

SAN Foundations

After completing this module, you will be able to:

- **Provide an overview of Fibre Channel and IP SANs**
 - Define a Storage Area Network (SAN)
 - List the features and benefits of implementing a SAN
 - Provide an overview of the underlying protocols used within a SAN
- **Discuss issues to consider when designing a SAN**
 - State the distinct characteristics of commonly deployed fabric topologies
 - Explain the basic operational details of Inter-Switch Links (ISL)
 - List performance and security related features relevant to a SAN
- **List the major product categories within the EMC Connectrix family**
 - State the features and benefits of the EMC Connectrix family
 - List the various software options for managing Fabric components
 - Identify Connectrix component types to be used, when designing a SAN

The objectives for this module are listed on the slide. Please take a moment to read them.



SAN Foundations

Storage Connectivity: Overview



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

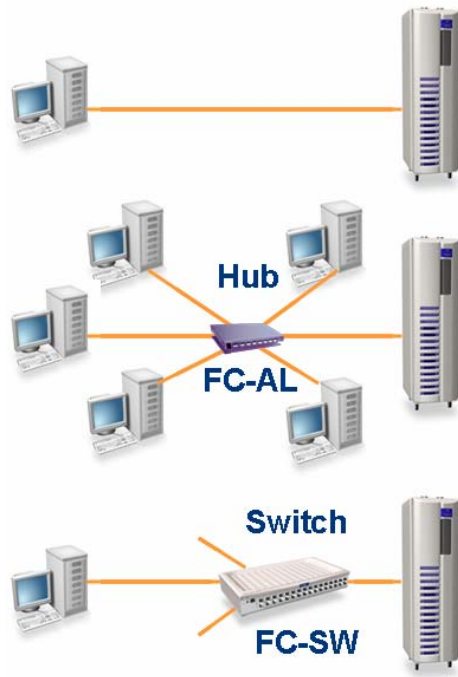


This section introduces the basic structure of a SAN. It highlights fundamental differences between a SAN and legacy connectivity architectures.



SAN Connectivity Methods

- There are three basic methods of communication using Fibre Channel infrastructure
 - Point to point (P-to-P)
 - A direct connection between two devices
 - Fibre Channel Arbitrated Loop (FC-AL)
 - A daisy chain connecting two or more devices
 - Fabric connect (FC-SW)
 - Multiple devices connected via switching technologies



© 2006 EMC Corporation. All rights reserved.

SAN Foundations



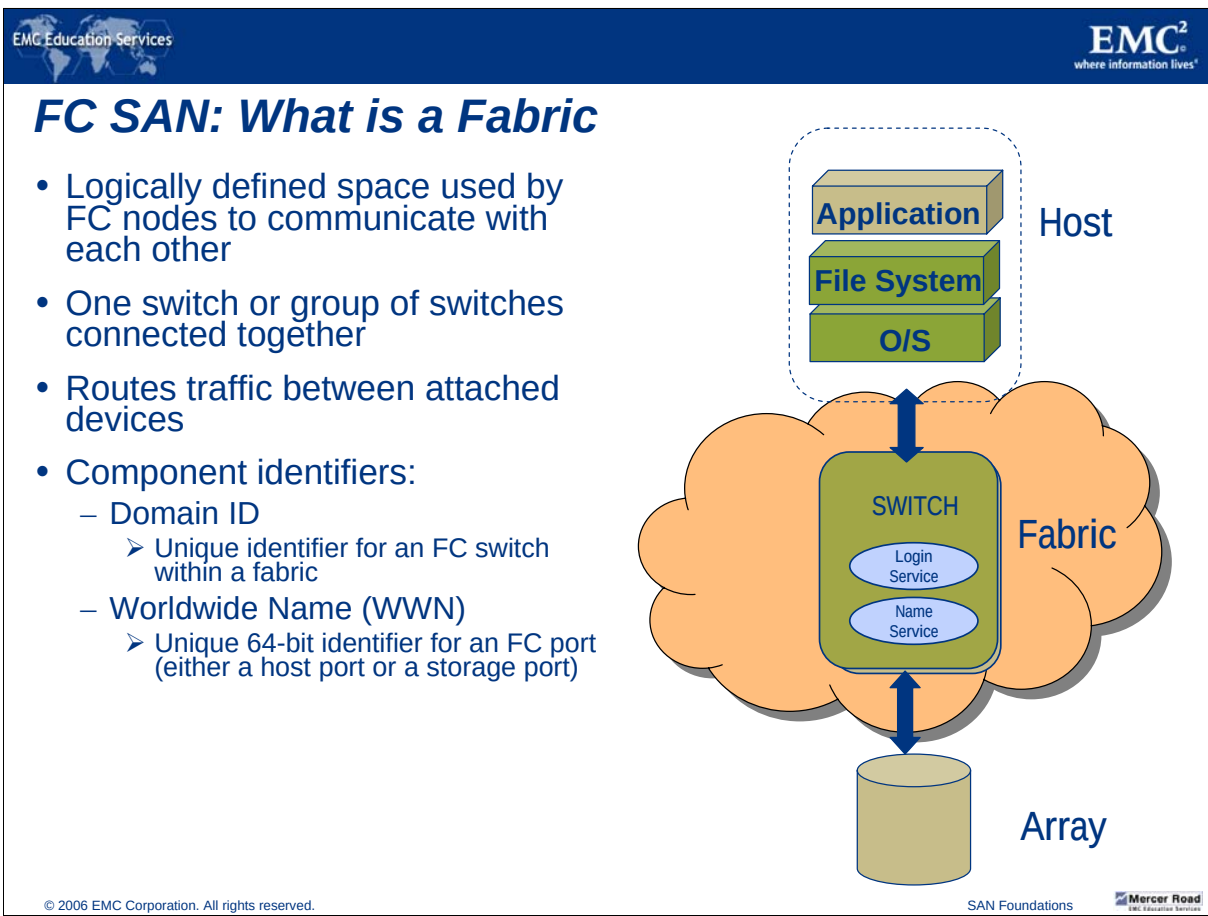
The slide shows the basic interconnectivity options supported with the Fibre Channel architecture:

- (1) Point to point
- (2) Fibre Channel Arbitrated Loop
- (3) Fabric Connect

FC-AL is a loop topology that does not require the expense of a Fibre Channel switch. In fact, even the hub is optional – it is possible to run FC-AL with direct cable connections between participating devices. However, FC-AL configurations do not scale well, for several reasons:

- (1) The topology is analogous to token ring. Each device has to contend for the loop via arbitration. This results in a shared bandwidth environment – since at any point in time, only one device can “own” the loop and transmit data.
- (2) Private arbitrated loops use 8-bit addressing. So there is a limit of 126 devices on a single loop.
- (3) Adding or removing devices on a loop results in a loop reinitialization, which can cause a momentary pause in all loop traffic.

For most typical SAN installations, Fabric connect via switches (FC-SW) is the appropriate choice of Fibre Channel topology. Unlike a loop configuration, a switched fabric provides scalability, and dedicated bandwidth between any given pair of inter-connected devices. FC-SW uses a 24-bit address (called the Fibre Channel Address) to route traffic, and can accommodate as many as 15 million devices in a single fabric. Adding or removing devices in a switched fabric does not affect ongoing traffic between other unrelated devices.



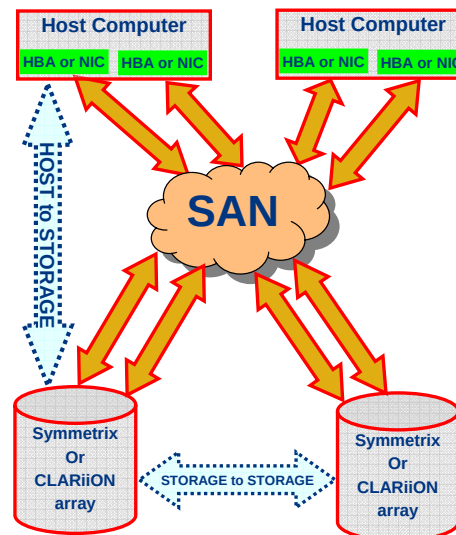
A fabric is a logically defined space in which Fibre Channel nodes can communicate with each other. A fabric can be created using just a single switch, or a group of switches connected together.

The primary function of the fabric is to receive FC data frames from a source port (device) and route them to the destination port (device) whose address identifier is specified in the FC frames. Each port (device) is physically attached through a link to the fabric.

Many models of switches can participate in only a single fabric. Some newer switches have the capability to participate simultaneously in multiple fabrics. Within a fabric, each participating switch must have a unique identifier called its Domain ID.

What a SAN Does

- SAN is a technology that addresses two critical storage connectivity problems:
 - *Host-to-storage connectivity*: so a host computer can access and use storage provisioned to it
 - *Storage-to-storage connectivity*: for data replication between storage arrays
- SAN technology uses block-level I/O protocols
 - As distinct from NAS, which uses file-level I/O protocols
 - The host is presented with raw storage devices: just as in traditional, direct-attached storage



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

A SAN provides two primary capabilities: block-level storage connectivity from a host to a storage frame or array, and block-level storage connectivity between storage frames or arrays.

For a storage array such as Symmetrix or CLARiiON, the LUN – which stands for Logical Unit Number – is the fundamental unit of block storage that can be provisioned. The host's disk driver treats the array LUN identically to a direct-attached disk spindle - presenting it to the operating system as a raw device or character device. This is the fundamental difference between SAN and NAS. A NAS appliance presents storage in the form of a filesystem, that the host can mount and use via network protocols such as NFS (Unix hosts) or CIFS (Windows hosts).

Some host software applications can use raw devices directly, e.g. relational database products. Most enterprise applications require, or prefer, the use of a filesystem. With SAN, the host can build a local, native filesystem on any presented raw devices.

SAN connectivity between storage frames or arrays enables the use of array-centric, block-level replication capabilities, e.g. SRDF (Symmetrix arrays) and MirrorView (CLARiiON arrays).

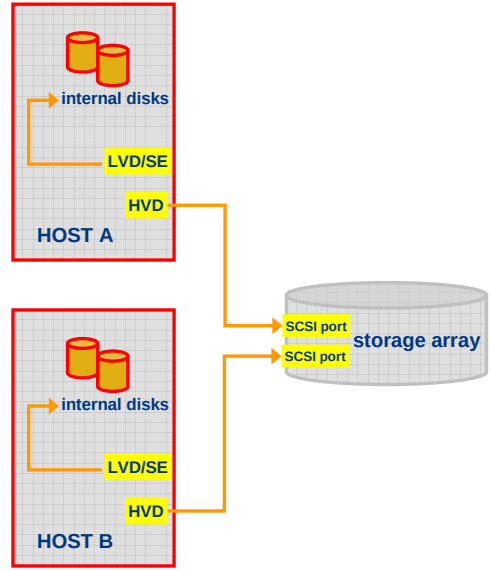




Legacy Storage Connectivity: DAS

- DAS (Direct-Attached Storage) is the legacy architecture for host-to-storage connectivity
 - Dedicated physical channel
 - Parallel transport
 - Examples
 - Parallel **SCSI** (pronounced “scuzzy”)
 - ESCON
 - Advantage: low protocol overhead
 - Very fast: rated bandwidth can be as high as 320 Mbytes/sec on a SCSI bus
 - Still appropriate, and universally used, for internal storage devices in host computers
- DAS is ill-suited to enterprise storage connectivity:
 - Static configuration
 - Distance limitations
 - Topology limitations
 - Scalability limitations

Legacy SCSI Solution



The diagram illustrates a Legacy SCSI Solution. It shows two hosts, Host A and Host B, each containing internal disks. Host A has an LVD/SE adapter connected to its internal disks and an HVD adapter connected to its SCSI port. Host B has an LVD/SE adapter connected to its internal disks and an HVD adapter connected to its SCSI port. The storage array has two SCSI ports, each connected to the HVD adapter of one of the hosts.

© 2006 EMC Corporation. All rights reserved.

SAN Foundations



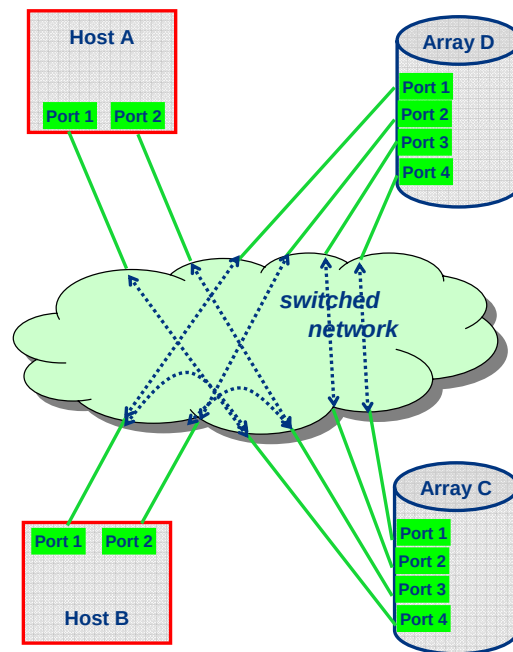
Traditionally, storage has been provisioned to hosts directly in the form of physical disk spindles, on a dedicated physical channel. Channel architectures provide fixed connections between a host and its peripheral devices. Host-to-storage connections are defined to the host operating system in advance. Tight integration between the transmission protocol and the physical interface minimizes protocol overhead. Parallel SCSI (in the open systems arena) and ESCON (in the mainframe world) are classic examples of channel architectures.

