

CLARiiON Foundations

- Draw and describe the basic architecture of a CLARiiON Disk Array
- Differentiate among the architectures of the various CLARiiON models
- List the data protection options available on CLARiiON
- Using a diagram, illustrate the relationship between CLARiiON physical disk drives and LUNs
- Illustrate some of the high availability features of the CLARiiON and how this potentially impacts data availability

These are the learning objectives for this training. This module is reinforced by lab exercises using the Navisphere Manager Simulator.



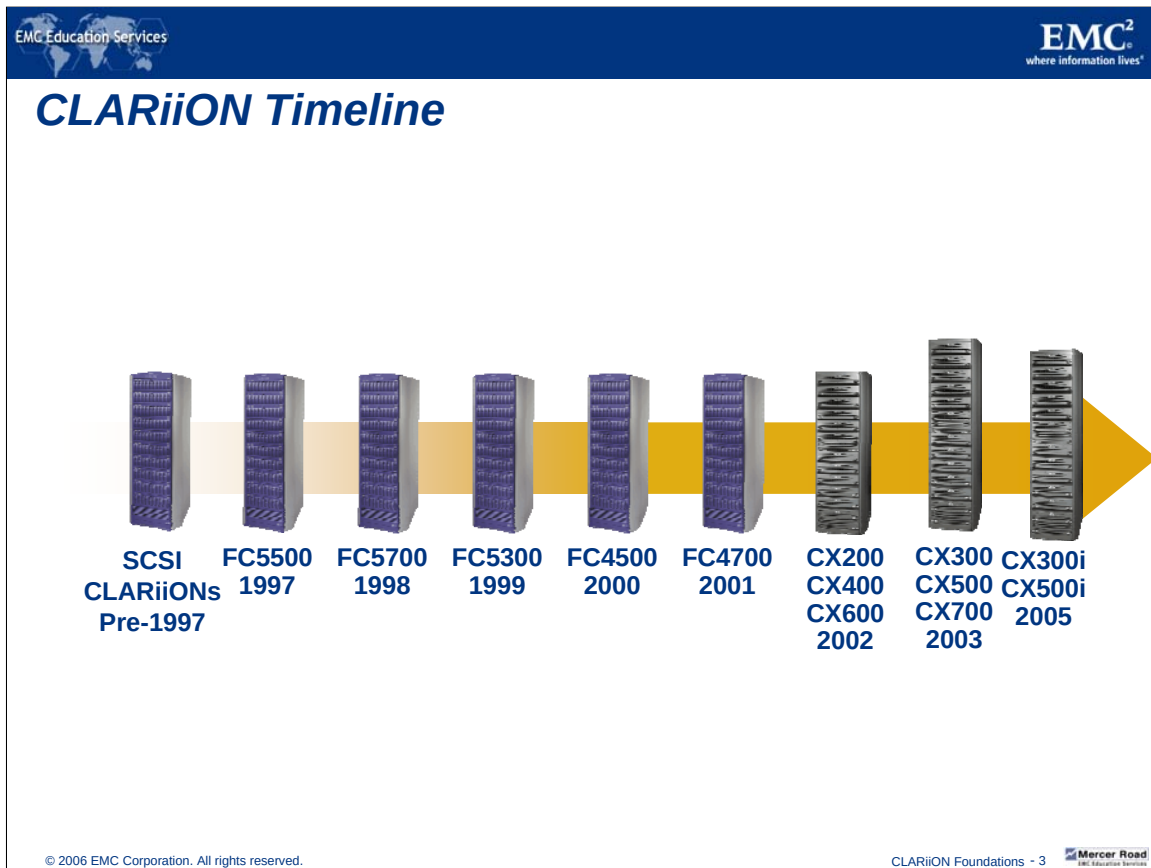
CLARiiON Foundations

CLARIION RANGE AND COMPONENTS

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 2 

First, we'll take a look at the range and components of CLARiiON.



The CLARiiON storage array is the EMC mid-tier storage offering, which delivers enterprise quality features and functionality. This does not mean that the CLARiiON range cannot meet with the most demanding of storage environments, but rather the CLARiiON range will meet with a different set of criteria than the traditional "high end" storage solution.

The CX series is the sixth generation of the CLARiiON storage systems. Each generation has added enhancements to performance, availability, and scalability, while the high level architecture has remained constant. This allows interoperability between generations and helps preserve the customer's investment in hardware, software, knowledge, and skills.

CLARiiON was the first full Fibre Channel storage system.



High-End Storage: The New Definition

High-End Then	High-End Today
• Simple redundancy – Automated Fail-over • Benchmark performance (IOPs and MB/s) – Single and/or simple workloads • Basic local and remote data replication – Backup windows, testing, and disaster recovery • Scalability – Capacity • Manage the storage system – Easy configuration, simple operation, minimal tuning	• Non-disruptive everything – Upgrades, operation, and service • Predictable performance... unpredictable world – Complex, dynamic workloads • Replicate any amount, any time, anywhere – Replicate any amount data, across any distance, without impact to service levels • Flexibility – Capacity, performance, multi protocol connectivity, workloads, etc. • Manage service levels – Centralized management of the storage environment

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 4



The decision of which storage product to purchase should always be driven by the application requirements. EMC offers multiple products that meet varied requirements for performance, availability, and functionality. The requirement for economics is the same for both mid-tier and high-end.

Performance: Mid-tier storage architectures are ideally suited for applications with predictable workloads that need fast response times and high sustained throughput. Mid-tier architectures can also be manually tuned and balanced to optimize performance for the applications they support.

Availability: For applications that require 99.999% availability, mid-tier architectures are designed with redundant components, support *(multiple levels of)* RAID, and should have the capability to continuously monitor the entire system looking for health status and latent errors.

Functionality: Mid-tier storage should also have functional capabilities to make point-in-time local replicas and do remote replication for business continuity. The management tools should be intuitive and seamlessly plug into other management frameworks.

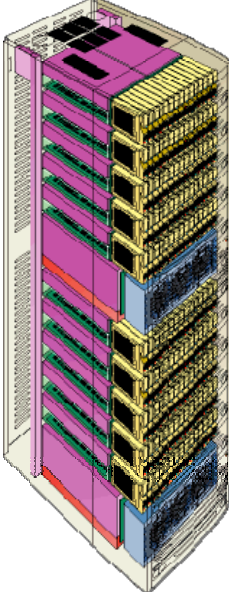
Economics: Lowest possible acquisition price, affordable upgrades, and investment protection.

EMC Education Services

EMC²
where information lives™

Flexible, High Availability Design

- Fully redundant architecture
 - Power, cooling, data paths, SPS
 - Non-stop operation
 - Online software upgrades
 - Online hardware changes
- Continuous diagnostics
 - Data and system integrity
 - CLARAlert Phone Home
- Dual I/O paths with non-disruptive failover
- Leader in data integrity
 - Mirrored write cache
 - Destage write cache to DISK upon power failure
 - SNiFF Verify
 - Background Verify-Per RAID Group



- SnapView and MirrorView replication software
- SAN Copy
- No single points of failure, modular architecture
- Fibre Channel and ATA RAID
- From 5 to 240 disks
- Flexibility
 - Individual Disk
 - RAID levels 0, 1, 1/0, 3, 5
 - Mix drive types
 - Mix RAID levels
- Up to 8 GB of memory
 - 4 GB per Storage Processor
 - Configurable read and write cache size

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 5

Mercer Road
EMC Education Services

Taking a high-level look at the architecture you will see why CLARiiON is so successful. Today, there are over 165,000 that have been sold worldwide.

Superior Availability: CLARiiON's design has no single point of failure—all the way down to the fans and power cords. It's designed for non-stop operation—software upgrades and hardware reconfigurations can be done while online. It's always checking the integrity of its hardware and software. Potential problems are escalated proactively to EMC through its ClarAlert phone home capability.

Scalability: A single CLARiiON can go from a complete 3/4 TB storage system in about 30 centimeters of rack space, all the way up to a 240 drive, 58 TB configuration (using Fibre Channel and ATA spindles).

Performance: CLARiiON was the first full Fibre Channel storage system. That leadership translates today into superior performance.

Flexibility: Customers can mix and match performance, capacity, cost, and data protection schemes, all in the same unit. Unlike storage systems that come from server vendors, CLARiiON is unique at supporting multiple operating systems and multi protocols while offering full interoperability between different servers, all at the same time. The ability to choose between Fibre Channel and ATA drives is only one example of CLARiiON's flexibility.

Replication: CLARiiON also supports storage-based local and remote data replication for backup and disaster recovery through its SnapView and MirrorView software.

CLARiiON CX Series

- Sixth generation, full Fibre Channel networked storage running FLARE Operating Environment
- Flexible connectivity and bandwidth
 - Up to 8 FC-AL or FC-SW host ports
 - 1 Gb/2 Gb Fibre Channel host connections
- Scalable processing power
 - Dual or Quad-processors supporting advanced storage-based functionality
- Industry-leading performance and availability
 - 1 GB, 2 GB, 4 GB, or 8 GB memory and dual or quad redundant 2 Gb back-end storage connections
- Cross-generational software support
- Non-disruptive hardware replacement and software upgrades



The market for mid range storage is growing year to year. With the CLARiiON CX series arrays, EMC is poised to become a major player in the mid tier space.

The EMC CLARiiON line of networked storage systems provides a full range of options in price, performance, connectivity, capacity, and more to suit your business and IT needs. EMC CLARiiON series offer entry level arrays with SATA disk drives to Enterprise arrays that scale to 240 drives, supporting FC or iSCSI.

EMC Education Services

EMC²
where information lives™

Modular Building Blocks

- Disk Array Enclosure (DAE)
 - CX family uses DAE2 with up to (15) 2Gb FC Drives
 - FC family uses DAE with up to (10) 1Gb FC drives
 - DAE2-ATA contains up to 15 ATA drives (Advanced Technology Attachment)
- Disk Array Enclosure (DAE2P)
 - CX family only on 300, 500, and 700 series
 - Replacement for DAE2 with code upgrade
 - Houses up to 15 2GB FC Drives
- Disk Processor Enclosure (DPE)
 - Some CX series use DPE2 that contains Storage Processors and up to (15) 2Gb FC Drives
 - FC family uses DPE that contains Storage Processors and up to (10) 1Gb FC Drives
- Storage Processor Enclosure (SPE)
 - Contains two dual CPU Storage Processors and no drives
- Standby Power Supply (SPS)
 - Provides battery backup protection



© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 7


EMC Education Services

The CLARiiON storage system is based on a modular architecture. The concept is “buy and build as you go”. The first building block of the architecture is the Disk Processor Enclosure, or DPE. The DPE houses the storage Processor(s) and the first Fibre Channel Arbitrated Loop (FC-AL) disks. Some older CX series arrays use the DPE2 which includes, in addition to the storage Processors, up to 15 2Gb dual-ported FC drives. Newer CX series namely the CX300, CX500 and CX 700 can use a DAE2P which provide for better fault isolation and reliability on the back end (BE) loops. The FC series DPE contains up to 10 1Gb dual-ported FC drives. In order to expand the capacity of the storage system, Disk Array Enclosures (DAEs) are interconnected using Link Control Cards (LCCs).

