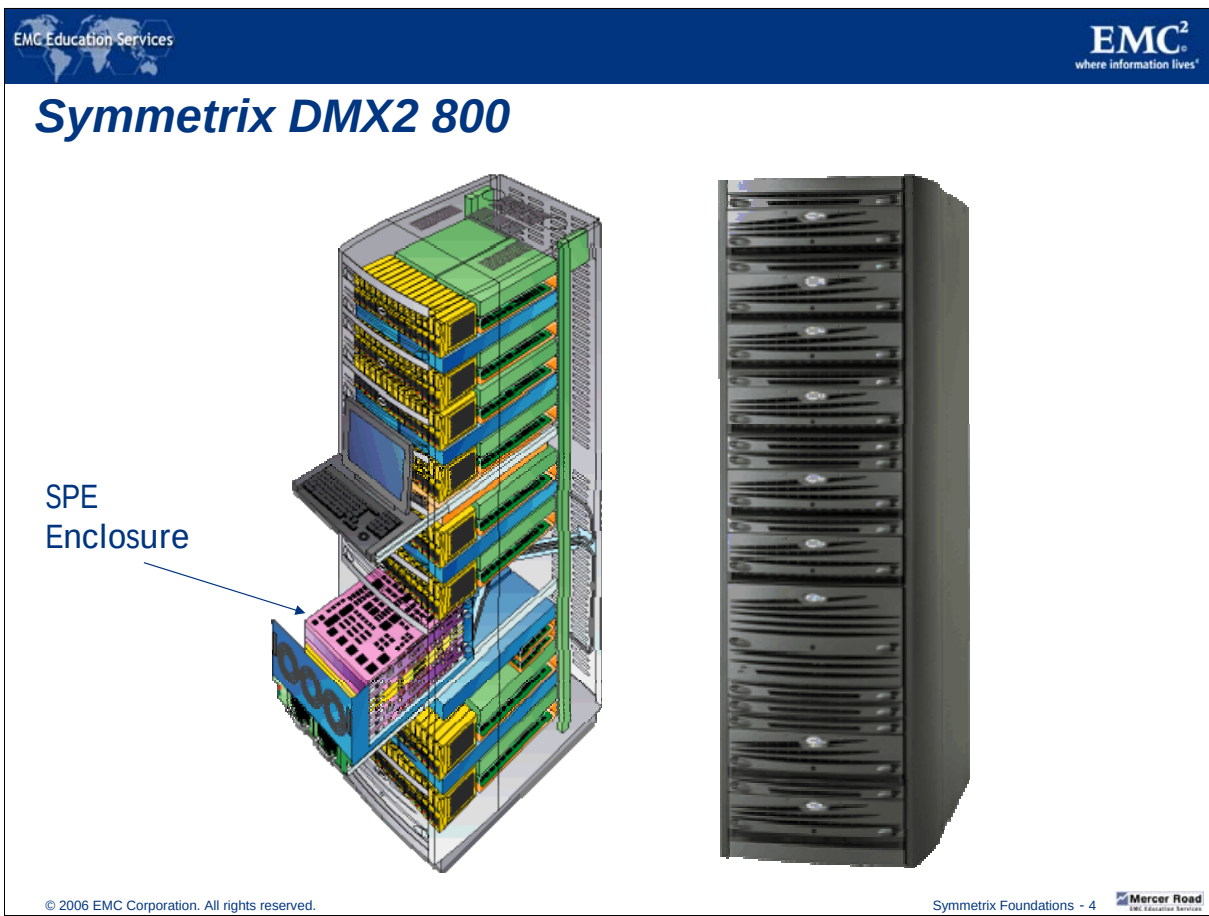
DMX2000



DMX3000



Symmetrix Direct Matrix (DMX) Architecture is storage array technology that employs a matrix of dedicated, serial point-to-point connections instead of traditional buses or switches. The Symmetrix DMX2 is a channel director specification for the DMX with faster processors and newer components. Symmetrix DMX800 is an incrementally scalable, high-end storage array which features modular disk array enclosures.



The physical layout of the DMX800 is very different than previous Symmetrix models. Directors, Memory, back adapter functionality, communications, and environmental functions are all in the Storage Processor Enclosure. The Storage Processor Enclosure Contains 2 to 4 Fibre director boards, up to 2 Multi Protocol Boards, 2 Memory boards, 2 Front-end Back-end adapters, Redundant Power Supplies, and Fan module.

The DMX800 does not contain disk drive cages; drives are in separate Disk Array Enclosures (DAEs). There are fifteen disks per each enclosure and a maximum of eight Disk Array Enclosures per frame which provide a maximum of 120 disks. Each Disk Array Enclosure has 2 Link Controller Cards (LCCs) and 2 Power Supplies.

The Service Processor is replaced by a 1U (1U = 1.75") Server. Batteries, or Standby Power Supplies (SPS), are in a separate 1U enclosure. Each Standby Power Supplies enclosure contains two Standby Power Supplies, and supports either two Disk Array Enclosures or one Storage Processor Enclosure. The Communication and Environmental functions are taken care of by Directors and Front-end Back-end Adapters.

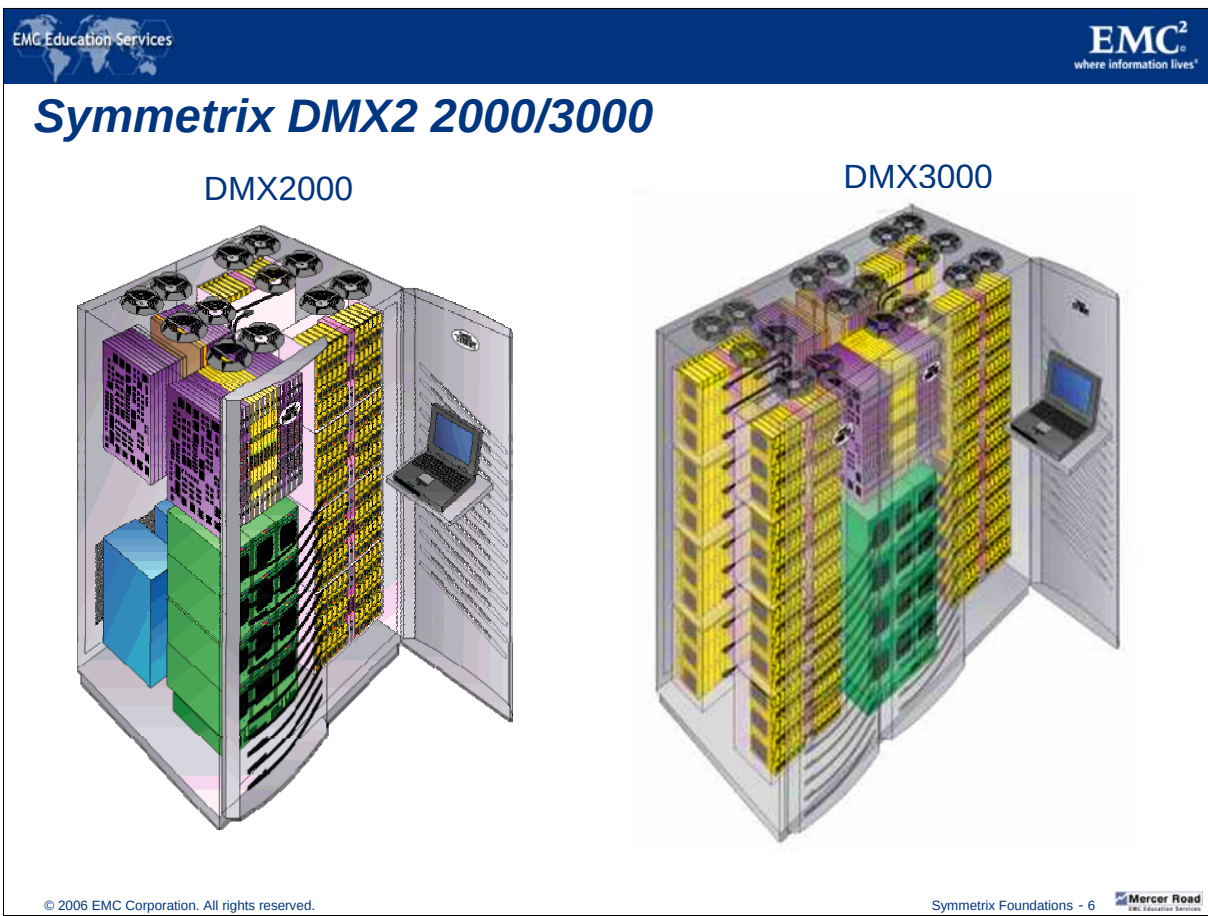
Symmetrix DMX2 1000



© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 5  Mercer Road
EMC Education Services

The DMX1000 system has a 12-slot midplane. Four slots in the center are reserved for global memory directors and the remaining eight slots are reserved for channel directors and disk directors. The Symmetrix DMX1000 can support a maximum of 144 disks. This single-bay system contains one power zone that can be populated with two-to-four redundant single phase power supply modules and two power line input modules.

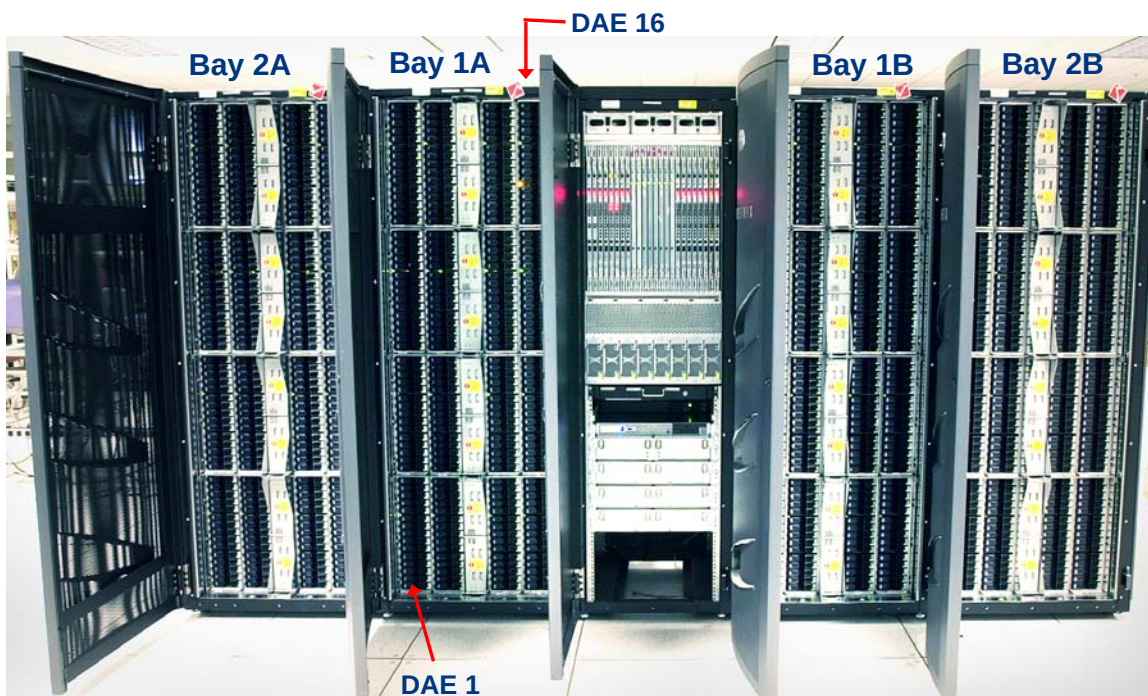


The Symmetrix DMX2000 and DMX3000 systems have a 24-slot midplane. On the front side, the eight slots in the center of the midplane are reserved for global memory directors and the remaining 16 slots are reserved for channel directors and disk directors. The DMX2000 can support 288 disks which are located in a disk bay while the directors and power are located in another bay. The DMX3000 has the same basic layout as the DMX2000 with an additional disk bay to accommodate a maximum of 576 disks.

The Symmetrix 2000/3000 systems have two power zones that provide redundancy in the event of a power loss and can house up to a maximum of 12 power supply modules providing 2(N+1) redundancy. The Symmetrix DMX2000/3000 systems also have two power line input modules.

The DMX2000 supports single phase and three phase power configurations while the DMX3000 supports only three phase power.

Symmetrix DMX3 Front View

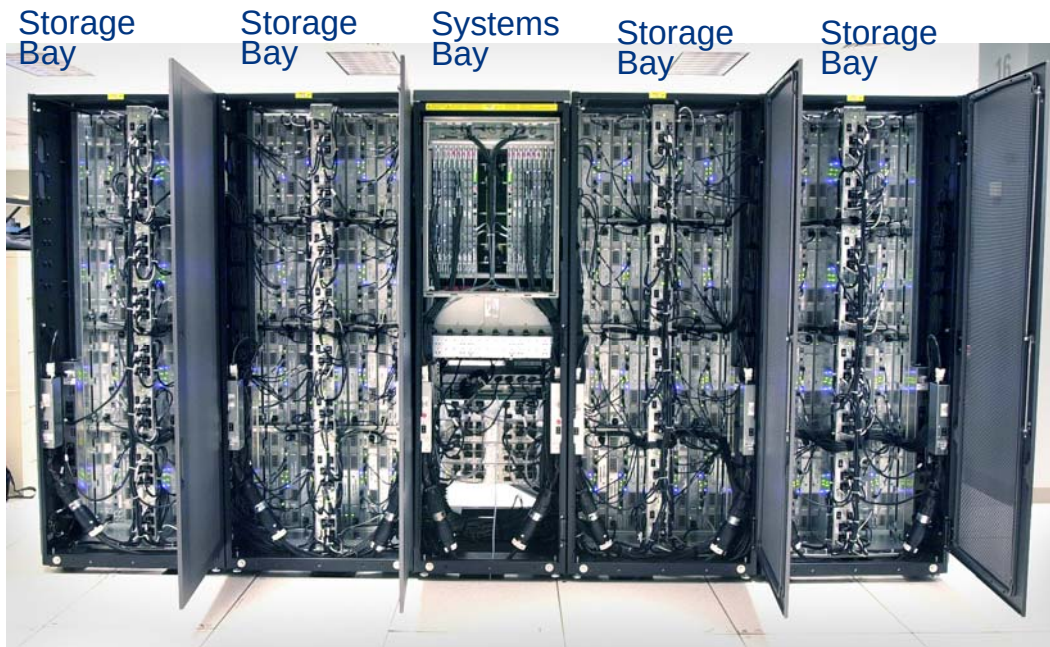


© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 7 

This is the front view of a DMX3 model DMX4500 with 1 System Bay and 4 Storage Bays. DMX3 will support configurations of up to 2,400 drives, and 160 Disk Array Enclosures. Each Storage Bay may contain up to 16 Disk Array Enclosures with up to 15 drives each for a total of 240 drives per Storage Bay, and 10 storage bays possibility for 2,400 drives in total. This configuration would involve daisy chaining Disk Array Enclosures. Each bay has two power zones which provide 2N redundancy. Power Zones A and B are provided and the bay can sustain itself on one power zone, A or B.

Symmetrix DMX3 model DMX 4500 Rear View



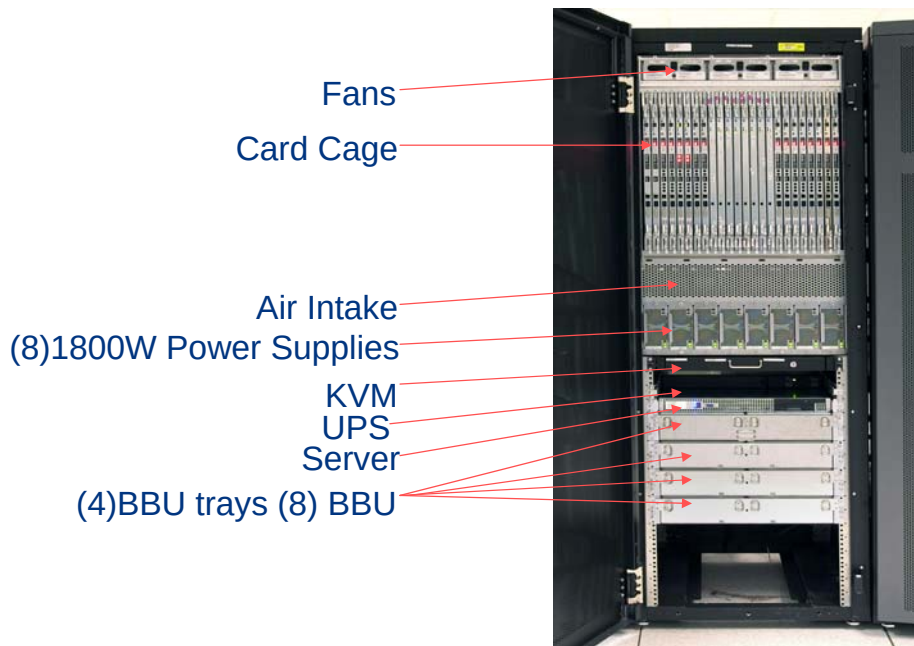
© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 8

Mercer Road
EMC Education Services

A fully configured system consists of one (1) System Bay and ten (10) Storage Bays for a total of 2,400 drives (shown in the example above is 1 System bay and 4 Storage Bays). The Storage Bay will be completely cabled at the factory and the only cabling needed at installation will be to the disk adapters in the System Bay or, in the case of a daisy chain bay, disk array enclosure link control card expansion ports in the storage bay.