SCSI - which is an acronym for Small Computer System Interface – is a peripheral interconnect standard that has existed and periodically evolved since the early 1980s. Parallel SCSI employs three distinct types of electrical bus signaling: Single-ended (SE), High-Voltage Differential (HVD) and Low-Voltage Differential (LVD). LVD and HVD devices are electrically incompatible, and cannot reside on the same SCSI bus. The host requires a SCSI controller (also called a SCSI host adapter, or initiator) to communicate with the attached SCSI storage devices (or targets). The host adapter can be an LVD/SE adapter or an HVD adapter, depending on the required signaling type. Typically, external storage devices such as arrays use HVD signaling due to the greater distances possible with HVD. Still, bus lengths beyond a few tens of meters can compromise signal integrity. Internal disk devices in modern hosts are invariably LVD.

In the picture, each of the two hosts has two different SCSI adapters – one LVD adapter to handle the internal LVD disk drives, and one HVD adapter to connect to a HVD/SCSI port on the storage array. Some hosts have one or more embedded SCSI controllers on the motherboard, thus eliminating the need for an add-on adapter card.

Motivations for Networked Storage

- The efficiency from isolating **physical** connectivity from **logical** connectivity
 - Topology limitations eliminated
- The ease of logically connecting a single array port to multiple host ports, and vice-versa
 - **Fan-out** (one storage port services multiple host ports)
 - **Fan-in** (one host port accesses storage ports on multiple arrays)
- Dynamic vs. static configuration
- Distance limits can be alleviated
- Provides better scalability



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Traditional DAS solutions such as Parallel SCSI were not really designed to scale to the requirements of modern enterprise-class storage.

Scalability issues with DAS include the following:

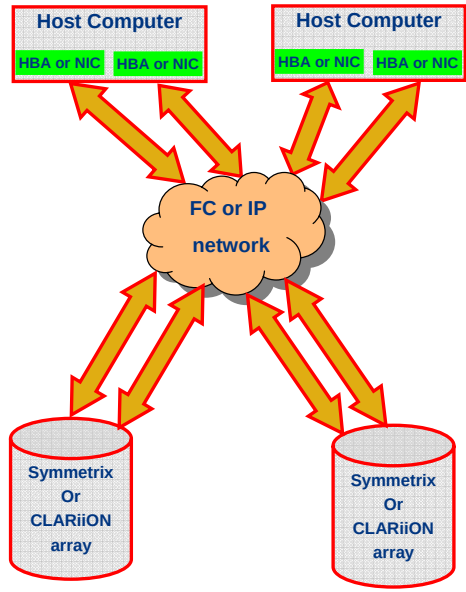
- (1) Distance limitations dictated by the underlying electrical signaling technologies.
- (2) With static configuration, the bus needs to be quiesced for every device reconfiguration. Every connected host would lose access to all storage on the bus during the process.
- (3) In parallel SCSI, devices on the bus must be set to a unique ID in the range of 0 to 15. Addition of new devices and/or initiators with parallel SCSI requires careful planning - ID conflicts can render the entire bus inoperational.
- (4) DAS requires an actual physical connection via cable for every logical connection from a host to a storage device or port. The only way to deploy new storage, or redeploy storage across hosts, is to modify the physical cabling suitably. In theory, multiple host initiators can be accommodated on a single bus. In practice, cabling issues rapidly become a challenge as the configuration grows.

In contrast, switched networked architectures (such as SAN fabrics) can service multiple logical connections to each device - via a single physical connection from that device to the infrastructure. In the picture, the storage array C can provide storage to both hosts A and B, since C's Port 4 is logically connected via the network to Port 1 on each of these hosts. Additionally, port 3 on the array is configured for a second redundant logical path to Port 2 of each host.



Basic Structure of a SAN

- SAN: a networked architecture that provides I/O connectivity between host computers and storage devices
- Communication over a SAN is at the block I/O level
- The storage network can be either:
 - A Fibre Channel network
 - Typically, a physical network of Fibre Channel connectivity devices: interconnected FC Switches and Directors
 - For transport, an FC SAN uses FCP
 - ❖ FCP is serial SCSI-3 over Fibre Channel
 - Or an IP network
 - Uses standard LAN infrastructure: interconnected Ethernet switches, hubs
 - For transport, an IP SAN uses iSCSI
 - ❖ iSCSI is serial SCSI-3 over IP



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

 Mercer Road
EMC Education Services

SANs (Storage Area Networks) combine the benefits of channel technologies and the benefits of a networked architecture. This results in a more robust, flexible and sophisticated approach to connecting hosts to storage resources. SANs overcome the limitations of Direct-Attached Storage, while using the same logical interface – SCSI - to access storage.

SANs use one of the following two data transport protocols:

- Serial SCSI-3 over Fibre Channel (FC). In the storage realm, this is widely referred to as simply the Fibre Channel Protocol, or FCP.
- Serial SCSI-3 over IP. This is commonly known as iSCSI.

Host to Storage communication in a SAN is block I/O – just as with DAS implementations. With parallel SCSI, the host SCSI adapter would handle block I/O requests. In a Fibre Channel SAN, block requests are handled by a Fibre Channel HBA or Host-Based Adapter. A Fibre Channel HBA is a standard PCI or Sbus peripheral card on the host computer, just like a SCSI adapter.

SAN versus DAS

- SANs eliminate the topology and distance limitations imposed by traditional DAS solutions
- SANs support non-disruptive provisioning of storage resources
- SANs allow multiple servers to easily share access to a storage array or frame
- SANs provide better infrastructure for multipathing
- SANs enable consolidation of storage peripherals
- SANs vastly increase scalability, as a net result of the above advantages

SANs make effective use of Fibre Channel networks and IP networks to solve the distance and connectivity problems associated with traditional DAS solutions such as parallel SCSI. In a SAN, a device can be added or removed without any impact on I/O traffic between hosts that do not participate in the configuration change. A host can reboot or disconnect from the SAN without affecting storage accessibility from other hosts. New arrays can be added to the SAN, and storage from them can be deployed selectively on some hosts only - without any impact on other hosts. Thus, SANs enable dynamic, non-disruptive provisioning of storage resources.

SAN architecture allows for multiple servers to easily share access to a single storage array port. This is technically possible with parallel SCSI too, via the use of daisy-chained cables. However, the setup is static, physically cumbersome, subject to practical constraints from requirements on signaling integrity, and difficult to establish and maintain.

SAN architecture also allows for a single host to easily connect to a storage frame via multiple physical and logical paths. In a multipathed configuration, and with the use of multipathing software such as Powerpath, the host experiences I/O failures only if every one of its logical paths to the storage array fails. Multipathing software can also help balance the host's I/O load over all available paths. Multipathing capability thus allows for the design of a high-performance, highly available, redundant host system.

SANs make it simple to consolidate multiple storage resources – such as disk arrays and tape libraries - within a single physical or logical infrastructure. These resources can be selectively shared across host computers. This approach can greatly simplify storage management, when compared to DAS solutions.

EMC's Connectrix Range



SAN Foundations

 Mercer Road

SAN Foundations - 10



EMC Connectrix Family

- Fibre Channel connectivity products
 - Enterprise Directors
 - Departmental Switches
- Multi-protocol routers
- Cabinets
 - Integrated Service Processor
 - Optimized airflow ensures high reliability
 - Cable management system
- In all the above categories, Connectrix products are available from three different vendors:
 - M-Series (McDATA)
 - B-Series (Brocade)
 - MDS-Series (Cisco)
- Management Software
 - Tools to manage Connectrix switches, directors and routers



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

 Mercer Road
EMC Education Services

EMC offers a complete range of SAN connectivity products under the Connectrix brand. Connectrix products are supplied by three different vendors: Brocade, McData and Cisco.

For Fibre Channel connectivity, the product family includes Enterprise Directors for data center deployments, and Departmental Switches for data center and workgroup deployment. Depending on the requirements for the SAN, such as number of Fibre Channel ports, redundancy and bandwidth requirements, the appropriate type and brand of switch can be selected.

Connectrix Fibre Channel switches and directors have several types of ports, each with a distinct function:

- (1) Fibre Channel ports, for block data transfer between inter-connected hosts and storage arrays;
- (2) One or more Ethernet (RJ45) ports, used for switch management via telnet, ssh or web browser; and switch monitoring via SNMP;
- (3) A serial port (COMM port), used for initial switch configuration, e.g. setting the IP address on the switch via CLI. Subsequent switch configuration, management and monitoring is typically done over the Ethernet port.

To support mixed iSCSI and Fibre Channel environments, multi-protocol routers are available. These routers have the capability of bridging FC SAN's and IP SAN's. Thus, they can provide connectivity between iSCSI host initiators and Fibre Channel storage targets. In addition, multi-protocol routers are required for extending Fibre Channel SAN's over long distances, via IP networks.



Connectrix: Departmental Switches vs. Enterprise Directors

• Departmental Switches – Limited hot-swappable components ➢ Redundant fans and redundant power supplies – High Availability through redundant deployment ➢ SAN can be designed to tolerate failure or decommissioning of an entire switch – Scalability through Inter-Switch links (ISLs) – Work group, departmental and data center deployment	• Enterprise Directors – Fully redundant components ➢ Optimal serviceability – Highest availability – Maximum scalability ➢ Can support large SANs – data center deployment
--	---

© 2006 EMC Corporation. All rights reserved.

SAN Foundations



Departmental Switches are less expensive compared to Directors, but they are smaller in capacity – i.e. have a limited of Fibre Channel ports - and offer limited availability. They are ideal for smaller environments where host connections are limited. SANs can be created with departmental switches but at the expense of a more complex architecture, requiring many more network devices and switch interconnects.

Connectrix Enterprise Directors on the other hand, offer greater levels of modularity, fault tolerance and expandability compared to Departmental Switches. Directors offer scalability and availability suitable for mission-critical SAN based applications, without sacrificing simplicity and manageability. Directors can be used to build larger SANs with simple topologies. Due to their relatively high port counts, they can help minimize, or completely avoid, the use of ISLs. Connectrix Directors have the following features:

- Redundant modular components supporting automated switchover triggered by hard or soft failures
- Pre-emptive hardware switchover powered by both automated periodic health checking and correlation of identified hardware failures
- On-line (non-disruptive) firmware update
- Hot-swappable hardware components

A combination of switches and directors from any given vendor (e.g. only B-series switches and directors) can usually interoperate. In single-vendor Fibre Channel networks, interoperability constraints (if any) arise from supported firmware revisions only.

EMC Education Services		EMC ² where information lives [®]	
Deployment: Switches vs. Directors			
Number of Hosts	Lowest Acquisition Cost	Highest Availability	Least Complexity
256–1024 ↑	 Switch  Director	 Director	 Director
64–256 ↓	 Switch	 Switch  Director	 Director
8–64 ↓	 Switch	 Switch	 Switch  Director
© 2006 EMC Corporation. All rights reserved. SAN Foundations 			

Enterprise Directors are deployed in High Availability and/or large scale environments. Connectrix Directors can have more than a hundred ports per device; when necessary, the SAN can be scaled further using ISLs.

Disadvantage of directors: higher cost, larger footprint.

Departmental Switches are used in smaller environments. SANs using switches can be designed to tolerate the failure of any one switch. This can be done by ensuring that any host/storage pair has at least two different paths through the network, involving disjoint sets of switches. Switches are ideal for workgroup or mid-tier environments. Large SANs built entirely with switches and ISLs require more connectivity components, due to the relatively low port-count per switch; therefore, there is more complexity in your SAN.

Disadvantage of departmental switches: Lower number of ports, limited scalability.

There are several widely-deployed Fibre Channel SAN topologies that can support a mix of switches and directors. A description of these topologies appears in the “Operational Details” section.

