

# Chapter 6

## Storage Area Networks

Organizations are experiencing an explosive growth in information. This information needs to be stored, protected, optimized, and managed efficiently. Data center managers are burdened with the challenging task of providing low-cost, high-performance information management solutions. An effective information management solution must provide the following:

- **Just-in-time information to business users:** Information must be available to business users when they need it. The explosive growth in online storage, proliferation of new servers and applications, spread of mission-critical data throughout enterprises, and demand for 24 × 7 data availability are some of the challenges that need to be addressed.
- **Integration of information infrastructure with business processes:** The storage infrastructure should be integrated with various business processes without compromising its security and integrity.
- **Flexible and resilient storage architecture:** The storage infrastructure must provide flexibility and resilience that aligns with changing business requirements. Storage should scale without compromising performance requirements of the applications and, at the same time, the total cost of managing information must be low.

### KEY CONCEPTS

Storage Consolidation

Fibre Channel (FC) Architecture

Fibre Channel Protocol Stack

Fibre Channel Ports

Fibre Channel Addressing

World Wide Names

Zoning

Fibre Channel Topologies

Direct-attached storage (DAS) is often referred to as a stovepiped storage environment. Hosts “own” the storage and it is difficult to manage and share resources on these isolated storage devices. Efforts to organize this dispersed data led to the emergence of the storage area network (SAN). SAN is a high-speed, dedicated network of servers and shared storage devices. Traditionally connected over Fibre Channel (FC) networks, a SAN forms a single-storage pool and facilitates data centralization and consolidation. SAN meets the storage demands efficiently with better economies of scale. A SAN also provides effective maintenance and protection of data.

This chapter provides detailed insight into the FC technology on which a SAN is deployed and also reviews SAN design and management fundamentals.

## **6.1 Fibre Channel: Overview**

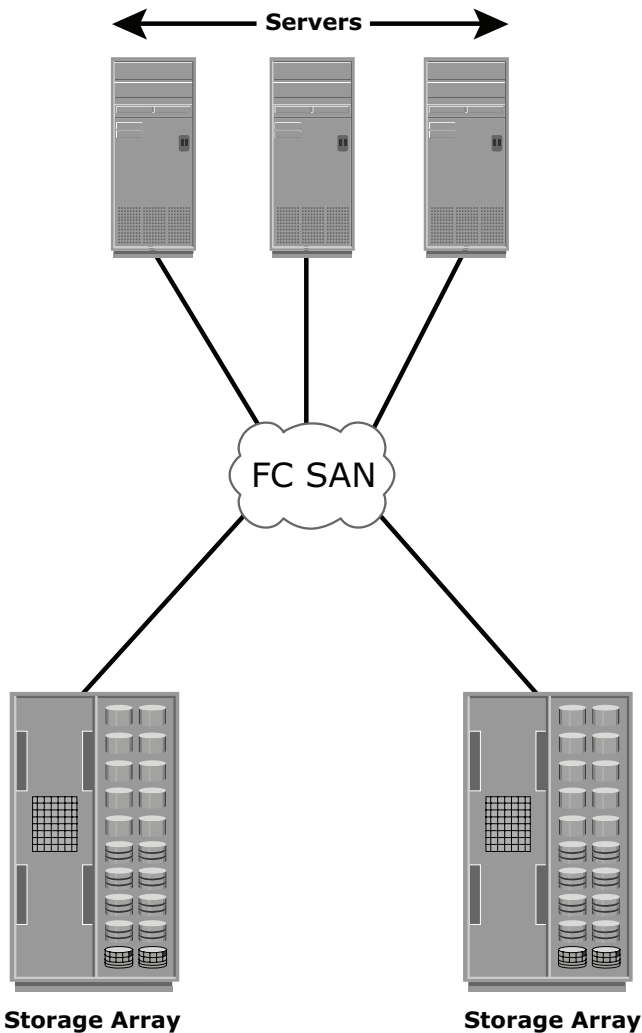
---

The FC architecture forms the fundamental construct of the SAN infrastructure. *Fibre Channel* is a high-speed network technology that runs on high-speed optical fiber cables (preferred for front-end SAN connectivity) and serial copper cables (preferred for back-end disk connectivity). The FC technology was created to meet the demand for increased speeds of data transfer among computers, servers, and mass storage subsystems. Although FC networking was introduced in 1988, the FC standardization process began when the American National Standards Institute (ANSI) chartered the Fibre Channel Working Group (FCWG). By 1994, the new high-speed computer interconnection standard was developed and the Fibre Channel Association (FCA) was founded with 70 charter member companies. Technical Committee T11, which is the committee within INCITS (International Committee for Information Technology Standards), is responsible for Fibre Channel interfaces. T11 (previously known as X3T9.3) has been producing interface standards for high performance and mass storage applications since the 1970s.

Higher data transmission speeds are an important feature of the FC networking technology. The initial implementation offered throughput of 100 MB/s (equivalent to raw bit rate of 1Gb/s i.e. 1062.5 Mb/s in Fibre Channel), which was greater than the speeds of Ultra SCSI (20 MB/s) commonly used in DAS environments. FC in full-duplex mode could sustain throughput of 200 MB/s. In comparison with Ultra-SCSI, FC is a significant leap in storage networking technology. Latest FC implementations of 8 GFC (Fibre Channel) offers throughput of 1600 MB/s (raw bit rates of 8.5 Gb/s), whereas Ultra320 SCSI is available with a throughput of 320 MB/s. The FC architecture is highly scalable and theoretically a single FC network can accommodate approximately 15 million nodes.

## 6.2 The SAN and Its Evolution

A *storage area network (SAN)* carries data between servers (also known as *hosts*) and storage devices through fibre channel switches (see Figure 6-1). A SAN enables storage consolidation and allows storage to be shared across multiple servers. It enables organizations to connect geographically dispersed servers and storage.



**Figure 6-1:** SAN implementation

A SAN provides the physical communication infrastructure and enables secure and robust communication between host and storage devices. The SAN management interface organizes connections and manages storage elements and hosts.

In its earliest implementation, the SAN was a simple grouping of hosts and the associated storage that was connected to a network using a hub as a connectivity device. This configuration of a SAN is known as a *Fibre Channel Arbitrated Loop (FC-AL)*, which is detailed later in the chapter. Use of hubs resulted in isolated FC-AL SAN islands because hubs provide limited connectivity and bandwidth.

The inherent limitations associated with hubs gave way to high-performance FC *switches*. The switched fabric topologies improved connectivity and performance, which enabled SANs to be highly scalable. This enhanced data accessibility to applications across the enterprise. FC-AL has been abandoned for SANs due to its limitations, but still survives as a disk-drive interface. Figure 6-2 illustrates the FC SAN evolution from FC-AL to enterprise SANs.

Today, Internet Protocol (IP) has become an option to interconnect geographically separated SANs. Two popular protocols that extend block-level access to applications over IP are iSCSI and Fibre Channel over IP (FCIP). These protocols are detailed in Chapter 8.

## **6.3 Components of SAN**

---

A SAN consists of three basic components: servers, network infrastructure, and storage. These components can be further broken down into the following key elements: node ports, cabling, interconnecting devices (such as FC switches or hubs), storage arrays, and SAN management software.

### **6.3.1 Node Ports**

In fibre channel, devices such as hosts, storage and tape libraries are all referred to as *nodes*. Each node is a source or destination of information for one or more nodes. Each node requires one or more ports to provide a physical interface for communicating with other nodes. These ports are integral components of an HBA and the storage front-end adapters. A port operates in full-duplex data transmission mode with a *transmit (Tx)* link and a *receive (Rx)* link (see Figure 6-3).