Symmetrix DMX3 System Bay (Front)



© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 9

Mercer Road
EMC Education Services

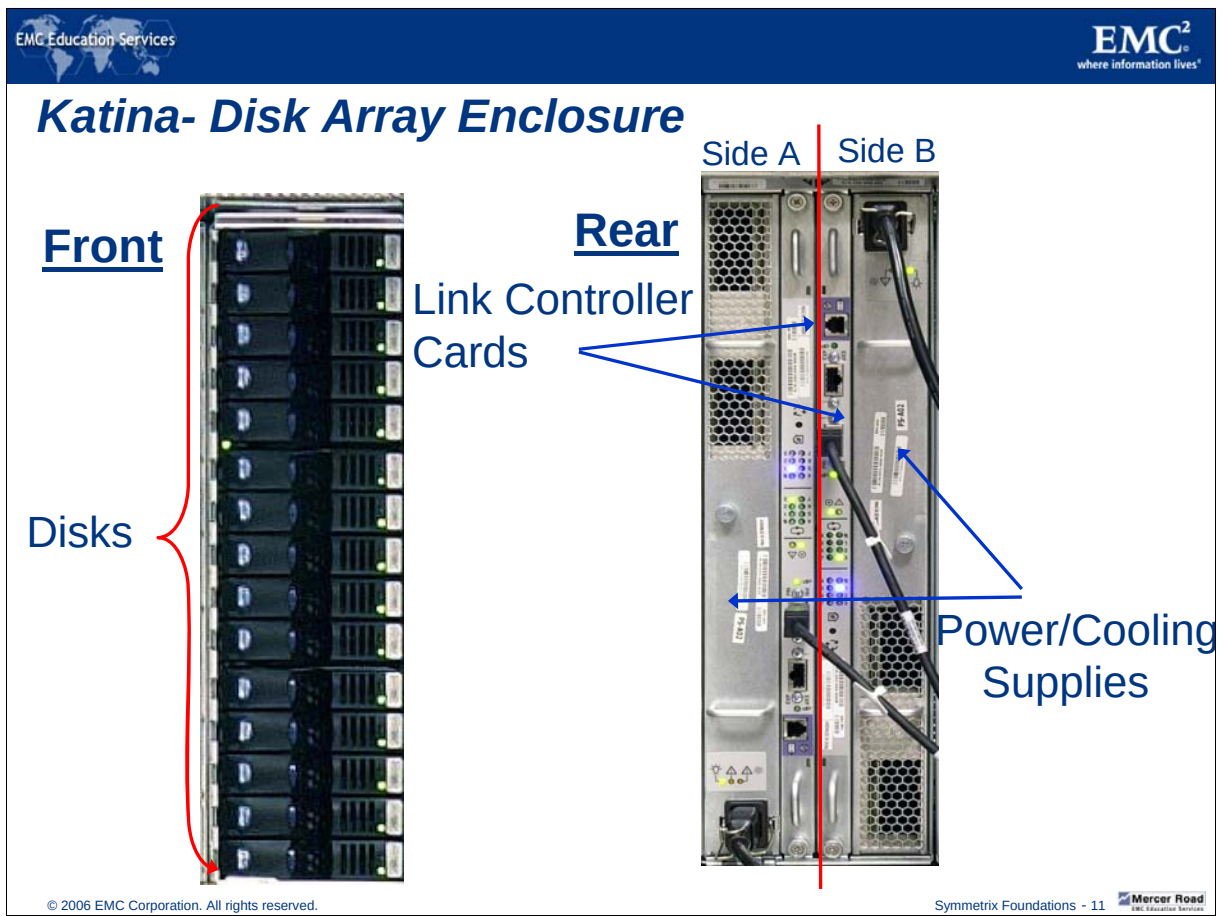
This slide shows the front view of the DMX3 System Bay. The System Bay contains a Keyboard Video Mouse, a 1U server with UPS, 3 cooling fan assemblies, 24-slot card cage, and up to 8 power supplies, each of which is connected to dedicated battery backup units. Air intake pulls air in from the front of the bay and vents it out the top.

Symmetrix DMX3 Unified Director Features

- The Unified director can hold different emulations (depending on the mezzanine cards) and therefore can be configured to support various interfaces
 - Escon (EA): Mainframe interface
 - Ficon (EF): Enhanced ESCON mainframe interface
 - Fibre(FA/DA): Open System host interface or Fibre Disk Adapter
 - GigE (RE): Multi-mode SRDF connection
 - iSCSI (SE): Multi-mode host connection

Note: There are 4 Mezzanine cards per Unified Director

The protocol depends on the mezzanine cards placed on the unified director. The mezzanine cards determine the functionality of each slice. Each board has 4 slices; previous architectures identified these as processors a through d.



Pictured above is a rear view of a disk array enclosure with two Link Control Cards and two Cooling/Power Supplies. Both Power Supply A and Link Control Card A are located on the left, and Power Supply B and Link Control Card B are on the right. Link Control Cards and cables attach DAEs to disk directors or other DAEs in the event of daisy chaining.

Symmetrix DMX Series Integrity Features

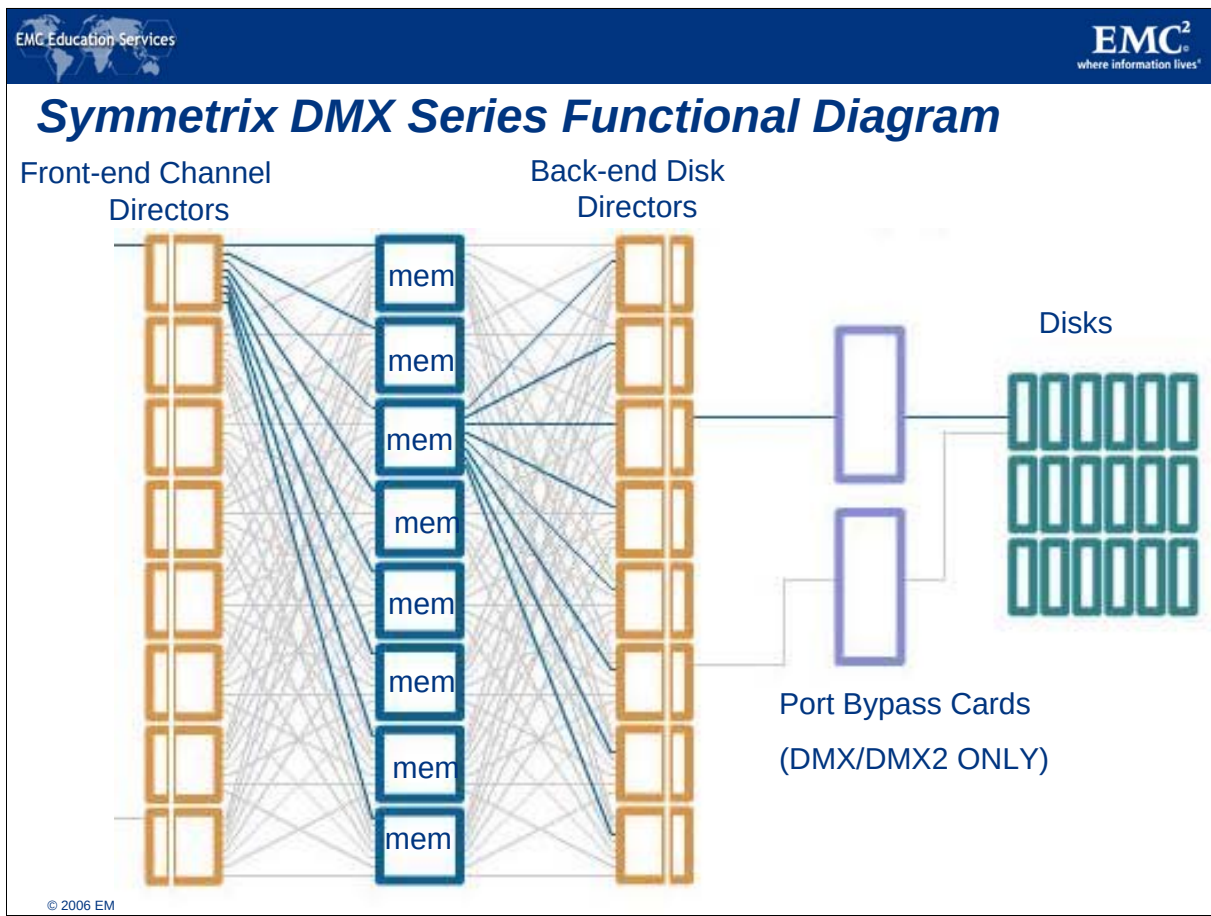
- Error checking, correction, and data integrity protection
- Global memory access path protection
- Global memory error correction and error verification
- Periodic system checks
- Remote support

Error verification prevents temporary errors from accumulating and resulting in permanent data loss. Symmetrix also evaluates the error verification frequency as a signal of a potentially failing component. The periodic system check tests all components as well as Engenuity integrity.

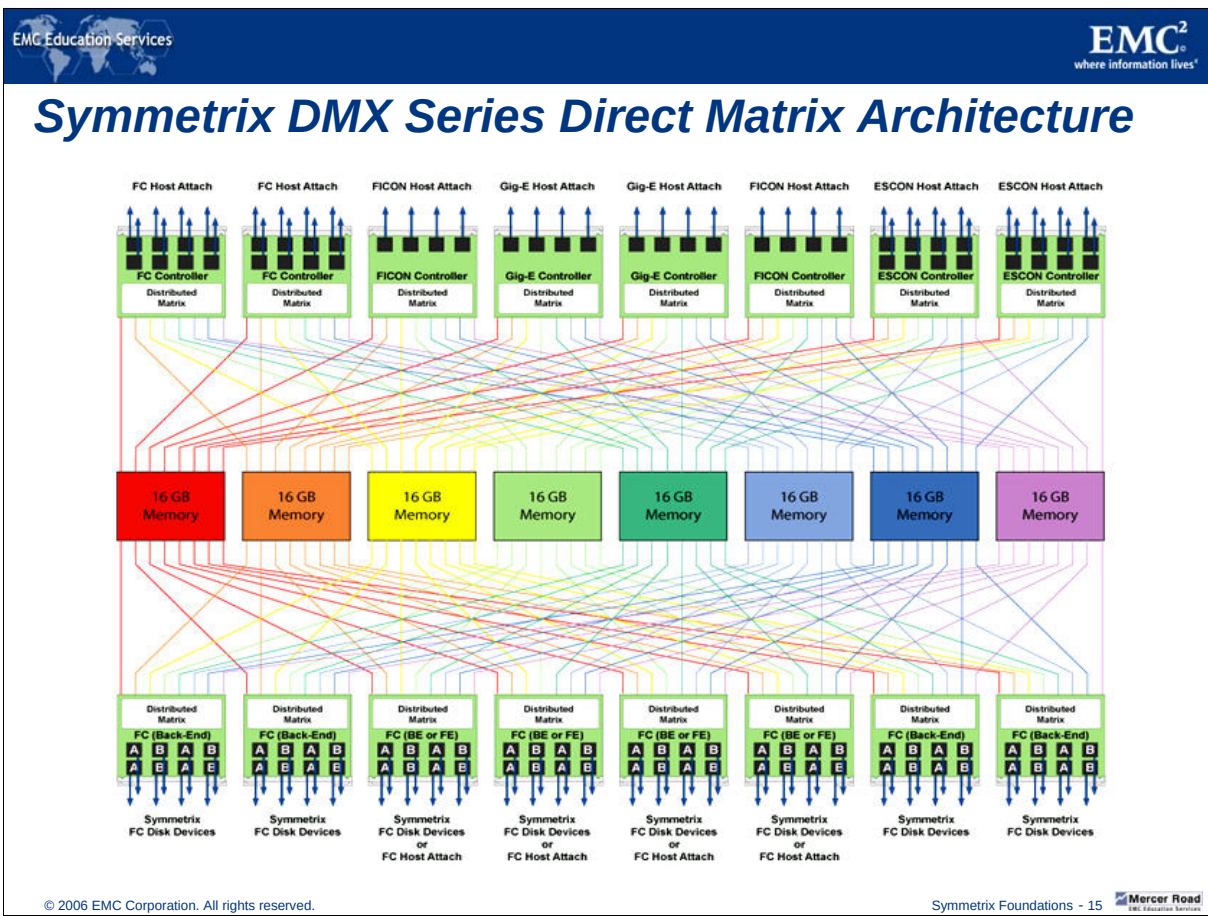
Symmetrix Foundations

Symmetrix Building Blocks and Architecture

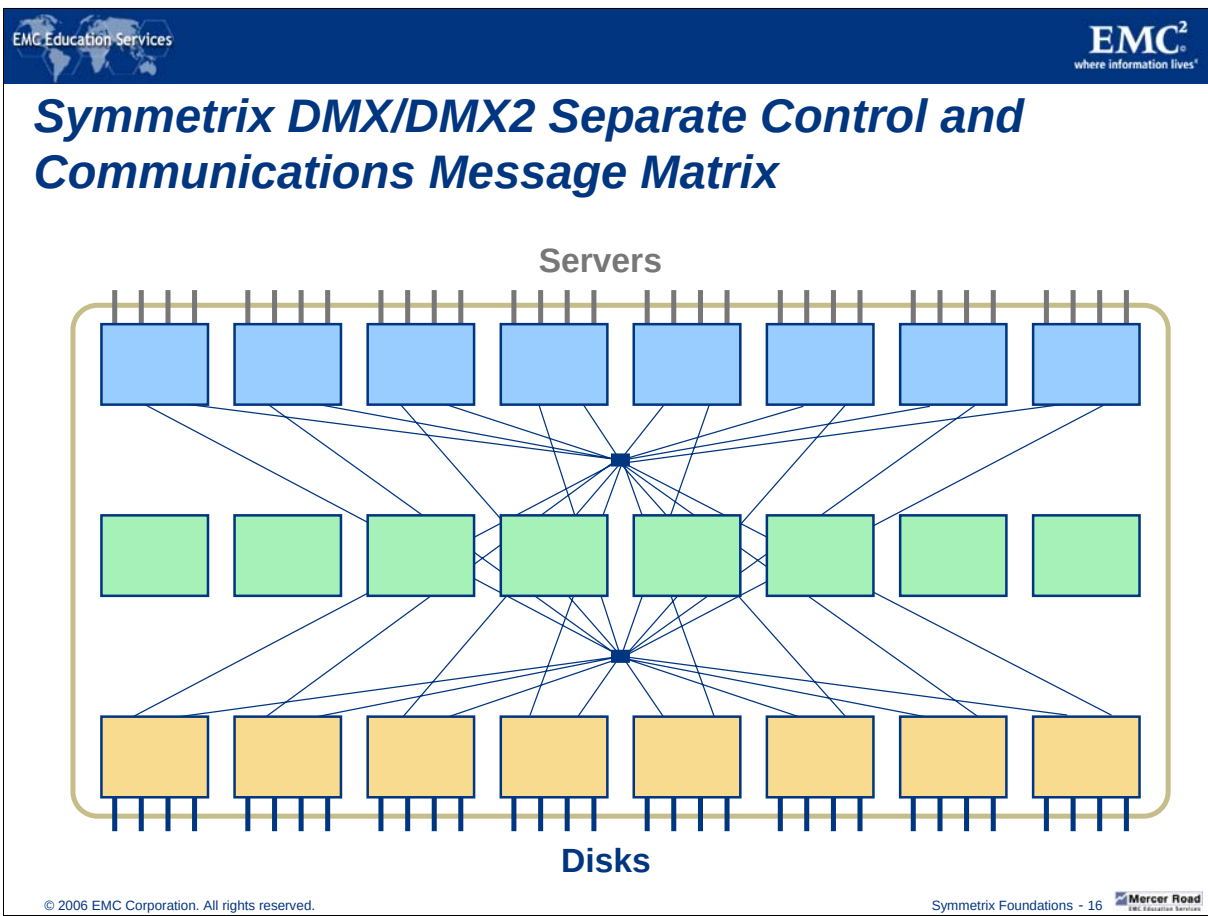
Next, we will discuss Symmetrix building blocks and architecture.



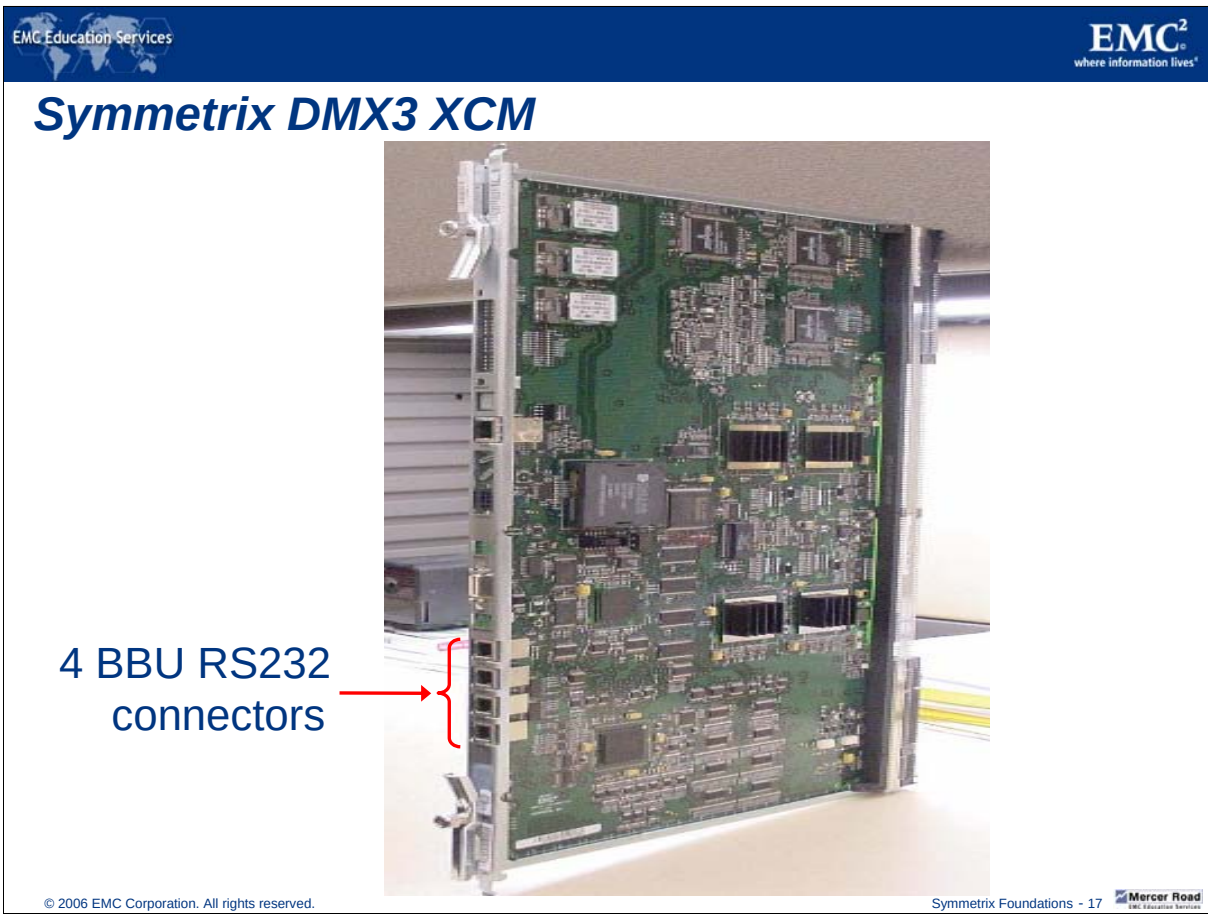
The DMX Series models' functional block diagram displays hosts connected to the back adapters of the front-end directors; they send their data to the Symmetrix DMX Series system's cache. The Point-to-Point matrix connection between cache and back-end disk directors allows for high-performance destaging to the drives, or retrieval of data from the disks into cache. Port Bypass cards are used by the Symmetrix DMX/DMX2 only.



