Information Storage and Management
Unit 3 – Direct Attached Storage and Introduction to SCSI

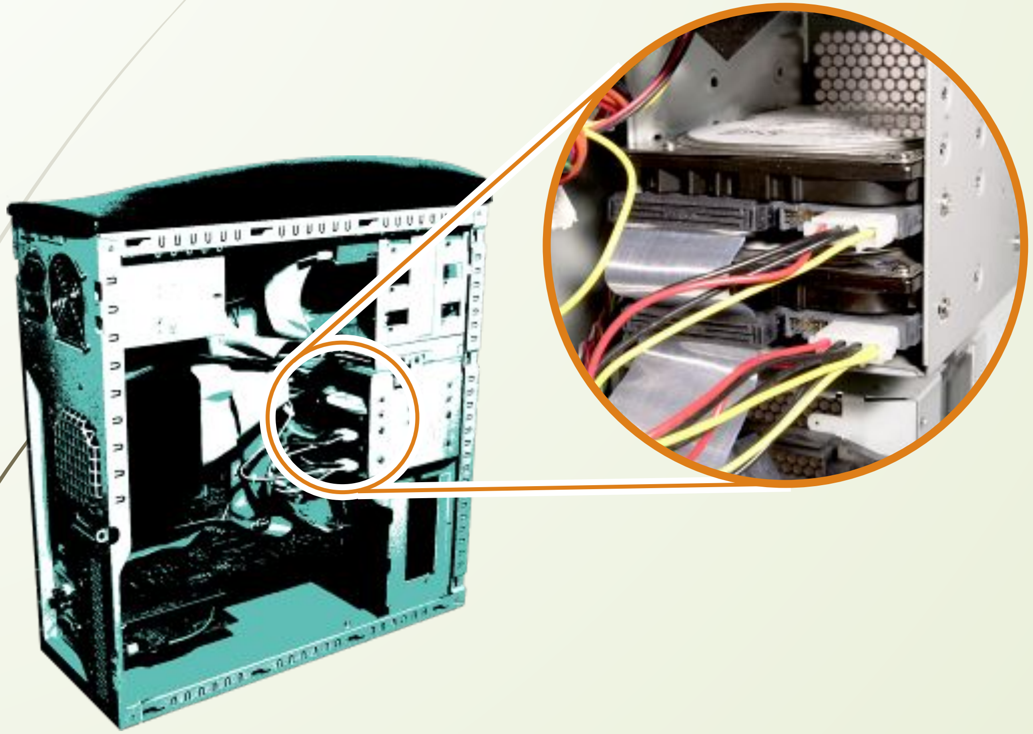
UNIT-III

- Direct-Attached Storage, SCSI, and Storage Area Networks: Types of DAS, DAS Benefits and Limitations, Disk Drive Interfaces, Introduction to Parallel SCSI, Overview of Fibre Channel, The SAN and Its Evolution, Components of SAN, FC Connectivity, Fibre Channel Ports, Fibre Channel Architecture, Zoning, Fibre Channel Login Types, FC Topologies

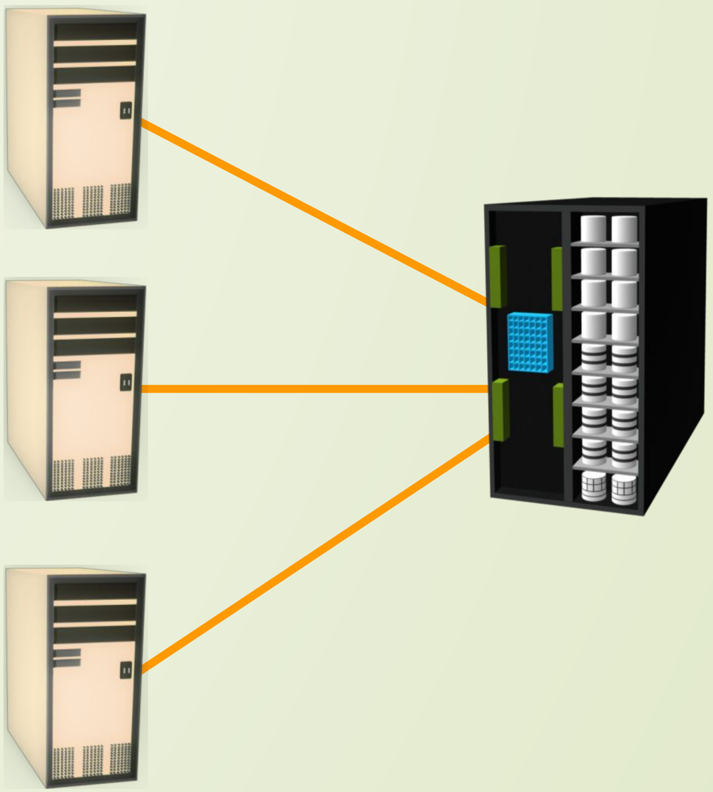
Direct-Attached Storage

- Direct – Attached storage (DAS) is a an architecture where storage connects directly to servers.
- Applications access data from DAS using block-level access protocols.
- DAS is ideal for localized data access and sharing in environments that have a small number of servers.
- Ex: small businesses, departments and workgroups do not share information across enterprises

What is DAS?



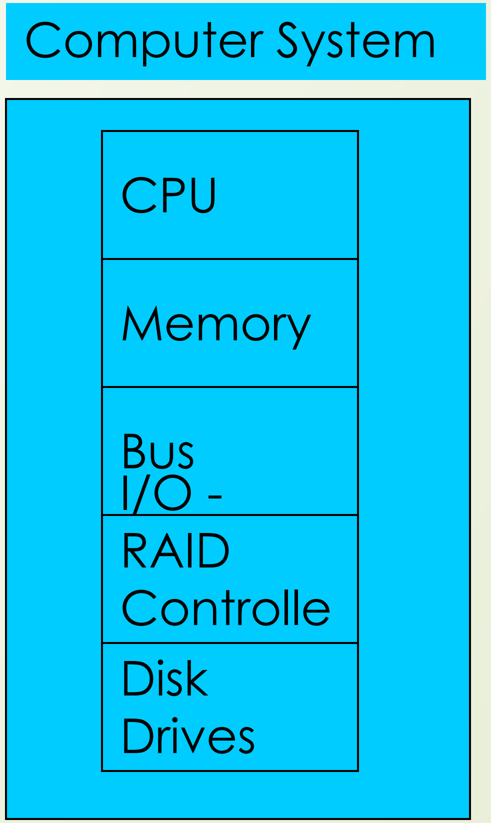
Internal Direct Connect



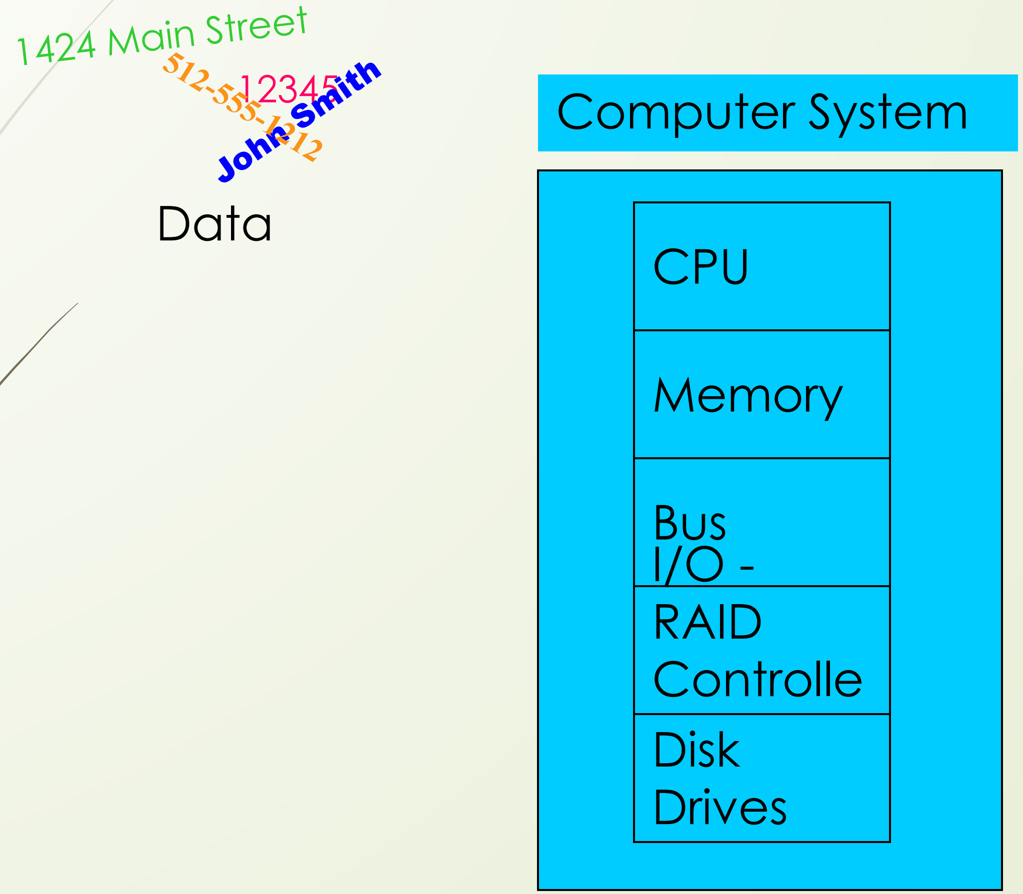
External Direct Connect

- Uses block level protocol for data access

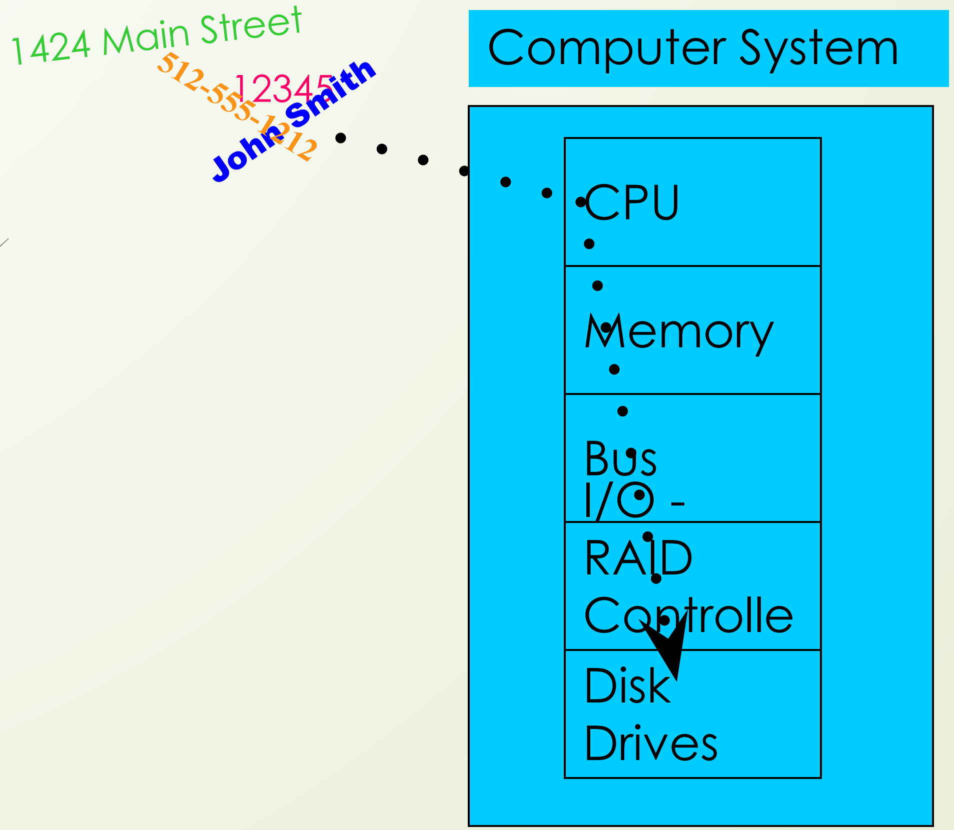
Direct Attached Storage (Internal)



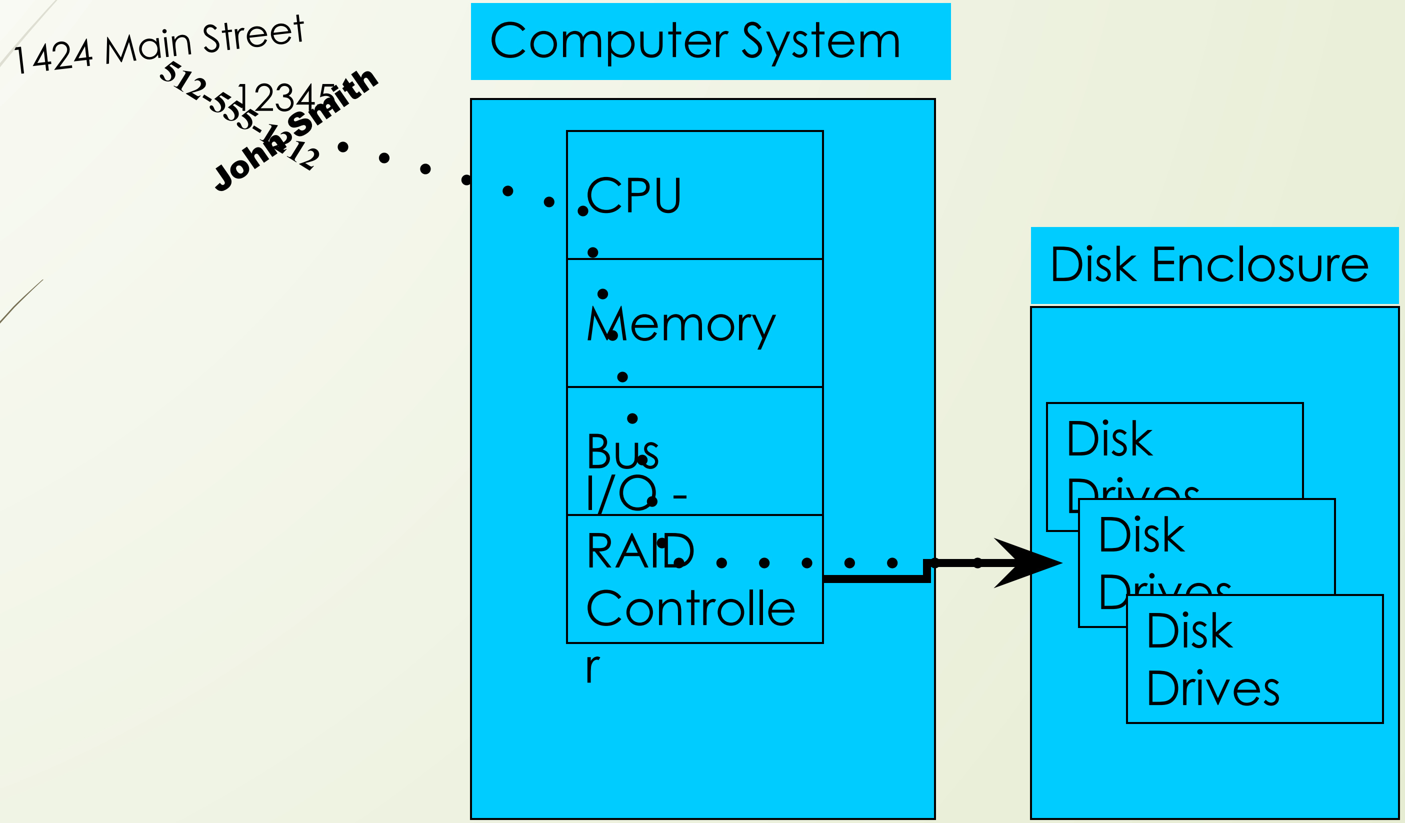
Direct Attached Storage (Internal)



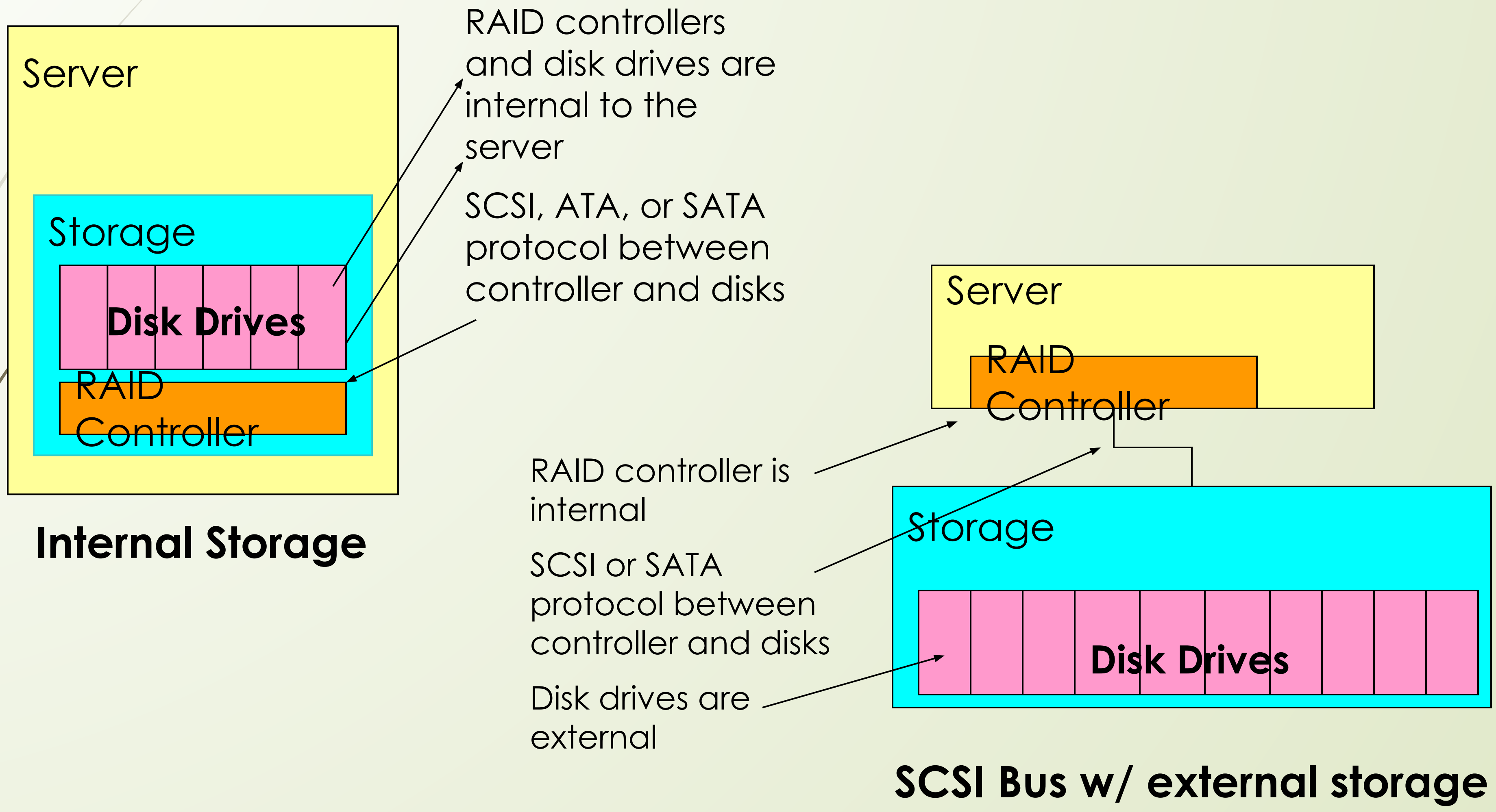
Direct Attached Storage (Internal)



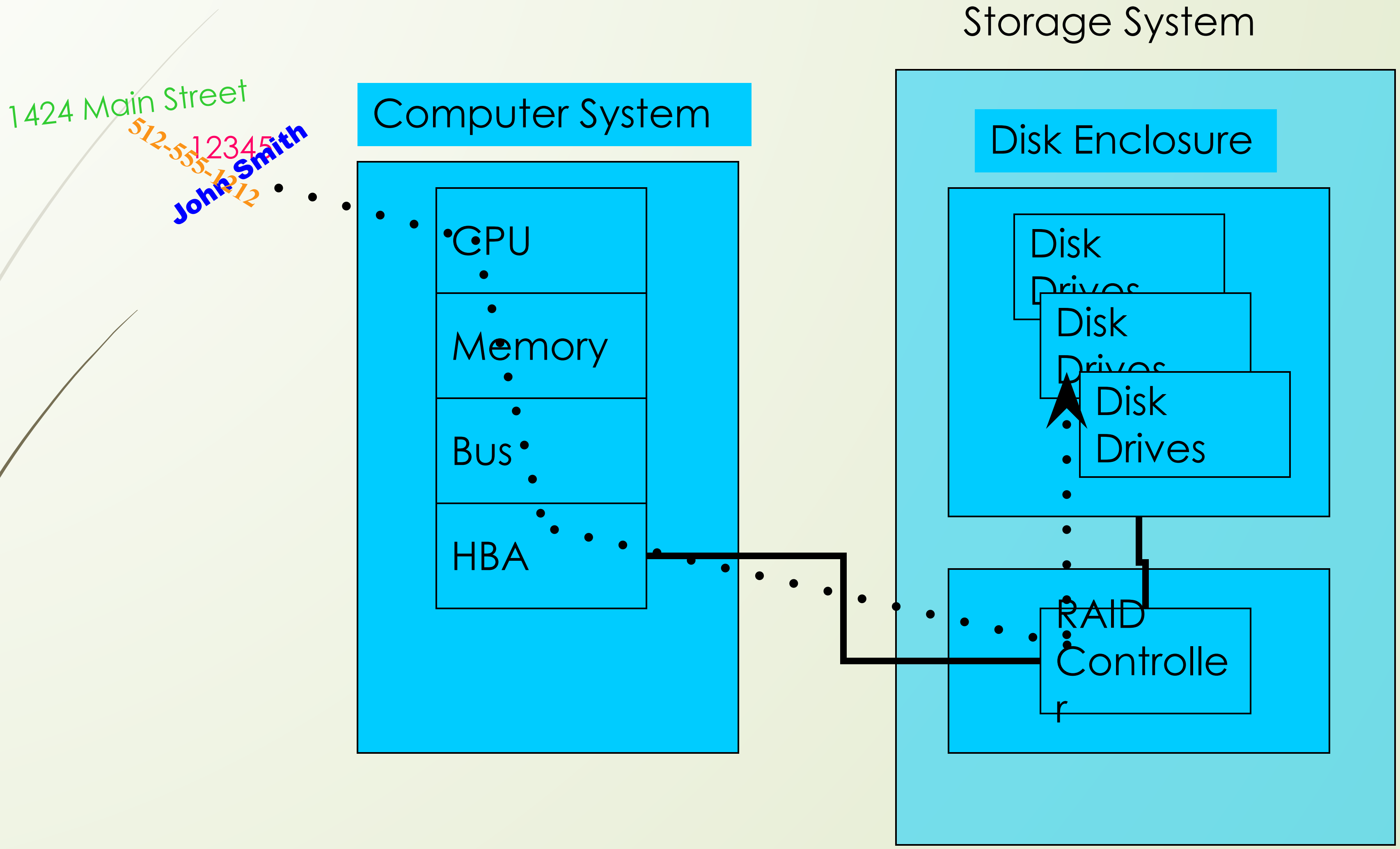
DAS w/ internal controller and external storage



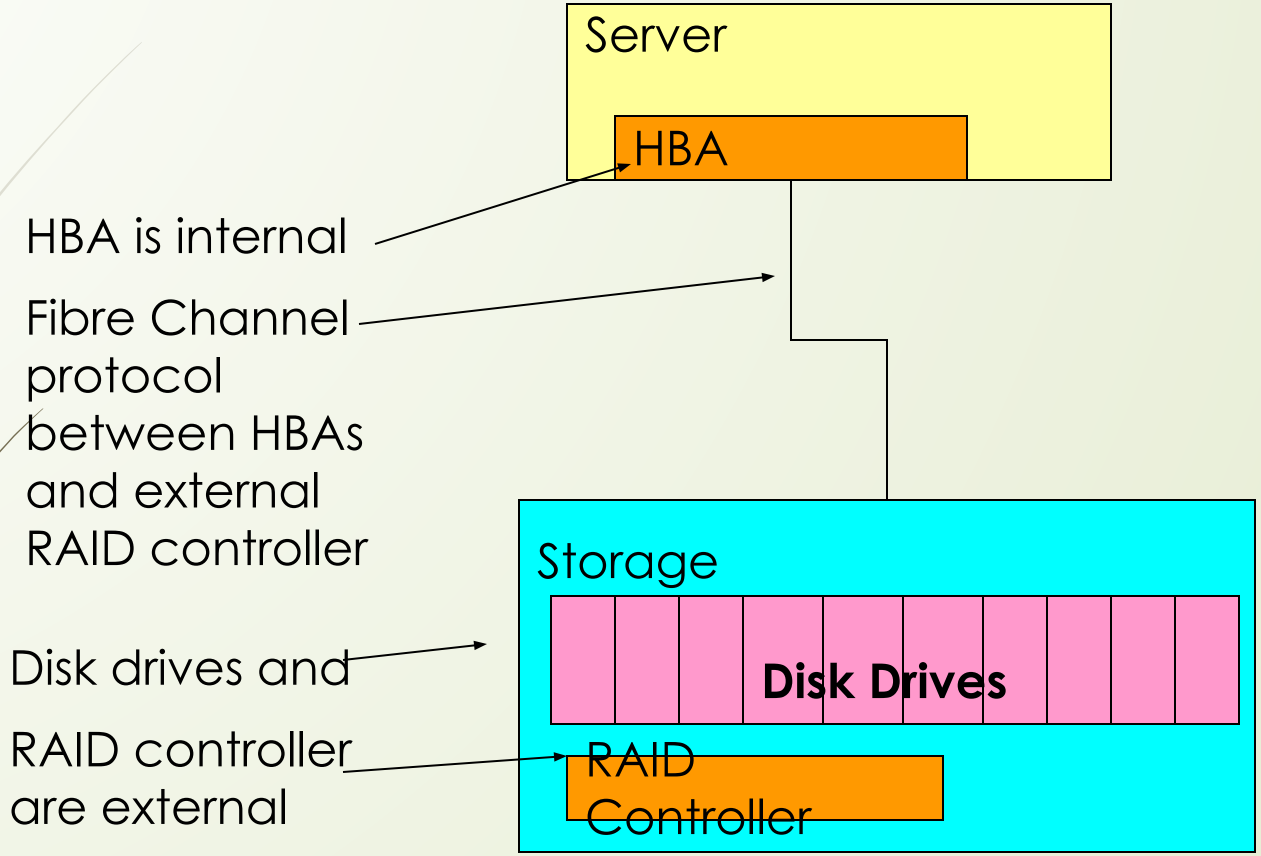
Comparing Internal and External Storage



DAS w/ external controller and external storage



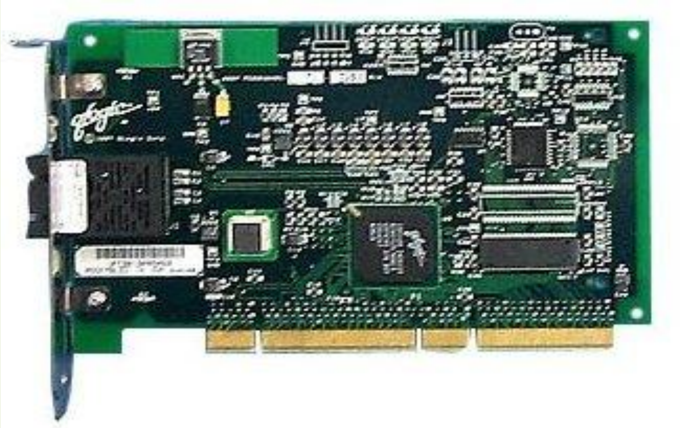
DAS over Fibre Channel



External SAN Array

I/O Transfer

- RAID Controller
  - Contains the “smarts”
  - Determines how the data will be written (striping, mirroring, RAID 10, RAID 5, etc.)
- Host Bus Adapter (HBA)
  - Simply transfers the data to the RAID controller.
  - Doesn’t do any RAID or striping calculations.
  - “Dumb” for speed.
  - Required for external storage.



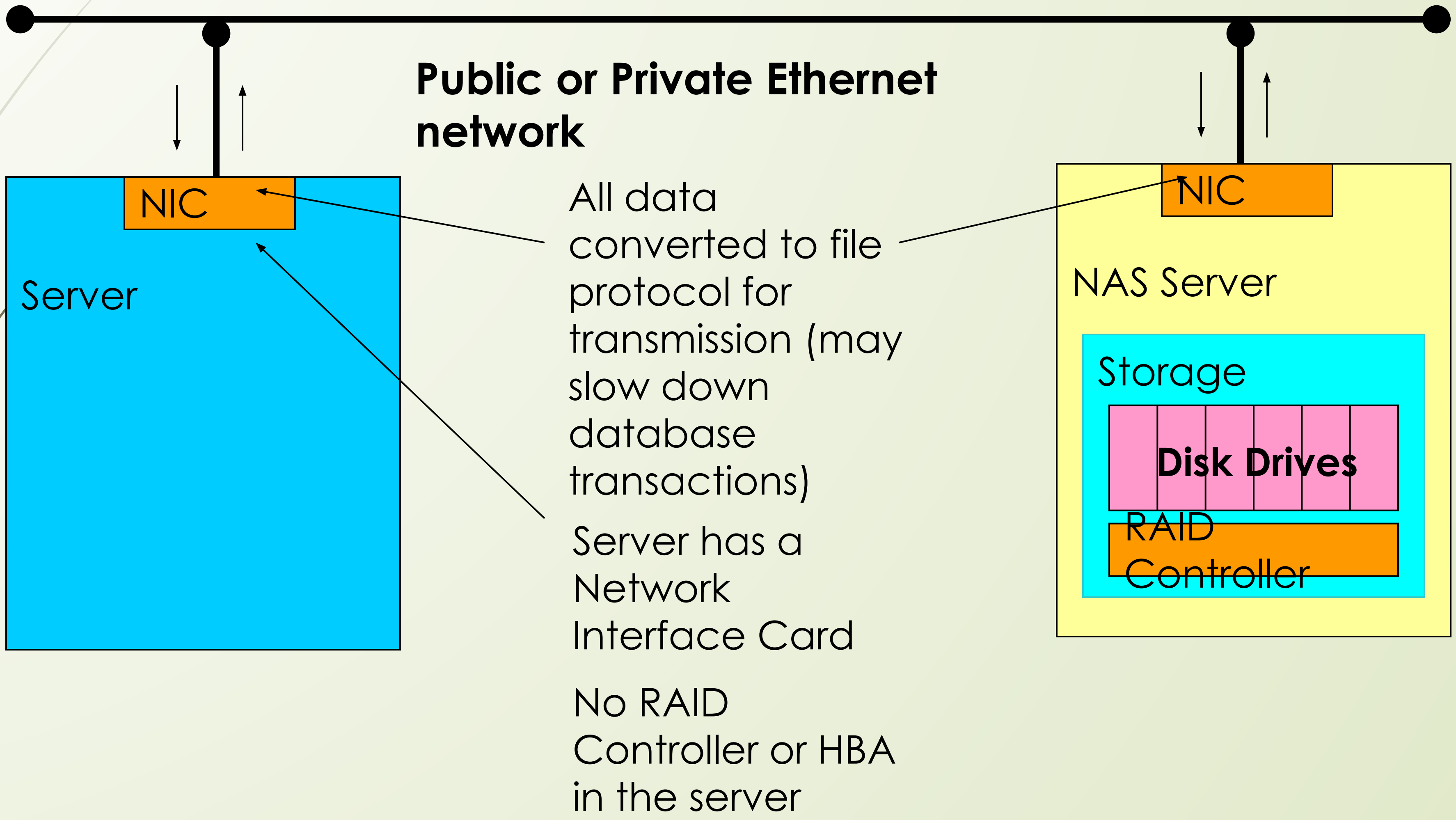
Storage types

- Single Disk Drive
- JBOD
- Volume
- Storage Array
- SCSI device
- DAS
- NAS
- SAN
- iSCSI

NAS: What is it?