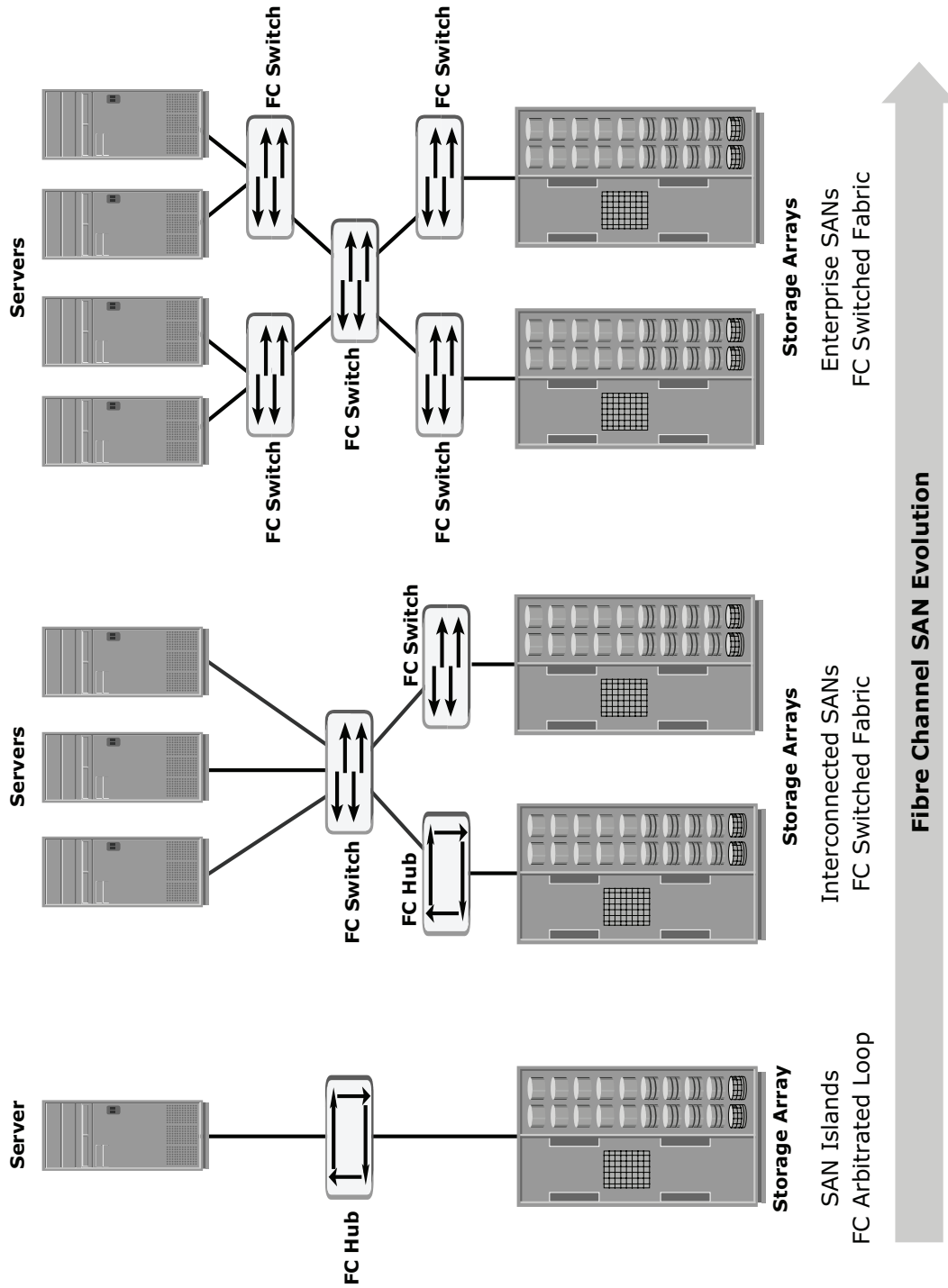
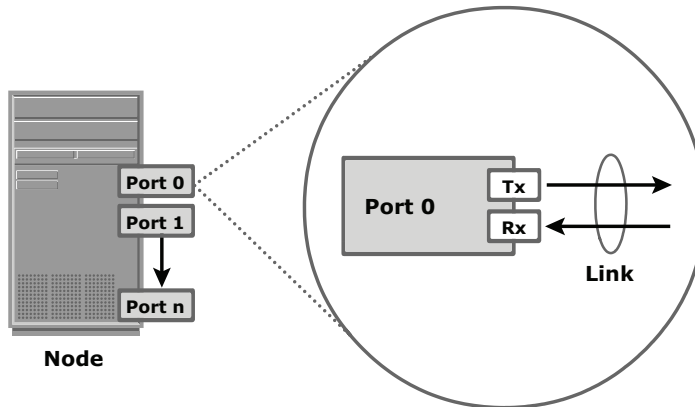


Figure 6-2: FC SAN evolution



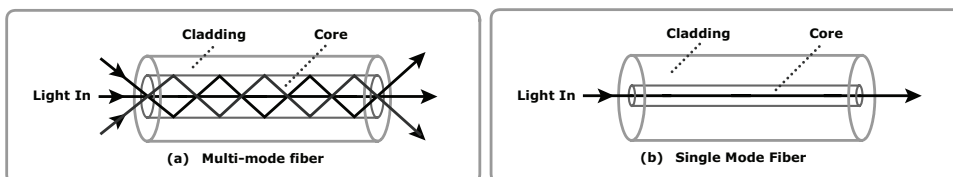
**Figure 6-3:** Nodes, ports, and links

### 6.3.2 Cabling

SAN implementations use optical fiber cabling. Copper can be used for shorter distances for back-end connectivity, as it provides a better signal-to-noise ratio for distances up to 30 meters. Optical fiber cables carry data in the form of light. There are two types of optical cables, multi-mode and single-mode.

Multi-mode fiber (MMF) cable carries multiple beams of light projected at different angles simultaneously onto the core of the cable (see Figure 6-4 (a)). Based on the bandwidth, multi-mode fibers are classified as OM1 (62.5 $\mu$ m), OM2 (50 $\mu$ m) and laser optimized OM3 (50 $\mu$ m). In an MMF transmission, multiple light beams traveling inside the cable tend to disperse and collide. This collision weakens the signal strength after it travels a certain distance — a process known as *modal dispersion*. An MMF cable is usually used for distances of up to 500 meters because of signal degradation (attenuation) due to modal dispersion.

Single-mode fiber (SMF) carries a single ray of light projected at the center of the core (see Figure 6-4 (b)). These cables are available in diameters of 7–11 microns; the most common size is 9 microns. In an SMF transmission, a single light beam travels in a straight line through the core of the fiber. The small core and the single light wave limits modal dispersion. Among all types of fibre cables, single-mode provides minimum signal attenuation over maximum distance (up to 10 km). A single-mode cable is used for long-distance cable runs, limited only by the power of the laser at the transmitter and sensitivity of the receiver.

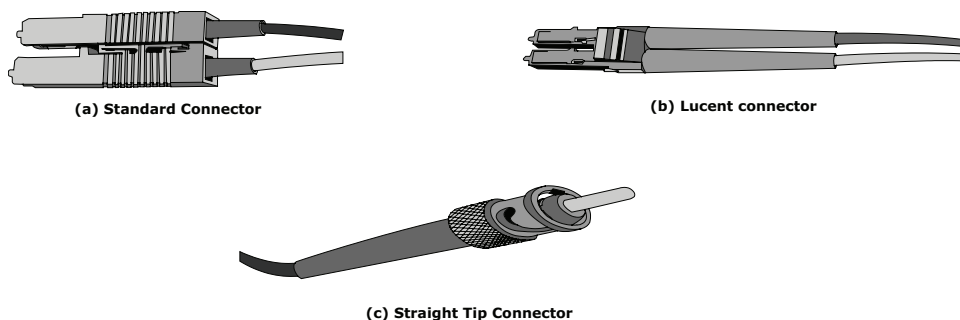


**Figure 6-4:** Multi-mode fiber and single-mode fiber

MMFs are generally used within data centers for shorter distance runs, while SMFs are used for longer distances. MMF transceivers are less expensive as compared to SMF transceivers.

A Standard connector (SC) (see Figure 6-5 (a)) and a Lucent connector (LC) (see Figure 6-5 (b)) are two commonly used connectors for fiber optic cables. An SC is used for data transmission speeds up to 1 Gb/s, whereas an LC is used for speeds up to 4 Gb/s. Figure 6-6 depicts a Lucent connector and a Standard connector.

A *Straight Tip (ST)* is a fiber optic connector with a plug and a socket that is locked with a half-twisted bayonet lock (see Figure 6-5 (c)). In the early days of FC deployment, fiber optic cabling predominantly used ST connectors. This connector is often used with Fibre Channel patch panels.



**Figure 6-5:** SC, LC, and ST connectors

The Small Form-factor Pluggable (SFP) is an optical transceiver used in optical communication. The standard SFP+ transceivers support data rates up to 10 Gb/s.

### 6.3.3 Interconnect Devices

Hubs, switches, and directors are the interconnect devices commonly used in SAN.

*Hubs* are used as communication devices in FC-AL implementations. Hubs physically connect nodes in a logical loop or a physical star topology. All the nodes must share the bandwidth because data travels through all the connection points. Because of availability of low cost and high performance switches, hubs are no longer used in SANs.

*Switches* are more intelligent than hubs and directly route data from one physical port to another. Therefore, nodes do not share the bandwidth. Instead, each node has a dedicated communication path, resulting in bandwidth aggregation.

*Directors* are larger than switches and are deployed for data center implementations. The function of directors is similar to that of FC switches, but directors have higher port count and fault tolerance capabilities.

### 6.3.4 Storage Arrays

The fundamental purpose of a SAN is to provide host access to storage resources. The capabilities of intelligent storage arrays are detailed in Chapter 4. The large storage capacities offered by modern storage arrays have been exploited in SAN environments for storage consolidation and centralization. SAN implementations complement the standard features of storage arrays by providing high availability and redundancy, improved performance, business continuity, and multiple host connectivity.

### 6.3.5 SAN Management Software

SAN management software manages the interfaces between hosts, interconnect devices, and storage arrays. The software provides a view of the SAN environment and enables management of various resources from one central console.

It provides key management functions, including mapping of storage devices, switches, and servers, monitoring and generating alerts for discovered devices, and logical partitioning of the SAN, called *zoning*. In addition, the software provides management of typical SAN components such as HBAs, storage components, and interconnecting devices.

#### FC SWITCH VERSUS FC HUB



**Scalability and performance are the primary differences between switches and hubs. A switched fabric uses 24-bit addressing, which supports over 15 million devices, whereas the FC-AL implemented in hubs supports only a maximum of 126 nodes.**

**Fabric switches provide full bandwidth between multiple pairs of ports in a fabric, resulting in a scalable architecture that can simultaneously support multiple communications.**

**Hubs provide shared bandwidth, and can support only single communication. They provide a low-cost connectivity expansion solution. Switches, conversely, can be used to build dynamic, high-performance fabrics through which multiple communications can take place simultaneously. Switches are more expensive than hubs.**

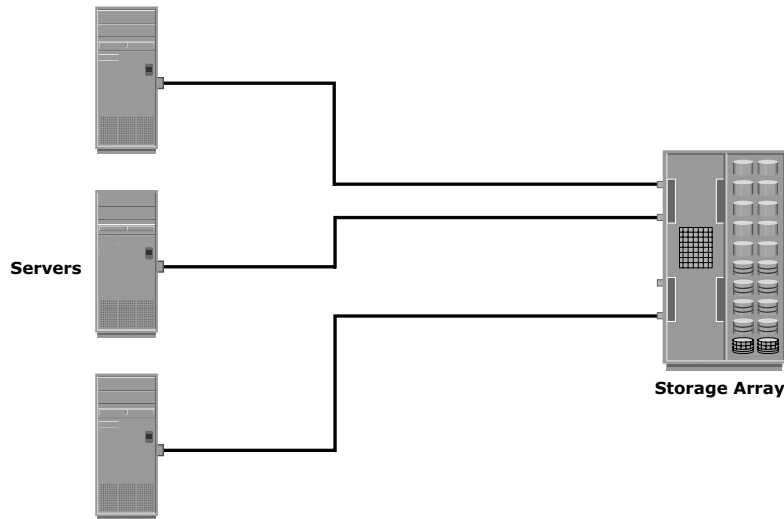


## 6.4 FC Connectivity

The FC architecture supports three basic interconnectivity options: point-to-point, arbitrated loop (FC-AL), and fabric connect.