Enhanced global memory technology supports multiple regions and sixteen connections on each global memory director, one to each director. The matrix midplane provides configuration flexibility through slot configuration. Each director slot port is hard-wired point-to-point to one port on each global memory director board. If a director is removed from a system, the usable bandwidth is not reduced. If a memory board is removed, the usable bandwidth is dropped.



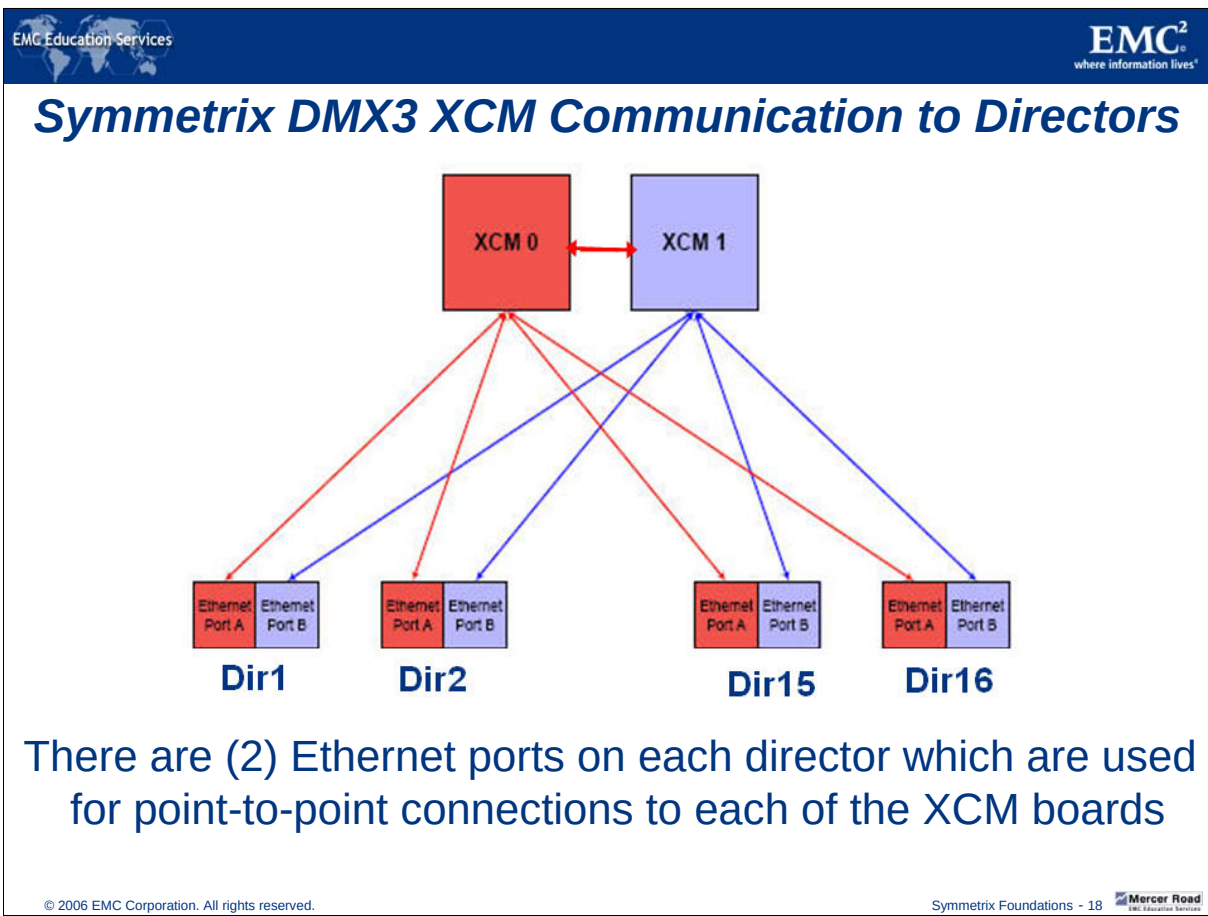
In the Symmetrix DMX/DMX2 Direct Matrix Architecture, contention is minimized because control information and commands are transferred across a separate and dedicated message matrix that enables communication between the directors, without consuming cache bandwidth.



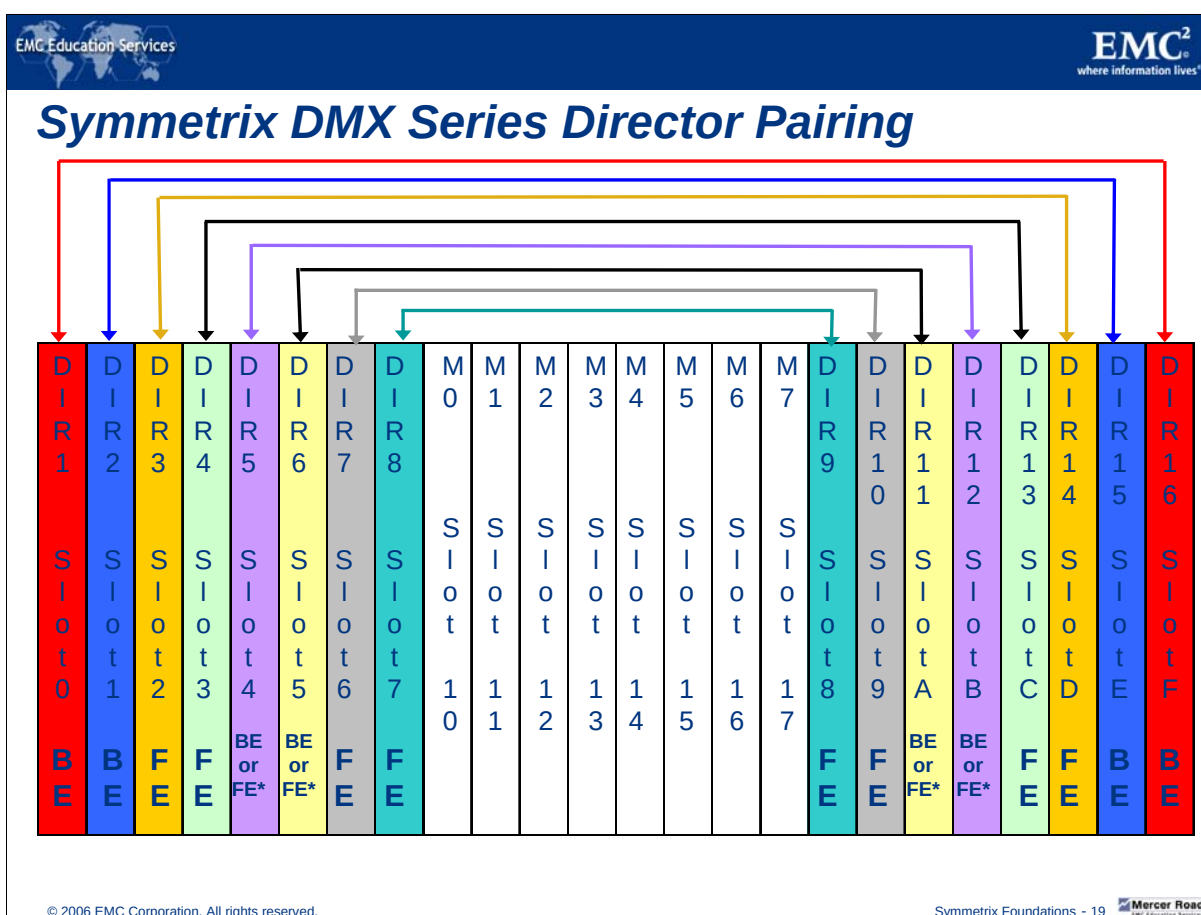
The XCM combines communication and environmental board capabilities. The XCM acts as a messaging fabric switch between the 16 Directors, and monitors environmental as well as logs errors. Four battery backup unit RS232 connectors provide paths between battery backup units and the XCM to send commands and receive status. The DMX3 has two XCM cards for redundancy and can run with one.

XCM handles the following four functions:

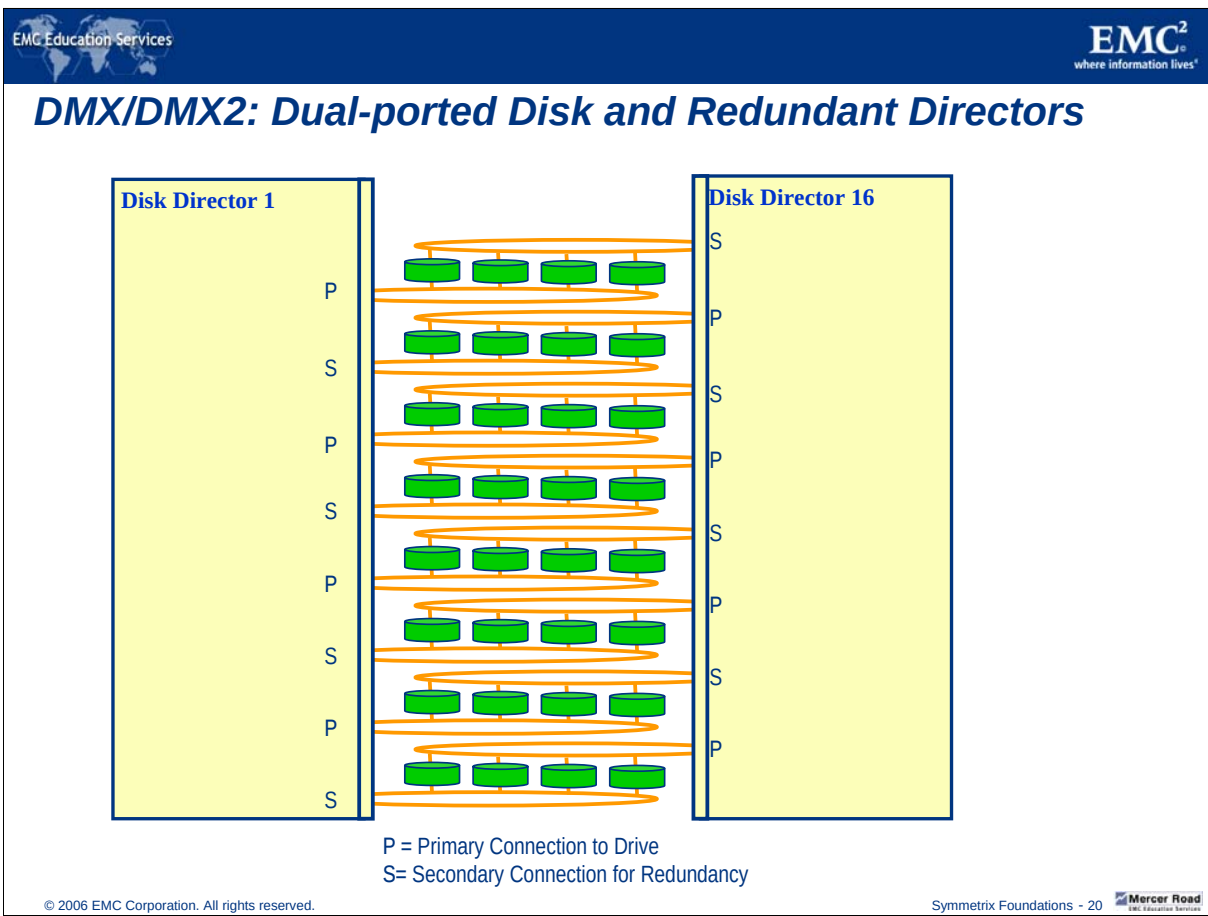
- Ethernet interface between Directors and Service Processor
- Messaging fabric switch between the 16 Directors
- Monitors environmental and log errors
- Sends and receives commands to SPS using the 4 RS232 SPS Connectors



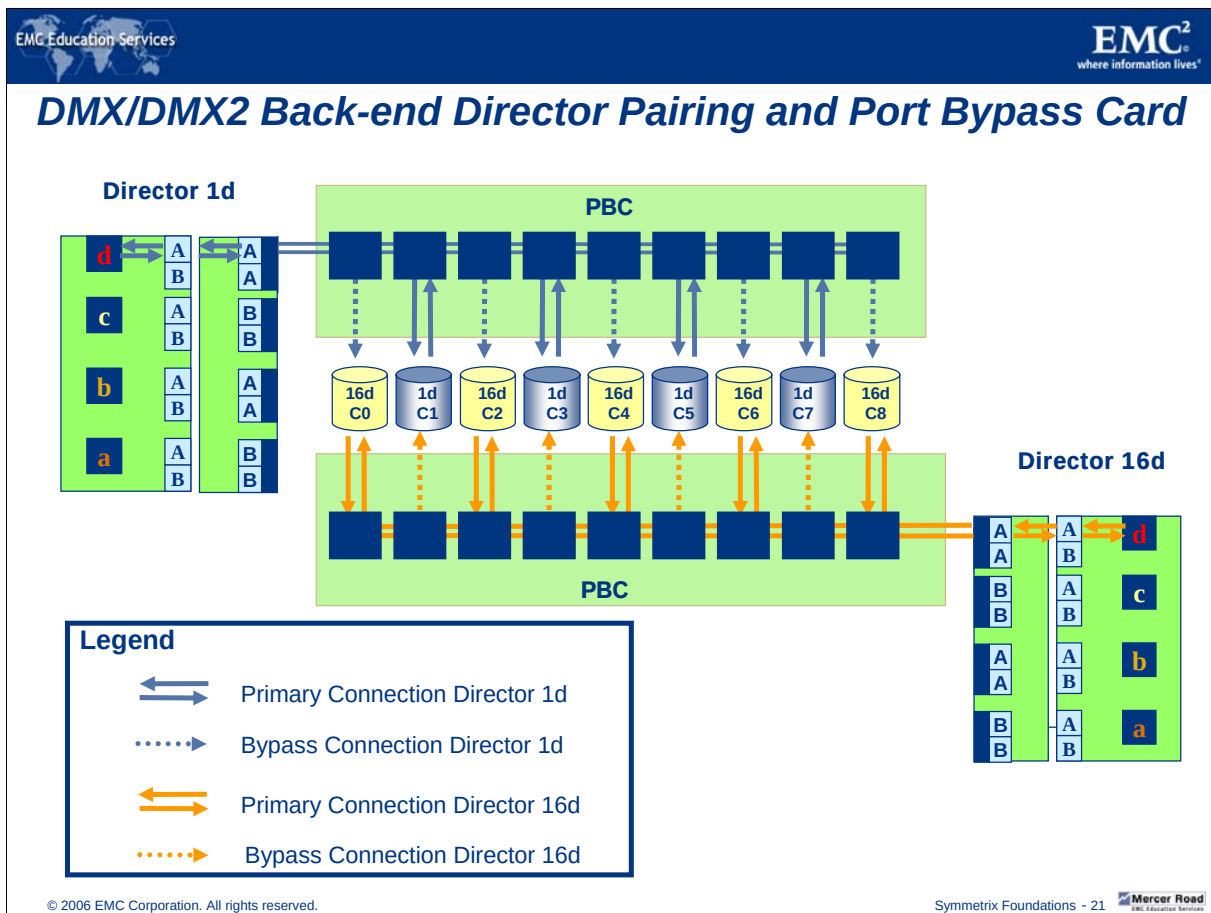
The DMX3 has two XCM cards for redundancy and can run with one. There is also an Ethernet connection between the XCM boards. These connections eliminate a single point of failure. XCM 0 and XCM 1 have connectivity to all components to eliminate single points of failure with the exception of memory boards and standby-power-supplies. XCM 0 has connectivity to memory positions 0 through 3 and all A standby-power-supplies; XCM 1 has connectivity to memory positions 4 through 7 and all B standby-power-supplies in the System Bay. In the event that either XCM must be replaced, the monitoring of memory boards and standby-power-supplies from that XCM is suspended until a replacement has been completed.



In the Symmetrix DMX series, Director pairing along with dual ported drives, provides redundancy for a disk director or drive path failure. Disk director pairing starts from the outside and works toward the center of the card cage, directors are paired processor-to-processor using the rule of 17. Notice in the diagram above, directors 1 and 16 are paired and directors 2 and 15 are paired. Front-end director pairing configuration is recommended, but not required. Specific director slots can be used for a front-end or back-end director giving the customer flexibility for enhanced back-end performance or additional connectivity.

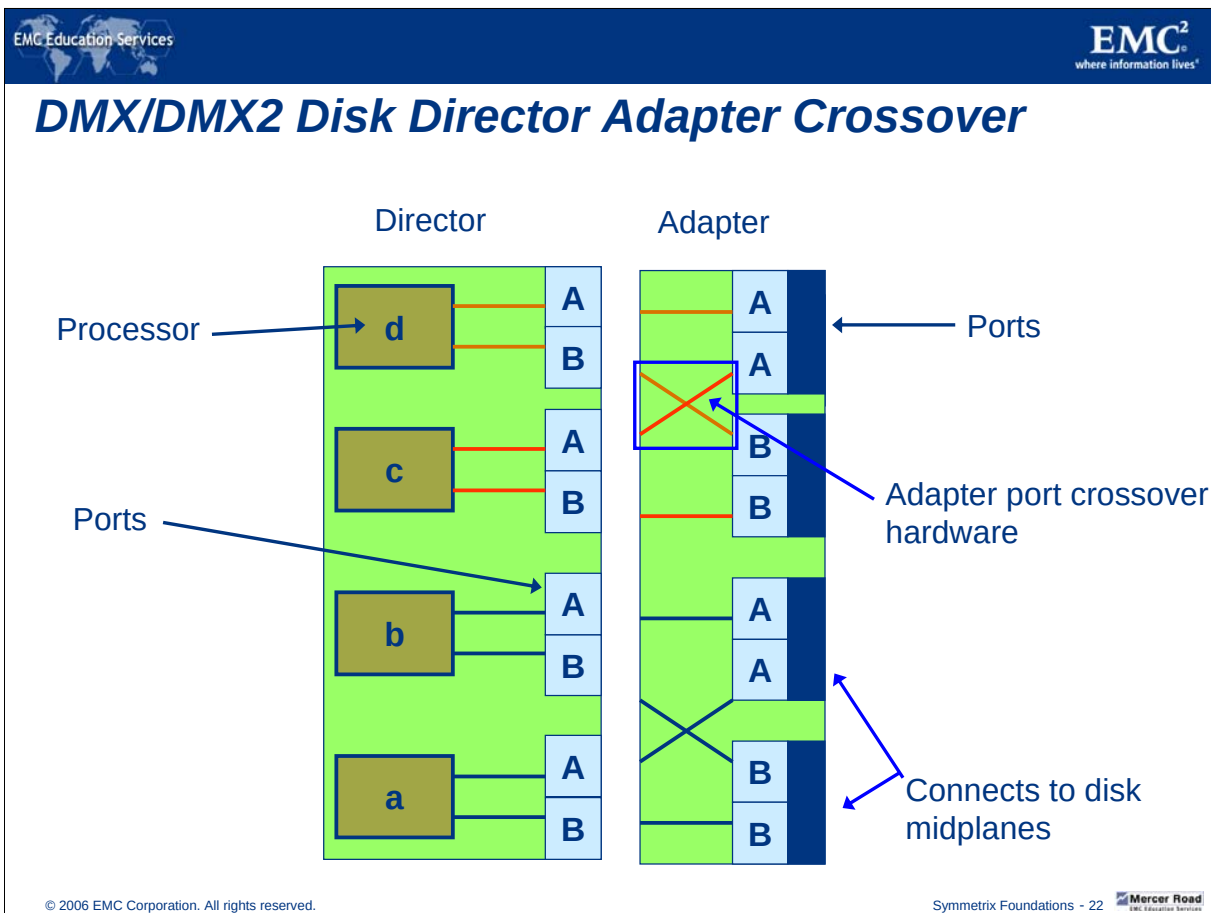


Symmetrix DMX/DMX2 back-end employs an arbitrated loop design and dual-ported disk drives. Here is an example of a 9 disk per loop configuration with 4 disks per loop. Each drive connects to two paired Disk Directors through separate Fibre Channel loops. Port Bypass Cards prevent a Director failure or replacement from affecting the other drives on the loop. Directors have four primary loops for normal drive communication and four secondary loops to provide alternate path if the other director fails.



