RAMAIAH
Institute of Technology

M.S. Ramaiah Institute of Technology
(Autonomous Institute, Affiliated to VTU)
Department of Computer Science and Engineering

Course Name: Storage Area Networks

RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

IP SAN and FCoE(Chapter 6)

- iSCSI (Small Computer Systems Interface (iSCSI)
- FCIP (Fibre Channel over IP)
- FCoE(Fibre Channel Over Ethernet)

RAMAIAH
Institute of Technology

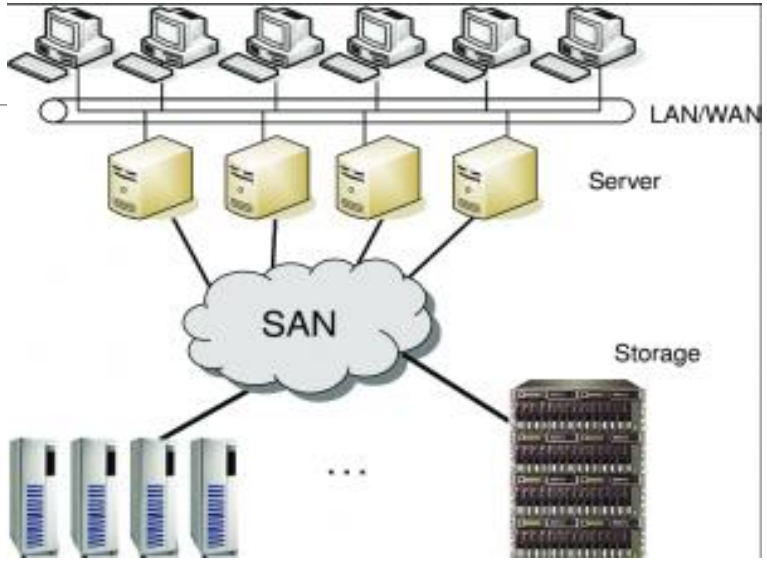
UNIT 3- Storage Networking Technologies

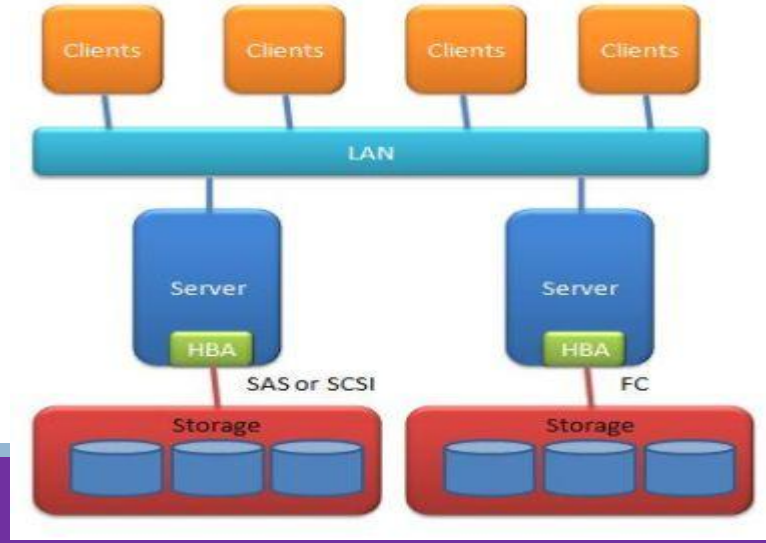
Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Introduction

- SAN is a high-speed network of servers and shared storage devices.
- SAN provides storage consolidation and facilitates centralized data management.
- Common SAN deployments are Fibre Channel (FC) SAN and IP SAN.
  - Fibre Channel SAN uses Fibre Channel protocol for the transport of data, commands, and status information between servers (or hosts) and storage devices.
  - IP SAN uses IP-based protocols for communication.





Explain the Fibre Channel Protocol Stack and key advantages of FCP.

Fiber Channel Overview

- Fibre channel(FC) is a technology that defines how data should be transmitted serially over copper and fiber optic media, from one node to another.
- FC networking was introduced in 1988
- 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 International Committee for Information Technology Standards (INCITS), is responsible for Fibre Channel interface standards.

Fiber Channel Overview

- High data transmission speed is an important feature of the FC networking technology.
- The initial implementation offered a throughput of 200 MB/s (equivalent to a raw bit rate of 1Gb/s).
- The latest FC implementations of 16 GFC (Fibre Channel) offer a throughput of 3200 MB/s (raw bit rates of 16 Gb/s), whereas Ultra640 SCSI is available with a throughput of 640 MB/s.

RAMAIAH
Institute of Technology

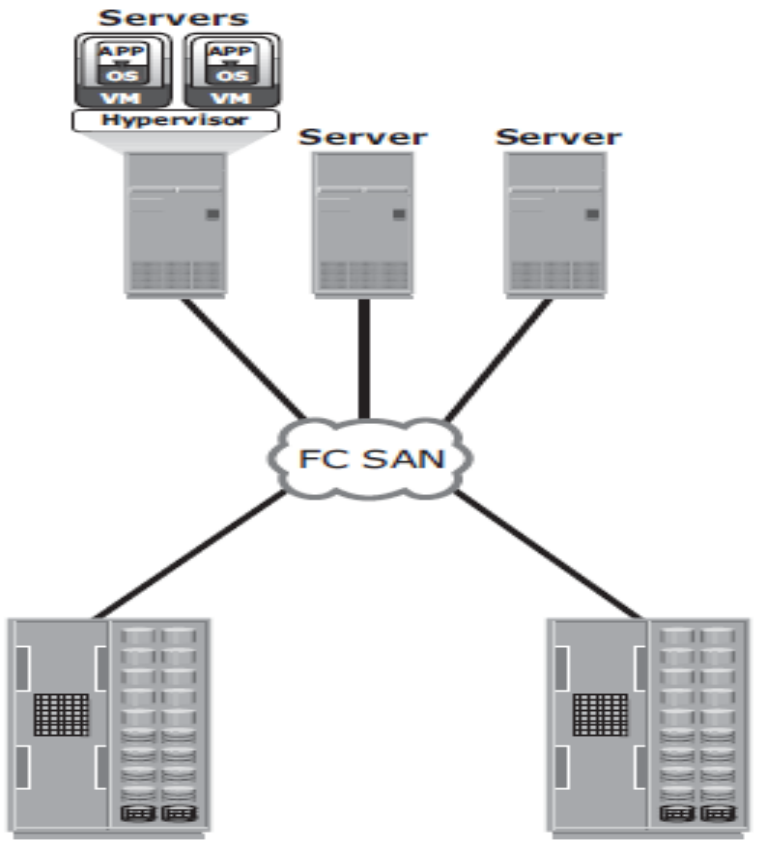
UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

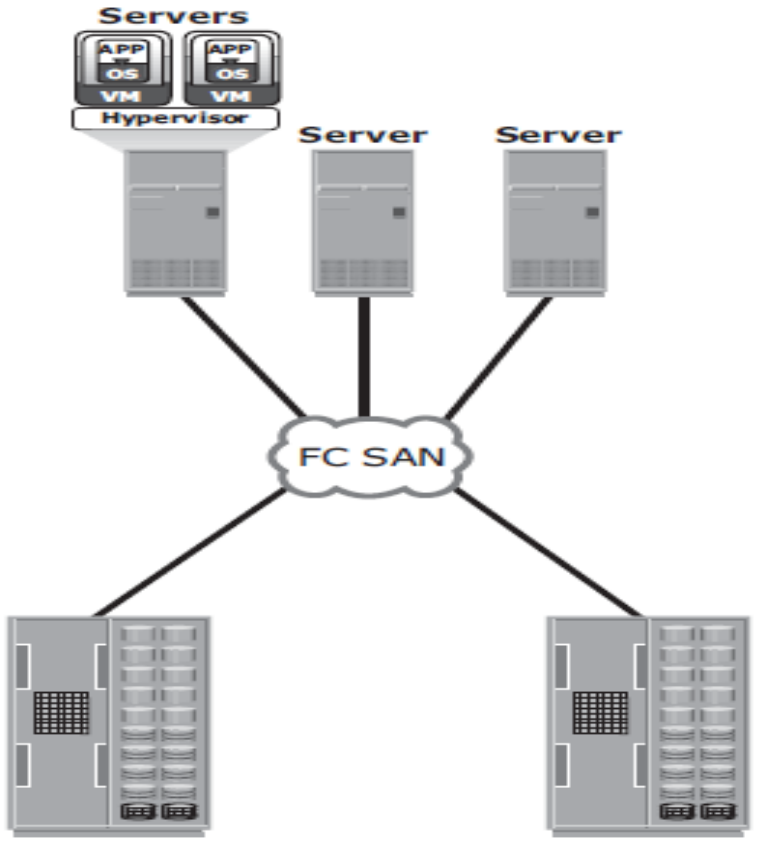
The SAN and Its Evolution

- A SAN carries data between servers (or *hosts*) and storage devices through Fibre Channel network .
- A SAN enables storage consolidation and enables storage to be shared across multiple servers.



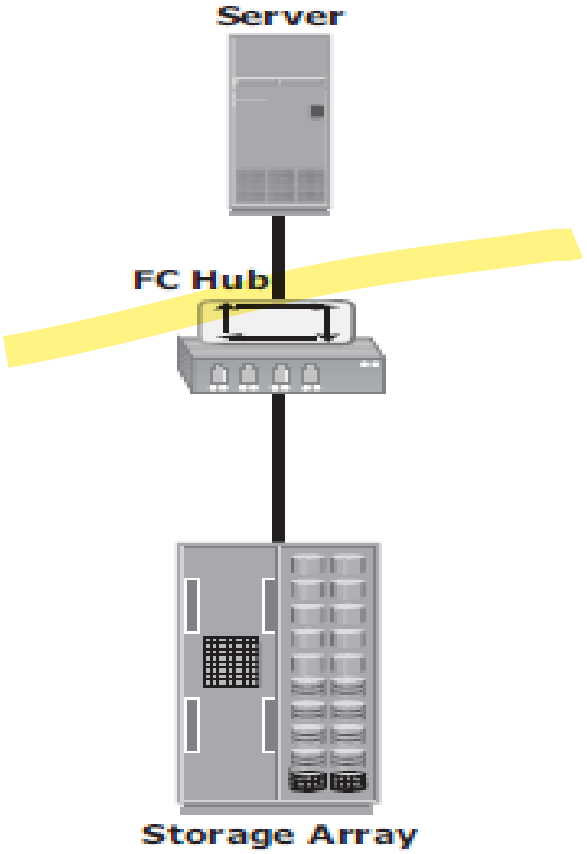
The SAN and Its Evolution

- This improves the
  - utilization of storage resources
  - reduces the total amount of storage an organization needs to purchase and manage.
  - storage management becomes centralized and less complex, which further reduces the cost of managing information.
  - SAN also enables organizations to connect geographically dispersed servers and storage.



The SAN and Its Evolution

- FC SAN : Grouping of hosts and storage devices connected to a network using an FC hub as a connectivity device.
- This configuration of an FC SAN is known as a Fibre Channel Arbitrated Loop (FC-AL).
- Use of hubs resulted in isolated FC-AL SAN islands because hubs provide limited connectivity and bandwidth.

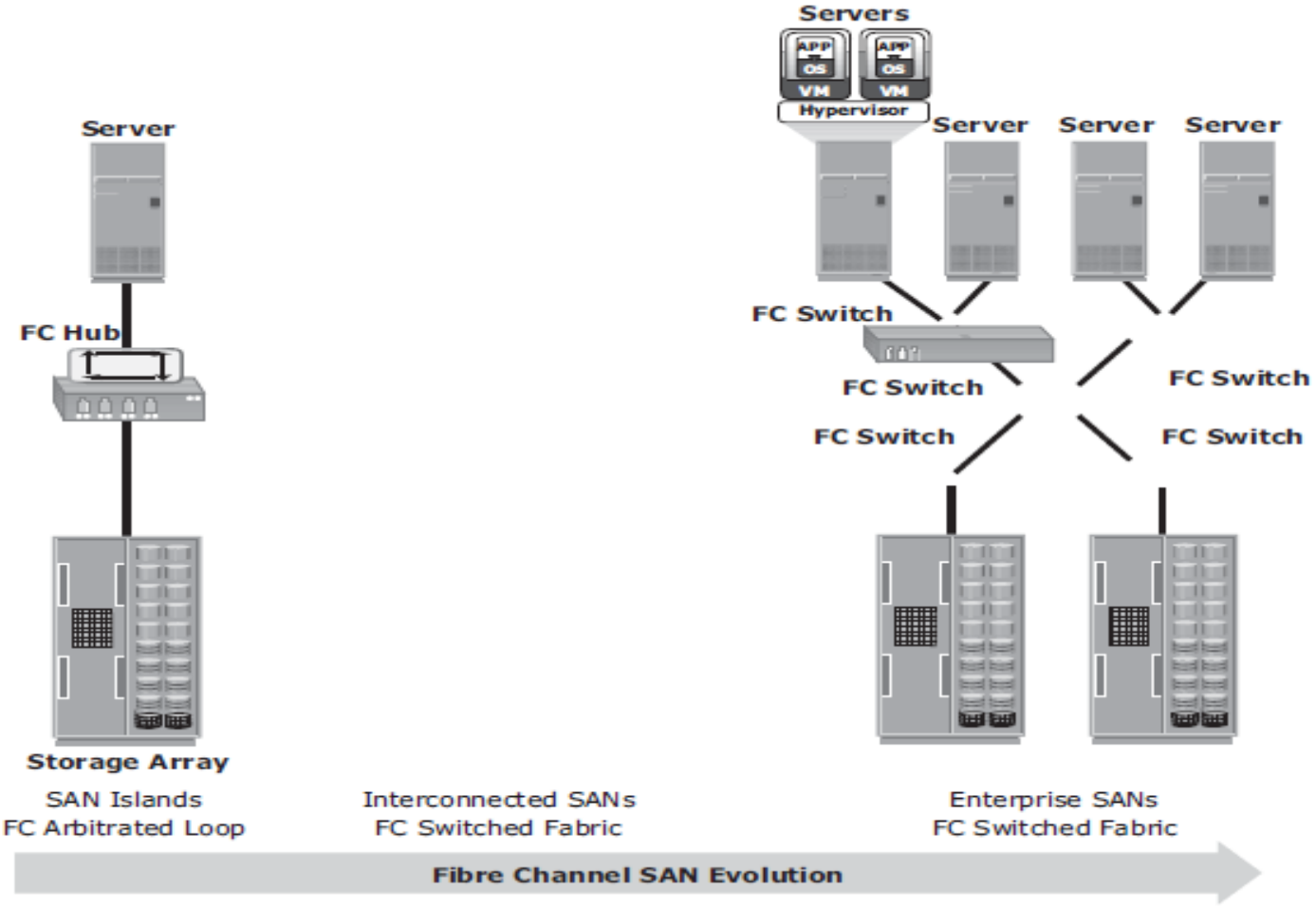


The SAN and Its Evolution

Limitations associated with hubs gave way to high-performance FC switches.

Use of switches in SAN improved connectivity and performance and enabled FC SANs to be highly scalable.

Figure illustrates the FC SAN evolution from FC-AL to enterprise SANs.



RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Components of FC SAN

- **Node Ports**
- **Cables and Connectors**
- **Interconnect Devices**
- **SAN Management Software**

Components of FC SAN

Node Ports

- In a Fibre Channel network, the end devices, such as hosts, storage arrays, and tape libraries, are to as nodes.
- Each node is a source or destination of information.
- Each node requires one or more ports to provide a physical interface for communicating with other nodes.
- These ports are integral components of host adapters, such as HBA, and storage front-end controllers or adapters.
- In an FC environment a port operates in full-duplex data transmission mode with a *transmit* (Tx) link and a *receive* (Rx) link (see Figure 5-3).

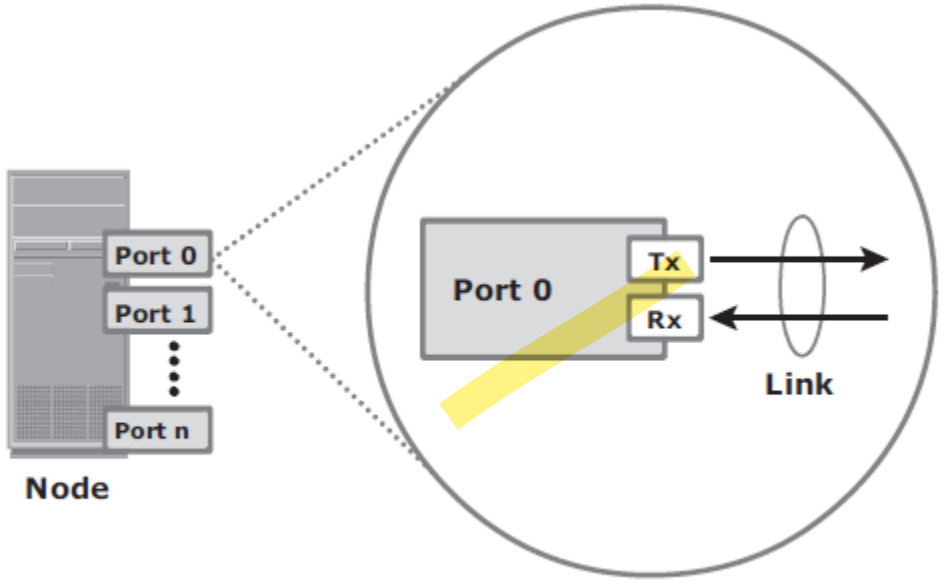


Figure 5-3: Nodes, ports, and links

Explain Multimode fiber and single-mode fiber

Components of FC SAN

Cables and Connectors

- SAN implementations use optical fiber cabling.
- Copper can be used for shorter distances for back-end connectivity because it provides an acceptable 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:
  - multimode and single-mode.

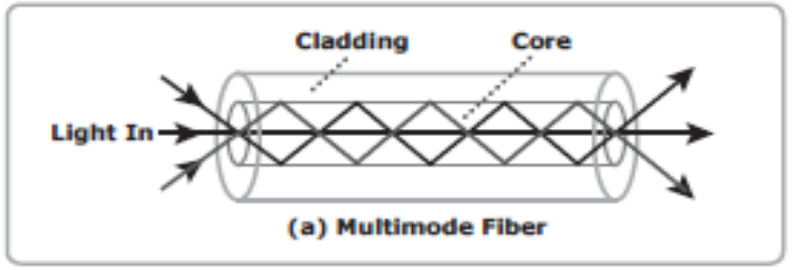


Figure 5-4: Multimode fiber

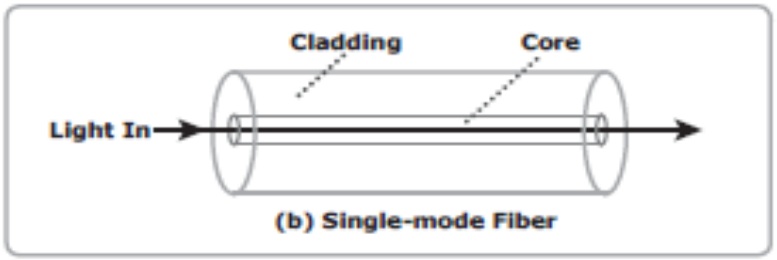


Figure 5-4: single-mode fiber

Components of FC SAN

Cables and Connectors

- Multimode fiber (MMF) cable carries multiple beams of light projected at different angles simultaneously onto the core of the cable.

  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 used for short distances because of signal degradation (attenuation) due to modal dispersion

Based on the bandwidth, multimode fibers are classified as

- OM1 (62.5µm core),
- OM2(50µm core),
- laser-optimized OM3 (50µm core).

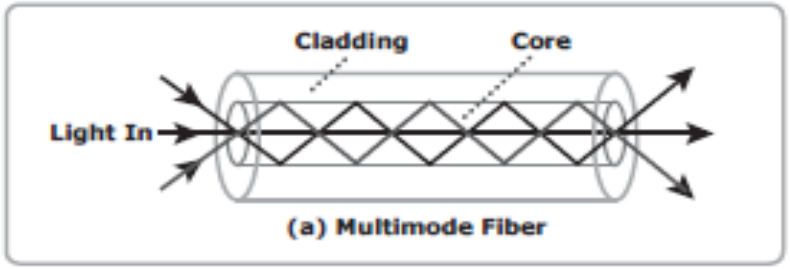


Figure 5-4: Multimode fiber

Components of FC SAN

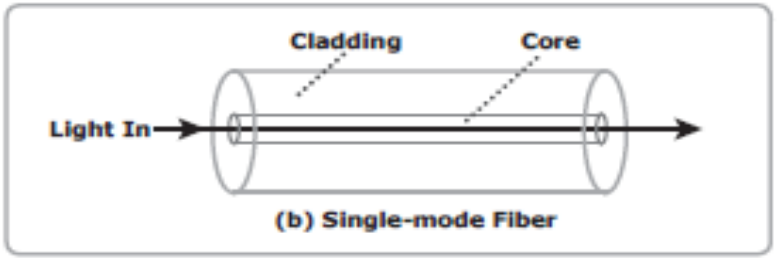


Figure 5-4: single-mode fiber

Cables and Connectors

- Single-mode fiber (SMF) carries a single ray of light projected at the center of the core.
- These cables are available in core diameters of 7 to 11 microns;the most common size is 9 microns.
- The small core and the single light wave help to limit modal dispersion.
- Singlemode provides minimum signal attenuation over maximum distance (up to 10km).
- A single-mode cable is used for long-distance cable runs, and distance usually depends on the power of the laser at the transmitter and sensitivity of the receiver.

The SAN and Its Evolution

Cables and Connectors

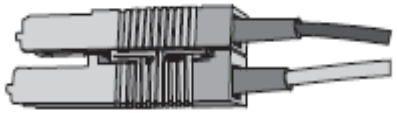
A connector is attached at the end of a cable to enable swift connection and disconnection of the cable to and from a port.

Two commonly used connectors for fiber optic cables.

- Standard connector (SC)
- Lucent connector (LC)

Another fiber-optic connector, used with fiber patch panels

- Straight Tip (ST)



(a) Standard Connector (SC)

(b) Lucent Connector (LC)

(c) Straight Tip Connector (ST)

Components of FC SAN

Interconnect Devices

Commonly used in FC SAN- FC hubs, switches, and directors

- Hubs are used as communication devices in FC-AL implementations
- Switches are more intelligent than hubs and directly route data from one physical port to another
- Directors are high-end switches with a higher port count and better fault tolerance capabilities.

Components of FC SAN

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 *zoning*

RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

FC Connectivity

The FC architecture supports three basic interconnectivity options:

1. point-to point
2. arbitrated loop
3. Fibre Channel switched fabric

FC Connectivity

1. Point-to point

- Point-to-point — two devices are connected directly to each other.
- This configuration provides a dedicated connection for data transmission between nodes.
- However, the point-to-point configuration offers limited connectivity, because only two devices can communicate with each other at a given time.
- Moreover, it cannot be scaled to accommodate a large number of nodes.

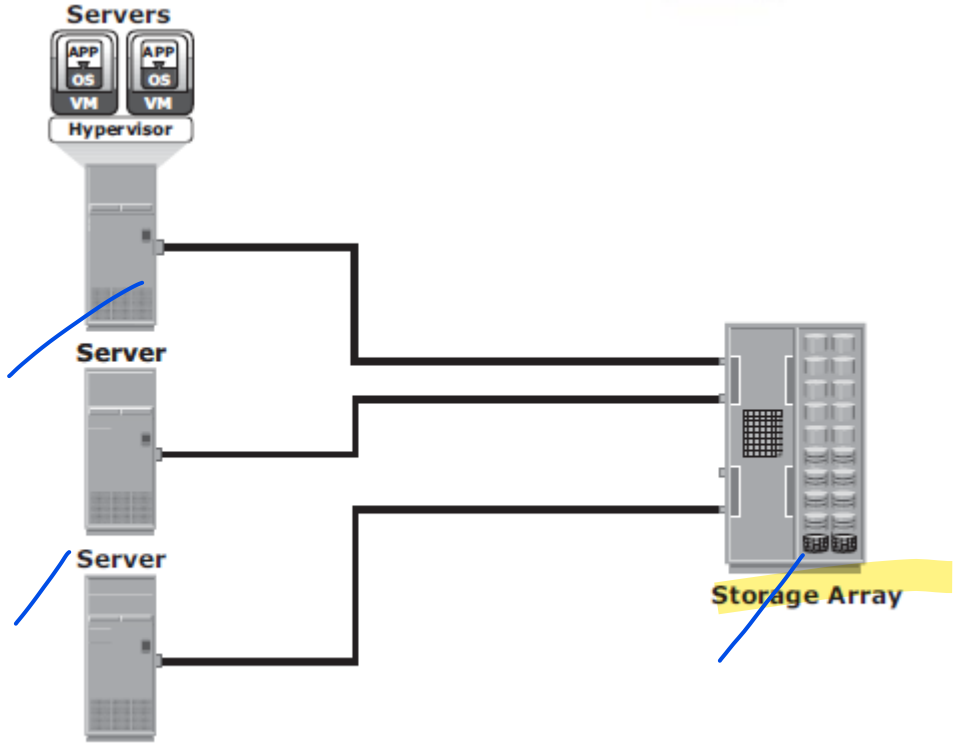


Figure 5-6: Point-to-point connectivity

FC Connectivity

2. Arbitrated loop(AL)

- Devices are attached to a shared loop
- FC-AL has the characteristics of a token ring topology and physical star topology.
- 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