Newer CLARiiON arrays have two processors per Storage Processor, and do not use a DPE. Instead, it utilizes an SPE or Storage Processor Enclosure. The SPE does not contain any disk modules so it must have at least one DAE2 and a maximum of 16 DAE2s.

DAEs typically exist in the same cabinet, or can be placed in one or more separate cabinets. The DAE2P requires a new Telco Rail kit and cables (HSSDC/HSSDC2) The DAE is the sheet metal enclosure and mid plane plus the following Field Replaceable components:

- 2 Link Control Cards (LCCs)
- Up to 15 disk modules
- 2 power supplies
- Two fan modules (part of Power Supply enclosure)

CLARiiON Cabinet Enclosure

- The CLARiiON Cabinet Enclosure is a pre-wired 19-inch-wide and can accept up to 40 NEMA units of storage system enclosures
 - NEMA Unit (U) is 1.75 inches tall
 - SPE= 4U
 - DPE2 = 3U
 - DAE2= 3U
 - DAE2P=3U
 - SPS = 1U
 - Fibre Channel Switches = 1U or 2U
- Each rack may include a mixture of storage system enclosures, as required. DAE2Ps require a Telco rack
- Multiple CLARiiON storage systems may be housed in a single rack



The DPE or SPE is what defines a CLARiiON storage system, not the rack or frame. In other words, if three DPEs are installed within one rack or frame, then there are three separate CLARiiONs within that rack or frame.

CLARiiON ATA (Advanced Technology Attachment)

- Lower \$/MB for backup or bulk storage
- Alternative to Fibre Channel HDAs
- Same Software capability as FC
- Uses FC interconnect
 - Mix FC and ATA enclosures
 - First shelf must be Fibre Channel
- Full HA features
 - Dual-ported access
 - Redundant power and LCCs
 - Hot swap capability



© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 9

Mercer Road
EMC Education Services

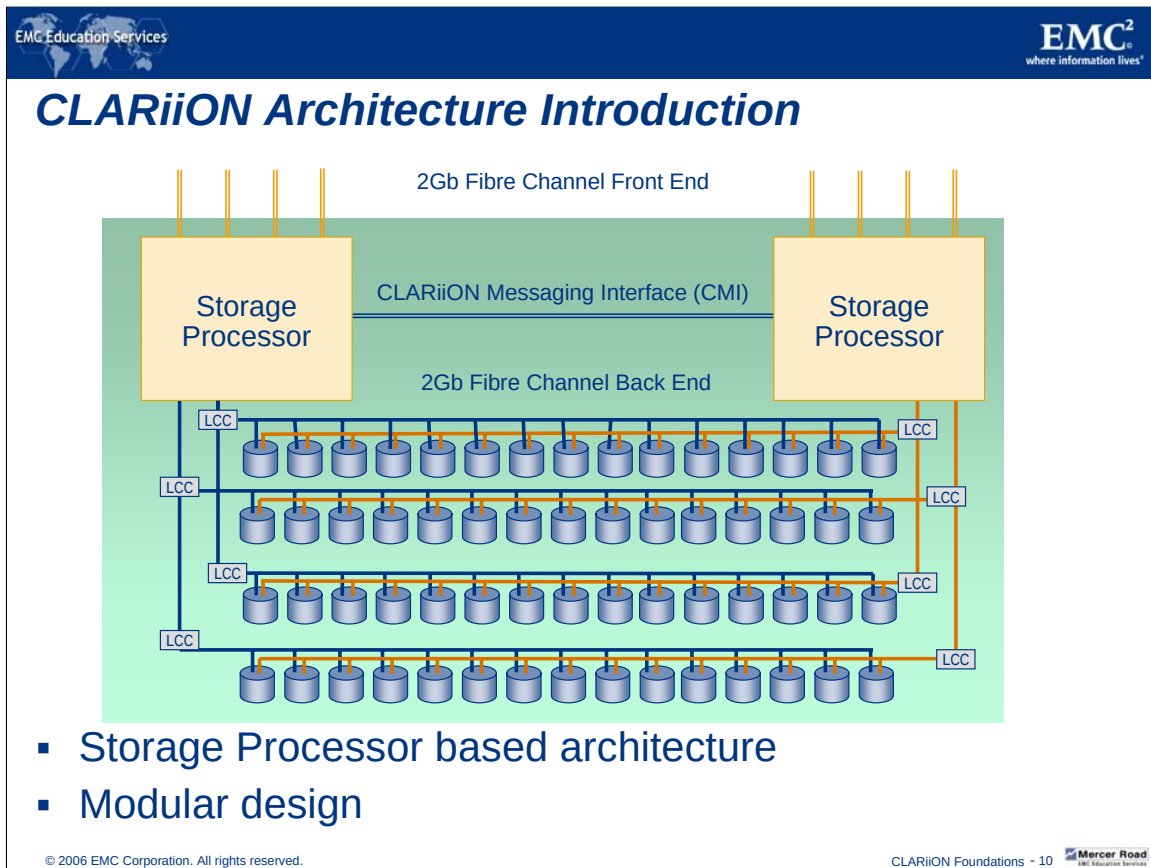
ATA disks are low-cost disk drives typically found in desktop PCs. ATA disk drives are ideal for applications requiring high capacity and lower performance levels than FC drives. IDE/ATA drives represent the vast majority of all drives manufactured worldwide, and therefore can be offered at a lower price point. They are also significantly slower than Fibre Channel disks in both rotational and seek speeds. CLARiiON supports mixing and matching of ATA DAEs and Fibre Channel DAEs in the same storage system. All CLARiiON software and functionality is supported with ATA drives. CLARiiON with ATA is targeted for backup applications with backup-to-disk functionality. CLARiiON with ATA also lowers the hardware implementation cost for SnapView BCVs (Clones) and MirrorView target sites.

Major ATA drive specification differences include:

- Rotational speed of 5400 rpm vs. 10,000 rpm
- Rotational latency of 5.5 ms vs. 2 to 2.99 ms
- Average seek times of 12 ms vs. 3.6 to 5.3 ms
- Transfer rates of 26 to 50 MB/s vs. 26 to 78 MB/s

There is anywhere from <25% to 50% difference in performance.

	36 GB 15K	73 GB 10K	146 GB 10K	250 GB ATA
Rotational Speed	15,000	10,000	10,000	5,400
Rotational Latency	2 ms	2.99 ms	2.99 ms	5.5 ms
Average Seek (Read)	3.6 ms	5.2 ms	4.7 ms	~ 12 ms
Average Seek (Write)	4.2 ms	6.0 ms	5.3 ms	13 ms
Transfer Rates (Internal)	51- 69 MB/s	26.7- 40.2 MB/s	43- 78 MB/s	26 - 50 MB/s
Transfer Rates (External)	200 MB/s	200 MB/s	200 MB/s	133 MB/s



The module architecture allows the customer to add drives as needed to meet capacity requirements. When more capacity is required, additional disk array enclosures (DAE or DAE2) containing disk modules can be easily added. LCC or Link Control Cards are used to connect shelves of disks. In addition, the LCC monitors the FRUs within the shelf and reports status information to the storage processor. The LCC contains bypass circuitry that allows continued operation of the loop in the event of port failure.

CLARiiON Architecture is based on intelligent Storage Processors that manage physical drives on the back end and service host requests on the front end, be it Fibre Channel or iSCSI protocols. Depending on the module, each Storage processor includes either one or two CPUs. Storage Processors communicate to each other over the CLARiiON Messaging Interface (CMI). Both the front-end connection to the host and the back-end connection to the physical storage is 2Gb Fibre channel.



CLARiiON Foundations

THEORY OF OPERATION

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 11 

Next, we will look at the theory of operations.

EMC Education Services

EMC²
where information lives™

Storage Processor Introduction

- Storage processors are configured in pairs for maximum availability
- One or two processors per Storage Processor board
- Two or four Fibre Channel front-end ports for host connectivity
 - 1Gb or 2Gb
 - Arbitrated loop or switched fabric
- Dual-ported Fibre Channel Disk drives at the back-end
 - Two or Four Arbitrated Loop connections
- Maximum of 4GB of memory per SP
 - Write Cache is mirrored between Storage Processors for availability using the CMI (CLARiiON Messaging Interface)
 - Write Caching accelerates host writes
- Ethernet connection for management

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 12 Mercer Road
EMC Education Services

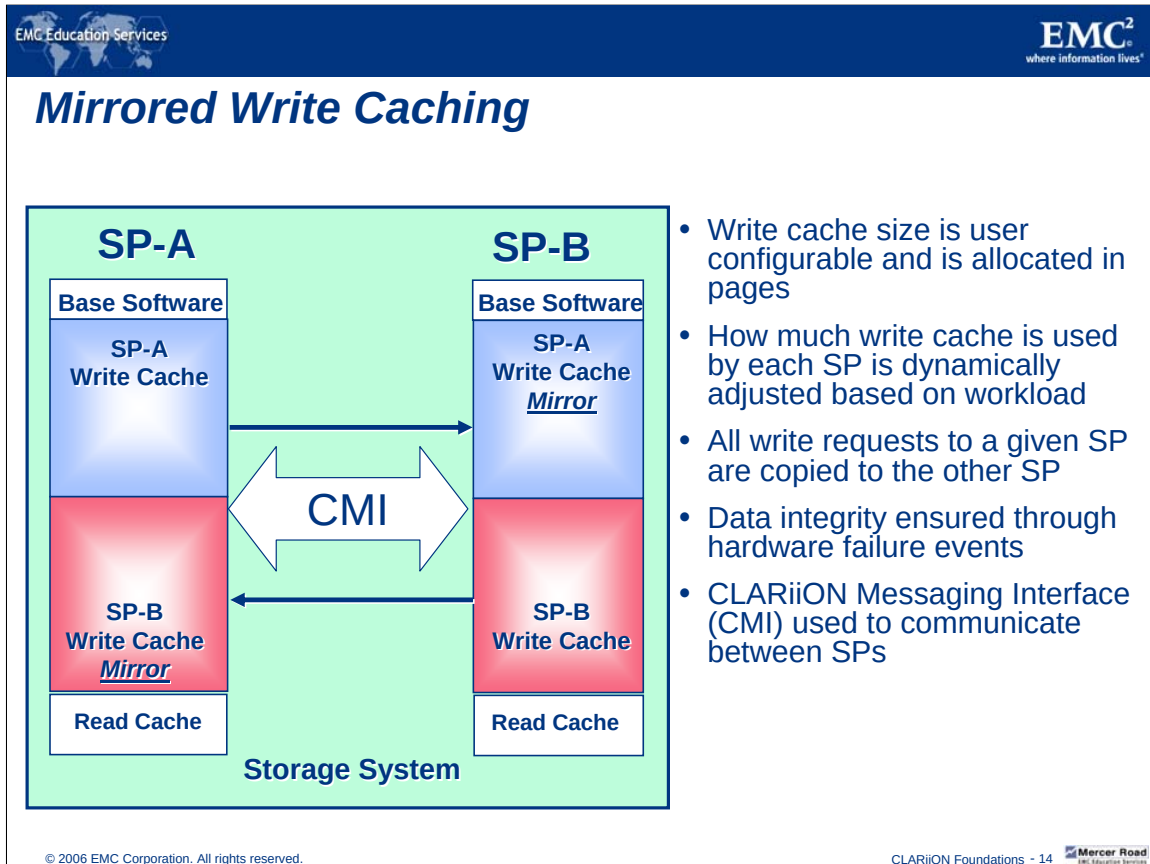
The main component in all CLARiiON series arrays is the Storage Processor. Storage Processors (SPs) are configured in pairs for maximum availability and are Field Replaceable Units (FRUs). SPs provide both front-end connectivity to the hosts and back-end connectivity to the physical disks. Each Storage Processor also includes up to 4 GB of memory, most of which is used for cache. Cache memory is segmented into read cache memory and write cache memory. Read cache memory is used for staging and prefetching read requests from the host. Write cache is used to accelerate host writes to the storage system. With write cache enabled, writes are mirrored to the write cache memory in the other storage processor over the CLARiiON Messaging Interface (CMI). The CMI is a Fibre Channel based link and operates at either 100MB/sec on FC series or 200MB/sec on CX series systems. Each storage processor also includes a TCP/IP connection that is used for configuration and management of the storage system.