- Network Attached Storage
- Utilizes a TCP/IP network to “share” data
- Uses file sharing protocols like Unix NFS and Windows CIFS
- Storage “Appliances” utilize a stripped-down OS that optimizes file protocol performance

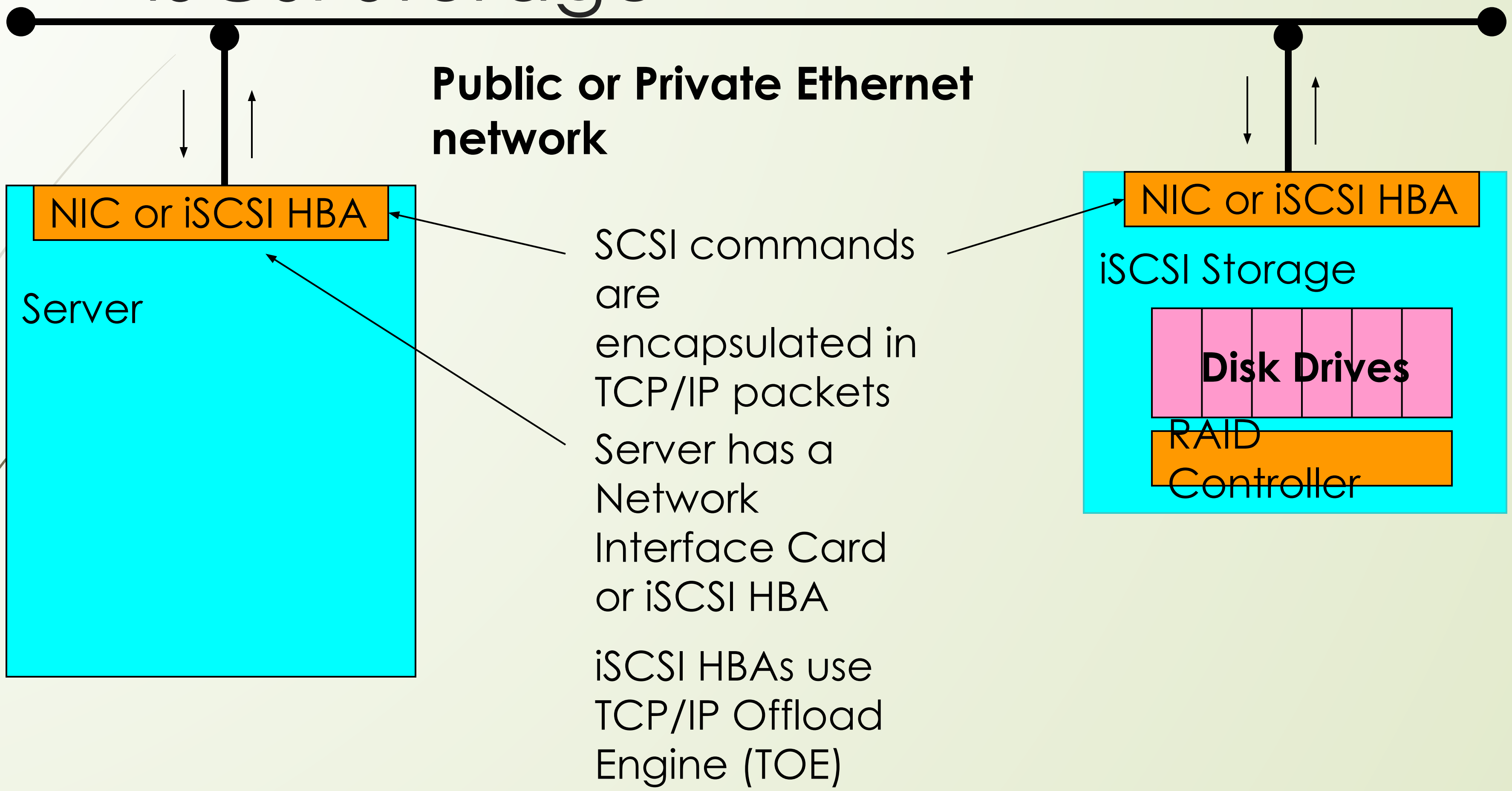
Networked Attached Storage



iSCSI: What is it?

- An alternate form of networked storage
- Like NAS, also utilizes a TCP/IP network
- Encapsulates native SCSI commands in TCP/IP packets
- Supported in Windows 2003 Server and Linux
- TCP/IP Offload Engines (TOEs) on NICs speed up packet encapsulation

iSCSI Storage



DAS Benefits

- Ideal for local data provisioning
- Quick deployment for small environments
- Simple to deploy
- Low capital expense
- Low complexity
- DAS needs lower initial investment than storage networking.
- Setup is managed using host-based tools like host OS.
- It requires fewer management tasks and less hardware and software elements to set up and operate.

DAS Challenges

- Scalability is limited
  - Number of connectivity ports to hosts
  - Difficulty to add more capacity
  - Limited bandwidth
  - Distance limitations
- Downtime required for maintenance with internal DAS
- Limited ability to share resources
  - Array front-end port
  - Unused resources cannot be easily re-allocated
  - Resulting in islands of over and under utilized storage pools

DAS Connectivity Options

- ATA (IDE) and SATA
  - Primarily for internal bus
- SCSI
  - Parallel (primarily for internal bus)
  - Serial (external bus)
- FC
  - High speed network technology
- Buss and Tag
  - Primarily for external mainframe
  - Precursor to ESCON and FICON

Types of DAS

- There are two types of DAS depending on the location of the storage device with respect to the host.
- Internal DAS
- External DAS

DAS Management

- Internal
  - Host provides:
    - Disk partitioning (Volume management)
    - File system layout
  - Direct Attached Storage managed individually through the server and the OS
- External
  - Array based management
  - Lower TCO (Total Cost of Ownership) for managing data and storage Infrastructure

Internal DAS

- In the internal DAS architecture, the storage device is internally connected to the host by a serial or parallel bus.
- The physical bus has distance limitations and can only be sustained over a shorter distance for high-speed connectivity.
- Most internal buses can support only a limited number of devices
- They occupy a large amount of space inside the host, making maintenance of other components difficult

External DAS

- In the external DAS architecture, the server connects directly to the external storage device.
- Communication between the host and the storage device takes place over SCSI (Simple Computer System Interconnect) or FC (Fiber Channel) protocol.

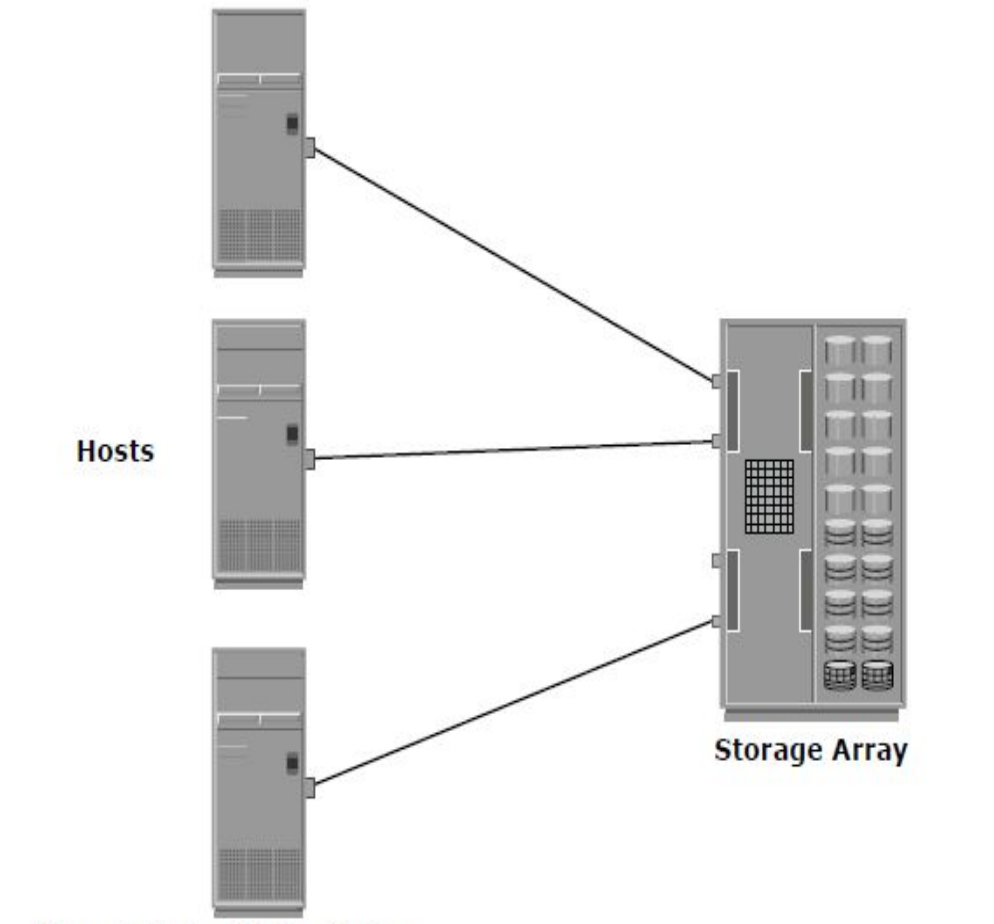


Figure 5-1: External DAS architecture

DAS Limitations

- A storage device has a limited number of ports.
- A limited bandwidth in DAS restricts the available I/O processing capability.
- The distance limitations associated with implementing DAS because of direct connectivity requirements can be addressed by using Fibre Channel connectivity.
- Unused resources cannot be easily re-allocated, resulting in islands of over-utilized and under-utilized storage pools.
- It is not scalable.
- Disk utilization, throughput, and cache memory of a storage device, along with virtual memory of a host govern the performance of DAS.
- RAID-level configurations, storage controller protocols, and the efficiency of the bus are additional factors that affect the performance of DAS.
- The absence of storage interconnects and network latency provide DAS with the potential to outperform other storage networking configurations.

Disk Drive Interfaces

- The host and the storage device in DAS communicate with each other by using predefined protocols such as IDE/ATA, SATA, SAS, SCSI, and FC.
- These protocols are implemented on the HDD controller. Therefore, a storage device is also known by the name of the protocol it supports.

IDE/ATA

- An Integrated Device Electronics/Advanced Technology Attachment (IDE/ATA) disk supports the IDE protocol.
- The IDE component in IDE/ATA provides the specification for the controllers connected to the computer's motherboard for communicating with the device attached.
- The ATA component is the interface for connecting storage devices, such as CD-ROMs, floppy disk drives, and HDDs, to the motherboard.
- IDE/ATA has a variety of standards and names, such as ATA, ATA/ATAPI, EIDE, ATA-2, Fast ATA, ATA-3, Ultra ATA, and Ultra DMA.
- The latest version of ATA—Ultra DMA/133—supports a throughput of 133 MB per second.

IDE/ATA

- In a master-slave configuration, an ATA interface supports two storage devices per connector. However, if the performance of the drive is important, sharing a port between two devices is not recommended.
- A 40-pin connector is used to connect ATA disks to the motherboard, and a 34-pin connector is used to connect floppy disk drives to the motherboard.
- An IDE/ATA disk offers excellent performance at low cost, making it a popular and commonly used hard disk.

SATA

- A SATA (Serial ATA) is a serial version of the IDE/ATA specification.
- SATA is a disk-interface technology that was developed by a group of the industry's leading vendors with the aim of replacing parallel ATA.
- A SATA provides point-to-point connectivity up to a distance of one meter and enables data transfer at a speed of 150 MB/s.
- Enhancements to the SATA have increased the data transfer speed up to 600 MB/s.
- A SATA bus directly connects each storage device to the host through a dedicated link, making use of low-voltage differential signaling (LVDS
- LVDS is an electrical signaling system that can provide high-speed connectivity over low-cost, twisted-pair copper cables. For data transfer, a SATA bus uses LVDS with a voltage of 250 mV.

SATA

- A SATA bus uses a small 7-pin connector and a thin cable for connectivity.
- A SATA port uses 4 signal pins, which improves its pin efficiency compared to the parallel ATA that uses 26 signal pins, for connecting an 80-conductor ribbon cable to a 40-pin header connector.
- SATA devices are hot-pluggable, which means that they can be connected or removed while the host is up and running.
- A SATA port permits single-device connectivity.
- Connecting multiple SATA drives to a host requires multiple ports to be present on the host.
- Single-device connectivity enforced in SATA, eliminates the performance problems caused by cable or port sharing in IDE/ATA.

Evolution of Parallel SCSI

- Developed by Shugart Associates & named as SASI (Shugart Associates System Interface)
- ANSI acknowledged SCSI as an industry standard
- SCSI versions
  - SCSI–1
    - Defined cable length, signaling characteristics, commands & transfer modes
    - Used 8-bit narrow bus with maximum data transfer rate of 5 MB/s
  - SCSI–2
    - Defined Common Command Set (CCS) to address non-standard implementation of the original SCSI
    - Improved performance, reliability, and added additional features
  - SCSI–3
    - Latest version of SCSI
    - Comprised different but related standards, rather than one large document

Evolution of Parallel SCSI

- SCSI was developed to provide a device-independent mechanism for attaching to and accessing host computers.
- SCSI also provided an efficient peer-to-peer I/O bus that supported multiple devices.
- SCSI is commonly used as a hard disk interface. However, SCSI can be used to add devices, such as tape drives and optical media drives, to the host computer without modifying the system hardware or software.
- Over the years, SCSI has undergone radical changes and has evolved into a robust industry standard.

Evolution of Parallel SCSI

- SCSI, first developed for hard disks, is often compared to IDE/ATA.
- SCSI offers improved performance and expandability and compatibility options, making it suitable for high-end computers.
- However, the high cost associated with SCSI limits its popularity among home or business desktop users.
- SCSI is available in a variety of interfaces. Parallel SCSI (referred to as SCSI) is one of the oldest and most popular forms of storage interface used in hosts.
- SCSI is a set of standards used for connecting a peripheral device to a computer and transferring data between them.
- Often, SCSI is used to connect HDDs and tapes to a host.
- SCSI can also connect a wide variety of other devices such as scanners and printers.
- Communication between the hosts and the storage devices uses the SCSI command set.
- The oldest SCSI variant, called SCSI-1 provided data transfer rate of 5 MB/s; SCSI Ultra 320 provides data transfer speeds of 320 MB/s.

Table 5-1: Comparison of IDE/ATA with SCSI

FEATURE	IDE/ATA	SCSI
Speed	100, 133, 150 MB/s	320 MB/s
Connectivity	Internal	Internal and external
Cost	Low	Moderate to high
Hot-pluggable	No	Yes
Performance	Moderate to low	High
Ease of configuration	High	Low to moderate
Maximum number of devices supported	2	16

SCSI - 1

- SCSI-1, renamed to distinguish it from other SCSI versions, is the original standard that the ANSI approved.
- SCSI-1 defined the basics of the first SCSI bus, including cable length, signaling characteristics, commands, and transfer modes.
- SCSI-1 devices supported only single-ended transmission and passive termination. SCSI-1 used a narrow 8-bit bus, which offered a maximum data transfer rate of 5 MB/s.
- SCSI-1 implementations resulted in incompatible devices and several subsets of standards.

SCSI – 2

- To control the various problems caused by the nonstandard implementation of the original SCSI, a working paper was created to define a set of standard commands for a SCSI device.
- The set of standards, called the common command set (CCS), formed the basis of the SCSI-2 standard.
- SCSI-2 was focused on improving performance, enhancing reliability, and adding additional features to the SCSI-1 interface, in addition to standardizing and formalizing the SCSI commands.
- The transition from SCSI-1 to SCSI-2 did not raise much concern because SCSI-2 offered backward compatibility with SCSI-1.

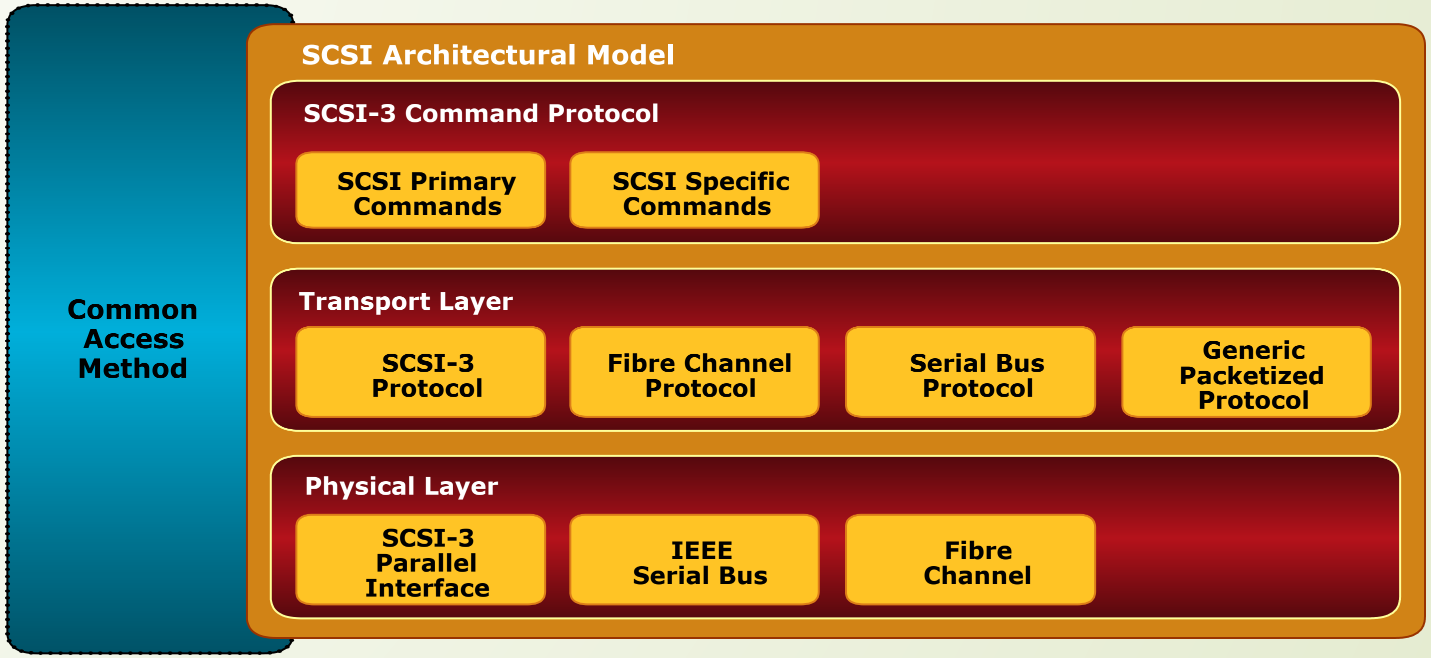
SCSI – 3

- In 1993, work began on developing the next version of the SCSI standard, SCSI-3.
- Unlike SCSI-2, the SCSI-3 standard document is comprised different but related standards, rather than one large document.

SCSI – 3 Architecture

- The SCSI-3 architecture defines and categorizes various SCSI-3 standards and requirements for SCSI-3 implementations.
- This architecture helps developers, hardware designers, and users to understand and effectively utilize SCSI.
- The three major components of a SCSI architectural model are as follows:
  - SCSI-3 command protocol: This consists of primary commands that are common to all devices as well as device-specific commands that are unique to a given class of devices.
  - Transport layer protocols: These are a standard set of rules by which devices communicate and share information.
  - Physical layer interconnects: These are interface details such as electrical signaling methods and data transfer modes.

SCSI–3 Architecture



- SCSI command protocol
  - Primary commands common to all devices
- Transport layer protocol
  - Standard rules for device communication and information sharing
- Physical layer interconnect
  - Interface details such as electrical signaling methods and data transfer modes

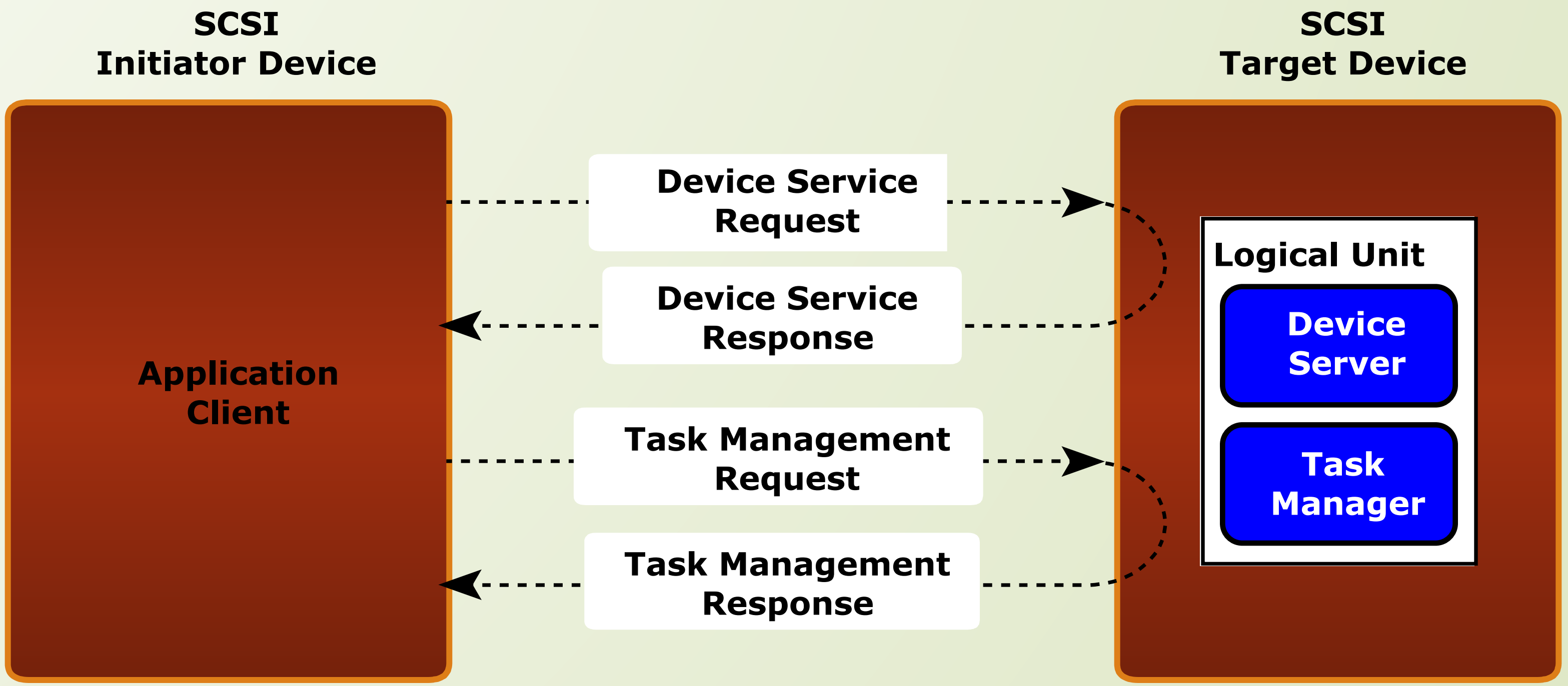
SCSI-3 client server model

- SCSI-3 architecture derives its base from the client-server relationship, in which a client directs a service request to a server, which then fulfills the client's request.
- In a SCSI-3 client-server model, a particular SCSI device acts as a SCSI target device, a SCSI initiator device, or a SCSI target/initiator device.
- Each device performs the following functions:
- **SCSI initiator device:** Issues a command to the SCSI target device, to perform a task. A SCSI host adaptor is an example of an initiator.
- acts as a target device. However, in certain implementations, the host adaptor can also be a**SCSI target device:** Executes commands to perform the task received from a SCSI initiator. Typically a SCSI peripheral device target device.

SCSI Device Model / client server model

SCSI communication involves:

- SCSI initiator device
  - Issues commands to SCSI target devices
  - Example: SCSI host adaptor
- SCSI target device
  - Executes commands issued by initiators
  - Examples: SCSI peripheral devices



SCSI-3 Device Model / client server model

- The SCSI initiator device is comprised of an application client and task management function, which initiates device service and task management requests.
- Each device service request contains a Command Descriptor Block (CDB).
- The CDB defines the command to be executed and lists command-specific input sand other parameters specifying how to process the command.

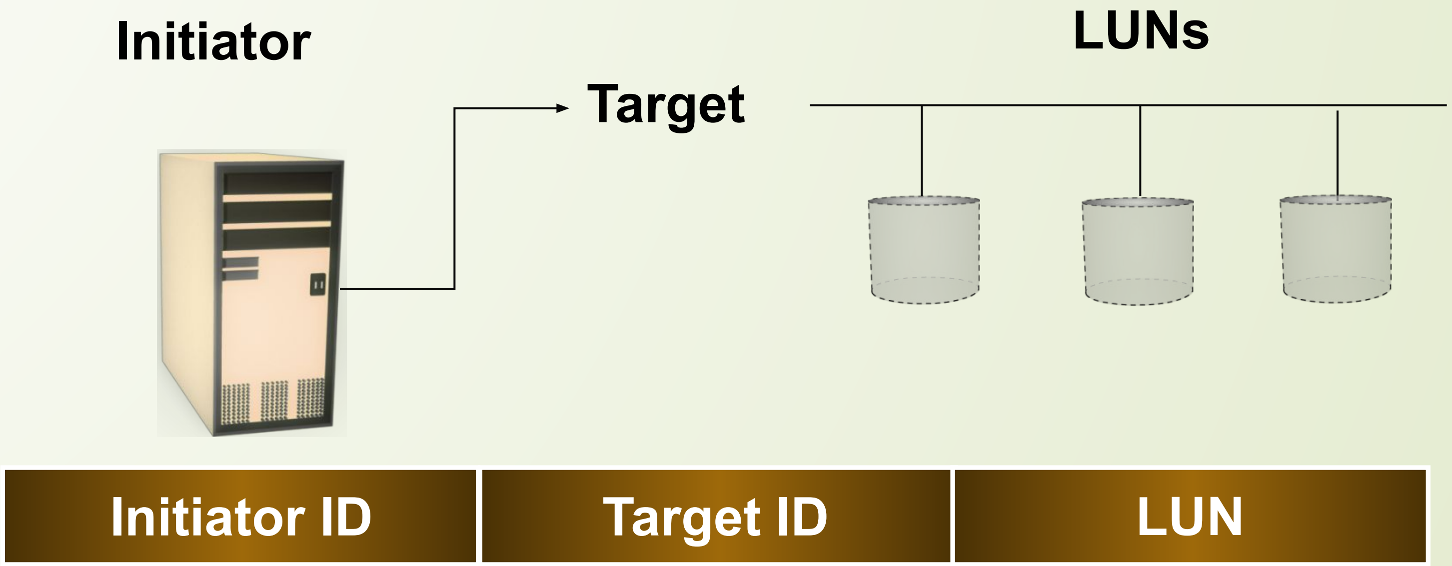
SCSI-3 Device Model / client server model

- The SCSI devices are identified by a specific number called a SCSI ID.
- In narrow SCSI (bus width=8), the devices are numbered 0 through 7; in wide (bus width=16) SCSI, the devices are numbered 0 through 15.
- These ID numbers set the device priorities on the SCSI bus.
- In narrow SCSI, 7 has the highest priority and 0 has the lowest priority.
- In wide SCSI, the device IDs from 8 to 15 have the highest priority, but the entire sequence of wide SCSI IDs has lower priority than narrow SCSI IDs.
- Therefore, the overall priority sequence for a wide SCSI is 7, 6, 5, 4, 3, 2, 1, 0, 15, 14, 13, 12, 11, 10, 9, and 8.