The Port Bypass Card contains the switch elements and control functions to allow intelligent management of the two FC-AL loops embedded in each disk cage midplane. There are two Port Bypass Cards per disk cage midplane. Each disk cage midplane can support 36 Fibre Channel drives. Each Processor has two ports, each with devices in the Front, as well as in the Back, Disk Midplane.

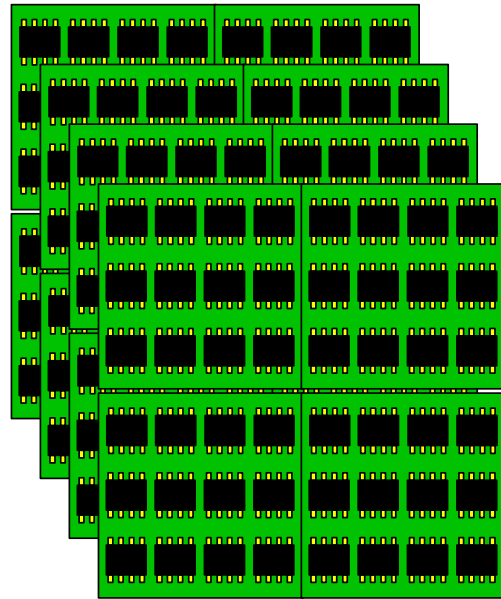
In the above slide, we are showing only one port from Director 1d, and one port from Director 16d. Notice that each director has the potential to access all drives in the loop (9-drive loop configuration in this example). Also notice that using the Port Bypass Card, each director is currently accessing only a portion of the drives (Director 1d has 4 drives; Director 16d has 5 drives). These directors will have an opposite configuration on their second port, which is connected to a different Port Bypass Card and Disk Midplane. With no component failure, each processor will manage 4 drives on one port and 5 drives on the other. These reside in Front and Back Disk Midplanes and are referred to as C and D Devices. If the processor on Director 1d fails, the processor on Director 16d will now access all 9 drives on this loop.



In order for each processor to access disks in the Front Disk Midplane and Back Disk Midplane, it is clear that each processor needs a physical path to two separate Disk Midplanes via two cables. The back adapter crossover feature will allow d processor port A and c processor port A to access the same Disk Midplane. All A ports from processors a, b, c and d access a Front Disk Midplane, all B ports from processors a, b, c and d access a Back Disk Midplane.

Symmetrix DMX/DMX2 Global Cache Directors

- Memory boards are now referred to as **Global Cache Directors** and contain global shared memory
- Boards are comprised of memory chips and divided into four addressable regions
- Symmetrix has a minimum of 2 memory boards and a maximum of 8. Generally installed in pairs
- Individual cache directors are available in 2 GB, 4 GB, 8 GB, 16 GB 32GB and 64 GB sizes
- Memory boards are Field Replaceable Units and “hot swappable”



© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 23  Mercer Road
EMC Education Services

DMX uses direct connections between directors and cache. When configuring cache for the Symmetrix DMX systems, five guidelines should be followed.

1. A minimum of four and a maximum of eight cache director boards are required for the DMX2000 and DMX3000 system configuration; a minimum of two and a maximum of four cache director boards are required for the DMX1000 system configuration.
2. Two-board cache director configurations require boards of equal size.
3. Cache directors can be added one at a time to configurations of two boards and greater.
4. A maximum of two different cache director sizes are supported, and the smallest cache director must be at least one-half the size of the largest cache director.
5. In cache director configurations with more than two boards, no more than one half of the boards can be smaller than the largest cache director.

Symmetrix DMX3 Redundant Global Memory

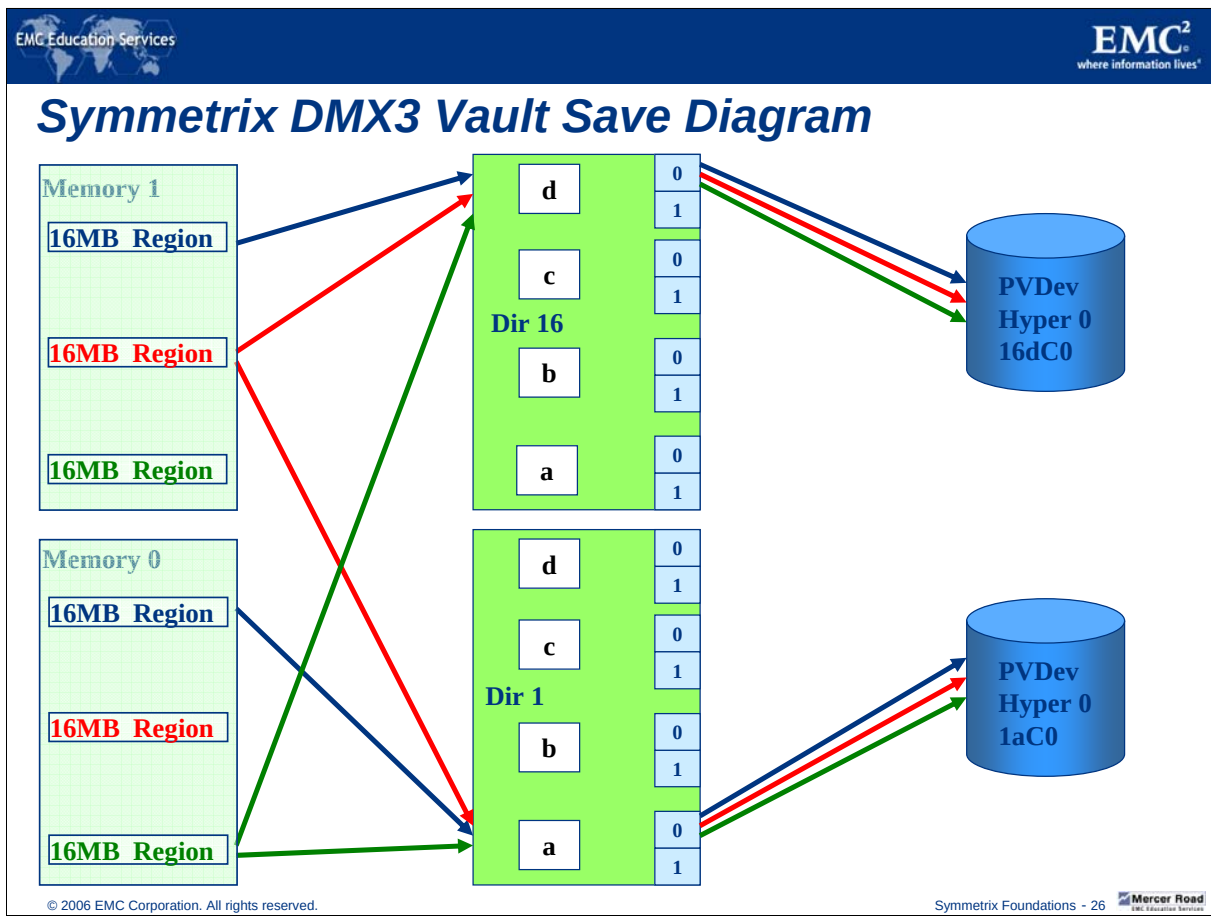
- Data written to Primary region then to Secondary region of the memory board pair
- All reads are from Primary region
- Algorithms in Enginuity will enable the Directors to take full advantage of all memory cards present when reading and writing
- Upon Primary or Secondary region board failure, all directors drop the failed board, and switch to non-dual write mode to the good board of the failed memory pair
- Striping between memory boards is default

Global Memory board pairs reside next to each other and memory is fully redundant. All writes are initially done to the primary region; writes are then carried out to the secondary region. Primary and secondary regions are distributed across all memory boards so, for example, memory board pairs in slots 0 and 1 will have alternating Primary and Secondary regions. Algorithms in Enginuity will enable the Directors to take full advantage of all memory cards present when reading and writing. Any failure condition of a memory board causes all directors to drop the failed board and switch to a normal write mode to the surviving board.

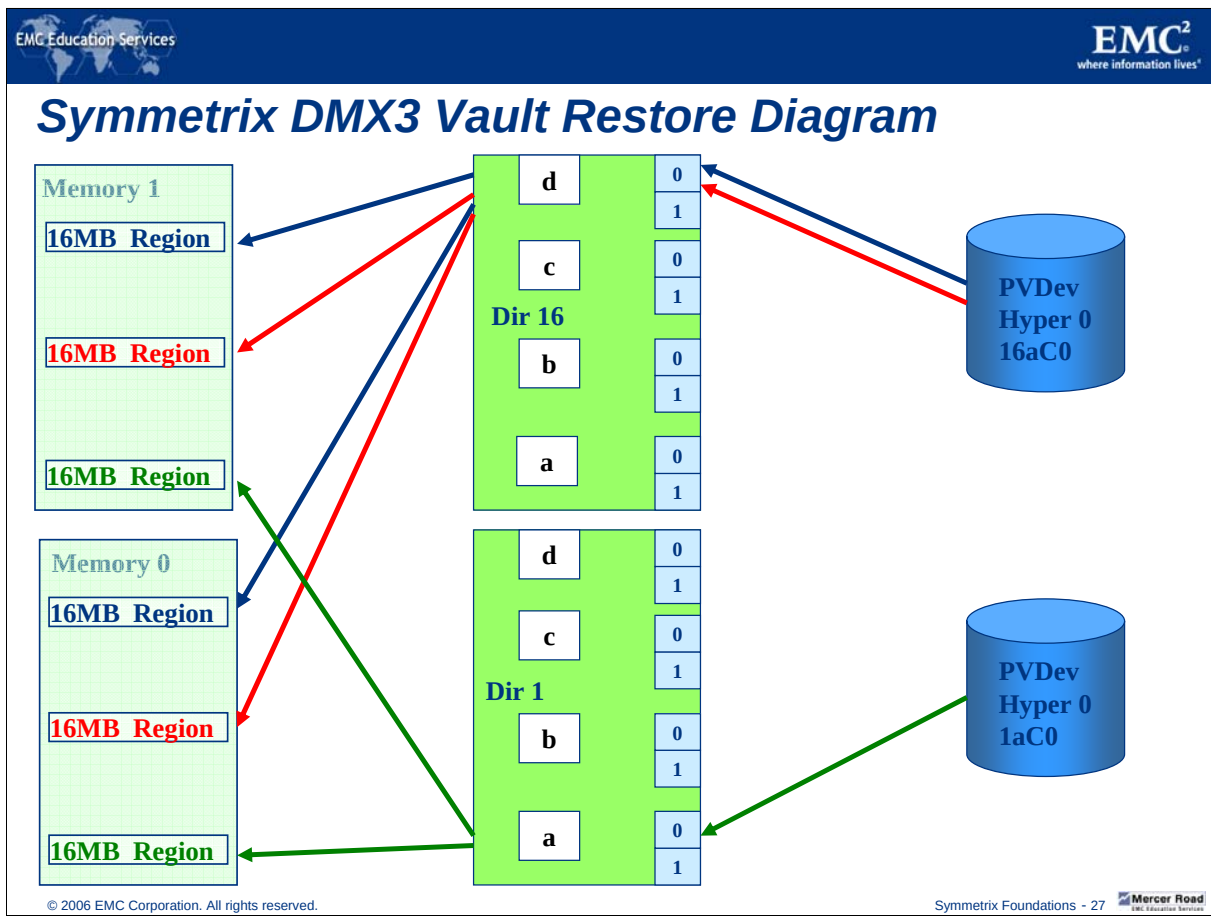
Symmetrix DMX3 Vaulting Overview

- The Vault Image saved is fully redundant (not mirrored)
- Data is written to mirrored memory boards in DMX3
 - Vault Save will write 2 copies of Global Memory to the Vault Disks
 - During the Vault Save, any region may be read from primary and secondary or twice from the same memory board
 - It's important to note that Vaulting & Mirrored Memory are independent features. The same region of memory may be saved twice because of contention or other issues
 - The 2 copies of mirrored memory will be restored from one copy of Vault

Data vaulting is a new feature only available with DMX3. As cache size, disk size and power requirements increase, the time required to destage data increases. Power vault was designed to limit the time necessary to power off the box on battery power. Power Vault will save global memory to specific vault devices on power down, then, on power up, the data will be loaded to cache so that it may be destaged to the correct location.



The vault image is fully redundant; contents of global memory above the vault line will be saved twice to independent disks. Vault Save will save each region on two separate power vault devices so there are two copies of global memory above the vault line saved. The power vault device memory region pair will attempt to be on opposite disk directors and power zones as above. Sixteen unprotected power vault devices will be automatically configured by SymmWin per each disk director.



A Vault Restore only occurs during a full IML of the system. During the IML in step 14, the entire vault image will be restored to global memory above the vault line. The restore is performed by the disk directors and restores only one copy of global memory which is written into both primary and secondary regions of memory. The Cyclic Redundancy Check of the vault region (16MB) being restored is calculated and checked against the Cyclic Redundancy Check recorded in the vault save table for the same region. The Cyclic Redundancy Check is also sanity checked with the other disk director's save table from Non-Volatile DRAM.

Field Replaceable Units

- **Symmetrix DMX/DMX2**
 - Port Bypass Card
 - Environmental Control Modules
 - Communication Control Modules
- **Symmetrix DMX3**
 - XCM Control Module
 - Disk Array Enclosure Link Control Card
 - Disk Array Enclosure Power Supply

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

Symmetrix DMX Series Field Replaceable Units

- Channel Director Boards and Disk Director Boards
- Global Memory Director Boards
- Disk Devices
- Power system components, and Batteries
- Service Processor
- Cooling Fan Modules

The Symmetrix DMX system features non-disruptive replacement of its major components.

Symmetrix Foundations

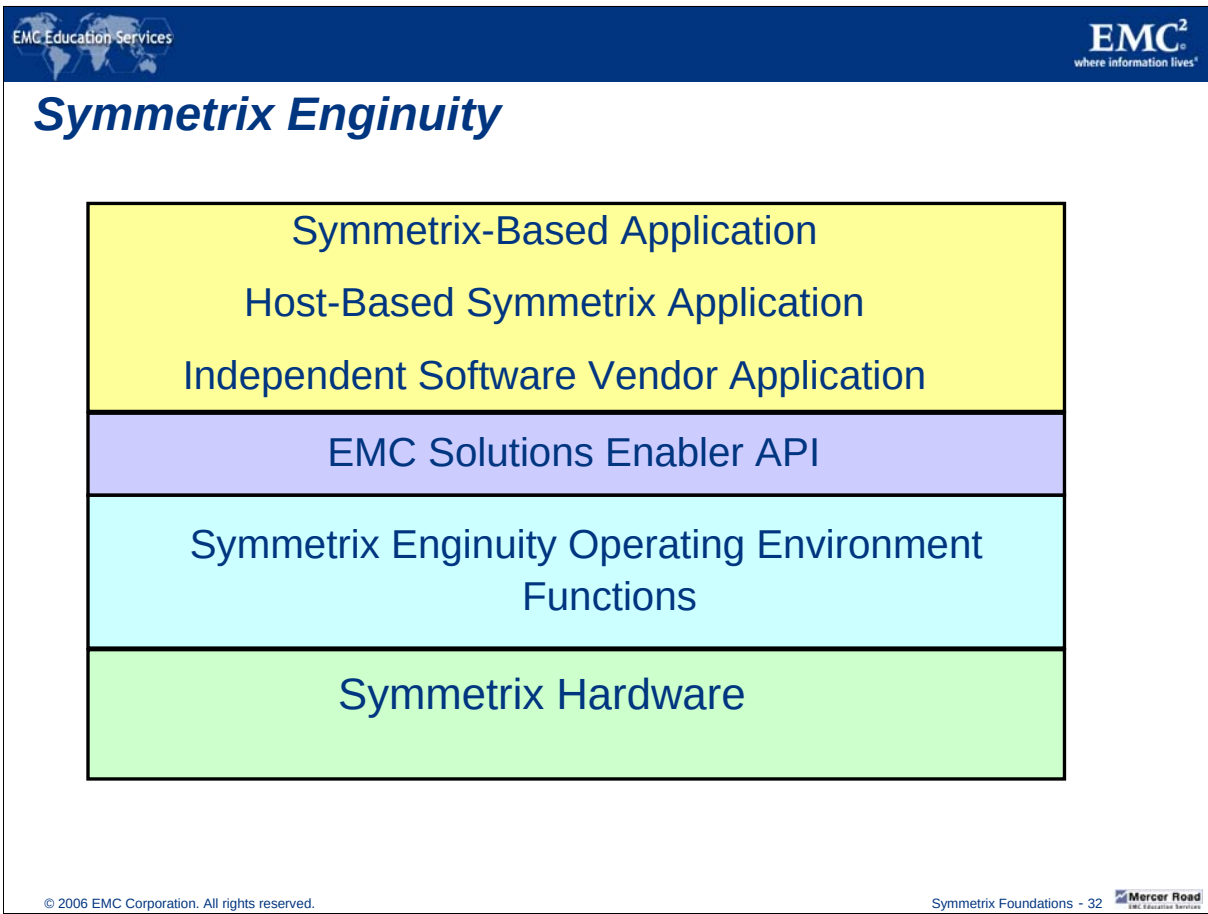
Software Operating Environment

The next three slides discuss the software-operating environment for Symmetrix.