### 6.4.1 Point-to-Point

*Point-to-point* is the simplest FC configuration — two devices are connected directly to each other, as shown in Figure 6-6. This configuration provides a dedicated connection for data transmission between nodes. However, the point-to-point configuration offers limited connectivity, as only two devices can communicate with each other at a given time. Moreover, it cannot be scaled to accommodate a large number of network devices. Standard DAS uses point-to-point connectivity.



**Figure 6-6:** Point-to-point topology

### 6.4.2 Fibre Channel Arbitrated Loop

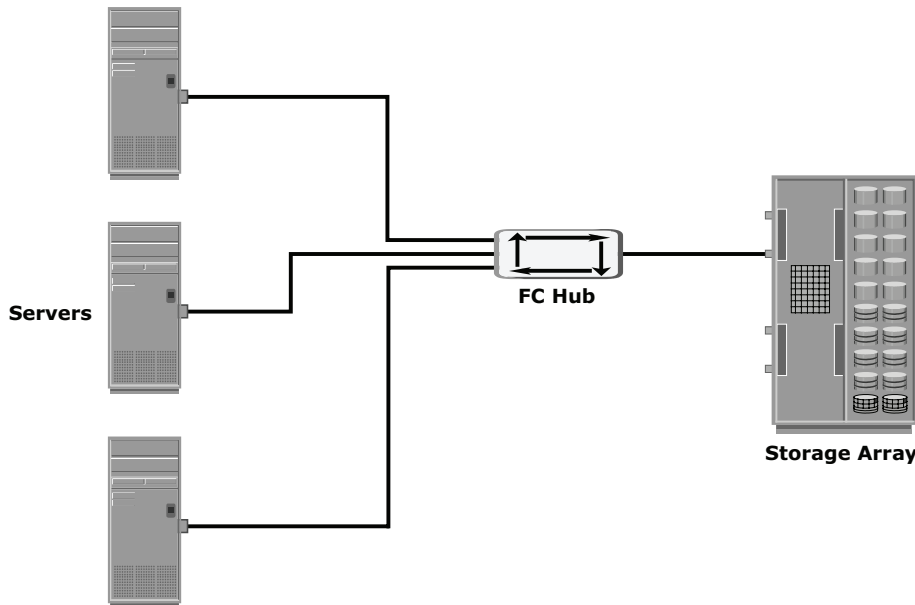
In the FC-AL configuration, devices are attached to a shared loop, as shown in Figure 6-7. FC-AL has the characteristics of a token ring topology and a physical star topology. In FC-AL, each device contends with other devices to perform I/O operations. Devices on the loop must “arbitrate” to gain control of the loop. At any given time, only one device can perform I/O operations on the loop.

As a loop configuration, FC-AL can be implemented without any interconnecting devices by directly connecting one device to another in a ring through cables.

However, FC-AL implementations may also use hubs whereby the arbitrated loop is physically connected in a star topology.

The FC-AL configuration has the following limitations in terms of scalability:

- FC-AL shares the bandwidth in the loop. Only one device can perform I/O operations at a time. Because each device in a loop has to wait for its turn to process an I/O request, the speed of data transmission is low in an FC-AL topology.



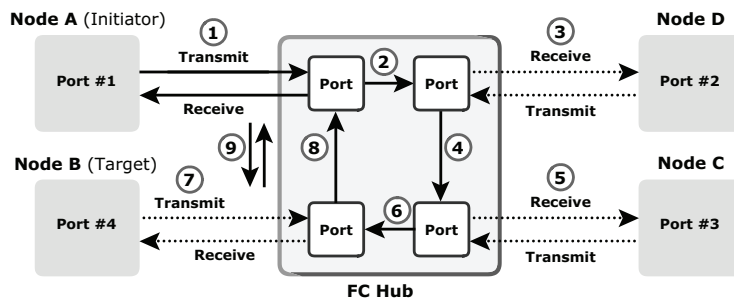
**Figure 6-7:** Fibre Channel arbitrated loop

- FC-AL uses 8-bit addressing. It can support up to 127 devices on a loop.
- Adding or removing a device results in loop re-initialization, which can cause a momentary pause in loop traffic.

### ***FC-AL Transmission***

When a node in the FC-AL topology attempts to transmit data, the node sends an *arbitration (ARB)* frame to each node on the loop. If two nodes simultaneously attempt to gain control of the loop, the node with the highest priority is allowed to communicate with another node. This priority is determined on the basis of Arbitrated Loop Physical Address (AL-PA) and Loop ID, described later in this chapter.

When the initiator node receives the ARB request it sent, it gains control of the loop. The initiator then transmits data to the node with which it has established a virtual connection. Figure 6-8 illustrates the process of data transmission in an FC-AL configuration.



#### Node A want to communicate with Node B

- ① High priority initiator, Node A inserts the ARB frame in the loop.
- ② ARB frame is passed to the next node (Node D) in the loop.
- ③ Node D receives high priority ARB, therefore remains idle.
- ④ ARB is forwarded to next node (Node C) in the loop.
- ⑤ Node C receives high priority ARB, therefore remains idle.
- ⑥ ARB is forwarded to next node (Node B) in the loop.
- ⑦ Node B receives high priority ARB, therefore remains idle and
- ⑧ ARB is forwarded to next node (Node A) in the loop.
- ⑨ Node A receives ARB back; now it gains control of the loop and can start communicating with target Node B.

**Figure 6-8:** Data transmission in FC-AL

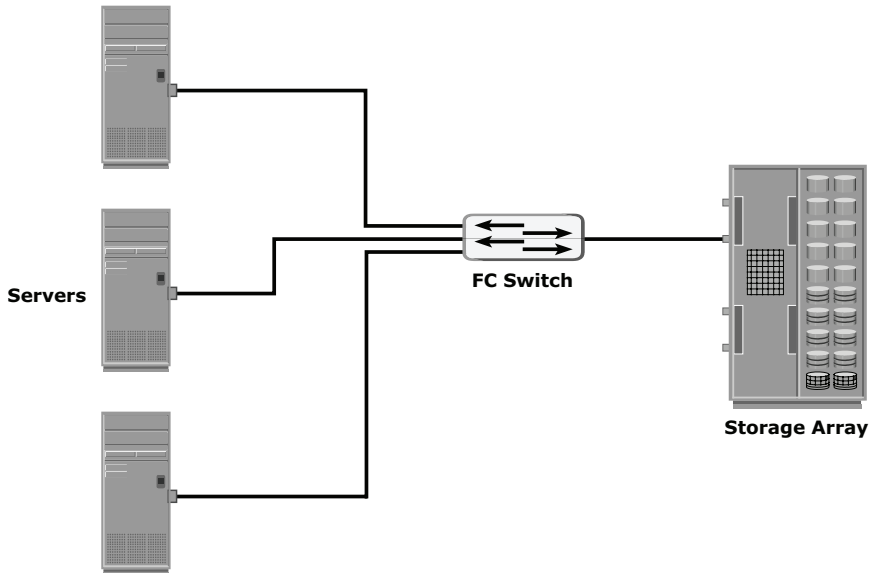
### 6.4.3 Fibre Channel Switched Fabric

Unlike a loop configuration, a Fibre Channel switched fabric (FC-SW) network provides interconnected devices, dedicated bandwidth, and scalability. The addition or removal of a device in a switched fabric is minimally disruptive; it does not affect the ongoing traffic between other devices.

FC-SW is also referred to as *fabric connect*. A fabric is a logical space in which all nodes communicate with one another in a network. This virtual space can be created with a switch or a network of switches. Each switch in a fabric contains a unique domain identifier, which is part of the fabric's addressing scheme. In FC-SW, nodes do not share a loop; instead, data is transferred through a dedicated path between the nodes. Each port in a fabric has a unique 24-bit fibre channel address for communication. Figure 6-9 shows an example of FC-SW.