SCSI Device Model / client server model

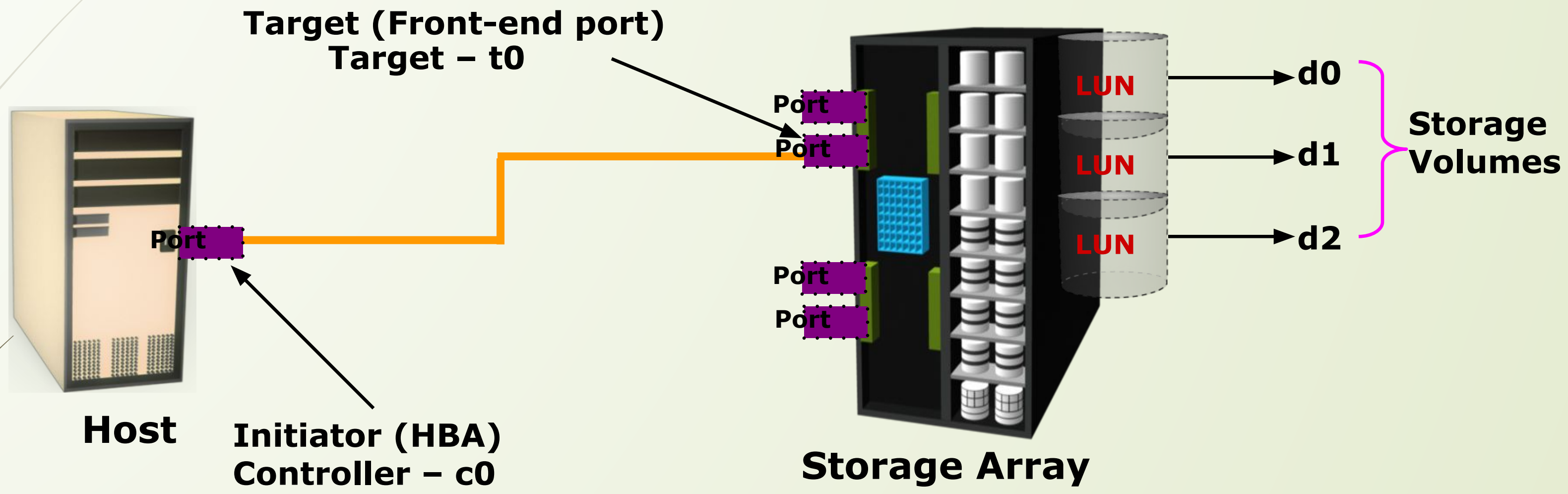
- Device requests uses Command Descriptor Block (CDB)
  - 8 bit structure
  - Contain operation code, command specific parameter and control parameter
- SCSI Ports
  - SCSI device may contain initiator port, target port, target/initiator port
  - Based on the port combination, a SCSI device can be classified as an initiator model, a target model, a target model with multiple ports or a combined model (target/initiator model). Example:
    - Target/initiator device contain target/initiator port and can switch orientations depending on the role it plays while participating in an I/O operation
  - To cater to service requests from multiple devices, a SCSI device may also have multiple ports (e.g. target model with multiple ports)

SCSI Addressing



- Initiator ID - a number from 0 to 15 with the most common value being 7.
- Target ID - a number from 0 to 15
- LUN - a number that specifies a device addressable through a target.

SCSI Addressing Example



Host Addressing:
Storage Volume 1 - c0t0d0
Storage Volume 2 - c0t0d1
Storage Volume 3 - c0t0d2

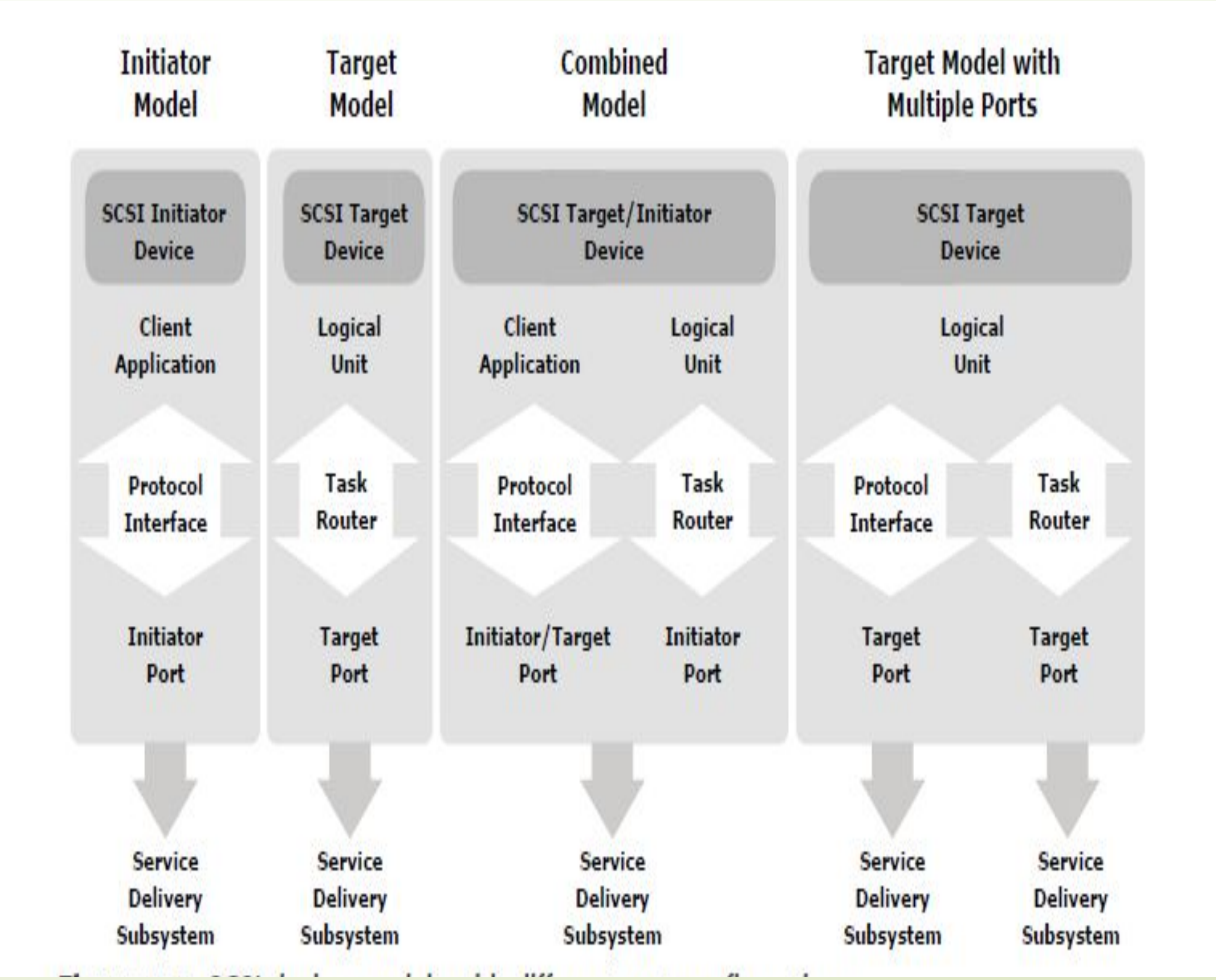
Initiator ID	Target ID	LUN
c0	t0	d0

Table 5-2: SCSI Interfaces

INTERFACE	STANDARD	BUS WIDTH	CLOCK SPEED	MAX THROUGHPUT	MAX DEVICES
SCSI-1	SCSI-1	8	5 MHz	5 MB/s	8
Fast SCSI	SCSI-2	8	10 MHz	10 MB/s	8
Fast Wide SCSI	SCSI-2; SCSI-3 SPI	16	10 MHz	20 MB/s	16
Ultra SCSI	SCSI-3 SPI	8	20 MHz	20 MB/s	8
Ultra Wide SCSI	SCSI-3 SPI	16	20 MHz	40 MB/s	16
Ultra2 SCSI	SCSI-3 SPI-2	8	40 MHz	40 MB/s	8
Ultra2 Wide SCSI	SCSI-3 SPI-2	16	40 MHz	80 MB/s	16
Ultra3 SCSI	SCSI-3 SPI-3	16	40 MHz DDR	160 MB/s	16
Ultra320 SCSI	SCSI-3 SPI-4	16	80 MHz DDR	320 MB/s	16
Ultra640 SCSI	SCSI-3 SPI-5	16	160 MHz DDR	640 MB/s	16

SCSI - Ports

- SCSI ports are the physical connectors that the SCSI cable plugs into for communication with a SCSI device.
- A SCSI device may contain target ports, initiator ports, target/initiator ports, or a target with multiple ports. Based on the port combinations, a SCSI device can be classified as an initiator model, a target model, a combined model, or a target model with multiple ports.



Fibre Channel
Storage Area Networks (SAN)

EMC Global Education

Module 3.3

Lesson: Fibre Channel SAN Overview

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

- Define a FC SAN.
- Describe the features of FC SAN based storage.
- Describe the benefits of an FC SAN based storage strategy.

Business Needs and Technology Challenges

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.

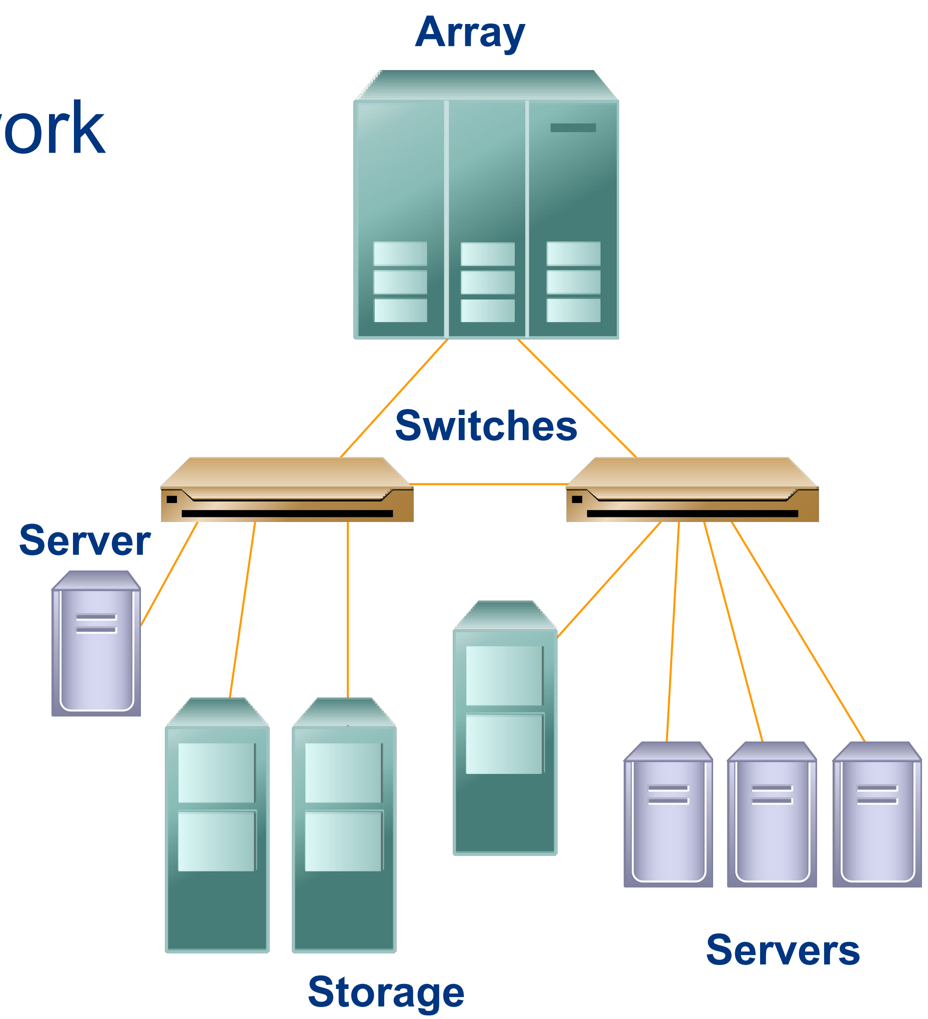
An effective information management solution must provide the following:

- **Just-in-time information to business users**
- **Integration of information infrastructure with business processes**
- **Flexible and resilient storage architecture**

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

What is a SAN?

- Dedicated storage network
- Organized connections among:
  - Storage
  - Communication devices
  - Systems
- Secure
- Robust



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.

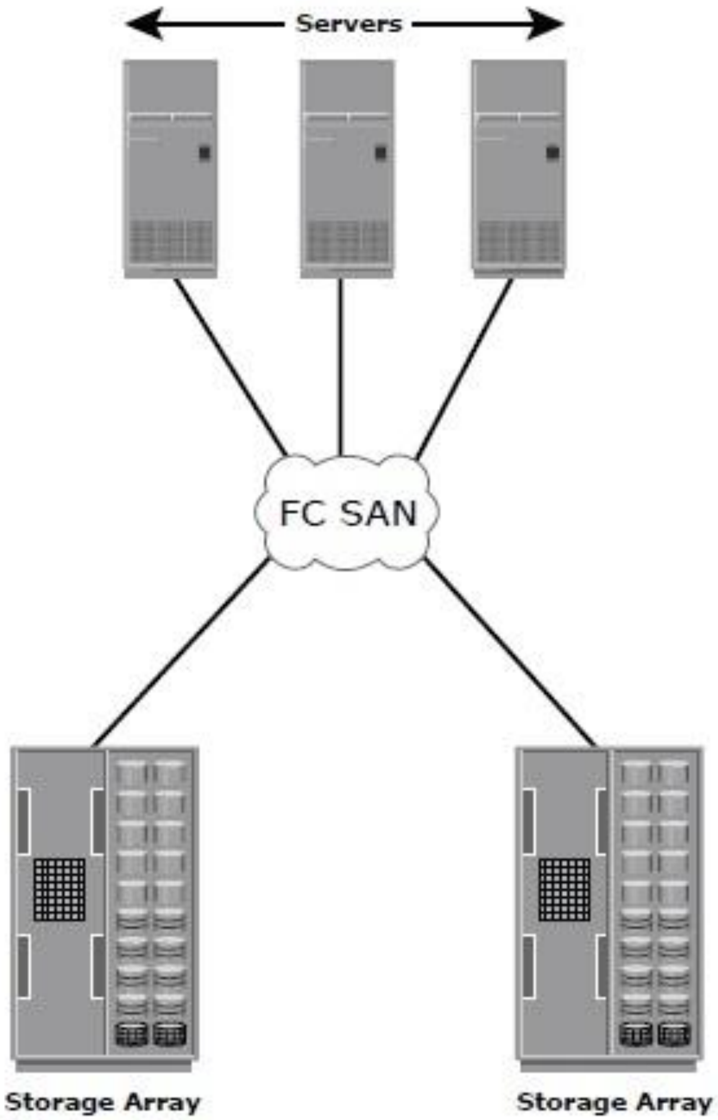


Figure 6-1: SAN implementation

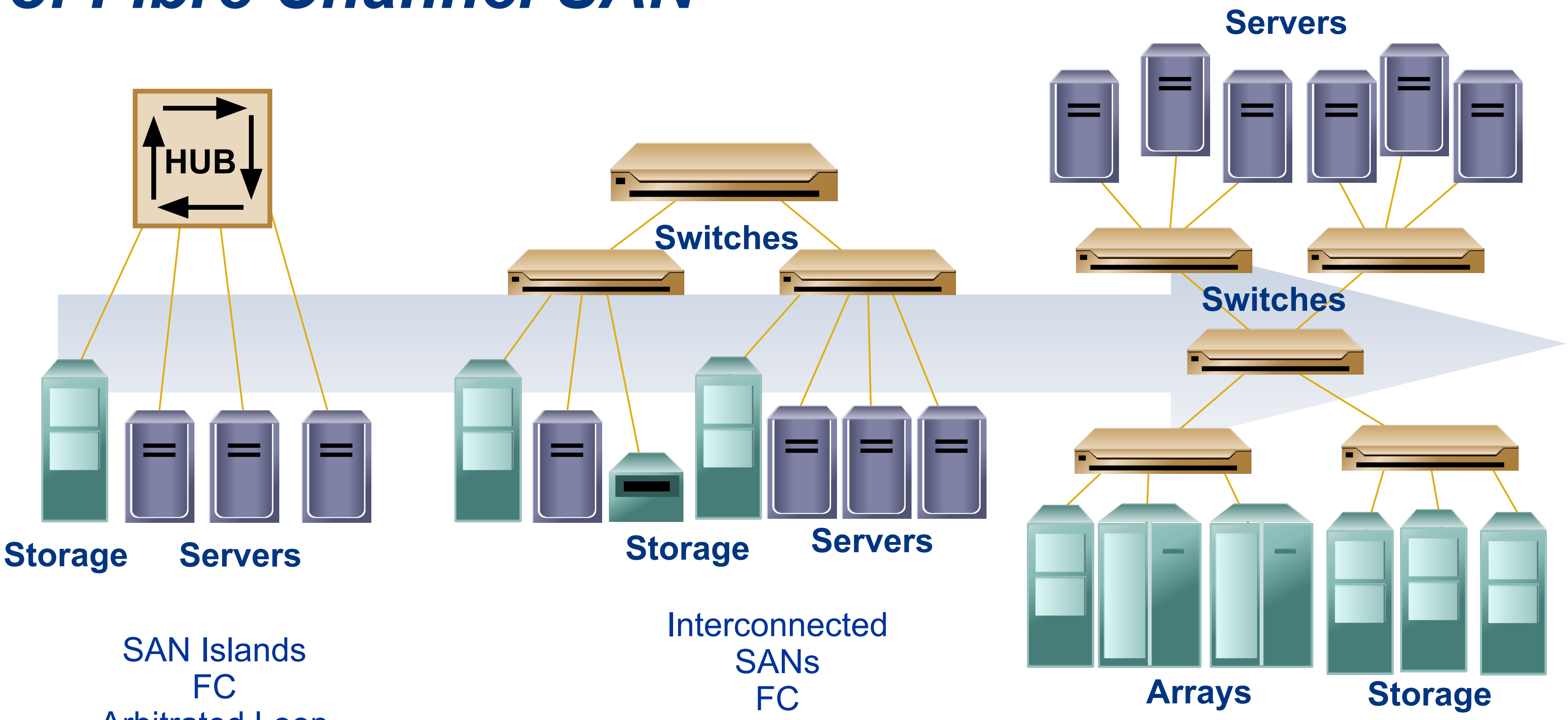
Components of SAN

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**
- **Interconnect Devices**
- **Storage Arrays**
- **SAN Management Software**

Evolution of Fibre Channel SAN



SAN Islands
FC
Arbitrated Loop

Interconnected
SANs
FC
Switched Fabric

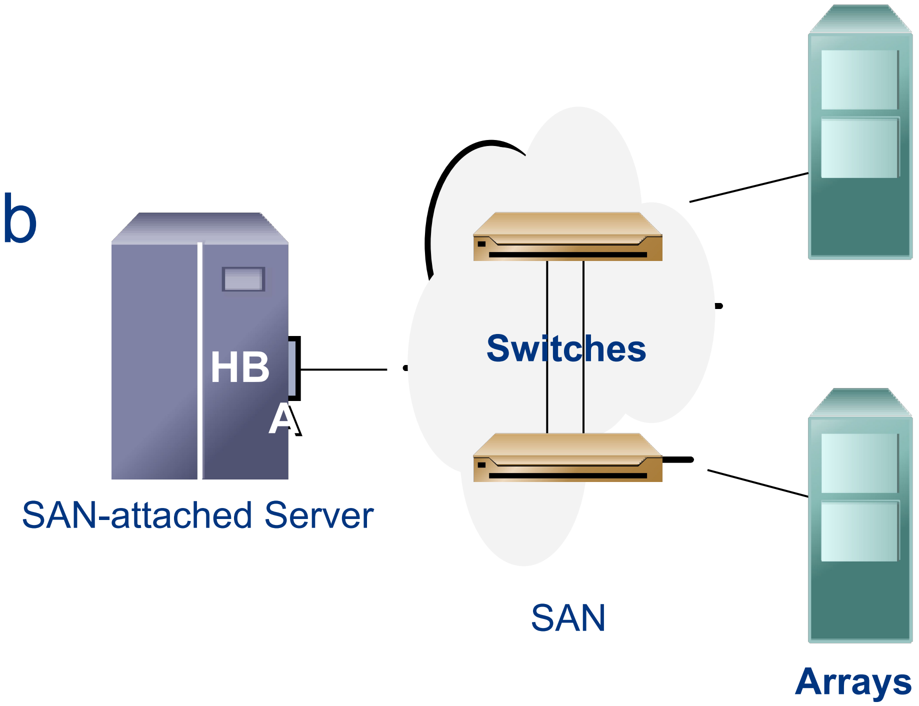
Enterprise SANs
FC Switched Fabric

Benefits of a SAN

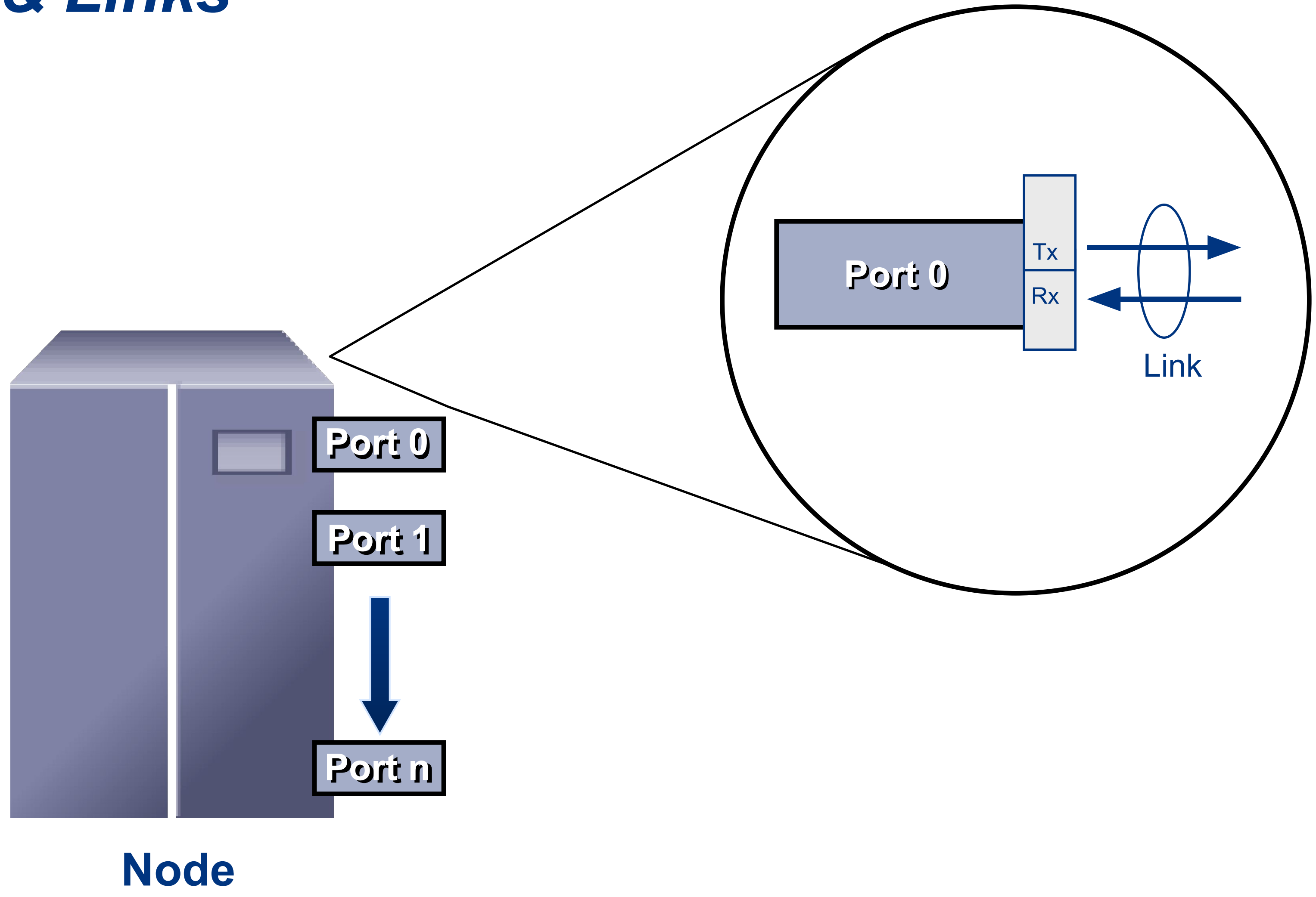
- High bandwidth
  - **Fibre Channel**
- SCSI extension
  - **Block I/O**
- Resource Consolidation
  - **Centralized storage and management**
- Scalability
  - **Up to 16 million devices**
- Secure Access
  - **Isolation and filtering**

Components of a Storage Area Network

- Host Bus Adapter (HBA)
- Fiber Cabling
- Fibre Channel Switch /Hub
- Storage Array
- Management System



Nodes, Ports, & Links



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 (62.5µm), OM2 (50µm) and laser optimized OM3 (50µm).

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)).

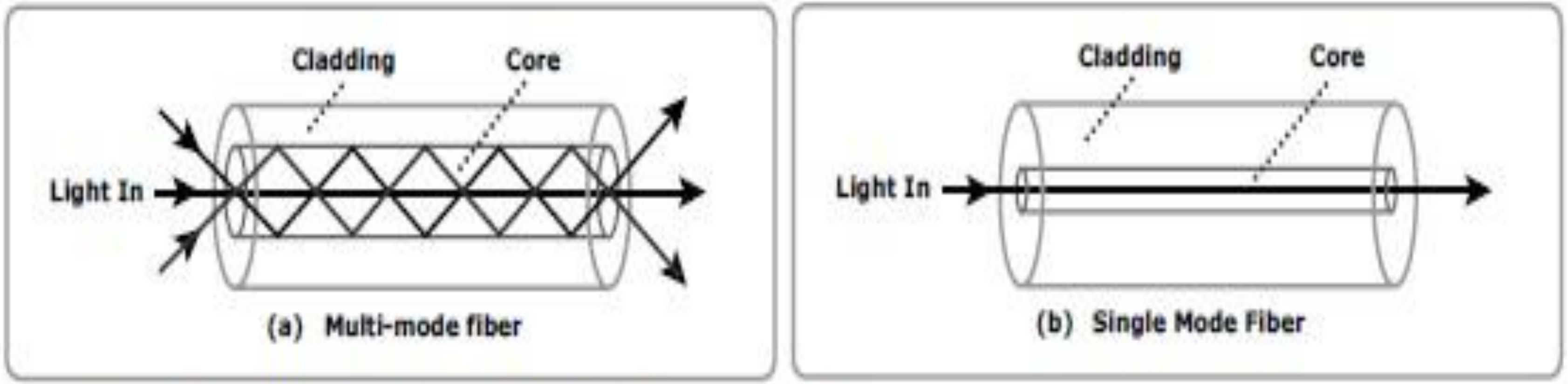
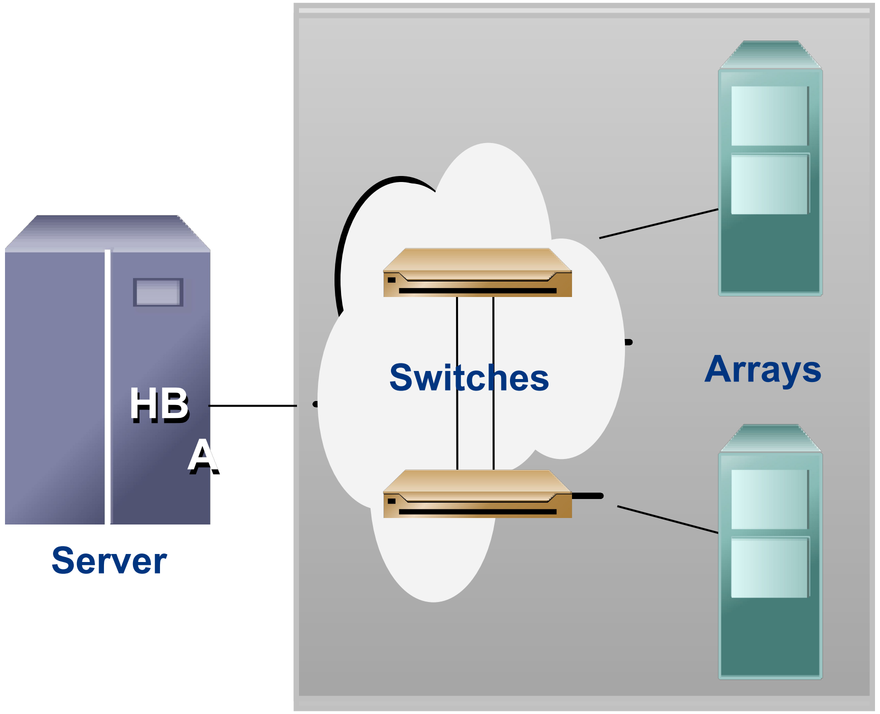


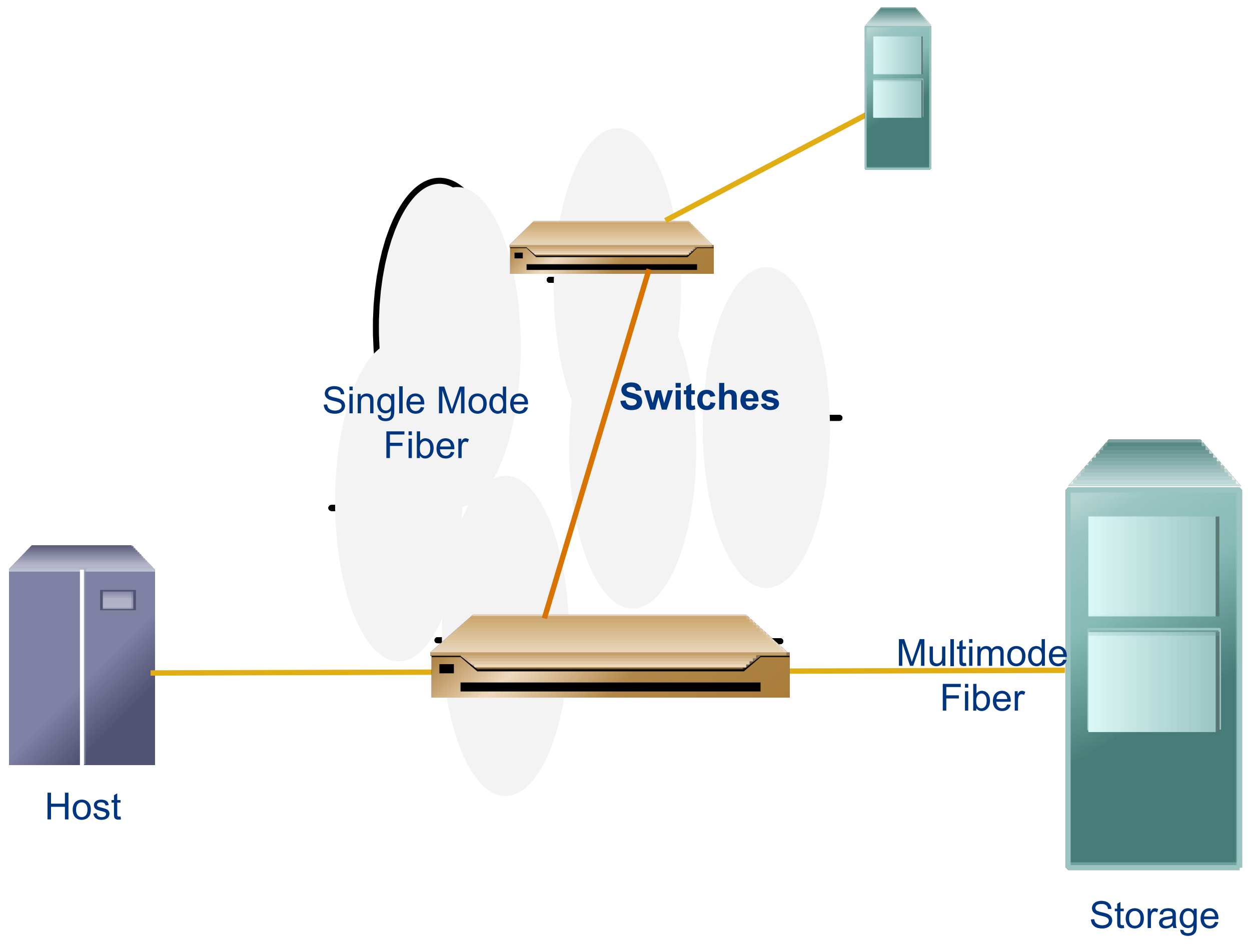
Figure 6-4: Multi-mode fiber and single-mode fiber

Host Bus Adapters

- HBAs perform low-level interface functions automatically to minimize the impact on host processor performance



Connectivity



Connectors

Node Connectors:

- SC Duplex Connectors
- LC Duplex Connectors

Patch panel Connectors

- ST Simplex Connectors

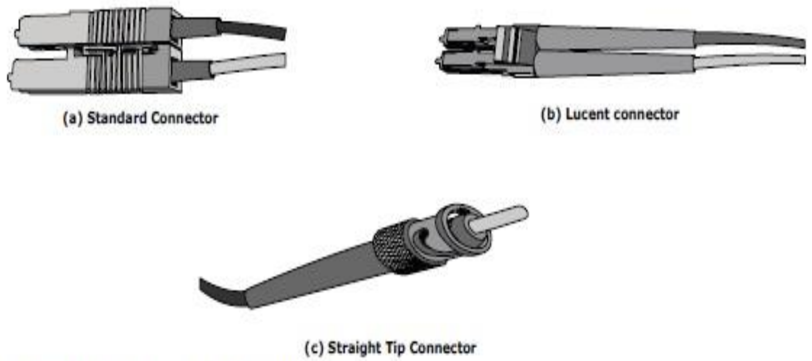


Figure 6-5: SC, LC, and ST connectors

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.