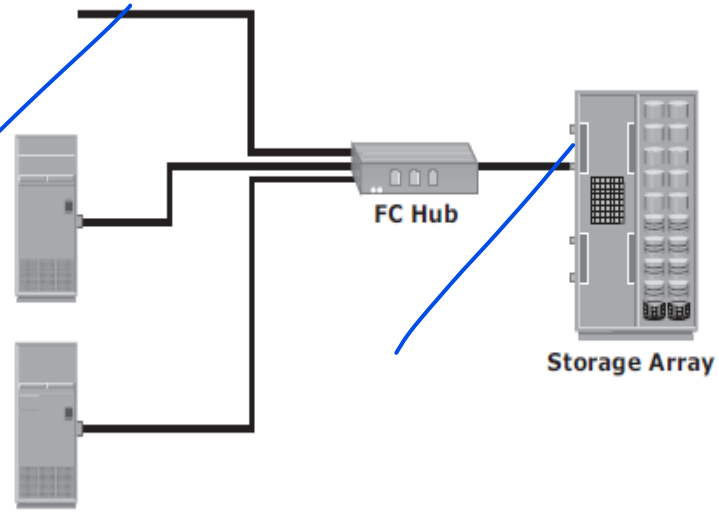
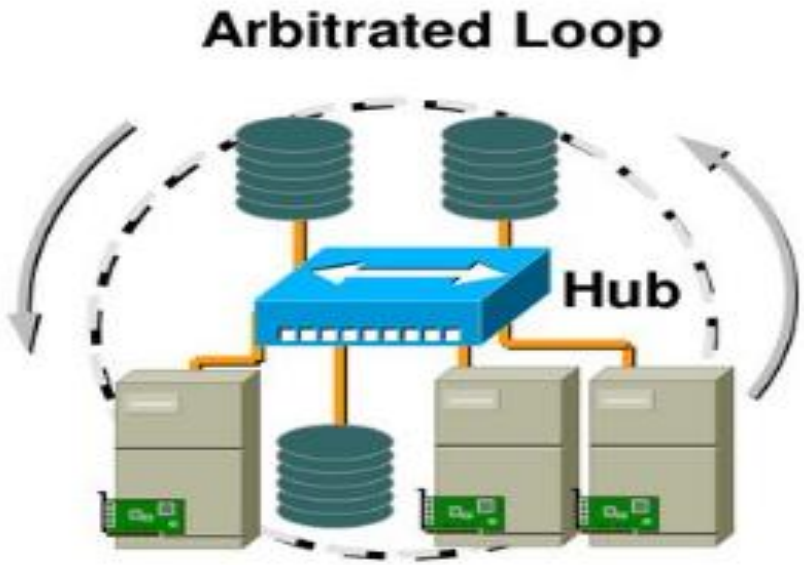
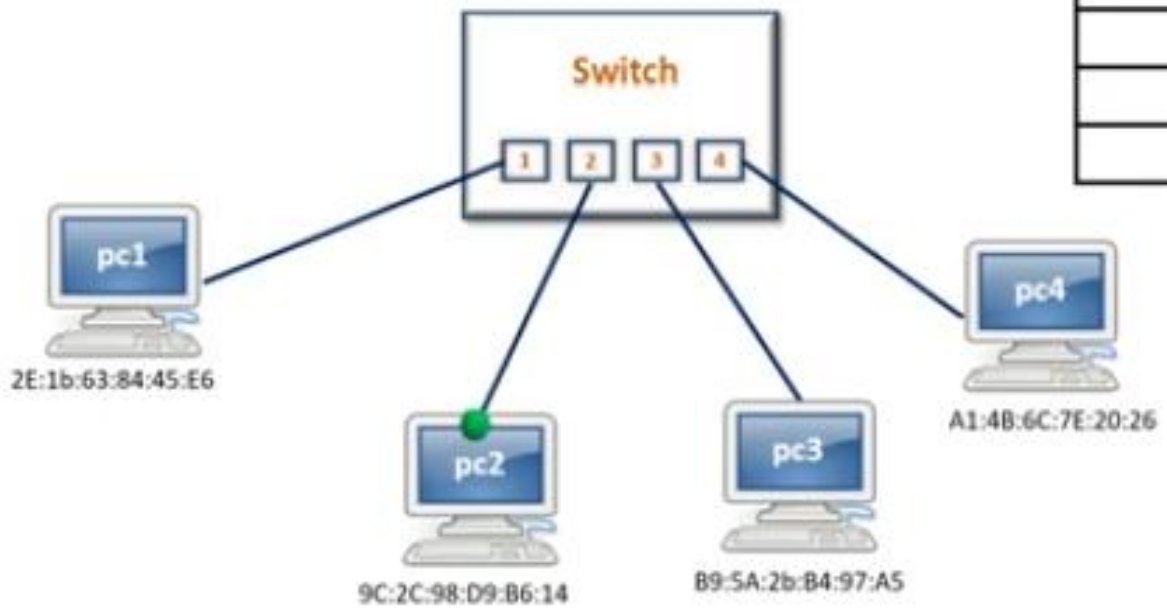


Figure 5-7: Fibre Channel Arbitrated Loop



FC Connectivity

Switch



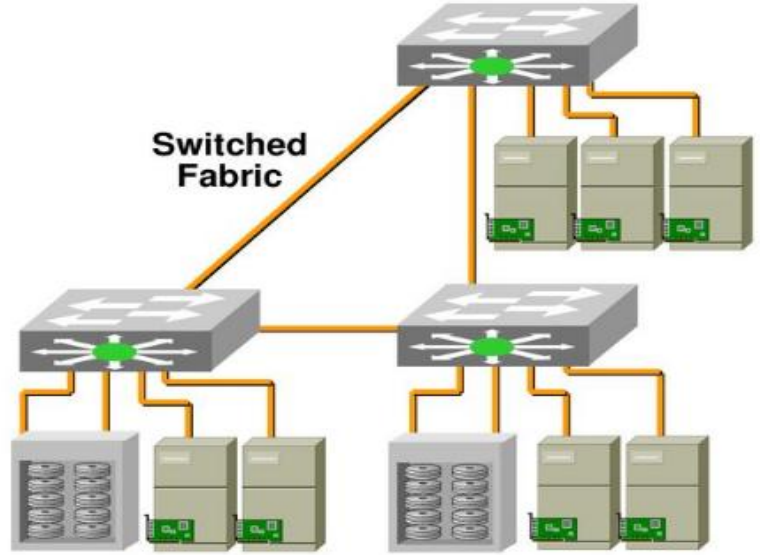
Port Numbers	MAC Addresses
1	2E:1b:63:84:45:E6
2	9C:2C:98:D9:B6:14
3	B9:5A:2b:B4:97:A5
4	A1:4B:6C:7E:20:26

RAMAIAH
Institute of Technology

FC Connectivity

3.Fibre Channel switched fabric(FC-SW)

- 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, 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.



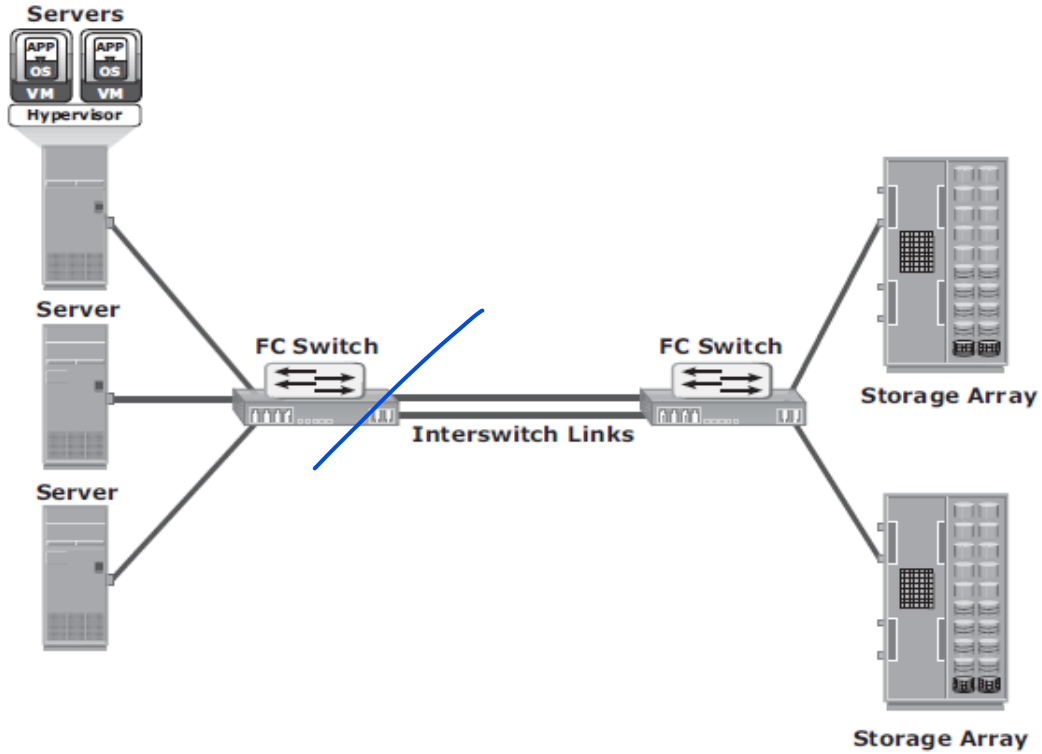


Figure 5-8: Fibre Channel switched fabric

FC Connectivity

3.Fibre Channel switched fabric(FC-SW)

In a switched fabric, the link between any two switches is called an Interswitch link (ISL).

ISLs enable switches to be connected together to form a single, larger fabric.

ISLs are used to transfer host-to-storage data and fabric management traffic from one switch to another.

By using ISLs, a switched fabric can be expanded to connect a large number of nodes.

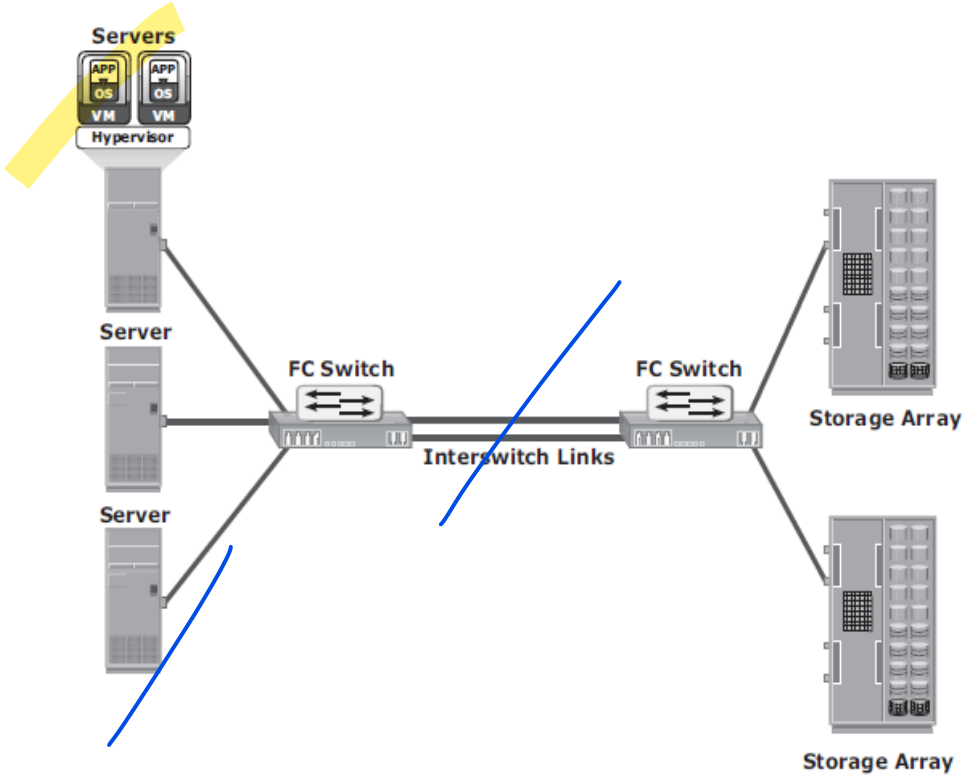


Figure 5-8: Fibre Channel switched fabric

FC Connectivity

FC-SW Transmission

- FC-SW uses switches that can switch data traffic between nodes directly through switch ports.
- Frames are routed between source and destination by the fabric.
- As shown in Figure 5-10, if node B wants to communicate with node D, the 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.

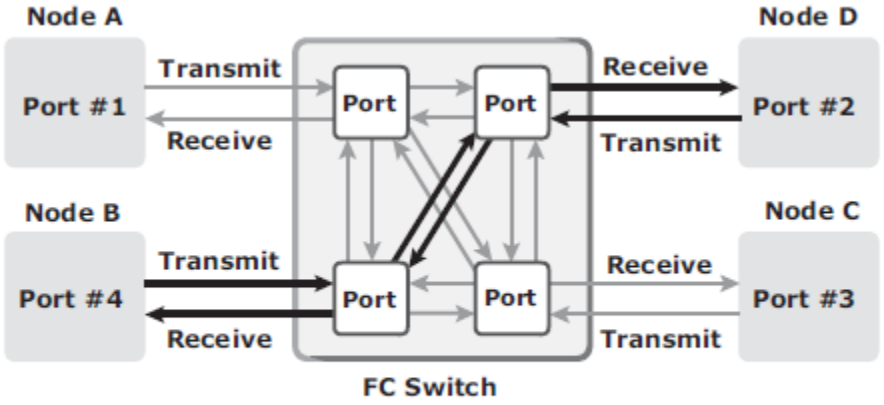


Figure 5-10: Data transmission in Fibre Channel switched fabric

RAMAIAH Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

discuss the different port types in a switched fabric.
Explain F-Port and N-Port

Switched Fabric Ports

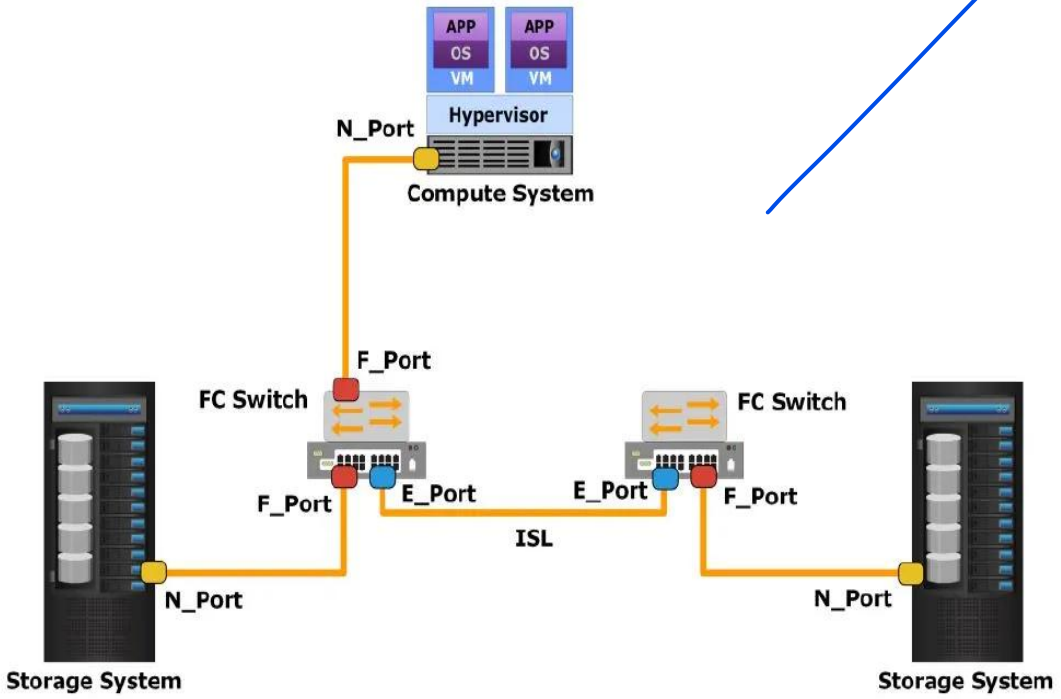
Ports in a switched fabric can be one of the following types:

N_Port: An end point in the fabric. This port is also known as the node port.

It is a host port (HBA) or a storage array port connected to a switch in a switched fabric.

E_Port: A 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 ISLs.



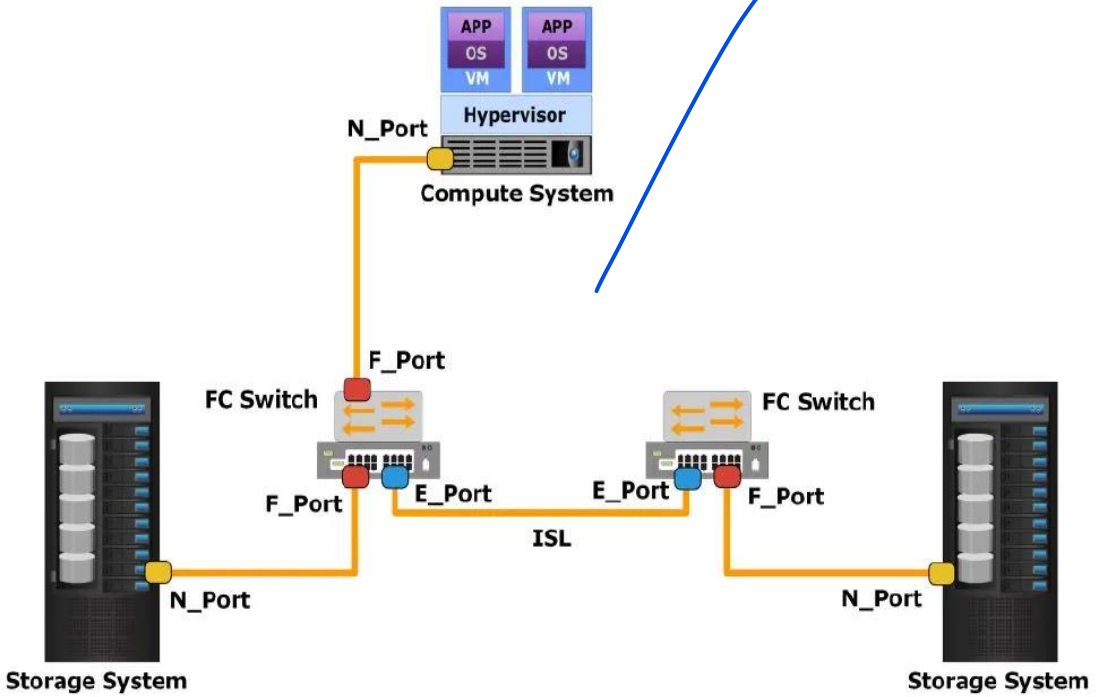
Switched Fabric Ports

Ports in a switched fabric can be one of the following types:

F_Port: A port on a switch that connects an N_Port.

It is also known as a fabric port.

G_Port: A generic port on a switch that can operate as an E_Port or an F_Port and determines its functionality automatically during initialization.



RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

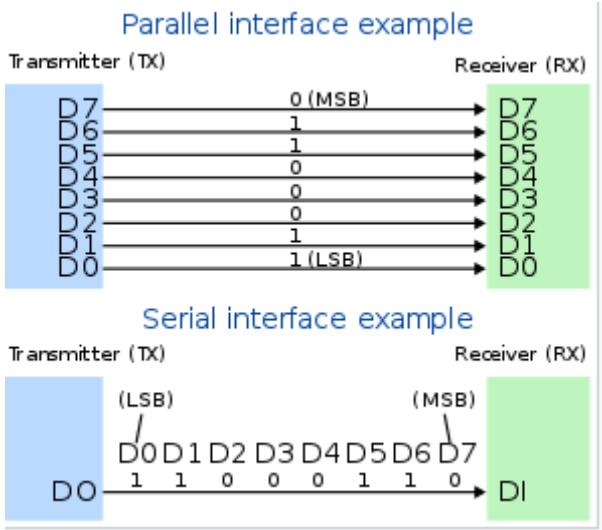
- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Fibre Channel Architecture

- **Fibre Channel Protocol Stack**
  - ***FC-4 Layer***
  - ***FC-2 Layer***
  - ***FC-1 Layer***
  - ***FC-0 Layer***
- **Fibre Channel Addressing**
- **World Wide Names**
- **FC Frame**
- **Structure and Organization of FC Data**
- **Flow Control**
- **Classes of Service**

Fibre Channel Architecture

- FC SAN uses the Fibre Channel Protocol (FCP) that provides both channel speed for data transfer with low protocol overhead and scalability of network technology.
- FCP forms the fundamental construct of the FC SAN infrastructure.
- Fibre Channel provides a serial data transfer interface that operates over copper wire and optical fiber.
- FCP is the implementation of serial SCSI over an FC network.
- In FCP architecture, all external and remote storage devices attached to the SAN appear as local devices to the host operating system.



RAMAIAH Institute of Technology

Fibre Channel Architecture

Fibre Channel Protocol Stack

FCP defines the communication protocol in five layers:

- *FC-4 Layer*
- *FC-2 Layer*
- *FC-1 Layer*
- *FC-0 Layer*

FC-0 through FC-4 (except FC-3 layer, which is not implemented). In a layered communication model

Figure 5-12 illustrates the Fibre Channel protocol stack.

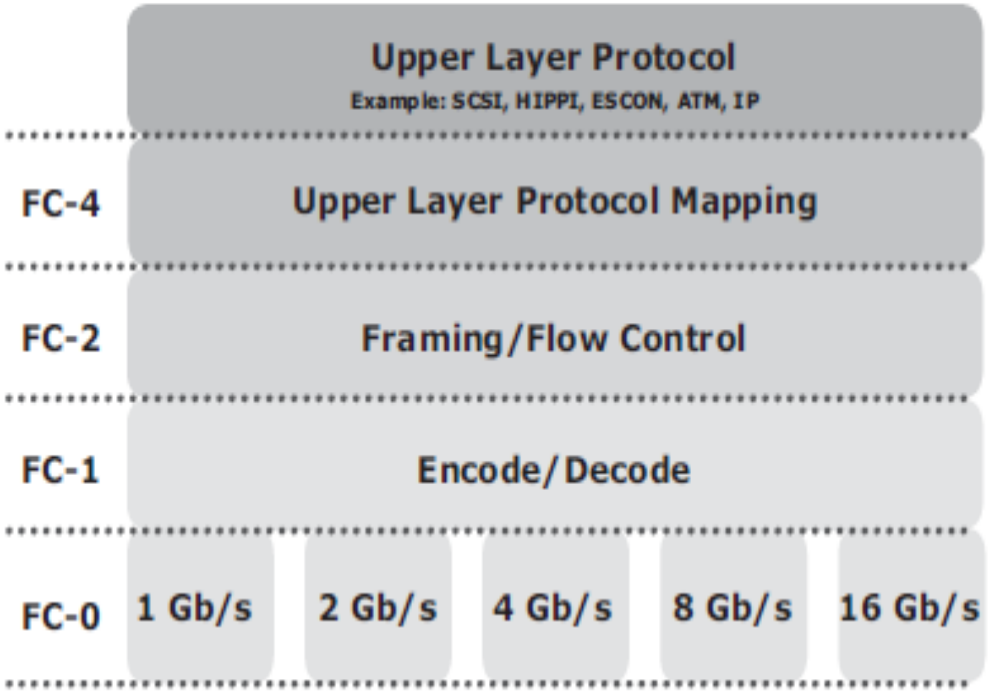


Figure 5-12: Fibre Channel protocol stack

Fibre Channel Architecture

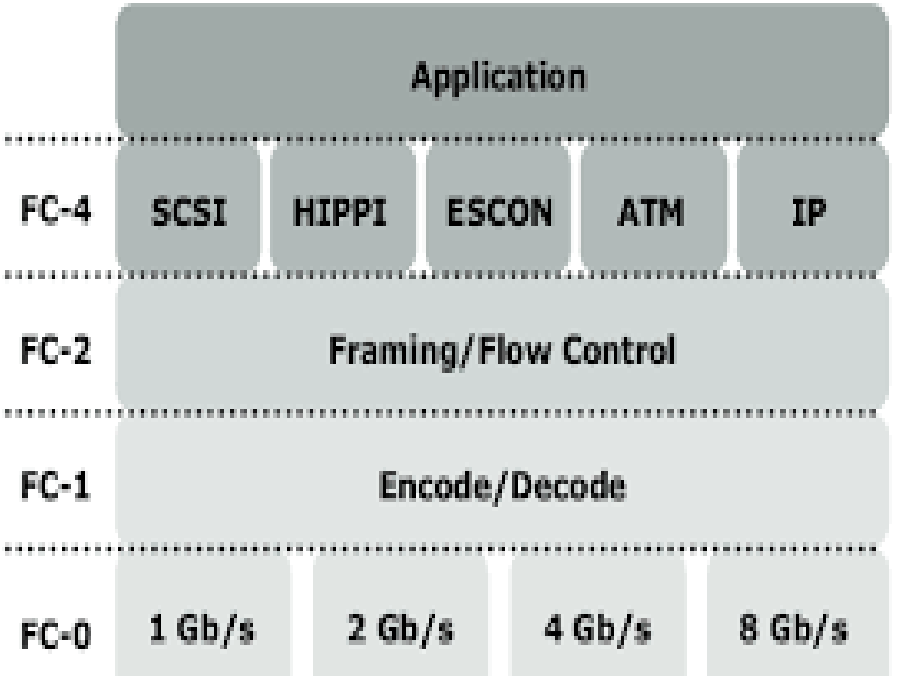
FC-4 Layer

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 . Some of the protocols include

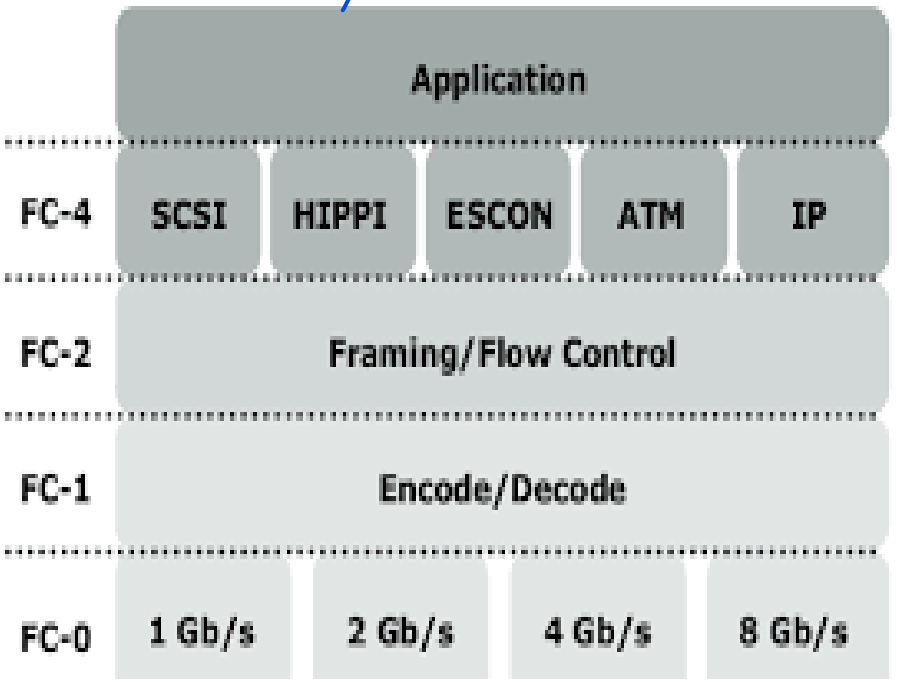
- SCSI,
- High Performance Parallel Interface (HIPPI) Framing Protocol, Enterprise Storage Connectivity (ESCON),
- Asynchronous Transfer Mode (ATM),
- IP.



Fibre Channel Architecture

FC- 2 Layer

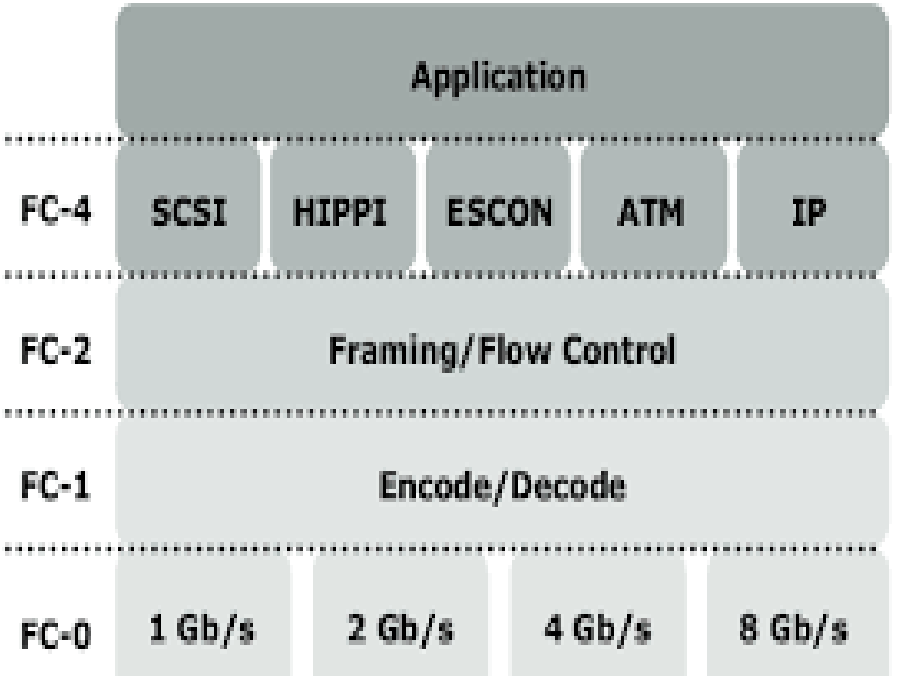
- The FC-2 layer provides Fibre Channel addressing, structure, and organization of data (frames, sequences, and exchanges).
- It also defines fabric services, classes of service, flow control, and routing.