SAN: Architecture and Components



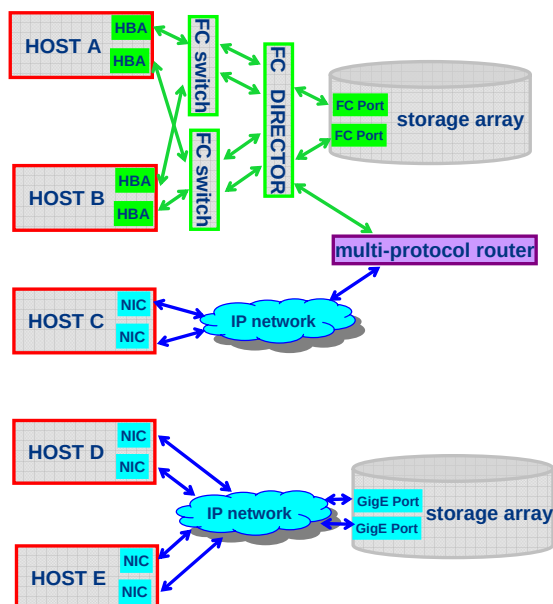
SAN Foundations

 Mercer Road

SAN Foundations - 14

SAN: Typical Connectivity Scenarios

- Fibre Channel SAN
 - Uses one or several inter-connected Fibre Channel switches and directors
 - Connects hosts and storage arrays that use Fibre Channel ports
- Bridged solution
 - Allows hosts to connect via iSCSI to Fibre Channel storage arrays
 - Requires use of a multi-protocol router
- IP SAN
 - Does not require any Fibre Channel gear (e.g. FC switches, HBAs)
 - Storage arrays must provide native support for iSCSI via GigE ports
- EMC's Connectrix family of products encompasses a range of Fibre Channel switches, directors and multi-protocol routers suitable for SAN deployments



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Physically, a Fibre Channel SAN can be implemented using a single Fibre Channel switch/director, or a network of inter-connected Fibre Channel switches and directors. The HBAs on each host, and the FC ports on each storage array, need to be cabled to ports on the FC switches or directors. Fibre Channel can use either copper or optics as the physical medium for the interconnect. All modern SAN implementations use fibre optic cables.

In the picture, Hosts A and B participate in a Fibre Channel SAN. These hosts can be readily provided access to any FC storage array on the SAN via the FC switches.

Bridging products such as multi-protocol routers enable hosts to use iSCSI over conventional network interfaces (NICs) to access Fibre Channel storage arrays. In the picture, Host C can be provided access via the multi-protocol router to the storage array with FC ports.

An IP SAN solution would use conventional networking gear, such as Gigabit Ethernet (GigE) switches, host NIC's and network cables. This eliminates the need for special-purpose FC switches, Fibre Channel HBAs and fibre optic cables. Such a solution becomes possible with storage arrays that can natively support iSCSI, via GigE ports on their front-end directors (Symmetrix) or on their SPs (CLARiiON). For performance reasons, it is typically recommended that a dedicated LAN be used to isolate storage network traffic from regular, corporate LAN traffic. In the picture, Hosts D and E are on an entirely IP-based SAN. Storage can be provisioned and made available to both hosts from the array with GigE ports.

FC SAN: Logical and Physical Components

- **Nodes and Ports:** A Fibre Channel SAN is a collection of “nodes”
 - A **node** is any addressable entity on a Fibre Channel network
 - A node can be: a host computer, storage array or other storage device
 - A node can have one or more **ports**
 - A **port** is a connection point to the Fibre Channel network
 - Examples of ports: host initiator i.e. a HBA port; or an FC port on a storage array
 - Every port has a **globally unique** identifier called the World Wide Port Name (WWPN), also called simply the World Wide Name (WWN)
 - ❖ WWN is 64 bits; in hexadecimal notation, it is a string of eight hex pairs
 - ★ For example: 10:00:08:00:88:44:50:ef
 - ❖ WWN is factory-set, i.e. “burned in” for a HBA
 - ❖ WWN may be software-generated for storage array ports
 - ❖ WWN of a port shall never change over time
- **Fibre Channel switches and directors**
 - There can be just one FC switch; or several inter-connected FC switches
- **Multi-protocol routers**
 - If deploying IP-based SAN extension
- **Management software**

A Fibre Channel SAN is a collection of fibre channel nodes that communicate with each other – typically via fibre-optic media. A node is defined as a member of the fibre channel network. A node is provided a physical and logical connection to the network by a physical port on a Fibre Channel switch. Every node requires the use of specific drivers to access the network. For example, on a host, one has to install an HBA and the corresponding drivers to implement FCP (Fibre Channel Protocol, i.e. SCSI-3 over FC). These operating system-specific drivers are responsible for translating fibre channel commands into something the host can understand (SCSI commands), and vice versa.

Fibre Channel nodes communicate with each other via one or more Fibre Channel switches, also called Fabric Switches. The primary function of a fabric switch is to provide a physical connection and logical routing of data frames between the attached devices.

When needed, Fibre Channel SANs can be extended over geographically vast distances. The inter-connection between geographically disparate SANs is achieved using an IP network. SAN extension via IP requires the use of one or more multi-protocol routers at each participating site. The IP-based protocols used for SAN extension will be covered briefly in a later section.

Services Provided by a Fabric

- **Login Service**
 - Used by every node when it performs a Fabric Login (FLOGI)
 - Tells the node about its physical location in the fabric
- **Name Service**
 - Node registers with this service by performing a Port Login (PLOGI)
 - Database of registered names, stored on every switch in the fabric
- **Fabric Controller**
 - Sends state change notifications to nodes (RSCN's)
- **Management Server**
 - Provides access point for all services, subject to configured zones

When a device logs into a fabric, its information is maintained in a database. Information required for it to access other devices, or changes to the topology, is provided by another database. The following are the common services found in a fabric:

- **Login Service:** The Login Service is used by all nodes when they perform a Fabric Login (FLOGI). For a node to communicate in a fabric, it has to register itself with this service. When it does so, it sends a Source Identifier (S_ID) with its ALPA ID (Arbitrated Loop Physical Address id). The login service returns a D_ID to the node with the Domain ID and port location information filled in. This gives the node information about its location in the fabric that it can now use to communicate with other nodes.
- **Name Service:** The Name Service stores information about all devices attached to the fabric. The node registers itself with the name server by performing a PLOGI. The name server stores all these entries in a locally resident database on each switch. Each switch in the fabric topology exchanges its Name Service information with other switches in the fabric to maintain a synchronized, distributed view of the fabric.
- **Fabric Controller:** The Fabric Controller service provides state change notification to all registered nodes in the fabric, using RSCNs (Registered State Change Notifications). The state of an attached node can change for a variety of reasons: for example, when it leaves or rejoins the fabric.
- **Management Server:** The role of this Server is to provide a single access point for all three services above, based on virtual “containers” called zones. A zone is a collection of nodes defined to reside in a closed space. Nodes inside a zone are aware of nodes in the zone they belong to, but not outside of it. A node can belong to any number of zones.

Fibre Channel (FC): Protocol Layers

- Fibre Channel SANs use SCSI-3 over FC for transport
- Fibre Channel is a serial protocol with defined standards
 - The standards are determined by the Fibre Channel Alliance
 - Available from <http://www.t11.org>
- The standards define a layered communications stack for FC
 - Similar to the OSI model used for IP

OSI layer #	name	TCP/IP	Fibre Channel
5-7	application	telnet, ftp, SCSI-3 (iSCSI)	IP, SCSI-3 (FCP)
4	transport	TCP, UDP	FC-4
3	network	IP, ICMP, IGMP	FC-3
2	data link	Ethernet, Token Ring	FC-2, most of FC-1
1	physical	media	FC-0

The table shows the layers of the TCP/IP stack, and their corresponding analogues (FC-0 through FC-4) in the Fibre Channel protocol specification.

EMC Education Services

EMC²
where information lives[®]

Fibre Channel Frame ↔ TCP Packet

Fibre Channel Frame Structure:

Idles	SOF	Header	Optional Headers & Payload	CRC	EOF	Idles
	1 TW	6 TW	0 - 528 Transmission Words (TW)	1 TW	1 TW	
			64 Bytes Optional Header 2048 Bytes Payload			
6 TW						6 TW
537 Transmission Words						

TCP Packet Structure:

TCP Header (20 bytes)					0-1460 TCP Data	
IP Header (20 bytes)					IP Data	
Ethernet Data					CRC	
Preamble	Destination Address	Source Address	Type or Length	Ethernet Data		CRC
8	6	6	2	0-1500		4
1514 bytes						

- Fibre Channel standard (FC-2 layer) defines the Fibre Channel frame
- Frame is the basic unit of data transfer within FC networks
- A frame in FC networks is analogous to a TCP packet in IP networks
 - FC frame: up to 2112 bytes of payload; 36 bytes of fixed overhead
 - TCP packet: up to 1460 bytes of payload; 66 bytes of fixed overhead
 - Overhead includes: TCP header, IP header; Ethernet addressing, preamble, CRC

© 2006 EMC Corporation. All rights reserved.
SAN Foundations
Mercer Road
EMC Education Services

FC-2 specifies the structure of the Frame, which is the basic unit of data transfer within an FC network.

Note: In the FC Frame picture, some sizes are specified in Transmission Words. A Transmission Word (TW) is four bytes long.

FC Protocol: Features

- Mechanisms within a SAN depend on FC features specified by the standards

<i>FC layer</i>	<i>Function</i>	<i>SAN-relevant features specified by FC layer</i>
FC-4	mapping interface	mapping Upper Layer Protocol (e.g. SCSI-3) to FC transport
FC-3	common services	(placeholder layer)
FC-2	routing, flow control	frames, topologies, ports, FC addressing, buffer credits
FC-1	encode/decode	8B/10B encoding, transmission protocol
FC-0	physical layer	connectors, cables, FC devices

Specific Fibre Channel features that the standards define and that are relevant to Fibre Channel SANs.

Physical Specifications (FC-0 layer)

- FC-0 specifies the physical connection
 - Standard allows for either copper or optics as physical medium
 - Modern SANs use fibre optic cabling
- Optical connector specifications
 - SC connector: 1 Gb/sec
 - LC connector: 2 Gb/sec
- Optical cable can be of several types
 - Multi-mode cable
 - “Multi-mode” means light is transmitted on different wavelengths simultaneously
 - ❖ impacted by modal dispersion, i.e. the various light beams lose shape over long cable runs
 - Has an inner diameter of either 62.5 microns or 50 microns
 - Can be used for short distances: 500 meters or less
 - Single-mode cable
 - Has an inner diameter of 9 microns
 - Always used with a long-wave laser
 - ❖ This significantly limits the effects of modal dispersion
 - Works for distances up to 10 km or more

Today, Fibre Channel over copper is mostly used for loop connectivity within storage arrays, between Fibre Channel disk drives and other internal components. Over short distances, copper can be superior to optics in some respects, for example, it can provide better signal-to-noise ratio.

Logical Specifications (FC-2 layer)

- FC topologies: Point-to-point, FC-AL and FC-SW
- Structure of a frame
- Fibre Channel Address
 - **Not** the same as the WWN, which can **never** change!
 - 24-bit address: in hexadecimal notation, of the form: XYZZZ
 - *Dynamically* assigned when node connects to switched fabric
 - Used to route frames from source to destination
 - **Will** change if re-cabled to another switch port
- Port Types
- Buffer Credits
 - Basic mechanism for flow control

This slide lists some of the key entities defined within layer 2 of the FC standard (FC-2).

Fibre Channel Address: A Fibre Channel address is a 24-bit identifier that is used to designate the source and destination of a frame in a Fibre Channel network. A fibre channel address is analogous to an Ethernet or Token Ring address. Unlike MAC addresses and Token Ring addresses however, these addresses are not “burned in”. They are assigned when the node is connected to a switched fabric, or enters a loop.

Port Type: Querying the fabric switches for negotiated port types is a useful diagnostic mechanism. A frequent cause of initial connectivity problems is a misconfigured host driver, which causes the wrong port type to be negotiated (FC-AL instead of FC-SW, and vice-versa). All connected host HBAs and storage array ports in a switched fabric should register as F-ports on the Fibre Channel switches. Ports used for Inter-Switch Links should register as E-ports on the switches at either end.

Buffer Credits: Specifies how many frames can be sent to a receiving port when flow control is in effect. The receiving port indicates its Buffer Credit. After sending this many frames, the sending port shall wait for a “Ready” indication. This parameter can be especially critical to the performance of long-distance ISLs (Inter-Switch Links). We shall examine this in greater detail during our coverage of ISLs.

Expanding SANs - Fabric Topologies

- Fabric topologies: different ways to connect FC switches to serve a specific function
 - Switches can be connected to each other using ISLs to create a single large fabric
 - A Fibre Channel SAN can be expanded by adding in one or more FC switches or directors
 - More FC ports become available for connecting hosts or storage frames
- Design considerations for a fabric topology:
 - Redundancy
 - Scalability
 - Performance

Switches can be connected in different ways to create a fabric. The type of topology to be used depends on requirements such Availability, Scalability, cost and performance. Typically, there is no single answer to the question as to which topology is best suited for an environment.

Topology: Storage Consolidation

- Fan-out ratio
 - Qualified maximum number of initiators that can access a single storage port through a SAN
- Allows storage to be consolidated and hence utilized more efficiently
- Ratio varies depending on HBA type and O/S
 - Check EMC Support Matrix



© 2006 EMC Corporation. All rights reserved.