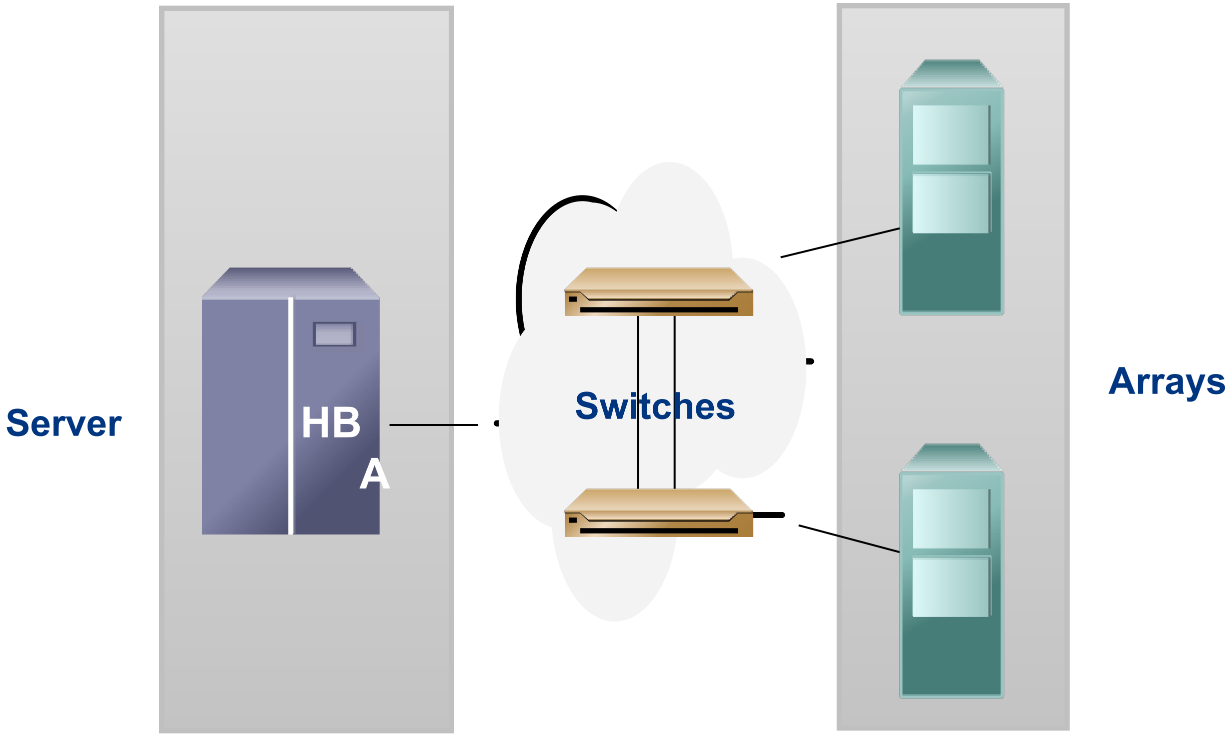
Because of availability of low cost and high performance switches, hubs are no longer used in SANs.

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.

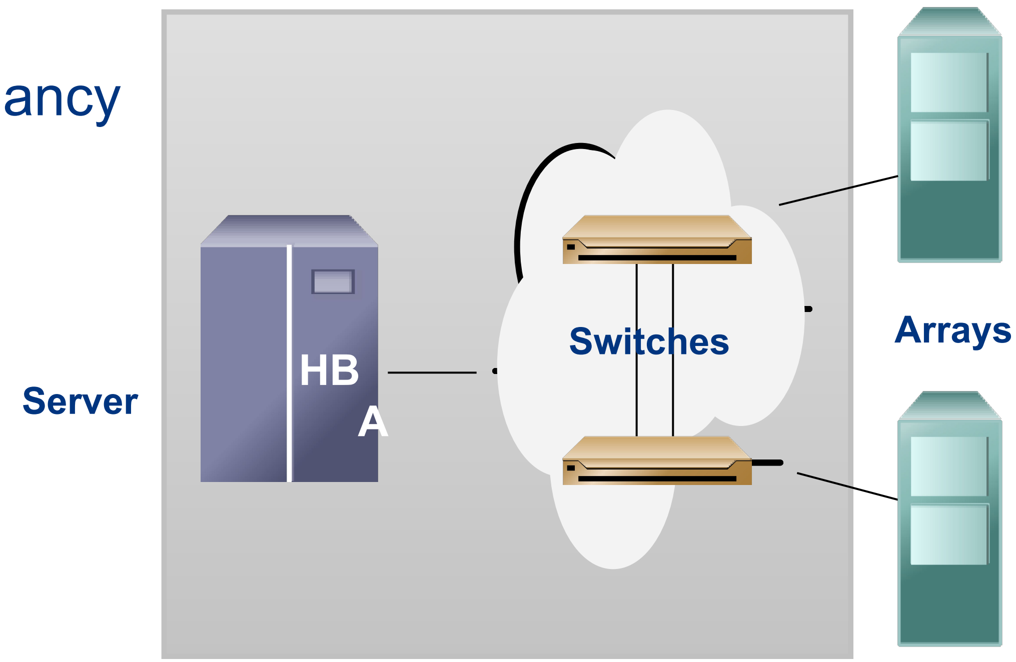
Connectivity Devices

- Basis for SAN communication
  - Hubs, Switches and Directors



Storage Resources

- Storage Array
  - Provides storage consolidation and centralization
- Features of an array
  - High Availability/Redundancy
  - Performance
  - Business Continuity
  - Multiple host connect



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.

SAN Management Software

SAN management software manages the interfaces between hosts, interconnect devices, and storage arrays.

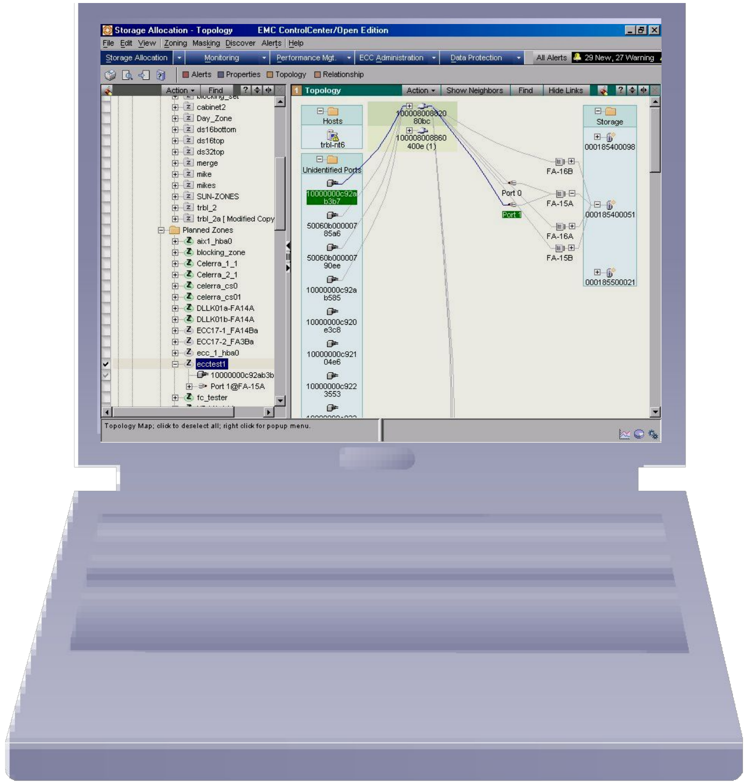
The software provides a view of the SAN environment and enables management of various resources from one central console.

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.

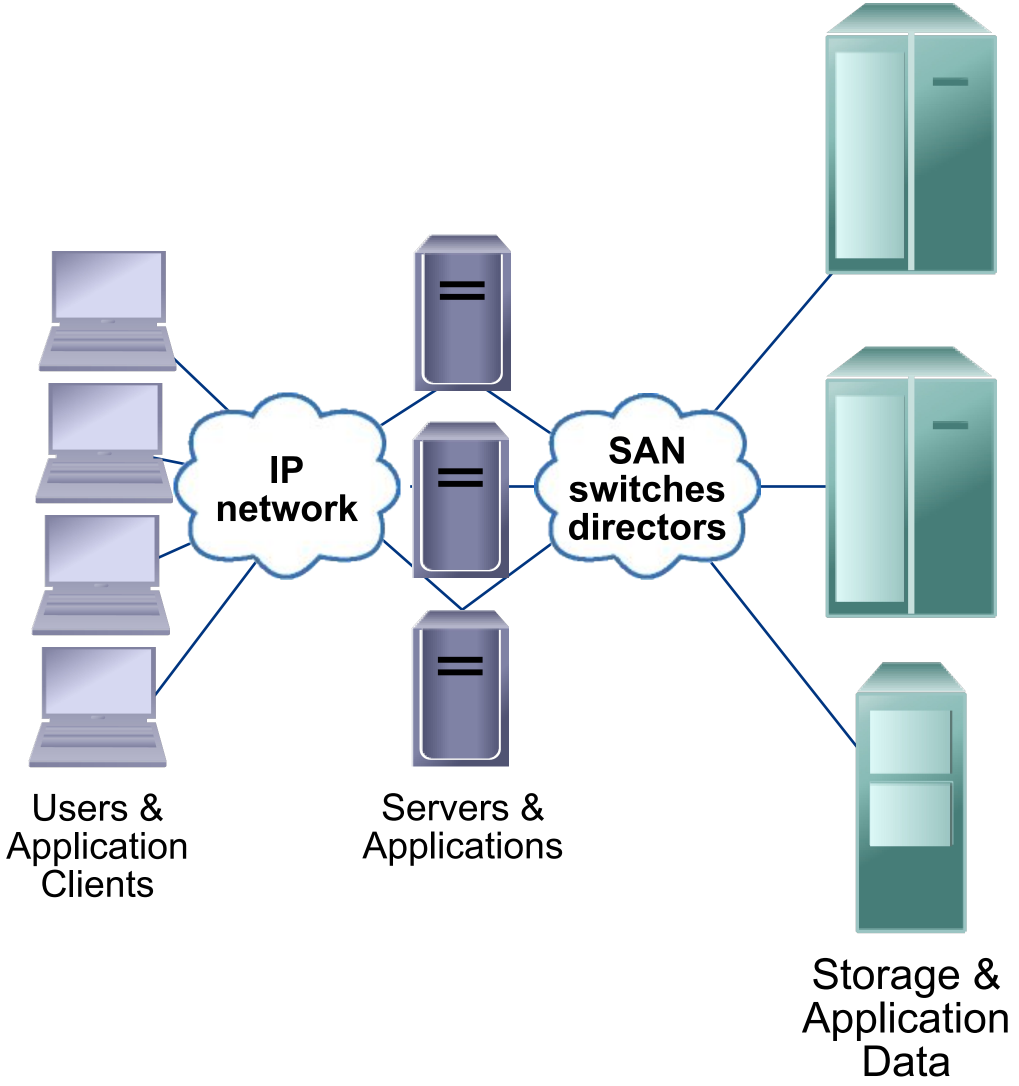
SAN Management Software

- A suite of tools used in a SAN to manage the interface between host and storage arrays.
- Provides integrated management of SAN environment.
- Web based GUI or CLI



Fibre Channel SAN Connectivity

- Core networking principles applied to storage
- Servers are attached to 2 distinct networks
  - Back-end
  - Front-end



What is Fibre Channel?

- SAN Transport Protocol
  - Integrated set of standards (ANSI)
  - Encapsulates SCSI
- A High Speed Serial Interface
  - Allows SCSI commands to be transferred over a storage network.
- Standard allows for multiple protocols over a single interface.

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.

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).

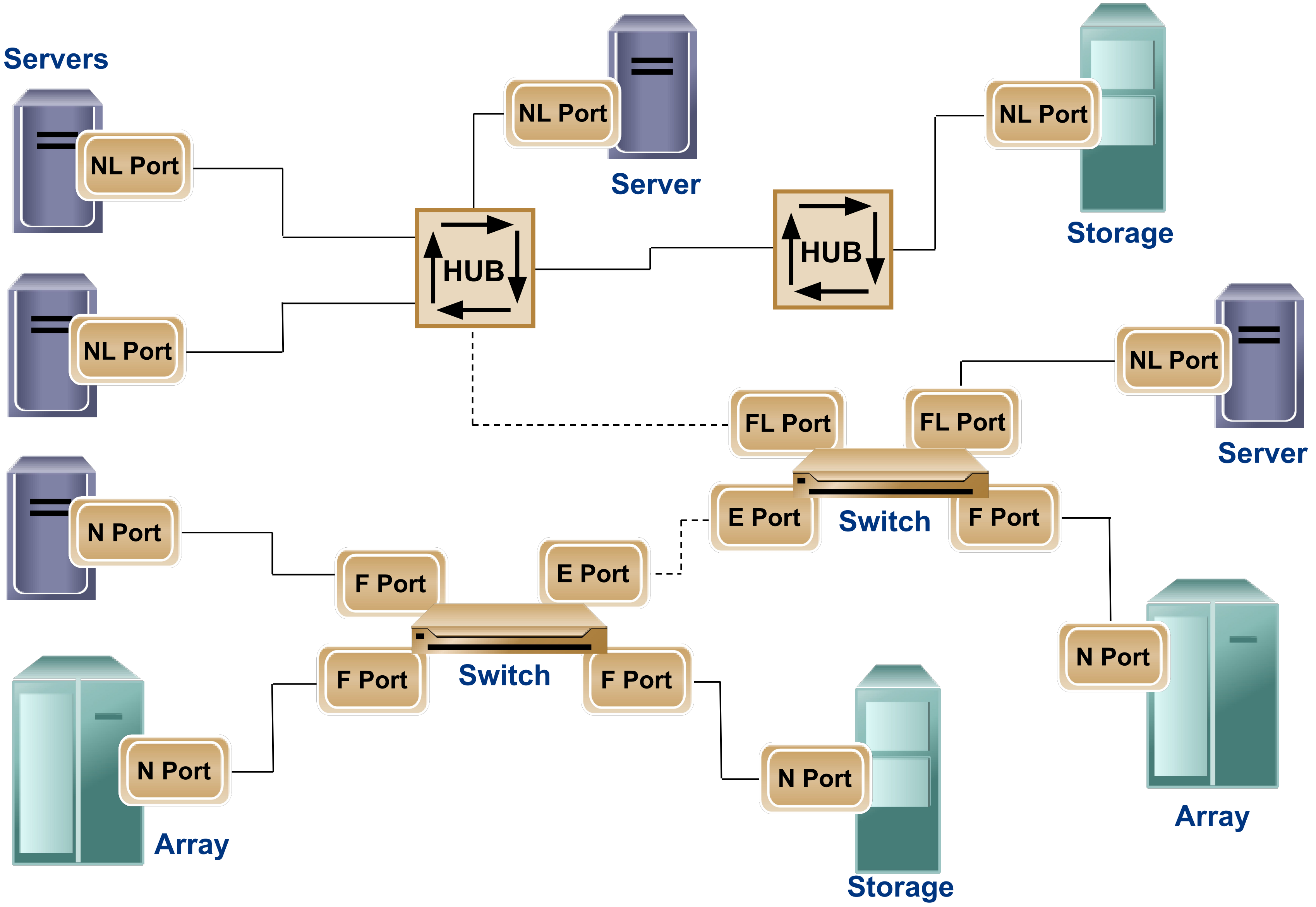
Fibre Channel Ports

Ports are the basic building blocks of an FC network.

Ports on the switch can be one of the following types:

- **N_port**
- **NL_port**
- **E_port**
- **F_port**
- **FL_port**
- **G_port**

Fibre Channel Ports



World Wide Names

- Unique 64 bit identifier.
- Static to the port.
  - Used to physically identify a port or node within the SAN
  - Similar to a NIC MAC address
- Additionally, each node is assigned a unique port ID (address) within the SAN
  - Used to communicate between nodes within the SAN
  - Similar in functionality to an IP address on a NIC

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.

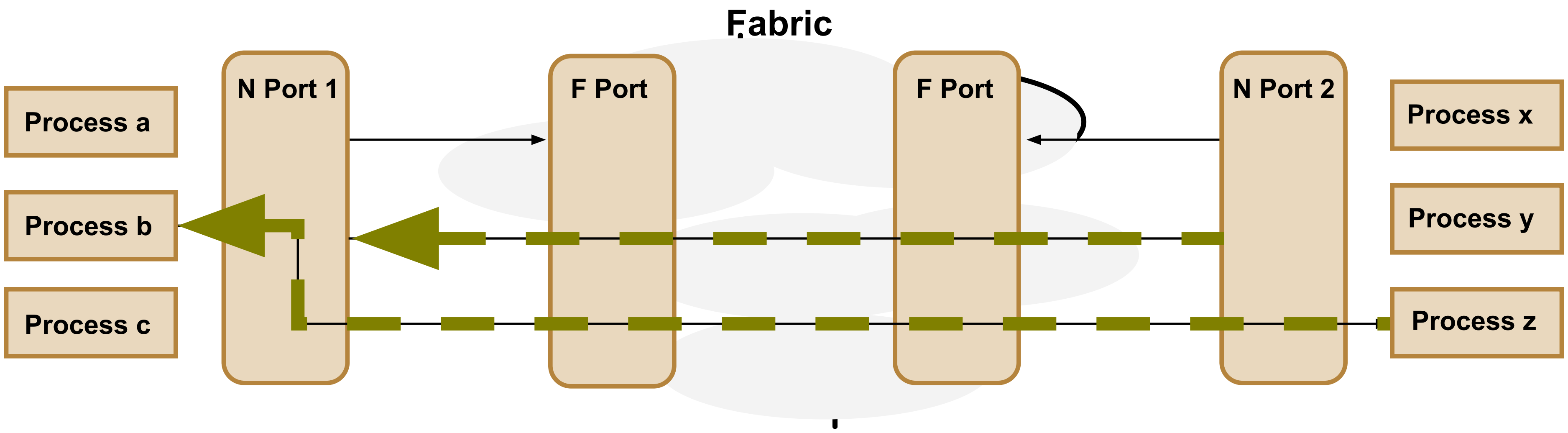
Figure 6-16 illustrates the WWN structure for an array and the HBA.

World Wide Names: Example

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 OUI 24 bits						Company Specific 24 bits					

Fibre Channel Logins

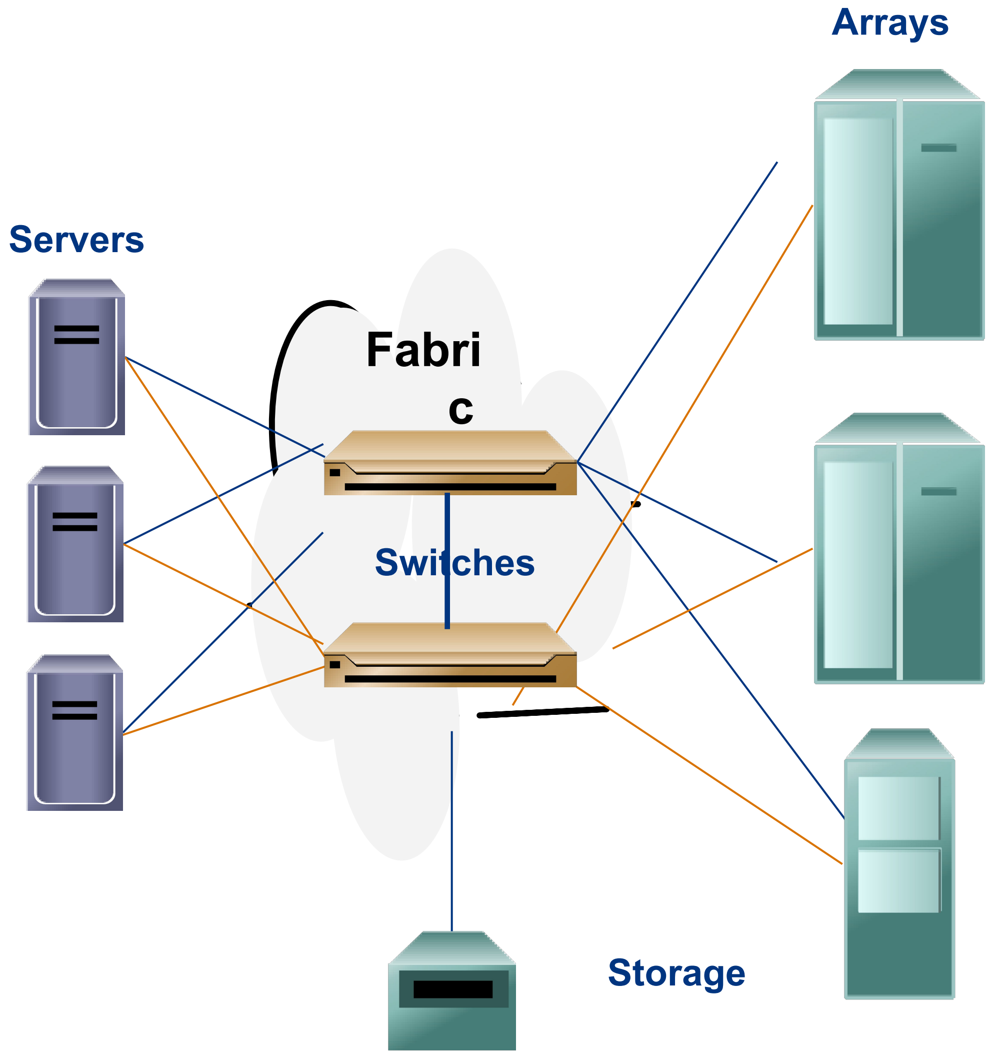


Fibre Channel Addressing

- Fibre Channel addresses are used for transporting frames from source ports to destination ports.
- Address assignment methods vary with the associated topology (loop vs switch)
  - Loop – self assigning
  - Switch – centralized authority
- Certain addresses are reserved
  - FFFFFC is Name Server
  - FFFFFE is Fabric Login

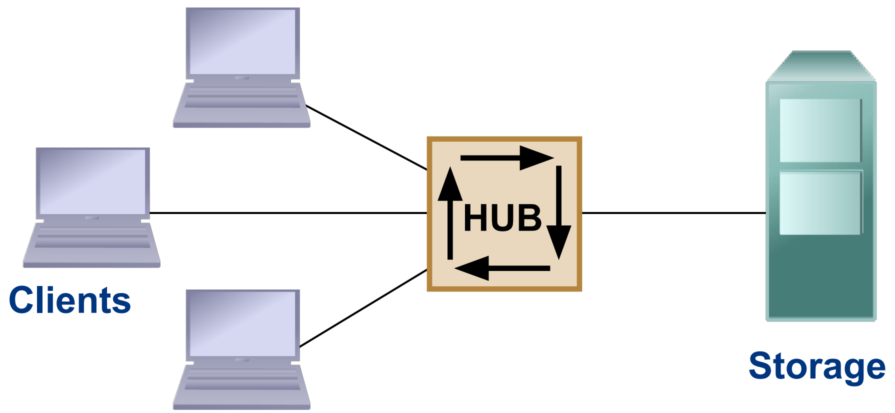
What is a Fabric?