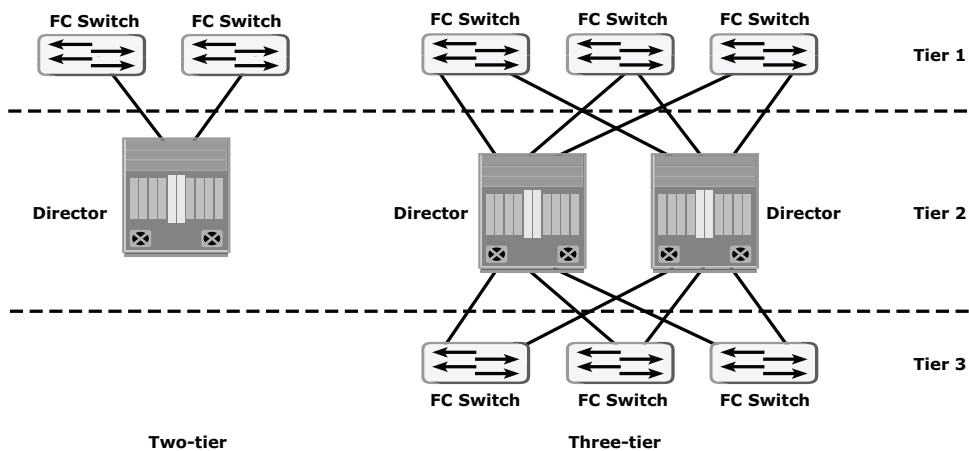
A fabric topology can be described by the number of tiers it contains. The number of tiers in a fabric is based on the number of switches traversed between two points that are farthest from each other. However, note that this number

is based on the infrastructure constructed by the fabric topology; it disregards how the storage and server are connected across the switches.



**Figure 6-9:** Fibre Channel switched fabric

When the number of tiers in a fabric increases, the distance that a fabric management message must travel to reach each switch in the fabric also increases. The increase in the distance also increases the time taken to propagate and complete a fabric reconfiguration event, such as the addition of a new switch, or a zone set propagation event (detailed later in this chapter). Figure 6-10 illustrates two-tier and three-tier fabric architecture.

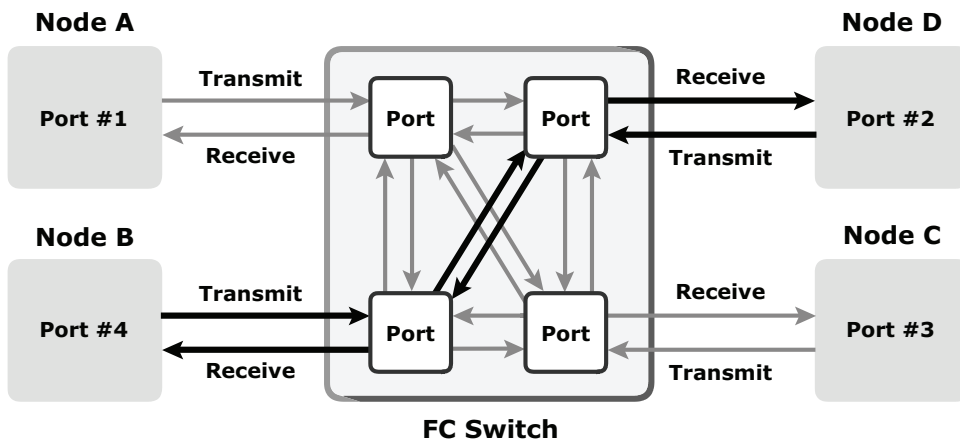


**Figure 6-10:** Tiered structure of FC-SW topology

## FC-SW Transmission

FC-SW uses switches that are intelligent devices. They can switch data traffic from an initiator node to a target node directly through switch ports. Frames are routed between source and destination by the fabric.

As shown in Figure 6-11, if node B wants to communicate with node D, Nodes should individually login first and then transmit data via the FC-SW. This link is considered a dedicated connection between the initiator and the target.



**Figure 6-11:** Data transmission in FC-SW topology

## 6.5 Fibre Channel Ports

Ports are the basic building blocks of an FC network. Ports on the switch can be one of the following types:

- **N\_port:** An end point in the fabric. This port is also known as the *node port*. Typically, it is a host port (HBA) or a storage array port that is connected to a switch in a switched fabric.
- **NL\_port:** A node port that supports the arbitrated loop topology. This port is also known as the *node loop port*.
- **E\_port:** An FC port that forms the connection between two FC switches. This port is also known as the *expansion port*. The E\_port on an FC switch connects to the E\_port of another FC switch in the fabric through a link, which is called an *Inter-Switch Link (ISL)*. ISLs are used to transfer host-to-storage data as well as the fabric management traffic from one

switch to another. ISL is also one of the scaling mechanisms in SAN connectivity.

- **F\_port:** A port on a switch that connects an N\_port. It is also known as a *fabric port* and cannot participate in FC-AL.
- **FL\_port:** A fabric port that participates in FC-AL. This port is connected to the NL\_ports on an FC-AL loop. A FL\_port also connects a loop to a switch in a switched fabric. As a result, all NL\_ports in the loop can participate in FC-SW. This configuration is referred to as a *public loop*. In contrast, an arbitrated loop without any switches is referred to as a *private loop*. A private loop contains nodes with NL\_ports, and does not contain FL\_port.
- **G\_port:** A generic port that can operate as an E\_port or an F\_port and determines its functionality automatically during initialization.

Figure 6-12 shows various FC ports located in the fabric.

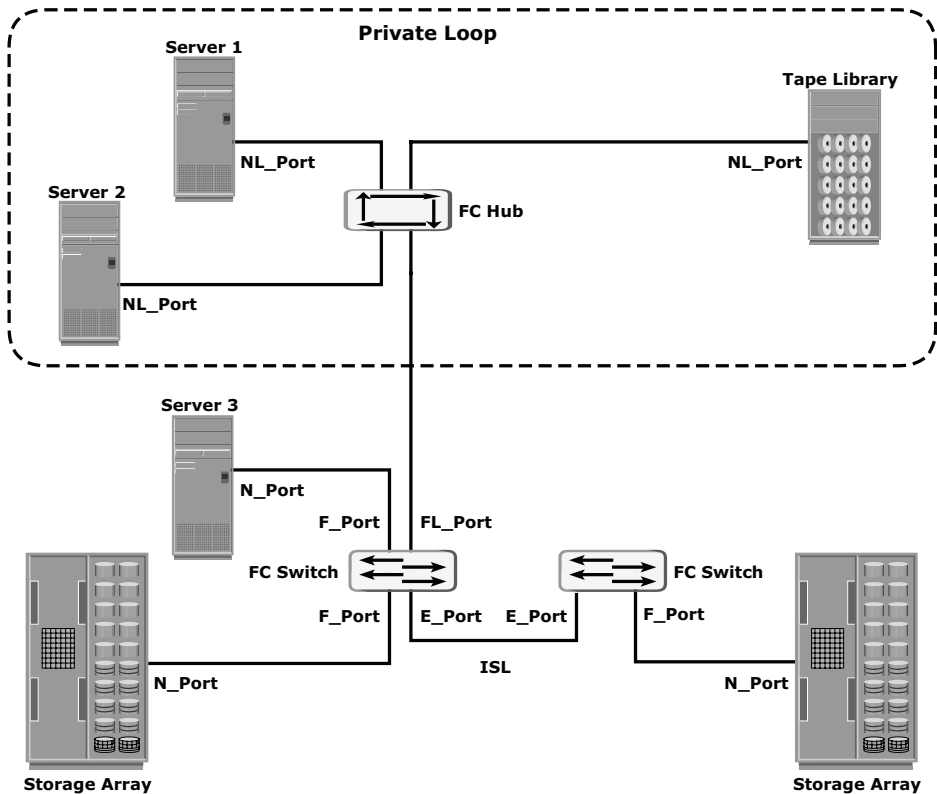


Figure 6-12: Fibre channel ports

## 6.6 Fibre Channel Architecture

The FC architecture represents true channel/network integration with standard interconnecting devices. Connections in a SAN are accomplished using FC. Traditionally, transmissions from host to storage devices are carried out over channel connections such as a parallel bus. Channel technologies provide high levels of performance with low protocol overheads. Such performance is due to the static nature of channels and the high level of hardware and software integration provided by the channel technologies. However, these technologies suffer from inherent limitations in terms of the number of devices that can be connected and the distance between these devices.

*Fibre Channel Protocol (FCP)* is the implementation of serial SCSI-3 over an FC network. In the FCP architecture, all external and remote storage devices attached to the SAN appear as local devices to the host operating system. The key advantages of FCP are as follows:

- Sustained transmission bandwidth over long distances.
- Support for a larger number of addressable devices over a network. Theoretically, FC can support over 15 million device addresses on a network.
- Exhibits the characteristics of channel transport and provides speeds up to 8.5 Gb/s (8 GFC).

### FICON



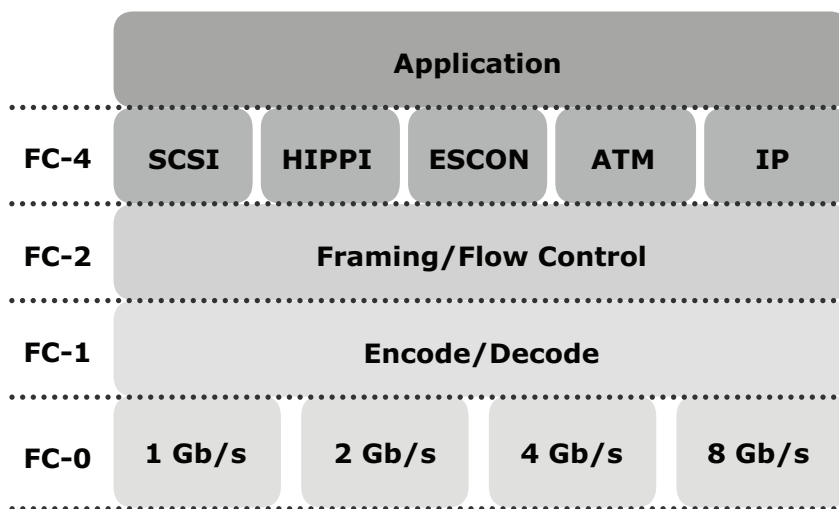
**Mainframe SANs use *FICON (Fibre Connectivity)* for a low-latency, high-bandwidth connection to the storage controller. FICON is an FC-4 type technology, and its place in the FC architecture is analogous to FCP. FICON was designed as a replacement for *ESCON (Enterprise System Connection)* to support mainframe attached storage systems.**

FCP is specified by standards produced by T10; FCP-3 is the last issued standard, and FCP-4 is under development. FCP defines a Fibre Channel mapping layer (FC-4) that uses the services defined by ANS X3.230-199X, Fibre Channel—Physical and Signaling Interface (FC-PH) to transmit SCSI commands, data, and status information between SCSI initiator and SCSI target. FCP defines Fibre Channel information units in accordance with the SCSI architecture model. FCP also defines how the Fibre Channel services are used to perform the services defined by the SCSI architecture model.