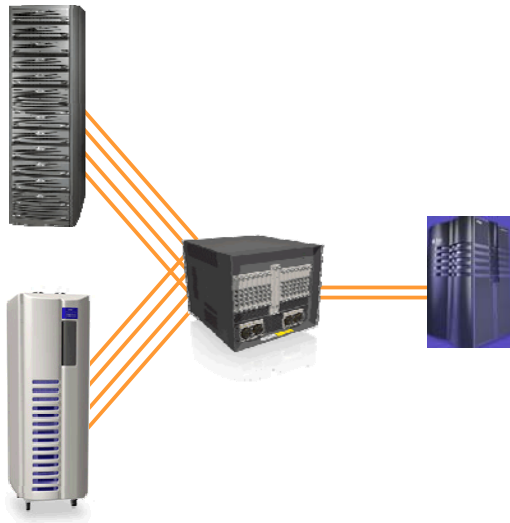
SAN Foundations

Mercer Road
EMC Education Services

Fan-Out ratio is a measure of the number of hosts that can access a Storage port at any given time. Storage consolidation enables customers to achieve the full benefits of using Enterprise Storage. This topology allows customers to map multiple host HBA ports onto a single Storage port, for example, a Symmetrix FA port.

The Fan-Out implementation is highly dependent on the I/O throughput requirements of customer applications. There are no hard-and-fast acceptable figures for the fan-out ratio. At least a rudimentary analysis of the anticipated workload from all participating hosts is required to establish acceptable fan-out for a given customer environment.

Topology: Capacity Expansion



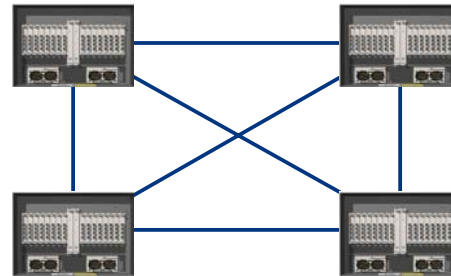
- Fan-In ratio
 - Qualified maximum number of storage ports that can be accessed by a single initiator through a SAN
- Solves the problem of capacity expansion
- Ratio varies depending on HBA type and O/S
 - Check EMC Support Matrix

Fan-In ratio is a measure of how many storage systems can be accessed by a single host at any given time. This allows a customer to expand connectivity by a single host across multiple storage units. There can be situations where a host requires additional storage capacity and additional space is carved from a new or existing storage unit that was previously used elsewhere. This topology then allows a host to see more storage devices.

As with fan-out, expanding the fan-in on a host requires careful consideration of the extra I/O load on the HBAs from accessing the newly-provisioned storage. Frequently, adding more HBAs on the host may become a requirement for performance reasons.

Topology: Mesh Fabric

- Can be either partial or full mesh
- All switches are connected to each other
- Pros/Cons
 - Maximum Availability
 - Medium to High Performance
 - Poor Scalability
 - Poor Connectivity



A full mesh topology has all switches connected to each other. A partial mesh topology is when there are some switches not interconnected. For example, consider the graphic above without the diagonal ISLs – this would be a partial mesh.

The path for traffic between any two end devices (hosts and storage) depends on whether they are localized or not. If a host and the storage it is communicating with are localized (i.e. they are connected to the same switch), traffic passes over the back plane of that switch only avoiding ISLs. If the devices are not localized, then traffic has to travel over at least one ISL (or a hop) to reach its destination, regardless of where they are located in the fabric. If a switch fails, an alternate path can be established using the other switches. Thus, a high amount of localization is needed to ensure that the ISLs don't get overloaded.

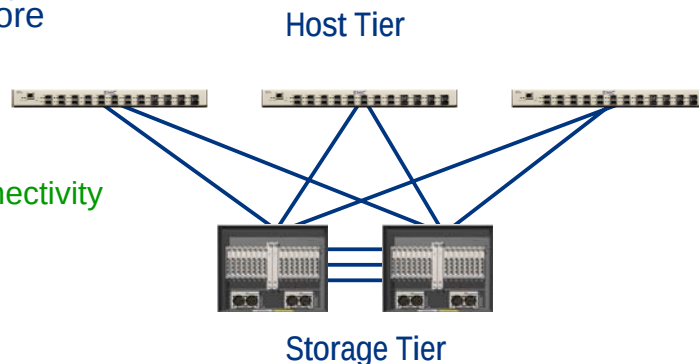
The full mesh topology provides maximum availability. However, this is done at the expense of connectivity – which can become prohibitively expensive with an increasing number of switches increases. For every switch that gets added, an extra ISL is needed to every one of the existing switches. This reduces the port count available for connecting hosts and storage.

Features of a Mesh topology:

- Maximum of one ISL hop for host to storage traffic
- Host and storage can be located anywhere in the fabric
- Host and storage can be localized to a single director or switch
- High level of localization results in ISLs used only for managing the fabric

Topology: Simple Core-Edge Fabric

- Can be two or three tier
- Single Core Tier
- One or two Edge Tiers
- In a two tier topology, storage is usually connected to the Core
- Benefits
 - High Availability
 - Medium Scalability
 - Medium to maximum Connectivity



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

This topology can have two variations: two-tier (one edge and one core) or three-tier (two Edge and one Core). In a two-tier topology – shown in the picture - all hosts are connected to the edge tier, and all storage is connected to the core tier. With three-tier, all hosts are connected to one edge; all storage is connected to the other edge; and the core tier is only used for ISLs.

In this topology, all node traffic has to traverse at least one ISL hop. There are two types of switch tiers in the fabric: Edge tier and the Core, or Backbone tier. The functions of each tier are:

Edge Tier

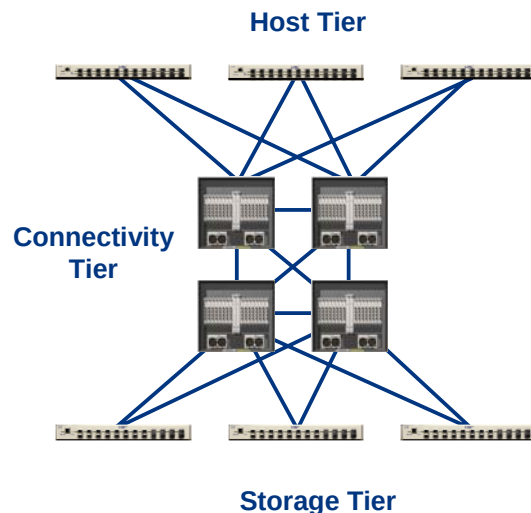
- Usually Departmental Switches; this offers an inexpensive approach to adding more hosts into the fabric
- Fans out from the Core tier
- Nodes on the edge tier can communicate with each other using the Core tier only
- Host to Storage Traffic has to traverse a single ISL (two-tier) or two ISLs (three-tier)

Core or Backbone Tier

- Usually Enterprise Directors; this ensures the highest availability since all traffic has to either traverse through or terminate at this tier
- Usually two directors/switches are used to provide redundancy
- With two-tier, all storage devices are connected to the core tier, facilitating fan-out
- Any hosts used for mission critical applications can be connected directly to the storage tier, thereby avoiding ISLs for I/O activity from those hosts
- If the storage and host tier are spread out across campus distances, the core tier can be extended using ISLs based on shortwave, longwave or even DWDM (Dense Wavelength Division Multiplexing)

Topology: Compound Core-Edge Fabric

- Core or Connectivity Tier is made up of switches configured in a full mesh topology
- Core Tiers are only used for ISLs
- Edge Tiers are used for host or storage connectivity
- Benefits
 - Maximum Connectivity
 - Maximum Scalability
 - High Availability
 - Maximum Flexibility



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

This topology is a combination of the Full Mesh and Core-Edge three-tier topologies. In this configuration, all host to storage traffic must traverse the Connectivity Tier. The Connectivity or Core tier is used for ISLs only. This permits stricter policies to be enforced, allowing distributed administration of the SAN. Fabrics of this size are usually designed for maximizing port count. This type of a topology is also found in situations where several smaller SAN “islands” are consolidated into a single large fabric, or where a lot of SAN-NAS integration requires everything to be plugged together for ease of management, or for backups.

Functions of the three tiers are:

Host Tier

- All hosts connected at the same hierarchical point in the fabric
- Fans out from the Connectivity Tier
- Minimum of two ISL hops for all host FC traffic to reach destination point
- Nodes on the edge tier can communicate with each other using the Core tier only

Connectivity Tier

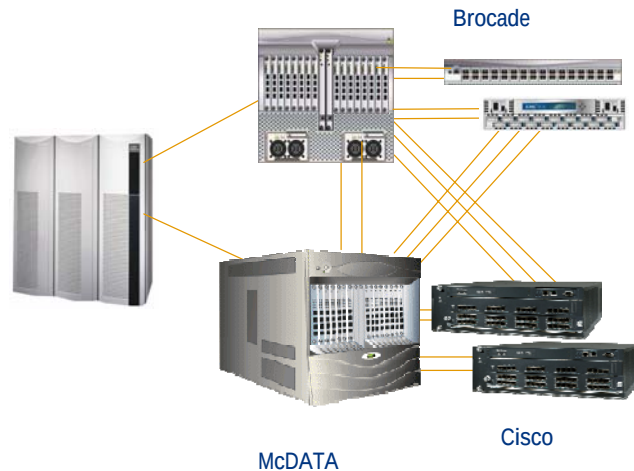
- Bridging point for all host and storage traffic
- No hosts or storage are located in this tier so it can be dedicated for ISL traffic

Storage Tier

- All storage can be connected to the same tier
- Fans out from the Connectivity Tier
- Nodes on the edge tier can communicate with each other using the Core tier only
- Storage and hosts used for mission critical applications can connect to the same tier if needed. Traffic need not traverse an ISL if it does not need to. However this is more of an exception than the rule.

Heterogeneous Fabrics

- Heterogeneous switch vendors within same fabric
- Limited number of switches in the fabric
- Limited number of ISL hops
- Refer to EMC Topology Guide



Usually topologies are designed using switches from the same vendor. This presents a problem when consolidating SANs made from different vendor switches. EMC supports a mode called Open Fabric to interconnect Brocade, Cisco and/or McDATA switches. This can be used in such special situations. The slide above provides an example of possible Open Fabric configurations. Technically speaking, Open Fabric is not really a topology but more of a supported configuration.



Expanding Fabric Connectivity: Inter-Switch Links (ISLs)

- Use
 - Expand fabric connectivity
 - Bandwidth expansion
 - Multiple ISLs aggregated to create a single logical ISL with higher bandwidth
- Factors influencing ISLs
 - Distance
 - Resilience to failure
 - Performance and redundancy
 - Availability and Accessibility
- Best practices
 - For Directors: Connect ISLs across different port cards
 - For Departmental switches: Connect ISLs to different switch ports, and/or different ASIC



ISL

- Oversubscription Ratio is a measure the theoretical utilization of an ISL
- EMC Support Matrix specifies ISL limits for individual switch vendors

© 2006 EMC Corporation. All rights reserved.

SAN Foundations



Switches are connected to each other in a fabric using Inter-switch Links (ISL). This is accomplished by connecting them to each other through an expansion port on the switch (E_Port). ISLs are used to transfer node-to-node data traffic, as well as fabric management traffic, from one switch to another. Thus, they can critically affect the performance and availability characteristics of the SAN. In a poorly-designed fabric, a single ISL failure can cause the entire fabric to fail. An overloaded link can cause an I/O bottleneck. Therefore, it is imperative to have a sufficient number of ISLs to ensure adequate availability and accessibility.

If at all possible, one should avoid using ISLs for host-to-storage connectivity whenever performance requirements are stringent. If ISLs are unavoidable, the performance implications should be carefully considered at the design stage.

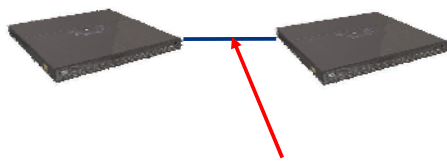
Distance is also a consideration when implementing ISLs. We explore the implications of distance in greater detail in the next slide.

Over subscription ratio as it applies to an ISL is defined as the number of nodes or ports that can contend for its bandwidth. This is calculated as the ratio of the number of initiator attached ports to the number of ISL ports on a switch. In general, a high oversubscription ratio can result in link saturation on the ISLs, leading to high I/O latency.

When adding ISLs in a fabric, there are some basic best practices such as, always connect each switch to at least two other switches in the fabric. This prevents a single link failure from causing total loss of connectivity to nodes on that switch. Also, for host-to-storage connectivity across ISLs, use a mix of equal-cost primary paths.