- Virtual space used by nodes to communicate with each other once they are joined.
- Component identifiers:
  - Domain ID
  - Worldwide Name (WWN)

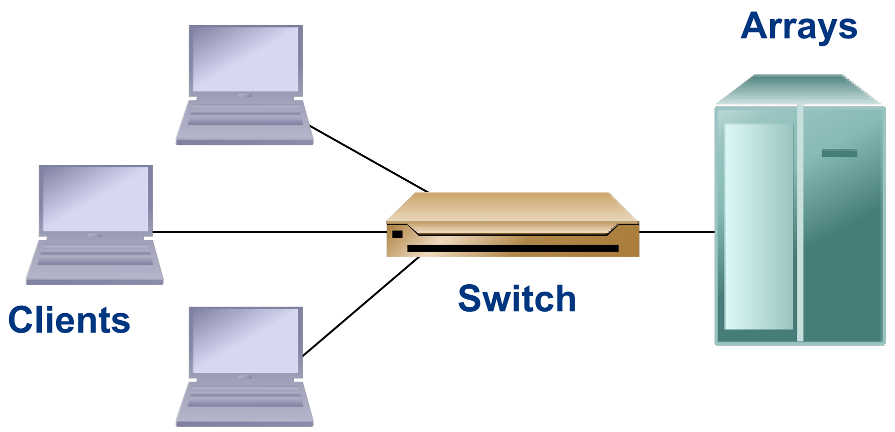


Fibre Channel Topologies

- Arbitrated Loop (FC-AL)
  - Devices attached to a shared "loop"
  - Analogous to Token Ring



- Switched Fabric (FC-SW)
  - All devices connected to a "Fabric Switch" – Analogous to an IP switch
  - Initiators have unique dedicated I/O paths to Targets

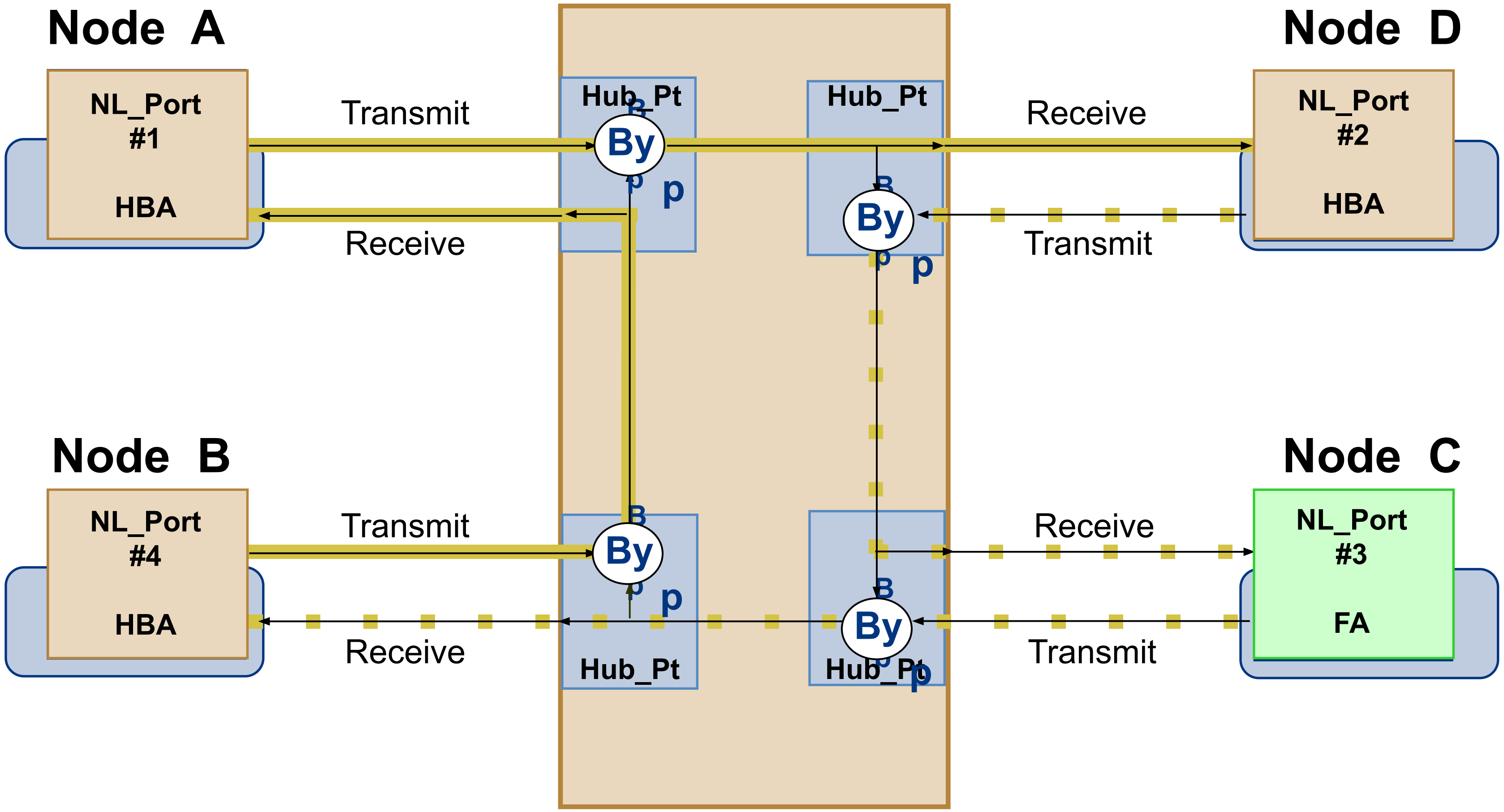


Switch versus Hub Comparison

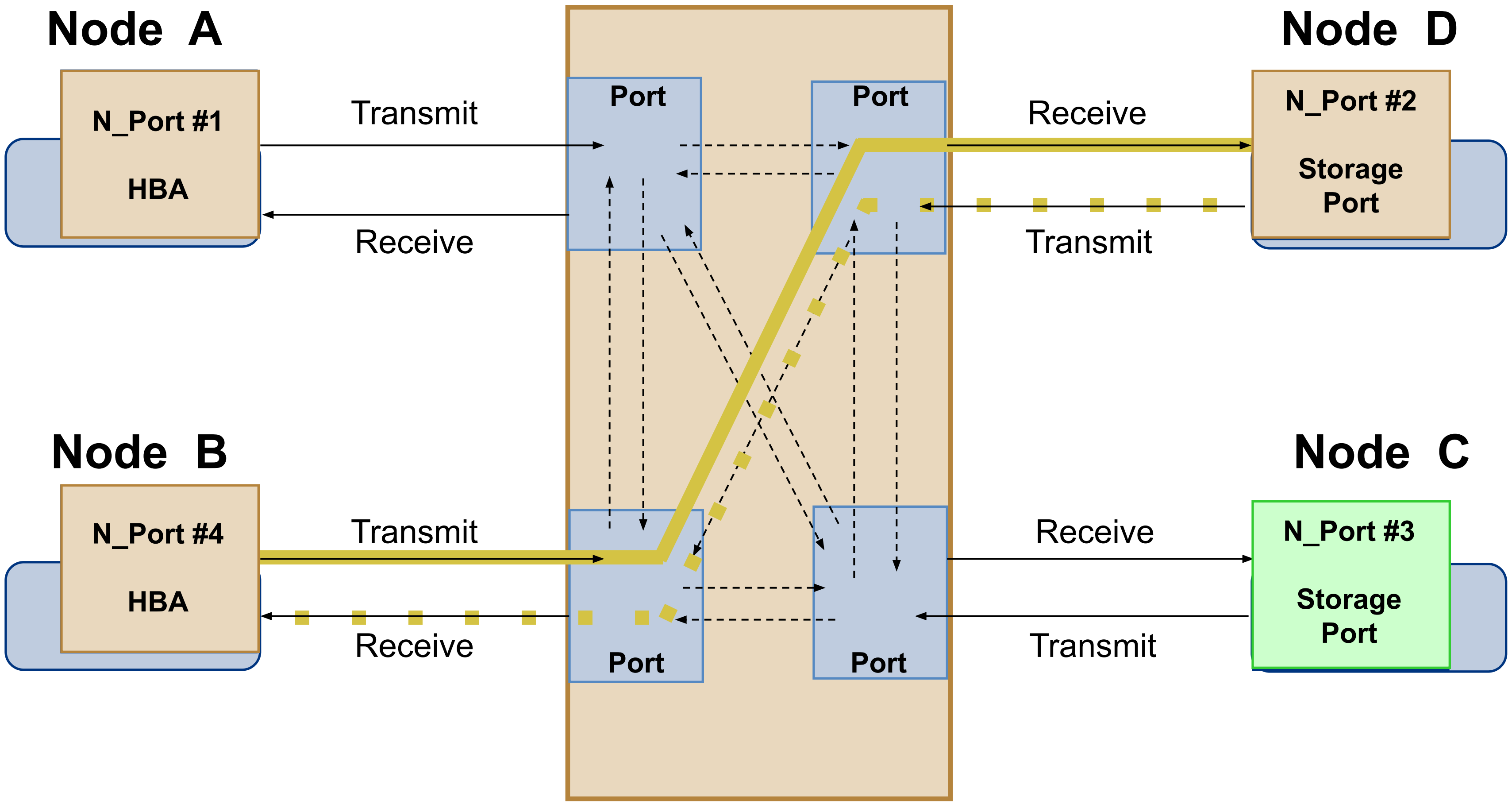
- Switches (FC-SW)
  - FC-SW architecture scalable to millions of connections.
  - Bandwidth per device stays constant with increased connectivity.
  - Bandwidth is scalable due to dedicated connections.
  - Higher availability than hubs.
  - Higher cost.

- Hubs (FC-AL)
  - FC-AL is limited to 127 connections (substantially fewer connections can be implemented for ideal system performance).
  - Bandwidth per device diminishes with increased connectivity due to sharing of connections.
  - Low cost connection.

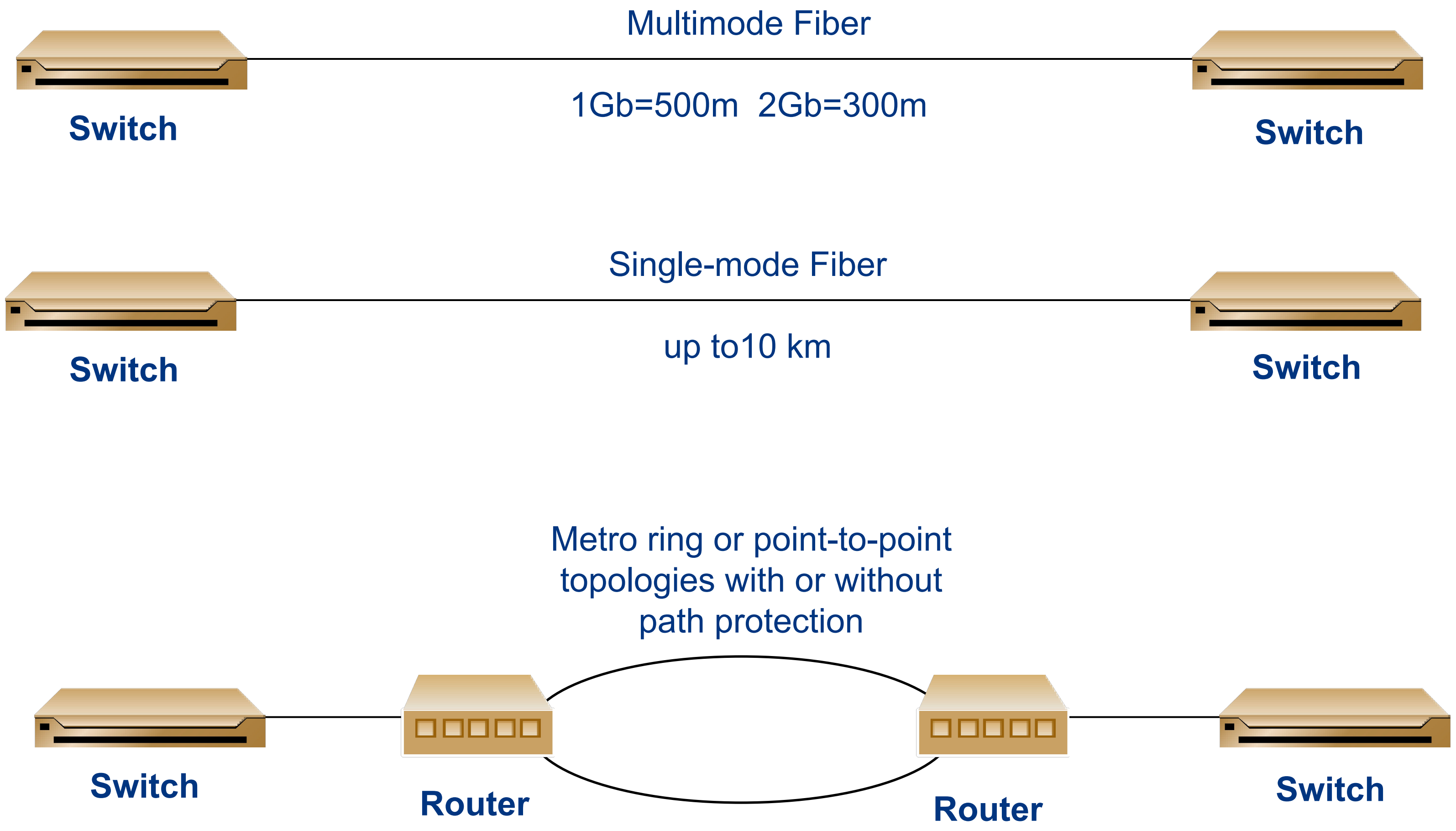
How an Arbitrated Loop Hub Works



How a Switched Fabric Works

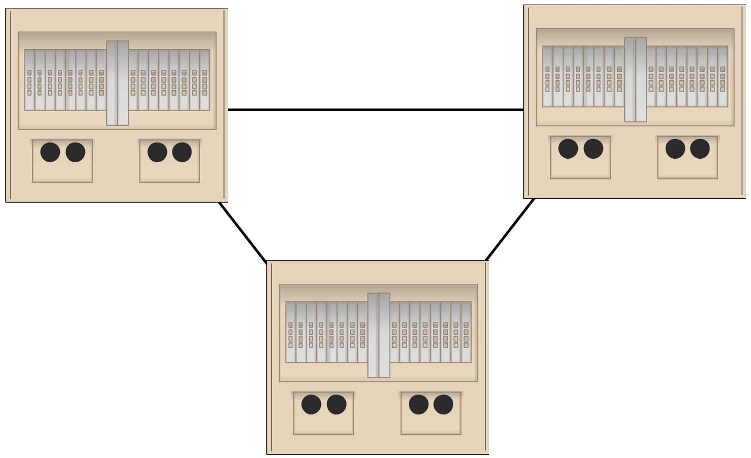


Inter Switch Links (ISLs)

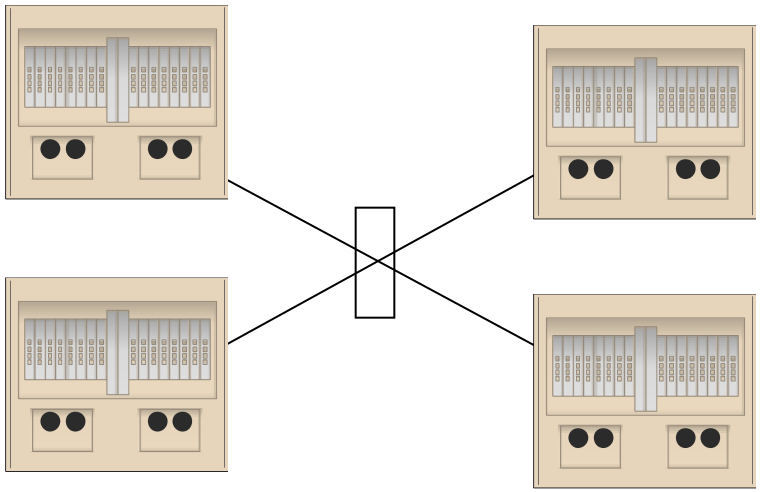


Topology: Mesh Fabric

- Can be either partial or full mesh
- All switches are connected to each other
- Host and Storage can be located anywhere in the fabric
- Host and Storage can be localized to a single switch



Partial Mesh



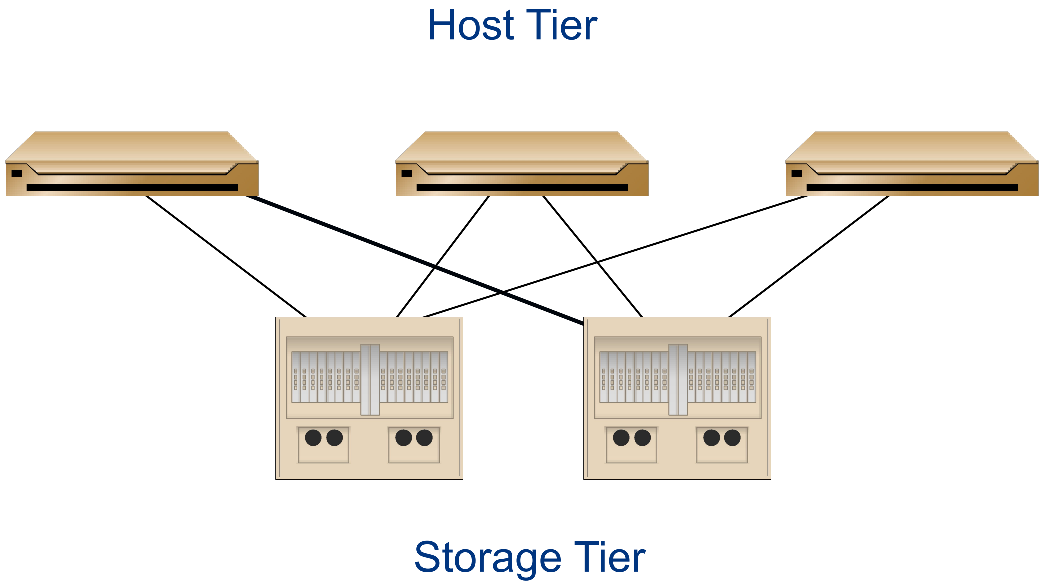
Full Mesh

Full Mesh Benefits

- Benefits
  - All storage/servers are a maximum of one ISL hop away.
  - Hosts and storage may be located anywhere in the fabric.
  - Multiple paths for data using the Fabric Shortest Path First (FSPF) algorithm.
  - Fabric management made simpler.

Topology: Simple Core-Edge Fabric

- Can be two or three tiers
  - 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



Core-Edge Benefits

- Simplifies propagation of fabric data.
  - One ISL hop access to all storage in the fabric.
- Efficient design based on node type.
  - Traffic management and predictability.
- Easier calculation of ISL loading and traffic patterns.

Lesson: Summary

Topics in this lesson included:

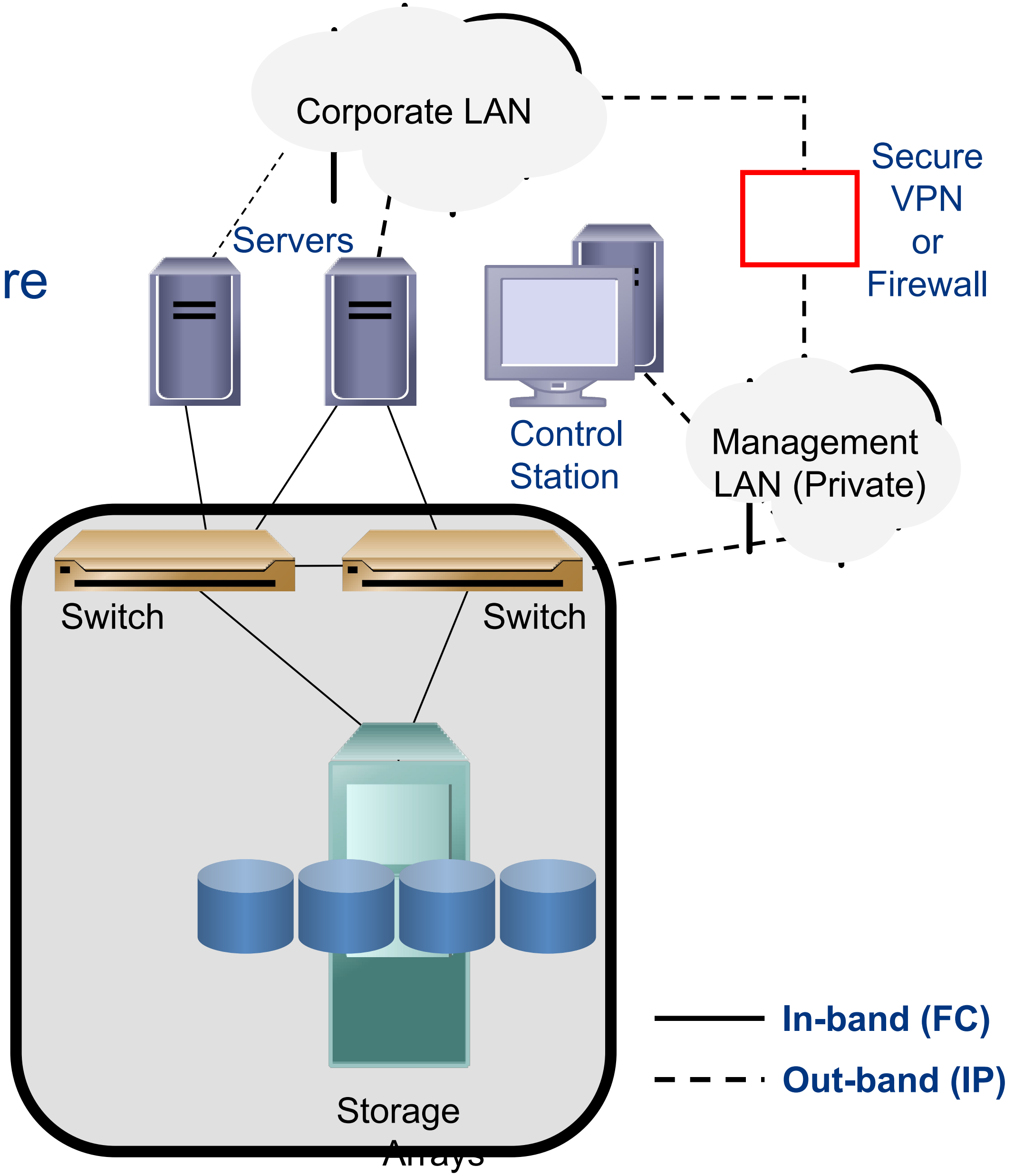
- The Fibre Channel SAN connectivity methods and topologies
- Fibre Channel devices
- Fibre Channel communication protocols
- Fibre Channel login procedures

SAN Management Overview

- Infrastructure protection
- Fabric Management
- Storage Allocation
- Capacity Tracking
- Performance Management

Infrastructure Security

- Physical security
  - Locked data center
- Centralized server and storage infrastructure
  - Controlled administrator access



Switch/Fabric Management Tools

- Vendor supplied management software
  - Embedded within the switch
  - Graphical User Interface (GUI) or Command Line Interface (CLI)
- Functionality
  - Common functions
    - Performance monitoring
    - Discovery
    - Access Management (Zoning)
  - Different “look and feel” between vendors
- Additional third party software add-ons
  - Enhanced functionality, such as automation

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.

Types of Zoning

Zoning can be categorized into three types:

- **Port zoning**
- **WWN zoning**
- **Mixed zoning**

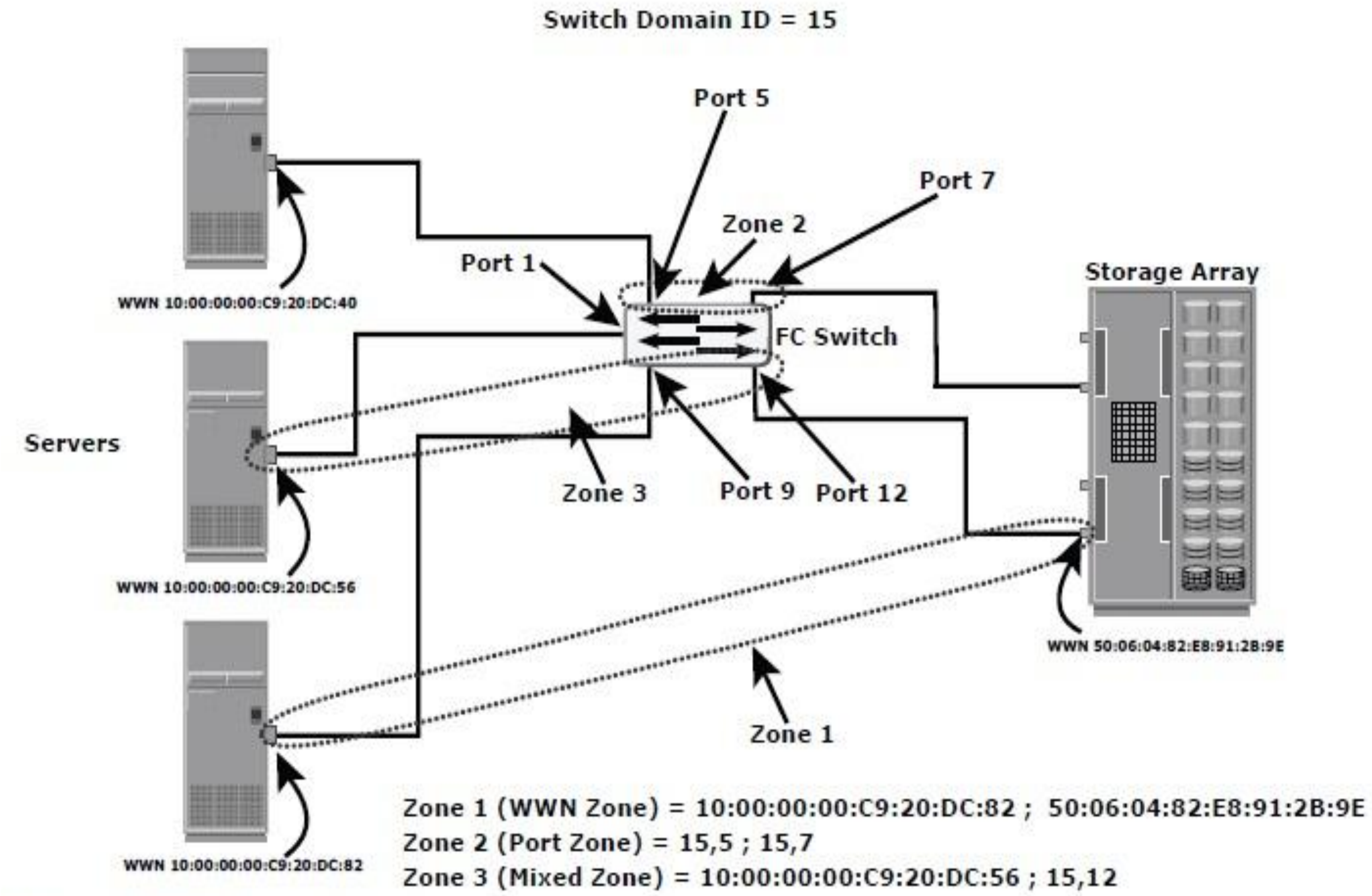


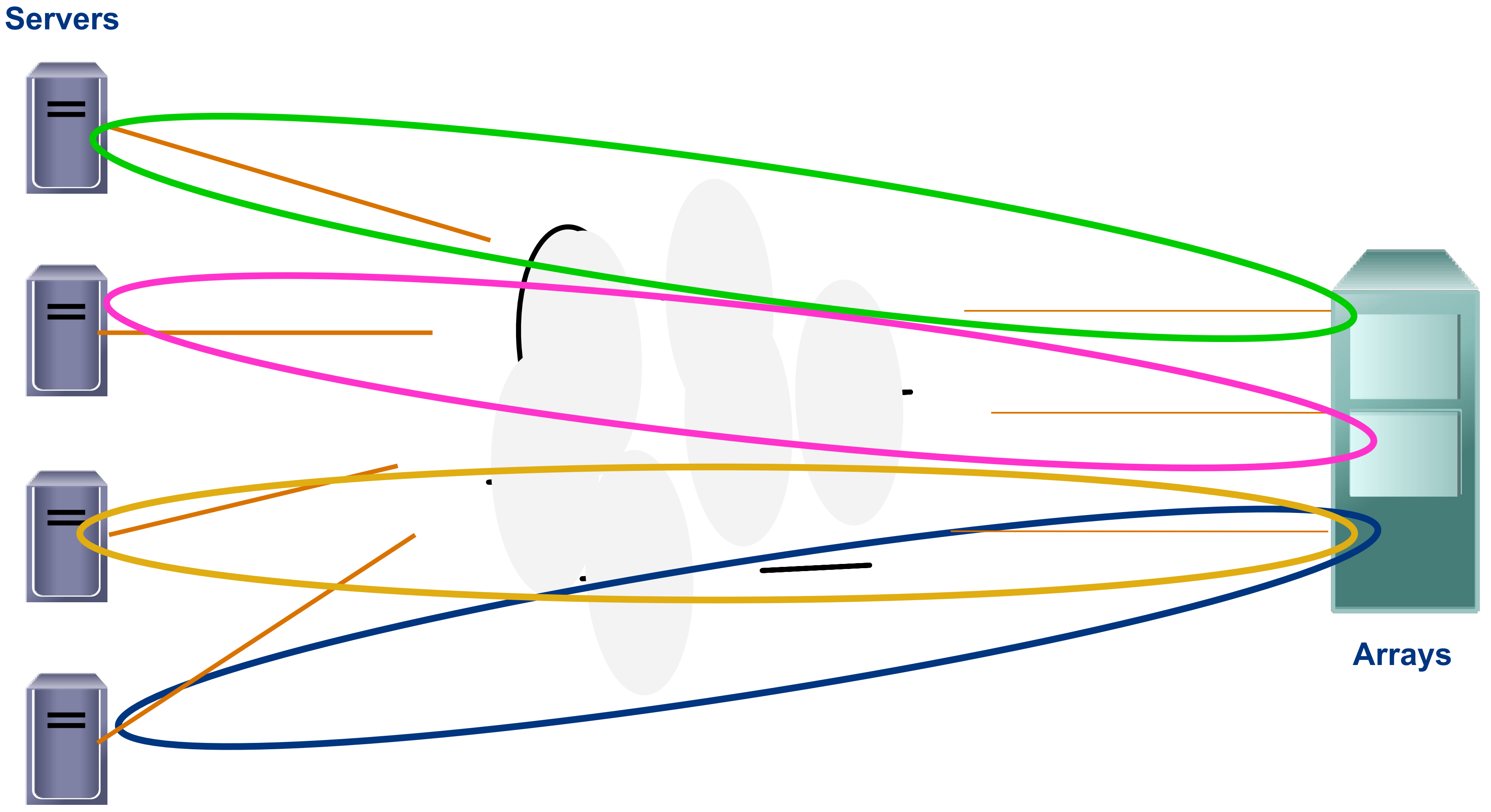
Figure 6-20: Types of zoning

Fibre Channel Login Types

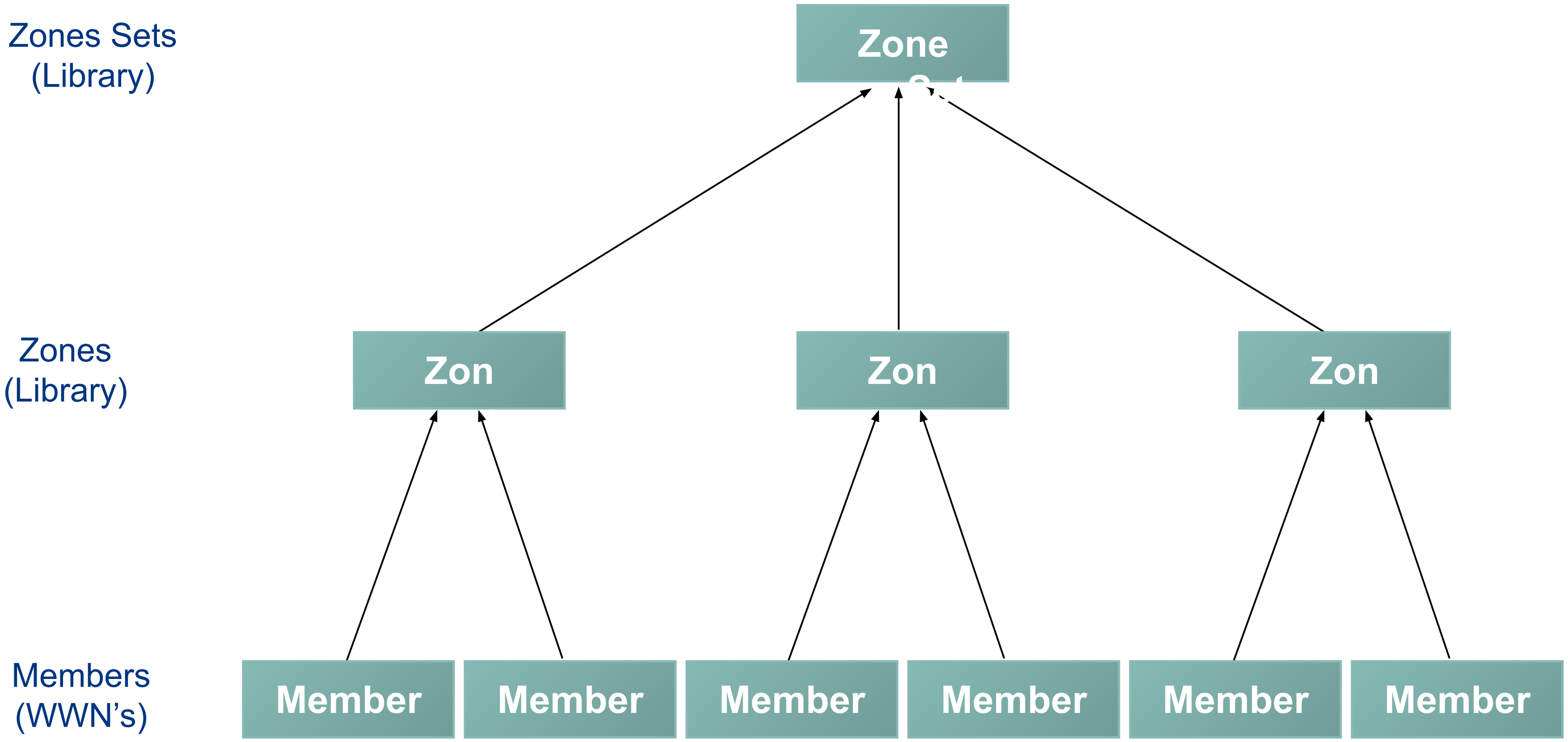
Fabric services define three login types:

- Fabric login (FLOGI) is performed between an N_port and an F_port.
- Port login (PLOGI) is performed between an N_port and another N_port to establish a session.
- Process login (PRLI) is also performed between an N_port and another N_port.