Symmetrix Enginuity Services

- Manage systems resources for I/O requirements
- Symmetrix component fault monitoring and detection
- Defines task priority
- Provides functional services for a suite of EMC storage application software

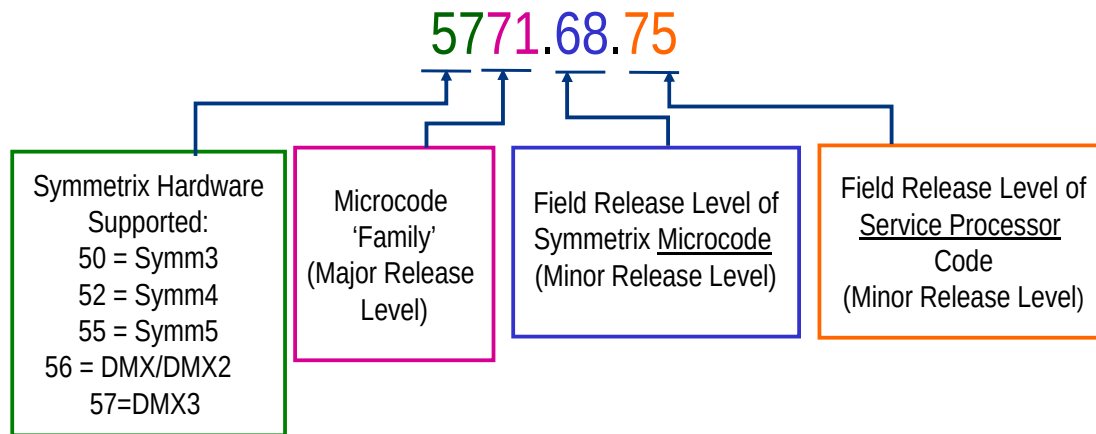
Symmetrix Enginuity is the operating environment for the Symmetrix DMX systems. Enginuity manages all Symmetrix operations from monitoring and optimizing internal data flow, to ensuring the fastest response to the user's requests for information, to protecting and replicating data.



EMC's solution enabler APIs are the storage management programming interfaces that provide an access mechanism for managing the Symmetrix third-party storage, switches, and host storage resources. They enable the creation of storage management applications that don't have to understand the management details of each piece within the total storage environment. Symmetrix DMX systems support platform software applications for data migration, replication, integration and more.

Enginuity Overview

- Operating Environment for Symmetrix
 - Each processor in each director is loaded with Enginuity
 - Enginuity is what allows the independent director processors to act as one Integrated Cached Disk Array
 - Also provides the framework for advanced functionality like SRDF, TimeFinder,...etc.



© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 33 

The numbers that define an Enginuity level have specific meaning. In this example the 57 represents the DMX3 hardware, 71 is the microcode family, 68 is the field release level to the microcode, and 75 is the field release to the service processor code.

Non-disruptive microcode upgrade and load capabilities are currently available for the Symmetrix. Symmetrix takes advantage of a multi-processing and redundant architecture to allow for hot loadability of similar microcode platforms.

The new microcode loads into the EEPROM areas within the channel and disk directors, and remains idle until requested for hot load in control storage. The Symmetrix system does not require manual intervention on the customer's part to perform this function. All channel and disk directors remain in an on-line state to the host processor, thus maintaining application access.

Symmetrix will load executable code at selected "windows of opportunity" within each director hardware resource, until all directors have been loaded. Once the executable code is loaded, internal processing is synchronized and the new code becomes operational.

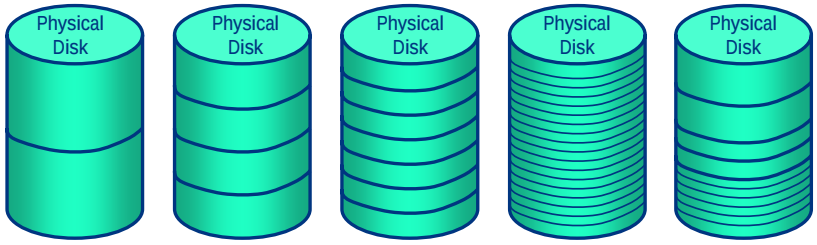
Symmetrix Foundations

Theory of Operation – Symmetrix Volumes

Let us look at the theory of operation for Symmetrix Volumes.



Defining Symmetrix Logical Volumes



Symmetrix Service Processor

Running SymmWin Application

- Symmetrix Logical Volumes are configured using the service processor and SymmWin interface/application
 - Generate configuration file (IMPL.BIN) that is downloaded from the service processor to each director
- Configuration changes can be performed online using the EMC ControlCenter Configuration Manager and Solutions Enabler Command Line Interface

© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 35 

Symmetrix logical volumes are defined by using the service processor and SymmWin interface. A disk is sliced into hypers or disk slices and protection schemes are then incorporated, creating the Symmetrix volume.

The Service Readiness Symmetrix Engenuity Configuration website is used to verify initial Symmetrix configuration and any subsequent changes to the configuration. They use time-honored extensive best practices and tools to configure Symmetrix.

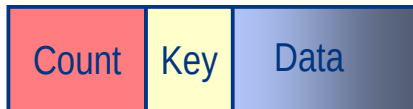
Symmetrix Logical Volume Types

- Open Systems hosts use Fixed Block Architecture (FBA)

- Each block is a fixed size of 512 bytes
- Volume size referred to by the number of Cylinders
- Each Cylinder has 15 tracks
- Each track has 64 blocks of 512bytes

Data Block
512 Bytes

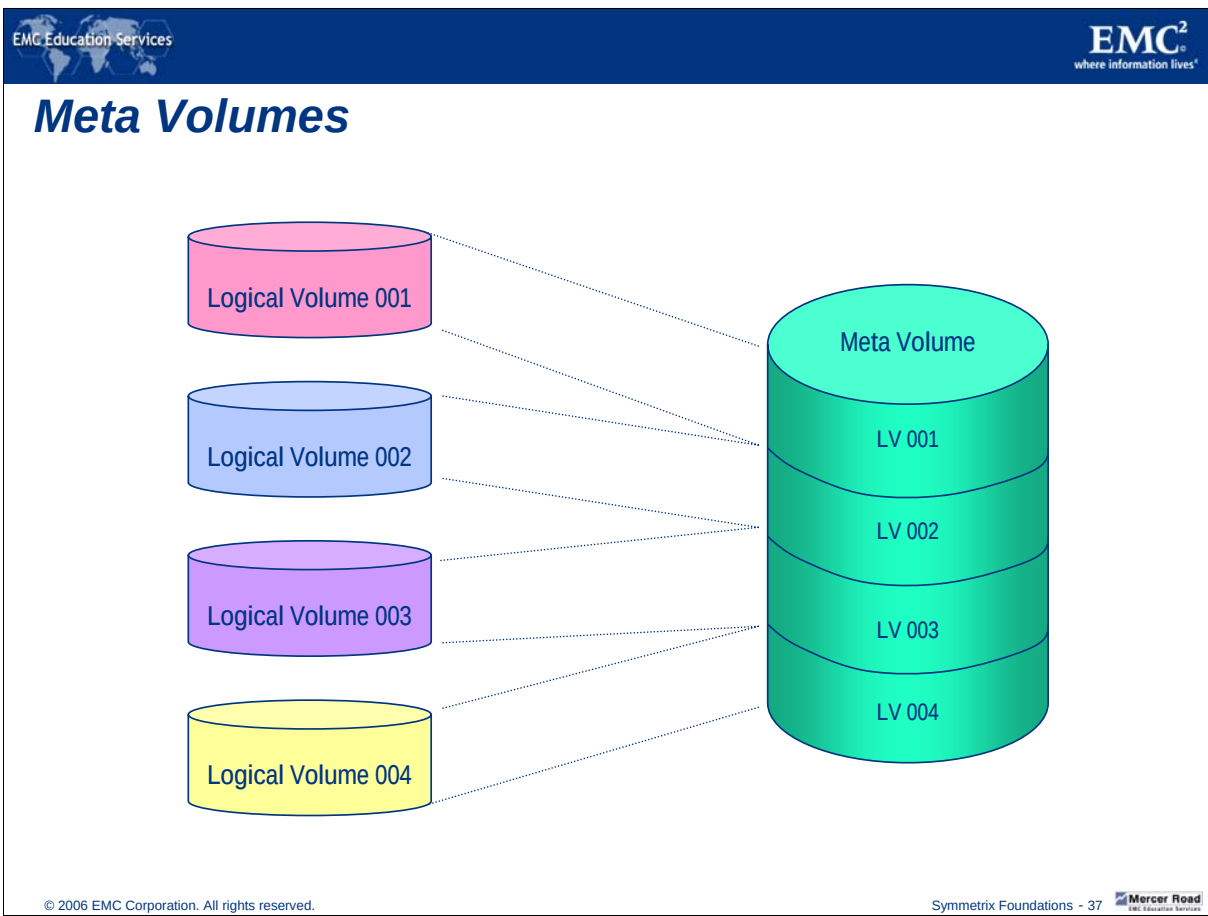
- Mainframes use Count Key Data (CKD)



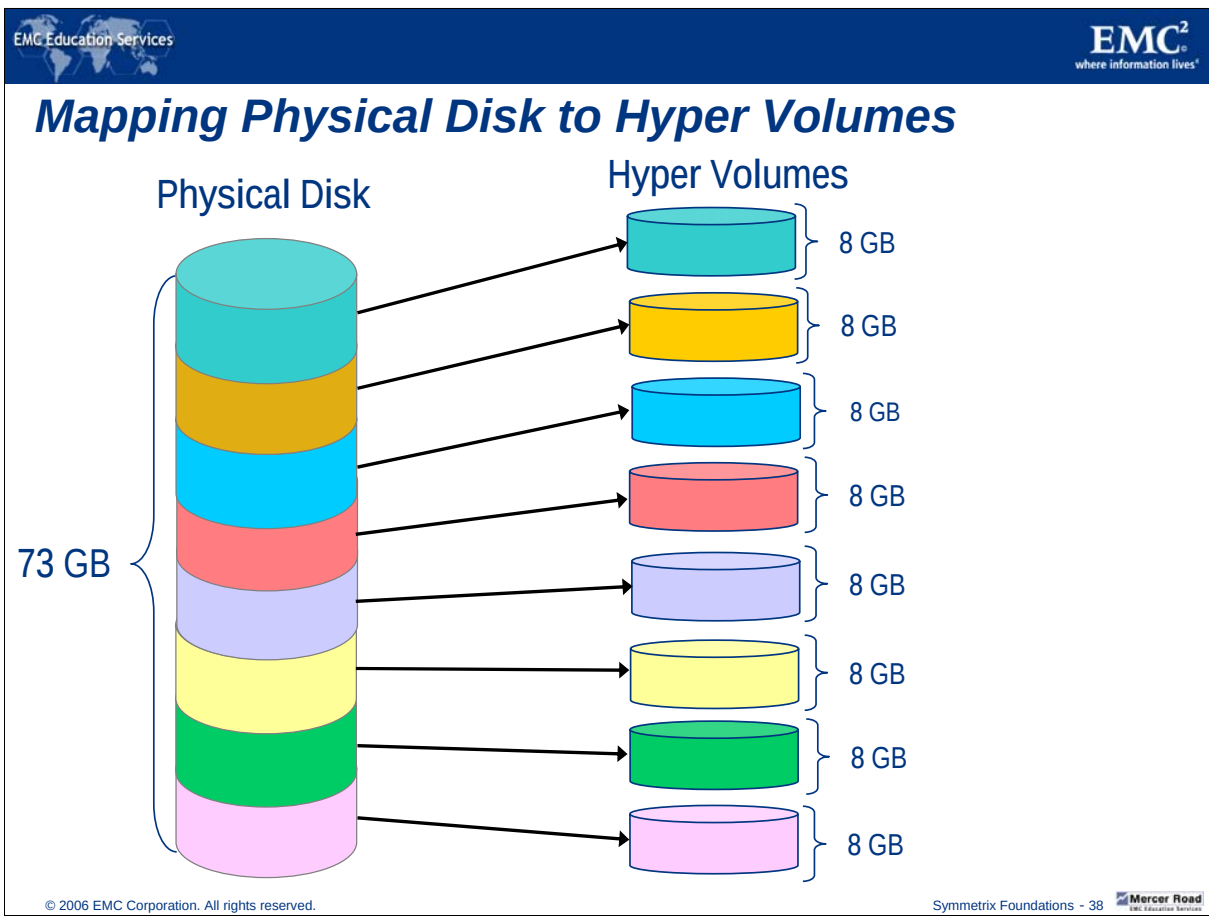
- Count field indicates the data record's physical location (cylinder and head) record number, key length, and data length
- Key field is optional and contains information used by the application
- Data field is the area which contains the user data

- Symmetrix stores data in cache in FBA and CKD and on physical disk in FBA 512 format

Mainframes use Count Key Data format while open systems use Fixed Block Architecture. Symmetrix stores data in cache in both formats and on physical disk in FBA 512 format.



Symmetrix Logical Volumes can be grouped into a Meta Volume configuration and presented to Open System hosts or Mainframes as a single disk. Data is striped or concatenated within open system Meta Volumes and striped only for CKD meta volumes. Meta Volumes allow customers to present larger Symmetrix Logical volumes than the current maximum hyper volume size and satisfies requirements for environments where there is a limited number of host addresses or volume labels available.



Symmetrix physical disk are split into logical hyper volumes. Hyper volumes are then defined as Symmetrix Logical Volumes and internally labeled with hexadecimal identifiers. A Symmetrix logical volume is the disk entity presented to a host via a Symmetrix channel director port. As far as the host is concerned, the Symmetrix Logical volume is a physical drive. Do not confuse Symmetrix Logical Volumes with host-based logical volumes. Symmetrix Logical Volumes are defined by the Symmetrix Configuration while Host-based logical are configured by customers through Logical Volume Manager software. A hyper volume could be used as an unprotected Symmetrix logical volume, a mirror of another hyper volume, a Business Continuance Volume (BCV), a member for Parity RAID, a remote mirror using SRDF, a Disk Reallocation Volume (DRV), and more.

How Symmetrix Logical Volumes Appear to a Host

- Symmetrix Logical Volumes are viewed by the hosts as disk devices
- Host is unaware of protection or other Symmetrix attributes
- Unix hosts access disk through device special files
 - Many hosts use CTD (Controller-Target-Device) format
 - Example `/dev/rdisk/c1t1d2`



- Other UNIX hosts assign logical names to disk devices
 - Example IBM-AIX uses hdisks (`/dev/hdisk2`)
 - NT accesses disk devices through a PHYSICALDRIVE name
 - ❖ Example: `\\.\PHYSICALDRIVE2`

A host views a Symmetrix Logical Volume in the same manner as it sees any other disk device. The host is unaware how the volume is configured in the Symmetrix, its protection scheme, or any other special attributes. Hosts assign disk devices logical device names that vary depending on the operating system.

Symmetrix Foundations

Theory of Operation – Data Protection Methodologies

Next, we will look at the wide range of data protection methodologies available with Symmetrix.

Data Protection

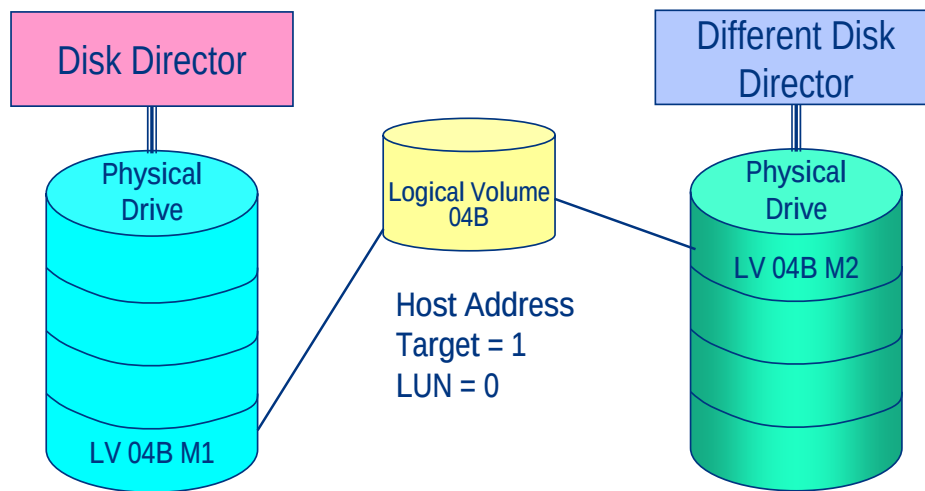
- **Mirroring (RAID 1)**
 - Highest performance, availability and functionality
 - Two hyper mirrors form one Symmetrix Logical Volume located on separate physical drives
- **Parity RAID (Not available on DMX3)**
 - 3 +1 (3 data and 1 parity volume) or 7 +1 (7 data and 1 parity volume)
- **Raid 5 Striped RAID volumes**
 - Data blocks are striped horizontally across the members of the RAID group (4 or 8 member group); parity blocks rotate among the group members
- **RAID 10 Mirrored Striped Mainframe Volumes**
- **Dynamic Sparing**
- **SRDF (Symmetrix Remote Data Facility)**
 - Mirror of Symmetrix logical Volume maintained in a separate Symmetrix

Data protection options are configured at the volume level and the same Symmetrix can employ a variety of protection schemes.

RAID stands for a Redundant Array of Independent Disks.

Mirroring: RAID-1

- Two physical “copies” or mirrors of the data
- Host is unaware of data protection being applied



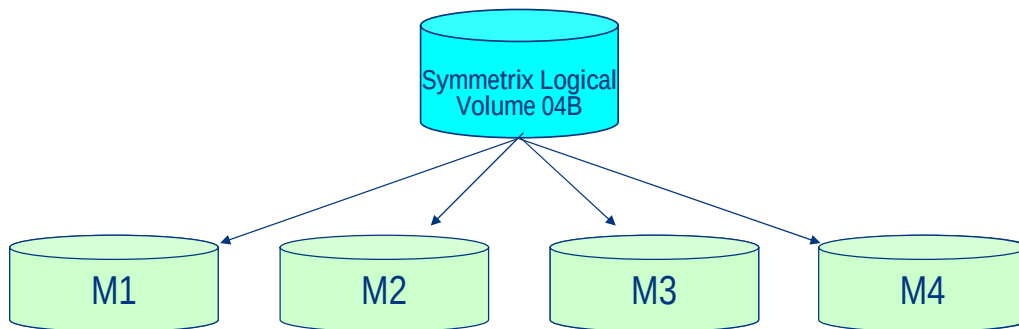
© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 42  Mercer Road
EMC Education Services

Mirroring provides the highest level of performance and availability for all applications. Mirroring maintains a duplicate copy of a logical volume on two physical drives. The Symmetrix maintains these copies internally by writing all modified data to both physical locations. The mirroring function is transparent to attached hosts, as the hosts view the mirrored pair of hypervisors as a single Symmetrix logical volume.