The FC standard enables mapping several existing *Upper Layer Protocols (ULPs)* to FC frames for transmission, including SCSI, IP, High Performance Parallel Interface (HIPPI), Enterprise System Connection (ESCON), and Asynchronous Transfer Mode (ATM).

### 6.6.1 Fibre Channel Protocol Stack

It is easier to understand a communication protocol by viewing it as a structure of independent layers. FCP defines the communication protocol in five layers: FC-0 through FC-4 (except FC-3 layer, which is not implemented). In a layered communication model, the peer layers on each node talk to each other through defined protocols. Figure 6-13 illustrates the fibre channel protocol stack.



**Figure 6-13:** Fibre channel protocol stack

#### ***FC-4 Upper Layer Protocol***

FC-4 is the uppermost layer in the FCP stack. This layer defines the application interfaces and the way Upper Layer Protocols (ULPs) are mapped to the lower FC layers. The FC standard defines several protocols that can operate on the FC-4 layer (see Figure 6-7). Some of the protocols include SCSI, HIPPI Framing Protocol, Enterprise Storage Connectivity (ESCON), ATM, and IP.



## FC-2 Transport Layer

The FC-2 is the transport layer that contains the payload, addresses of the source and destination ports, and link control information. The FC-2 layer provides Fibre Channel addressing, structure, and organization of data (frames, sequences, and exchanges). It also defines fabric services, classes of service, flow control, and routing.

## FC-1 Transmission Protocol

This layer defines the transmission protocol that includes serial encoding and decoding rules, special characters used, and error control. At the transmitter node, an 8-bit character is encoded into a 10-bit transmissions character. This character is then transmitted to the receiver node. At the receiver node, the 10-bit character is passed to the FC-1 layer, which decodes the 10-bit character into the original 8-bit character.

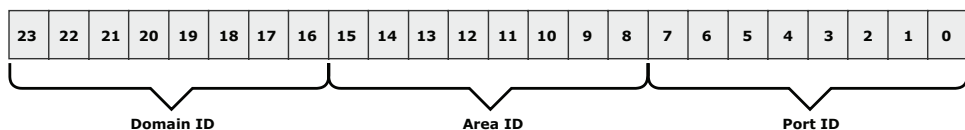
## FC-0 Physical Interface

FC-0 is the lowest layer in the FCP stack. This layer defines the physical interface, media, and transmission of raw bits. The FC-0 specification includes cables, connectors, and optical and electrical parameters for a variety of data rates. The FC transmission can use both electrical and optical media.

## 6.6.2 Fibre Channel Addressing

An FC address is dynamically assigned when a port logs on to the fabric. The FC address has a distinct format that varies according to the type of node port in the fabric. These ports can be an N\_port and an NL\_port in a public loop, or an NL\_port in a private loop.

The first field of the FC address of an N\_port contains the domain ID of the switch (see Figure 6-14). This is an 8-bit field. Out of the possible 256 domain IDs, 239 are available for use; the remaining 17 addresses are reserved for specific services. For example, FFFFFFFC is reserved for the name server, and FFFFFFFE is reserved for the fabric login service. The maximum possible number of N\_ports in a switched fabric is calculated as 239 domains  $\times$  256 areas  $\times$  256 ports = 15,663,104 Fibre Channel addresses.



**Figure 6-14:** 24-bit FC address of N\_port

The area ID is used to identify a group of F\_ports. An example of a group of F\_ports would be a card on the switch with more than one port on it. The last field in the FC address identifies the F\_port within the group.

### FC Address of an NL\_port

The FC addressing scheme for an NL\_port differs from other ports. The two upper bytes in the FC addresses of the NL\_ports in a private loop are assigned zero values. However, when an arbitrated loop is connected to a fabric through an FL\_port, it becomes a public loop. In this case, an NL\_port supports a fabric login. The two upper bytes of this NL\_port are then assigned a positive value, called a *loop identifier*, by the switch. The loop identifier is the same for all NL\_ports on a given loop.

Figure 6-15 illustrates the FC address of an NL\_port in both a public loop and a private loop. The last field in the FC addresses of the NL\_ports, in both public and private loops, identifies the AL-PA. There are 127 allowable AL-PA addresses; one address is reserved for the FL\_port on the switch.

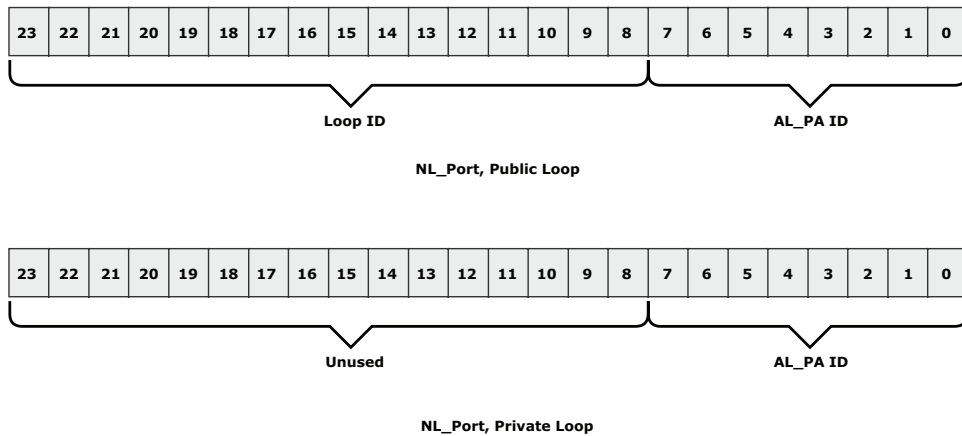



Figure 6-15: 24-bit FC address of NL\_port

**N\_PORT ID VISUALIZATION (NPIV)**



**NPIV is a Fibre Channel configuration that enables multiple N\_port IDs to share a single physical N\_port.**

## World Wide Names

Each device in the FC environment is assigned a 64-bit unique identifier called the *World Wide Name* (WWN). The Fibre Channel environment uses two types of WWNs: World Wide Node Name (WWNN) and World Wide Port Name (WWPN). Unlike an FC address, which is assigned dynamically, a WWN is a static name for each device on an FC network. WWNs are similar to the Media Access Control (MAC) addresses used in IP networking. WWNs are *burned* into the hardware or assigned through software. Several configuration definitions in a SAN use WWN for identifying storage devices and HBAs. The name server in an FC environment keeps the association of WWNs to the dynamically created FC addresses for nodes. Figure 6-16 illustrates the WWN structure for an array and the HBA.

World Wide Name - Array															
5	0	0	6	0	1	6	0	0	0	6	0	0	1	B	2
0101	0000	0000	0110	0000	0001	0110	0000	0000	0000	0110	0000	0000	0001	1011	0010
Company ID 24 bits							Port	Model Seed 32 bits							

World Wide Name - HBA															
1	0	0	0	0	0	0	0	c	9	2	0	d	c	4	0
Reserved 12 bits			Company ID 24 bits						Company Specific 24 bits						

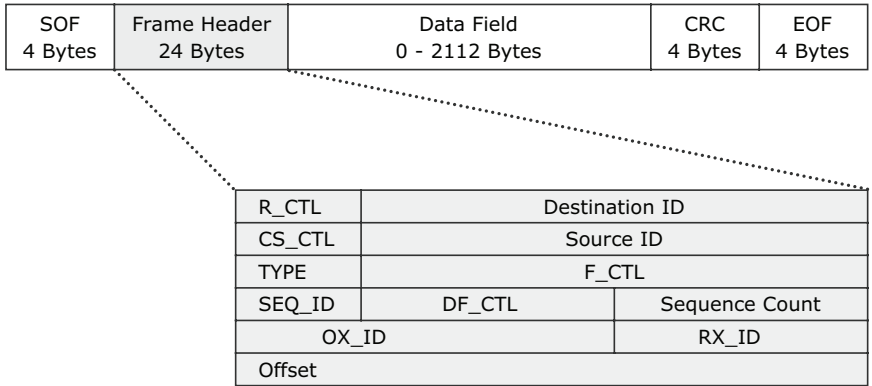
**Figure 6-16:** World Wide Names

### 6.6.3 FC Frame

An FC frame (Figure 6-17) consists of five parts: *start of frame (SOF)*, *frame header*, *data field*, *cyclic redundancy check (CRC)*, and *end of frame (EOF)*.

The SOF and EOF act as delimiters. In addition to this role, the SOF is a flag that indicates whether the frame is the first frame in a sequence of frames.

The frame header is 24 bytes long and contains addressing information for the frame. It includes the following information: Source ID (S\_ID), Destination ID (D\_ID), Sequence ID (SEQ\_ID), Sequence Count (SEQ\_CNT), Originating Exchange ID (OX\_ID), and Responder Exchange ID (RX\_ID), in addition to some control fields.