Each storage system ships with a complete copy of FLARE software installed on the first four disks on back-end loop 0. Disks 0_0 and 0_2 store mirrored copies of the software for SP A, and disks 0_1 and 0_3 store mirrored copies of the software for SP B. When you power up the storage system, each SP boots and executes FLARE software.

Cache Memory

- Cache memory on an SP performs two tasks
 - Staging: Temporary buffering of current read and write data
 - Always performed on each I/O
 - Storage: Repository for frequently accessed data
 - Maintaining copies of read and write data (uses least-recently used algorithm to manage storage cache)
 - User must explicitly enable this (for both read and write)
- Benefits of caching
 - Burst Smoothing - Absorb bursts of writes without becoming “disk bound”
 - Write cache optimization
 - Locality - Merge several writes to the same area into a single operation
 - Increases write performance
 - Immediacy - Satisfy user requests without going to the disks
 - Read cache optimization prefetching of data for sequential reads

The CLARiiON uses cache in a manner similar to how cache memory is used in traditional workstations and servers. With servers and workstations, application code and application data locality and immediacy are the focus of the design. A characteristic of many applications (including file systems) is to cache new data locally, then periodically flush the data to the actual storage device. This “lazy write” approach would result in large bursts of large I/Os to the storage systems - a perfect fit with the “burst smoothing” benefit of CLARiiON’s caching. Cache page size is perhaps the most influential parameter on cache performance. The division between staging and storage memory is not formally defined by addresses, but more by functionality. As data arrives, it is always staged, and as needed, it may be marked for storage as well. The Partition Memory dialog in Navisphere Manager enables part of the cache to be used for storage operations. Unlike the Symmetrix, certain writes may bypass cache memory.



The amount of cache memory allocated to write operations can be dynamically adjusted, based on workload. Navisphere Analyzer can be used to baseline actual workloads and to understand efficiency of cache utilization. When write cache is enabled, a write operation received by one storage processor is sent to the other storage processor and a confirmation that it was successfully stored is returned before write complete is sent to the host. The mirroring insures data is protected from a single SP failure. The connection between the SPs, called the CMI, carries data and status information between SPs. On FC-series arrays, the CMI is a single connection, running at 100 MB/s. On CX-series arrays, the CMI is a dual connection, with each connection running at 200 MB/s.

Managing Write Cache

- To maintain free space in write cache for I/O bursts, pages are flushed from cache to the drives
 - Only the least-recently used pages are flushed
- Three levels of flushing:
 - Idle - Low I/Os to the LUN; user I/Os continue
 - Watermark - Priority depends on cache fullness; user I/Os continue
 - Forced - Cache has no free space; user I/Os queue
- For maximum performance:
 - Provide a “cushion” of unused cache for I/O bursts
 - Minimize/avoid forced flushes

The flushing algorithm is usually referred to as the Least Recently Used (or “LRU”) algorithm. The LRU approach focuses on write cache immediacy (as the most recent used data remains available), and less on locality (as the location of the data is considered only to produce larger flushes of data).

Idle flushing keeps some free space in write cache when I/O activity is relatively low. If data immediacy were most important, idle flushing would be sufficient. If idle flushing cannot maintain free space, though, forced flushes create space for new I/Os. Forced flushes maintain the desired free space, though it will impact overall performance as all read and write operations are halted to clear space in the write cache.

Write Cache Protected by “Vault”

- The “vault” is a reserved area found on specific protected disks
- At the first sign of an event which could potentially compromise the integrity of the data in write cache, cache data is dumped to the vault area
- After the data is dumped to the vault, it will be migrated to the LUNs where it belongs
 - When power is restored data is migrated from the vault back to cache (if an SPS has a charged battery)
- Other failures such as a SP or Cache failure will disable write cache

The vault is a reserved area found on the first 9 disks of the DPE on the FC series and the first 5 drives on a CX series system. For this reason, only the DPE needs to be kept alive in the event of a power failure. At the first sign of an event which could potentially compromise the integrity of the write cache data, that data is dumped to the vault area. It is protected there by the non-volatile nature of disk storage, as well as the RAID level (RAID-3) implemented for the vault.

After the data is dumped to the vault, it will be migrated to the LUNs where it belongs. Until all data has been migrated, I/Os check the vault because the most recent copy of data may be located there. Once the migration process is complete, the vault is marked as empty, and I/Os need no longer check it.

Persistent Storage Manager (PSM)

- PSM is a hidden LUN that records configuration information
- Both SPs access a single PSM so environmental records are in sync
 - If one SP receives new configuration info, that info is written to the PSM and the other SP instantaneously updates itself
- If one SP needs to be replaced, the new one can easily find the unique environmental information on the PSM
- Enables SPs to be completely field-replaceable

The PSM is created at initialization of the array via Navisphere, and resides on the first loop of drives. The default size is 1GB on CX systems, 512 MB on the FC4700. This LUN is hidden from end-user view in the CX series, but may be seen if using Navisphere Manager on an FC4500 or FC4700 series array.

Host Information	Array Information	Other
Drive Mapping	SP IP Addresses	Polling Rates
Privileged Users (Host)	Privileged Users (Array)	Core OS & FW Revisions
Host Information	ALPAs (Arbitrated Loop Physical Address)	
Initiator ID information	Pre-upgrade Packages (Snap, Core OS, etc.)	
	Pre-upgrade Packages (Snap, Core OS, etc.)	

EMC Education Services

EMC²
where information lives™

Data Units on a Disk

- **Sector**
 - 520 Bytes
 - 512 bytes of user data
 - 8 bytes of administrative data
- **Element**
 - Element size is the number of blocks written to single disk in a RAID group before moving to the next disk
- **Stripe**
 - A group of sequential elements starting with the first disk in a RAID group and ending with the last, including any parity or mirrored elements.

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 18

Mercer Road
EMC Education Services

On the CLARiiON, sectors are formatted to 520 bytes each. Of those, 512 bytes are dedicated to host data, and the other 8 bytes are used for CLARiiON housekeeping purposes. The final 2 bytes are the LRC, or Longitudinal Redundancy Checksum, which allows CLARiiON to detect errors as small as a single bit on a sector.

One or more sectors make up an element, which is the largest amount of data which can be written to a disk before moving to the next disk.

A stripe consists of one element per disk across all the disks.

Element Tips

- **Element size is defined for each LUN**
 - Large elements are more efficient for reading from disk
 - Good for sequential data
 - Smaller elements scatter the data across the RAID group
 - Good for random data
 - Element size cannot be changed without unbinding the LUN
 - Unbinding a LUN destroys the data
 - Default element size of 64K is optimal for most applications
- **Read and write operations that cross an element boundary degrade performance**
 - Element size set as a multiple of the average unit of data requested
 - For 2k I/O size - element should be 2k, 4k, 8k,...etc
 - Monitor LUNs using Navisphere Analyzer

The element size is configurable on a per LUN basis and must be set up when the LUN is created.

To a large degree, the element size will determine the performance of the LUN in the environment. For RAID level 5, where the disks can operate independently, small element sizes are most efficient when I/Os are random. The data is then scattered across all the disks, and the chances are good that multiple disks can be used simultaneously to fetch data. In the case of data which is largely sequential, large element sizes are written to contiguous areas of the disk and therefore result in little or no head movement. This increases performance – one of the major factors that influences performance is the seek time of the disks.

If the average I/O is a multiple of the element size, then each read or write operation will involve 2 or more disks, and this will reduce performance, particularly write performance. Ideally, the element size should be an even multiple of the average I/O size.

Navisphere Analyzer tracks these events, and calls them “disk crossings” or “stripe crossings”. If the percentage of disk crossings is high for a LUN, it may be an indication that the element size needs to be increased. Bear in mind that this process is destructive.

These parameters should only be changed by experienced and trained individuals, if advised to do so by CLARiiON Performance Engineering.



CLARiiON RAID Options

Step 0 - Planning

- Disk
 - No protection
 - JBOD
- RAID-0: Stripe
 - No protection
 - Performance JBOD
- RAID-1: Mirroring
 - Some performance gain by splitting read operations
 - Protection against single disk failure
 - Minimum performance hit during failure

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 20 

RAID-0 spreads the data across the drives in the array, one segment at a time. Characteristics include Data spreading across the drives in the array by system block size. Data is striped across all of the drives in the RAID Group without generating any redundant data. A RAID-0 Group consists of three to sixteen disk modules. The number of drives and groups is scaleable. In RAID-0, there is no error recovery over or above what is normally on a disk drive. Advantages include high I/O rate (small block size) and transparent to host system software. Disadvantages include a somewhat slower transfer rate than RAID 3 and no parity disk or redundant disk.

RAID Level 1 is data mirroring. Any time data is written to the LUN, a write occurs to a mirror disk. This is done transparently to the attached host. Characteristics include a traditional approach for improving reliability of disk storage as this technology is well understood. It is an expensive option because all disk drives are duplicated. Every write to a data disk is also a write to the mirror disk. A RAID-1 group consists of 2 disk modules. Data on a replaced drive is rebuilt from the mirror drive. In terms of Error Recovery, in RAID Level 1, if a disk fails, the disk array controller uses the mirror drive for data recovery and continuous operation. Advantages include high data availability and high I/O rate (small block size). Disadvantages include a total number of disks in the array equaling 2 times the data (useable) disks. The overhead cost equals 100%, while usable storage capacity is 50%. Applications that write sequentially in small I/Os like Microsoft Exchange; transaction logs are a good choice for RAID-1 disks.