Mirror Positions

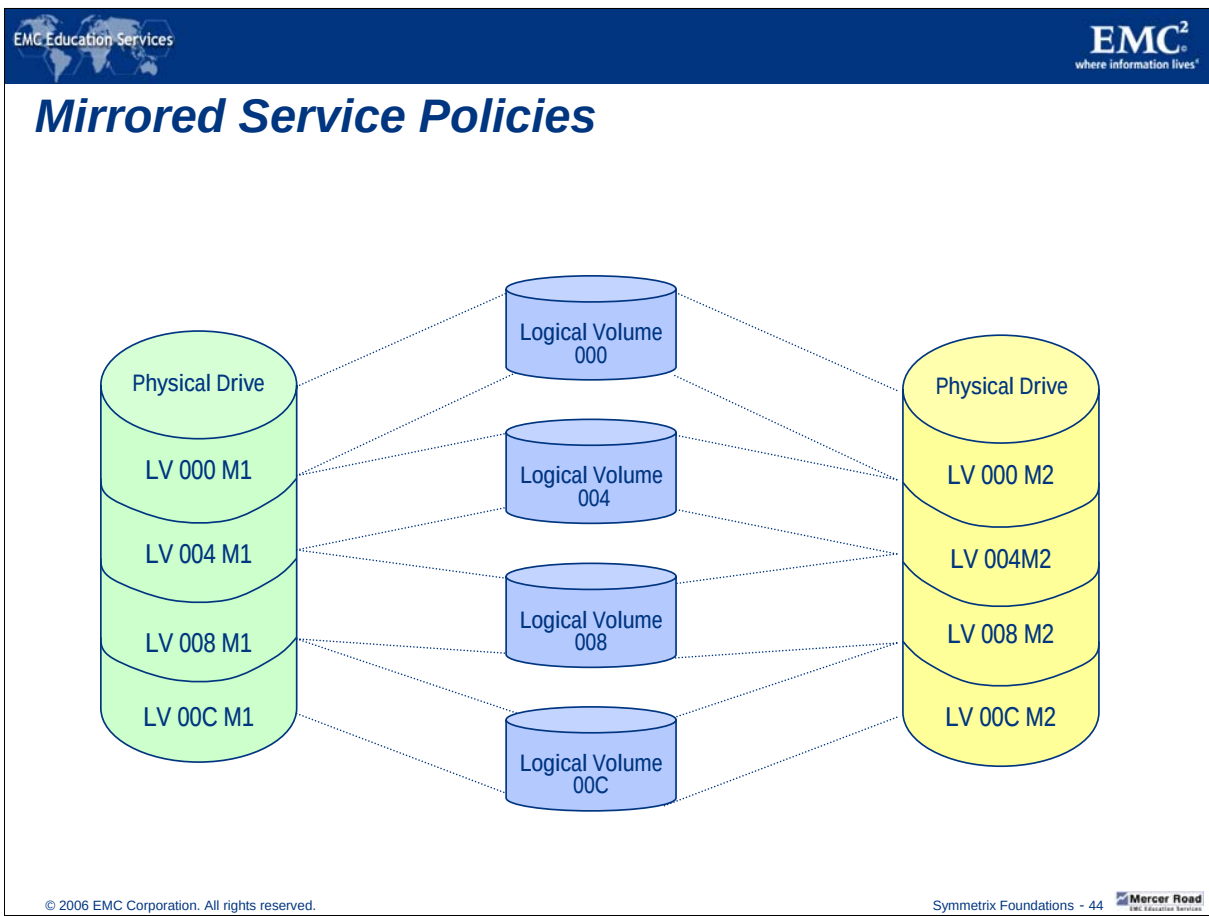
- Internally each Symmetrix Logical Volume is represented by four mirror positions – M1, M2, M3, M4
- Mirror position are actually data structures that point to a physical location of a mirror of the data and status of each track
- Each mirror positions represents a mirror copy of the volume or is unused



© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 43  Mercer Road
EMC Education Services

Within the Symmetrix, each logical volume is represented by four mirror positions – M1, M2, M3, and M4. These Mirror Positions are actually data structures that point to a physical location of a data mirror and the status of each track of data. Each position either represents a mirror or is unused. For example, an unprotected volume will only use the M1 position to point to the only data copy. A RAID-1 protected volume will use the M1 and M2 positions. If this volume was also protected with SRDF, three mirror positions would be used, and if we add a BCV to this SRDF protected RAID-1 volume, all four mirror positions would be used.

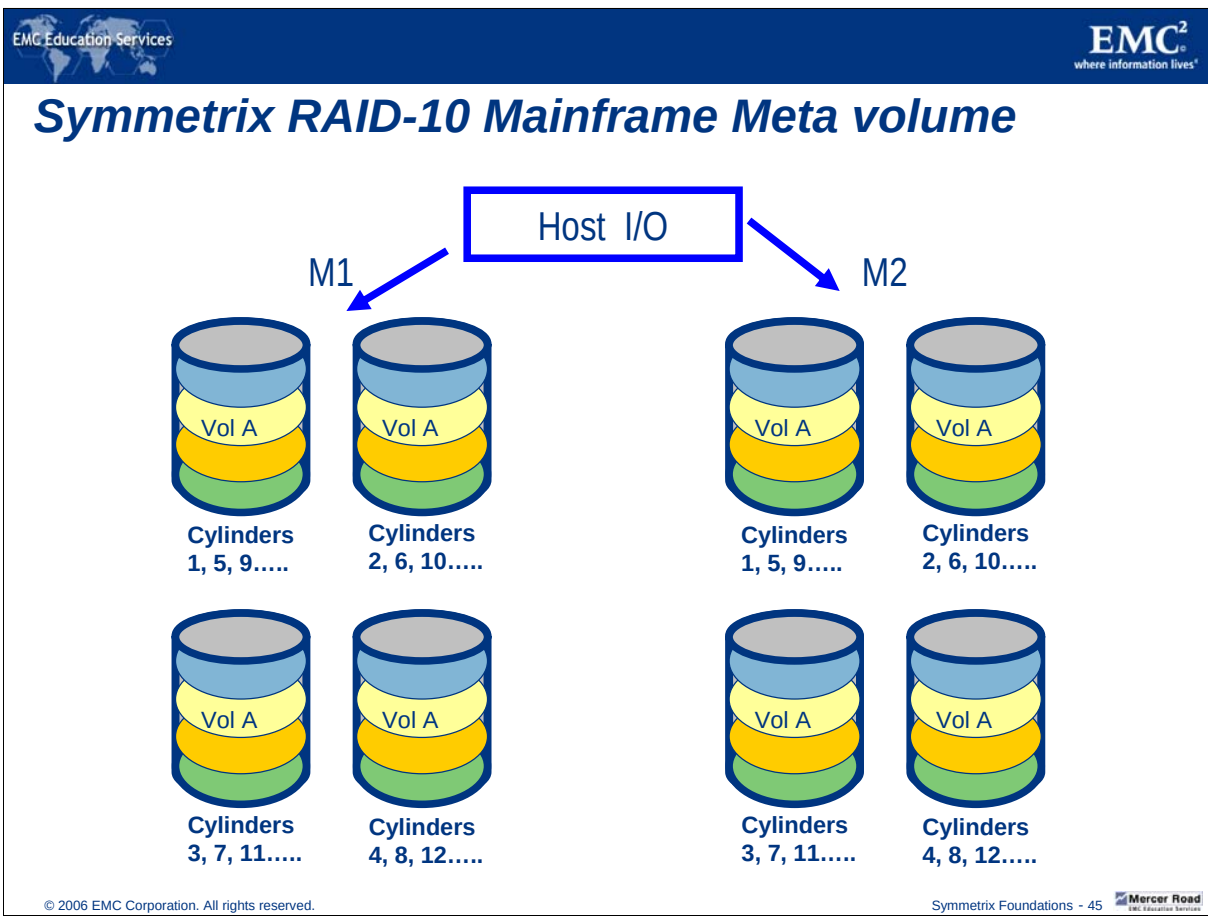


Symmetrix performance algorithms for read operations choose the best hyper in the mirrored pair based upon three service policies: Interleave Service Policy, Split Service Policy, or Dynamic Mirror Service policy.

Interleave Service Policy - Shares the read operations of the mirrored pair by reading tracks from both logical hypers in an alternating method: a number of tracks from the primary volume (M1) and a number of tracks from the secondary volume (M2). The interleave policy is designed to achieve maximum throughput.

Split Service Policy - Differs from the interleave policy because read operations are assigned to either the M1 or the M2 logical volume, but not to both. Split is designed to minimize head movement.

Dynamic Mirror Service policy (DMSP) - Utilizes both Interleave and split for maximum throughput and minimal head movement. Dynamic Mirror Service policy adjusts each logical volume dynamically, based on access patterns detected. This is the default mode within the Enginuity operating system.



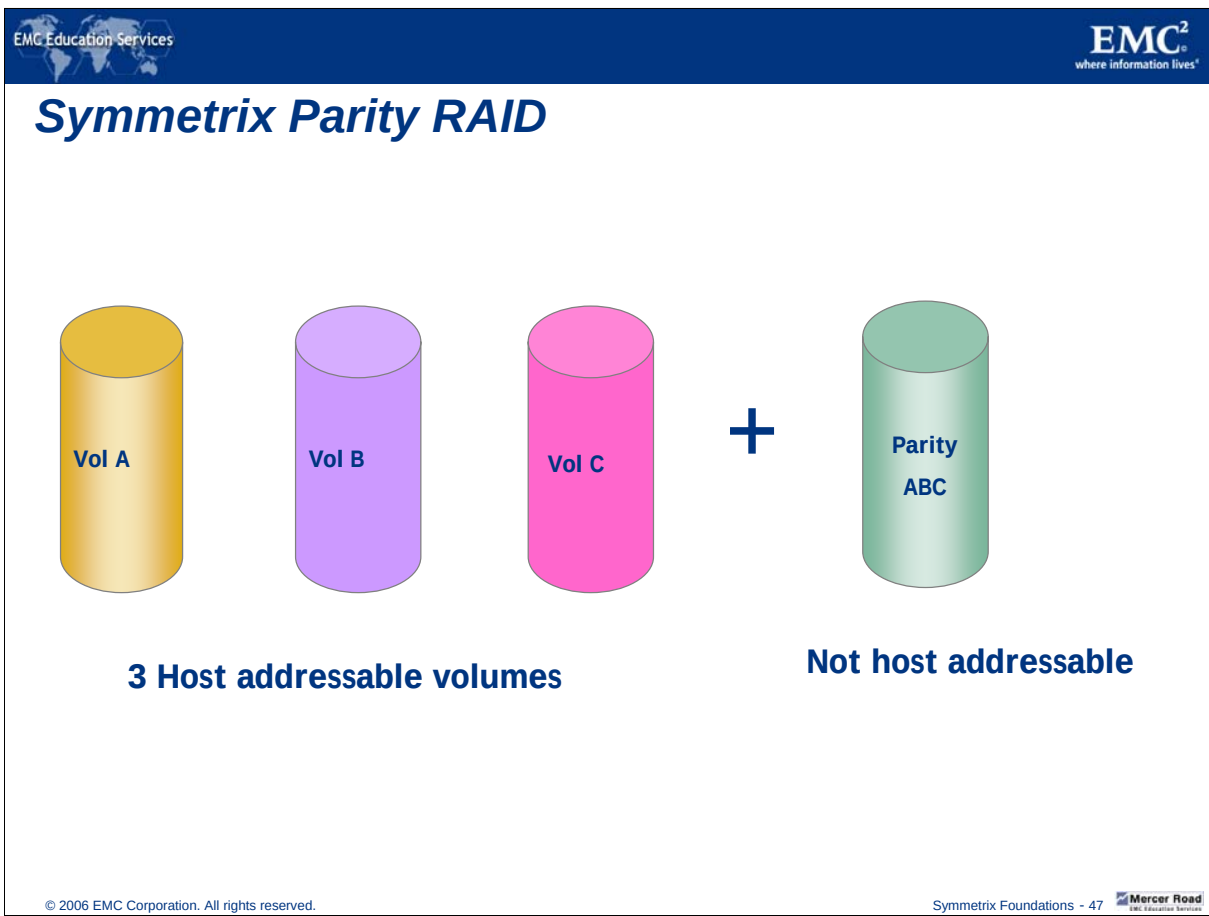
This is a diagram of a RAID-10 stripe mainframe meta volume group. Each RAID-10 stripe group consists of four stripes distributed across four volumes. These are mirrored to consist of eight total volumes. The stripe group is constructed by alternately placing one cylinder across each of the four volumes. These volumes cannot be on the same disk director. The eight volumes are distributed across the Symmetrix back end for additional availability and improved performance.

Symmetrix DMX/DMX2 Parity RAID Advantages

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

Parity RAID is not available on the Symmetrix DMX3.

Compared to a RAID-1 mirrored Symmetrix system, Parity RAID offers more usable capacity than a mirrored system containing the same number of disk drives.



A Parity RAID rank is the set of logical volumes related to each other for parity protection. A data volume is presented to the host operating system and defined as a separate unit address to the host. All data volumes within a rank must be the same size and emulation such as Fixed Block Architecture or Count Key Data.

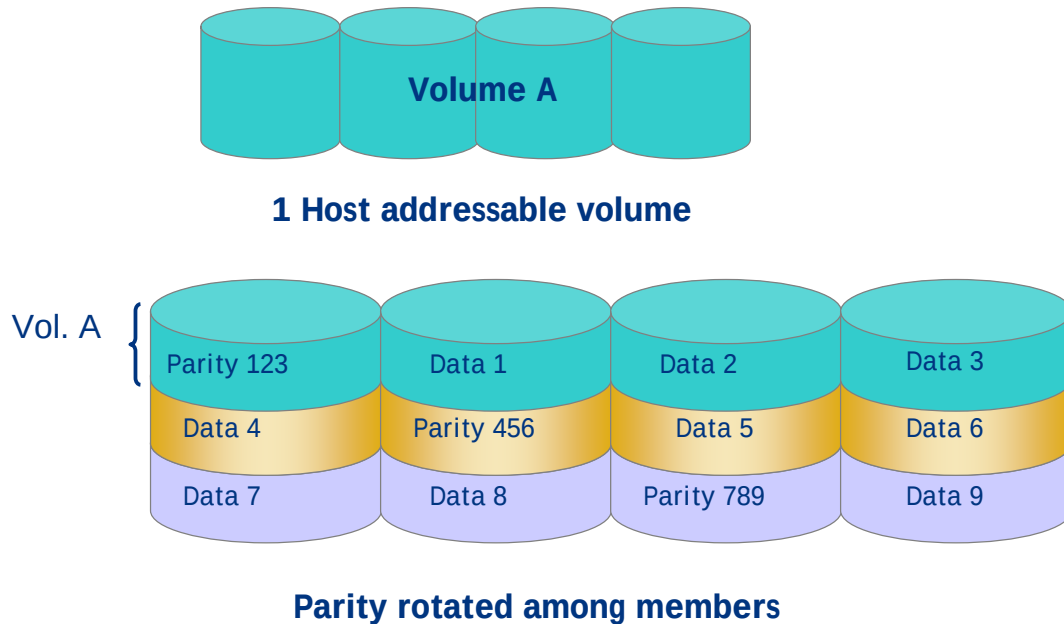
Parity RAID employs the EXCLUSIVE OR (XOR) Boolean operation and XOR hardware assist, built into the Symmetrix global memory directors, distributes the XOR function throughout the system to improve performance in regeneration mode. Parity RAID is available in (3+1) or (7+1) configurations, but both of these cannot exist within the same Symmetrix. This graphic illustrates a (3+1) configuration.

Symmetrix RAID-5 Volume Attributes

- RAID-5 track size is 32KB for open system and 57KB for mainframes (Enginuity 5771 uses 64KB track size)
- Data blocks are striped horizontally across the members of a RAID-5 group, each member owns some data tracks and some parity tracks
- There is no separate parity volume in a RAID-5 group. Instead, parity blocks rotate among the group members.
- RAID-5 groups can be:
 - Four members per logical volume, RAID 5(3+1)
 - Eight members per logical volume, RAID 5(7+1)

For ease of identifying the RAID-5 groups, EMC uses 3RAID 5 to describe the four-member group otherwise identified as RAID 5(3+1). Likewise, 7RAID 5 refers to the eight-member group otherwise identified as RAID 5(7+1).

Symmetrix 3RAID-5 (4 Members)

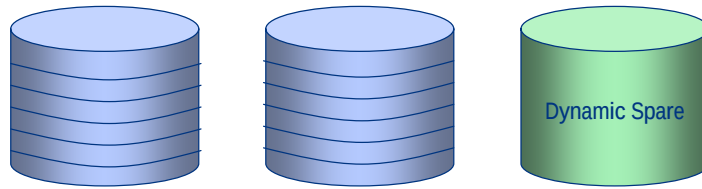


© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 49 

Symmetrix DMX Series RAID-5 optimizes performance for large sequential write workloads as there is no need to read the parity from disks. Since many sequential tracks are written, they are all in Symmetrix global memory. The parity is calculated in global memory and information is written to the disk in one stroke without requiring the use of an expensive disk-level read-XOR –write operation. RAID-5 is available in (3+1) or (7+1) member configurations, but both of these cannot exist within the same Symmetrix. This graphic illustrates a (3+1) member configuration. Engenuity 5670 or higher is required for a Symmetrix RAID-5 configuration.