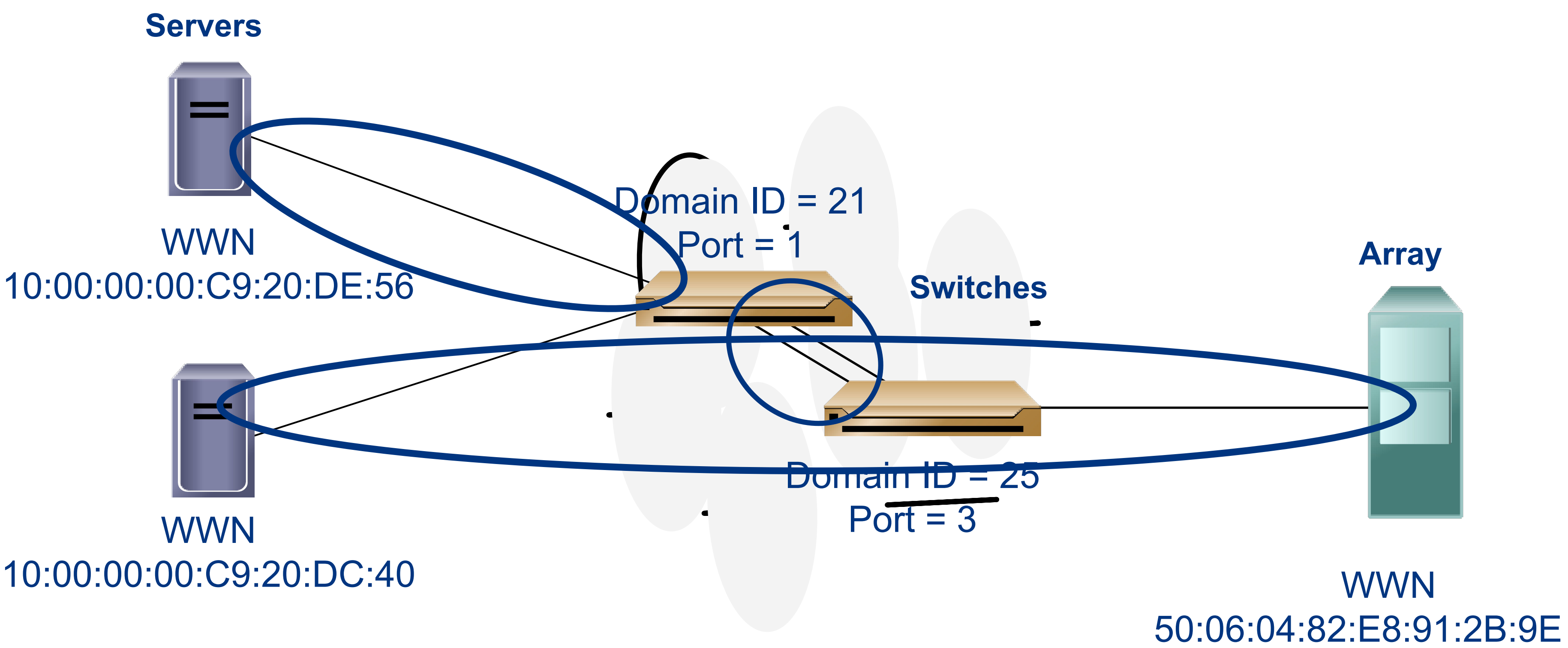
Fabric Management: Zoning



Zoning Components



Types of Zoning



Examples:

WWN Zone 1 = 10:00:00:00:C9:20:DC:40; 50:06:04:82:E8:91:2B:9E

Port Zone 1 = 21,1; 25,3

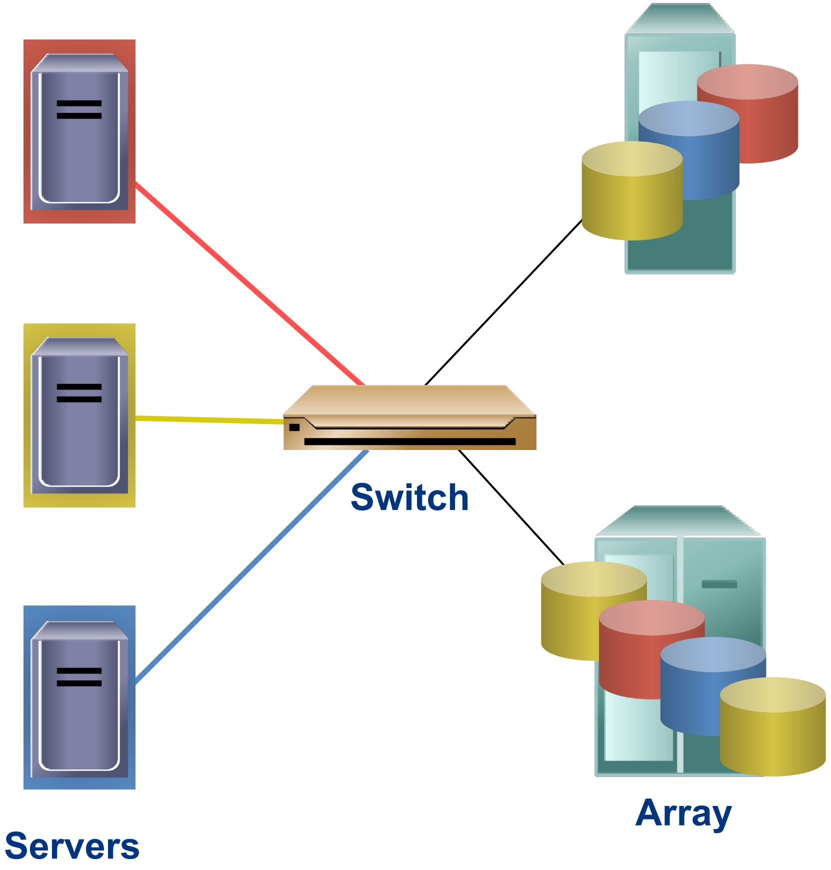
Mixed Zone 1 = 10:00:00:00:C9:20:DE:56; Port 21/1

Single HBA Zoning

- Optimally, one HBA per zone.
  - Nodes can only “talk” to Storage in the same zone
- Storage Ports may be members of more than one zone.
- HBA ports are isolated from each other to avoid potential problems associated with the SCSI discovery process.
  - Also known as “chatter”
- Decreases the impact of a changes in a Fabric by reducing the amount of nodes that must communicate.

Provisioning: LUN Masking

- Restricts volume access to specific hosts and/or host clusters.
- Servers can only access the volumes that they are assigned.
- Access controlled in the storage and not in the fabric
  - Makes distributed administration secure
- Tools to manage masking
  - GUI
  - Command Line



Capacity Management

- Tracking and managing assets
  - Number of ports assigned
  - Storage allocated
- Utilization profile
  - Indicates resource utilization over time
  - Allows for forecasting
- SAN management software provides the tools
  - Inventory databases
  - Report writers

Performance Management

- What is it?
  - Capturing metrics and monitoring trends
  - Proactively or Reactively responding
  - Planning for future growth
- Areas and functions
  - Host, Fabric and Storage Performance
  - Building baselines for the environment

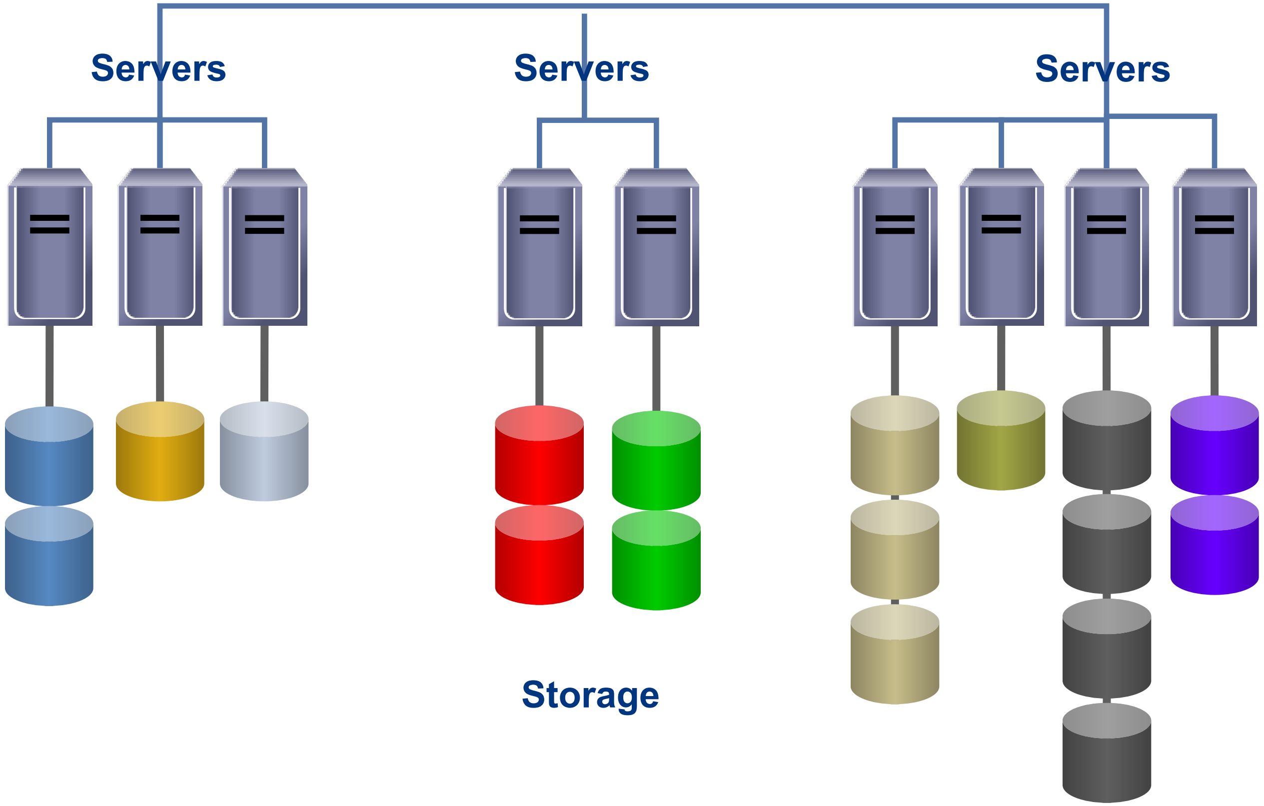
Lesson: Summary

- Topics in this lesson included:
  - Infrastructure protection
  - Provisioning
  - Capacity Management
  - Performance Management

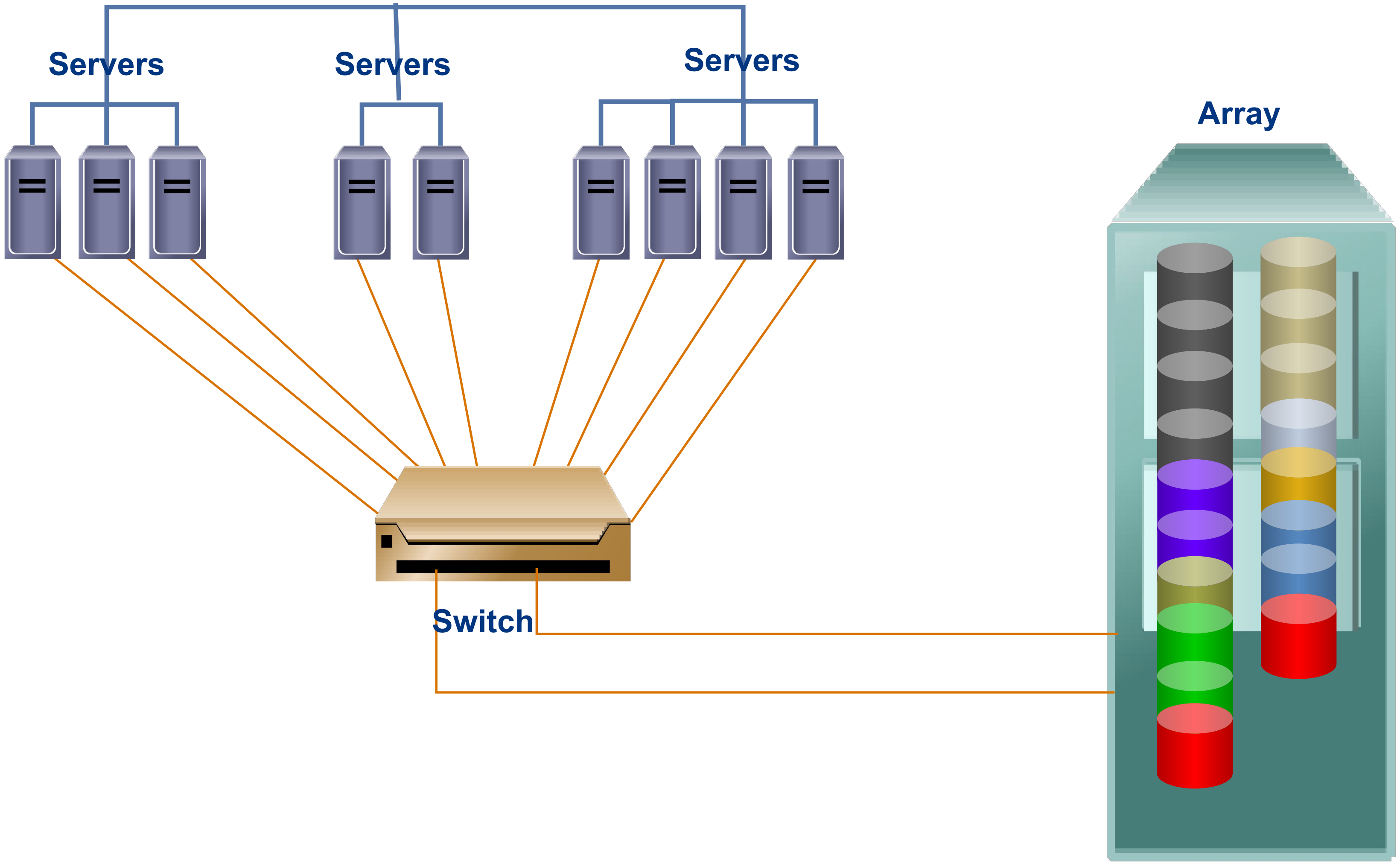
When Should a SAN be Used?

- SANs are optimized for high bandwidth block level I/O
- Suited for the demands of real time applications
  - Databases: OLTP (online transaction processing)
  - Video streaming
- Any applications with high transaction rate and high data volatility
  - Stringent requirements on I/O latency and throughput
- Used to consolidate heterogeneous storage environments
  - Physical consolidation
  - Logical consolidation

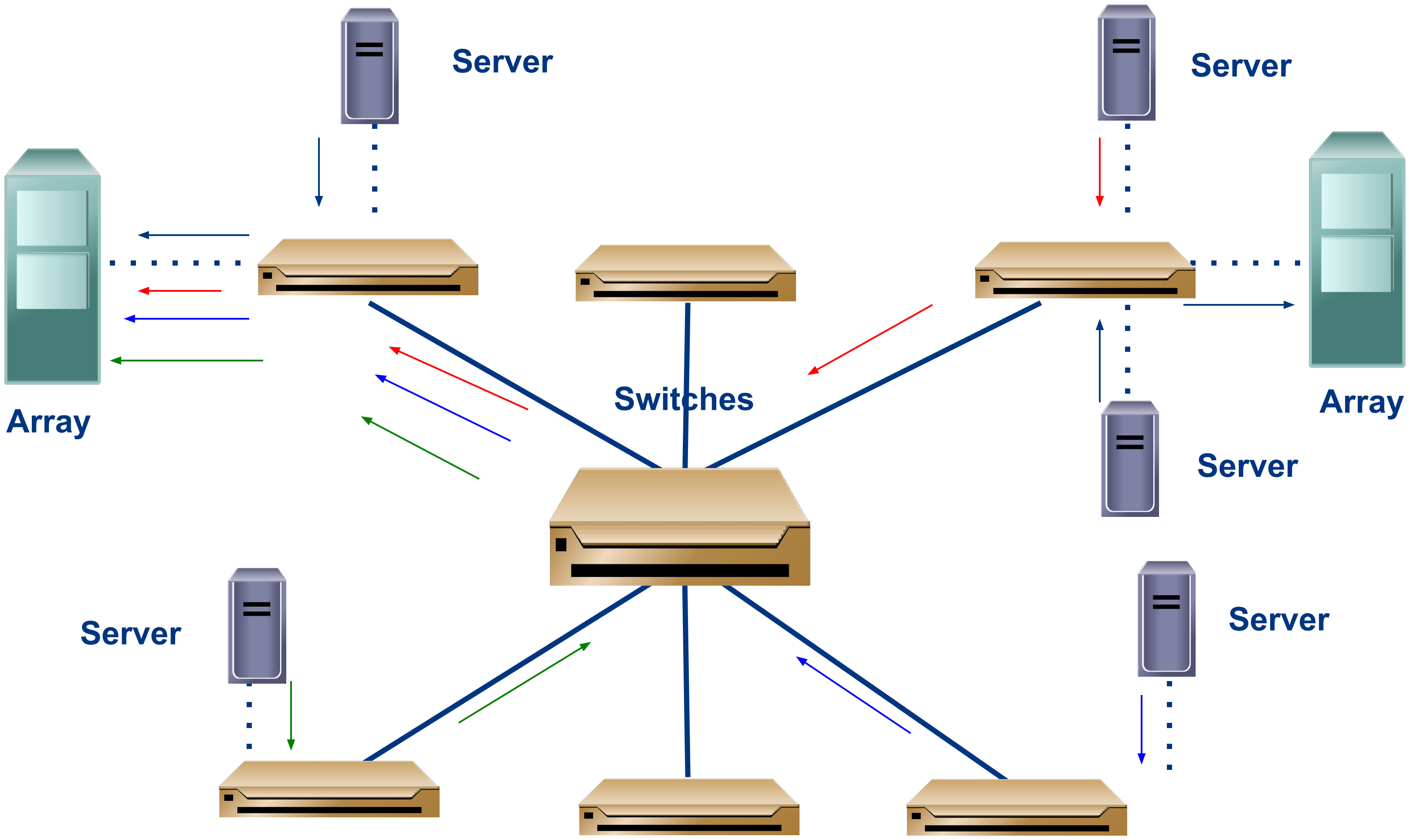
Consolidation Example: DAS Challenge



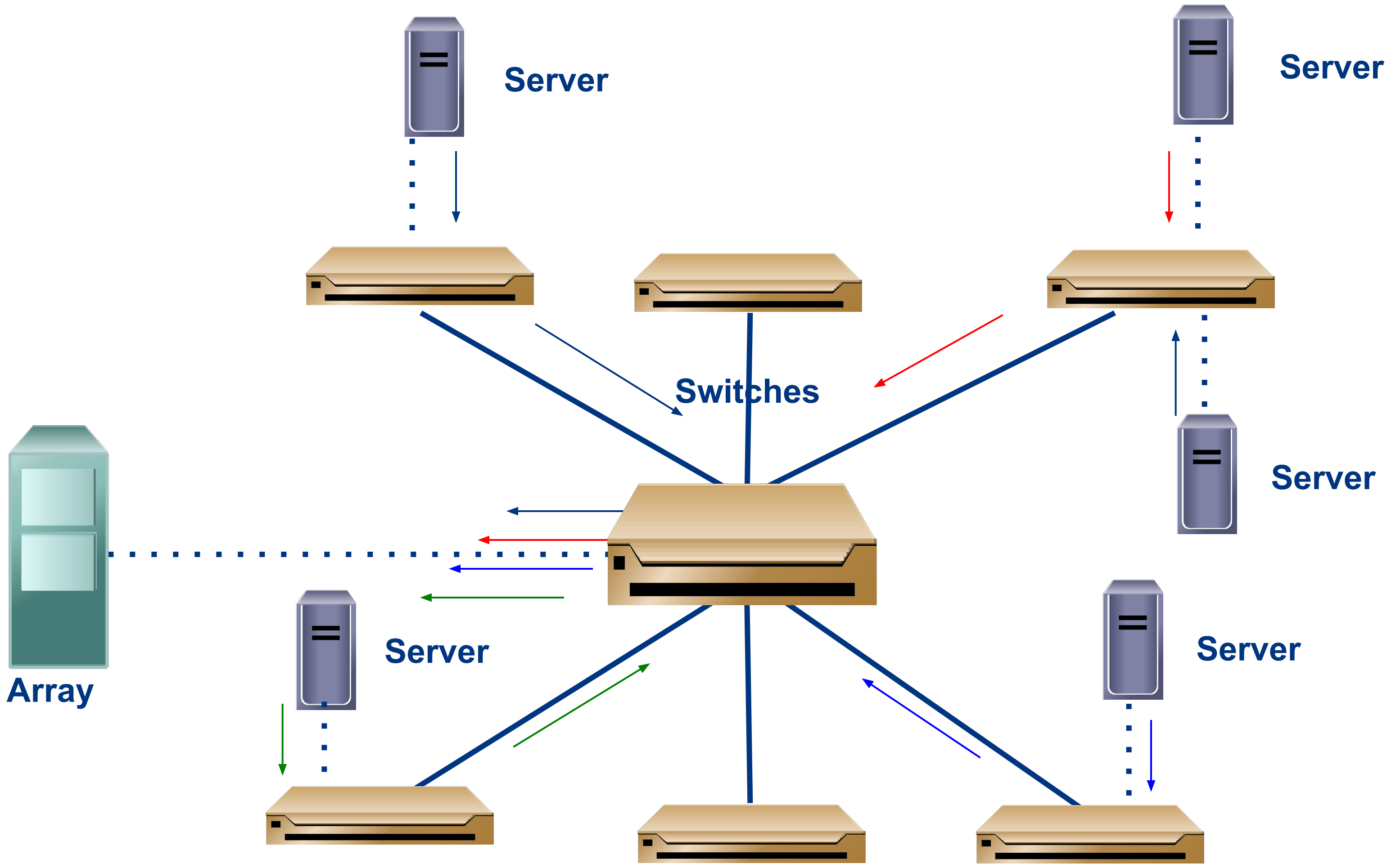
Consolidation Example: SAN Solution



Connectivity Example: Challenge



Connectivity Example: Solution



FC SAN Challenges

- Infrastructure
  - New, separate networks are required.
- Skill-sets
  - As a relatively new technology, FC SAN administrative skills need to be cultivated.
- Cost
  - Large investments are required for effective implementation.

FC Connectivity

The FC architecture supports three basic interconnectivity options:

- **point-to point**
- Fibre Channel **Arbitrated loop (FC-AL)**
- **fabric connect**

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.

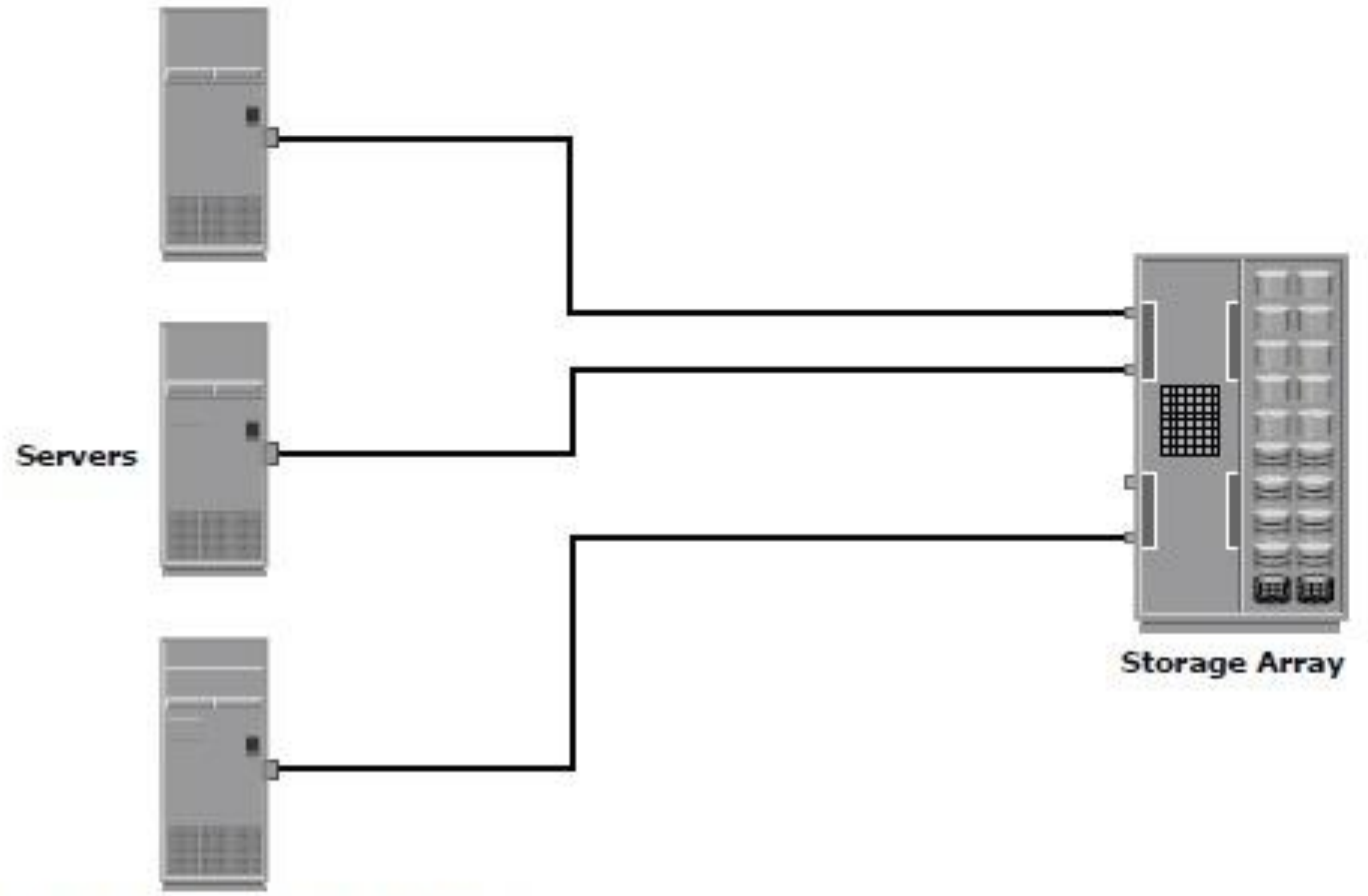


Figure 6-6: Point-to-point topology

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.

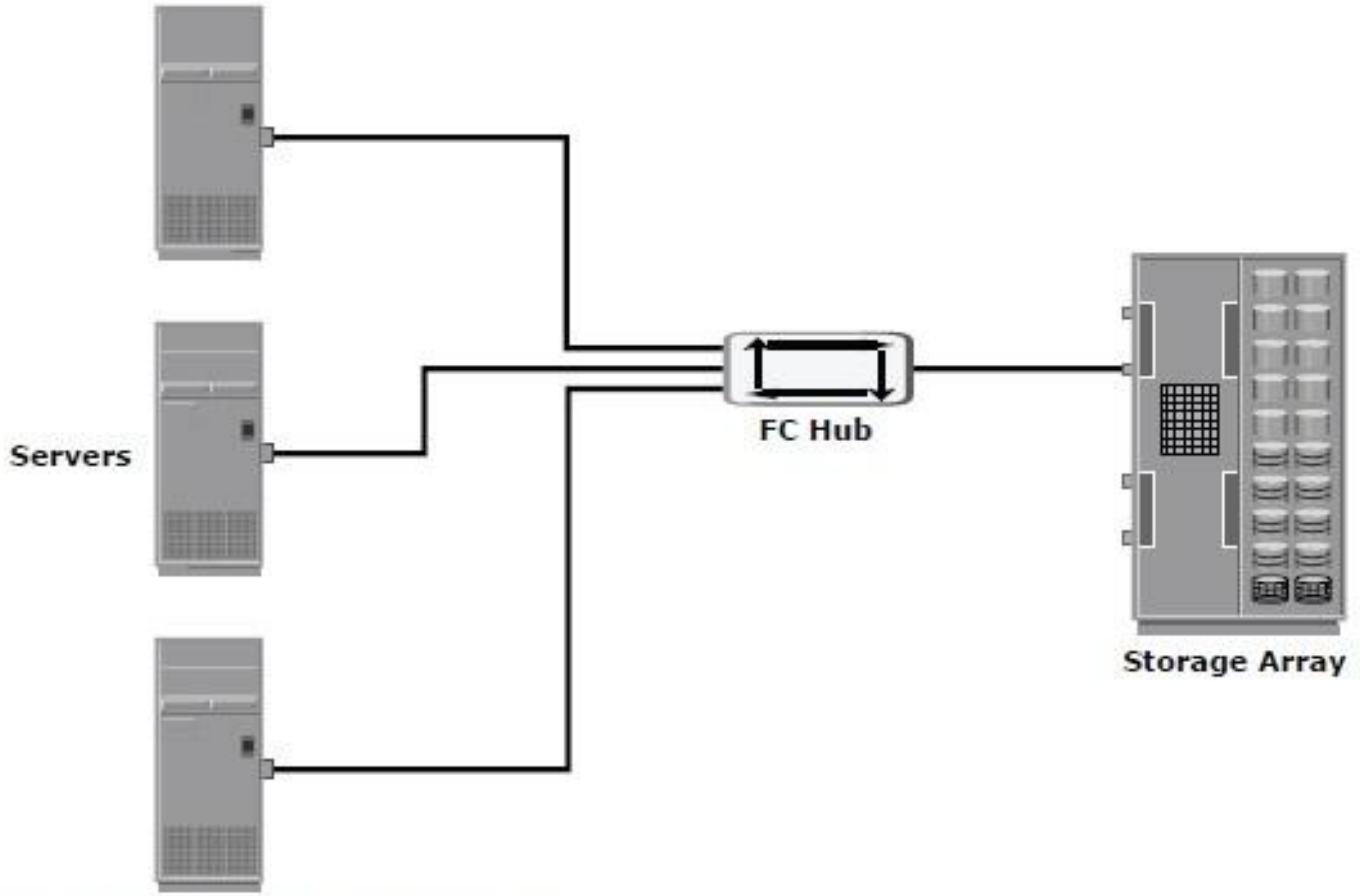


Figure 6-7: Fibre Channel arbitrated loop

- 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.
- 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.