**Figure 6-17:** FC frame

The S\_ID and D\_ID are standard FC addresses for the source port and the destination port, respectively. The SEQ\_ID and OX\_ID identify the frame as a component of a specific sequence and exchange, respectively.

The frame header also defines the following fields:

- **Routing Control (R\_CTL):** This field denotes whether the frame is a link control frame or a data frame. Link control frames are nondata frames that do not carry any payload. These frames are used for setup and messaging. In contrast, data frames carry the payload and are used for data transmission.
- **Class Specific Control (CS\_CTL):** This field specifies link speeds for class 1 and class 4 data transmission.
- **TYPE:** This field describes the upper layer protocol (ULP) to be carried on the frame if it is a data frame. However, if it is a link control frame, this field is used to signal an event such as “fabric busy.” For example, if the TYPE is 08, and the frame is a data frame, it means that the SCSI will be carried on an FC.
- **Data Field Control (DF\_CTL):** A 1-byte field that indicates the existence of any optional headers at the beginning of the data payload. It is a mechanism to extend header information into the payload.
- **Frame Control (F\_CTL):** A 3-byte field that contains control information related to frame content. For example, one of the bits in this field indicates whether this is the first sequence of the exchange.

The data field in an FC frame contains the data payload, up to 2,112 bytes of original data — in most cases, SCSI data. The biggest possible payload an FC frame can deliver is 2,112 bytes of data with 36 bytes of fixed overhead. A link control frame, by definition, has a payload of 0 bytes. Only data frames carry a payload.

The CRC checksum facilitates error detection for the content of the frame. This checksum verifies data integrity by checking whether the content of the frames was received correctly. The CRC checksum is calculated by the sender before encoding at the FC-1 layer. Similarly, it is calculated by the receiver after decoding at the FC-1 layer.

#### 6.6.4. Structure and Organization of FC Data

In an FC network, data transport is analogous to a conversation between two people, whereby a frame represents a word, a sequence represents a sentence, and an exchange represents a conversation.

- **Exchange operation:** An exchange operation enables two N\_ports to identify and manage a set of information units. This unit maps to a sequence. Sequences can be both unidirectional and bidirectional depending upon the type of data sequence exchanged between the initiator and the target.
- **Sequence:** A sequence refers to a contiguous set of frames that are sent from one port to another. A sequence corresponds to an information unit, as defined by the ULP.
- **Frame:** A frame is the fundamental unit of data transfer at Layer 2. Each frame can contain up to 2,112 bytes of payload.

#### 6.6.5 Flow Control

Flow control defines the pace of the flow of data frames during data transmission. FC technology uses two flow-control mechanisms: buffer-to-buffer credit (BB\_Credit) and end-to-end credit (EE\_Credit).

##### ***BB\_Credit***

FC uses the *BB\_Credit* mechanism for hardware-based flow control. *BB\_Credit* controls the maximum number of frames that can be present over the link at any given point in time. In a switched fabric, *BB\_Credit* management may take place between any two FC ports. The transmitting port maintains a count of free receiver buffers and continues to send frames if the count is greater than 0. The *BB\_Credit* mechanism provides frame acknowledgment through the *Receiver Ready (R\_RDY)* primitive.

##### ***EE\_Credit***

The function of end-to-end credit, known as *EE\_Credit*, is similar to that of *BB\_Credit*. When an initiator and a target establish themselves as nodes communicating with each other, they exchange the *EE\_Credit* parameters (part of Port Login).

The EE\_Credit mechanism affects the flow control for class 1 and class 2 traffic only.

### 6.6.6 Classes of Service

The FC standards define different classes of service to meet the requirements of a wide range of applications. The table below shows three classes of services and their features (Table 6-1).

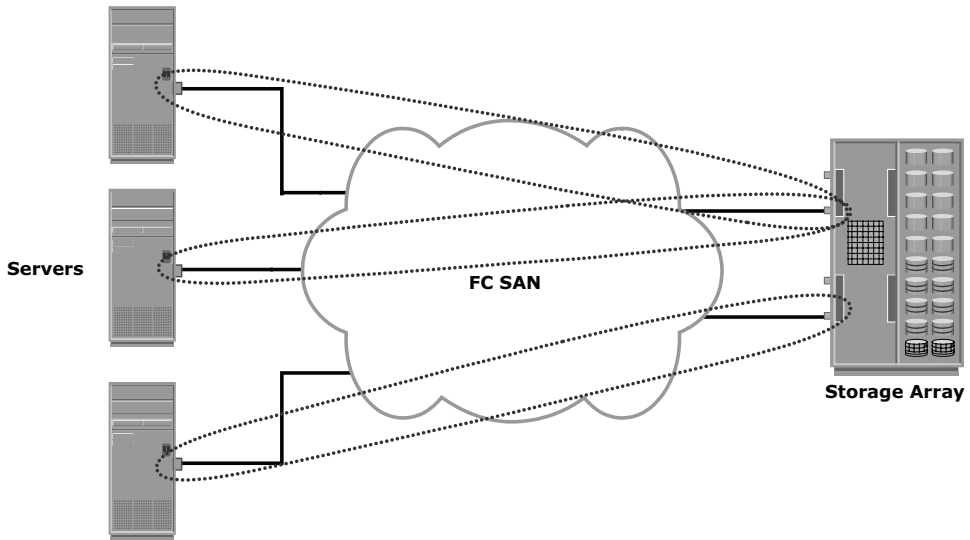
**Table 6-1:** FC Class of Services

	CLASS 1	CLASS 2	CLASS 3
Communication type	Dedicated connection	Nondedicated connection	Nondedicated connection
Flow control	End-to-end credit	End-to-end credit B-to-B credit	B-to-B credit
Frame delivery	In order delivery	Order not guaranteed	Order not guaranteed
Frame acknowledgement	Acknowledged	Acknowledged	Not acknowledged
Multiplexing	No	Yes	Yes
Bandwidth utilization	Poor	Moderate	High

Another class of services is *class F*, which is intended for use by the switches communicating through ISLs. Class F is similar to Class 2, and it provides notification of nondelivery of frames. Other defined Classes 4, 5, and 6 are used for specific applications. Currently, these services are not in common use.

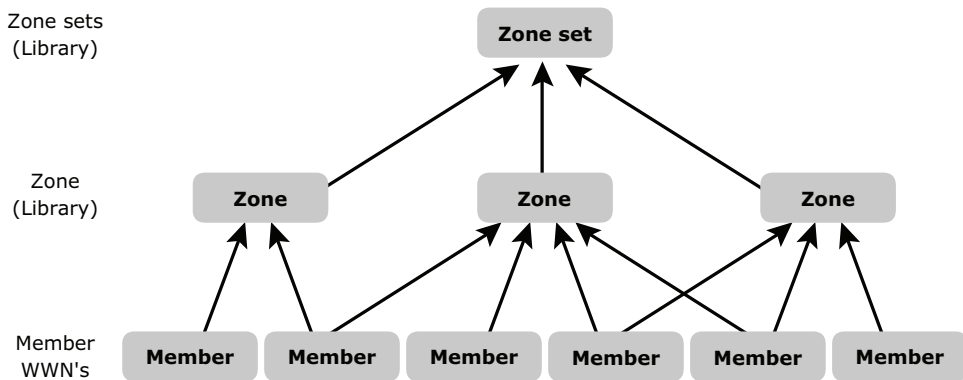
## 6.7 Zoning

Zoning is an FC switch function that enables nodes within the fabric to be logically segmented into groups that can communicate with each other (see Figure 6-18). When a device (host or storage array) logs onto a fabric, it is registered with the name server. When a port logs onto the fabric, it goes through a device discovery process with other devices registered in the name server. The zoning function controls this process by allowing only the members in the same zone to establish these link-level services.



**Figure 6-18:** Zoning

Multiple zone sets may be defined in a fabric, but only one zone set can be active at a time. A zone set is a set of zones and a zone is a set of members. A member may be in multiple zones. Members, zones, and zone sets form the hierarchy defined in the zoning process (see Figure 6-19). *Members* are nodes within the SAN that can be included in a zone. *Zones* comprise a set of members that have access to one another. A port or a node can be a member of multiple zones. *Zone sets* comprise a group of zones that can be activated or deactivated as a single entity in a fabric. Only one zone set per fabric can be active at a time. Zone sets are also referred to as *zone configurations*.