ISL - Distance and Cables

- Operating distances decrease when moving from 1Gbps to 2Gbps
- Media options
 - Multi-mode
 - Single-mode
 - DWDM



Multi-mode 1Gb=500m 2Gb=300m
Single-mode > 10Km
DWDM < 200Km

- ISL design parameters
 - Capacity
 - Distance
 - Signal loss
 - Throughput
 - Power

Fibre Optic Glass Filament Core	Port Speed	Operating Distance
50 micron Multimode	1Gbps	500m
50 micron Multimode	2Gbps	300m
62.5 micron Multimode	1gbps	~300m
62.5 micron Multimode	2Gbps	~150m
9 micron Single mode	1Gbps	>10km
9 micron Single mode	2Gbps	>10km

There are three media options available when implementing an ISL:

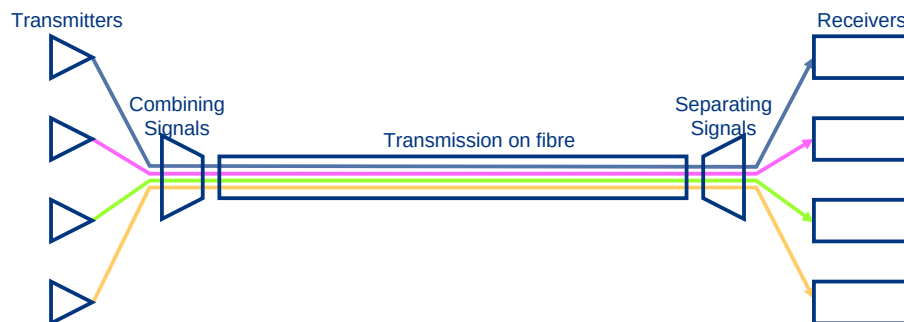
- Multimode ISL: For distances up to 500m
- Single-mode ISL: For distances up to 35 km; depends on switch port transceiver technology
- DWDM ISL: DWDM (Dense Wavelength Division Multiplexing) is typically used for distances up to 200 km. DWDM can be configured with multi-mode or single-mode.

Variables that affect Supportable Distance:

- Propagation and Dispersion Losses: For best possible long-distance results with a conventional Fibre Channel link, use long wave laser over single-mode 9-micron cable. This is the least susceptible to modal dispersion, thereby enabling distances up to 35 km.
- Port speed: Multi-mode cable exhibits increased susceptibility to modal dispersion as the port speed increases. From the table above, with 50-micron multi-mode, the maximum operating distance decreases from 500m to 300m when port speed is increased from 1 Gbps to 2 Gbps. In contrast, single-mode fiber is not affected by higher port speeds.
- Buffer-to-Buffer Credit: Throughput on long links can degrade quickly if not enough frames are on the link. The longer the link, the more frames must be sent contiguously down the link to prevent this degradation. This is because the signal itself propagates at the speed of light. Due to this, a standard Fibre Channel frame of 2 KB is approximately 4 km long. Transmitting a 2 KB frame across a 100 km link is similar to a 4 km-long train on a 100 km track – the track remains vastly under-utilized with only a single train on the track. Configuring sufficient Buffer Credits ensures that the pipe is used efficiently.
- Optical power- Sufficient signal power is needed at the transmitting and receiving end to account for signal loss due to environmental conditions.

DWDM

- Data are carried at different wavelengths over fiber links
- Different data formats can be transmitted together (e.g. IP, ESCON SRDF, Fibre Channel SRDF)
- DWDM topologies include Point-to-Point and Ring configurations



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

DWDM (Dense Wavelength Division Multiplexing) is a protocol in which different channels of data are carried at different wavelengths over the same pair of fiber links. This is in contrast to a conventional fiber optic link, in which just one channel is carried over a single wavelength over a single fiber.

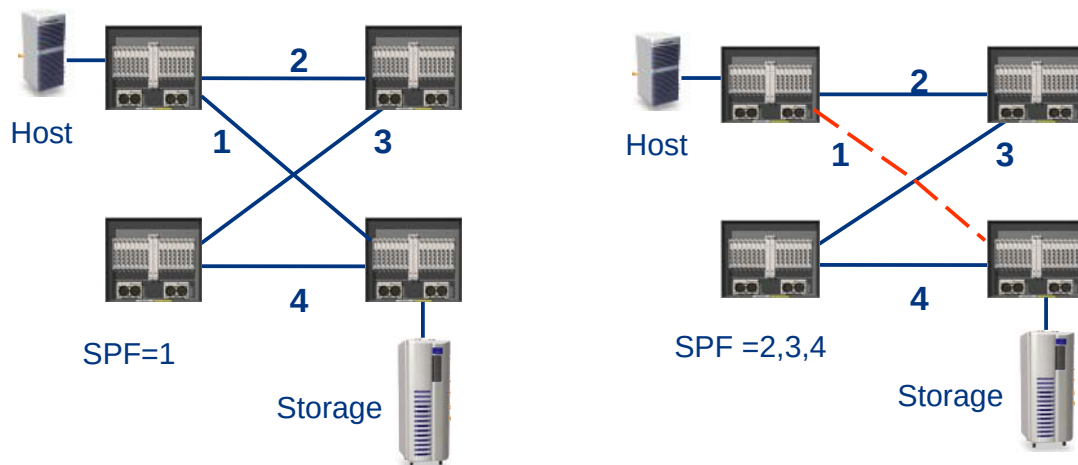
Using DWDM, several separate wavelengths (or channels) of data can be multiplexed into a light stream transmitted on a single optical fiber. Each wavelength can carry a signal at any bit rate less than an upper limit defined by the electronics, typically up to several Gigabits per second.

Different data formats can be transmitted simultaneously on different channels. Examples of protocols that can be transmitted are IP, FCIP, iFCP, ESCON, Fibre Channel SRDF, SONET and ATM.

For EMC customers, DWDM enables multiple SRDF channels and Fibre Channel ISLs to be implemented using one pair of long-distance fiber links, along with traditional network traffic. This is especially important where fiber links are at a premium. For example, a customer may be leasing fiber, so the more traffic they can run over a single link, the more cost effective the solution.

Routing of Frames

- A Routing Table algorithm calculates the lowest cost Fabric Shortest Path First (FSPF) route for a frame
- Recalculated at each change in topology
- ISLs may remain unused



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Fibre Channel Frames are routed across the fabric via an algorithm that uses a combination of lowest cost metric and Fabric shortest-path-first (FSPF).

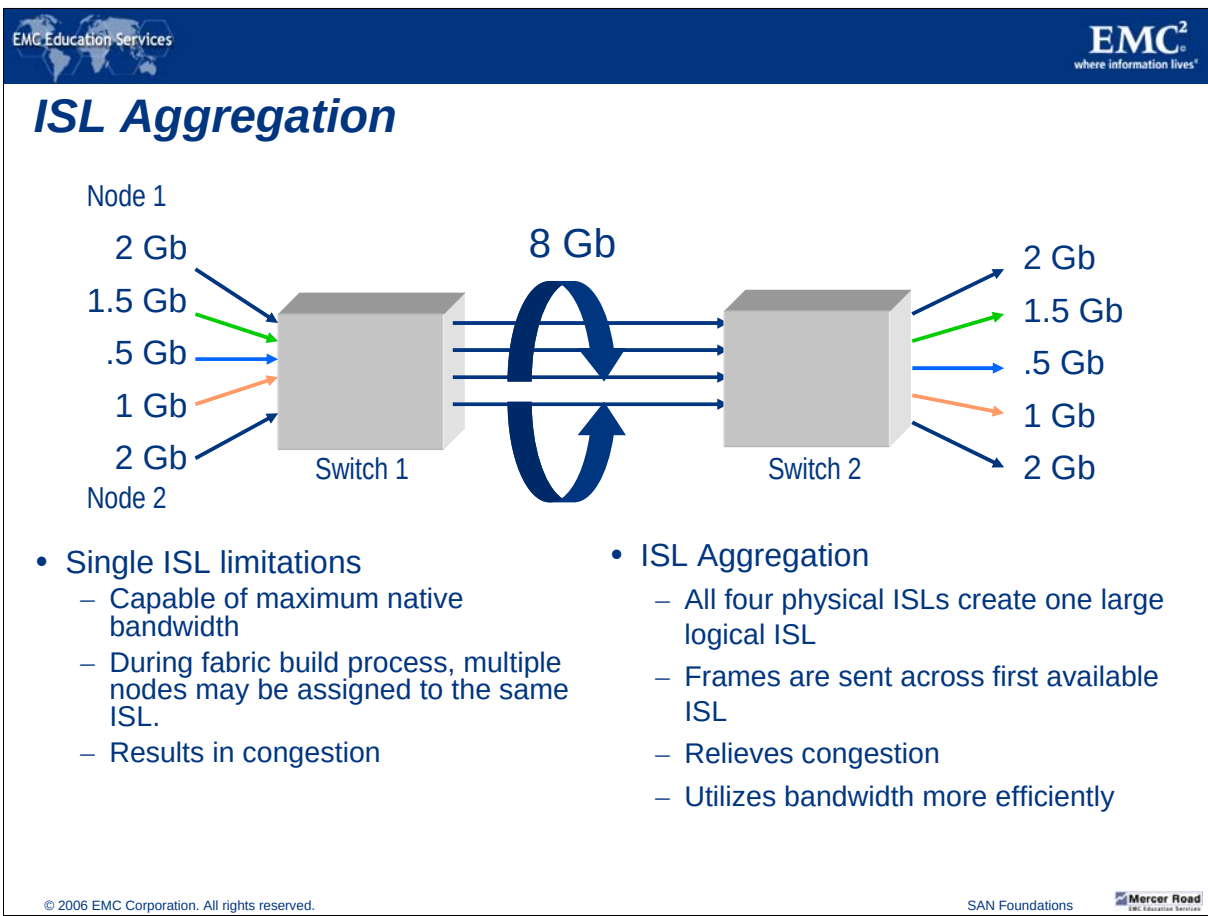
Lowest cost metric refers to the speed of the links in the routes. As the speed of the link increases, the cost of the route decreases. FSPF refers to the number of ISLs or “hops” between the host and its storage.

EMC strongly recommends that a fabric be constructed so that it has multiple equal, lowest-cost, shortest-path routes between any combination of host and storage. Routes that are not the shortest, lowest-cost path will not be used at all - until there is an event in the fabric that causes them to become the shortest, lowest-cost path. This is true even if currently active routes are close to peak utilization.

Routes are assigned to devices for each direction of the communication. The route one way may differ from the return route. Routes are assigned in a round-robin fashion after the device is logged into the fabric. These routes are static for as long as the device is logged in.

Routing tables on each switch are updated during events that change the status of links in the system. The calculation of routes, and the switch’s ability to perform this function in a timely fashion, is important for fabric stability.

For this reason, as well as the fact that every ISL effectively removes two ports that would otherwise be available for connecting storage or hosts, EMC recommends using reasonable limits on the number of ISLs in a fabric. For a reliable estimate of required ISLs, ISL utilization should be periodically monitored, and the level of actual protection from link failures critically examined.



ISL Aggregation is a capability supported by some vendors to enable distribution of traffic over the combined bandwidth of two or more ISLs.

ISL aggregation ensures that all links are used efficiently, eliminating congestion on any single link, while distributing the load across all the links in a trunk. Each incoming frame is sent across the first available ISL. As a result, transient workload peaks for one system or application are much less likely to impact the performance of other parts of a SAN.

In the example portrayed above, four ISLs are combined to form a single logical ISL with a total capacity of 8Gbps. The full bandwidth of each physical link is available for use and hence bandwidth is efficiently allocated.

SAN Foundations

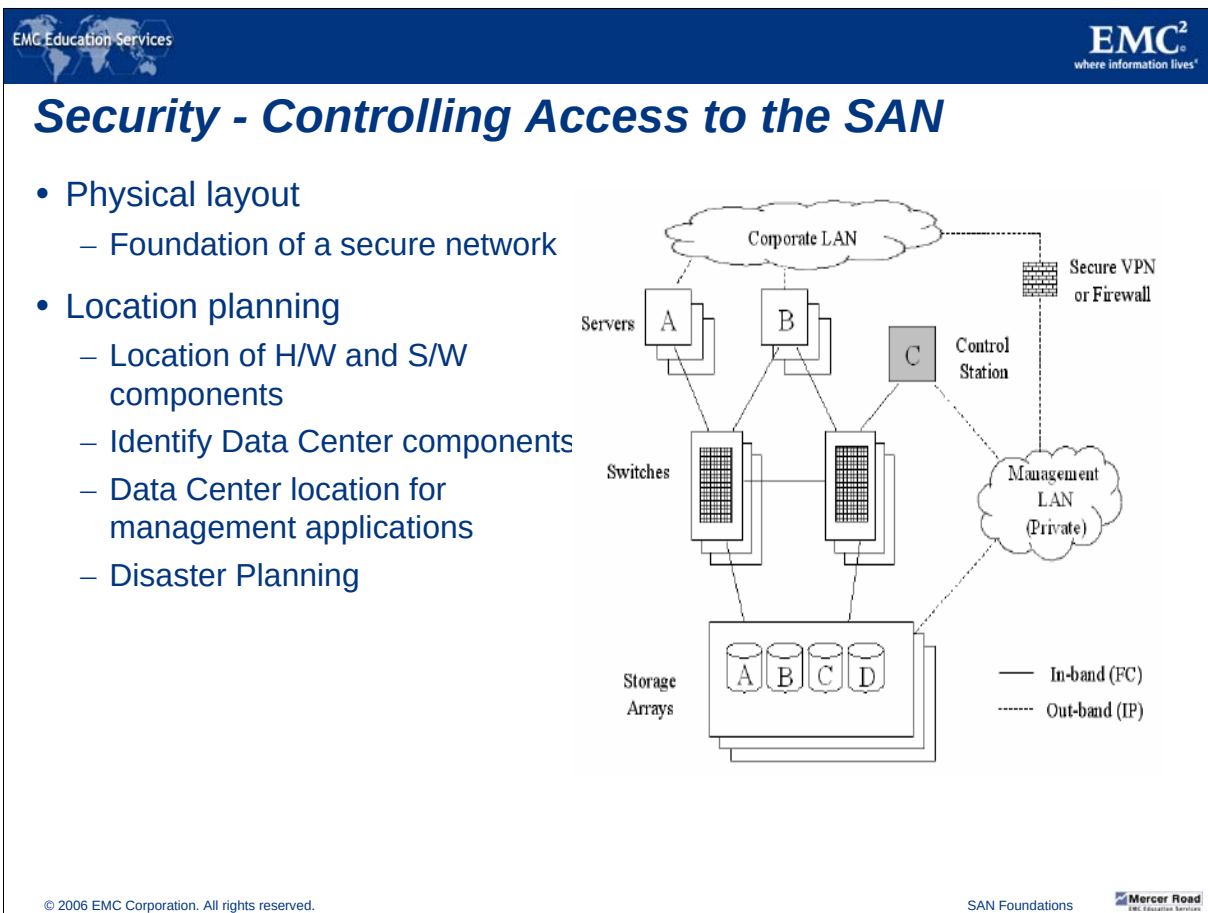
Securing a SAN

© 2006 EMC Corporation. All rights reserved.