CLARiiON RAID Options

Step 0 - Planning

- **RAID-1/0: Striped Mirrors**
 - Performance of stripes combined with split read operations
 - Protection against single disk failure
 - Minimum performance hit in failure mode
- **RAID-3: Striped Elements**
 - Each data element striped across disks - parity kept on the last disk in the RAID group
 - Extremely fast read access from the disk
 - Used for streaming media
 - Parity protection against single disk failure
 - Performance penalty during failure

RAID Level 1/0 is essentially a striped RAID-1 group. It combines the speed advantage of RAID-0 with the redundancy advantage of RAID-1. This is a more expensive option compared to other parity protection schemes, because all disk drives are duplicated. Every write to a data disk is also a write to the mirror disk. RAID-1/0 Groups consists of 4, 6, 8, 12, 14 or 16 disk modules. In RAID Level 1/0, if a disk fails, the disk array controller uses either its mirror drive or the data drive for data recovery and continuous operation. If there is a hot spare, data is rebuilt onto the hot spare from the mirrored copy. When the failed disk is replaced, data from the other mirror (copy) is used to rebuild the replaced disk. Advantages include high data availability, high I/O rate (small block size), and the ability to withstand multiple drive failures as long as they occur on different mirrors. Disadvantages include that the total number of disks equals two times the data disks, with overhead cost equaling 100%. RAID-1/0 is a good solution for sites that normally have RAID-1 but need the added performance boost that RAID-1/0 provides.

RAID Level 3 is an array of disk drives transferring data in parallel. RAID-3 combines 5 or 9 disk drives to act as a large virtual drive with high throughput (two or more times that of a conventional drive). A parity drive is added to the Group to store parity information. Information on the parity drive allows data to be reconstructed, should a drive fail. Parity is the Exclusive_OR (XOR) of data on drives 1, 2, 3, and 4. Parallel data paths are provided to the drives. With Error Recovery, if a disk fails, the controller dynamically recovers the drive's data, using data on the other data drives and the parity drive. The controller uses data from the good drives in the normal positions within the data block. The controller combines the first byte of each sector in the good drives and the parity drive to regenerate the first byte of the sector on the failed drive. This process is repeated to generate each byte of the failed drive. Advantages include total number of disks equals 1.25 times the data disks for group of 5, overhead cost is 20% for group of 5, and a high bandwidth (large block size). There is good throughput on large data transfer. Disadvantages include poor media efficiency of small blocks. RAID Level 3 is not good for transaction processing systems. Data is lost if multiple drives fail within the same RAID-3 Group. RAID-3 is a good choice for applications that require moderate to high I/O that is sequential, such as MirrorView targets, Clones, and on-line backups.

CLARiiON RAID Options

- RAID-5: Striping with Parity

- Performance of striping
- Protection from single disk failure
- Parity distributed across member drives within the RAID Group
- Write performance penalty
- Performance impact if a disk fails in RAID Group

Step 0 - Planning

- Hot Spare

- Takes the place of failed disk within a RAID group
- Must have equal or greater capacity than the disk it replaces
- Can be located anywhere except on Vault disks
- When failing disk is replaced, the hot spare restores the data to the replacement disk and returns to the hot spare pool

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 22 

RAID-5 is an independent disk array, meaning it does not read and write data to parallel disks like RAID-3, but it performs independent or parallel read and write operations. There is no dedicated parity drive. The result is a transfer rate equal to that of a single drive, but with a high overall I/O rate.

Characteristics: Independent data paths to the drives. A RAID-5 Group usually consists of 5 disk modules. The number of drives and groups is scaleable. Data is striped by system block size and distributes the data and parity information per sector across all disks (no parity disk).

In RAID Level 5, if a disk fails, the controller rebuilds its data from the data and parity from the remaining drives. For example, to rebuild sector 0 on drive 5, the controller uses the parity information from drive 1 and data block information from drives 2, 3, and 4. The controller combines the information of each sector in the remaining drives and the parity information to regenerate the first block of the sector on the failed drive. The process is repeated to generate each block of the failed drive.

Advantages include it being good for parallel processing (multi-tasking) applications or environments. Also, the total number of disks equals 1.25 times the number of data disks for a group of five, or 1.1 times the number of data disks for a group of 10, while the overhead cost is 20% for groups of 5 and 10% for groups of 10. It has a high I/O rate (small block size).

Disadvantages include somewhat slower transfer rate than RAID-3. There is degradation in performance in recovery and reconstruction modes and data loss if multiple drives within the same group are lost.

RAID-5 offers great price/performance.

Which RAID Level Is Right

Step 0 - Planning

- RAID 0 – Data striping
 - No parity protection, least-expensive storage
 - Applications using read-only data that require quick access, such as data down-loading
- RAID 1 – Mirroring between two disks
 - Excellent availability, but expensive storage
 - Transaction, logging or record keeping applications
- RAID 1/0 – Data striping with mirroring
 - Excellent availability, but expensive storage
 - Provides the best balance of performance and availability
- RAID 3 – Data striping with dedicated parity disk
- RAID 5 – Data striping/parity spread across all drives
 - Very good availability and inexpensive storage
- Support mixed types of RAID in the same chassis

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 23 

Use a RAID 0 group where the best overall performance is important. A RAID 0 group is less available than an individual unit. A RAID 0 group is useful for applications using short-term data to which you need quick access.

Use a RAID 1 group for logging or record keeping applications because it requires fewer modules than a RAID 0 group and provides high availability and fast write access. Or you could use it to store daily updates to a database that resides on a RAID 5 group, and then, during off-peak hours, copy the updates to the database on the RAID 5 group.

A RAID 1/0 group provides the best balance of performance and availability. You can use it very effectively for any of the RAID 3 or RAID 5 applications. A RAID 1/0 group requires a minimum of four disks.

Use a RAID 3 group with a single-task application that uses large I/O transfers (more the 64 Kb).

Use a RAID 5 group for a database repository or database server that uses a normal or less-than-normal percentage of write operations (writes are 33% or less of all I/O operations). Use a RAID 5 group where multi-tasking applications perform I/O transfers of different sizes. Write caching can significantly enhance the write performance of a RAID 5 group.



Storage Configuration and Provisioning

- Understanding application and server requirements and planning configuration is critical!
- RAID Group is a collection of physical disks
 - RAID Protection level is assigned to all disks within the RAID group
- Binding LUNs is the creation of Logical Units from space within a RAID Group
- Storage groups are collections of LUNs that a host or group of hosts have access to

Step 0 - Planning

Step 1 – Create RAID Groups

Step 2 – Bind LUNs

Step 3 – Create Storage Groups

Step 4 – Add LUNs to Storage Groups

Step 5 – Connect Hosts with Storage Groups

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 24 

The first step in configuring a CLARiiON Storage System is to understand application and server requirements as they pertain to availability, capacity, access control, and other features such as clustering and local and remote replication. While this step is only mentioned here, it will take more time and effort than all the other steps combined. If this step is skipped, reconfiguration will most likely be required.

In Step 0, you will have identified the availability and performance goals and, from that, should have determined the appropriate RAID level. A RAID Group is a collection of physical disk that share the same RAID protection scheme or RAID level. The actual protection scheme does not get set until you “bind” the first LUN.

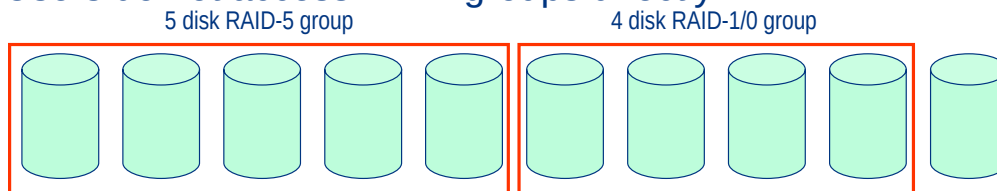
In step 0, you will also have determined the number and size of Logical Units (LUNs) that are required. Each RAID group can contain one or more LUNs.

Access control is implemented in the CLARiiON using the Access Logix code that runs in the Storage Processor. Access Logix uses the concept of Storage Groups. LUNs are assigned to a Storage Group and hosts are associated with the Storage Group to allow access. Again, planning is critical here as well. In Step 0, you should have identified which hosts need access to which LUNs and any special requirements such as shared access that would be required if the hosts were part of a cluster.

Creating RAID Groups

- RAID protection levels are set through a RAID group
- Physical disks part of one RAID group only
 - Drive types cannot be mixed in the RAID Group
- May include disks from any enclosure
- RAID types may be mixed in an array
- RAID groups may be expanded
- Users do not access RAID groups directly

Step 1 – Create
RAID Groups



© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 25

Mercer Road
EMC Education Services

The physical disks on the CLARiiON are configured into RAID groups. Any RAID Group should consist of all FC disks or all ATA disks, but NOT a mix of both FC and ATA.

Binding a LUN

- Binding is the process of building LUNs onto RAID Groups
- May be up to 128 LUNs in a RAID Group
- May be up to 2048 LUNs per CLARiiON array
- LUNs are assigned to one SP at a time
 - The SP owns the LUN
 - The SP manages the RAID protection of the LUN
 - The SP manages access to the LUN
- The LUN uses part of each disk in the RAID Group
 - Same sectors on each disk

Step 2 – Bind LUNs

A logical unit (LUN) is a grouping of one or more disks into one span of disk storage space. A LUN looks like an individual disk to the server's OS. It has a RAID type and properties that define it. The process of building a LUN onto a RAID Group is called binding. When you bind a LUN on a RAID Group, you specify how much of the Group's user space (contiguous free space) you want the LUN to use. The LUN is distributed equally across all the disks in the RAID Group.

The characteristics of a LUN are:

- 128 LUNs per RAID Group is the maximum for all RAID types
- Up to 2048 LUNs per array depending on the model
- An SP owns a LUN, and different SPs can own different LUNs in the same RAID Group
- All LUNs in the RAID Group have the same RAID type
- Each LUN in a RAID Group can have a different element size where applicable

Bind Operation - Setting Parameters

Step 2 – Bind LUNs

- **Fixed Bind parameters**
 - Disk numbers, RAID type, LUN #, element size
 - Can't be changed without unbinding and rebinding
- **Variable parameters**
 - Cache enable, rebuild time, verify time, auto assignment
 - Can change without unbinding
- **Bind Operation**
 - Fastbind is the almost instantaneous bind achieved on a factory system

Once the LUN has been bound to a RAID Group, its characteristics can be viewed in the properties screens. From there, the variable parameters, such as cache enable, rebuild time, verify time, and auto assign can be changed without unbinding and rebinding the LUN.

If the fixed parameters, such as Disk numbers, RAID type, LUN #, or element size needs to be changed, the LUN must be unbound and rebound. Unbinding a LUN causes data loss.

When the array arrives from the factory, the disks are described as “Zeroed”. This means that the “pointer” is set to the zero position on the physical platter of the drive. If a LUN is bound to the disk, the pointer is moved to the end position of that LUN. If the LUN is deleted, the pointer is not reset and the next LUN will be started from the end position of the deleted LUN.

However, the bind operation will have to zero out all the space previously used by the now deleted LUN. This means that a LUN bind to a modified system takes a lot longer than the “fastbind” of a factory system. This is important from a customer perspective when new LUNs are being created and system downtime is critical. The pointers can be reset from the CLI, but the operation takes the same amount of time as the zeroing of the deleted LUN space. This functionality makes it very important to do LUN planning and implementation carefully upfront when implementing a new frame.