Fibre Channel Architecture

FC- 1 Layer

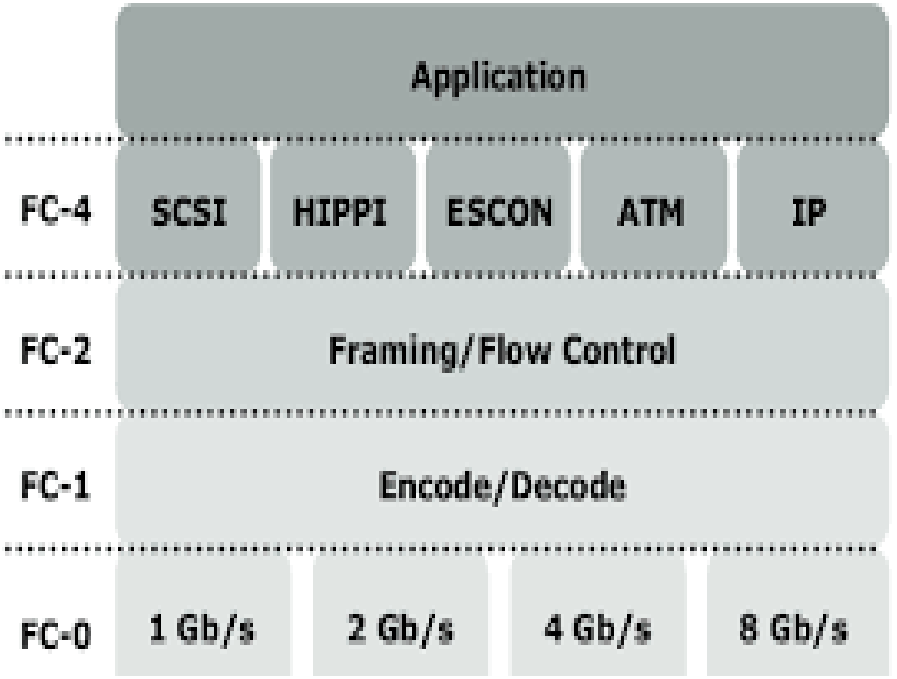
- The FC-1 layer defines how data is encoded prior to transmission and decoded upon receipt.
  - At the transmitter node, an 8-bit character is encoded into a 10-bit transmissions character. This character is then transmitted to the receiver node.
  - At the receiver node, the 10-bit character is passed to the FC-1 layer, which decodes the 10-bit character into the original 8-bit character.
- FC links with speeds of 10 Gbps and above use 64-bit to 66-bit encoding algorithms.
- The FC-1 layer also defines the transmission words, such as FC frame delimiters, which identify the start and end of a frame and primitive signals that indicate events at a transmitting port.
- FC-1 layer performs link initialization and error recovery.



Fibre Channel Architecture

FC- 0 Layer

- FC-0 is the lowest layer in the FCP stack.
- This layer defines the physical interface,media, and transmission of 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.



RAMAIAH Institute of Technology

Fibre Channel Architecture

Fibre Channel Addressing

An FC address is dynamically assigned when a node port logs on to the fabric.

The FC address has a distinct format, as shown in Figure 5-13.

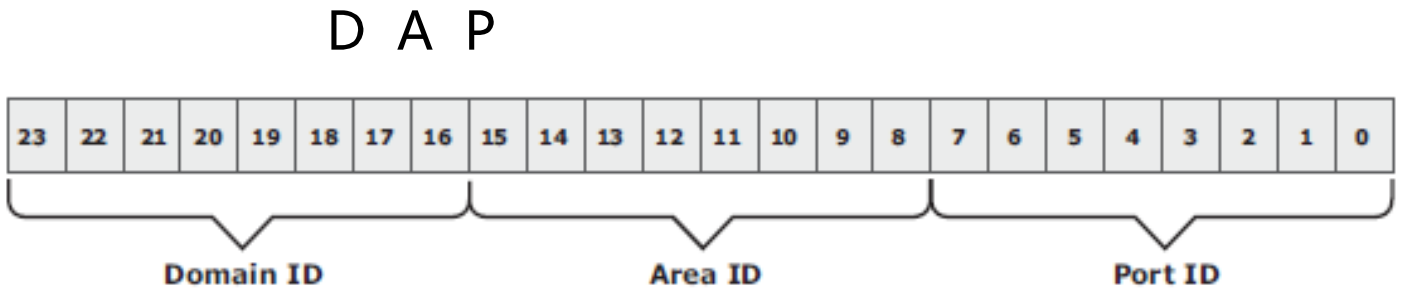


Figure 5-13: 24-bit FC address of N_Port

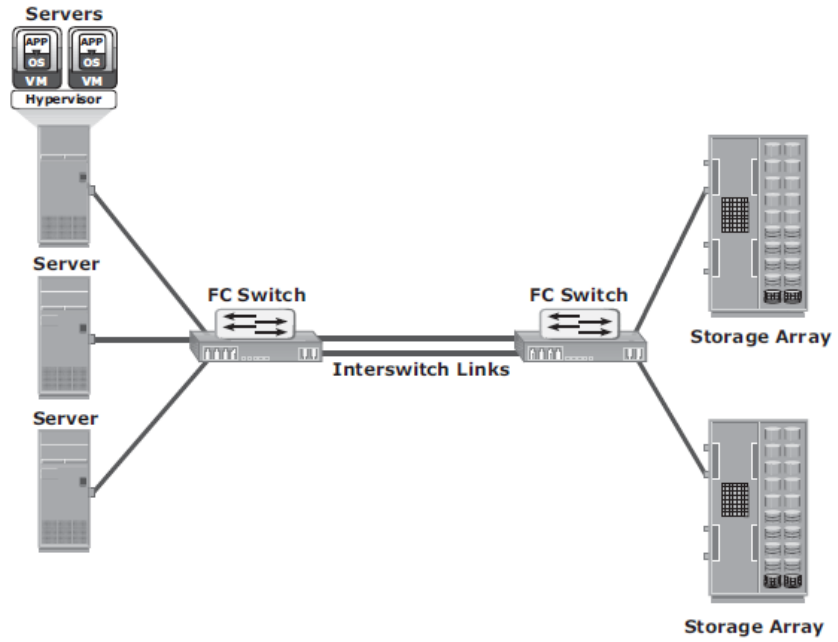


Figure 5-8: Fibre Channel switched fabric

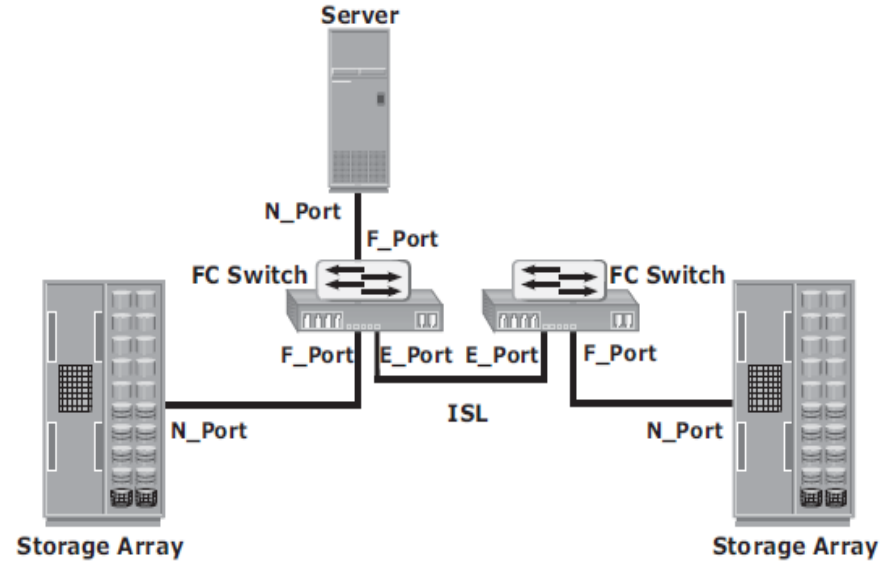


Figure 5-11: Switched fabric ports

RAMAIAH Institute of Technology

Fibre Channel Architecture

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)
  - World Wide Port Name (WWPN).

WWN is a static name for each node on an FC network.

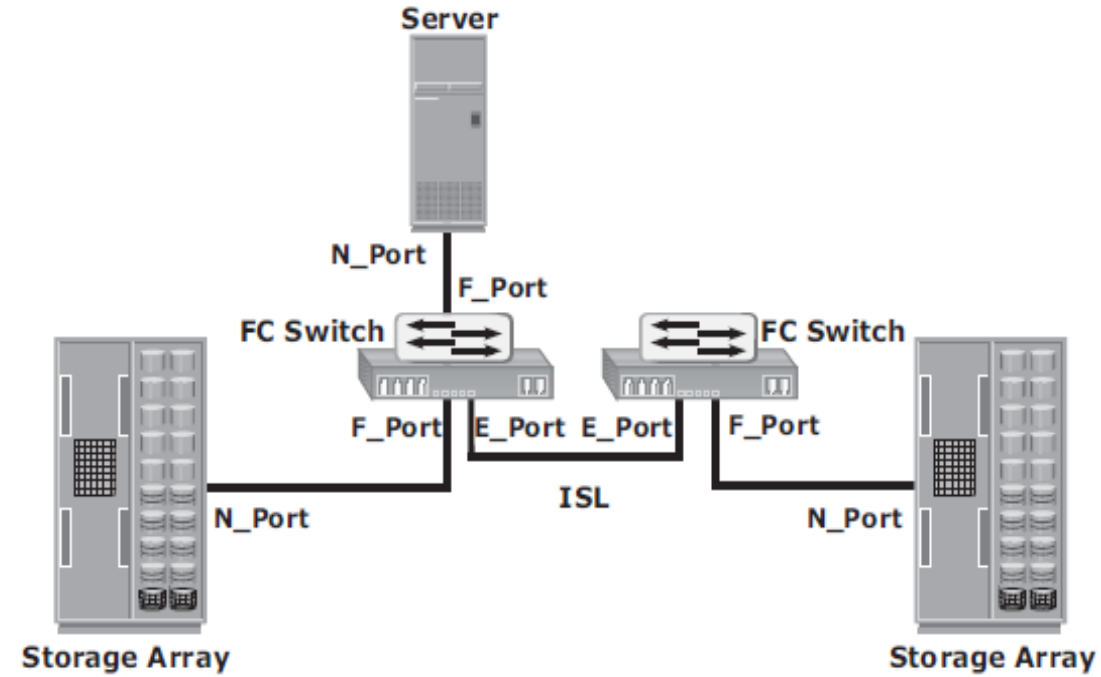


Figure 5-11: Switched fabric ports

RAMAIAH Institute of Technology

Fibre Channel Architecture

World Wide Names

- 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.
- Figure 5-14 illustrates the WWN structure examples for an array and an 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
Format Type	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
Format Type	Reserved 12 bits			Company ID 24 bits						Company Specific 24 bits					

Figure 5-14: World Wide Names

RAMAIAH Institute of Technology

Fibre Channel Architecture

FC Frame

An FC frame consists of five parts:

1. start of frame (SOF),
2. frame header,
3. data field,
4. cyclic redundancy check (CRC),
5. end of frame (EOF).

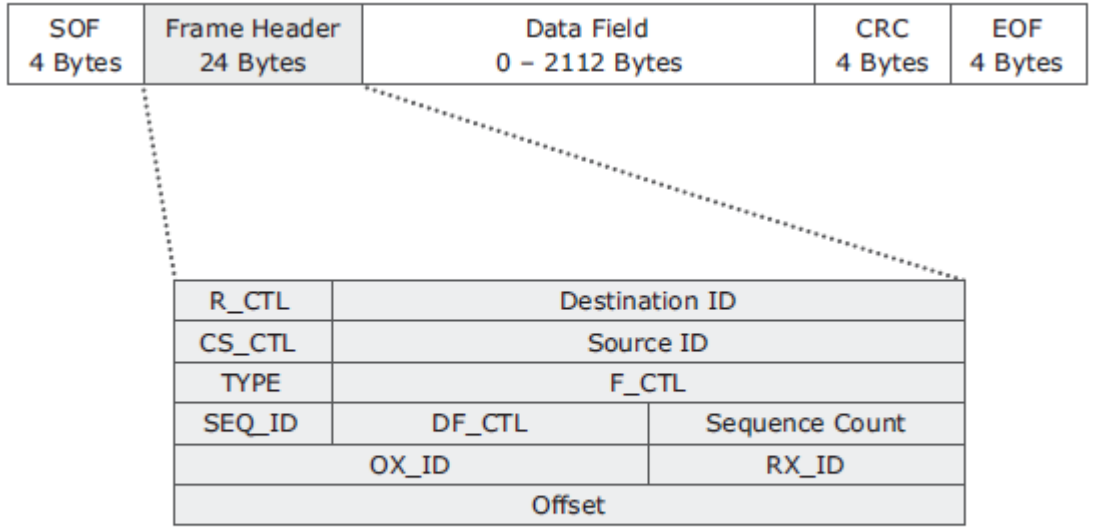


Figure 5-15: FC frame

Fibre Channel Architecture

FC Frame

- The SOF and EOF act as delimiters.
- SOF also indicates whether the frame is the first frame in a sequence of frames.
- The data field in an FC frame contains the data payload, up to 2,112 bytes of actual data with 36 bytes of fixed overhead.
- The CRC checksum facilitates error detection for the content of the frame.

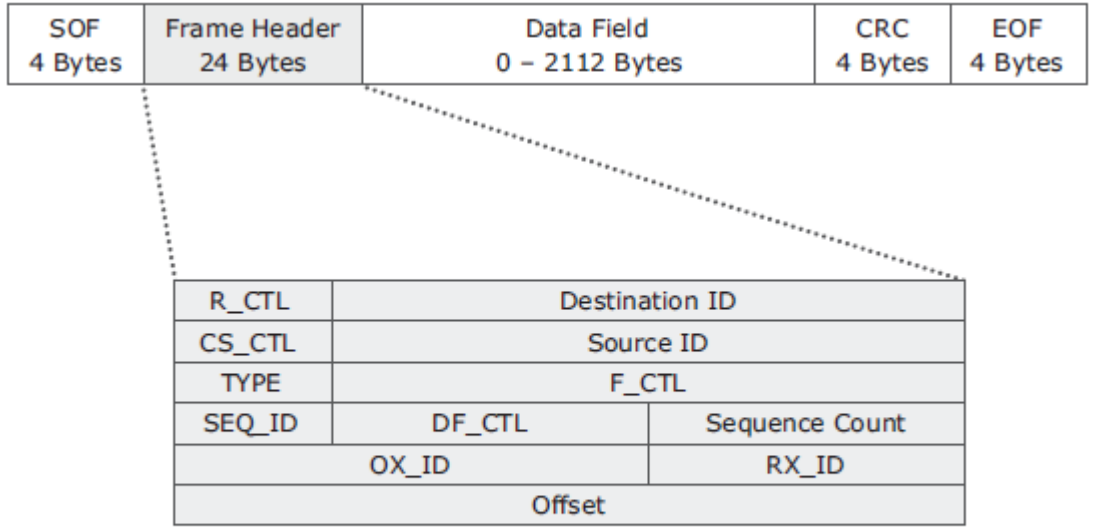


Figure 5-15: FC frame

Fibre Channel Architecture

FC Frame

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),
- Responder Exchange ID (RX_ID),

Additionally control fields.

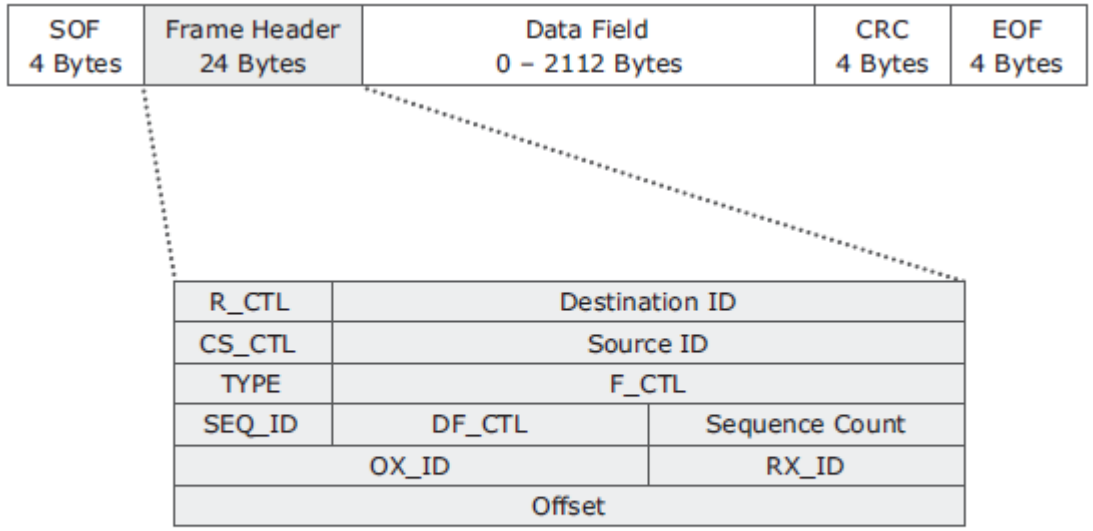


Figure 5-15: FC frame

Fibre Channel Architecture

FC Frame

The frame header 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 frames that do not carry any user data. These frames are used for setup and messaging.

Data frames carry the user data.

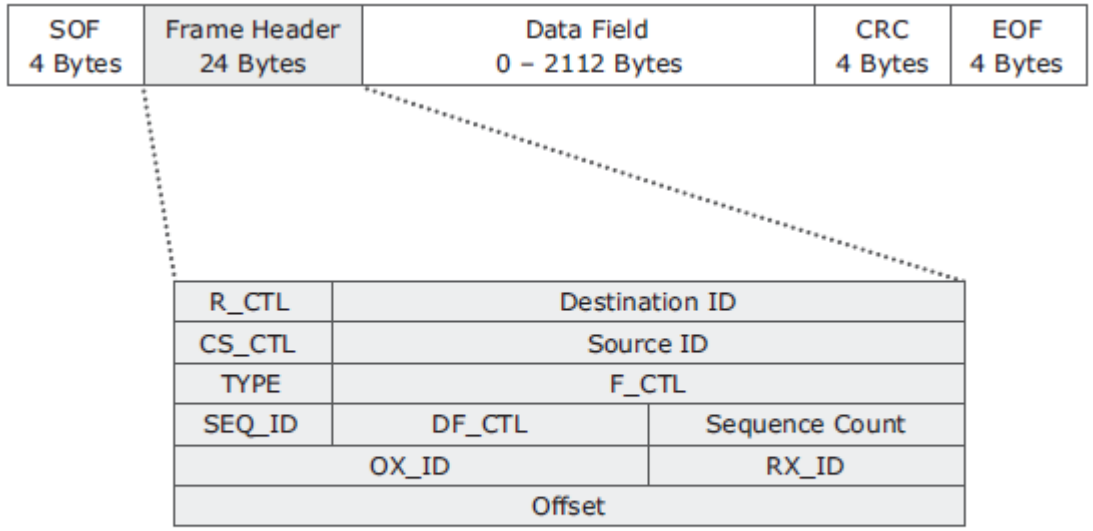
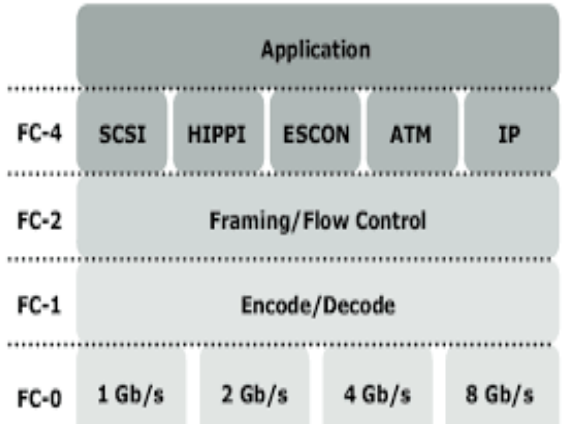


Figure 5-15: FC frame

RAMAIAH Institute of Technology

Fibre Channel Architecture



FC Frame

The frame header also defines the following fields:

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.

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.

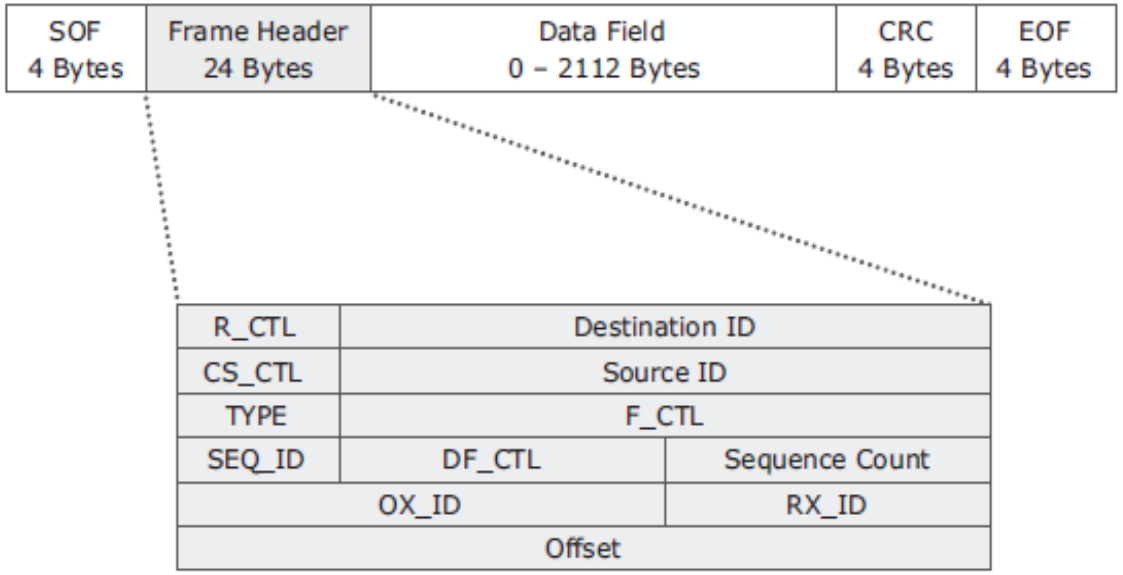


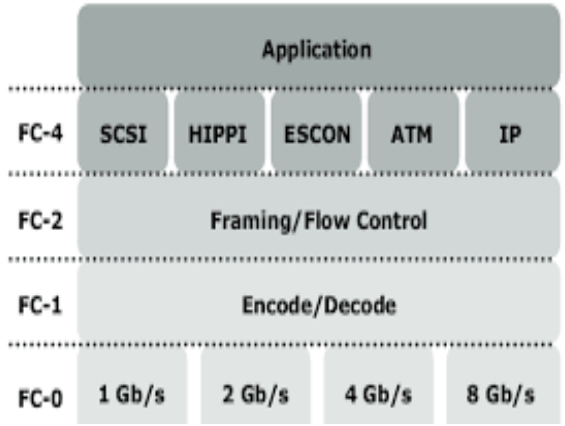
Figure 5-15: FC frame

Fibre Channel Architecture

Structure and Organization of FC Data

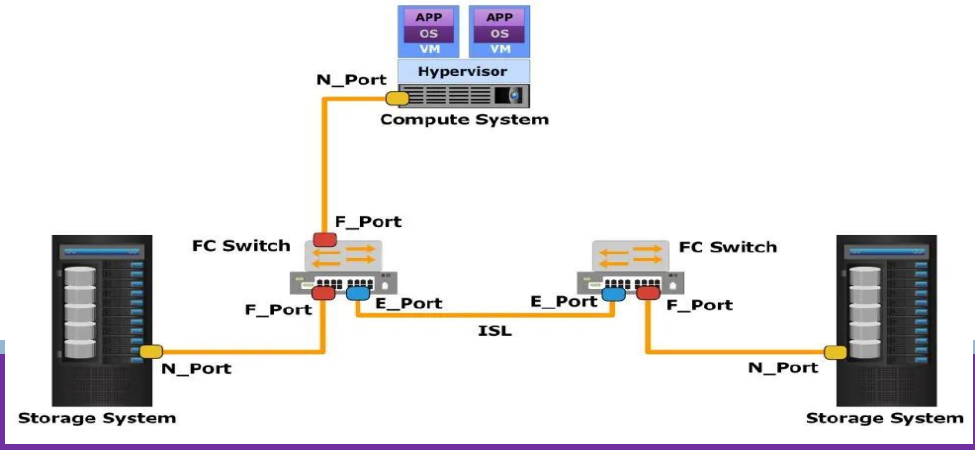
- In an FC network, data transport is analogous to a conversation between two people,
  - frame represents a word,
  - sequence represents a sentence,
  - exchange represents a conversation.

Fibre Channel Architecture



Structure and Organization of FC Data

- Exchange:
  - An exchange operation enables two node ports to identify and manage a set of information units.
  - Each upper layer protocol has its protocol-specific information that must be sent to another port to perform certain operations.
  - This protocol-specific information is called an information unit. The structure of these information units is defined in the FC-4 layer.
  - This unit maps to a sequence.
  - An exchange is composed of one or more sequences.



Fibre Channel Architecture

Structure and Organization of FC Data

- 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.

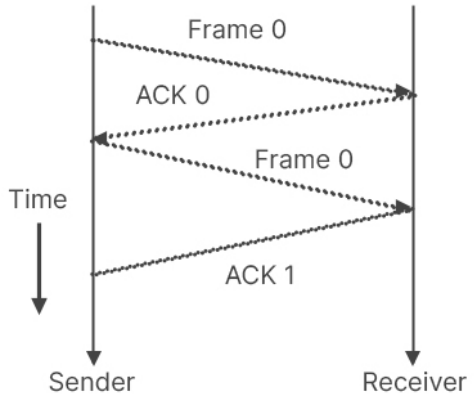
Fibre Channel Architecture

Flow Control

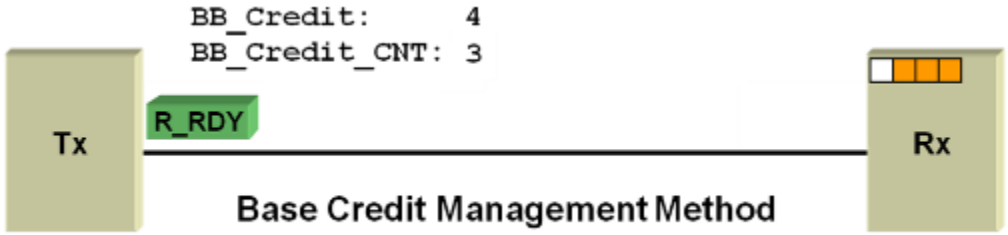
Flow control defines the pace of the flow of data frames during data transmission.