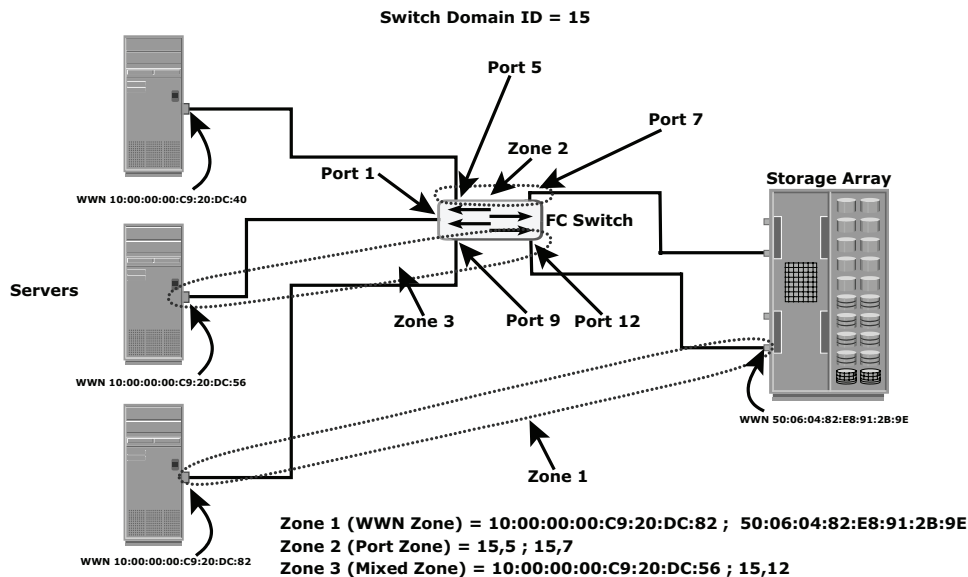
**Figure 6-19:** Members, zones, and zone sets

## Types of Zoning

Zoning can be categorized into three types:

- **Port zoning:** It uses the FC addresses of the physical ports to define zones. In port zoning, access to data is determined by the physical switch port to which a node is connected. The FC address is dynamically assigned when the port logs on to the fabric. Therefore, any change in the fabric configuration affects zoning. Port zoning is also called *hard zoning*. Although this method is secure, it requires updating of zoning configuration information in the event of fabric reconfiguration.
- **WWN zoning:** It uses World Wide Names to define zones. WWN zoning is also referred to as *soft zoning*. A major advantage of WWN zoning is its flexibility. It allows the SAN to be recabled without reconfiguring the zone information. This is possible because the WWN is static to the node port.
- **Mixed zoning:** It combines the qualities of both WWN zoning and port zoning. Using mixed zoning enables a specific port to be tied to the WWN of a node.

Figure 6-20 shows the three types of zoning on an FC network.



**Figure 6-20:** Types of zoning

Zoning is used in conjunction with LUN masking for controlling server access to storage. However, these are two different activities. Zoning takes place at the fabric level and LUN masking is done at the array level.



## 6.8 Fibre Channel Login Types

Fabric services define three login types:

- Fabric login (FLOGI) is performed between an N\_port and an F\_port. To log on to the fabric, a device sends a FLOGI frame with the World Wide Node Name (WWNN) and World Wide Port Name (WWPN) parameters to the login service at the well-known FC address FFFFFFFE. In turn, the switch accepts the login and returns an Accept (ACC) frame with the assigned FC address for the device. Immediately after the FLOGI, the N\_port registers itself with the local name server on the switch, indicating its WWNN, WWPN, and assigned FC address.
- Port login (PLOGI) is performed between an N\_port and another N\_port to establish a session. The initiator N\_port sends a PLOGI request frame to the target N\_port, which accepts it. The target N\_port returns an ACC to the initiator N\_port. Next, the N\_ports exchange service parameters relevant to the session.
- Process login (PRLI) is also performed between an N\_port and another N\_port. This login relates to the FC-4 ULPs such as SCSI. N\_ports exchange SCSI-3-related service parameters. N\_ports share information about the FC-4 type in use, the SCSI initiator, or the target.

### FAN-OUT AND FAN-IN



***Fan-out*** enables multiple server ports to communicate to a single storage port. A four server connection to a single storage port results in a fan-out ratio of 4. Typically, there is an architectural limit based on the array's capability to record and manage several initiators that connect to a port, as the hardware capabilities determine the possible aggregate IOPS and MB/s per port.

***Fan-in*** specifies accessibility of a host port to storage ports on multiple arrays. Like fan-out, the restrictions on fan-in are also based on an architectural limit.

## 6.9 FC Topologies

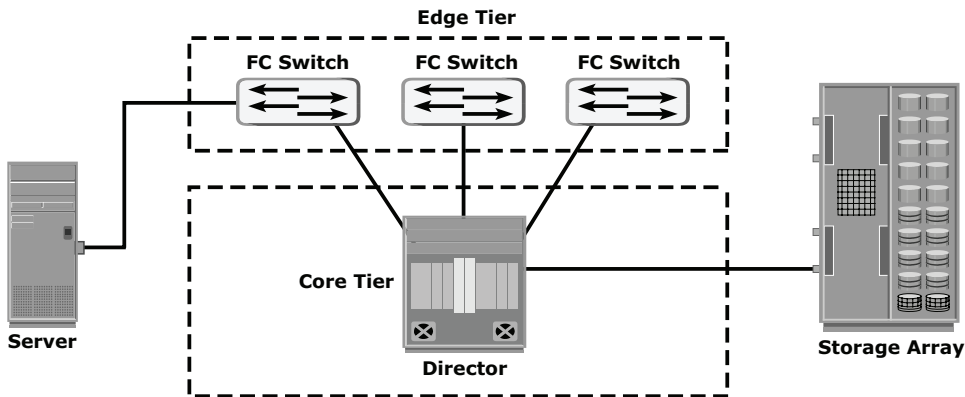
Fabric design follows standard topologies to connect devices. Core-edge fabric is one of the popular topology designs. Variations of core-edge fabric and mesh topologies are most commonly deployed in SAN implementations.

## 6.9.1 Core-Edge Fabric

In the *core-edge fabric* topology, there are two types of switch tiers in this fabric. The *edge tier* usually comprises switches and offers an inexpensive approach to adding more hosts in a fabric. The tier at the edge fans out from the tier at the core. The nodes on the edge can communicate with each other.

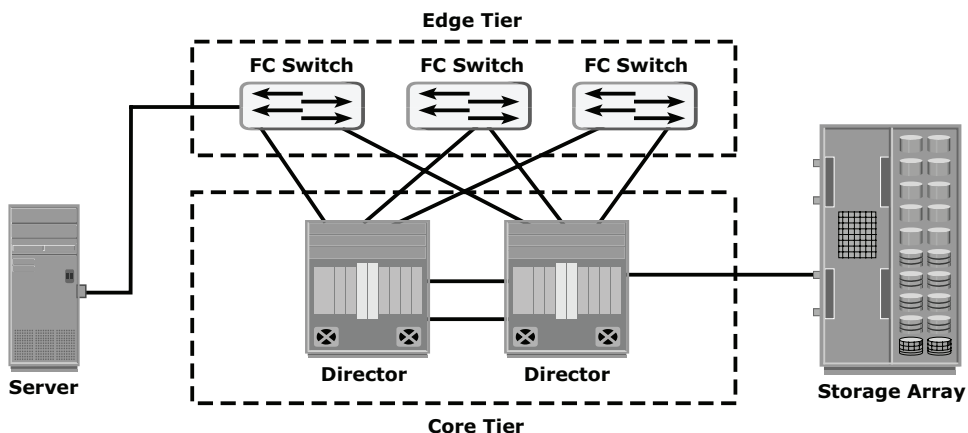
The *core tier* usually comprises enterprise directors that ensure high fabric availability. Additionally all traffic has to either traverse through or terminate at this tier. In a two-tier configuration, all storage devices are connected to the core tier, facilitating fan-out. The host-to-storage traffic has to traverse one and two ISLs in a two-tier and three-tier configuration, respectively. Hosts used for mission-critical applications can be connected directly to the core tier and consequently avoid traveling through the ISLs to process I/O requests from these hosts.

The core-edge fabric topology increases connectivity within the SAN while conserving overall port utilization. If expansion is required, an additional edge switch can be connected to the core. This topology can have different variations. In a *single-core topology*, all hosts are connected to the edge tier and all storage is connected to the core tier. Figure 6-21 depicts the core and edge switches in a single-core topology.



**Figure 6-21:** Single core topology

A *dual-core topology* can be expanded to include more core switches. However, to maintain the topology, it is essential that new ISLs are created to connect each edge switch to the new core switch that is added. Figure 6-22 illustrates the core and edge switches in a dual-core topology.



**Figure 6-22:** Dual-core topology

### ***Benefits and Limitations of Core-Edge Fabric***

The core-edge fabric provides one-hop storage access to all storage in the system. Because traffic travels in a deterministic pattern (from the edge to the core), a core-edge provides easier calculation of ISL loading and traffic patterns. Because each tier's switch is used for either storage or hosts, one can easily identify which resources are approaching their capacity, making it easier to develop a set of rules for scaling and apportioning.

A well-defined, easily reproducible building-block approach makes rolling out new fabrics easier. Core-edge fabrics can be scaled to larger environments by linking core switches, adding more core switches, or adding more edge switches. This method can be used to extend the existing simple core-edge model or to expand the fabric into a compound or complex core-edge model.

However, the core-edge fabric may lead to some performance-related problems because scaling a core-edge topology involves increasing the number of ISLs in the fabric. As more edge switches are added, the domain count in the fabric increases. A common best practice is to keep the number of host-to-storage hops unchanged, at one hop, in a core-edge. Hop count represents the total number of devices a given piece of data (packet) passes through. Generally a large hop count means greater the transmission delay between data traverse from its source to destination.

As the number of cores increases, it may be prohibitive to continue to maintain ISLs from each core to each edge switch. When this happens, the fabric design can be changed to a compound or complex core-edge design.