Storage Groups

- Storage Groups are a feature of Access Logix and used to implement LUN Masking
 - Storage Groups define the LUNs each host can access
 - A Storage Group contains a subset of LUNs grouped for access by one or more hosts and inaccessible to other hosts
 - Without Storage Groups, all host can access all LUNs
 - Storage groups can be viewed as a “bucket” that has dedicated and/or shared LUNs accessible by a server or servers
 - Access Logix controls which hosts have access to a storage group
 - Host access the array and provide information through the Initiator Registration Records process
- Storage Group planning is required

Step 3 – Create
Storage Groups

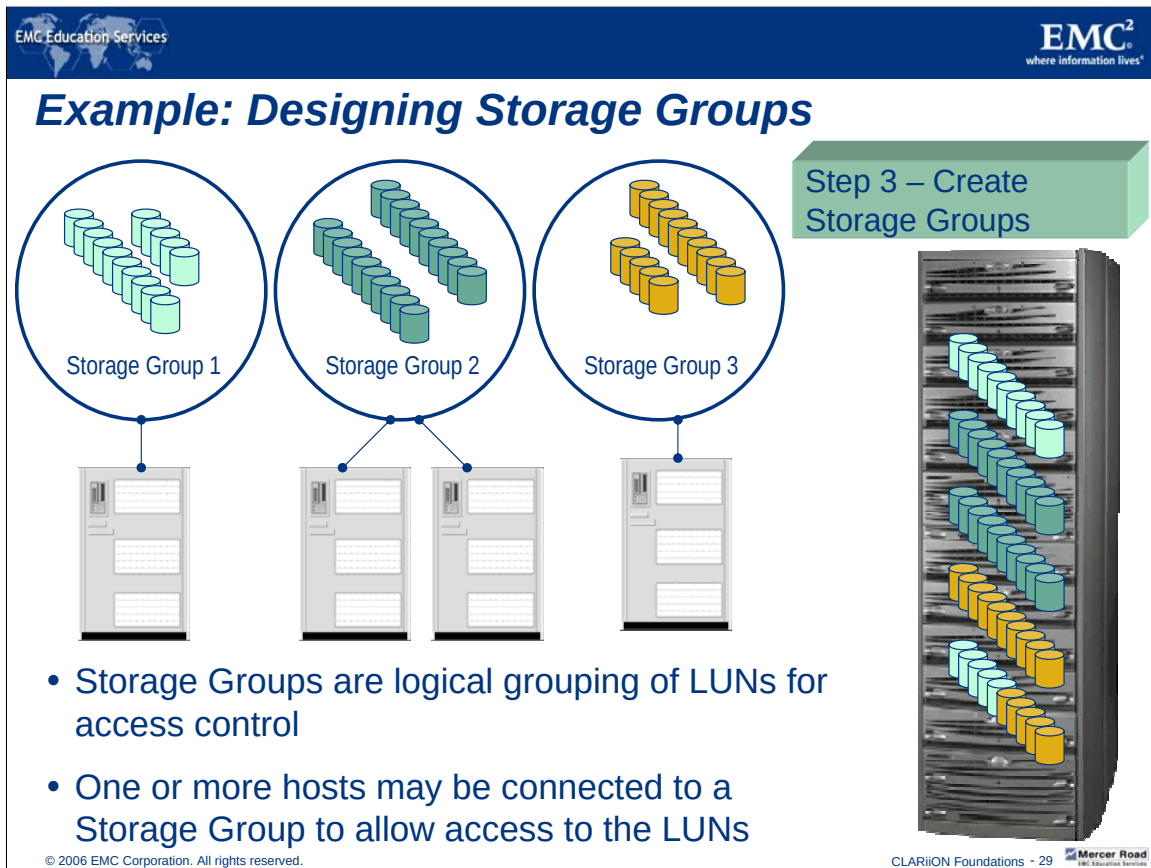
Storage groups are another key feature. Without Storage Groups, all hosts can access all LUNs. These Storage Groups provide Data Access Control by defining which LUNs each host can access.

Storage Groups are a combination of LUNs assigned to a particular host.

An initiator can only be connected to a single Storage Group on any storage system. The exception would be in a clustered environment where 2 or more hosts share the same storage group.

There is also a new LUN mapping table applied to the LUN numbers in Standard SCSI commands. Each entry includes the LUN alias, or the number presented to the host and the FLARE LUN number which is created during binding.

These pieces are key to Access Logix and shared environments.

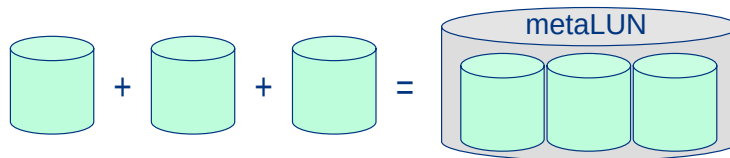


Here is a picture of what we're talking about with storage groups and how they can be set up. Unless all hosts require access to the same LUNs, it is easier to create a storage group per server. It will allow the flexibility of adding and removing LUNs on the fly.

In the graphic above, Storage Group 2 is being accessed by two servers. While this is an acceptable configuration, it is usually only recommended when the servers are in a clustered environment. An example of this is when a NAS head is front-ending a CLARiiON array. Additionally, the LUNs in each of the Storage Groups need not be unique. For example, LUNs 1, 2, and 3 could be present in both SG 1 and in SG 3. Care should be exercised when doing so since access to the same volumes, at the same time, may corrupt the data.

metaLUNs

- A metaLUN is created by combining LUNs
 - Dynamically increase LUN capacity
 - Can be done on-line while host I/O is in progress
 - A LUN can be expanded to create a metaLUN and a metaLUN can be further expanded by adding additional LUNs
 - Striped or concatenated
 - Data is restriped when a striped metaLUN is created
- Appears to host as a single LUN
 - Added to storage group like any other LUN
 - Can be used with MirrorView, SnapView, or SAN Copy
- Supported only on CX family with Navisphere 6.5+



© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 30 

MetaLUNs are supported only on CX-Series storage systems.

The metaLUN feature lets you dynamically expand the capacity of a single LUN (base LUN) into a larger unit called a metaLUN. You do this by adding LUNs to the base LUN. You can also add LUNs to a metaLUN to further increase its capacity.

Like a LUN, a metaLUN can belong to a Storage Group, and can participate in SnapView, MirrorView and SAN Copy sessions.

During the expansion process, the host is able to process I/O to the LUN or metaLUN and access any existing data on the Base LUN. It does not, however, have access to any added capacity until the expansion is complete. Depending on the operating system, a reboot of the host or the use of a disk expansion utility, such as diskpart in Windows, may be used to claim the additional space.

Each set of striped LUNs is called a component. All metaLUNs contain at least one component which includes the base LUN and one or more LUNs. Any data that gets written to a metaLUN component is striped across all the LUNs in the component. You can expand a LUN or metaLUN in two ways — stripe expansion or concatenate expansion. A stripe expansion takes the existing data on the LUN or metaLUN, and restripes (redistributes) it across the existing LUNs and the new LUNs you are adding. The stripe expansion may take a long time to complete and will affect performance while the expansion is in process. A concatenate expansion creates a new metaLUN component that includes the new LUNs and appends this component to the end of the existing LUN or metaLUN. There is no restriping of data between the original storage and the new LUNs. The concatenate operation completes immediately.

Advanced Availability: LUN Ownership Model

- Only one Storage Processor “owns” a LUN at any point in time
 - Assigned when LUN is created but can also be changed using Navisphere Manager or CLI
- If the Storage Processor, Host Bus Adapter, cable, or any component in the I/O path fails, ownership of the LUN can be moved to the surviving SP
 - Process is called LUN Trespassing
 - CLARiiON originally used host based ATF (Application Transparent Failover) to automate path-failover. Today, EMC PowerPath provides this function
- For maximum availability, careful design of I/O path for no Single-Point-Of-Failure is required

CLARiiON uses the single ownership model. At any time, only 1 SP owns a LUN, although this ownership may be changed easily with Navisphere Manager or the Navisphere CLI.

When optional software such as ATF or PowerPath is used on the host, temporary changes can be automatically made when certain failures occur. An SP failure causes all LUNs to be failed over to the surviving SP.

EMC Education Services

EMC²
where information lives™

Host Connectivity Redundancy PowerPath – Failover Software

- Host resident program for automatic detection and management of failed paths
- Host will typically be configured with multiple paths to LUN
- If HBA, cable or Switch fails, PowerPath will redirect I/O over surviving path
- If Storage Processor fails, PowerPath will “Trespass” LUN to surviving Storage Processor and redirect I/O
- Dynamic load balancing across HBA and Fabric – Not Storage Processors

The diagram illustrates the PowerPath architecture for host connectivity redundancy. At the top, an 'Application' box connects to a 'PowerPath' box. Below 'PowerPath', two 'Request' boxes are shown. The left 'Request' box connects to an 'sd' box, which then connects to an 'HBA' box. The right 'Request' box is shown with a red starburst icon, indicating a failure. Both 'Request' boxes connect to two 'FC Switch' boxes. The 'FC Switch' boxes are connected to two 'Storage Processor' (SP) boxes, labeled 'SP A' and 'SP B'. Each SP box has two ports labeled '0' and '1'. A 'LUN' icon is shown at the bottom, connected to the 'SP B' box. The diagram is set against a yellow background with a blue border.

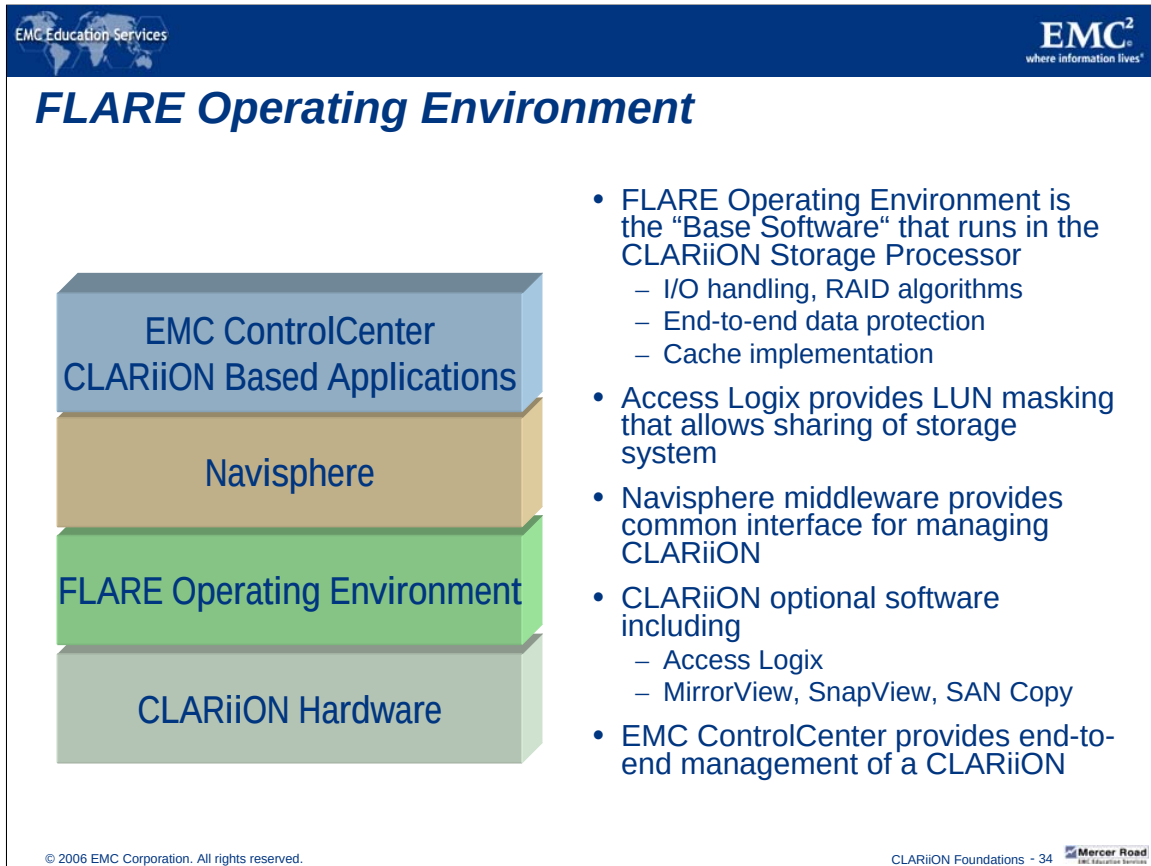
© 2006 EMC Corporation. All rights reserved. CLARiiON Foundations - 32 Mercer Road

To support the redundancy of the Hardware and Software of the CLARiiON arrays, host connectivity is also protected with PowerPath. PowerPath provides automatic load balancing of Host I/O to the array and is customer configurable. The CLARiiON Trespass policy will enable the PowerPath software to interact with the Access Logix software on the array to change to the surviving path, and temporarily change the SP ownership of a LUN in the event of a path failure. Administrators can improve the server’s ability to manage heavy storage loads through continuous and intelligent I/O load balancing.