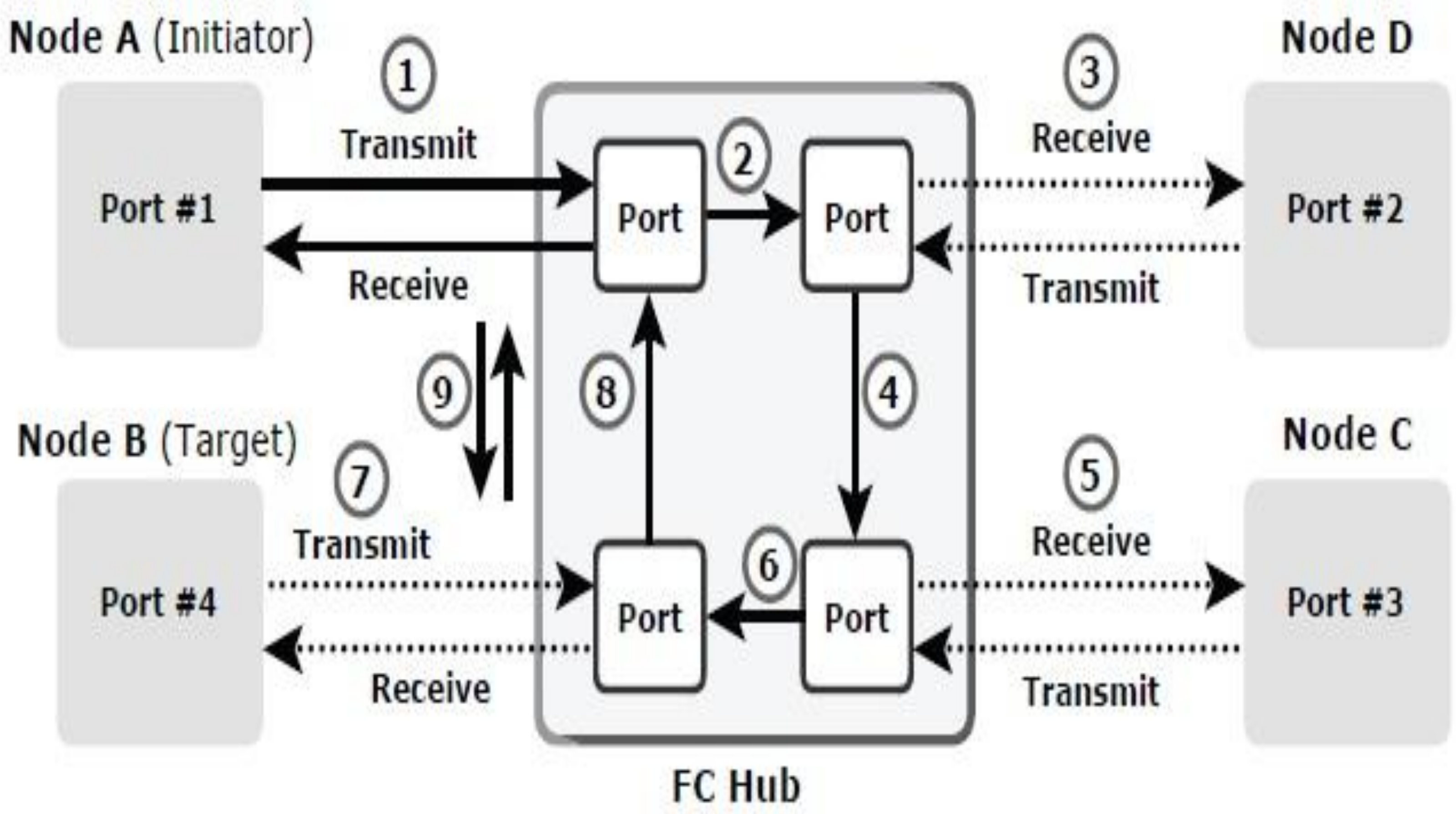


Figure 6-8: Data transmission in FC-AL

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.

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.

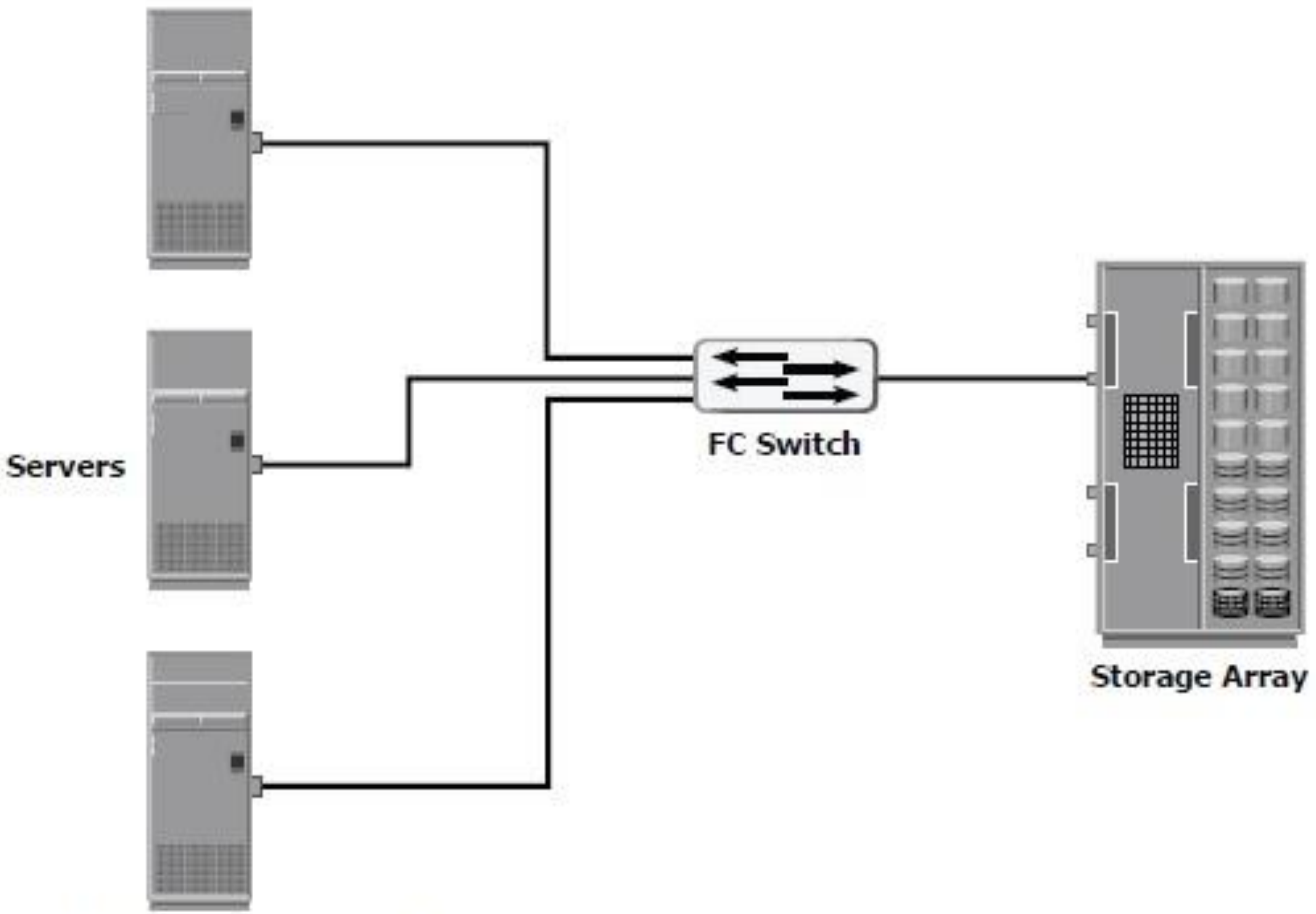


Figure 6-9: Fibre Channel switched fabric

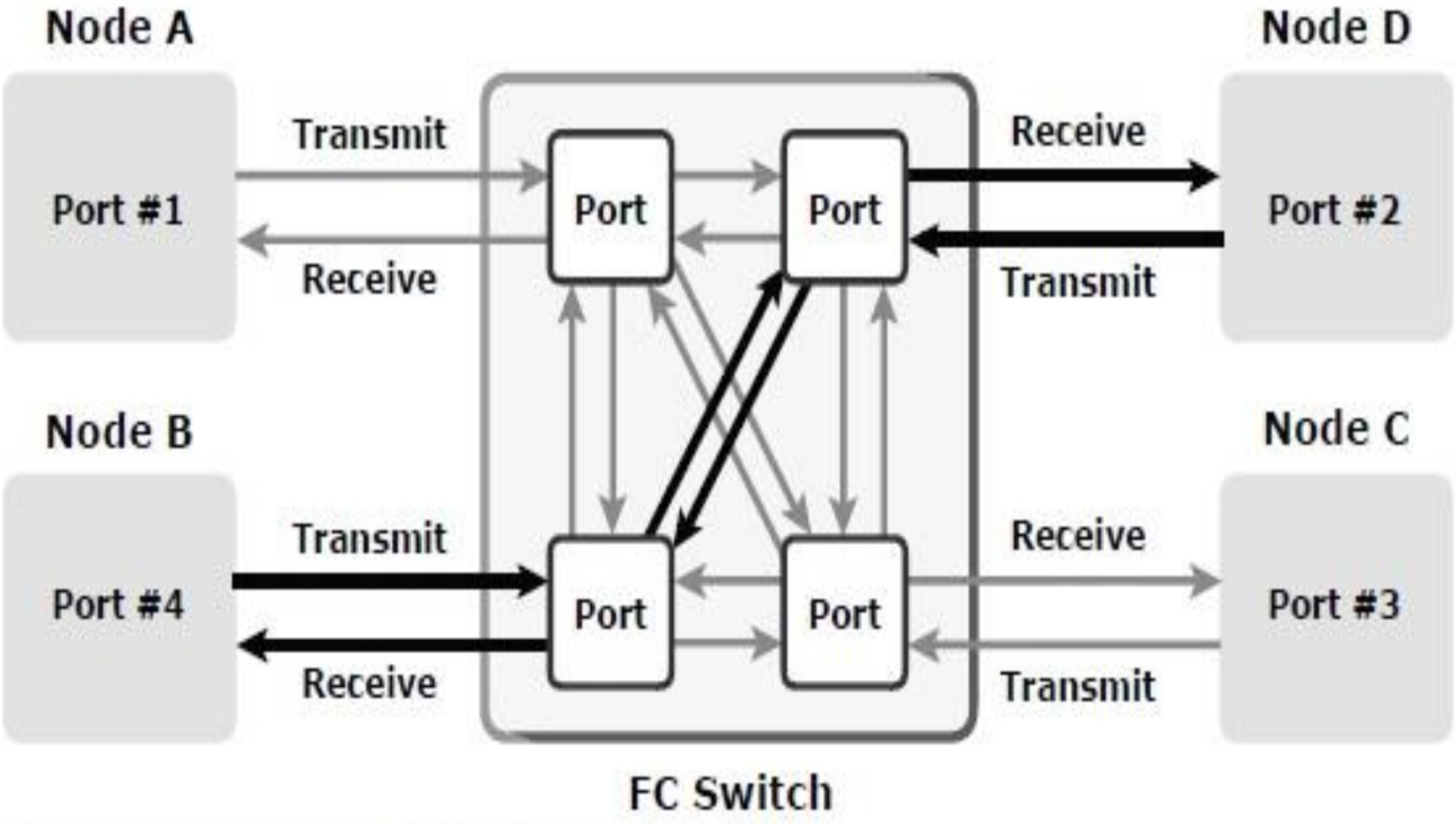


Figure 6-11: Data tansmission in FC-SW topology

Fibre Channel Architecture

- 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).

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.

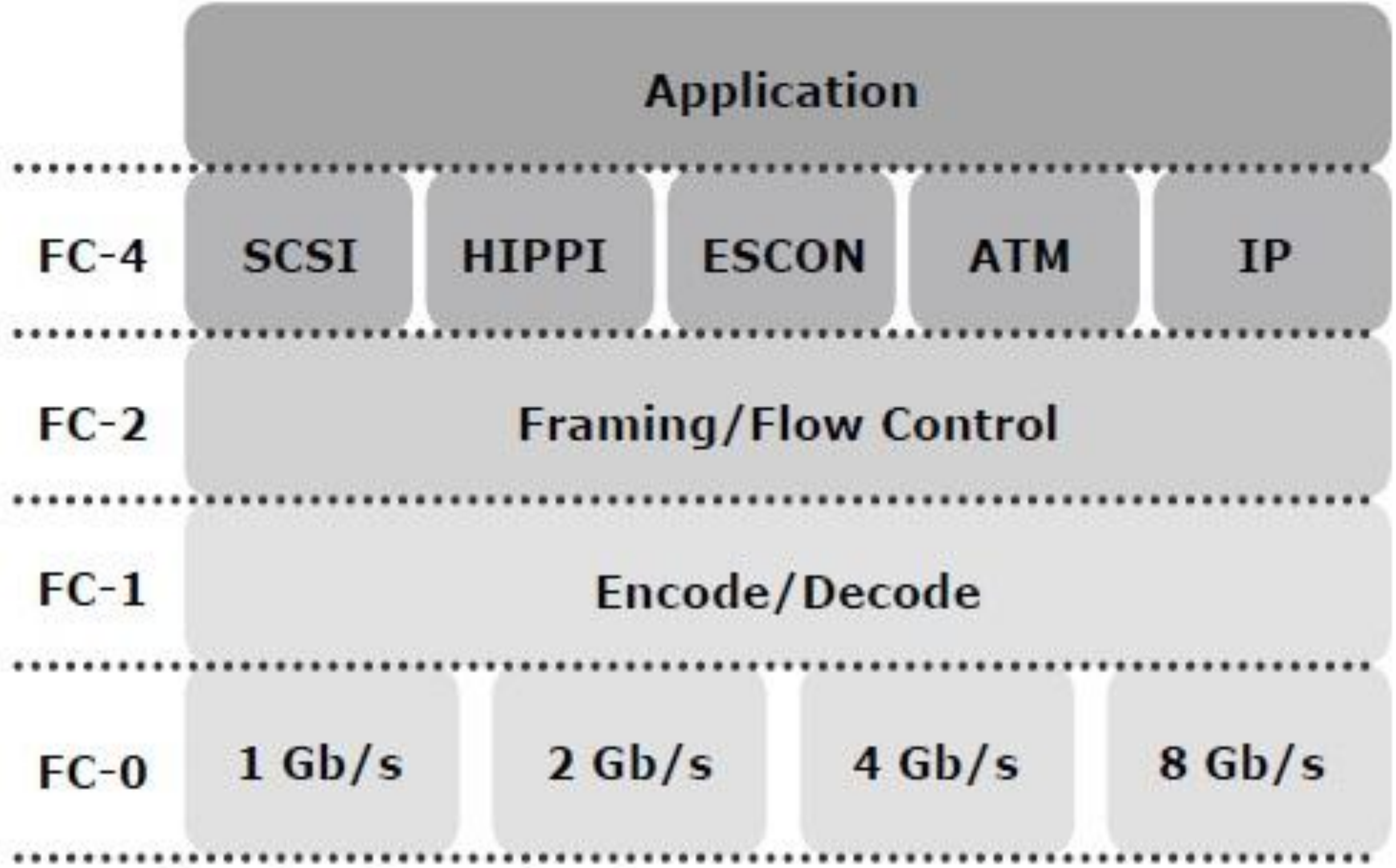


Figure 6-13: Fibre channel protocol stack

1. **FC-4 Upper Layer Protocol**

   SCSI, HIPPI Framing Protocol, Enterprise Storage Connectivity (ESCON), ATM, and IP.

2. **FC-2 Transport Layer**

   fabric services, classes of service, flow control, and routing.

3. **FC-1 Transmission Protocol**

4. **FC-0 Physical Interface**

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, FFFFFC is reserved for the name server, and FFFFFE is reserved for the fabric login service.

The maximum possible number of N_ports in a switched fabric is calculated as 239 domains

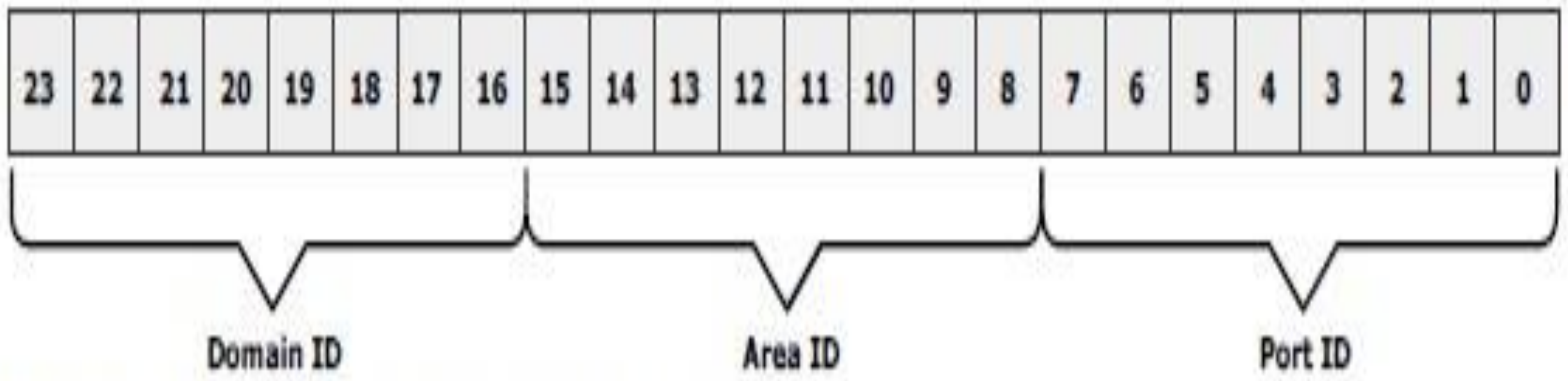


Figure 6-14: 24-bit FC address of N_port

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.

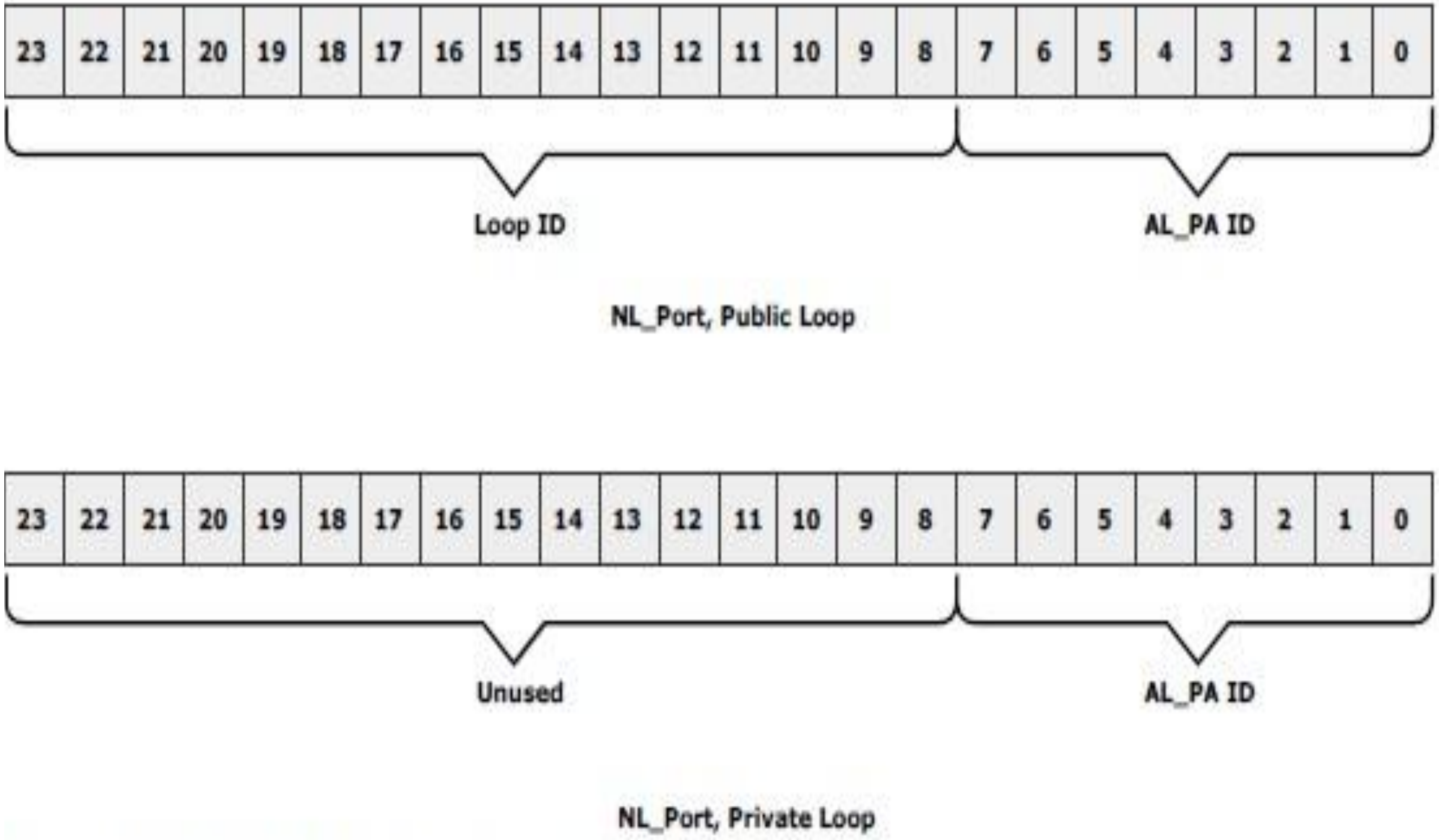


Figure 6-15: 24-bit FC address of NL_port

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.

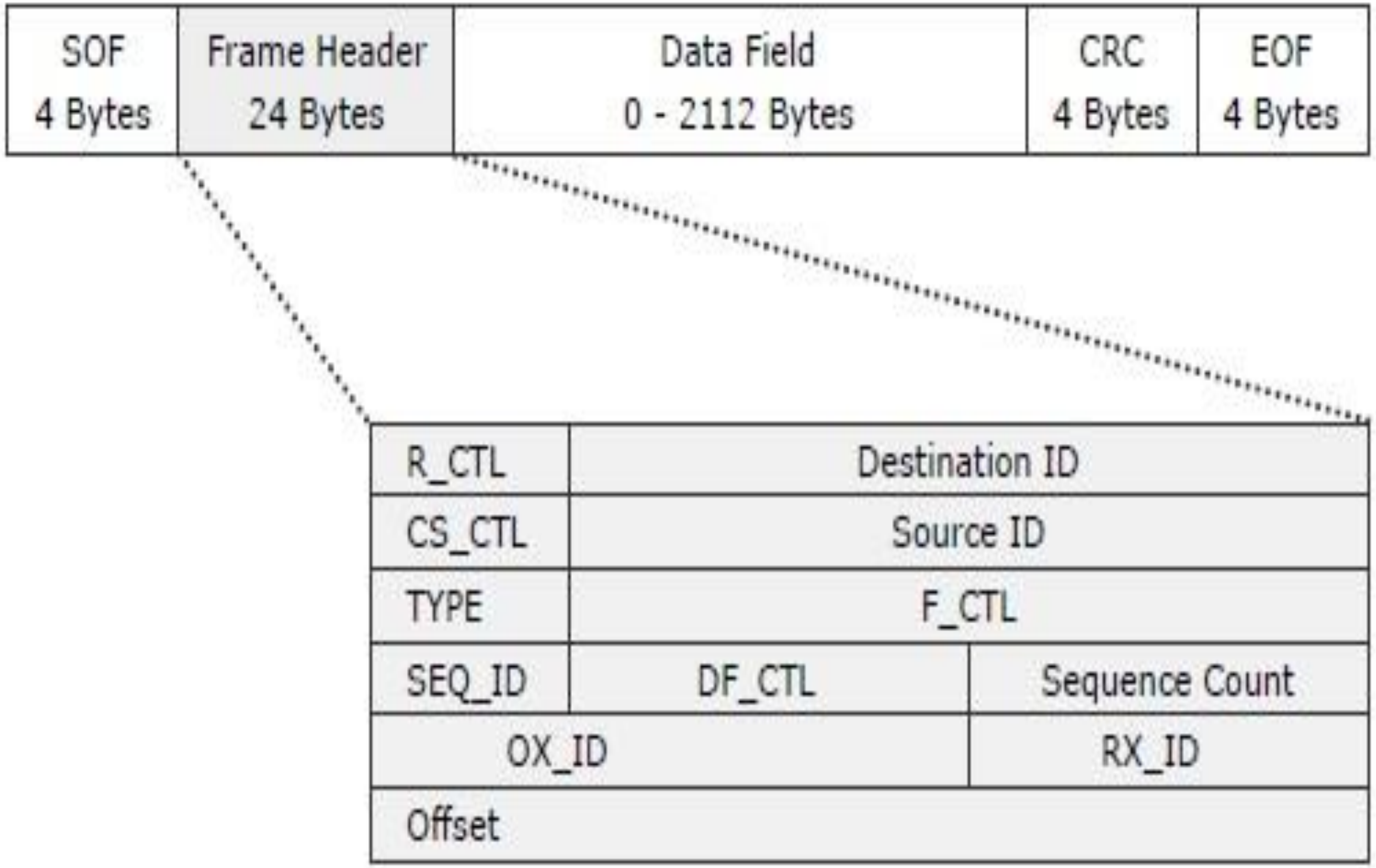


Figure 6-17: FC frame

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)**
- **Class Specific Control (CS_CTL)**
- **TYPE**
- **Data Field Control (DF_CTL)**
- **Frame Control (F_CTL)**

Structure and Organization of FC Data

- **Exchange operation**

An exchange operation enables two N_ports to identify and manage a set of information units.

- **Sequence**

A sequence refers to a contiguous set of frames that are sent from one port to another.

- **Frame**

A frame is the fundamental unit of data transfer at Layer 2.

Each frame can contain up to 2,112 bytes of payload.

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)**
- **end-to-end credit (EE_Credit)**

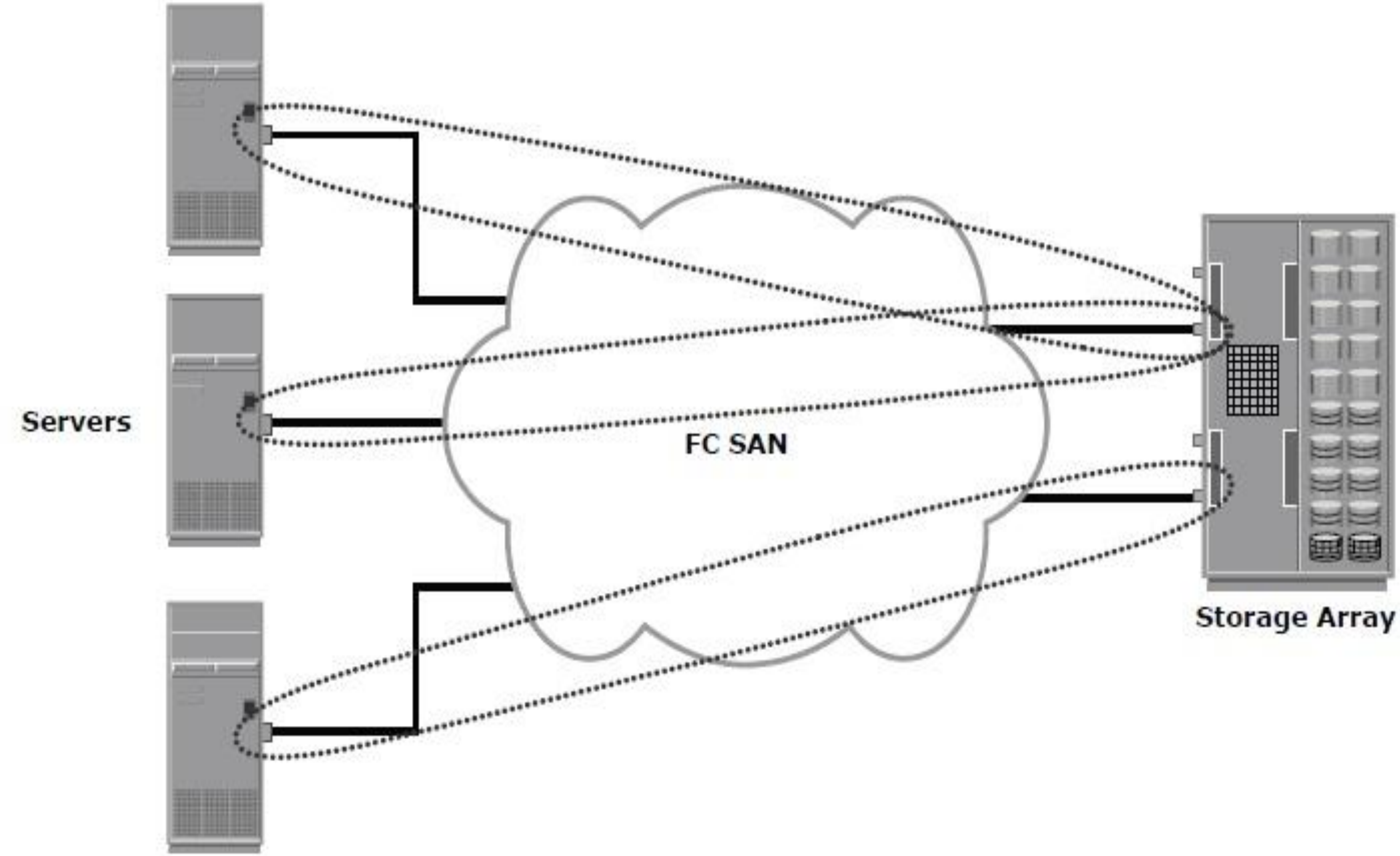


Figure 6-18: Zoning

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).