SAN Foundations


Mercer Road
EMC Education Services

Security mechanisms available within a Fibre Channel SAN.



Planning the physical location of all components is an essential part of storage network security. Building a physically secure data center is only half the challenge; deciding where hardware and software components need to reside is the other, more difficult, half.

Critical components such as storage arrays, switches, control stations and hosts running management applications should reside in the same data center. With physical security implemented, only authorized users should have the ability to make physical or logical changes to the topology (for example, move cables from one port to another, reconfigure access, add/remove devices to the network etc.).

Planning should also take into account environmental issues such as cooling, power distribution and requirements for disaster recovery.

At the same time, one has to ensure that the IP networks that are used for managing various components in the SAN are secure and not accessible to the entire company. It also makes sense to change the default passwords for all the various devices to prevent unauthorized use. Finally, it helps to create various administration hierarchies in the management interface so that responsibilities can be delegated.

Fabric Security - Zoning

- Zone
 - Controlled at the switch layer
 - List of nodes that are made aware of each other
 - A port or a node can be members of multiple zones
- Zone Set
 - A collection of zones
 - Also called zone config
- EMC recommends Single HBA Zoning
 - A separate zone for each HBA
 - Makes zone management easier when replacing HBAs
- Types of zones:
 - Port Zoning (Hard Zoning)
 - Port-to-Port traffic
 - Ports can be members of more than one zone
 - Each HBA only “sees” the ports in the same zone
 - If a cable is moved to a different port, zone has to be modified
 - WWN based Zoning (Soft Zoning)
 - Access is controlled using WWN
 - WWNs defined as part of a zone “see” each other regardless of the switch port they are plugged into
 - HBA replacement requires the zone to be modified
 - Hybrid zones (Mixed Zoning)
 - Contain ports and WWNs

© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Zoning is a switch function that allows devices within the fabric to be logically segmented into groups that can communicate with each other. When a device logs into a fabric, it is registered by the name server. When a port logs into the fabric, it goes through a device discovery process with other devices registered as SCSI FCP in the name server. The zoning function controls this process by only letting ports in the same zone establish these link level services.

A collection of zones is called a zone set. The zone set can be active or inactive. An active zone set is the collection of zones currently being used by the switched fabric to manage data traffic.

Single HBA zoning consists of a single HBA port and one or more storage ports. A port can reside in multiple zones. This provides the ability to map a single Storage port to multiple host ports. For example, a Symmetrix FA port or a CLARiiON SP port can be mapped to multiple single HBA zones. This allows multiple hosts to share a single storage port.

The type of zoning to be used depends on the type of devices in the zone and site policies.

- In port zoning, only the ports listed in the zone are allowed to send Fibre Channel frames to each other. The switch software examines each frame of data for the Domain ID of the switch, and the port number of the node, to ensure it is allowed to pass to another node connected to the switch. Moving a node that is zoned by a port zoning policy to a different switch port may effectively isolate it. On the other hand, if a node is inadvertently plugged into a port that is zoned by a port zoning policy, that port will gain access to the other ports in the zone.
- WWN zoning creates zones by using the WWNs of the attached nodes (HBA and storage ports). WWN zoning provides the capability to restrict devices, as specified by their WWPNS, into zones. This is more flexible, as moving the device to another physical port with the fabric cannot cause it to lose access to other zone members.

Zoning - Hard vs. Soft Zoning

	Advantages	Disadvantages
Port Zoning	– More Secure – Simplified HBA replacement	– Reconfiguration
WWPN Zoning	– Flexibility – Reconfiguration – Troubleshooting	– “Spoofing” – HBA replacement

© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Port zoning advantages: Port zoning is considered more secure than WWN zoning, because zoning configuration changes must be performed at the switch. If physical access to the switch is restricted, the potential for unauthorized configuration changes is greatly reduced. Also, HBAs can be replaced without requiring modification of zone configurations.

Port zoning disadvantages: Switch port replacement and the use of spare ports require manual changes to the zone configuration. If the domain ID changes – e.g. when a set of independent switches are linked to form a multi-switch fabric - the zoning configuration becomes invalid. Replacing an HBA requires reconfiguration of the volume access control settings on the storage subsystem. This minimizes the benefit of hard zoning, because manual configuration changes will still be necessary to get things working again.

WWN zoning advantages: The zone member identification will not change if the fiber cable connections to switch ports are rearranged. Fabric changes such as switch addition or replacement do not require changes to zoning.

WWN zoning disadvantages: It is possible to change an HBA's WWN to match the current WWN of another HBA (commonly referred to as “spoofing” *). Replacement of a damaged HBA requires the user to update the zoning information and the volume access control settings.

* HBA spoofing implies that a compromise of security has already been made at the root level on the host in question. Once this compromise has been completed, the host is vulnerable to HBA spoofing and other types of data interception. However, HBA spoofing should also be considered a serious risk to any other host attached to either the SAN or array in the environment.

Fabric Security - Vendor Specific Access Control

- Most vendors have proprietary access control mechanisms
- These mechanisms are not governed by the Fibre Channel standard
- Examples of vendor features:
 - McDATA
 - Port Binding
 - SANtegrity
 - Brocade
 - Secure FabricOS

McDATA has developed Port Binding and SANtegrity to add further security to a Fabric:

- Port binding uses the WWN of a device to create an exclusive attachment to a port. When port binding is enabled, the only device that can attach to a port is the one specified by its WWN.
- SANtegrity enhances security in SANs that contain a large and mixed group of fabrics and attached devices. It can be used to allow or prohibit switch attachment to fabrics and device attachment to switches. This prevents Fibre Channel traffic from being directed to the incorrect port, device or domain thereby enforcing the policy for that SAN.

Brocade has developed the Secure FabricOS environment. In this environment, in addition to device based access control, switch to switch trusts can be set up.

EMC Education Services

EMC²
where information lives[®]

Security: Volume Access Control (LUN Masking)

- Restricts volume access to specific hosts and/or host clusters
- Policies set based on functions performed by the host
- Servers can only access volumes that they are permitted to access
- Access controlled in the Storage Array - **not** in the fabric
 - Makes distributed administration secure
- Tools to manage masking
 - GUI
 - Command Line

© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Device (LUN) Masking ensures that volume access to servers is controlled appropriately. This prevents unauthorized or accidental use in a distributed environment.

A zone set can have multiple host HBAs and a common storage port. LUN Masking prevents multiple hosts from trying to access the same volume presented on the common storage port. LUN Masking is a feature offered by EMC Symmetrix and CLARiiON arrays.

When servers log into the switched fabric, the WWNs of their Host Bus Adapters (HBAs) are passed to the storage fibre adapter ports that are in their respective zones. The storage system records the connection and builds a filter listing the storage devices (LUNs) available to that WWN, through the storage fibre adapter port. The HBA port then sends I/O requests directed at a particular LUN to the storage fibre adapter. Each request includes the identity of their requesting HBA (from which its WWN can be determined) and the identity of the requested storage device, with its storage fibre adapter and logical unit number (LUN). The storage array processes requests to verify that the HBA is allowed to access that LUN on the specified port. Any request for a LUN that an HBA does not have access to returns an error to the server.

LUNs can be masked through the use of bundled tools. For EMC platforms these include ControlCenter; Navisphere or Navicli for CLARiiON; and Solutions Enabler (SYMCLI) for a Symmetrix.

Host Considerations for Fabric-Attach

- Host Bus Adapters should have a supported firmware version, and, a supported driver for the operating system
 - EMC Support Matrix provides exhaustive data for server models from specific manufacturers, HBA models, and for each storage array model
- Persistent Binding must be used if the operating system requires it
 - Prevents controller IDs/device names from changing, when new storage targets become visible to the host
- Multipathing software (e.g. Powerpath) can provide high availability and better performance
 - Protects against HBA failures, storage port failures or path failures
 - Can also distribute I/O load from the host over all available, active paths

HBA options: EMC supports a variety of Emulex and Qlogic fibre Channel HBAs on several operating systems, including: Windows Server, Solaris, and Linux.

AIX (IBM) and HP-UX (Hewlett-Packard) servers typically use factory-supplied HBAs with native OS drivers. The EMC Support Matrix lists the qualified driver versions on these boards.

Host Connectivity Guides are available on Powerlink for all supported host operating systems.

SAN Foundations

IP-Based SANs and SAN Extensions

© 2006 EMC Corporation. All rights reserved.

SAN Foundations

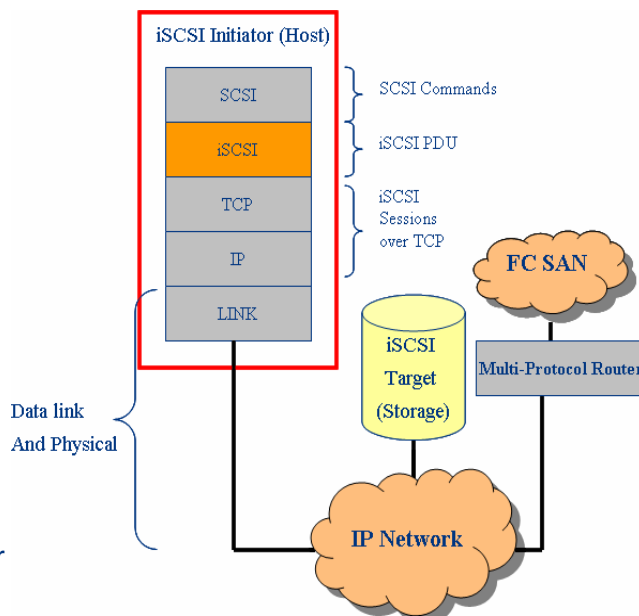
Mercer Road
EMC Education Services



This section covers iSCSI, and IP-based SAN extension via FCIP or iFCP.

IP SANs: Overview

- IP SANs use iSCSI
 - Serial SCSI-3 over IP
 - Uses TCP/IP for transport
 - Block-level I/O
 - Standard SCSI command set
- iSCSI concepts:
 - Network Entity
 - Network Portal
 - ❖ Initiator - Software or HBA
 - ❖ Target - Storage port
 - iSCSI Node
 - Portal group
 - Internet Storage Name Server (iSNS)



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

iSCSI is becoming popular in the new generation Storage Area Networks. Unlike Fibre Channel SANs, IP SANs use the iSCSI protocol over standard IP networks for host-to-storage communications. iSCSI is also becoming an increasingly popular mechanism to bridge disparate SAN islands and fabrics into a single large fabric. These advantages allow companies to leverage their existing investment in IP technologies to grow their Storage networks.

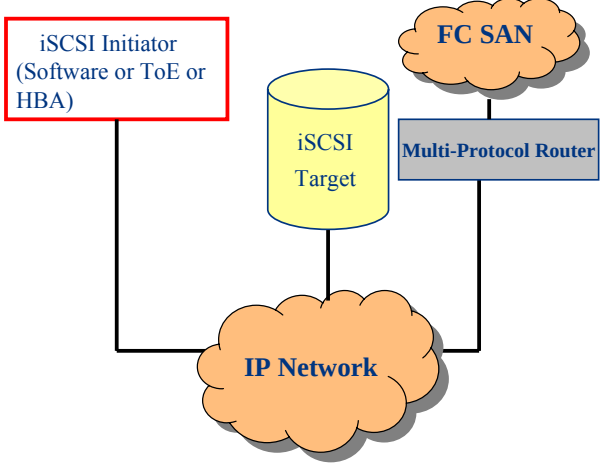
In an IP SAN, hosts communicate with Storage Arrays using Serial SCSI-3 over IP. Gigabit Ethernet (GigE) is a commonly used medium for connectivity. This eliminates the need for a Fibre Channel HBA on the host. Modern server-class hosts typically ship with two network ports (NICs) in their factory configuration, with at least one port being GigE-capable. So no extra hardware may be needed on the host for iSCSI connectivity.