FC technology uses two flow-control mechanisms:

1. buffer-to-buffer credit (BB_Credit)
2. end-to-end credit (EE_Credit).



Fibre Channel Architecture



Flow Control

1. buffer-to-buffer credit (BB_Credit)

- FC uses the BB_Credit mechanism for 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.

RAMAIAH Institute of Technology

Fibre Channel Architecture

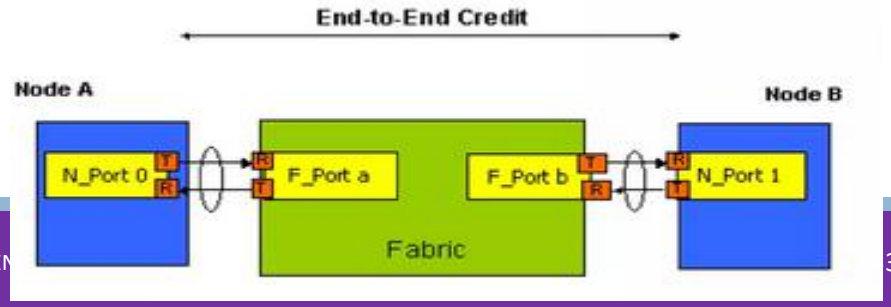
Flow Control

2. end-to-end credit (EE_Credit)

It is the maximum number of data frames a source port can send to destination port without receiving an acknowledgment(ACK).

This credit is granted during N_Port login and is replenished with the return of an ACK Response.

The EE_Credit mechanism provides the flow control for class 1 and class 2 traffic only.



RAMAIAH Institute of Technology

Fibre Channel Architecture

Classes of Service

The FC standards define different classes of service to meet the requirements of a wide range of applications.

Table shows three classes of services and their features.

	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 acknowledgment	Acknowledged	Acknowledged	Not acknowledged
Multiplexing	No	Yes	Yes
Bandwidth utilization	Poor	Moderate	High

RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Fabric Services

All FC switches provide a common set of services as defined in the Fibre Channel standards.

These services are available at certain predefined addresses.

Some of these services are

- Fabric Login Server
- Fabric Controller
- Name Server
- Management Server

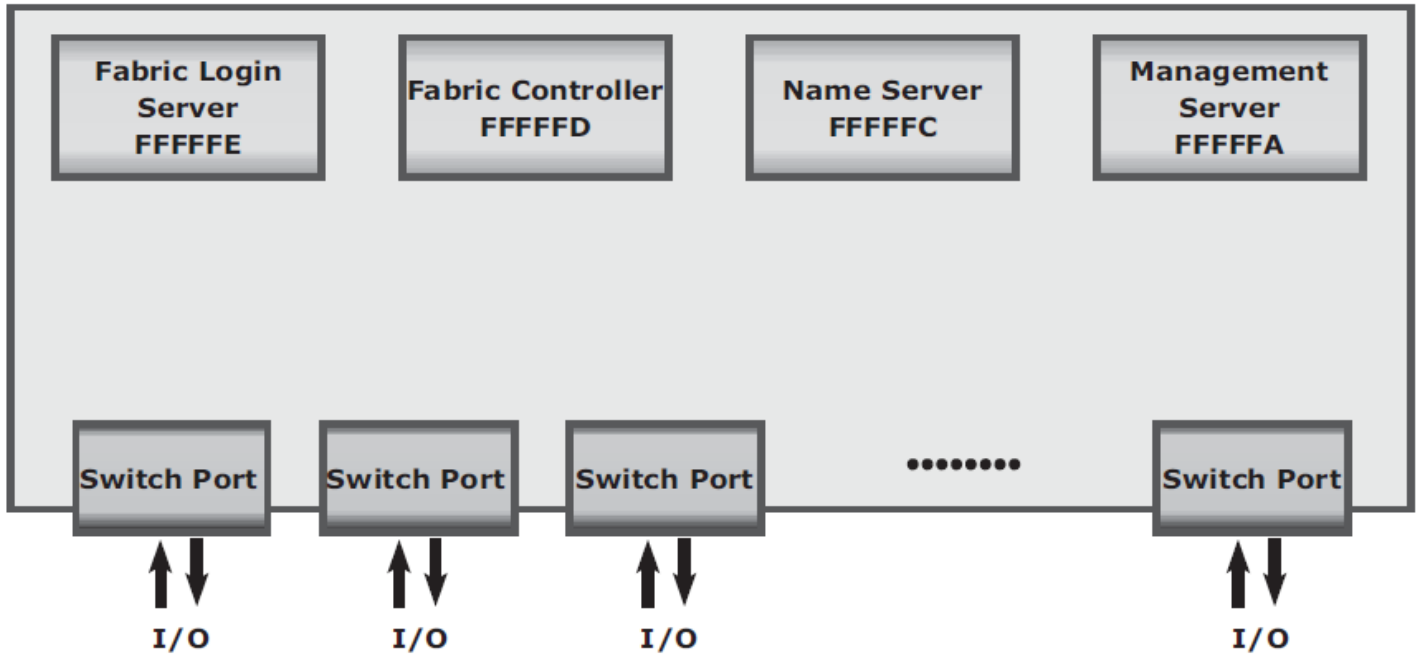


Figure 5-16: Fabric services provided by FC switches

RAMAIAH Institute of Technology

Fabric Services

Fabric Login Server

The Fabric Login Server is located at the predefined address of FFFFFE and is used during the initial part of the node's fabric login process.

Name Server

The Name Server (Distributed Name Server) is located at the predefined address FFFFFC and is responsible for name registration and management of node ports.

Each switch exchanges its Name Server information with other switches in the fabric to maintain a synchronized, distributed name service.

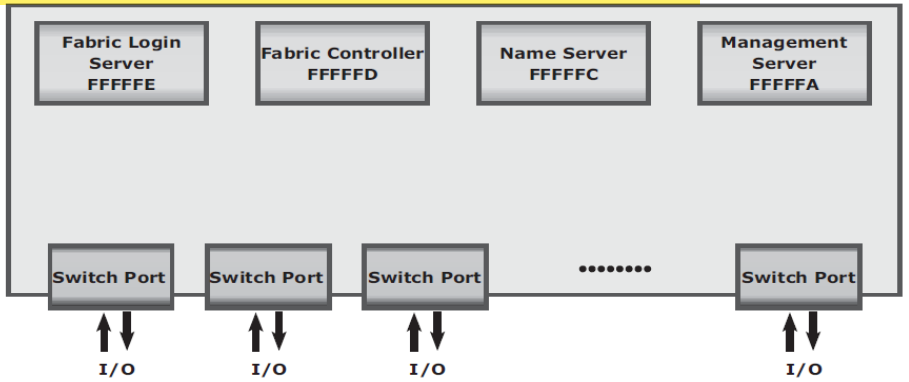


Figure 5-16: Fabric services provided by FC switches

Fabric Services

Fabric Controller

Each switch has a Fabric Controller located at the predefined address FFFFFD.

The Fabric Controller provides services to both node ports and other switches.

The Fabric Controller is responsible for managing and distributing Registered State Change Notifications (RSCNs) to the node ports registered with the Fabric Controller.

If there is a change in the fabric, RSCNs are sent out by a switch to the attached node ports.

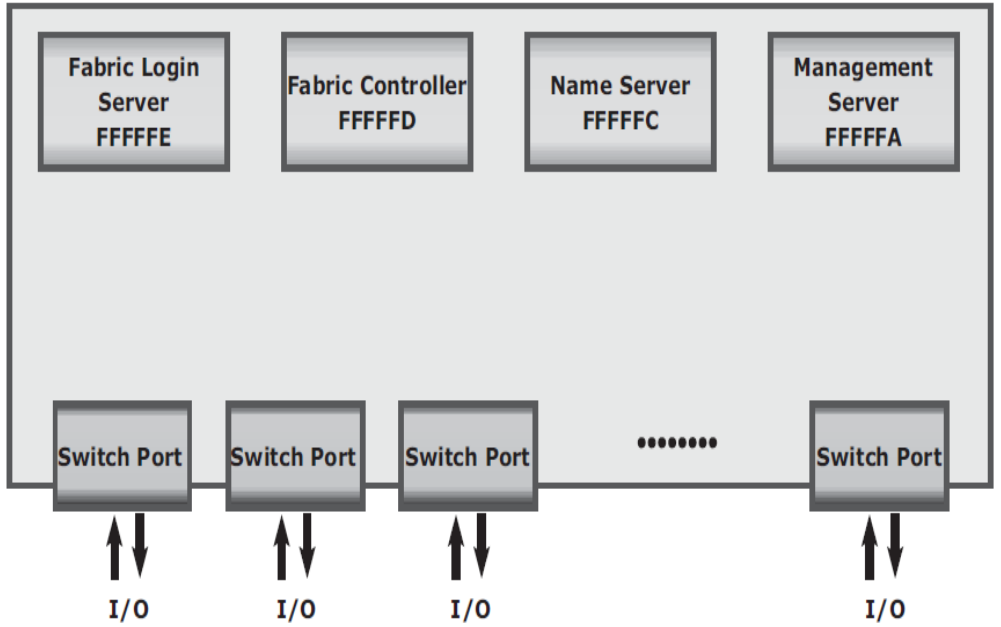


Figure 5-16: Fabric services provided by FC switches

Fabric Services

Fabric Controller

The Fabric Controller also generates Switch Registered State Change Notifications (SW-RSCNs) to every other domain (switch) in the fabric.

These RSCNs keep the name server up-to-date on all switches in the fabric.

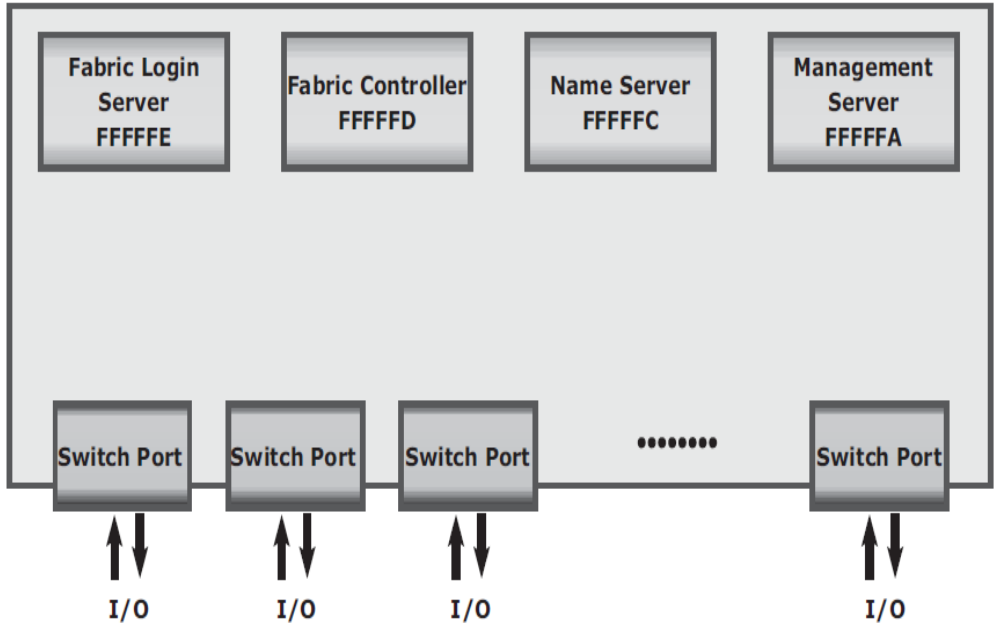


Figure 5-16: Fabric services provided by FC switches

RAMAIAH Institute of Technology

Fabric Services

Management Server

- FFFFFA is the Fibre Channel address for the Management Server.
- The Management Server is distributed to every switch within the fabric.
- The Management Server enables the FC SAN management software to retrieve information and administer the fabric.

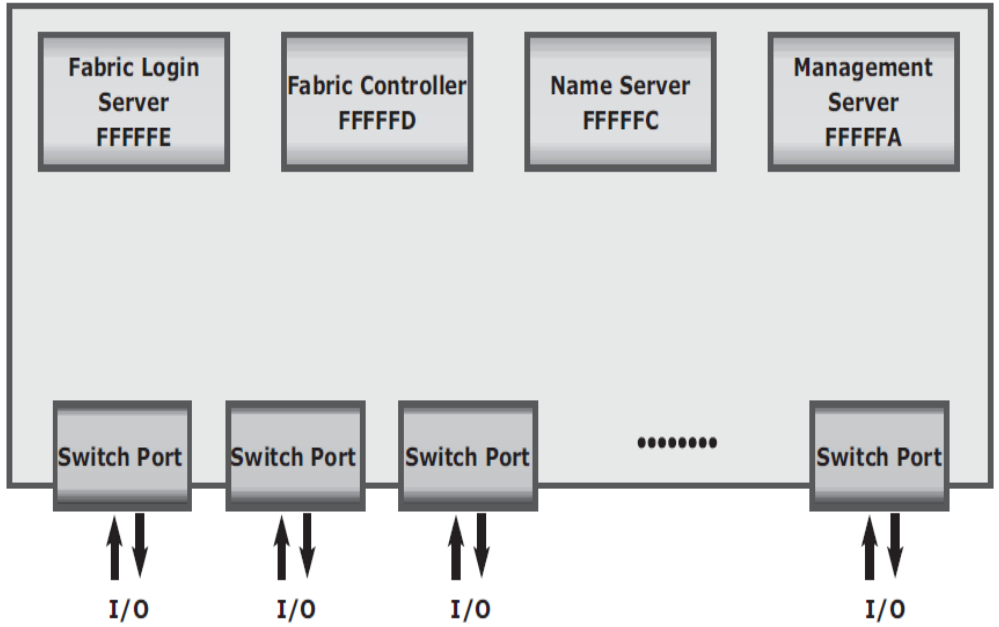


Figure 5-16: Fabric services provided by FC switches

RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Switched Fabric Login Types

Device Login is a series of process which happens when a new node or device is connected to a Fabric.

Two steps

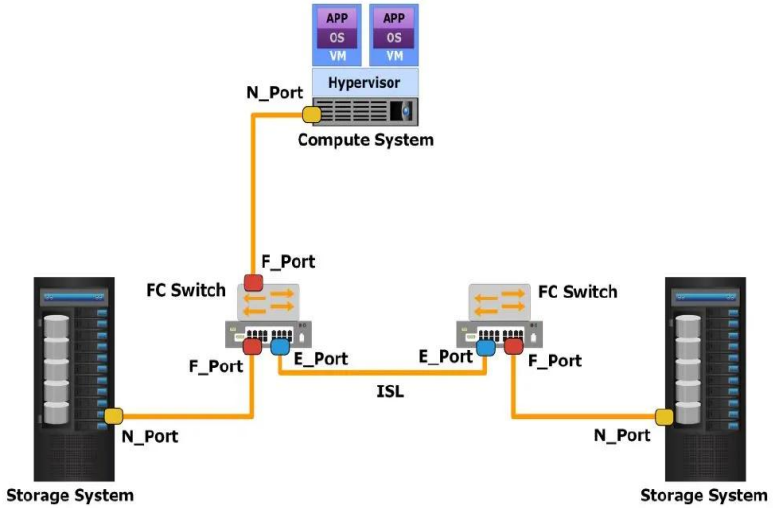
- Fabric login (FLOGI)
- Port login (PLOGI)
- Process login (PRLI)

* Host – These are the server which are connected to Switch
* Array – These are the storage Array connected to Switch
* SAN Switch – A SAN switch having FC Ports
* Any other device which connects to SAN switch with FC protocol.

RAMAIAH
Institute of Technology

Switched Fabric Login Types

* **FLOGI Process** – When a new device is connected to switch or fabric. The device send it first frame to the switch where it is connected.
* The first frame contains some information like WWPN and buffer to buffer credits (B2B credits)
* The Switch assign a 24 bit Unique address to device called as FCID.
* The device also confirms FCID allocation. 0x120100



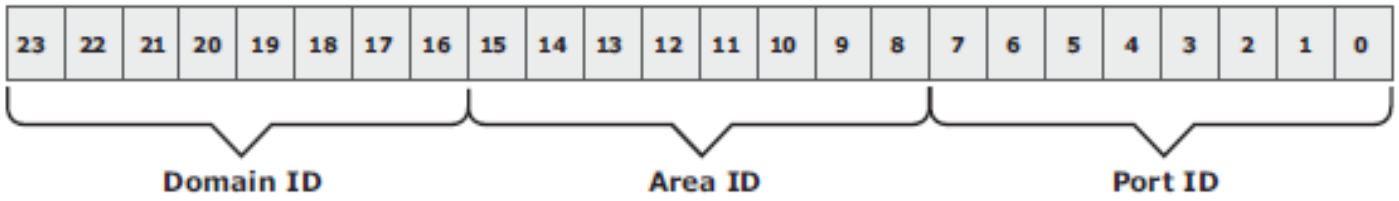


Figure 5-13: 24-bit FC address of N_Port

RAMAIAH Institute of Technology

Switched Fabric Login Type

Fabric login (FLOGI):

- Performed between an N_Port and an F_Port.
- To log on to the fabric, a node sends a FLOGI frame with the WWNN and WWPN parameters to the login service at the predefined FC address FFFFFE (Fabric Login Server).
- Switch accepts the login and returns an Accept (ACC) frame with the assigned FC address for the node.

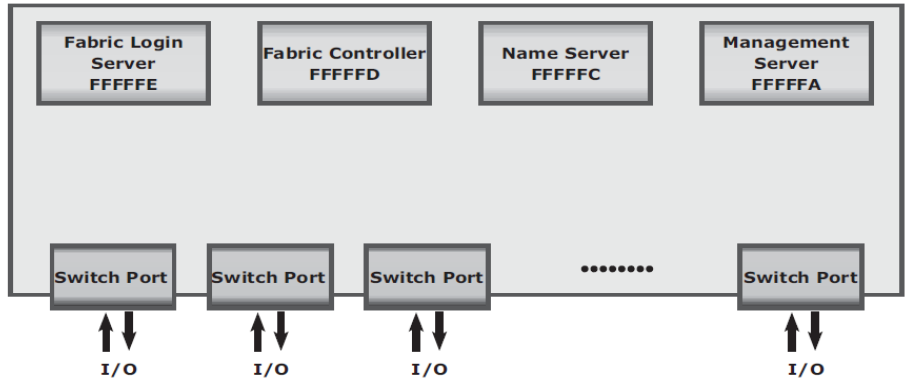
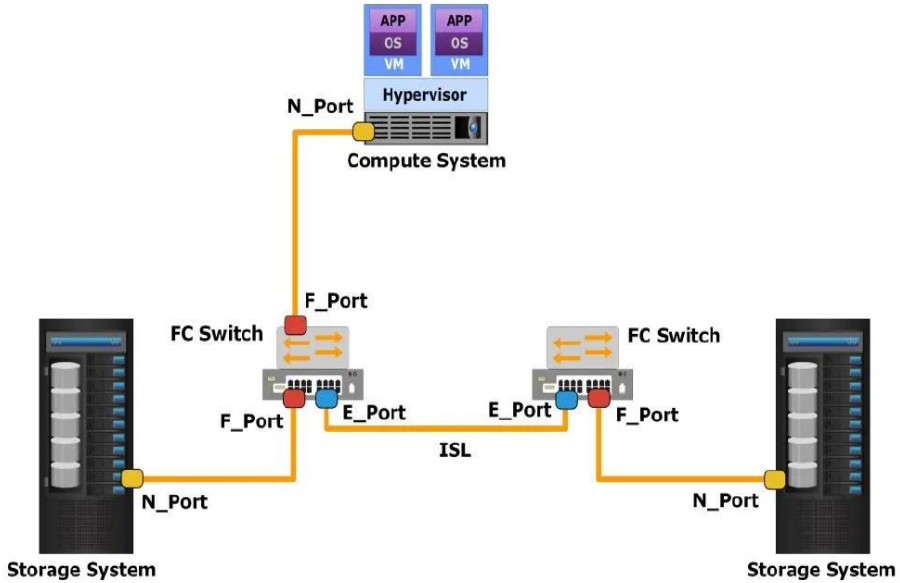


Figure 5-16: Fabric services provided by FC switches



Switched Fabric Login Types

* **PLOGI** – In Plogi two process happens port initialization and registration in Name Server.
* During Initialization information like port type and port speed are exchanged.
* During Name Server registration the node register its information in Name Server of the Fabric. This provides information about all other device registered in SAN fabric that it can communicate to.

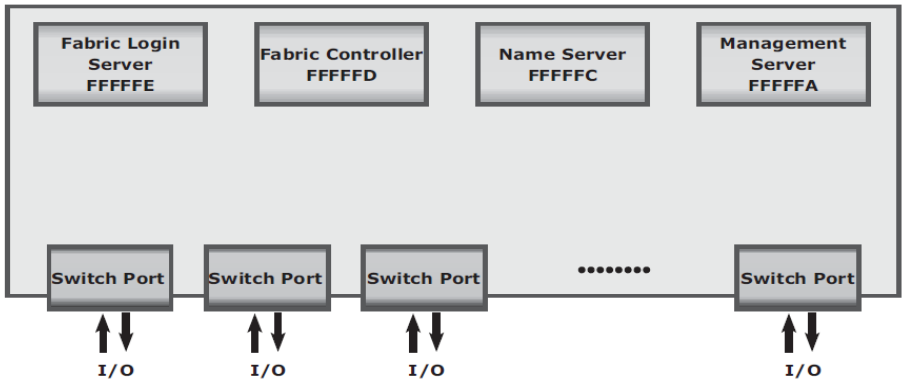
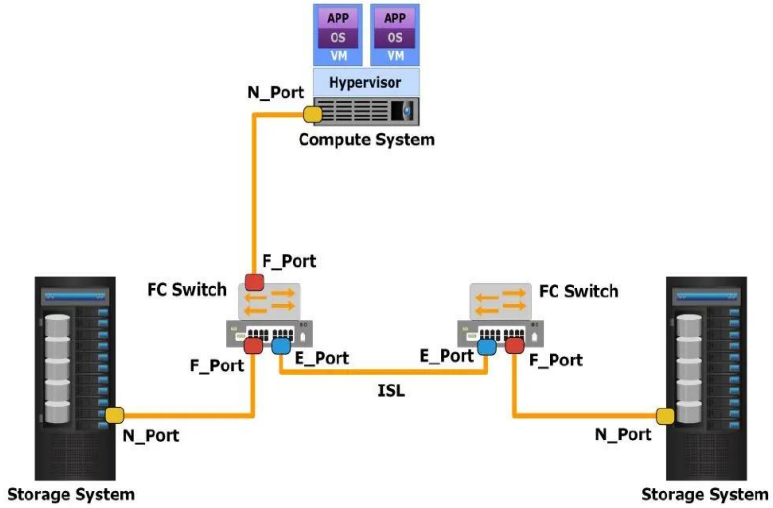
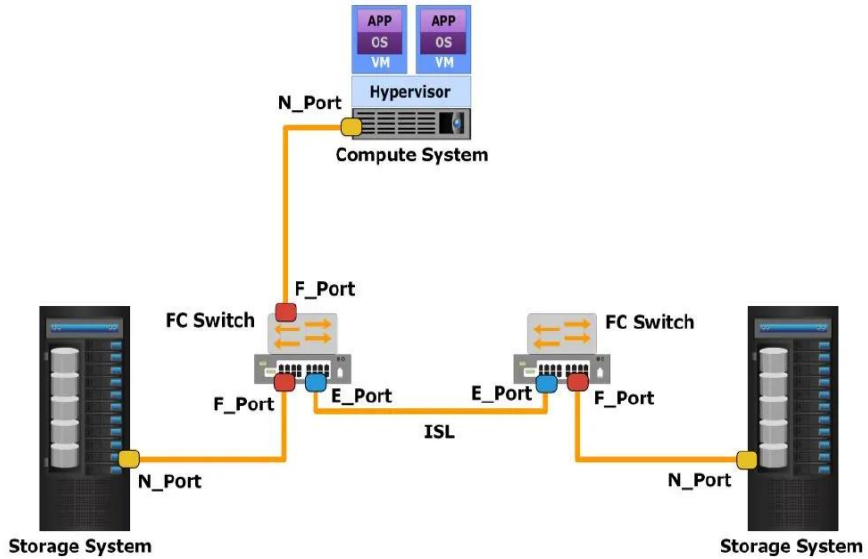


Figure 5-16: Fabric services provided by FC switches



Switched Fabric Login Types



Port login (PLOGI):

- Performed between two N_Ports 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):

- Also performed between two N_Ports.
- This login relates to the FC-4 ULPs, such as SCSI. If the ULP is SCSI, N_Ports exchange
- SCSI-related service parameters.