Dynamic Sparing



- Increases protection of all volumes from loss of data
- Dedicated spare disk(s) protect storage
- Ensures that the spare copy is identical to the original
- Resynchronizes a new disk device with the dynamic spare after repair of the defective device is complete
- Increases data availability of all volumes in use without loss of any data capacity
- Dynamic Sparing is transparent to the host and requires no user intervention

Dynamic Sparing is used as additional protection for volumes already protected by RAID-1 mirroring, Parity RAID, RAID-5, or SRDF options. Dynamic Sparing provides incremental protection against failure of a second disk during the time a disk is taken offline and when it is ultimately replaced and resynchronized. Every Symmetrix logical volume has four mirror positions. When sparing is necessitated, Hyper Volumes on the spare disk devices take the next available mirror position for the logical volumes present on the failing volume. All of these Dynamic Spare Hyper Volumes are marked as having all tracks invalid in the respective mirror positions of the logical volumes. It is now the responsibility of the Symmetrix to copy all tracks over to the Dynamic Spare. Dynamic Sparing occurs at the physical drive level, since a physical drive is the Field Replaceable Unit in the Symmetrix. In other words, you can't just replace a failed Hyper Volume, only the disk it resides on. However, the actual data migration from the volumes on the failed drive to the Dynamic Spare occurs at the logical volume level.



Symmetrix Foundations

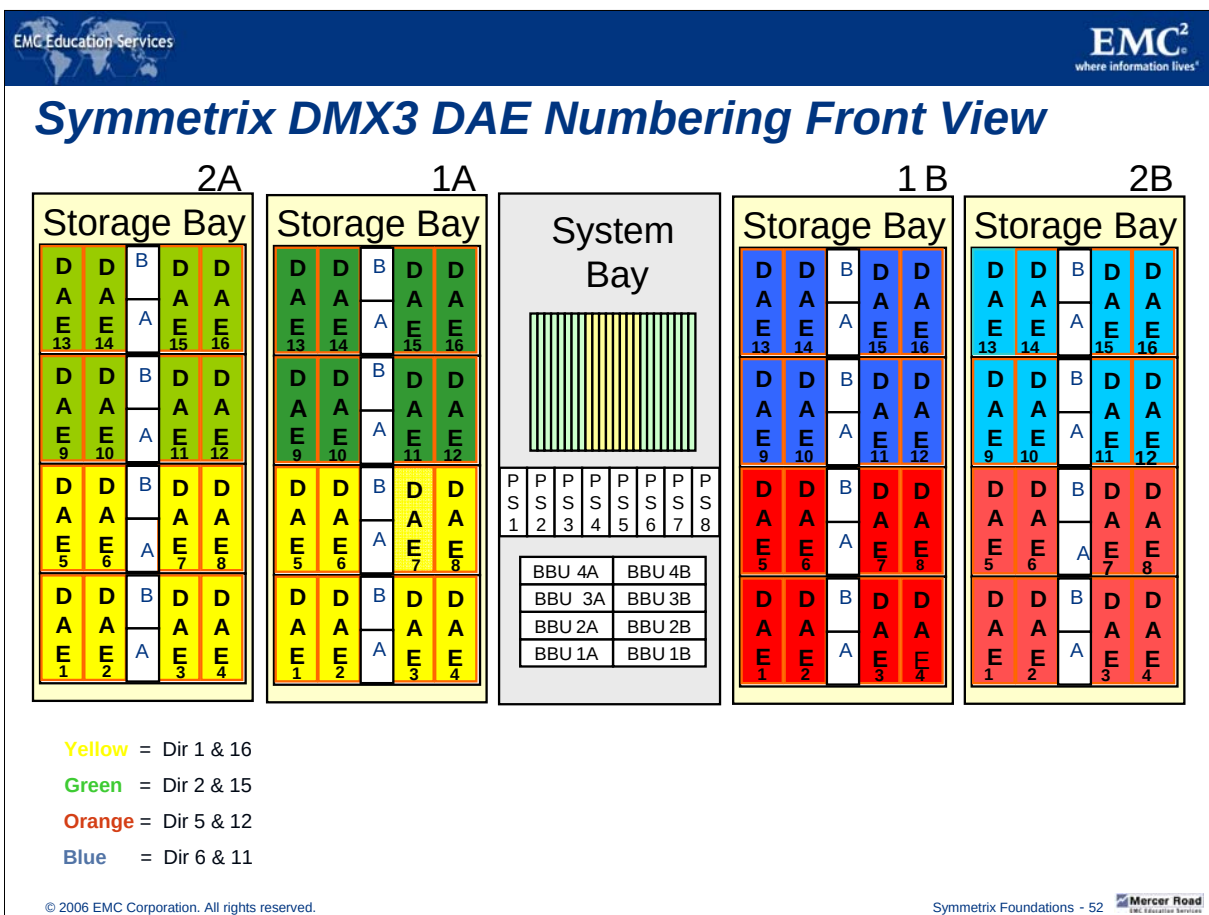
Symmetrix Configuration Fundamentals



© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 51  Mercer Road
EMC Education Services

The next few slides will explain Symmetrix configuration fundamentals.



Here is an example of a DMX 4500, Disk Array Enclosures are numbered from 1 – 16 in each Storage Bay. Disk Array Enclosure 1 is on the bottom left, counting across, 2, 3 and 4. Disk Array Enclosure 16 is located on the top right. There are sixteen direct connect disk array enclosures in Storage Bay 1A and 1B, configured with a 15 Drive loop. There are sixteen daisy chained disk array enclosures in Storage Bay 2A and 2B, which will expand the 15 drive loops to 30 drive loops. There are eight battery backup units in each storage bay, 4 for power zone A and 4 for power zone B. The storage bays are color coded to match the directors used to configure them. Yellow represents directors 1 and 16, green represents directors 2 and 15, orange represents directors 5 and 12, while blue represents directors 6 and 11. Configurations can now contain up to 10 storage bays (240 disk per bay) for a maximum disks count of 2,400 disks.

EMC Education Services
EMC²
where information lives[®]

Symmetrix DMX2 Director Configuration Information

Director Map : TEMP\$\$\$ BIN

File Tasks Windows Help

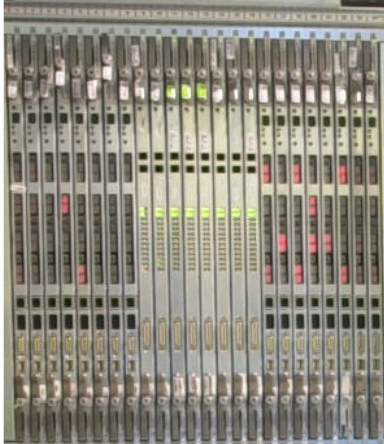
SymmWin Director Map Configuration

Init DirMap DirEdit Hosts Disks Algo

(BOARD 0 0 "FIBER 4x2" opus 0, portsPerCpl 2, HcMode 2)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
D	DF	DF	FA	FA	DF	DF	FA	FA	FA	FA	DF	DF	FA	FA	DF	DF
C	DF	DF	FA	FA	DF	DF	FA	FA	FA	FA	DF	DF	FA	FA	DF	DF
B	DF	DF	FA	FA	DF	DF	FA	FA	FA	FA	DF	DF	FA	FA	DF	DF
A	DF	DF	FA	FA	DF	DF	FA	FA	FA	FA	DF	DF	FA	FA	DF	DF
brd	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2	FIBER 4x2

Symmetrix Director
Hardware



© 2006 EMC Corporation. All rights reserved.
Symmetrix Foundations - 53
Mercer Road
EMC Education Services

Here is an Example of a Symmetrix DMX3000 with 8 disk directors, SymmWin is a graphics-based tool for configuring and monitoring a Symmetrix system. Symmetrix configuration information includes physical hardware installed, the number and type of directors, memory size, and mapping of addresses to front-end directors along with operational parameter bit settings for front-end director adapter to host connectivity. Configuration information created with SymmWin GUI is stored in the IMPL.bin file. Changes made to the bin file must first be made to the IMPL.bin on the Service Processor and then downloaded to the directors over the internal Ethernet LAN. Configuration changes can also be made using EMC ControlCenter Configuration Manager GUI and Solutions Enabler CLI.

EMC Education Services
EMC²
where information lives[®]

Symmetrix DMX2 Disk/Volume Configuration Information

SymmWin Disk Map Displaying Volumes

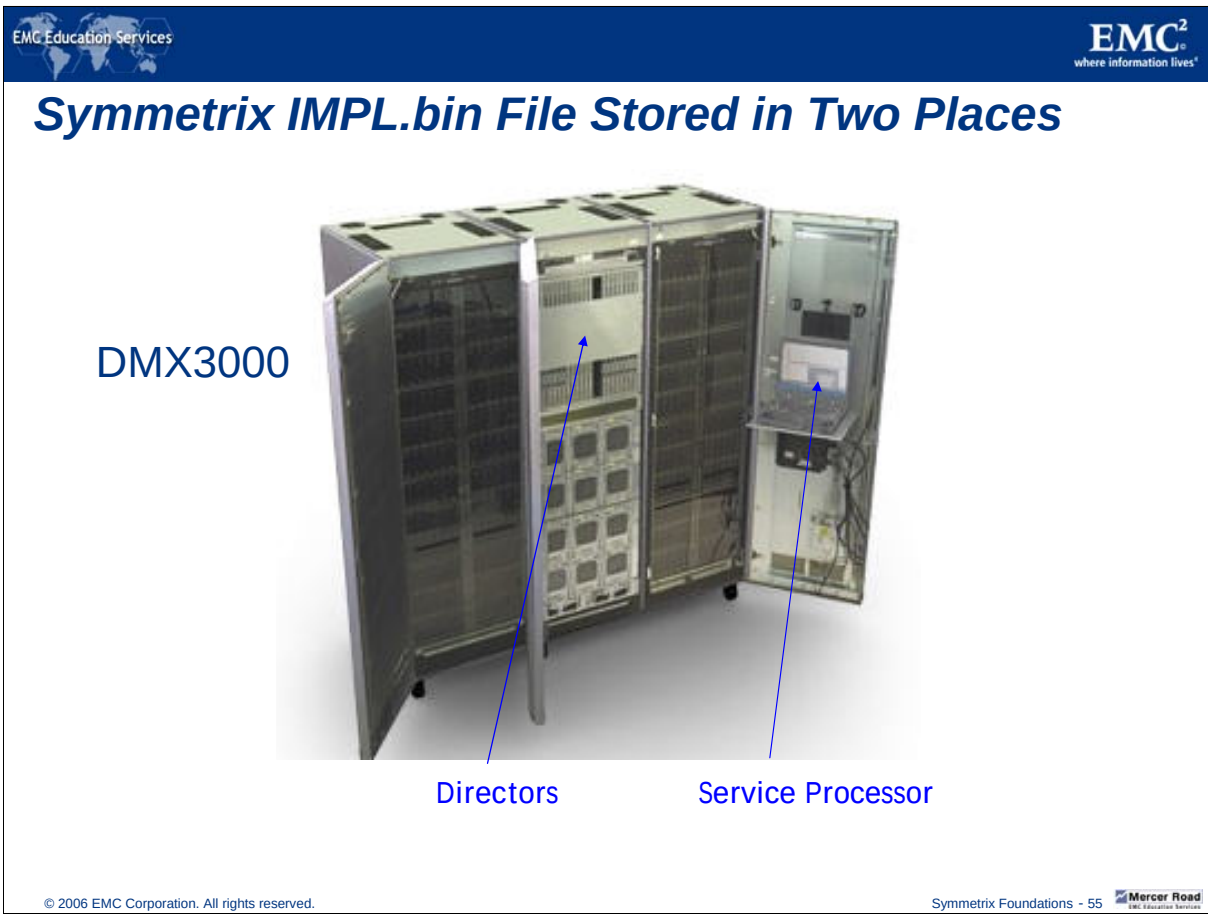
Configured Disk

Logical Volumes

Physical Disk Hardware

© 2006 EMC Corporation. All rights reserved.
Symmetrix Foundations - 54

Here is a SymmWin GUI representation of the disk in the Symmetrix. The logical volumes configured on the highlighted disk are displayed on the right.



Both Channel and Disk directors have a local copy of the configuration file stored in EEPROM. This enables Channel Directors to be aware of the Disk Directors that are managing the physical copies of Symmetrix Logical Volumes and vice versa. The IMPL.bin file also allows Channel Directors to map host requests to a channel address, or target and LUN to the Symmetrix Logical Volume.

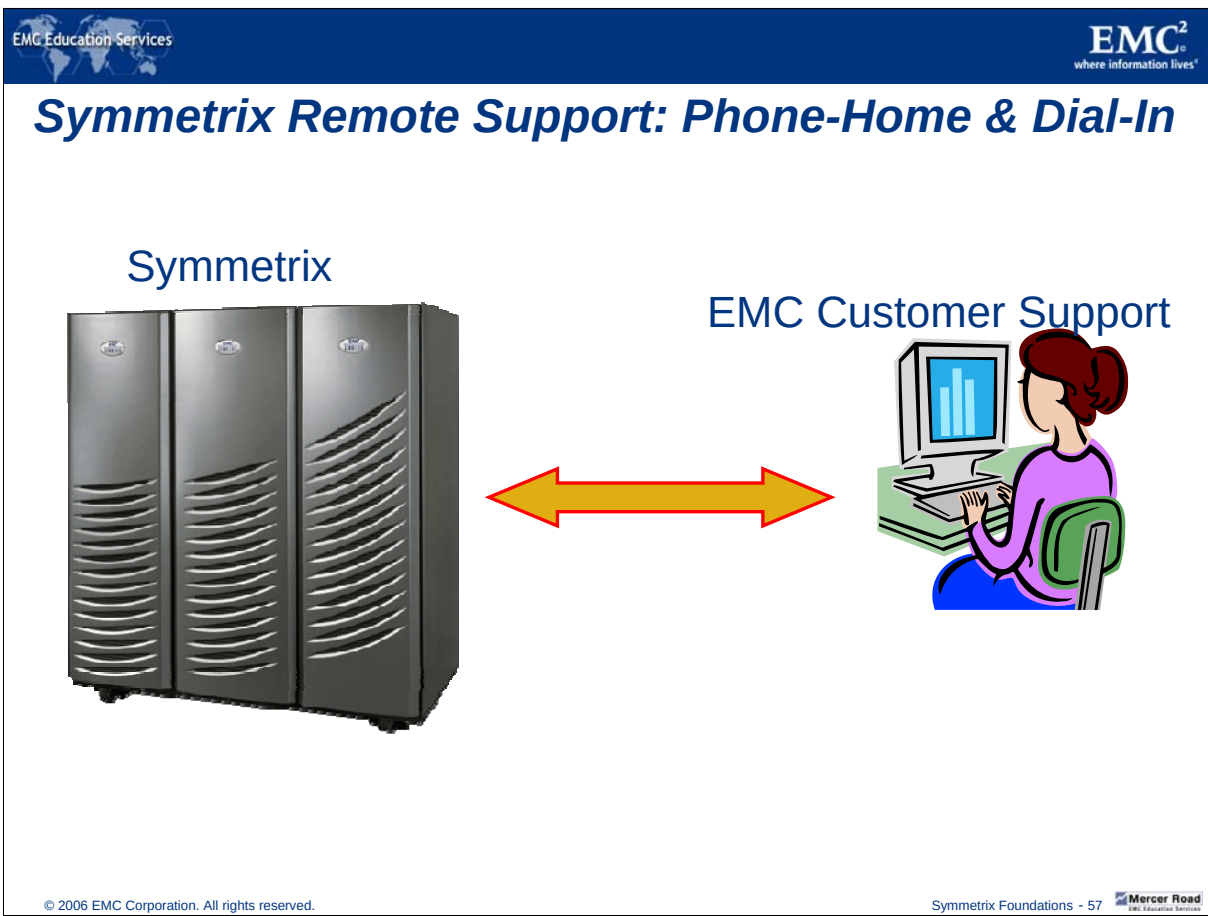
Configuration Considerations

- Understand the applications on the host connected to the Symmetrix system
 - Capacity requirements
 - I/O rates
 - Read/Write ratios
 - Read/Write - Sequential or Random
- Understand special host considerations
 - Maximum drive and file system sizes supported
 - Consider Logical Volume Manager (LVM) on the host and the use of data striping
 - Device sharing requirements - Clustering
- Determine Volume size and appropriate level of protection
 - Symmetrix provides flexibility for different sizes and protection within a system
 - Standard sizes make it easier to manage
- Determine connectivity requirements
 - Number of channels available from each host
- Distribute workloads from the busiest to the least busy

The best possible performance will only be achieved if all the resources within the system are being equally utilized. This is much easier said than done, but through careful planning, you will have a better chance for success. Planning starts with understanding the host and application requirements.

Within the Symmetrix bin file, the emulation type, size in cylinders, count, number of mirrors, and special flags (like BCV, DRV, Dynamic Spare) are defined. Each Symmetrix logical volume is assigned a hexadecimal identifier. The bin file also tells the Channel director which volumes are presented on which port, and the address used to access it.

From the Host's perspective, when a device discovery process occurs, the information provided back to the Operating System appears to be referencing a series of disk drives. The host is unaware of the bin file, RAID protection, remote mirroring, BCV mirrors, dynamic sparing, etc. In other words, the host "thinks it's getting" an entire physical drive.