A network entity is a device (a client, server or gateway) that is connected to an IP network. It contains one or more network portals. A network portal is a component within a network entity that is responsible for the TCP/IP protocol stack. Network portals consist of an initiator portal that is identified by its IP address, and a target portal that is identified by its IP address and listening port. An initiator makes a connection to the target at the specified port, creating an iSCSI session. An iSCSI initiator or target identified by its iSCSI address is known as an iSCSI node. A portal group is a set of network portals that support an iSCSI session that is made up of multiple connections over different network portals. iSCSI supports multiple TCP connections within a session. Each session can be across multiple network portals. Similar to DNS in the IP world, iSNS acts like a query database in the iSCSI world. iSCSI initiators can query the iSNS and discover iSCSI targets.



IP SANs (continued.)

- iSCSI Initiators can be
 - Software based
 - TCP Offload Engine (ToE)
 - iSCSI Host Bus Adapters
- All iSCSI nodes identified by an iSCSI name or address
- iSCSI addressing
 - iSCSI Qualified Name (iQN)
 - IEEE Naming convention (EUI)



The diagram illustrates the architecture of an IP SAN. An iSCSI Initiator (Software or ToE or HBA) is connected to an iSCSI Target via an IP Network. The iSCSI Target is also connected to an FC SAN via a Multi-Protocol Router.

© 2006 EMC Corporation. All rights reserved.

SAN Foundations



Initiators can be implemented using one of three approaches, listed here in order of decreasing host-side CPU overhead:

- Software based drivers where all processing is performed by the host OS.
- TCP offload engines (ToE) where TCP/IP processing is performed at the controller level.
- iSCSI HBA, where all processing is performed by the controller. This requires a supported driver provided by the HBA manufacturer.

The problem with the more high-performance approaches – the ToE or the iSCSI HBA – is the significantly increased cost relative to a generic NIC. iSCSI HBAs and Fibre Channel HBAs are comparable in price.

All iSCSI nodes are identified by an iSCSI name. An iSCSI name is neither the IP address nor the DNS name of an IP host. iSCSI addresses can be one of two types - iSCSI Qualified Name (iQN) or IEEE Naming convention (EUI).

iQN format - iqn.ccyy-mm.com.xyz.aabbccddeeffgghh where

- iqn - Naming convention identifier
- ccyy-nn - Point in time when the .com domain was registered
- com.xyz - Domain of the node backwards
- aabbccddeeffgghh - Device identifier (can be a WWN or the system name or any other vendor implemented standard)

EUI format - eui.64-bit WWN

- eui - Naming prefix
- 64-bit WWN - FC WWN of the host.

IP SAN: Components

- **iSCSI host initiators**
 - Typically use Ethernet ports (NIC's), with a software implementation of iSCSI initiator on the host
- **iSCSI targets**
 - Storage arrays with GigE ports and native iSCSI support
- **Ethernet LAN for IP storage network**
 - Interconnected Ethernet switches and hubs
- **Multi-protocol routers**
 - If bridging to Fibre Channel arrays from iSCSI initiators is required
- **Management software**

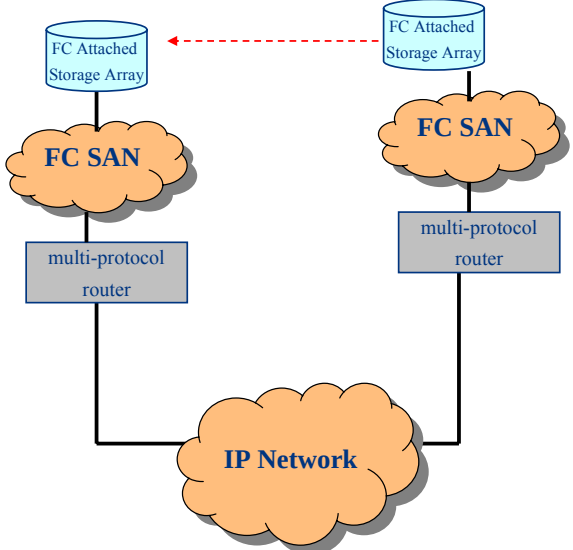
The typical components of an IP SAN are listed above.

Strictly speaking, an IP SAN requires no Fibre Channel components. In practice, however, bridging to existing Fibre Channel devices such as storage arrays is frequently a requirement. One or more multi-protocol routers are required for this purpose.



IP-Based SAN Extension: the FCIP and iFCP Protocols

- For SAN extension over vast distances
 - Geographically disparate sites, well beyond the limits of DWDM
- Primarily used for disaster recovery and array-based replication
 - Array-to-array connectivity is the principal application
- FCIP
 - Tunnels Fibre Channel frames over a TCP/IP network
 - Merges FC fabrics over long distances, to form a single fabric
- iFCP
 - Wraps FC data in IP packets
 - Maps IP addresses to individual FC devices
 - Fabrics are **not** merged



© 2006 EMC Corporation. All rights reserved. SAN Foundations 

With the use of multi-protocol routers, it is possible to extend traditional Fibre Channel SANs over long distances via an IP network. FCIP and iFCP are the two widely-used protocols for IP-based SAN extensions. SAN extension technology is primarily used for disaster recovery functions such as SRDF and MirrorView.

Fibre Channel over IP (FCIP) is a tunneling protocol. It allows one to merge two FC fabrics at two physically distant locations - well beyond the limits of DWDM - into a single large fabric.

Unlike FCIP, iFCP is a gateway-to-gateway protocol. iFCP wraps Fibre Channel data in IP packets, but maps IP addresses to individual Fibre Channel devices. Storage targets at either end can be selectively exposed to each other, by configuring the multi-protocol routers that serve as the gateways. However, the two fabrics are not merged.

When iFCP creates the IP packets, it inserts information that is readable by network devices, and routable within the IP network. Because the packets contain IP addresses, customers can use IP network management tools to manage the flow of Fibre Channel data using iFCP.

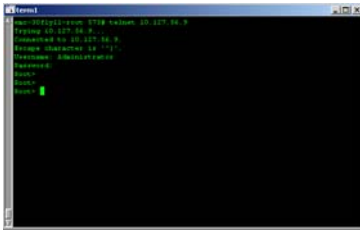
 Mercer Road
K-12 Education Services Mercer Road
EMC Education Services

EMC Education Services

EMC²
where information lives[®]

Connectrix: Management Tools

- Individual switch management:
 - Command line interface
 - Via Serial port, or
 - Via IP (telnet, ssh)
 - Required for initial configuration
 - Facilitates automation
 - Browser-based interface
- Fabric-wide management and monitoring:
 - Vendor-specific tools for each:
 - B-series, M-series, MDS-series
 - SAN Manager
 - Part of EMC ControlCenter
 - SNMP-based third-party software



© 2006 EMC Corporation. All rights reserved.SAN FoundationsMercer Road
EMC Education Services

There are several ways to monitor and manage Fibre Channel switches in a fabric:

- If the switches in the fabric are contained in a cabinet with a Service Processor (SP), console software loaded on the SP can be used to manage them.
- Some switches also offer a console port, which is used for serial connection to the switch for initial configuration using a Command Line Interface (CLI). This is typically used to set the management IP address on the switch. Subsequently, all configuration and monitoring can be done via IP. Telnet or ssh may be used to log into the switch over IP, and issue CLI commands to it. The primary purpose of the CLI is to automate management of a large number of switches/directors with the use of scripts – although the CLI may be used interactively, too. In addition, almost all models of switches support a browser-based graphical interface for management.
- There are vendor-specific tools and management suites that can be used to configure and monitor the entire fabric. They include:
 - M-Series Connectrix Manager
 - B-Series WebTools
 - MDS-Series Fabric Manager

SAN Manager, an integral part of EMC ControlCenter, provides some management and monitoring capabilities for devices from all three vendors.

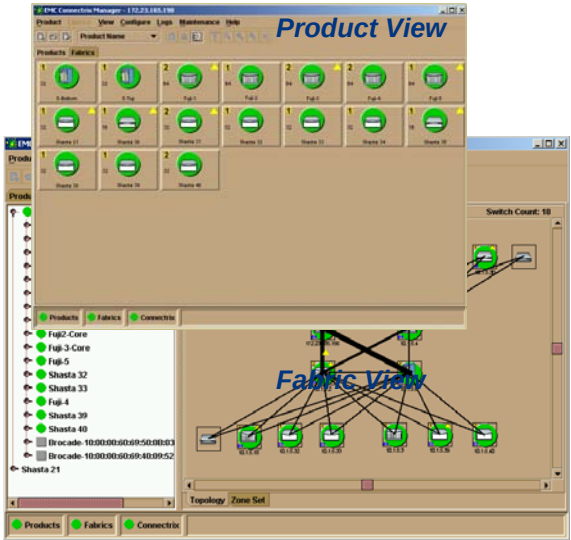
A final option is to deploy a third-party management framework such as Tivoli. Such frameworks can use SNMP (Simple Network Management Protocol) to monitor all fabric elements.

EMC Education Services

EMC²
where information lives[®]

Connectrix: Connectrix Manager (M-Series)

- Manage multiple M-Series Directors and/or Switches from a single Service Processor
- Network-wide fabric and device management
- Scalable
- Network focused tools
 - Performance
 - Availability
 - Capacity
- Topology snap shot feature
- Ability to set and identify operating speeds and hardware



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

Mercer Road
EMC Education Services

Connectrix Manager is widely used for the management of M-series (McDATA) switches. It can be run locally on the Connectrix Service Processor, or remotely on any network-attached workstation. Since this application is Java-based, IT administrators can run it from virtually any type of client device.

Connectrix Manager provides the following views:

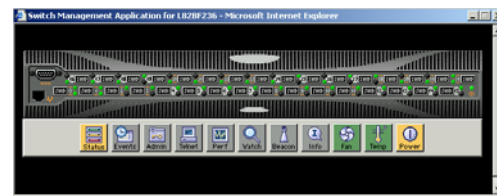
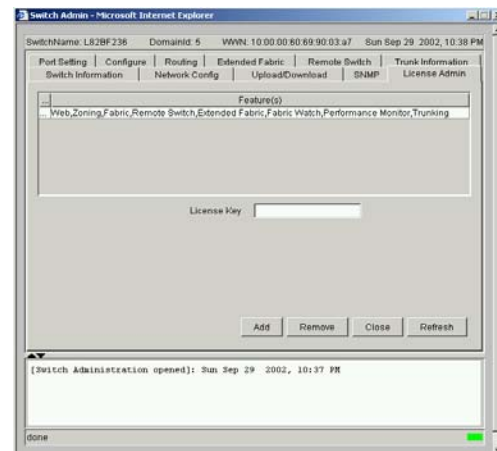
- **Product View:** An intuitive graphical view of all the devices on the network with mini-icons that display information about the device - such as the device name or IP address, number of ports, switch speed and health.
- **Fabric View:** A logical view of the fabric (known as tree control) and tabs for topology and zone sets. The elements in the tree control context menus allow single-click administration, and display a visual status of fabric health for immediate problem identification.
- **Hardware View:** Used to manage individual switches.

All M-series switches also have an Embedded Web Server (EWS). This can be used when the switch is not being managed by a Service Processor. All that EWS requires is that the switch be configured with a management IP address, and available on the network. EWS can be used to perform all functions on an M-series switch - including hardware configuration and zoning management.

Connectrix: Web Tools (B-Series)

- Browser-based management application for B-Series switches and directors
- Provides zoning, fabric, and switch management
 - Supports aliases
 - Provides fabric-wide and detailed views
 - Firmware upgrades
- Accessible through Ethernet using any desktop browser, such as Internet Explorer

Administration View



Switch View

SAN Foundations

Mercer Road
EMC Education Services

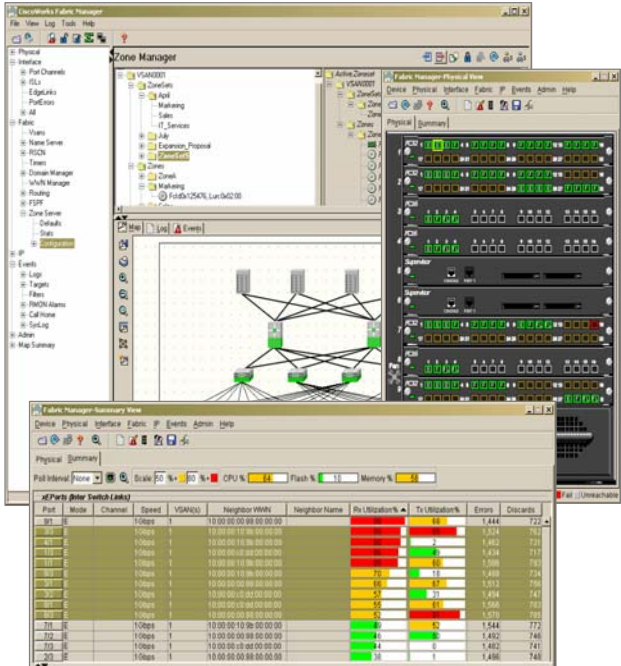
© 2006 EMC Corporation. All rights reserved.