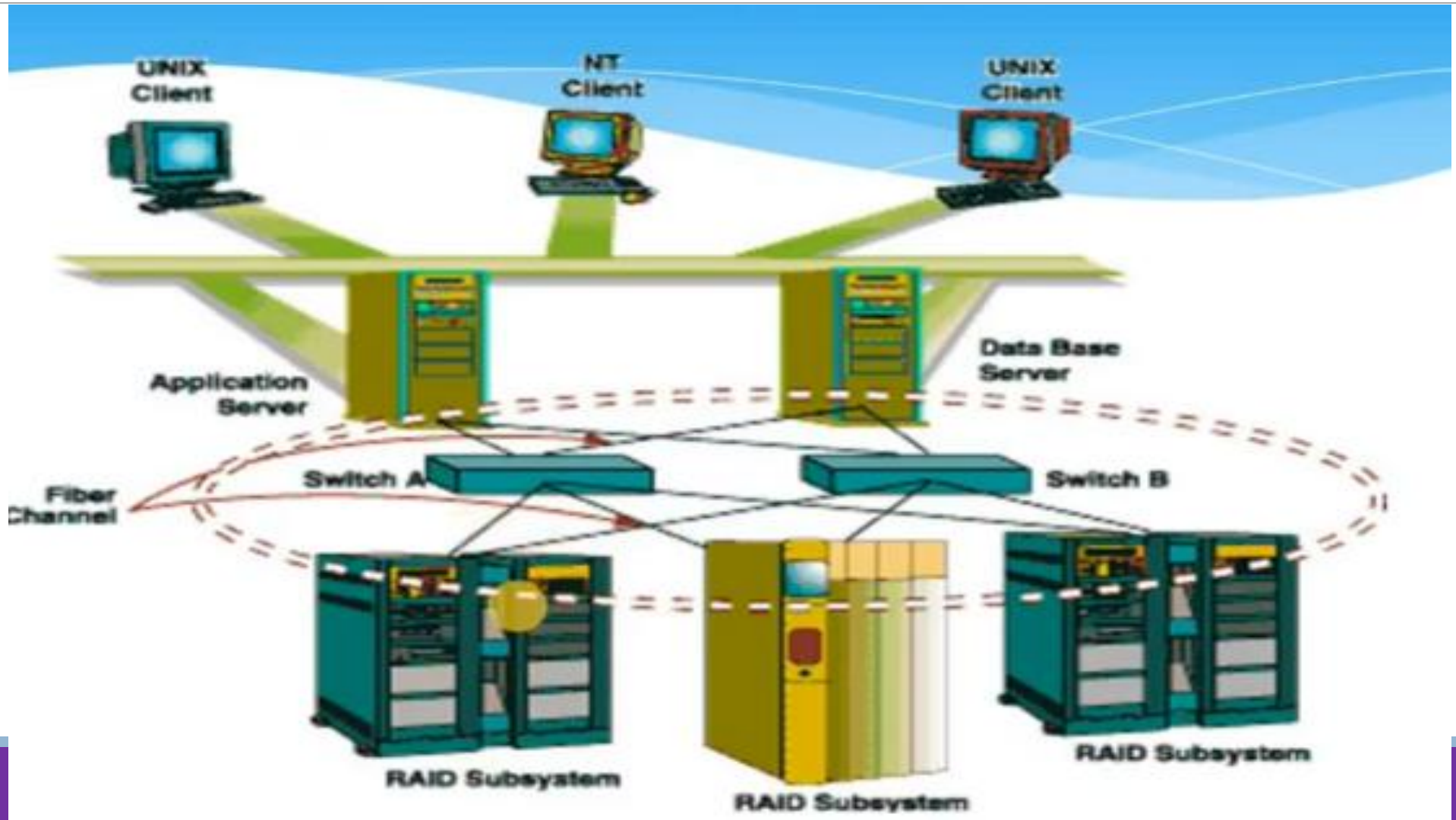
RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Zoning


UNIX
Client
NT
Client
UNIX
Client
Application
Server
Data Base
Server
Fiber
Channel
Switch A
Switch B
RAID Subsystem
RAID Subsystem
RAID Subsystem

Briefly explain the process of zoning in an FC network and illustrate the different types of zoning with a diagram.

Zoning

Zoning is an FC switch function that enable node ports within the fabric to be logically segmented into groups and to communicate with each other within the group (see Figure 5-17).

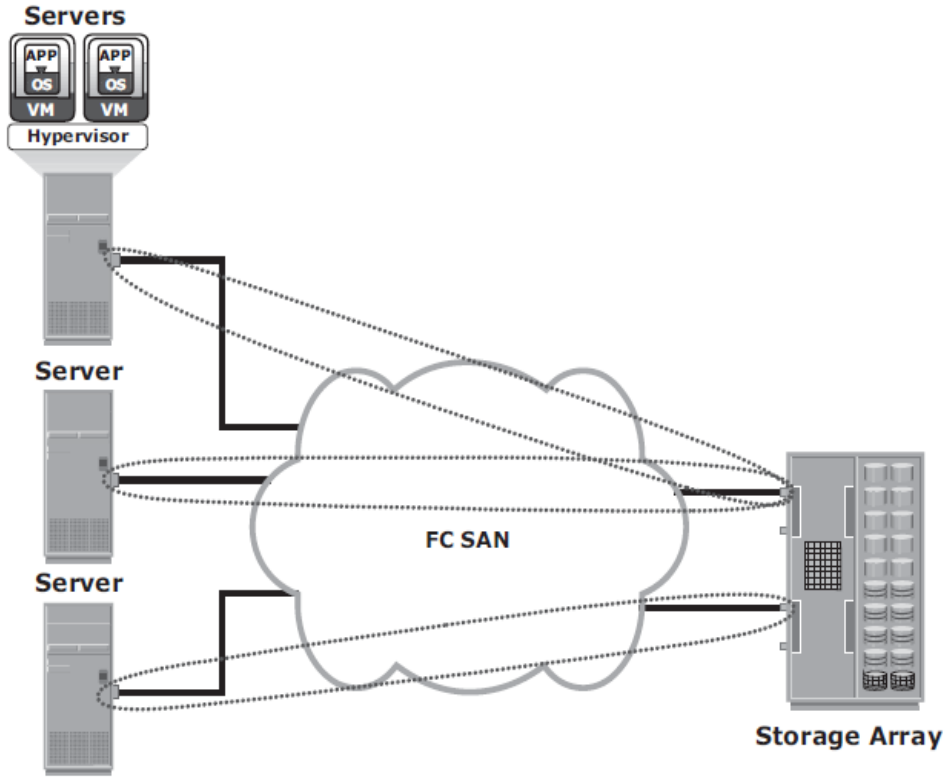


Figure 5-17: Zoning

Briefly explain the process of zoning in an FC network and illustrate the different types of zoning with a diagram.

Zoning

Zone members, zones, and zone sets form the hierarchy defined in the zoning process (see Figure 5-18).

A zone set is composed of a group of zones that can be activated or deactivated as a single entity in a fabric.

Multiple zone sets may be defined in a fabric, but only one zone set can be active at a time.

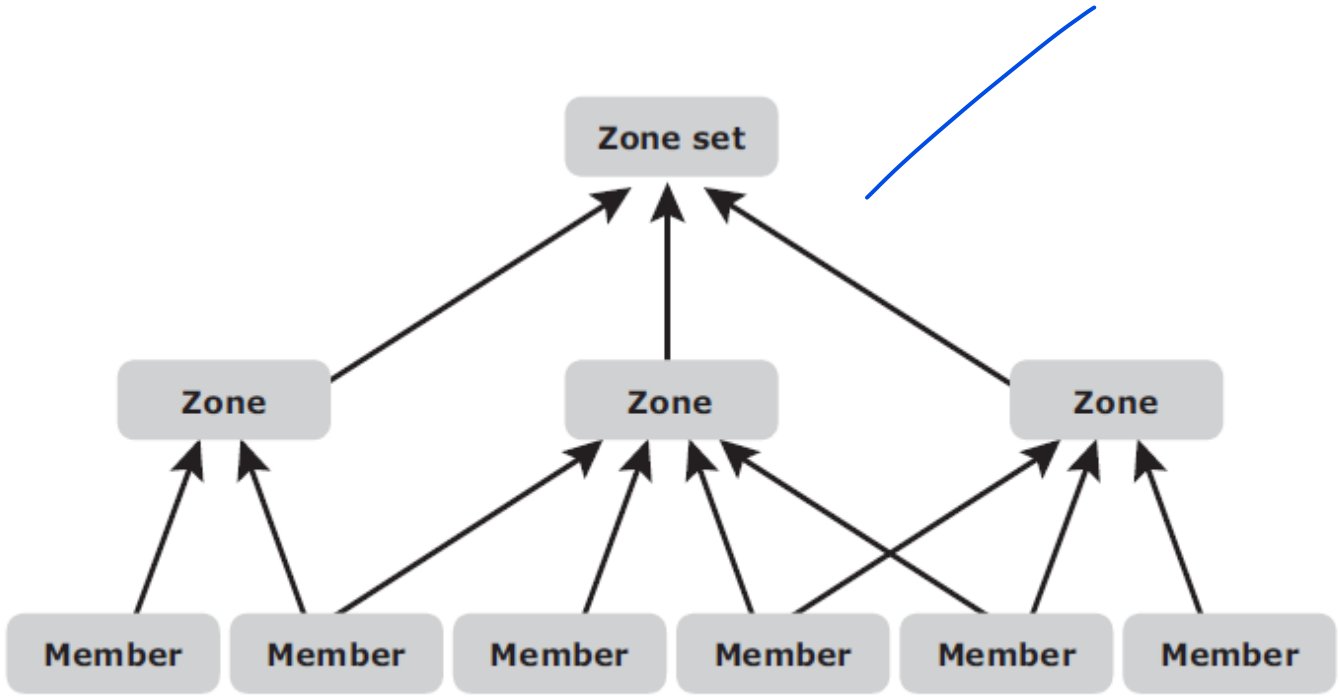


Figure 5-18: Members, zones, and zone sets

Zoning

Zone Members are nodes within the SAN that can be included in a zone.

Switch ports, HBA ports, and storage device ports can be members of a zone.

A port or node can be a member of multiple zones.

Nodes distributed across multiple switches in a switched fabric may also be grouped into the same zone.

Zone sets are also referred to as zone configurations.

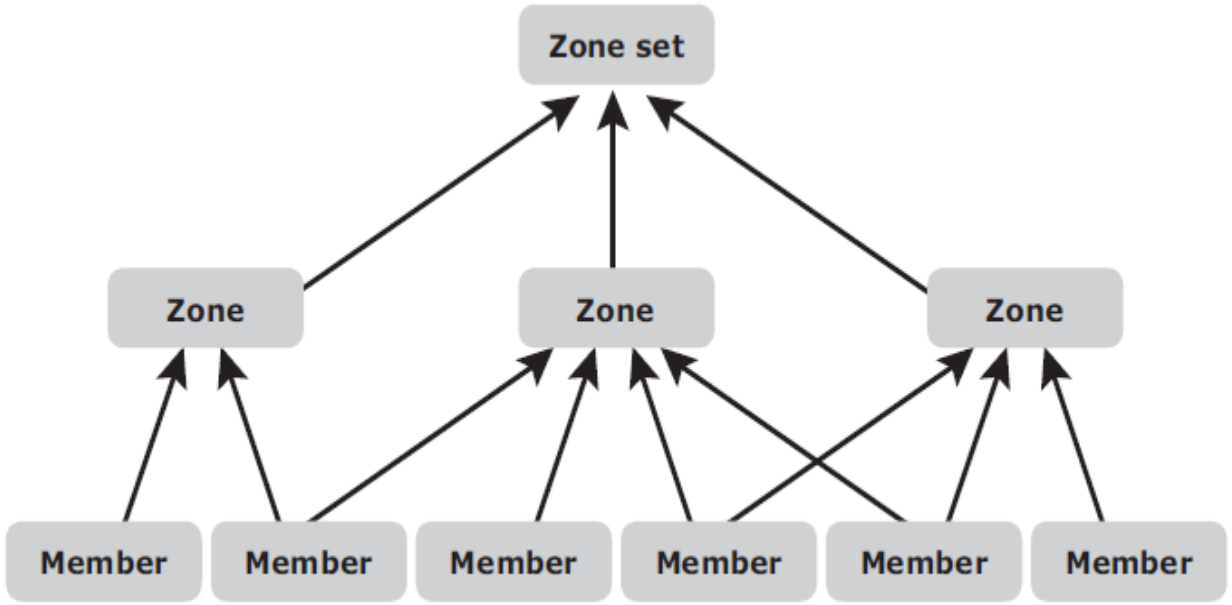


Figure 5-18: Members, zones, and zone sets

Types of Zoning

Zoning can be categorized into three types:

Port zoning:

Uses the physical address of switch ports to define zones.

In port zoning, access to node is determined by the physical switch port to which a node is connected.

The zone members are the port identifier (switch domain ID and port number) to which HBA and its targets (storage devices) are connected.

If a node is moved to another switch port in the fabric, then zoning must be modified to allow the node, in its new port, to participate in its original zone. However, if an HBA or storage device port fails, an administrator just has to replace the failed device without changing the zoning configuration.

Types of Zoning

Zoning can be categorized into three types:

WWN zoning:

Uses World Wide Names to define zones.

The zone members are the unique WWN addresses of the HBA and its targets (storage devices).

A major advantage of WWN zoning is its flexibility.

WWN zoning allows nodes to be moved to another switch port in the fabric and maintain connectivity to its zone partners without having to modify the zone configuration.

This is possible because the WWN is static to the node port.

Types of Zoning

Zoning can be categorized into three types:

Mixed zoning:

Combines the qualities of both WWN zoning and port zoning.

Using mixed zoning enables a specific node port to be tied to the WWN of another node.

RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

FC SAN Topologies

- Fabric design follows standard topologies to connect devices.
- Most commonly deployed in FC SAN implementations are
  - **Mesh Topology**
  - **Core-Edge Fabric**

FC SAN Topologies

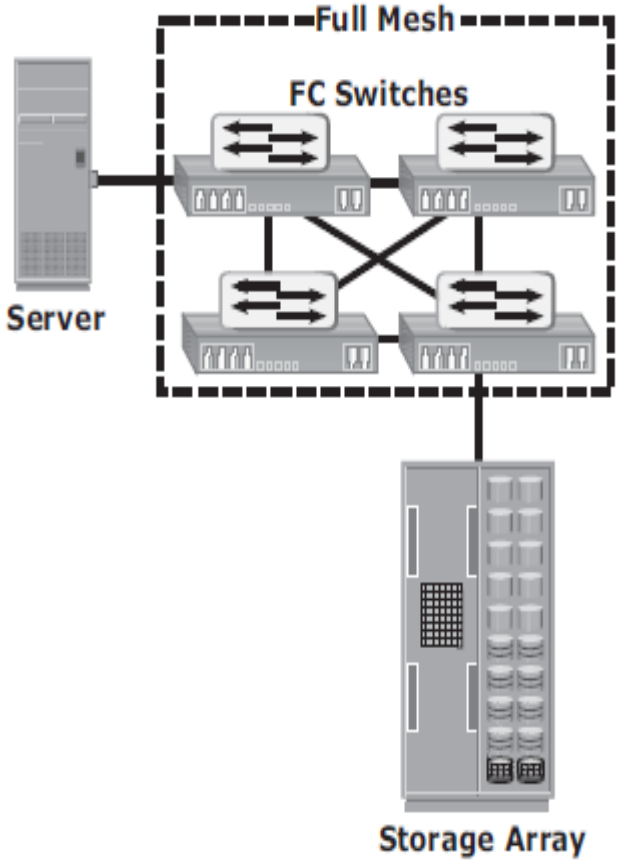
Mesh Topology

A mesh topology may be one of the two types:

- full mesh
- partial mesh.

FC SAN Topologies

- Full mesh
- Every switch is connected to every other switch in the topology.
- A full mesh topology may be appropriate when the number of switches involved is small.
- In a full mesh topology, a maximum of one ISL or hop is required for host-to-storage traffic.
- However, with the increase in the number of switches, the number of switch ports used for ISL also increases. This reduces the available switch ports for node connectivity.



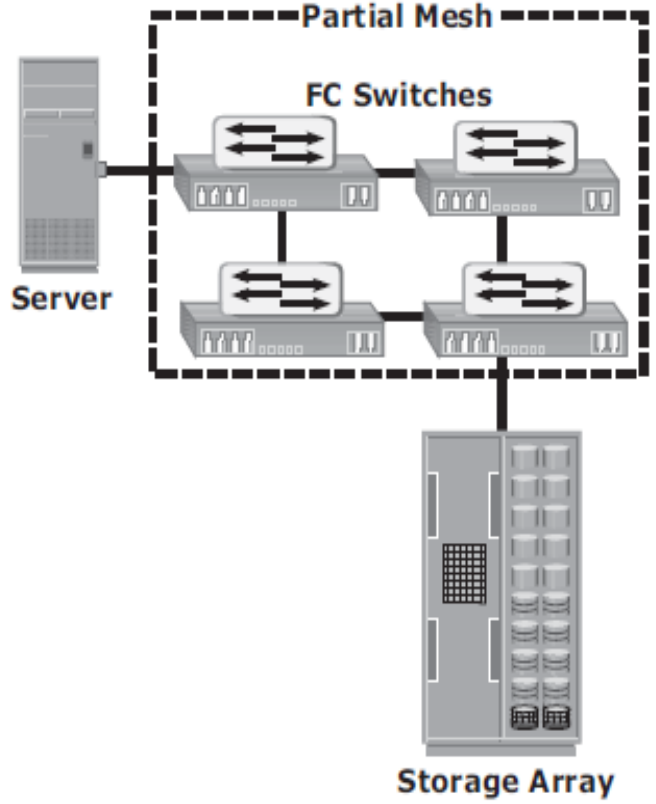
FC SAN Topologies

Partial mesh topology

Several hops or ISLs may be required for the traffic to reach its destination.

Partial mesh offers more scalability than full mesh topology.

Traffic management in a partial mesh fabric is complicated and ISLs could become overloaded due to excessive traffic aggregation.



Explain the various variations of the Core-Edge Fabric topology with the help

FC SAN Topologies

Core-Edge Fabric

The core-edge fabric topology has two types of switch tiers.

- Edge tier
- Core tier

FC SAN Topologies

Core-Edge Fabric – Edge tier

- The edge tier is usually composed of switches and offers an inexpensive approach to adding more hosts in a fabric.
- Each switch at the edge tier is attached to a switch at the core tier through ISLs.

Core-Edge Fabric – Core tier

- The core tier is usually composed of enterprise directors that ensure high fabric availability.
- In addition, typically all traffic must either traverse this tier or terminate at this tier.
- In this configuration, all storage devices are connected to the core tier, enabling host-to-storage traffic to traverse only one ISL.
- Hosts that require high performance may be connected directly to the core tier and consequently avoid ISL delays.

FC SAN Topologies

Core-Edge Fabric

Based on the number of core-tier switches, this topology has different variations, such as, single-core topology and dual-core topology

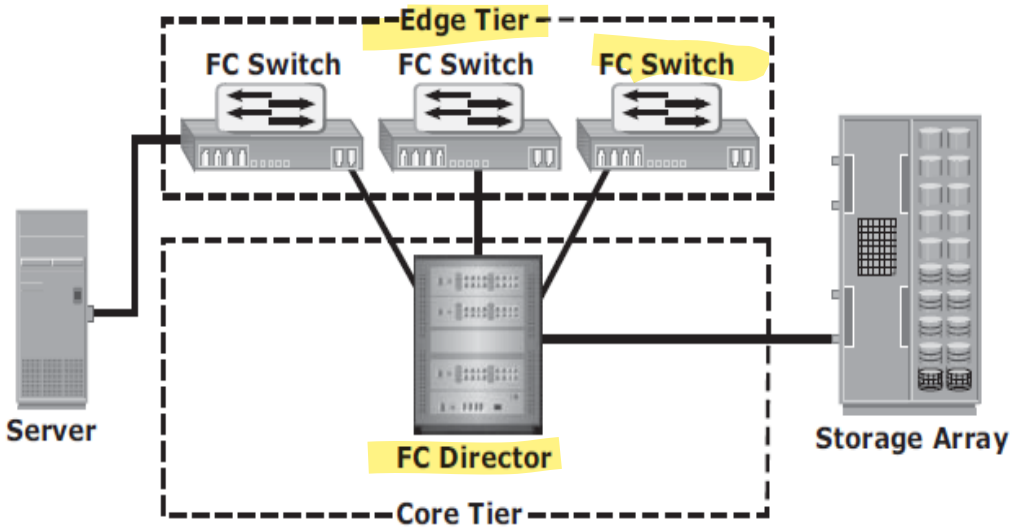


Figure 5-21: Single-core topology

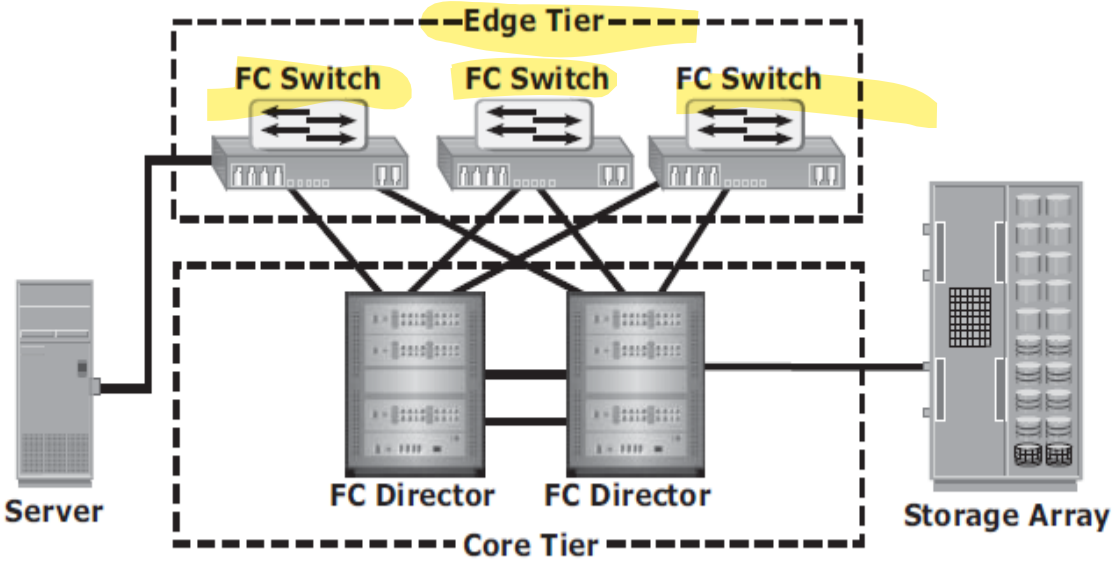


Figure 5-22: Dual-core topology

FC SAN Topologies

Benefits and Limitations of Core-Edge Fabric

- The core-edge fabric provides maximum one-hop storage access to all storage devices in the system.
- Core-edge provides easier calculation of the ISL load and traffic patterns
- Easy to identify which network resources are approaching their capacity
- Makes easier to develop a set of rules for scaling and apportioning.
- As the number of cores increases, it is prohibitive to continue to maintain ISLs from each core to each edge switch.

RAMAIAH
Institute of Technology

UNIT 3- Storage Networking Technologies

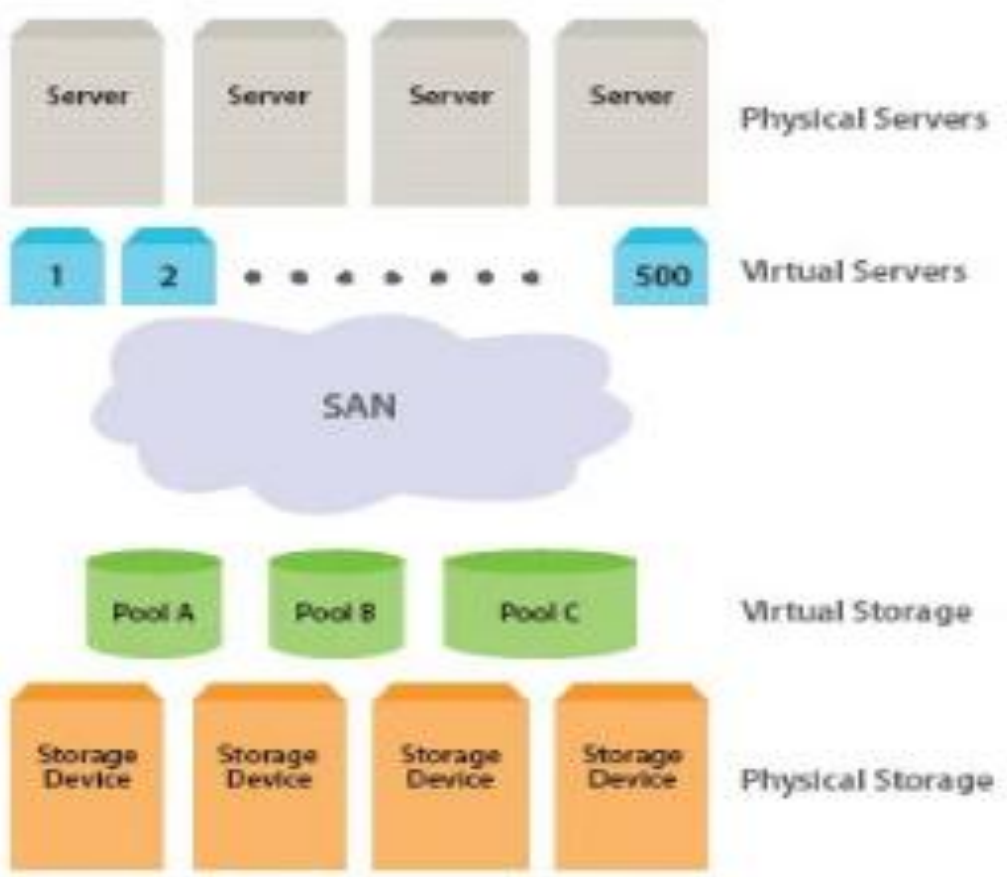
Fiber Channel Storage Area Networks(Chapter 5)

- Fiber Channel Overview
- The SAN and Its Evolution
- Components of FC SAN,
- FC Connectivity,
- Switched Fabric Ports ,
- Fibre Channel Architecture,
- Fabric Services
- Switched Fabric Login Types,
- Zoning,
- FC SAN Topologies,
- Virtualization in SAN.

Virtualization in SAN

Two network-based virtualization techniques in a SAN environment:

- Block-level storage virtualization
- Virtual SAN (VSAN)



Block-level Storage Virtualization

- Figure 5-24 illustrates a virtualized environment.
- It shows two physical servers, each of which has one virtual volume assigned.
- These virtual volumes are used by the servers.
- These virtual volumes are mapped to the LUNs in the storage arrays.
- When an I/O is sent to a virtual volume, it is redirected through the virtualization layer at

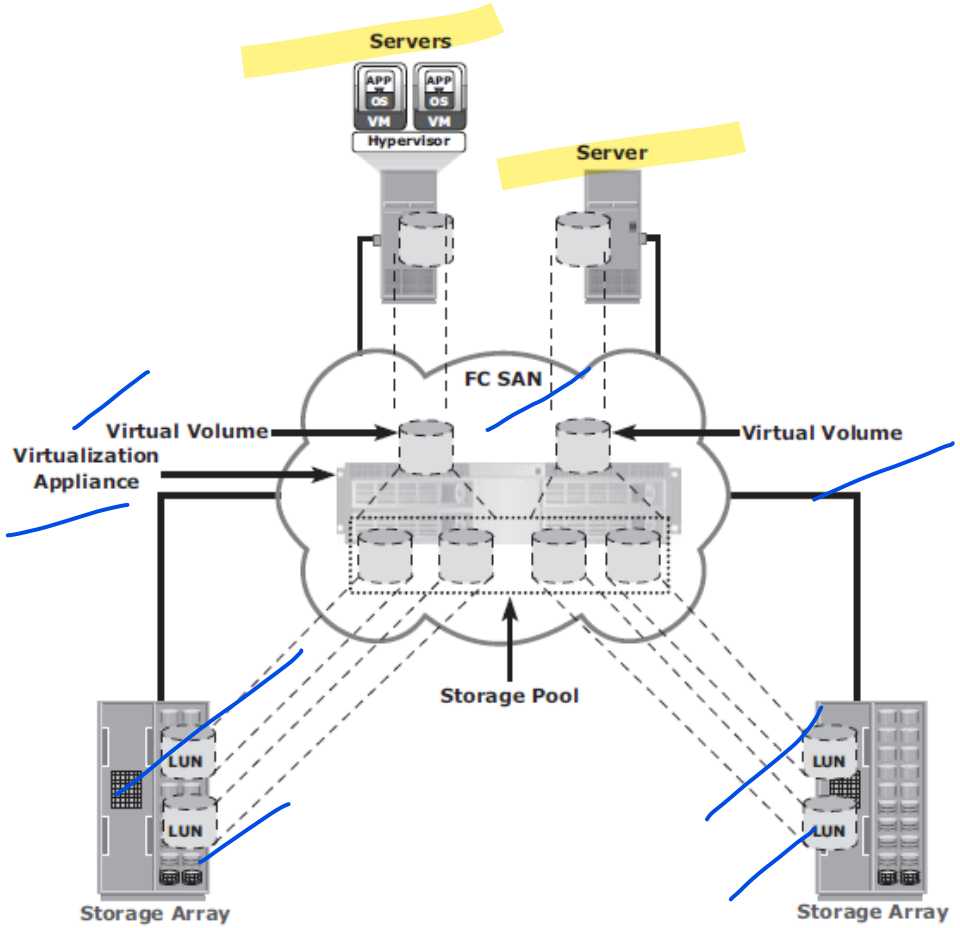


Figure 5-24: Block-level storage virtualization

RAMAIAH
Institute of Technology

Block-level Storage Virtualization

- Block-level storage virtualization enables extending the storage volumes online to meet application growth requirements.
- It consolidates heterogeneous storage arrays and enables transparent volume access.