CLARiiON Foundations

SOFTWARE AND MANAGEMENT ENVIRONMENT

This section will look at the management environment of CLARiiON.



Let's briefly introduce the software components, as most functionality is based in software and supported by the hardware.

FLARE software manages all functions of the CLARiiON storage system. Each storage system ships with a complete copy of FLARE software installed. When you power up the storage system, each SP boots and executes FLARE software.

- FLARE performs provisioning and resource allocation
- Memory budgets for caching and for snap sessions, mirrors, clones, copies
- Process Scheduling
- Boot Management

Access Logix software is optional software that runs within the FLARE operating environment on each storage processor (SP). Access Logix provides access control and allows multiple hosts to share the storage system. This “LUN Masking” functionality is implemented using Storage Groups. A Storage Group is one or more LUNs within a storage system that are reserved for one or more hosts and are inaccessible to other hosts. When you power up the storage system, each SP boots and executes its Access Logix software.

Navisphere Management software is a suite of tools that allows centralized management of CLARiiON storage systems. Navisphere provides a centralized tool to monitor, configure, and analyze performance. CLARiiON can also be managed as part of EMC ControlCenter, allowing full end-to-end management.

CLARiiON Management Options

- There are two CLARiiON management interfaces
 - CLI (Command Line Interface)
 - navicli commands can be entered from the command line and can perform all management functions
 - GUI (Graphical User Interface)
 - Navisphere Manager is the graphical interface for all management functions to the CLARiiON array

Navisphere CLI allows a command line interface for storage management. Provided the host has the agent file configured, a user can perform all the functions needed to manage the array.

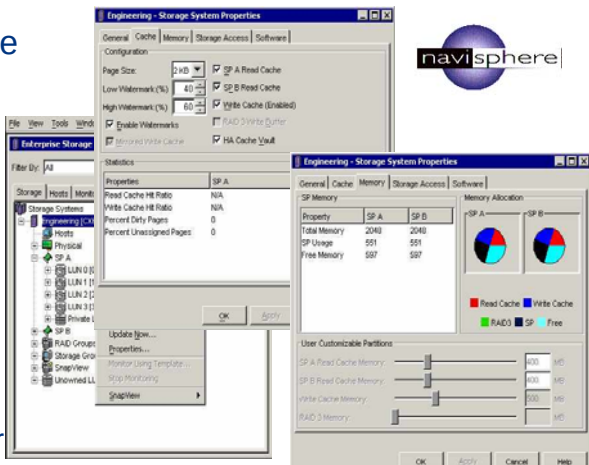
Navisphere GUI is an intuitive, easy to use graphical user interface for all management functions on the array.

EMC Education Services

EMC²
where information lives™

EMC Navisphere Management Software

- Centralized management for CLARiiON storage throughout the enterprise
 - Centralized management means more effective staff
- Allows user to quickly adapt to business changes
- Keeps business-critical applications available
- Key features
 - Java based interface has familiar look and feel
 - Multiple server support
 - EMC ControlCenter integration
 - Management framework integration



- Navisphere Software Suite
 - Navisphere Manager
 - Navisphere Analyzer
 - Navisphere CLI

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 36 Mercer Road
EMC Education Services

Navisphere Management software is a suite of tools that allows centralized management of CLARiiON storage systems. Navisphere provides a centralized tool to monitor, configure, and analyze performance. The Navisphere Suite includes:

- **Navisphere Manager:** Allows graphical user interface (GUI) management and configuration of a single or multiple system, and is also the center for management and configuration of system-based access and protection software including Access Logix, SnapView, and MirrorView applications
- **Analyzer:** A performance analysis tool for CLARiiON hardware components
- **Agent:** Provides the management communication path to the system. Enables CLI access

Navisphere can be launched from EMC ControlCenter. In the event that you already have a Symmetrix and ControlCenter, ControlCenter will allow you to manage all your information from one central location.

Navisphere stores initiator and host information on the array. It is used to create the relationships used for access control, as well as information presented in Navisphere. It also can automatically set any unique features required to support an operating system, enabling it to support a heterogeneous environment on a single port.

EMC Education Services

EMC²
where information lives™

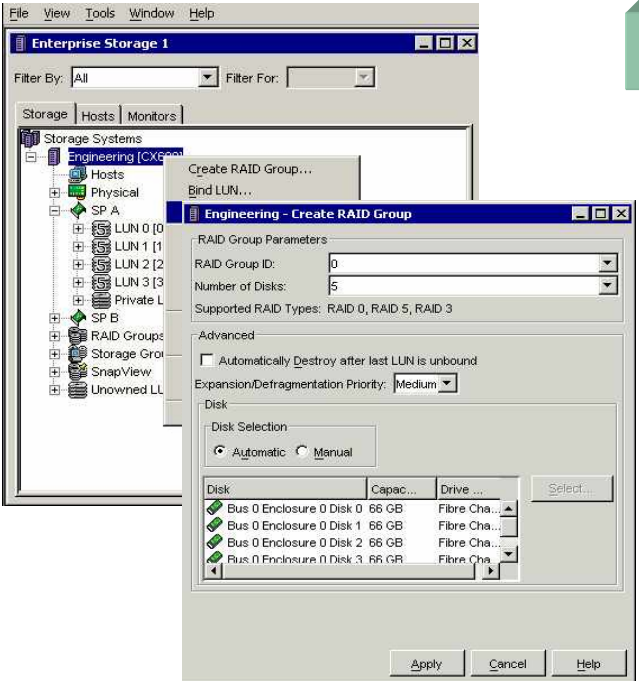
Navisphere Manager

- Discover
 - Discovers all managed CLARiiON systems
- Monitor
 - Show status of storage systems, Storage Processors, disks, snapshots, remote mirrors, and other components
 - Centralized alerting
- Apply and provision
 - Configure volumes and assign storage to hosts
 - Configure snapshots and remote mirrors
 - Set system parameters
 - Customize views via Navisphere Organizer
- Report
 - Provide extensive performance statistics via Navisphere Analyzer

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 37 Mercer Road
EMC Education Services

Navisphere Manager provides speed and flexibility using the familiar Microsoft windows interface. It lowers cost of management/ownership, including training and administrative costs, and has a proactive focus on addressing potential problems. It reduces personnel requirements by its ability to increase the productivity of staff by managing larger amounts of storage with fewer resources.



Step 1 – Create RAID Groups

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 38

 Mercer Road
EMC Education Services

Before a LUN can be created, you must create the RAID group on which you will bind the LUN. A RAID Group is one or more Physical Drives that share the same RAID type.

Each RAID Group has a unique number to identify it. By default, the next available RAID Group ID is automatically assigned, however, you have the capability of selecting from the list of available identifiers.

The actual physical drives that make up the RAID group can be assigned automatically by simply specifying the number of disks. The next available drives in sequence will be used or, if you select Manual, you will be allowed to choose the disks that make up the RAID Group.

Under RAID Group Parameters, in the Expansion/Defragmentation Priority list, click the priority for the new RAID Group. Select the Automatically Destroy check box to enable automatic dissolution of the RAID Group when the last LUN is unbound, or clear the check box to disable automatic dissolution.

When you create a RAID Group and specify the number of drives, the supported RAID type will be displayed. However, the actual RAID type for a RAID Group is not set until the first LUN is created within the group.

EMC Education Services

EMC²
where information lives™

Binding a LUN

Step 2 – Bind LUNs

© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 39

Mercer Road
EMC Education Services

This shows how a LUN is bound in a RAID Group. The following steps are used to bind a LUN:

- For each LUN you create, select the type of RAID protection you would like from the RAID Type box
- For the RAID Group, select a number for the LUN
- Under LUN properties, select a LUN ID, Element Size, Rebuild Priority, and Verify Priority
- Enable the Read or Write Cache by checking the appropriate box
- Specify the LUN Size, Number of LUNs to Bind, and which SP will own the LUN (Default Owner)

EMC Education Services

EMC²
where information lives™

LUN Properties - General

Engineering : LUN 0 : LUN Properties

General | Cache | Prefetch | Statistics | Hosts | Disks

Identity

LUN Name: LUN 0
LUN ID: 0 (0x00)
Unique ID: 60.06.01:E3.2C.09.00:00:82.01:91:77:5F:F4:D7:11
Current State: Normal

Miscellaneous

RAID Type: RAID5
Drive Type: Fibre Channel
RAID Group: 0

Capacity

User Blocks: 2097152
User Capacity: 1 GB (1,073,741,824 bytes)
Raw Capacity: 1.25 GB (1,342,177,280 bytes)

Advanced

Percent Bound: 100
Percent Rebuilt: 100
Rebuild Priority: ASAP
Verify Priority: ASAP