WebTools is an easy-to-use, browser-based application for switch management and is included with all Connectrix B-Series products. WebTools simplifies switch management by enabling administrators to configure, monitor, and manage switch and fabric parameters from a single online access point. WebTools supports the use of aliases for easy identification of zone members. With WebTools, firmware upgrade is a one-step process. The Switch View allows you to check the status of a switch in the fabric. The LED icon for the port reporting an issue will change color.

EMC Education Services
EMC²
where information lives[®]

Connectrix: Fabric Manager (MDS-Series)

- Switch-embedded
- Java-based application
- Switch configuration
- Discovery
- Topology mapping
- Monitoring
- Alerts
- Network diagnostics
- Security (SNMPv3, SSH, RBAC)
- Fabric, Summary and Physical Views



© 2006 EMC Corporation. All rights reserved.
SAN Foundations
Mercer Road
EMC Education Services

MDS Fabric Manager and device manager are included with all MDS Directors and switches. This Java-based tool simplifies management of the MDS Series through an integrated approach to fabric administration, device discovery, topology mapping, and configuration functions for the switch, fabric, and port.

Features of MDS Fabric Manager include:

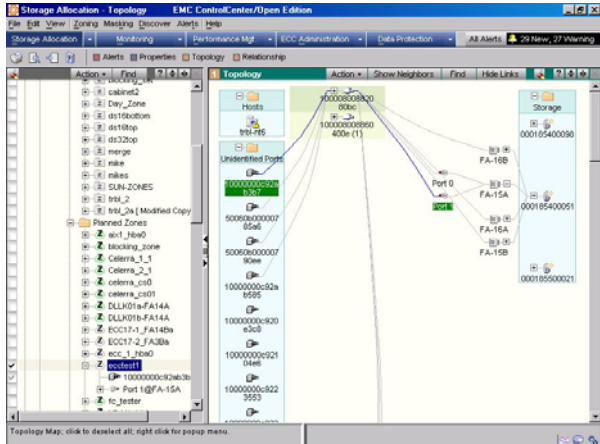
- Fabric visualization: Automatic discovery, zone and path highlighting
- Comprehensive configuration across multiple switches
- Powerful configuration analysis including real-time monitoring, alerts, zone merge analysis, and configuration checking
- Network diagnostics: Probes network and switch health, enabling administrators to pinpoint connectivity and performance issues
- Comprehensive security: Protection against unauthorized management access with Simple Network Management Protocol Version 3 (SNMPV3), Secure Shell Protocol (SSH), and role-based access control (RBAC)
- Traffic Management: Congestion control mechanism (FCC) can throttle back traffic at its origin
- Quality of Service allows traffic to be intelligently managed; Low-priority traffic is throttled at source; High-priority traffic is not affected

EMC Education Services

EMC²
where information lives[®]

Connectrix: SAN Manager (EMC ControlCenter)

- Integrated in ControlCenter
- Single interface
 - Switch zoning Brocade and McDATA
 - Device Masking Symmetrix, CLARiiON
 - View Cisco switches
- Discovers heterogeneous SAN elements
 - Servers
 - SAN devices
 - Storage



© 2006 EMC Corporation. All rights reserved.

SAN Foundations

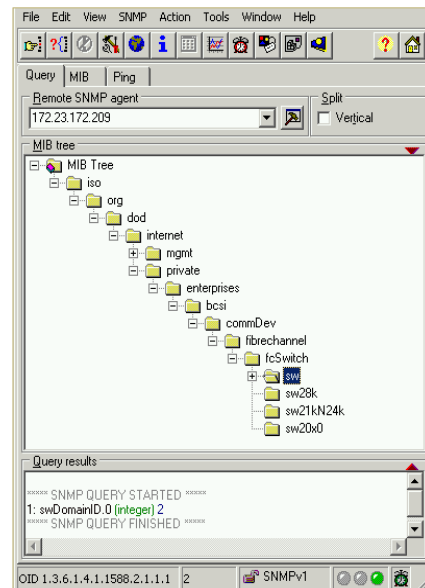
Mercer Road
EMC Education Services

SAN Manager provides a single interface to manage LUN Masking, switch zoning, device monitoring and management. The integration of SAN Manager into ControlCenter provides a distributed infrastructure allowing for remote management of a SAN. It offers reporting and monitoring features such as threshold alarms, state change alerts and component failure notifications for devices in the SAN.

SAN Manager has capabilities to automatically discover, map and display the entire SAN topology at a level of detail desired by the administrator. It can also display specific physical and logical information about each object in the fabric. Administrators can view details on physical components such as host bus adapters, Fibre Channel switches and storage arrays as well as logical components such as zones and LUN masking policies. SAN Manager offers support for non-EMC arrays such as HDS Lightning, HP StorageWorks and IBM Shark.

Connectrix: SNMP Management

- All Connectrix devices support SNMP
- Allows third-party management tools to manage Connectrix devices
- Management Information Base (MIB) support
 - FibreAlliance
 - Fabric Element (FE)
 - Switch (SW)



SW-MIB

Support for SNMP (Simple Network Management Protocol) is available for all members of the Connectrix family. SNMP is an industry standard for managing networks, and is used mostly for monitoring the status of the network to identify problems. SNMP is also used to gather performance and poll real-time usage from fabric elements.

Each vendor product has a specific SNMP MIB (Management Information Base) associated with it. The FibreAlliance MIB is an actively evolving standard MIB specifically designed with multi-vendor fabrics in mind. A MIB is just a numerical representation of the status information that is accessed via SNMP from a management station.

Examples of SNMP based Software:

- IBM Tivoli
- HP OpenView
- CA UniCenter



SAN Foundations

SAN: Technical Positioning

© 2006 EMC Corporation. All rights reserved.

SAN Foundations



This section highlights the key benefits that are attainable by using SAN technology.

When Should Storage Area Networks Be Used

- SANs are optimized for high bandwidth block level I/O
- Suited for the demands of real time applications with stringent requirements on I/O latency and throughput, such as:
 - Databases: OLTP (online transaction processing)
 - Video streaming
 - Any applications with high transaction rate and high data volatility
- Used to consolidate heterogeneous storage environments
 - Gain efficiencies in the management of storage resources including capacity, performance and connectivity
 - Physical consolidation
 - Logical consolidation
- For highly available host-to-storage connectivity, where multipathing and/or host-based clustering are mandatory

Storage Area Networks can handle large amounts of block level I/O and are suited to meet the demands of high performance applications that need access to data in real time.

In several environments, these applications have to share access to storage resources and implementing them in a SAN allows efficient use of these resources. When data volatility is high, a host's needs for capacity and performance can grow or shrink significantly over time. The SAN architecture is flexible, so existing storage can be rapidly redeployed across hosts - as needs change - with minimal disruption.

SANs are also used to consolidate storage within an enterprise. Consolidation can be at a physical or logical level.

Physical consolidation involves the physical relocation of resources to a centralized location. Once these resources are consolidated, one can make more efficient use of facility resources such as HVAC (heating, ventilation and air conditioning), power protection, personnel, and physical security. Physical consolidations have a drawback in that they do not offer resilience against a site failure.

Logical consolidation is the process of bringing components under a unified management infrastructure and creating a shared resource pool. Since SANs can be extended to span vast distances physically, they do not strictly require that logically related entities be physically close to each other. Logical consolidation does not allow one to take full advantage of the benefits of site consolidation. But it does offer some amount of protection against site failure, especially if well planned.

Deploying a New SAN

- More choices to consider than in the past
 - Fibre Channel SANs
 - iSCSI SANs
 - Bridged SANs, with mixed iSCSI and Fibre Channel hosts and storage arrays
 - Bridging mandates the use of a multiprotocol router; cost of this must be factored in
 - the router can also serve a second purpose: extend Fibre Channel SANs over long distances
 - ❖ may be a critical consideration if disaster recovery across sites is a factor

With the advent of IP SANs and IP-based SAN extensions, SAN design presents considerably more choices than in the past.

The feature-set of chosen storage array models, host system qualification, application performance requirements, availability requirements, disaster recovery requirements, and cost are some of the key factors that influence the design process.

Upgrading an Existing SAN

- Specific motivation for the upgrade should help make suitable product choices:
 - scaling up connectivity – to add more hosts and/or storage
 - improving performance and/or availability – resolve known bandwidth issues with existing ISLs
 - choices available will depend on existing topology
 - phase in or out a switch product vendor
 - heterogeneous, multi-vendor fabric can be an interim solution
 - extend existing Fibre Channel SAN over long distance
 - design choices: traditional long-wave over single-mode, DWDM, IP-based extensions
 - provide connectivity to existing FC storage from new hosts
 - iSCSI hosts with a router may be a viable option
 - consolidate disparate SAN islands within existing enterprise
 - interoperability of gear within the various islands must be analyzed

This is a short list of the likely reasons why customers may wish to upgrade their SANs, possibly using new Connectrix gear. As with a new SAN installation, upgrade options have become more numerous as well.

Block Storage over IP: Value Proposition

- **Leverage IP Network Investment for Storage Connectivity**
 - Reduced acquisition costs
 - Single NIC for network and storage I/O
 - Leverage high-speed backbone (1GB+)
 - Economies of scale (cables, connectors, installation, etc.)
 - Reduced operation and management costs
 - Common management paradigm with LAN and internet
 - Cable plant installation and management
 - Vendor independence
 - Total interoperability
 - ❖ Network interface cards, storage arrays, and devices
 - ★ Network infrastructure (switches, routers, etc.)
 - Infrastructure management software/framework

The promise of these standards/protocols/approaches is really about reducing TCO (total cost of ownership) for storage networking, particularly when positioned/compared to Fibre Channel-based networks.

In this case, the TCO value proposition can be broken into three areas:

- Reduced acquisition costs
 - Cheaper host connections (don't need a separate/dedicate card for Fibre Channel)
 - Use existing high-speed "backbone" network (don't need a separate net for storage)
 - Cheaper deployment (uses same network components as rest of Ethernet infrastructure)
- Reduced management/operation costs
 - Don't have to learn a totally new network paradigm
 - Don't have to manage a separate network
- Vendor independence
 - Interoperability has been a major challenge for Fibre Channel, both within the network, and at the end-points
 - While Fibre Channel network hardware often supports SNMP, full element management still requires a completely separate management console. The storage over IP promise is that this will be combined into a single management environment.

SANs and ILM

- **SANs add value to the Information LifeCycle Management (ILM) strategy of a company**
 - SAN-based storage arrays can hold data during the “high access rate, high-performance” stage of its lifecycle
 - Data migration across storage arrays of differing classes is easy
 - Hosts and all participating storage frames can reside on the same SAN infrastructure
 - Inherent access control features of a SAN allow for shared storage across hosts, without compromising security
 - Data migration across storage frames can be driven either by a host-based application, or using array-centric replication features
 - NAS gateway products can share SAN storage with hosts
 - An ILM strategy involving SAN-to-NAS data migration is feasible

Implementation of an ILM strategy mandates convenient migration of data, as it progresses through its lifecycle, through different tiers of storage. Each storage tier has distinct price-versus-performance characteristics. In general, the highest tiers are the most expensive per Gbyte of capacity - but best suited for high transaction rates.

Typically, data needs to be available in a high transaction rate environment during the early stages of its existence – thus it would need to reside on relatively high-cost, high-end storage arrays. As data ages, it can move to lower tiers of storage successively – with less stringent I/O performance requirements as time progresses. A carefully designed and implemented ILM strategy can therefore result in efficient and cost-effective use of available storage resources.

SANs add key value to the ILM proposition.

Simple, scalable and secure connectivity makes it possible to have multiple tiers of block-oriented storage – e.g. a mix of Symmetrix and CLARiiON arrays - on the same SAN. These arrays can be made selectively accessible by multiple hosts. Data migration between the storage arrays is facilitated by the ease of connectivity. Migration can be achieved using either host-based applications, or array-to-array replication features.

Second, it is possible to apportion storage within a SAN to multiple hosts, as well as to NAS gateways – such as the Celera gateway products. This facilitates the use of NAS as an additional storage tier within the ILM design, whenever appropriate.

Module Summary

Key points covered in this module:

- Overview of Fibre Channel and IP SANs
 - Definitions, features and benefits
 - Overview of the underlying protocols used within a SAN
- Issues to consider when designing a SAN
 - Distinct characteristics of commonly deployed fabric topologies
 - Basic operational details of Inter-Switch Links
 - performance and security related features
- Major product categories within the EMC Connectrix family
 - Features and benefits of the EMC Connectrix family
 - Software options for managing Fabric components
 - Connectrix component types to be used, when designing a SAN

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