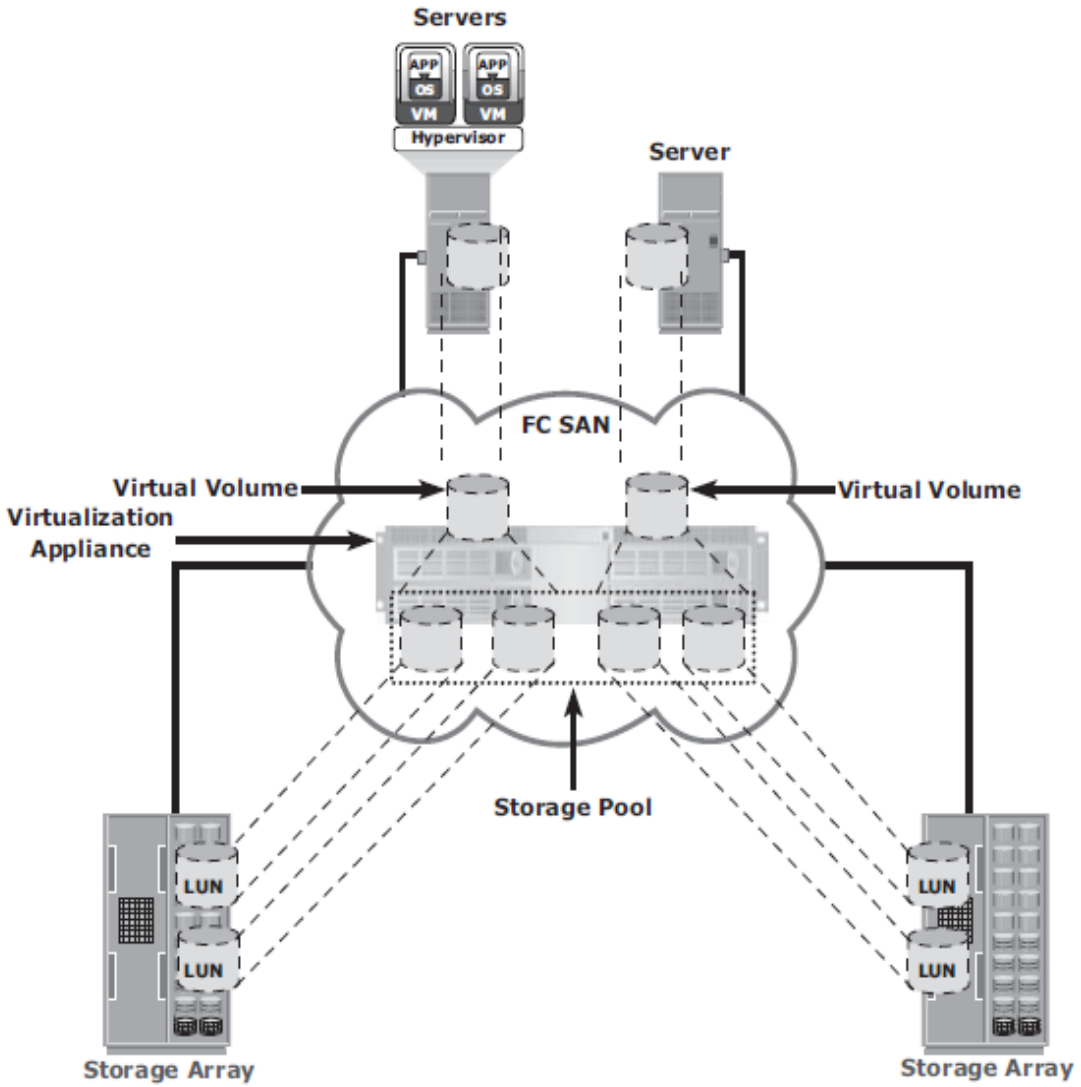


Figure 5-24: Block-level storage virtualization

Block-level Storage Virtualization

Block-level storage virtualization provided nondisruptive data migration only within a data center.

The new generation of block-level storage virtualization enables nondisruptive data migration both within and between data centers.

It provides the capability to connect the virtualization layers at multiple data centers.

The connected virtualization layers are managed centrally and work as a single virtualization layer stretched across data centers (Figure 5-25).

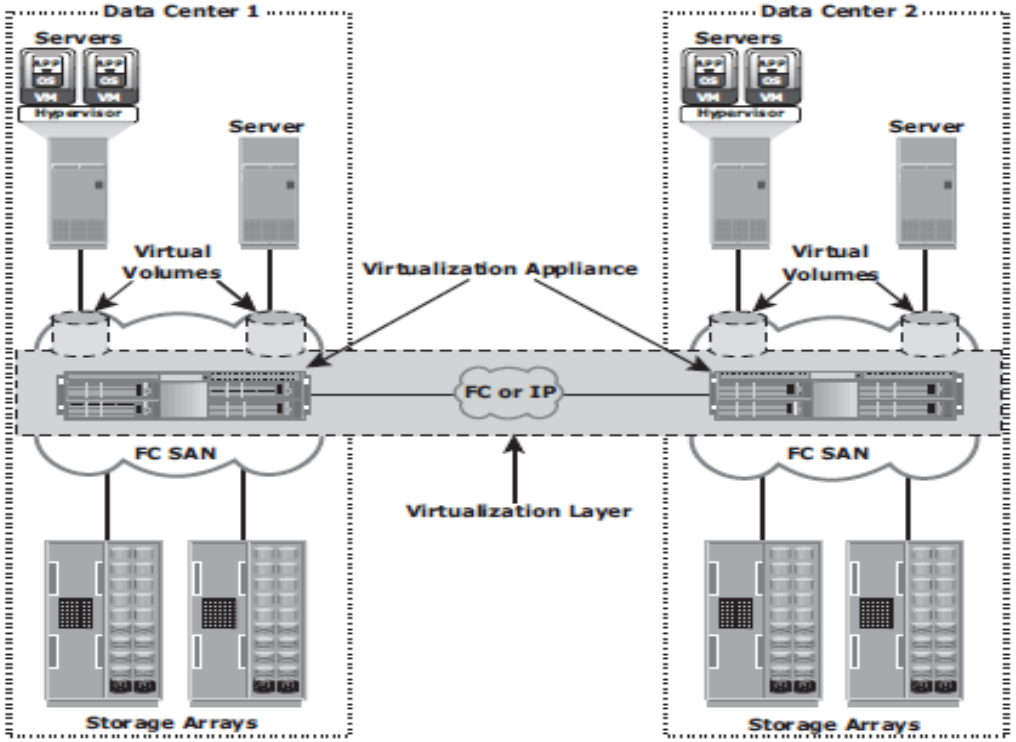


Figure 5-25: Federation of block storage across data centers

Virtual SAN (VSAN)

- Virtual SAN (also called virtual fabric) is a logical fabric on an FC SAN, which enables communication among a group of nodes regardless of their physical location in the fabric.
- In a VSAN, a group of hosts or storage ports communicate with each other using a virtual topology defined on the physical SAN.
- Multiple VSANs may be created on a single physical SAN. Each VSAN acts as an independent fabric with its own set of fabric services, such as name server, and zoning.
- Fabric-related configurations in one VSAN do not affect the traffic in another.

RAMAIAH Institute of Technology

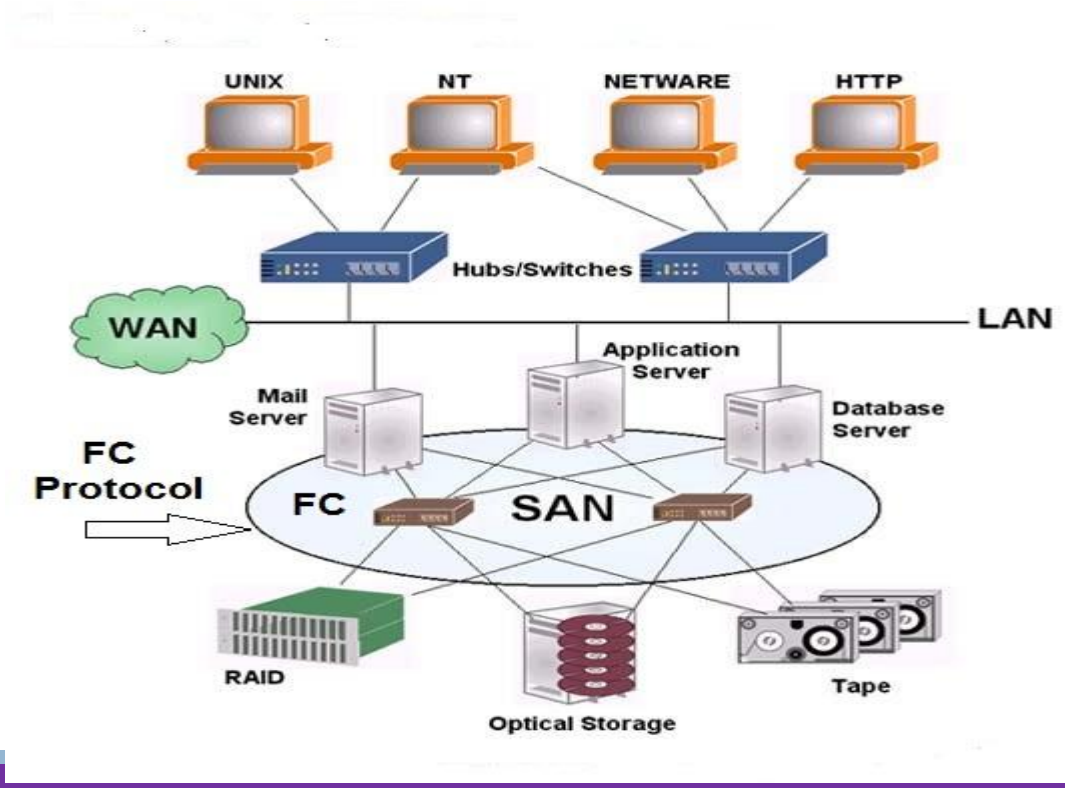
UNIT 3- Storage Networking Technologies

IP SAN and FCoE(Chapter 6)

- iSCSI (Small Computer Systems Interface (iSCSI)
- FCIP (Fibre Channel over IP)
- FCoE(Fibre Channel Over Ethernet)

Introduction

Traditional SAN enables the transfer of block I/O over Fibre Channel and provides high performance and scalability.



Introduction

Organizations typically have an existing Internet Protocol (IP)-based infrastructure, which could be leveraged for storage networking.

Advancements in technology have enabled IP to be used for transporting block I/O over the IP network.

This technology of transporting block I/Os over an IP is referred to as IP SAN.

IP is a mature technology, so use IP as a storage networking

Introduction

Two primary protocols that leverage IP as the transport mechanism are

1. Internet SCSI (iSCSI)
2. Fibre Channel over IP (FCIP)

Emerging protocol - Fibre Channel over Ethernet (FCoE).

iSCSI (Small Computer Systems Interface (iSCSI)

iSCSI is an IP based protocol that establishes and manages connections between host and storage over IP, as shown in Figure 6-1.

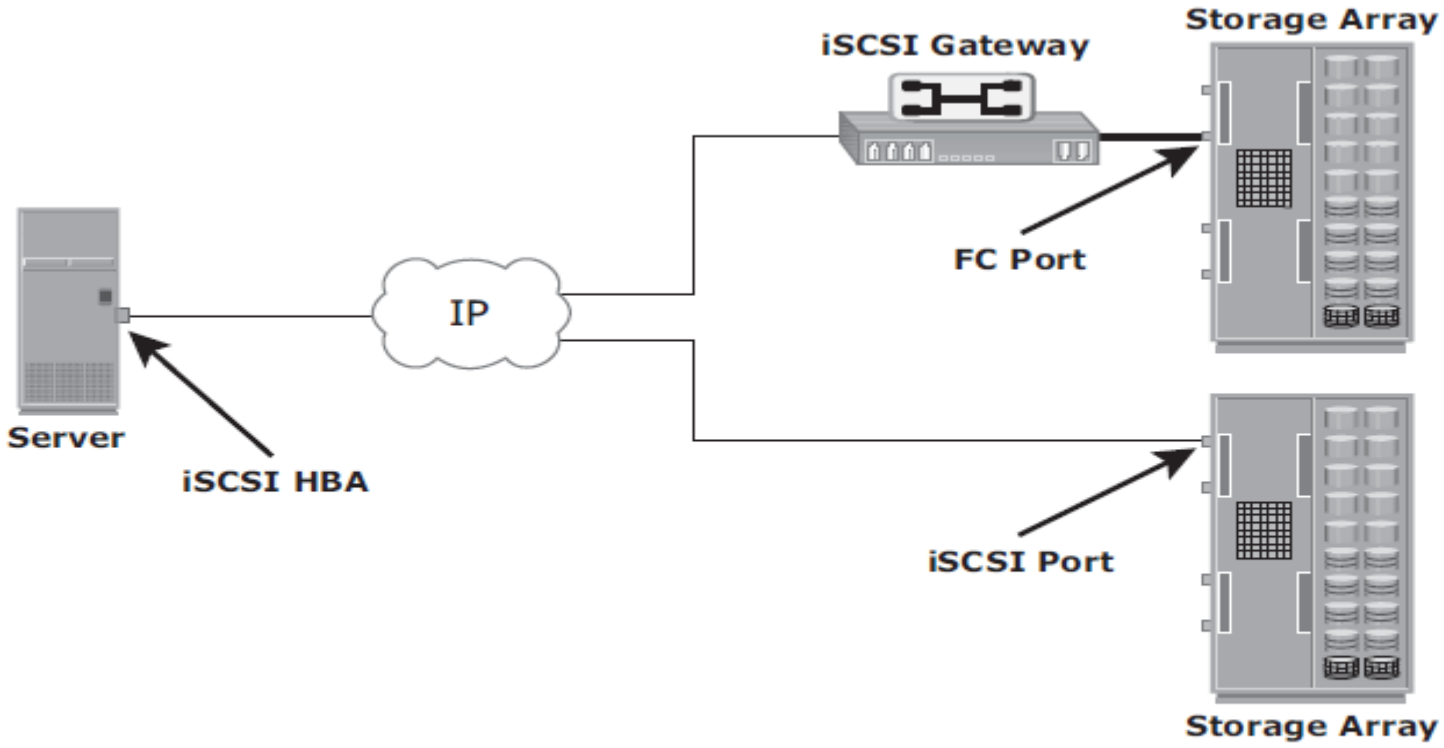


Figure 6-1: iSCSI implementation

iSCSI (Small Computer Systems Interface (iSCSI)

iSCSI is an IP based protocol that establishes and manages connections between host and storage over IP, as shown in Figure 6-1.

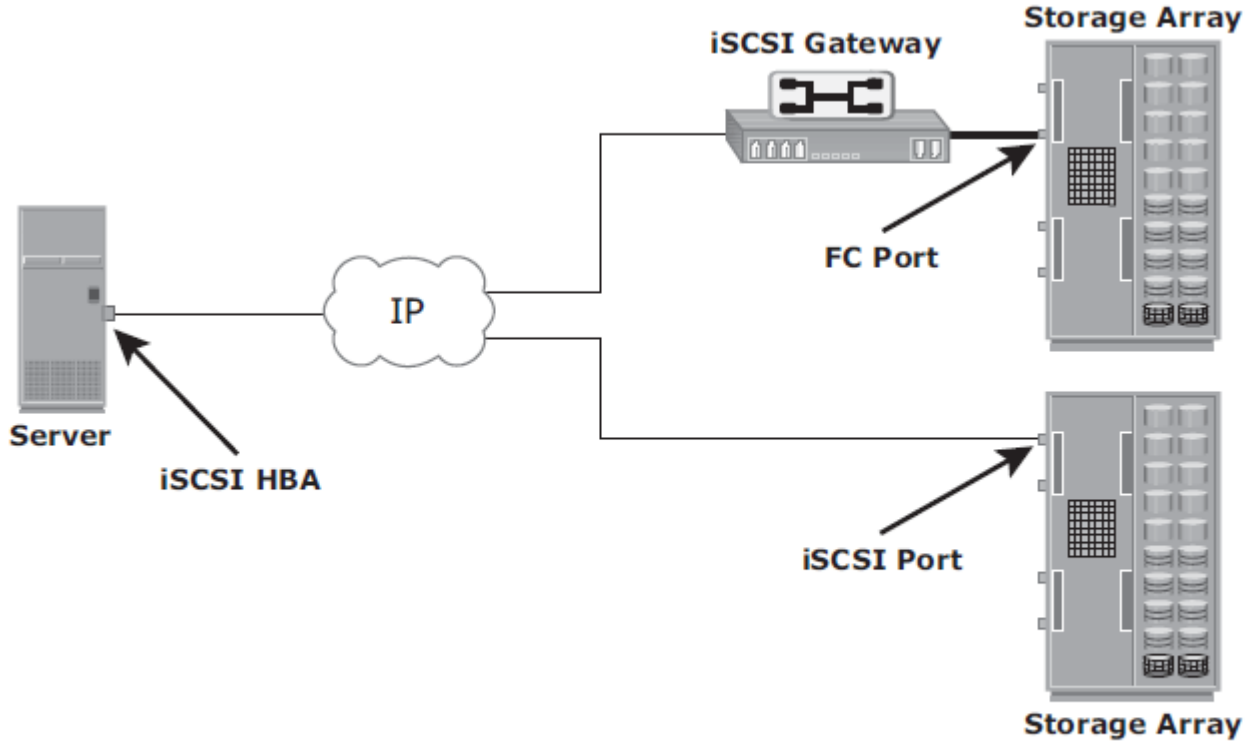


Figure 6-1: iSCSI implementation

iSCSI (Small Computer Systems Interface (iSCSI)

iSCSI encapsulates SCSI commands and data into an IP packet and transports them using TCP/IP.

iSCSI is widely adopted for connecting servers to storage because it is relatively inexpensive and easy to implement, especially in environments in which an FC SAN does not exist.

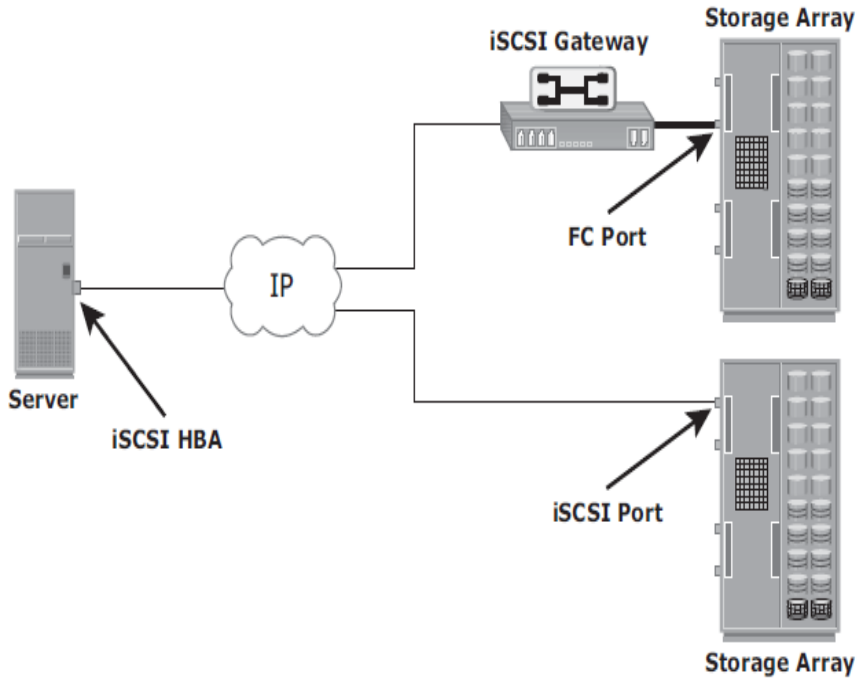


Figure 6-1: iSCSI implementation

iSCSI (Small Computer Systems Interface (iSCSI)

- Components of iSCSI
- iSCSI Host Connectivity
- iSCSI Topologies
  - Native iSCSI Connectivity
  - Bridged iSCSI Connectivity
- iSCSI Protocol Stack
- iSCSI PDU
  - iSCSI Discovery
  - iSCSI Names
  - iSCSI Session
  - iSCSI Command Sequencing

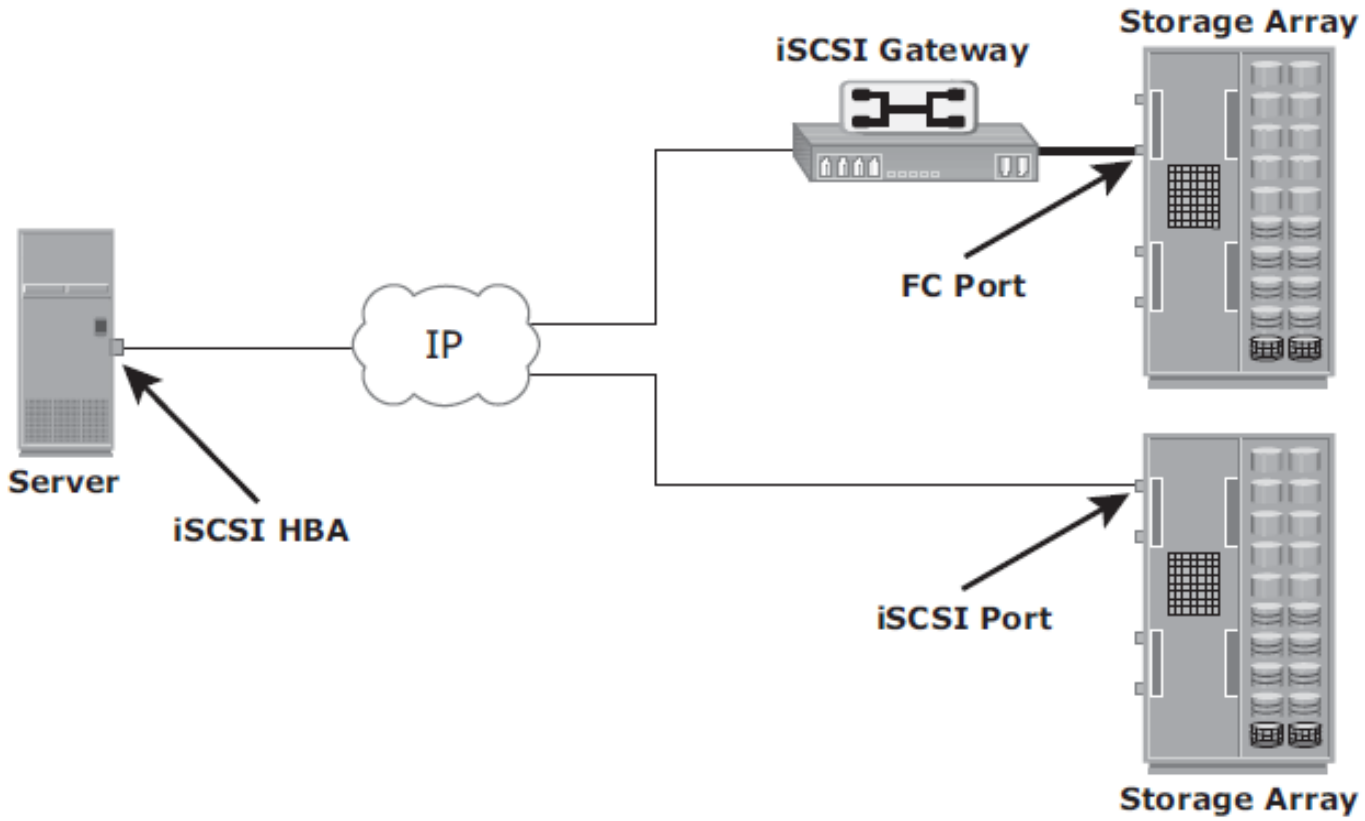


Figure 6-1: iSCSI implementation

Components of iSCSI

key iSCSI components

- An initiator (host)- The function is to route the SCSI commands over an IP network.
- target (storage or iSCSI gateway),
- IP-based network

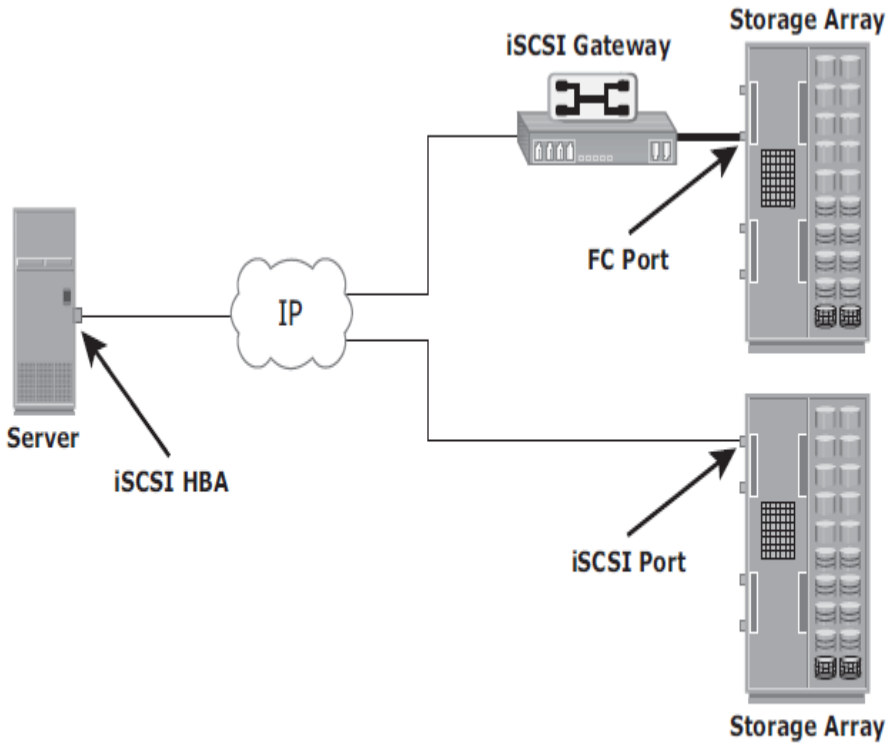


Figure 6-1: iSCSI implementation

Components of iSCSI

If an iSCSI-capable storage array is deployed, then a host with the iSCSI initiator can directly communicate with the storage array over an IP network.

However, in an implementation that uses an existing FC array for iSCSI communication, an iSCSI gateway is used.