Element Size: 128
Alignment Offset: 0

Ownership

☐ Auto Assignment Enabled
Current Owner: SP A

Default Owner
☒ SP A ☐ SP B

OK

Apply

Cancel

Help

Step 2 – Bind LUNs

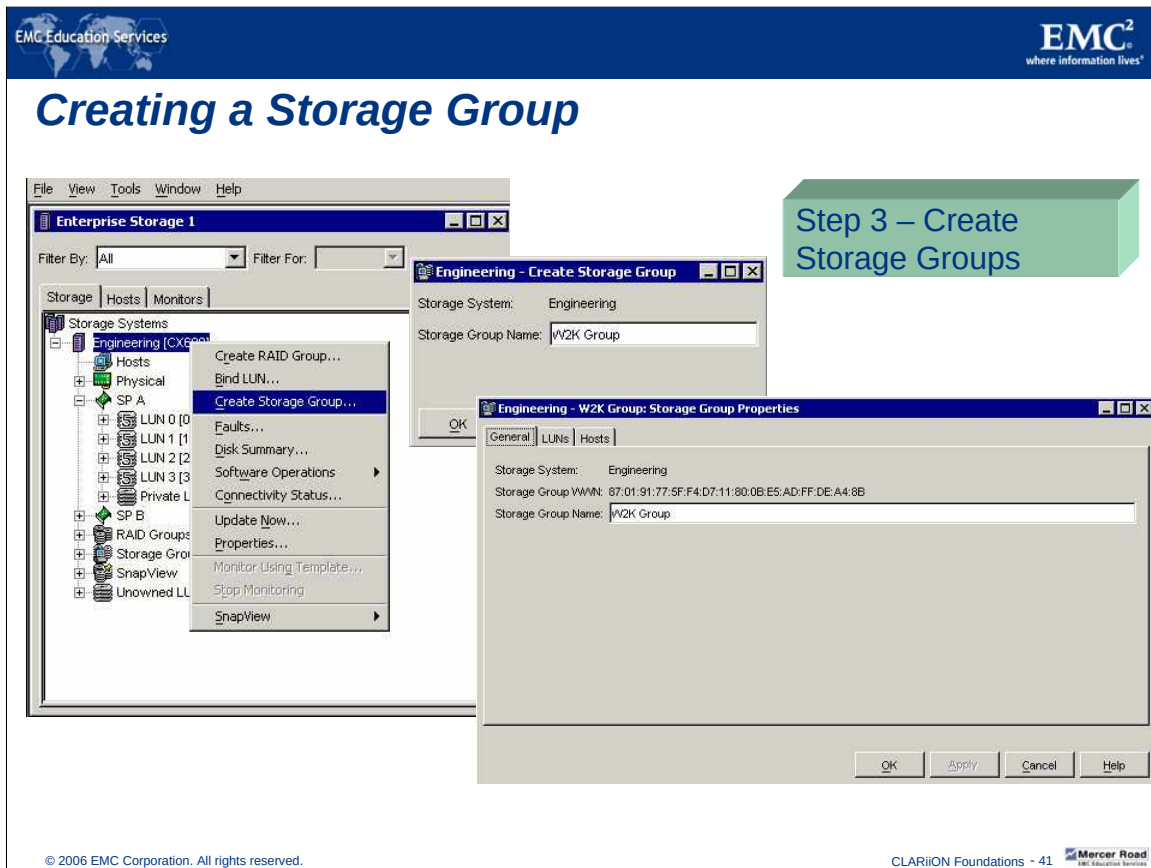
© 2006 EMC Corporation. All rights reserved. CLARiiON Foundations - 40 Mercer Road
EMC Education Services

This shows the LUN Properties screen with the General tab selected. The characteristics of the LUN are shown.

To change the Unique ID, Element, or RAID Type, you must unbind and rebind the LUN with the new characteristics. Unbinding is disruptive and will result in all data being lost on that LUN.

The Rebuild Priority, Verify Priority, and Auto Assign can be changed on this screen without affecting the data stored on the LUN.

The default owner is the SP that assumes ownership of a LUN after power is turned off, then on again. You may want to transfer ownership of a LUN to balance LUNs between SPs. Ownership is changed on this screen by selecting SP A or SP B under Default Owner.



This is an example of creating a Storage Group in Navisphere Manager. When a Storage Group is created, a unique name is assigned. The Storage Group Properties – General Tab will display this name, as well as the Storage Group Fibre Channel World Wide Name that is used internally to control access to LUNs within the Storage Group.

EMC Education Services

EMC²
where information lives™

Storage Group Properties - LUNs

Step 4 – Add LUNs to Storage Groups

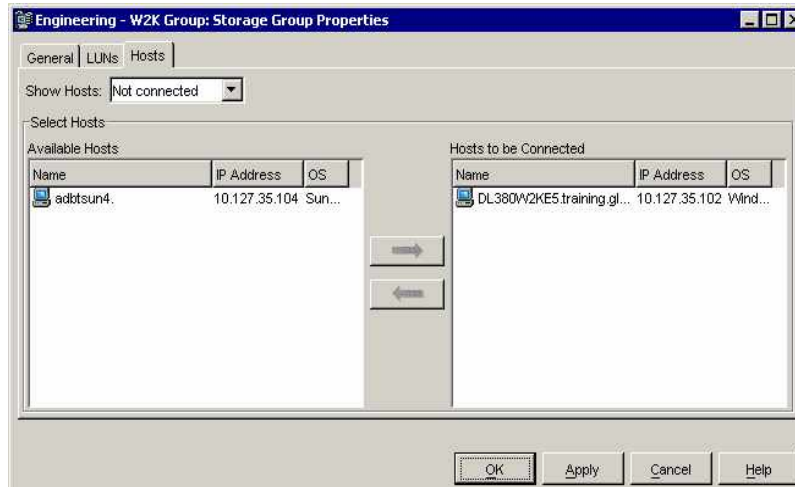
© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 42 Mercer Road
EMC Education Services

The Storage Group Properties – LUNs tab in Navisphere Manager is used to assign LUNS to a storage group. Hosts that are “connected” to a Storage Group will have full read-write access to all LUNs that are selected as part of the Storage Group.

Storage Group Properties - Hosts

Step 5 – Connect Hosts to Storage Groups



The Storage Group Properties – Host tab in Navisphere Manager is used to connect hosts to a storage group. When a host is “connected” to a Storage Group, it will have full read-write access to all LUNs in the group. Non-connected hosts will have no access to these LUNs.

CLARiiON Foundations – Appendix A

CX3 Series Differences

Appendix A will look at the differences between CX Series arrays and the CX3 Series array.

CLARiiON CX3 Series

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

- Describe the basic differences between the CX3 Series existing CLARiiON CX Series Arrays
- Describe the features of the CX3 Series arrays
- Describe the CLARiiON CX technical positioning in the marketplace



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

Fibre Channel Speed Support

CX Series

- Front End supports 1/2Gb speeds
 - Can only be set to fixed speed
- Back End only supports 2Gb
- Link LEDs
 - Green for 1/2Gb connection

CX3 Series

- Front End supports 1/2/4Gb speeds
 - Can be set to fixed speed, or auto-negotiation
- Back End supports 2Gb and 4Gb speeds
 - Auto-configured per bus at installation, persistent thereafter
 - Buses can be subsequently rescanned, and their speed changed, via a Navi wizard
 - Bus speed is independent per BE bus
 - A and B loops of a given bus change speed simultaneously
- Link LEDs
 - Green for 1/2Gb connection
 - Blue for 4Gb connection
 - Alternating Green/Blue for faulted port/cable

The 3rd generation of CX3 Series arrays differs from the older CX Series in a number of ways. The front end ports which attach to the hosts now support speeds of up to 4Gb per second and must be set to “auto-negotiate”. Back end buses now support both 2Gb and 4Gb per second speeds, giving the CX3 Series the first end to end 4Gb solution in the industry. Built-in algorithms allow the back end buses to be rescanned and changed independently, on a per bus level. An additional LED has been added to accommodate the 4Gb speed and faults.

Fibre Channel Connectivity

CX Series

- Front End ports consist of “embedded” SFF optical transceivers
 - Optical cables use “LC” style connector
 - 300 Meter maximum length
- Back End ports use HSSDC connectors
 - HSSDC-to-HSSDC Copper cables
 - 1 Meter, 5 Meter and 10 Meter lengths

CX3 Series

- SFP cages support FE and BE connectivity
- Front End ports use “pluggable” SFP optical transceivers (FRUs)
 - Optical cables use “LC” style connector
 - 100 Meter maximum length
- Back End ports use empty SFP cage
 - SFP-to-HSSDC2 Copper cables plug directly into cage
 - 2 Meter 5 Meter lengths

The CX3 Series now uses Small Form factor pluggable (SFP) for connectivity. The SFPs are LC to LC style connectors, whereas the back end ports use LC to HSSDC2 cables.

Serviceability/Diagnosability

CX Series

- Single Management LAN port per SP
- Entire Storage Processor is a FRU
- Storage Processors can survive a single fan or blower fault, but shut down within 2 minutes of cooling module removal

CX3 Series

- Separate Customer and Service LAN ports per SP
- New SP “sub-FRUs”
 - CPU Module
 - I/O Modules
 - DIMMs
 - Optical SFP transceivers
 - Error Handling firmware and software in place to diagnose hardware faults to one or more sub-FRUs, and call them out in the logs
- Storage Processors can run indefinitely, in the event of a single blower fault or removal

The CX3 Series now supports two management LAN ports per Storage Processor/Blade. The Storage Processors/Blades contain “sub FRUs”. The sub FRUs can be individually replaced without replacing the entire Blade. Previous versions of CX Series SPs would shut down within two minutes upon a fan or blower failure. CX3 Series will now continue to operate indefinitely under those conditions.



CLARiiON CX3 Series Differences Summary

- **Better performance**
 - End-to-end 4-Gbit platform (architecture optimized for full 4Gbit throughput)
 - Increased memory
- **Greater scalability**
 - Low cost FC Disks
 - Up to 2X capacity using increased disk counts
- **Enhanced serviceability**
 - Cross-generational software support
 - FRUs and CRUs
 - Service tools to help customers service and maintain the storage environment
- **Enhanced reliability**
 - System-level intelligence to maximize availability
- **Data in place upgrades**
 - From previous generation CX systems
 - Throughout CX3 series

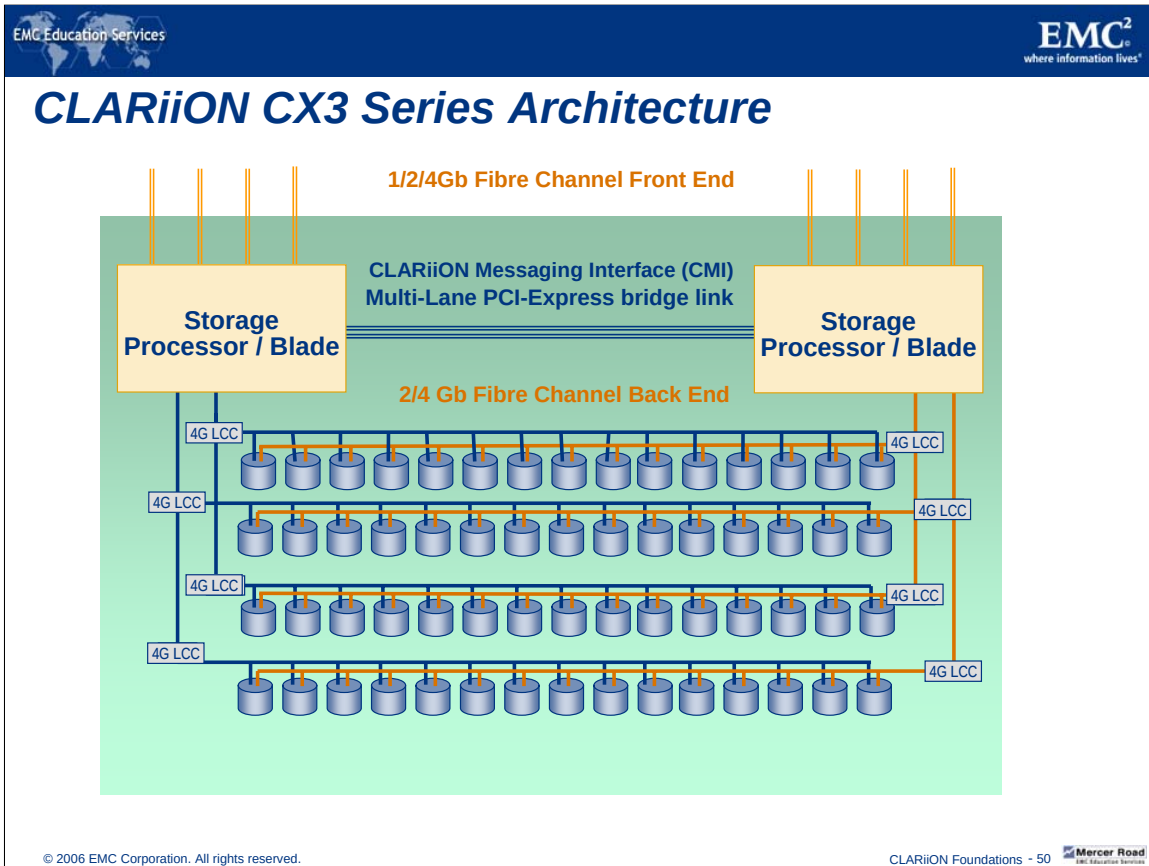
© 2006 EMC Corporation. All rights reserved.