### BLADE SERVER



**Blade server architecture deployment has rapidly increased server count in modern data centers.**

**In blade server architecture, the backplane hosts the server blades and the I/O modules. High-end blade servers have up to 16 server blades and 8 I/O modules configured. The server blades are hot pluggable. The FC switch module in the chassis takes the place of the edge switch in standard core-edge fabrics, which also reduces the required cable count. Blade servers use mezzanine cards instead of HBAs for FC connectivity. The mezzanine cards connect the internal ports on the switch module through the backplane.**

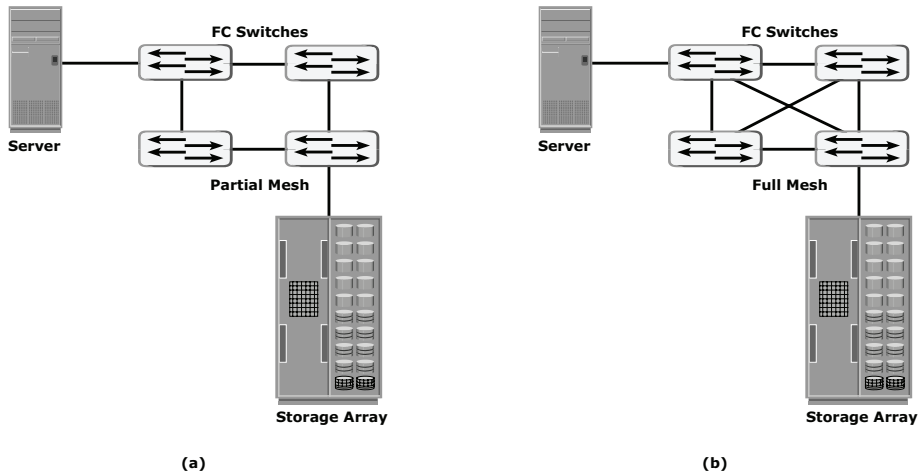
## 6.9.2 Mesh Topology

In a *mesh topology*, each switch is directly connected to other switches by using ISLs. This topology promotes enhanced connectivity within the SAN. When the number of ports on a network increases, the number of nodes that can participate and communicate also increases.

A mesh topology may be one of the two types: full mesh or partial mesh. In a *full mesh*, every switch is connected to every other switch in the topology. Full mesh topology may be appropriate when the number of switches involved is small. A typical deployment would involve up to four switches or directors, with each of them servicing highly localized host-to-storage traffic. In a full mesh topology, a maximum of one ISL or hop is required for host-to-storage traffic.

In a *partial mesh* topology, several hops or ISLs may be required for the traffic to reach its destination. Hosts and storage can be located anywhere in the fabric, and storage can be localized to a director or a switch in both mesh topologies.

A full mesh topology with a symmetric design results in an even number of switches, whereas a partial mesh has an asymmetric design and may result in an odd number of switches. Figure 6-23 depicts both a full mesh and a partial mesh topology.



**Figure 6-23:** Partial mesh and full mesh topologies

## 6.10 Concepts in Practice: EMC Connectrix

This section discusses the Connectrix connectivity products offered by EMC that provide connectivity in large-scale, workgroup, mid-tier, and mixed iSCSI and FC environments. For the latest information, visit <http://education.EMC.com/ismbook>.

FC switches and directors are key components of the SAN environment. EMC offers the following connectivity products under the Connectrix brand (see Figure 6-24):

- Enterprise directors
- Departmental switches
- Multiprotocol routers



**Enterprise Director**



**Departmental Switches**



**Multiprotocol Router**

**Figure 6-24:** EMC Connectrix

Enterprise directors are ideal for large enterprise connectivity. They offer high port density and high component redundancy. Enterprise directors are deployed in high-availability or large-scale environments. Connectrix directors offer several hundred ports per domain. Departmental switches are best suited for workgroup, mid-tier environments. Multiprotocol routers support mixed iSCSI and FC environments. They can bridge FC SAN and IP SAN, a feature that enables these routers to provide connectivity between iSCSI host initiators and FC storage targets. They can extend FC SAN over long distances through IP networks.

In addition to FC ports, Connectrix switches and directors have Ethernet ports and serial ports for communication and switch management functions. Connectrix management software enables configuration, monitoring, and management of Connectrix switches.

### **6.10.1 Connectrix Switches**

B-series and MDS make up the Connectrix family of switches offered by EMC. These switches offer scalability up to 80 ports and are designed to meet workgroup, department-level, and enterprise-level requirements. They are designed with a nonblocking architecture and can operate in heterogeneous environments. The features of these switches that ensure their high availability are their nondisruptive software, port upgrade, redundant, and hot-swappable components. These switches can be managed through CLI, Web Tools, and Fabric Manager.

### **6.10.2 Connectrix Directors**

EMC offers the high-end Connectrix family of directors. Their modular architectural design offers scalability up to 528 ports. They are suitable for server and storage consolidation across enterprises. These directors have redundant components for high availability and they provide multi-protocol connectivity for both mainframe and open system environments. Connectrix directors offer high speeds (up to 10 Gb/s) and high system bandwidth (up to 2.2 Tb/s). They also support ISL trunking and in-band and out-of-band management functionality. The connectrix director also offers a virtual SAN feature for fabric management and security. Like switches, directors can be managed through CLI or other GUI tools.

### **6.10.3 Connectrix Management Tools**

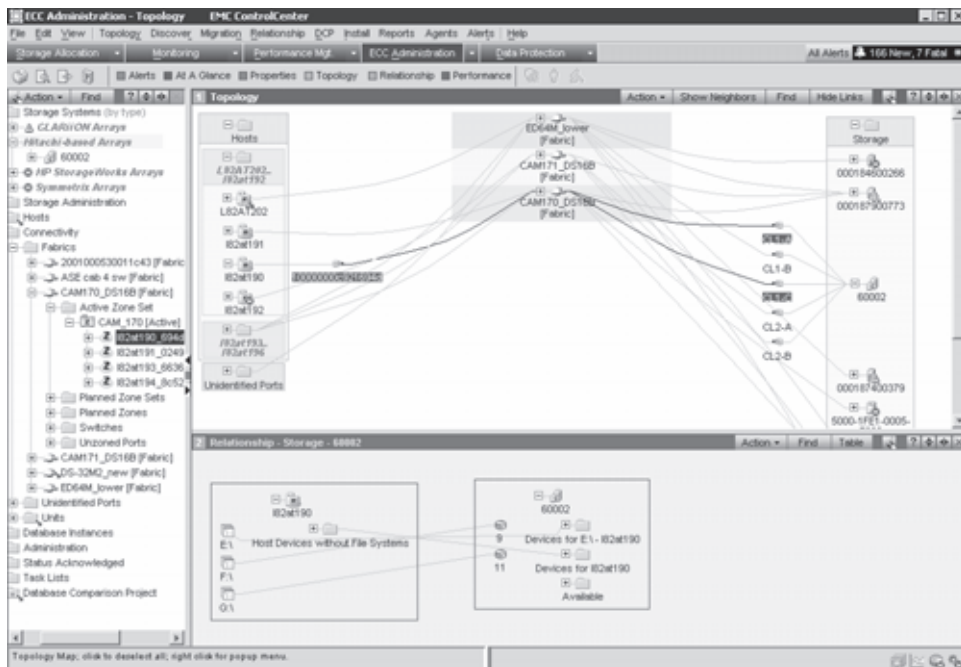
There are several ways to monitor and manage FC switches in a fabric. Individual switch management is accomplished through the console port, using CLI or browser-based tools.

Command-line utilities such as Telnet and SSH may be used to log on to the switch over IP and issue CLI commands. The primary purpose of the CLI is to automate the management of a large number of switches or directors with the use of scripts. The third option is to use browser-based tools that provide GUIs. These Java-based tools can also display the topology map.

Fabricwide management and monitoring is accomplished by using vendor-specific tools and Simple Network Management Protocol (SNMP)-based, third-party software.

EMC ControlCenter SAN Manager provides a single interface for managing Storage Area Network. With SAN Manager one can discover, monitor, manage, and configure complex heterogeneous SAN environments faster and easier. It streamlines and centralizes SAN management operations across multi-vendor storage networks and storage devices. It enables storage administrators to manage SAN zones and LUN masking consistently across multi-vendor SAN arrays and switches. EMC ControlCenter SAN Manager also supports virtual environments including VMware, Symmetrix Virtual Provisioning, and Virtual SANs.

Figure 6-25 illustrates EMC ControlCenter SAN Manager interface.



**Figure 6-25:** Managing FC switches through SAN Manager

## Summary

---

The SAN has enabled the consolidation of storage and benefited organizations by lowering the cost of storage service delivery. SAN reduces overall operational cost and downtime and enables faster application deployment.

SANs and tools that have emerged for SANs enable data centers to allocate storage to an application and migrate workloads between different servers and storage devices dynamically. This significantly increases server utilization.

SANs simplify the business-continuity process because organizations are able to logically connect different data centers over long distances and provide cost-effective, disaster recovery services that can be effectively tested.

The adoption of SANs has increased with the decline of hardware prices and has enhanced the maturity of storage network standards. Small and medium-size enterprises and departments that initially resisted shared storage pools have now begun to adopt SANs.

This chapter detailed the components of a SAN and the FC technology that forms its backbone. FC meets today's demands for reliable, high-performance, and low-cost applications.

The interoperability between FC switches from different vendors has enhanced significantly compared to early SAN deployments. The standards published by a dedicated study group within T11 on SAN routing, and the new product offerings from vendors, are now revolutionizing the way SANs are deployed and operated.

Although SANs have eliminated islands of storage, their initial implementation created islands of SANs in an enterprise. The emergence of the iSCSI and FCIP technologies, detailed in Chapter 8, has pushed the convergence of the SAN with IP technology, providing more benefits to using storage technologies.

### EXERCISES

- 1. What is zoning? Discuss a scenario,**
  - (i) where soft zoning is preferred over hard zoning.**
  - (ii) where hard zoning is preferred over soft zoning.**
- 2. Describe the process of assigning FC address to a node when logging in to the network for the first time.**
- 3. Seventeen switches, with 16 ports each, are connected in a mesh topology. How many ports are available for host and storage connectivity if you create a high-availability solution?**
- 4. Discuss the advantage of FC-SW over FC-AL.**
- 5. How flow control works in FC network.**
- 6. Why is class 3 service most preferred for FC communication?**