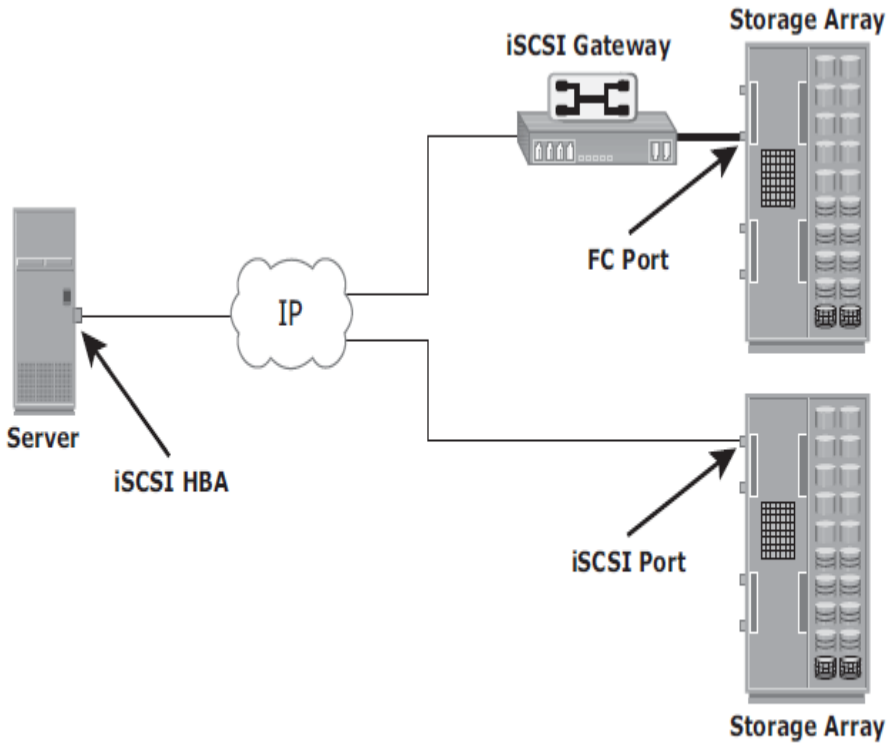


Figure 6-1: iSCSI implementation

Components of iSCSI

These devices perform the translation of IP packets to FC frames and vice versa, thereby bridging the connectivity between the IP and FC environments.

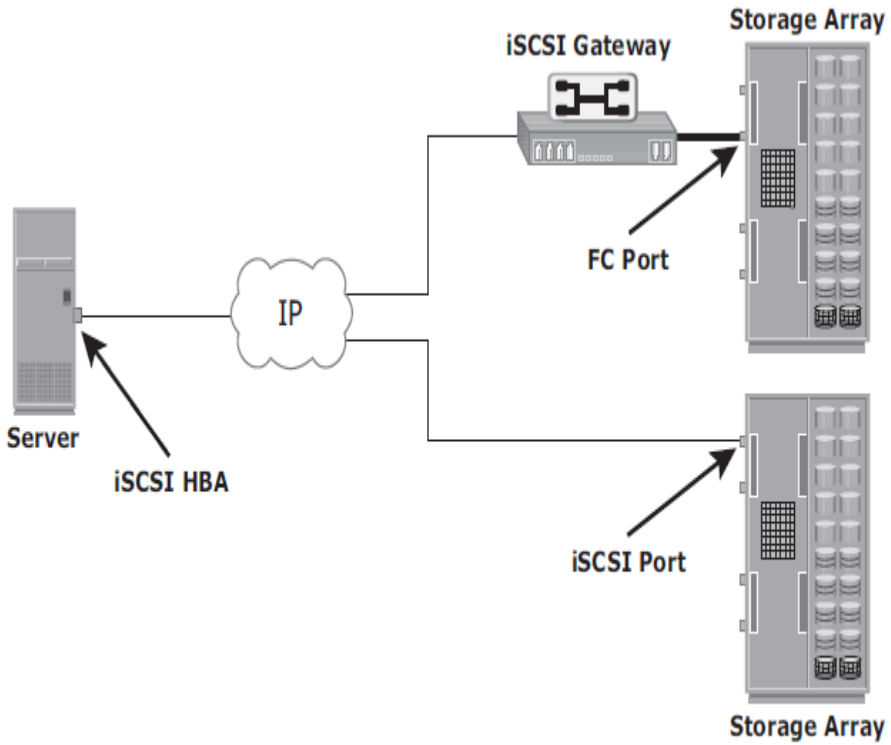


Figure 6-1: iSCSI implementation

iSCSI Host Connectivity

1. NIC with software iSCSI initiator

- It is easy to implement because most servers come with at least one or two embedded NICs.
- Because NICs provide standard IP function, encapsulation of SCSI into IP packets and decapsulation are carried out by the host CPU.
- This places additional overhead on the host CPU.
- If a standard NIC is used in heavy I/O load situations, the host CPU might become a bottleneck.

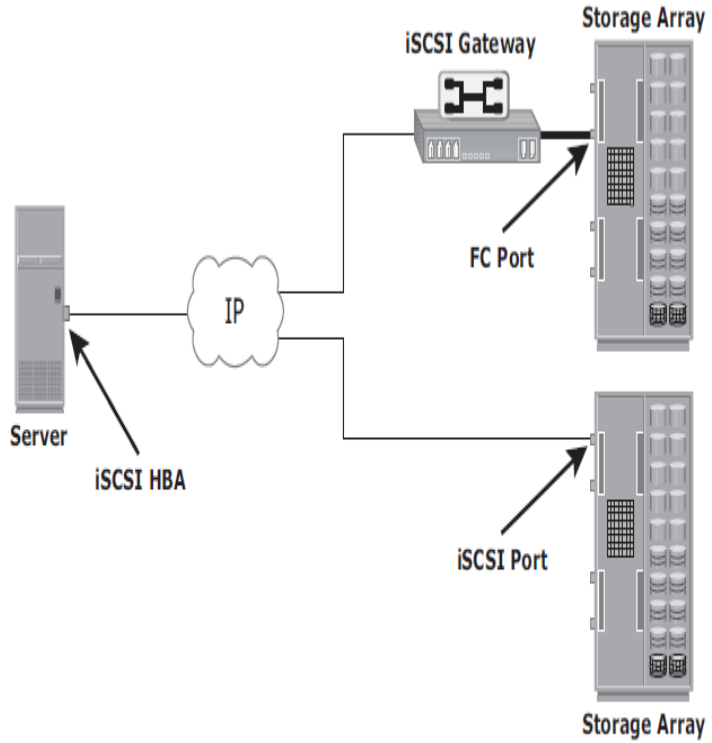


Figure 6-1: iSCSI implementation

iSCSI Host Connectivity

2. TCP offload engine (TOE) NIC with software iSCSI initiator

- TOE NIC helps alleviate this burden.
- A TOE NIC offloads TCP management functions from the host and leaves only the iSCSI functionality to the host processor.
- The host passes the iSCSI information to the TOE card, and the TOE card sends the information to the destination using TCP/IP.
- Although this solution improves performance, the iSCSI functionality is still handled by a software initiator that requires host CPU cycles.

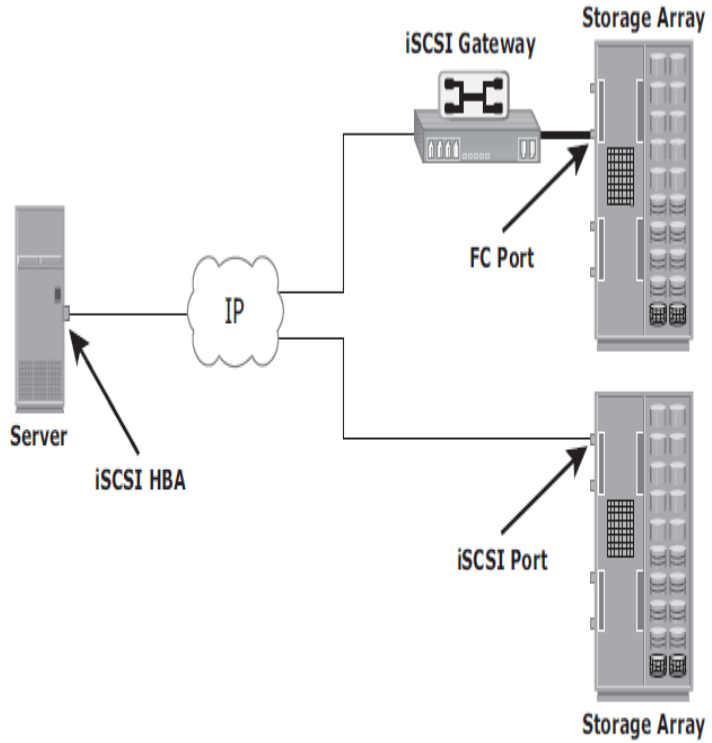


Figure 6-1: iSCSI implementation

iSCSI Host Connectivity

3. iSCSI HBA

- An iSCSI HBA is capable of providing performance benefits because it offloads the entire iSCSI and TCP/IP processing from the host processor.
- The use of an iSCSI HBA is also the simplest way to boot hosts from a SAN environment via iSCSI.
- If there is no iSCSI HBA, modifications must be made to the basic operating system to boot a host from the storage devices because the NIC needs to obtain an IP address before the operating system loads.

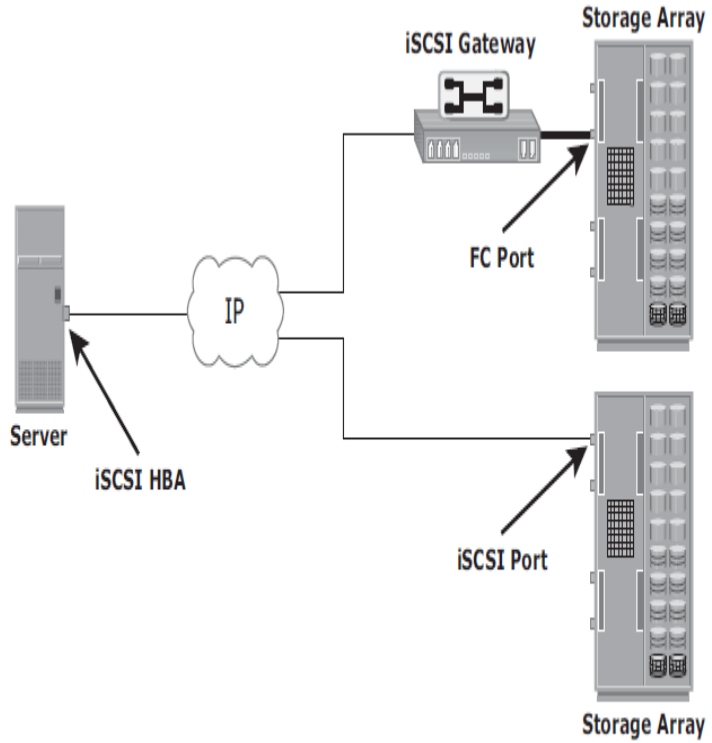


Figure 6-1: iSCSI implementation

iSCSI Topologies

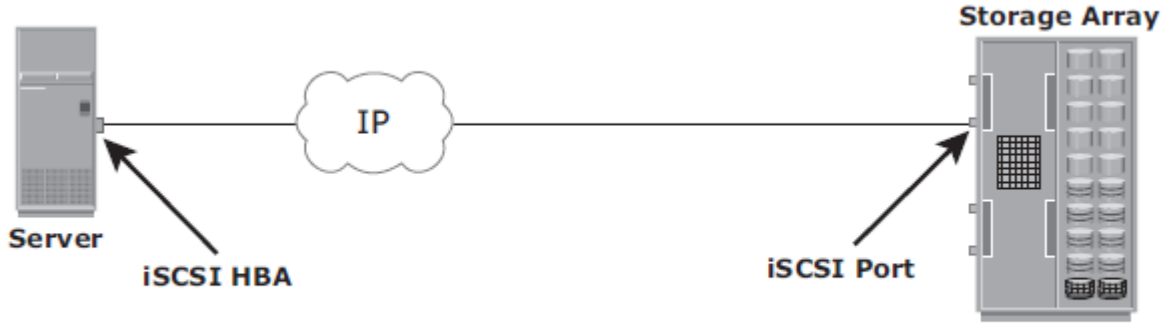
Two topologies of iSCSI implementations are

1. Native
2. bridged

iSCSI Topologies - Native iSCSI Connectivity

Native topology does not have FC components. The initiators may be either directly attached to targets or connected through the IP network.

In Figure 6-2 (a), the array has one or more iSCSI ports configured with an IP address and is connected to a standard Ethernet switch.



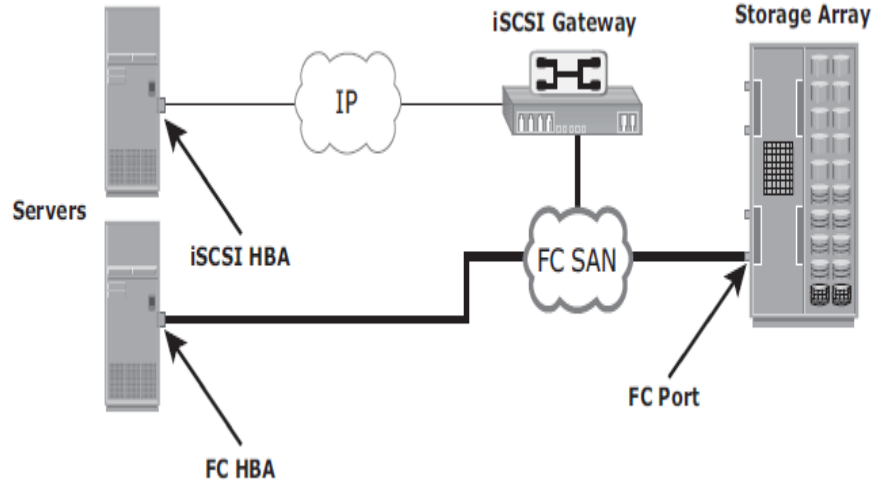
(a) Native iSCSI Connectivity

RAMAIAH Institute of Technology

iSCSI Topologies - Bridged iSCSI Connectivity

- Bridged topology enables the coexistence of FC with IP by providing iSCSI-to-FC bridging functionality.
- Figure 6-2 (b) illustrates iSCSI host connectivity to an FC storage array.
- In this case, the array does not have any iSCSI ports. Therefore, an external device, called a gateway or a multiprotocol router, must be used to facilitate the communication between the iSCSI host and FC storage.

The gateway converts IP packets to FC frames and vice versa.

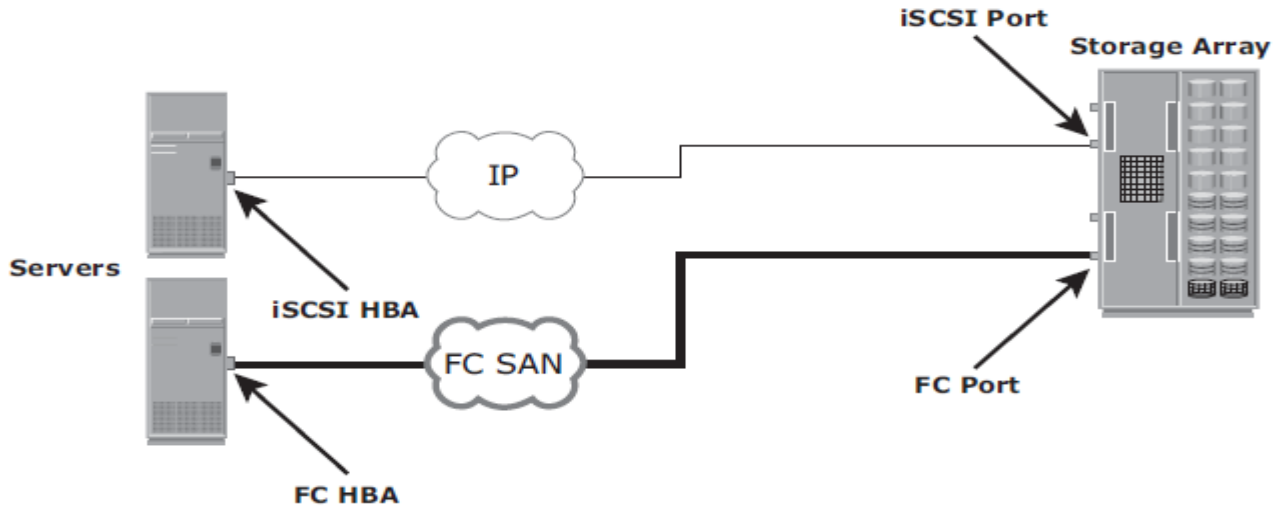


(b) Bridged iSCSI Connectivity

iSCSI Topologies - Bridged iSCSI Connectivity

The most common topology is a combination of FC and native iSCSI.

Typically, a storage array comes with both FC and iSCSI ports that enable iSCSI and FC connectivity in the same environment, as shown in Figure 6-2 (c).



(c) Combining FC and Native iSCSI Connectivity

Figure 6-2: iSCSI Topologies

iSCSI Protocol Stack

Figure 6-3 displays a model of the iSCSI protocol layers and depicts the encapsulation order of the SCSI commands for their delivery through a physical carrier.

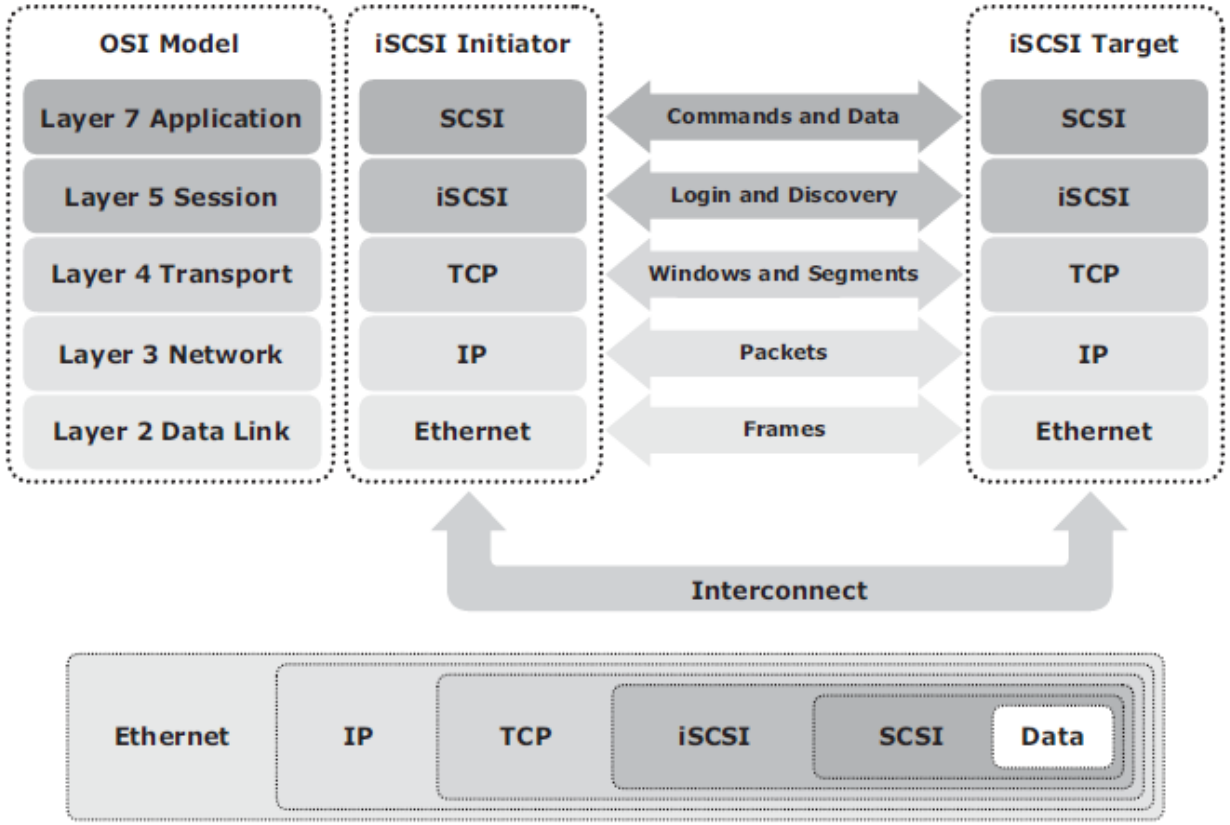


Figure 6-3: iSCSI protocol stack

iSCSI Protocol Stack

SCSI is the command protocol that works at the application layer of the Open System Interconnection (OSI) model.

The initiators and targets use SCSI commands and responses to talk to each other.

The SCSI command descriptor blocks, data, and status messages are encapsulated into TCP/IP and transmitted across the network between the initiators and targets.

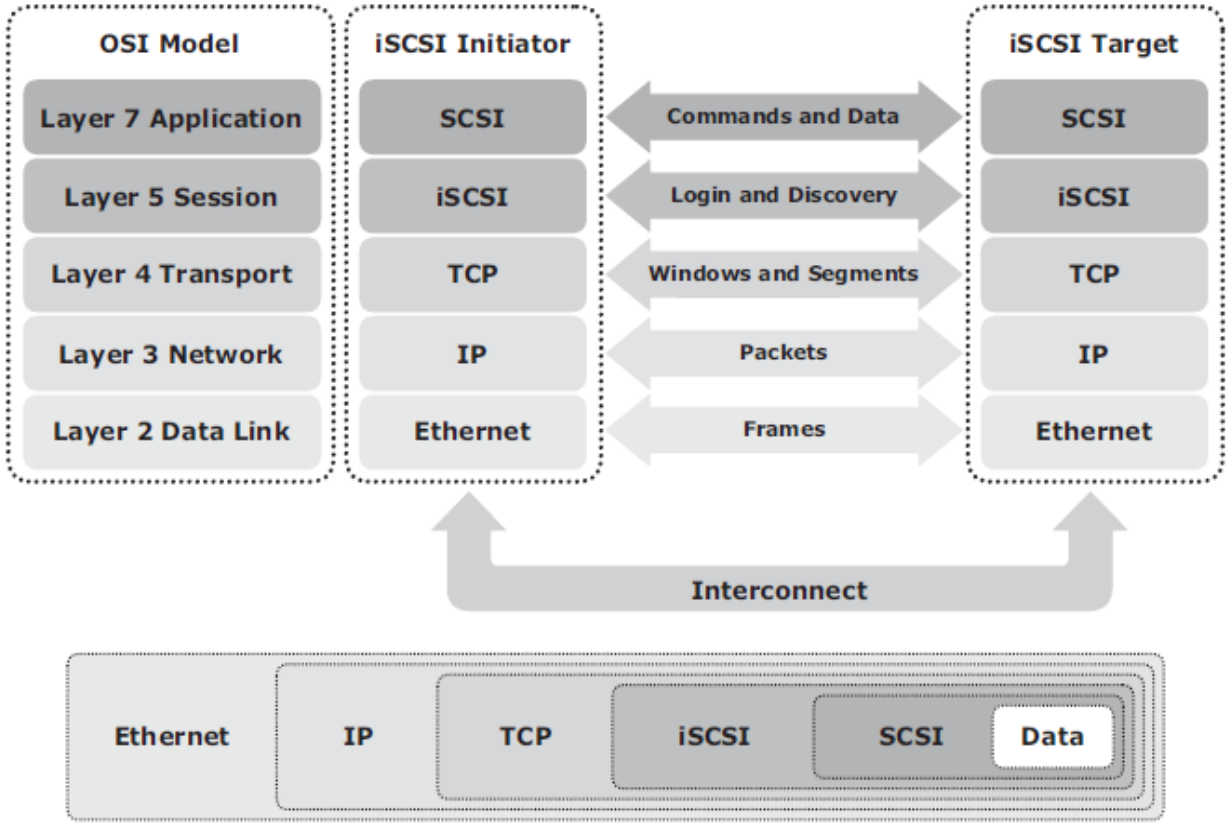


Figure 6-3: iSCSI protocol stack

iSCSI Protocol Stack

iSCSI is the session-layer protocol that initiates a reliable session between devices that recognize SCSI commands and TCP/IP.

The iSCSI session-layer interface is responsible for handling login, authentication, target discovery, and session management.

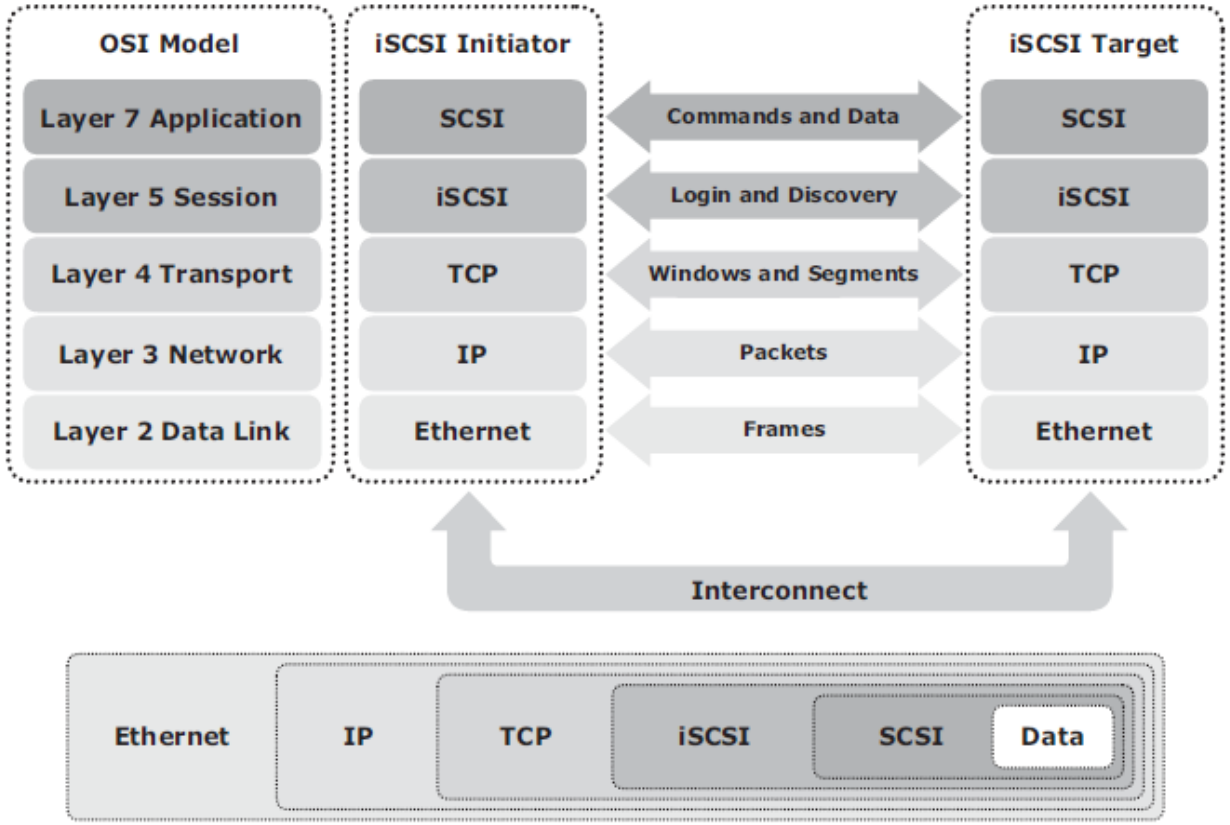


Figure 6-3: iSCSI protocol stack

iSCSI Protocol Stack

TCP is used with iSCSI at the transport layer to provide reliable transmission.