CLARiiON Foundations - 49 

The CX3 series product line delivers performance and scalability by providing the industries first storage array with end to end 4Gbit throughput.

The CX3 Series also introduces support for a new high capacity/low price disk drive which enables the CX3 series to deliver the performance and economics that customers require for backup-to-disk applications. A unique benefit of the Low Cost – Fibre Channel (LC-FC) drives is that it can be placed in any drive slot in the array, with the exception of the first 5 drives. (reserved for vault space). This maximizes array capacity, while minimizing price/capacity. To improve serviceability, the CX3 has been designed with more components that are available as Field Replaceable Units (FRU's), as well as some Customer Replaceable Units (CRUs).

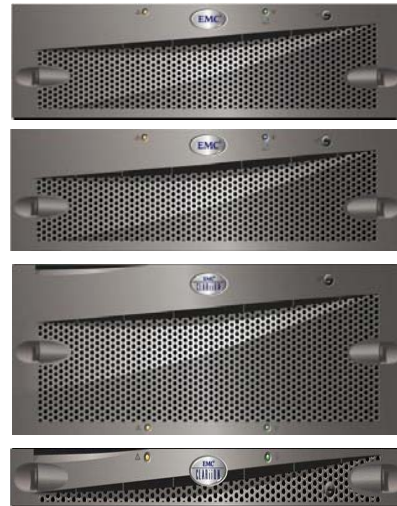
Enhancements have been made to existing tools and new tools have been implemented for customer ease of use.



The CX3 Series module architecture allows the customer to add drives, as needed, to meet capacity requirements. CLARiiON architecture is based on intelligent Storage Processors/Blades that manage physical drives on the back end and service host requests on the front end, be it Fibre Channel or iSCSI protocols. Depending on the model, each Storage Processor or Blade includes either one or two CPUs. Storage Processors communicate to each other over the CLARiiON Messaging Interface (CMI); however, unlike previous CX series arrays, the CX3 Series will be the first platform to use PCI-Express as the high-speed CMI path. PCI Express architecture delivers advance I/O technology to new levels of performance and scalability.

Modular Building Blocks

- DAE3P Disk Enclosure
 - Supports CX3 platform
 - Supports up to 15 low-profile, 2/4Gb Fibre Channel disk drives
 - Can operate at either 4Gb or 2Gb speeds
 - Drive fillers must be installed in empty slots
- Uses same Chassis and Power Supply/Cooling module as DAE2P Enclosure
- Uses same cables as DAE2P Enclosure
 - 8 Meter maximum between DAE3Ps
 - 5 Meter maximum between SPEs and DAEs
- Two 4G Link Control Cards (LCCs)



The new CX3 Series platform supports the new DAE3P disk enclosure. The DAE3P can house up to 15 low-profile, 2/4Gb Fibre Channel disk drives, and can operate at either 2Gb or 4Gb speeds.

The DAE3P uses the same chassis, power supply/cooling module as the DAE2P enclosure. The same cables used for the DAE2P are also supported by the DAE3P. DAE3P uses 4Gb LCCs.

CLARiiON CX3 Series Hardware

- Fully redundant architecture
 - Dual Storage Processors
 - Power, cooling, data paths, SPS
- Separate Customer and Service LAN ports per SP
- Customer Maintainable Components
 - Customer Replaceable Disks
 - SPE Power Suppliers and Blowers
 - Small Form Factor Pluggable (SFP's)
- Compatible with DAE2P and DAE3P
 - 15 disk drive capacity
- PCI-Express computing technology
 - Peer-to-peer communication
 - Enhanced system performance
- Flexibility
 - Mix drive types
 - Mix RAID levels
 - Mix DAE types

The CX3 Series will utilize a combination of a diskless SPE and external DAEs in the SAN offering. They are compatible with DAE2P and DAE3P enclosures. CLARiiON uses these building blocks to offer an innovative single storage processor enclosure, leveraging PCI-Express computing technology for peer-to-peer communication and enhanced system performance.

Disk support is via a separate 3U rack-mountable Disk Array Enclosure (DAE3P or DAE2P), each with 15-disk drive capacity supporting Fibre Channel drives. Fibre Channel and SATA drives cannot be mixed within the same enclosure.

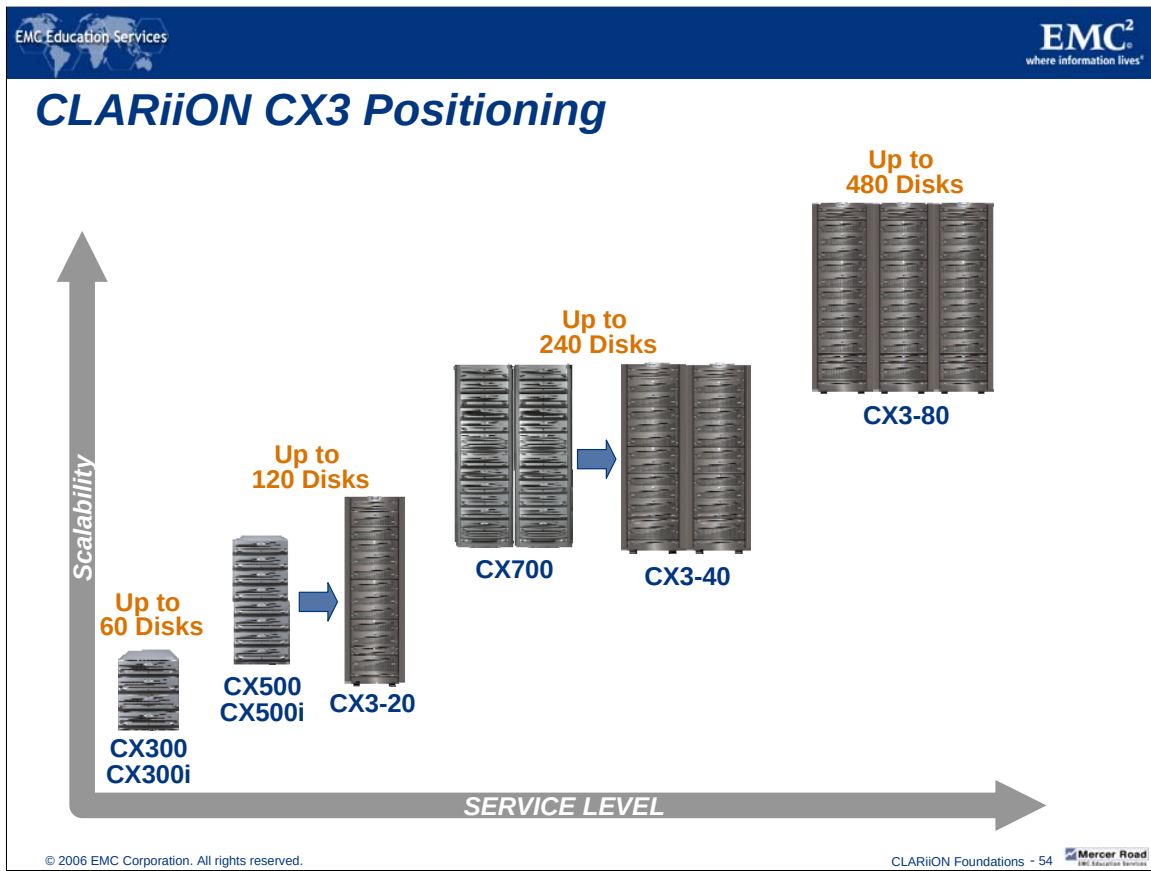
Storage Processor Enclosure (SPE)

- **Front End supports 1/2/4Gb speeds**
 - Can be set to fixed speed, or auto-negotiation
- **Back End supports 2Gb and 4Gb speeds**
 - Auto-configured per bus at installation, persistent thereafter
 - Buses can be subsequently rescanned, and their speed changed, via a Navisphere wizard
 - Bus speed is independent per Back End (BE) bus
 - A and B loops of a given bus change speed simultaneously
- **Greater FRU capability**
 - Write Cache Card (WCC)
 - Greater performance in the event of an SP failure.

The CX3 Series Storage Processor enclosure houses two hot-swappable Storage Processors referred to as “blades”. Each blade contains replaceable Dual In-line Memory Modules (DIMM) and Small Form Factor Pluggable (SFP), as well as four hot-swappable power supply/blower modules (two modules per storage processor).

The SAN management card exists to provide various management functionalities contained on a three board set. A Processor/Memory Board, Front End I/O Module (multiple personality GbE, FC), Back End I/O Module (multiple personality).

The Write Cache Card (WCC) is a hot-repairable card for the SAN system (high end CX3 series). It allows for greater performance in the event of an SP failure. With the introduction of the WCC, an SP will be able to continue write caching operation in the event of a single SP failure by mirroring data to the WCC. This will limit the effect a single SP failure has on a system.



The CLARiiON CX3 Series is the next generation midrange storage arrays. It provides a new architecture designed to optimize 4Gbit performance as well as provide scalability and optimum value for the cost-sensitive midrange storage market segment.

There are 3 new arrays in the series: the CX3-20 which supports 120 drives, the CX3-40 which supports 240 drives, and the CX3-80 which support 480 drives. The larger disk counts enable customer's to satisfy their increasing storage needs to support business continuity and archiving.

The approach is to add new hardware under a stable software platform and add software features and serviceability tools. The CX3 Series is completely software compatible with the existing CX series. Additionally, all EMC infrastructure software products are compatible with the CX3 series. The CLARiiON CX series offers a reliable solutions for storage consolidation, from the entry-level for small to medium businesses or the departments/branch offices of enterprises. The top-of-the-line CLARiiON CX 3 Series provides high performance and large-scale consolidation of demanding application environments.

Module Summary

- Identified the basic differences between the CX3 Series and the existing CLARiiON CX Series arrays
- Identified the features of the CX3 Series arrays
- Identified the CLARiiON CX technical positioning in the marketplace

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