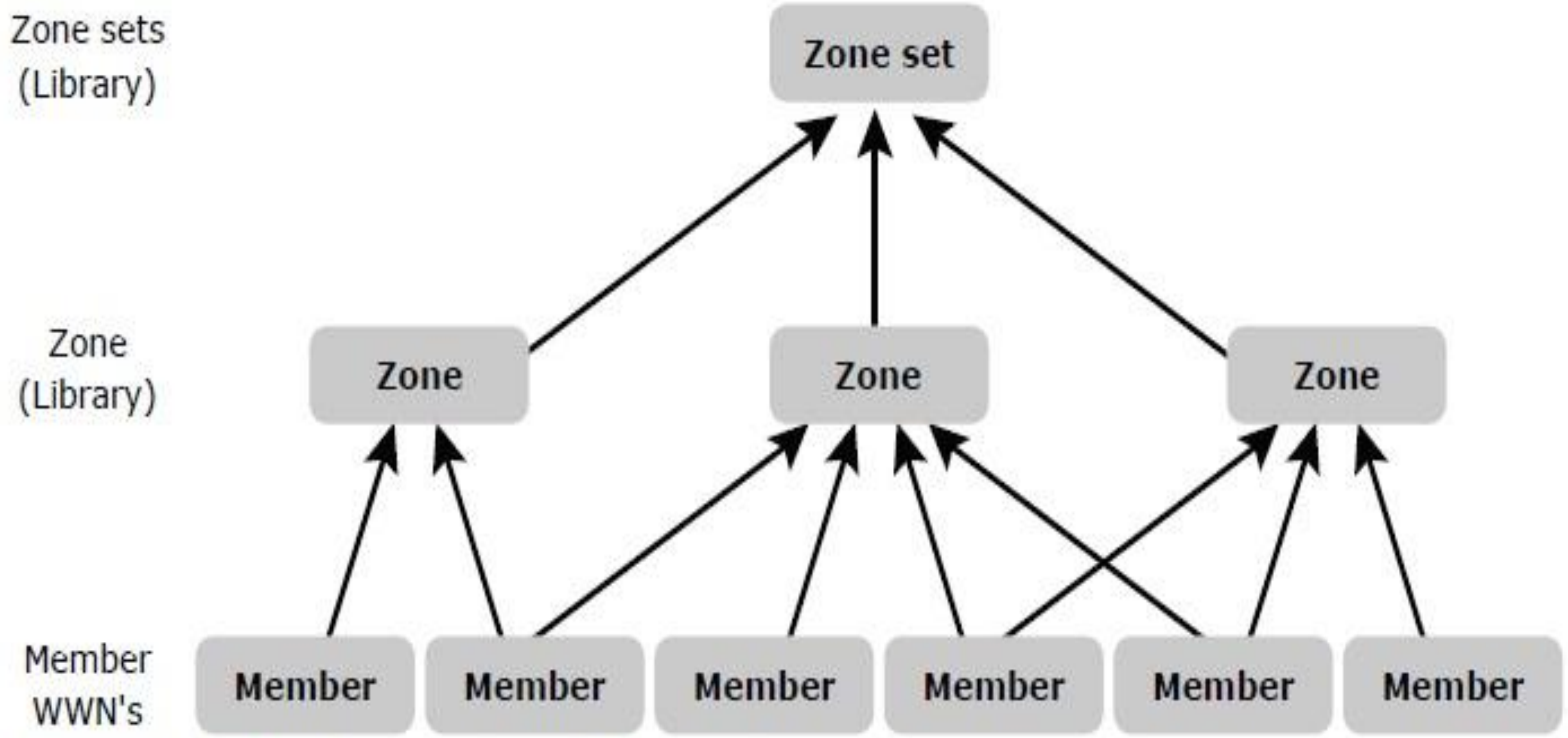


Figure 6-19: Members, zones, and zone sets

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.

- **Core-Edge Fabric**
- **Mesh Topology**

Core-Edge Fabric

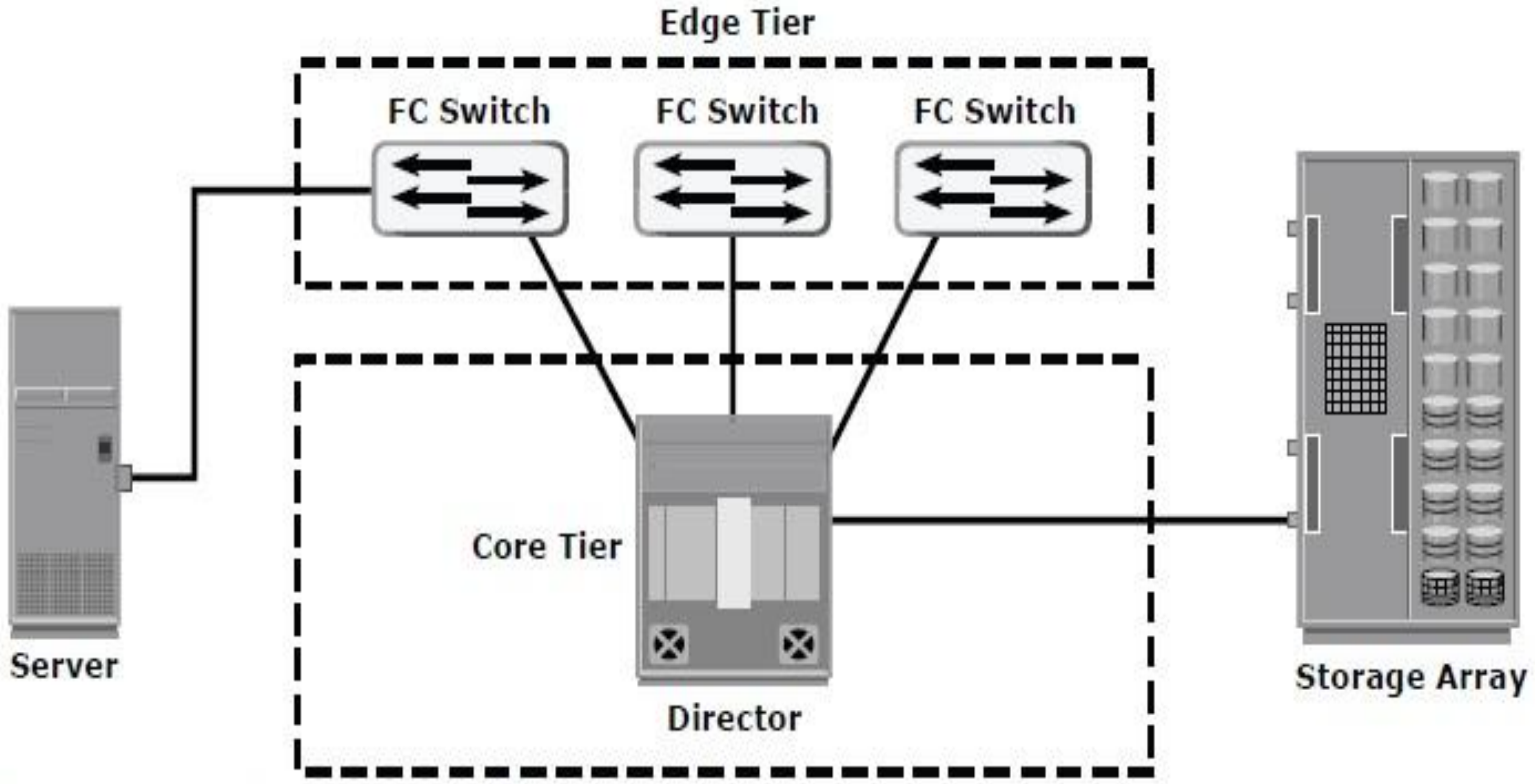


Figure 6-21: Single core topology

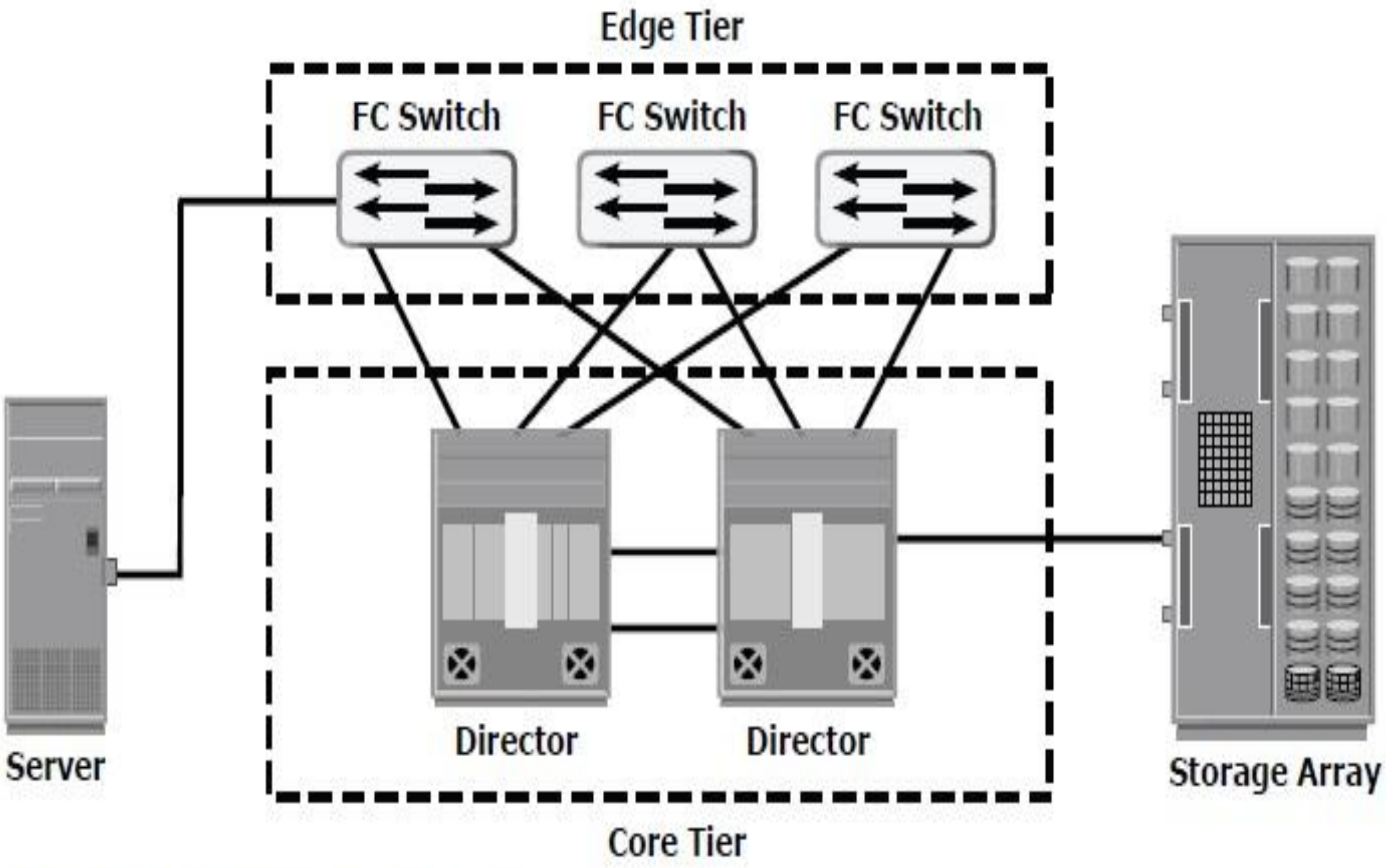


Figure 6-22: Dual-core topology

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.

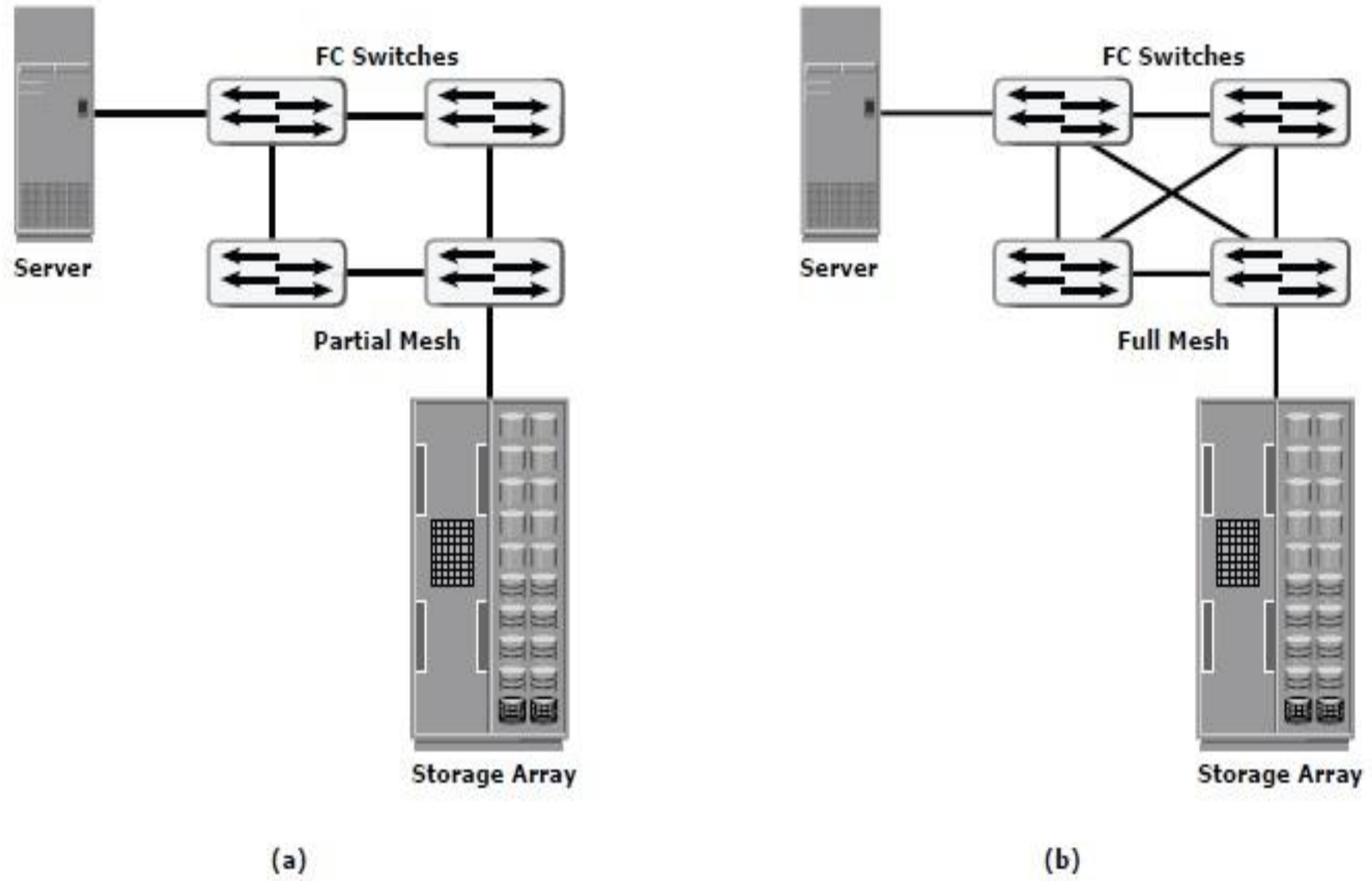


Figure 6-23: Partial mesh and full mesh topologies

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.

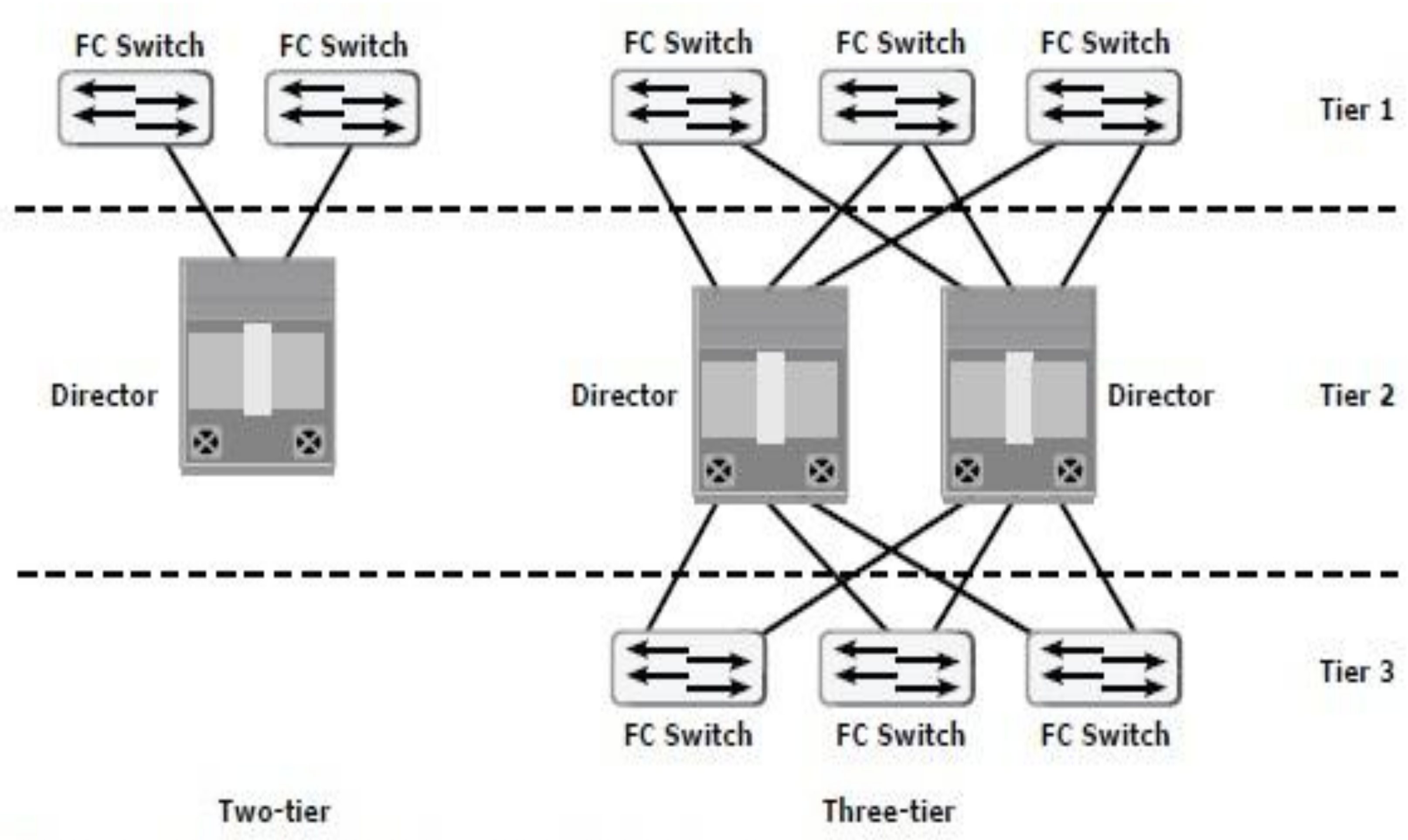


Figure 6-10: Tiered structure of FC-SW topology

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.

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 acknowl-edgement	Acknowledged	Acknowledged	Not acknowledged
Multiplexing	No	Yes	Yes
Bandwidth utilization	Poor	Moderate	High

Lesson: Summary

Topics in this lesson included:

- Common SAN deployment considerations.
- SAN Implementation Scenarios
  - Consolidation
  - Connectivity
- SAN Challenges

Apply Your Knowledge…

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

- Describe EMC's product implementation of the Connectrix™ Family of SAN Switches and Directors.

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.

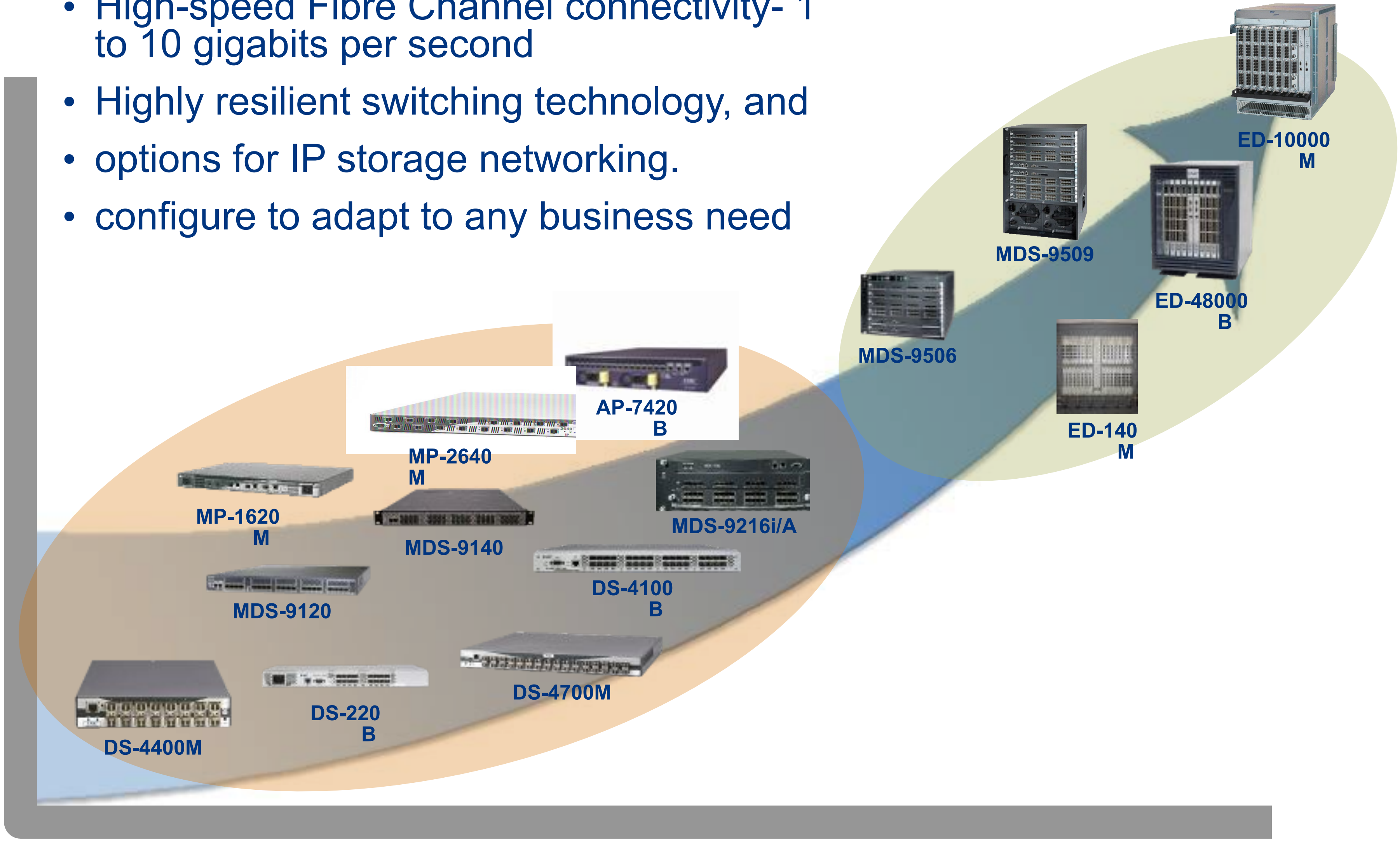
- **Connectrix Switches**
- **Connectrix Directors**
- **Connectrix Management Tools**

EMC offers the following connectivity products under the Connectrix brand :

- **Enterprise directors**
- **Departmental switches**
- **Multiprotocol routers**

The Connectrix Family

- High-speed Fibre Channel connectivity- 1 to 10 gigabits per second
- Highly resilient switching technology, and
- options for IP storage networking.
- configure to adapt to any business need



Switches versus Directors

- Connectrix Switches
  - High availability through redundant deployment
  - Redundant fans and power supplies
  - Departmental deployment or part of Data Center deployment
  - Small to medium fabrics
  - Multi-protocol possibilities
- Connectrix Directors
  - "Redundant everything" provides optimal serviceability and highest availability
  - Data center deployment
  - Maximum scalability
  - Maximum performance
  - Large fabrics
  - Multi-protocol

Connectrix Switch - DS-220B

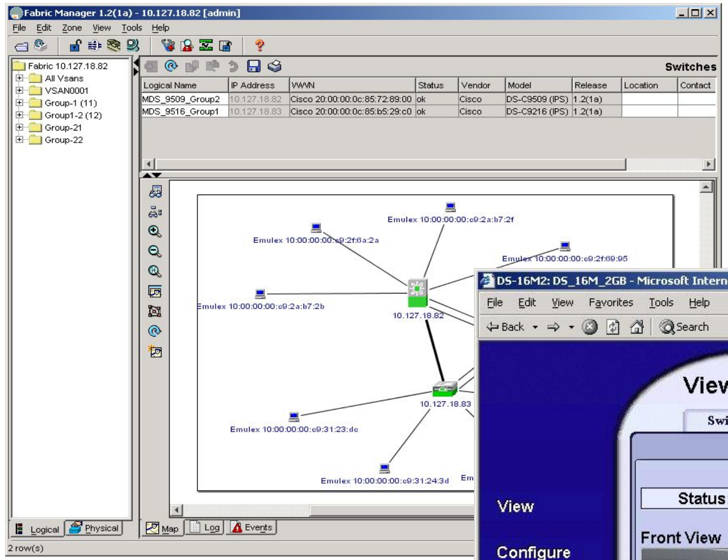
- Provides eight, 12, or 16 ports
  - Auto-detecting 1, 2, and 4 Gb/s Fibre Channel ports
  - Single, fixed power supply
  - Field-replaceable optics
  - Redundant cooling
- Simplified setup—no previous SAN experience needed
  - Eliminates the need for advanced skills to manage IP addressing or Zoning

Connectrix Director – MDS 9509

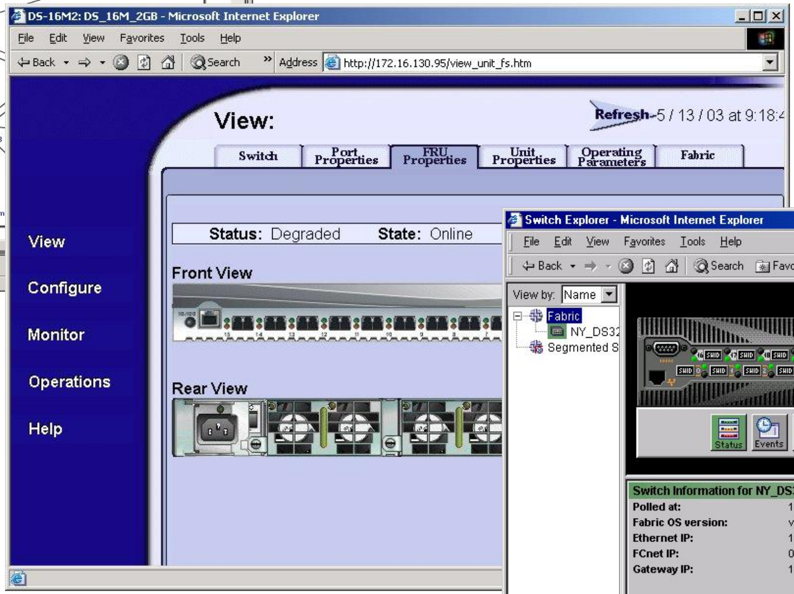
- Multi-transport switch—Fibre Channel, FICON, iSCSI, FCIP
  - 16 to 224 Fibre Channel ports
  - 4–56 Gigabit Ethernet ports for iSCSI or FCIP
  - Non-blocking fabric
  - 1 / 2 Gb/s auto-sensing ports
- All components are fully redundant

MDS-9509

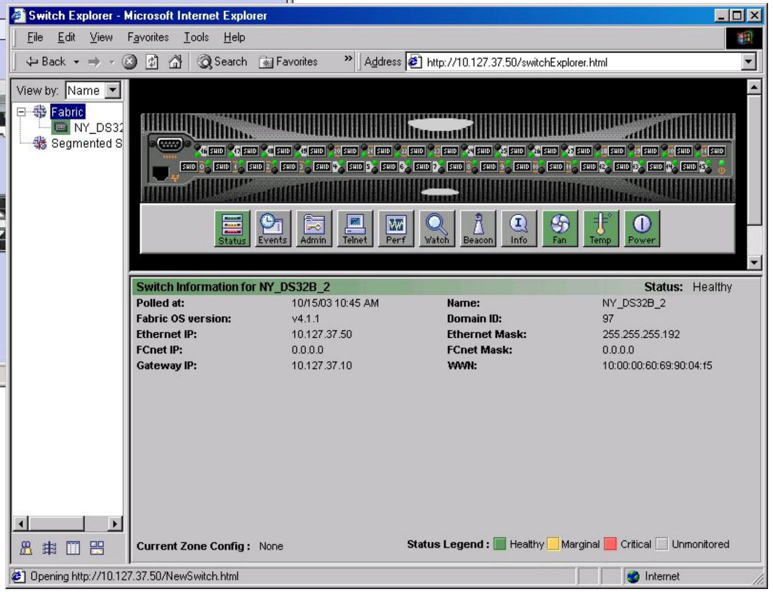
Connectrix Management Interfaces



MDS-Series Fabric Manager



M-Series Web Server



B-Series Web Tools

Module Summary

The Connectrix Family of Switches and Directors;

- Has three product sets:
  - Connectrix B-Series
  - Connectrix MDS 9000 Series
  - Connectrix M-Series
- Provides highly available access to storage.
- Connects a wide range of host and storage technologies.

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, has pushed the convergence of the SAN with IP technology, providing more benefits to using storage technologies.