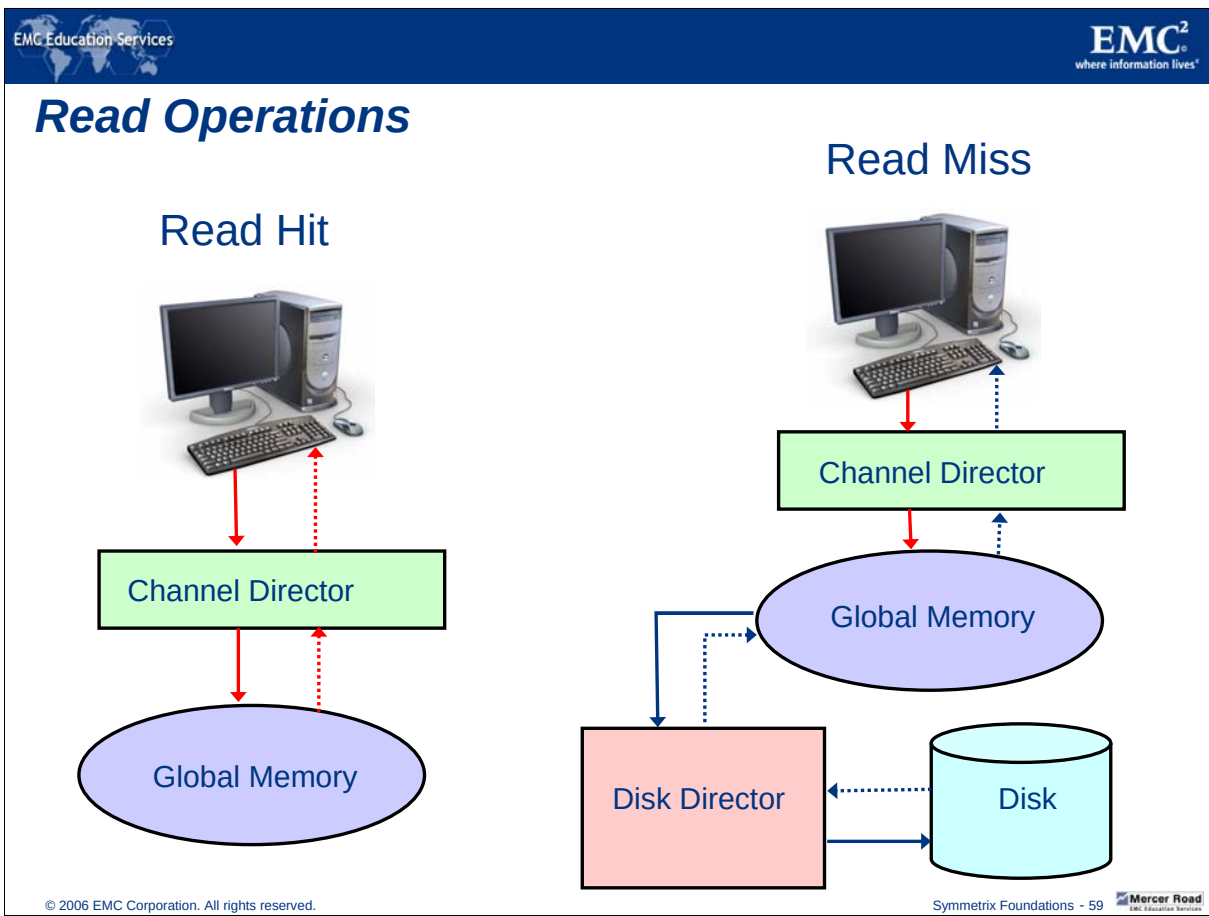


Using EMC Remote and SymmWin software on the service processor or server, the Symmetrix is configured to phone home and alert EMC Customer Support of a failure or potential failure. Remote access can be done through network or phone technologies. When required, a Customer Engineer will be dispatched to the Symmetrix to replace hardware or perform other maintenance.

Symmetrix Foundations

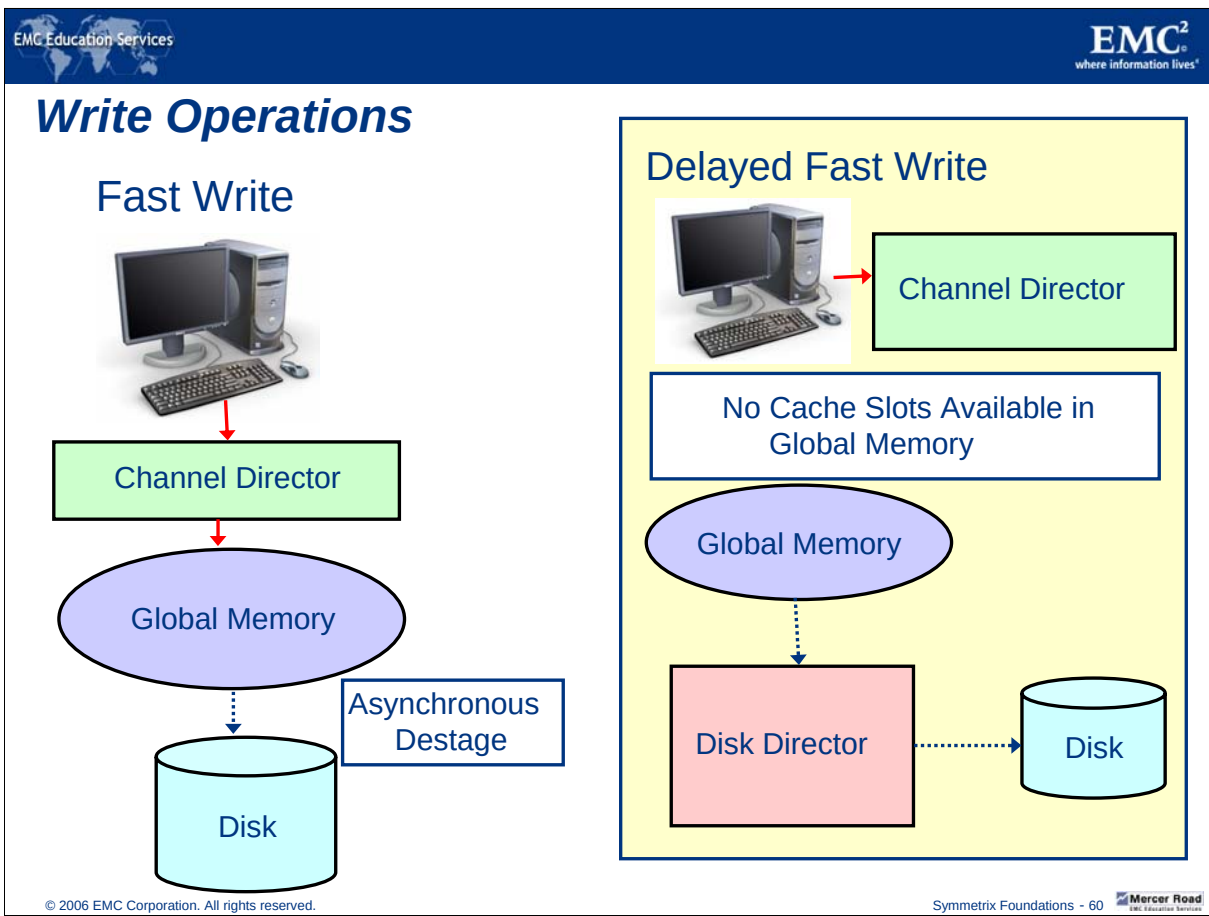
Host Data Access to Symmetrix Data

Now that we have an understanding of Symmetrix architecture, we will discuss host access to Symmetrix data.



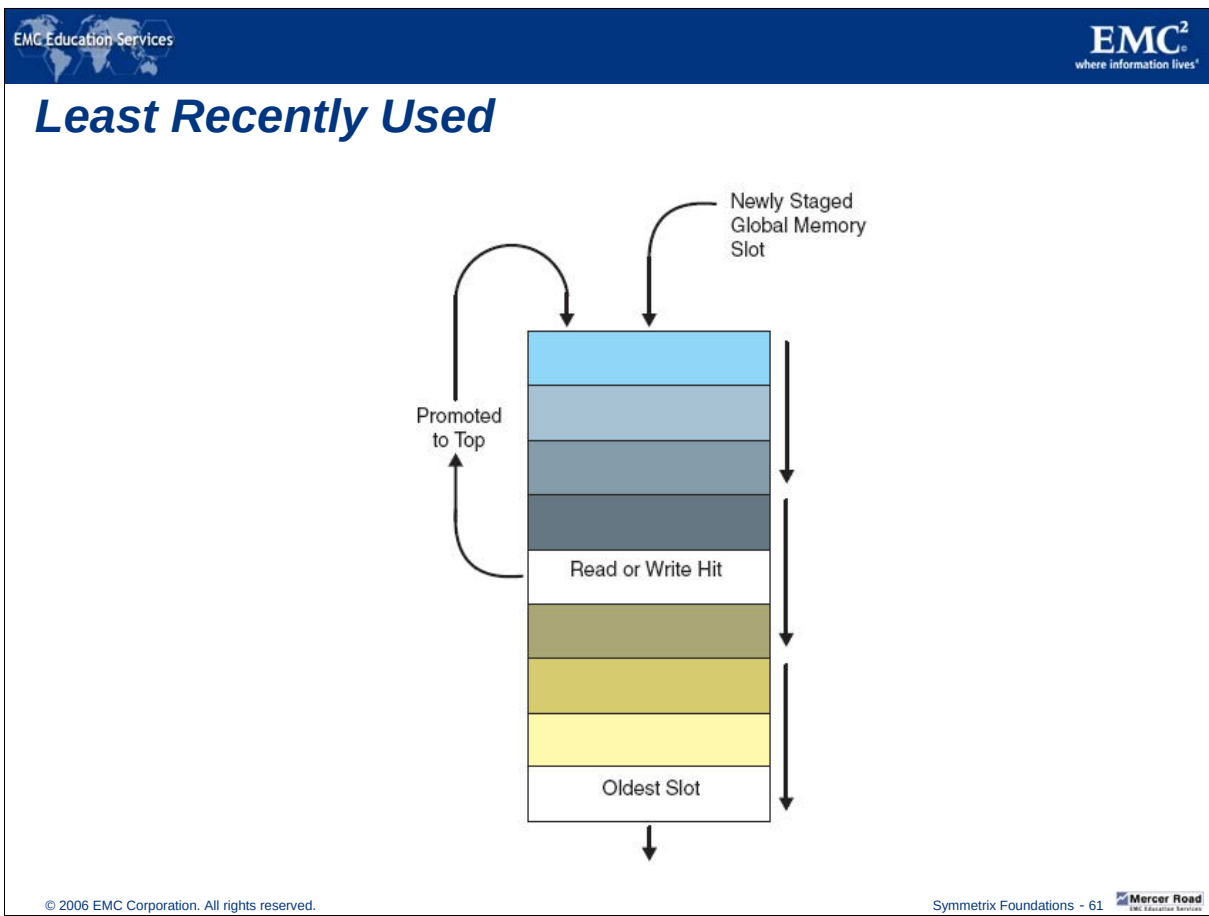
In a Read hit operation, the requested data resides in global memory. The channel director transfers the requested data through the channel interface to the host and updates the global memory directory. Since the data is in global memory, there are no mechanical delays due to seek and latency.

In a read miss operation the requested data is not in global memory and must be retrieved from a disk device. While the channel director creates space in the global memory, the disk director reads the data from the disk device. The disk director stores the data in global memory and updates the directory table. The channel director then reconnects with the host and transfers the data. Because the data is not in global memory, the Symmetrix system must search for data on the disk and then transfer it to the channel, this adds seek and latency times to the operation.



A fast write occurs when the percentage of modified data in global memory is less than the fast write threshold. On a host write command, the channel director places the incoming block(s) directly into global memory. For fast write operations, the channel director stores the data in global memory and sends a “channel end” and “device end” to the host computer. The disk director then asynchronously destages the data from global memory to the disk device.

A delayed fast write occurs only when the fast write threshold has been exceeded. That is, the percentage of global memory containing modified data is higher than the fast write threshold. If this situation occurs, the Symmetrix system disconnects the channel directors from the channels. The disk directors then destage the Least Recently Used data to disk. When sufficient global memory space is available, the channel directors reconnect to their channels and process the host I/O request as a fast write. The Symmetrix system continues to process read operations during delayed fast writes. With sufficient global memory present, this type of global memory operation rarely occurs.



The Symmetrix System supports two different mechanisms for Least Recently Used (LRU): the traditional double-linked list, and Tag Based Caching (TBC).

The Least Recently Used is a data structure that keeps the slots in the order the system accesses them. The Least Recently Used algorithm determines which slot was least recently used. This slot loses its association with the track/data that is stored.

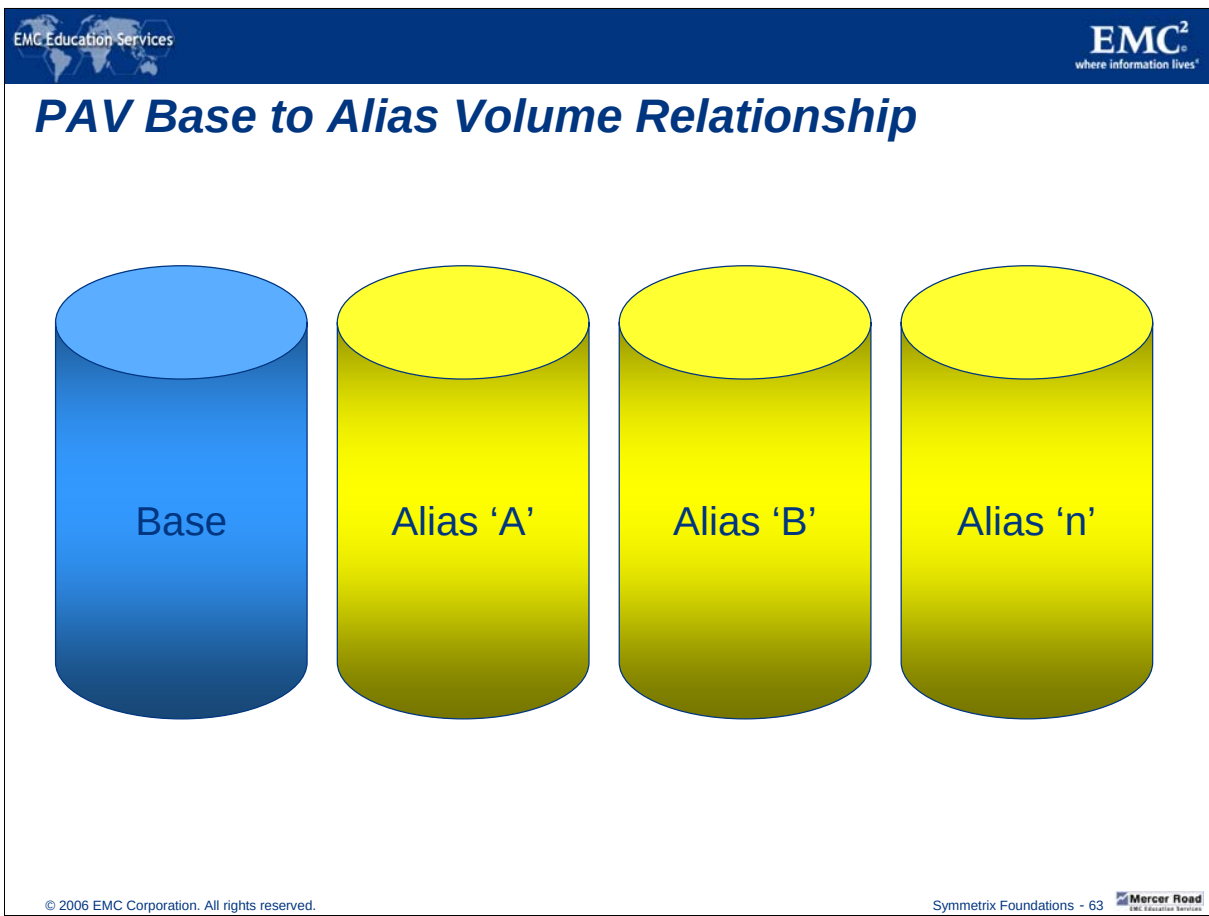
Tag Based Caching is the default cache management algorithm used in Enginuity 5670 and higher, and divides global memory into groups of several hundred slots called Tag Based Cache groups. In the Tag Based Cache data structure, two bytes represent each slot. The two bytes contain information about the last time the system most recently accessed this slot, and whether the slot is write pending. The bytes that represent the slots of a Tag Based Cache group are contiguous in global memory. All the CPUs in a Symmetrix system access all the Tag Based Cache groups with each CPU accessing each Tag Based Cache group in a different order. The system manipulates the Tag Based Cache groups under lock.

The diagram above represents data flow with the Least Recently Used algorithm. Each time a read hit or write hit occurs, the Symmetrix System marks that memory slot as most recently used and promotes it to the top of the Least Recently Used list. For each write, a written-to flag is set on the initial write to each global memory block and is cleared when the global memory block is destaged. The Least Recently Used global memory slot appears at the bottom of the Least Recently Used list.

Prefetch to Memory

- Process is used to avoid a global memory read miss
- Continually monitor I/O activity and look for patterns
- Sequential prefetch process is invoked when a sequential I/O to a track occurs
- Sequential process discontinues when the host processor uses a random I/O pattern

The intelligent, adaptive prefetch algorithm reduces response time and improves performance by transferring data into memory before a host requests it. The prefetch algorithm maintains, per each logical volume, an array of statistics and parameters based on the latest sequential patterns observed on the logical volume. Prefetch dynamically adjusts based on workload demand across all resources in the back-end of the Symmetrix system. This algorithm also ensures that global memory resources are never overly consumed in order to maintain optimal performance.



The Symmetrix systems support Compatible Parallel Access Volumes, an IBM feature that improves response time by reducing device contention, resulting in higher performance and throughput. Parallel Access Volumes is a mainframe-exclusive feature that resolves operating system limitation allowing only one outstanding I/O operation to a device.

Parallel Access Volume technology allows a single mainframe host to simultaneously process multiple I/O operations to the same logical volume. A Base volume can be thought of as the real physical disk space, with its own unique sub-channel ID. Alias volumes are mapped against the Base's physical space. Parallel Access Volumes essentially present multiple addresses for the same logical device within the operating system. By presenting multiple Unit Control blocks for the same device, I/Os can be queued instead of being rejected and access time can be reduced. Enginuity adds dynamic support to the existing PAV implementation which enables management utilities to dynamically reassign aliases to a base, improving the opportunity for parallel I/O operations.

Application Considerations for Host Connectivity

- It is not just about physical access to data; it is about how the data is to be used
 - How often does the data change
 - Performance considerations
 - Sharing considerations
 - Capacity requirements
 - Availability requirements
 - Distance between host and storage
 - Skill level of administration team

It is really the application that determines the appropriate connectivity technology. Here are just a few of the issues that should be addressed when assessing an environment while architecting a storage infrastructure.

Symmetrix Foundations

Environmental Integration

Now we will look at environmental integration through host connectivity.



Symmetrix DMX Series Enterprise Connectivity



- Fibre Channel**
 - UNIX, Windows, Netware, Linux, IBM iSeries
 - Direct and SAN attach, SRDF Family links
- ESCON**
 - Mainframe and SRDF Family links
- FICON**
 - High performance for mainframe
- Native Gigabit Ethernet**
 - For SRDF replication
 - Supports compression
- Native iSCSI**
 - Industry's first high-end implementation
- NAS gateway**
 - Celerra CNS
 - NSxxxG NAS gateway (where xxx is the model)

© 2006 EMC Corporation. All rights reserved.

Symmetrix Foundations - 66 

The Symmetrix DMX provides the widest range of connectivity options in the industry. It supports Fibre Channel, ESCON, FICON, and Native Gigabit Ethernet. Additional supported connectivity options are iSCSI and Celerra NAS gateway.

Module Summary

Key points covered in this module:

- Front-end directors, back-end directors, cache and disk location in a Symmetrix DMX
- The relationship between Symmetrix physical disk and Symmetrix logical volumes
- Volume protection options available on the Symmetrix
- The I/O path through Symmetrix cache
- Symmetrix DMX connectivity options
- Symmetrix DMX3 Vaulting Overview

These are the key points covered in this training. Please take a moment to review them.