TCP controls message flow, windowing, error recovery, and retransmission.

It relies upon the network layer IP of the OSI model to provide global addressing and connectivity.

The Layer 2 protocols at the data link layer of this model enable node-to-node communication through a physical network.

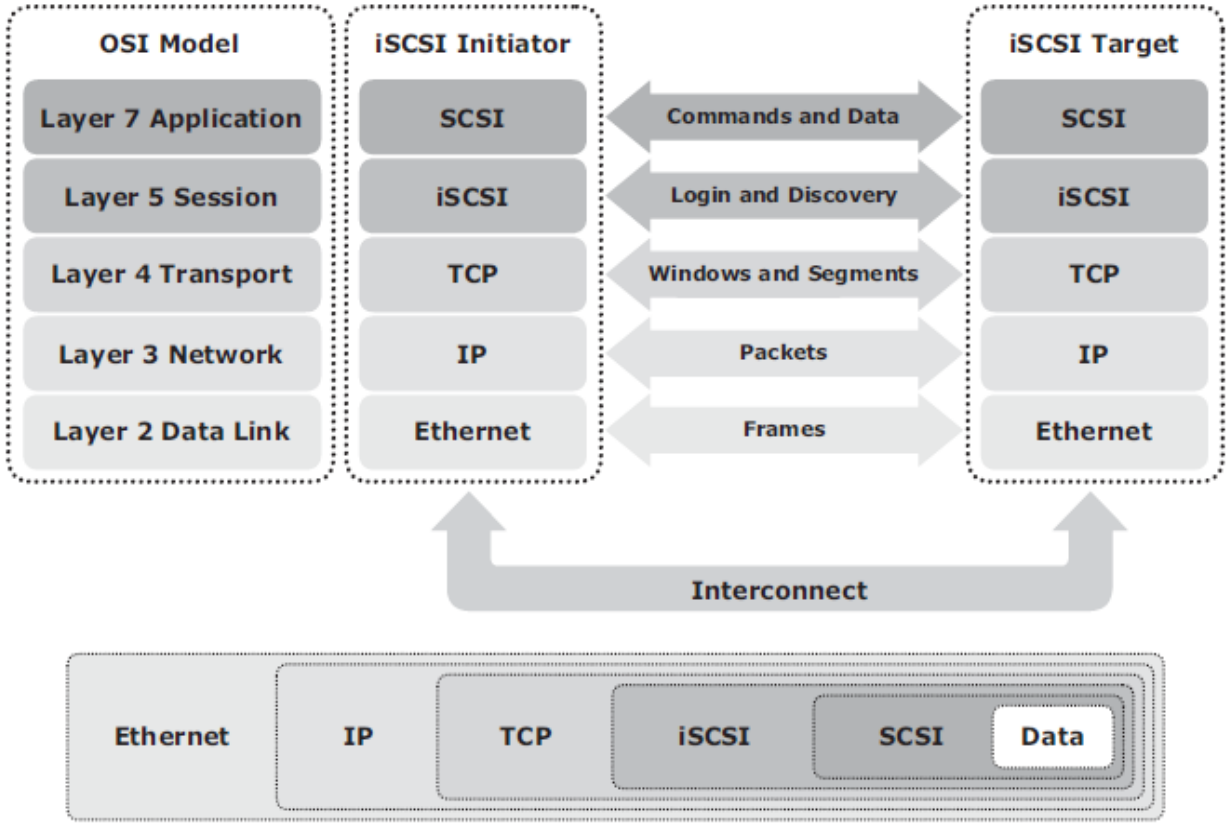


Figure 6-3: iSCSI protocol stack

iSCSI PDU

- A protocol data unit (PDU) is the basic “information unit” in the iSCSI environment.
- The iSCSI initiators and targets communicate with each other using iSCSI PDUs.
- This communication includes establishing
  - iSCSI connections
  - iSCSI sessions,
  - performing iSCSI discovery,
  - sending SCSI commands and data, and receiving SCSI status.
- All iSCSI PDUs contain one or more header segments followed by zero or more data segments.
- The PDU is then encapsulated into an IP packet to facilitate the transport.

iSCSI PDU

- A PDU includes the components shown in Figure 6-4.
- The IP header provides packet-routing information to move the packet across a network.
- The TCP header contains the information required to guarantee the packet delivery to the target.
- The iSCSI header (basic header segment) describes how to extract SCSI commands and data for the target.
- iSCSI adds an optional CRC, known as the digest, to ensure datagram integrity. This is in addition to TCP checksum and Ethernet CRC.
- The header and the data digests are optionally used in the PDU to validate integrity and data placement.

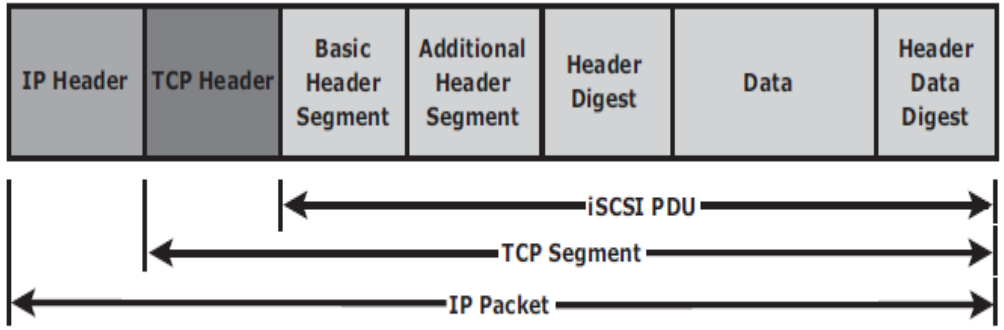


Figure 6-4: iSCSI PDU encapsulated in an IP packet

iSCSI PDU

- As shown in Figure 6-5, each iSCSI PDU does not correspond in a 1:1 relationship with an IP packet.
- Depending on its size, an iSCSI PDU can span an IP packet with another PDU in the same packet.
- To achieve the 1:1 relationship between the IP packet and the iSCSI PDU, the maximum transmission unit (MTU) size of the IP packet is modified.

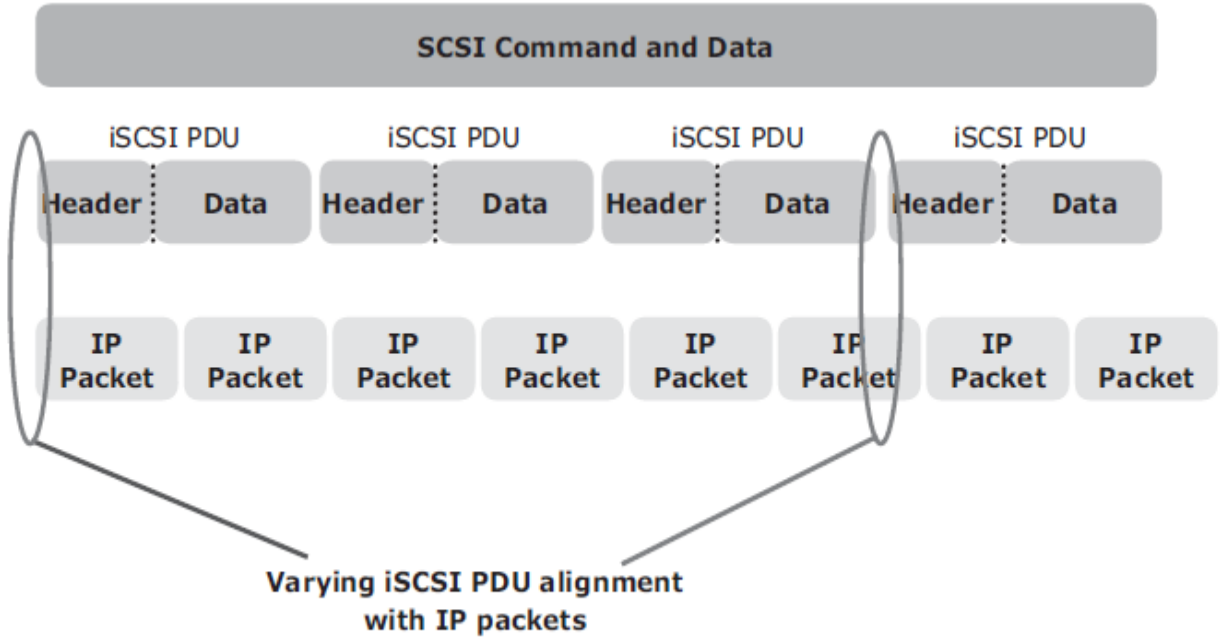


Figure 6-5: Alignment of iSCSI PDUs with IP packets

iSCSI Discovery

An initiator must discover the location of its targets on the network and the names of the targets available to it before it can establish a session.

This discovery can take place in two ways:

1. SendTargets discovery
2. internet Storage Name Service (iSNS).

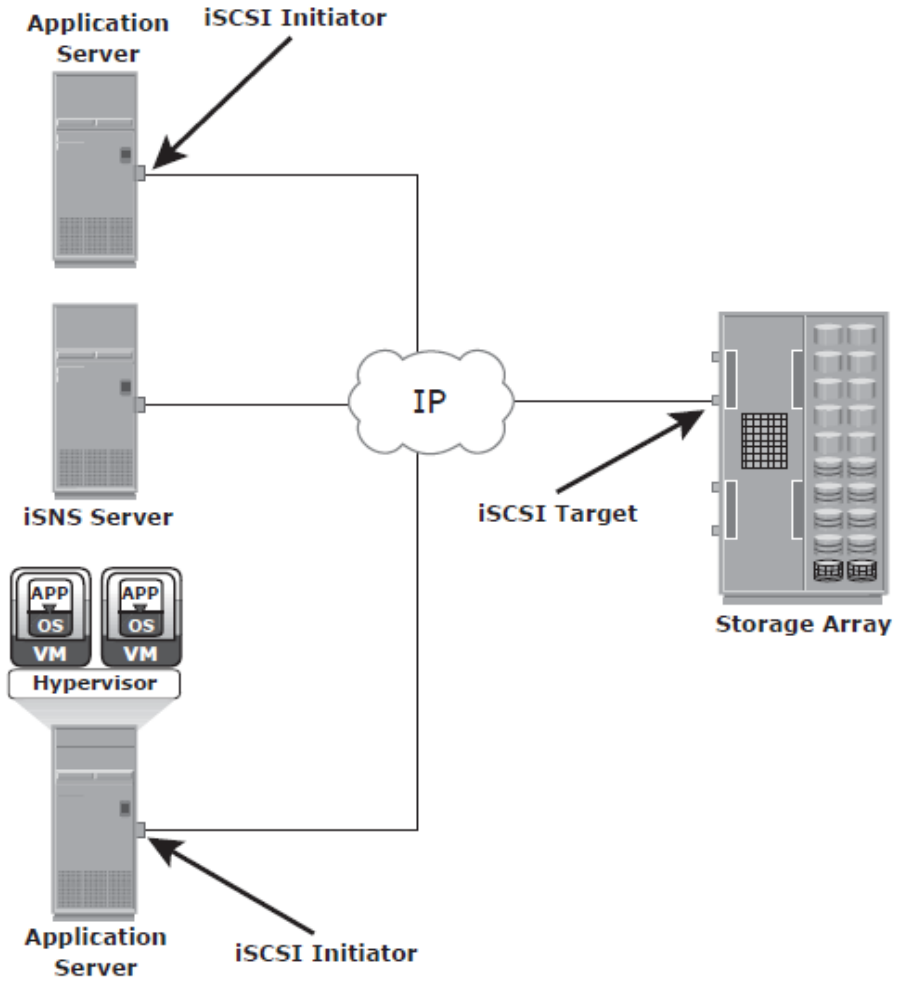


Figure 6-6: Discovery using iSNS

iSCSI Discovery

1. SendTargets discovery

- In SendTargets discovery, the initiator is manually configured with the target's network portal to establish a discovery session.
- The initiator issues the SendTargets command, and the target network portal responds with the names and addresses of the targets available to the host.

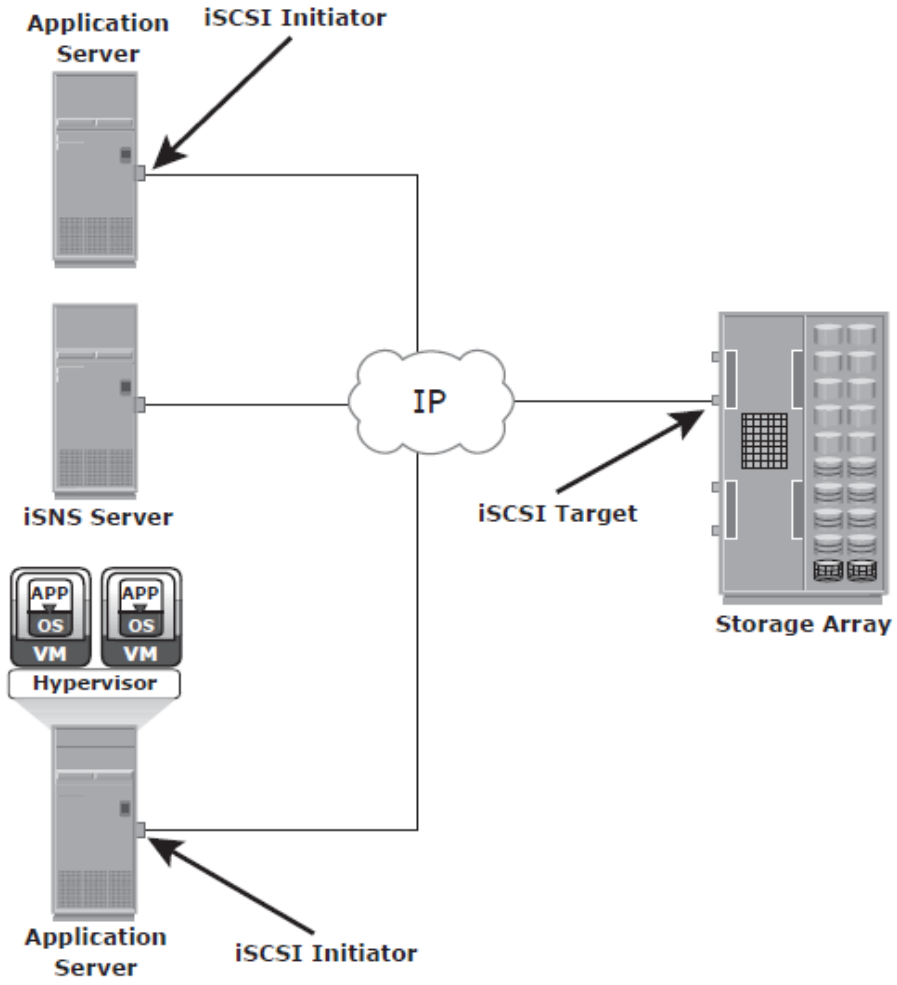


Figure 6-6: Discovery using iSNS

iSCSI Discovery

2. internet Storage Name Service (iSNS).

- iSNS (see Figure 6-6) enables automatic discovery of iSCSI devices on an IP network.
- The initiators and targets can be configured to automatically register themselves with the iSNS server.
- Whenever an initiator wants to know the targets that it can access, it can query the iSNS server for a list of available targets.
- The discovery can also take place by using service location protocol (SLP).

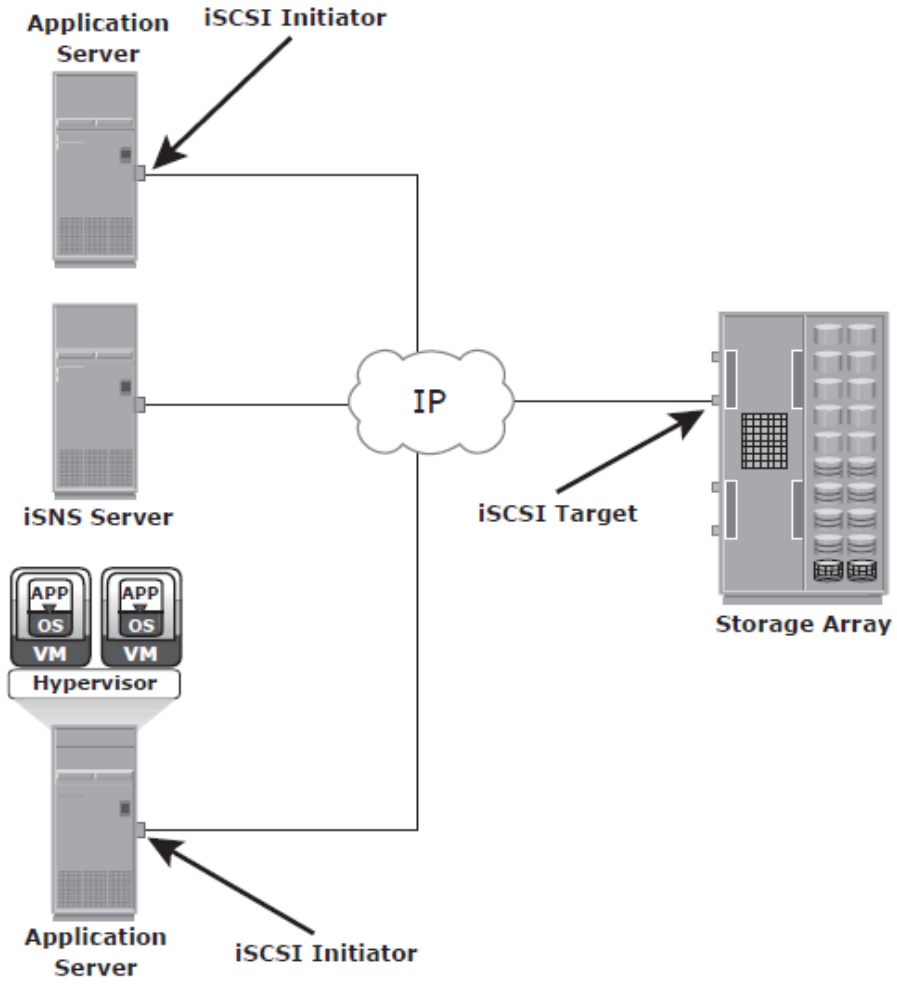


Figure 6-6: Discovery using iSNS

iSCSI Names

A unique worldwide iSCSI identifier is used to identify the initiators and targets within an iSCSI network to facilitate communication.

The unique identifier can be a combination of the names of the department, application, or manufacturer, serial number, asset number, or any tag that can be used to recognize and manage the devices.

Following are two types of iSCSI names commonly used:

1. iSCSI Qualified Name (IQN)
2. Extended Unique Identifier (EUI)

iSCSI Names

1. iSCSI Qualified Name (IQN)

- An iSCSI Qualified Name enables storage administrators to assign meaningful names to iSCSI devices.
- An organization must own a registered domain name to generate iSCSI Qualified Names.
- A date is included in the name to avoid potential conflicts caused by the transfer of domain names.
- An example of an IQN is

iqn.2008-02.com.example:optional_string

- The optional_string provides a serial number, an asset number, or any other device identifiers.

iSCSI Names

2. Extended Unique Identifier (EUI)

An EUI is a globally unique identifier based on the IEEE EUI-64 naming standard.

An EUI is composed of the eui prefix followed by a 16-character hexadecimal name, such as

eui.0300732A32598D26

In either format, the allowed special characters are dots, dashes, and

blank spaces.

iSCSI Session

An iSCSI session is established between an initiator and a target, as shown in Figure 6-7.

A session is identified by a session ID (SSID), which includes part of an initiator ID and a target ID.

The session can be intended for one of the following:

- The discovery of the available targets by the initiators and the location of a specific target on a network
- The normal operation of iSCSI (transferring data between initiators and targets)

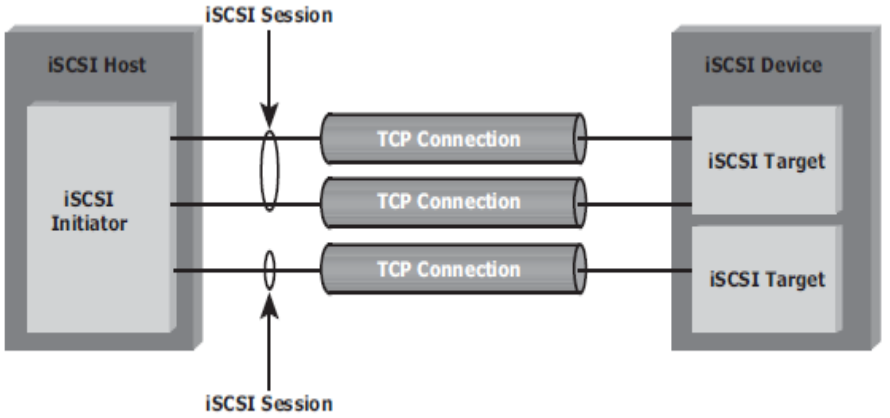


Figure 6-7: iSCSI session

iSCSI Command Sequencing

- The iSCSI communication between the initiators and targets is based on the request-response command sequences.
- Each of the sequence numbers is stored locally as an unsigned 32-bit integer counter defined by iSCSI.

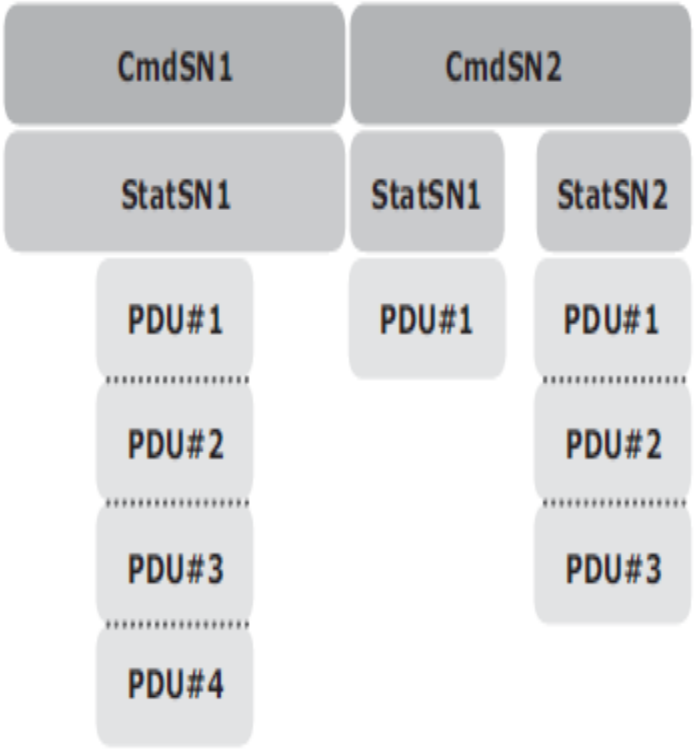


Figure 6-8: Command and status sequence number

iSCSI Command Sequencing

- A command sequence may generate multiple PDUs.
- A command sequence number (CmdSN) within an iSCSI session is used for numbering all initiator-to-target command PDUs belonging to the session.
- This number ensures that every command is delivered in the same order in which it is transmitted, regardless of the TCP connection that carries the command in the session.

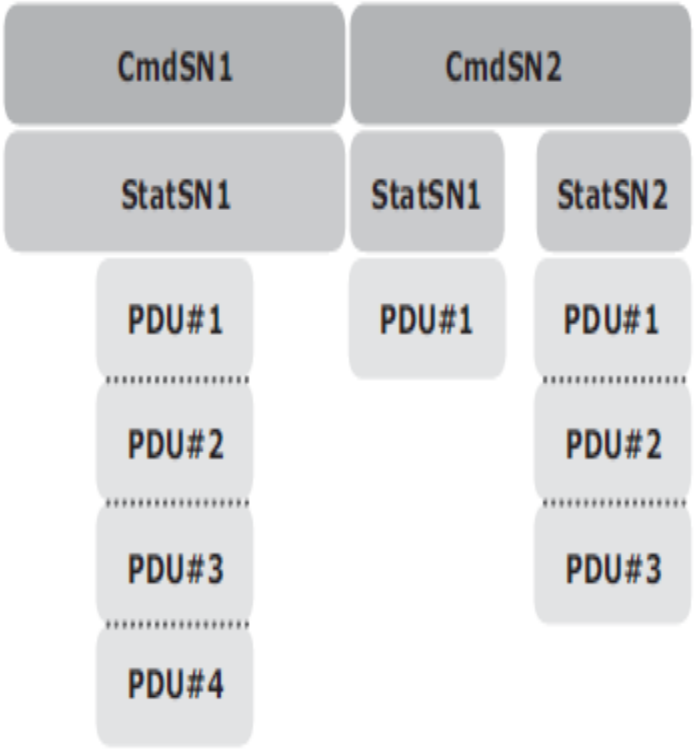


Figure 6-8: Command and status sequence number

iSCSI Command Sequencing

A status sequence number (StatSN) is used to sequentially number status responses, as shown in Figure 6-8.

These unique numbers are established at the level of the TCP connection.

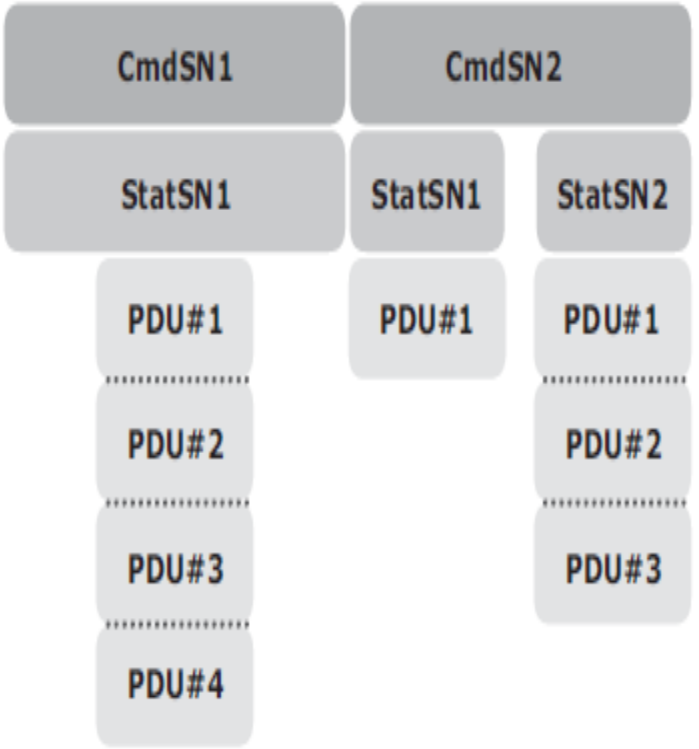


Figure 6-8: Command and status sequence number

iSCSI Command Sequencing

A target sends request-to-transfer (R2T) PDUs to the initiator when it is ready to accept data.

A data sequence number (DataSN) is used to ensure in-order delivery of data within the same command.

The DataSN and R2TSN are used to sequence data PDUs and R2Ts, respectively.

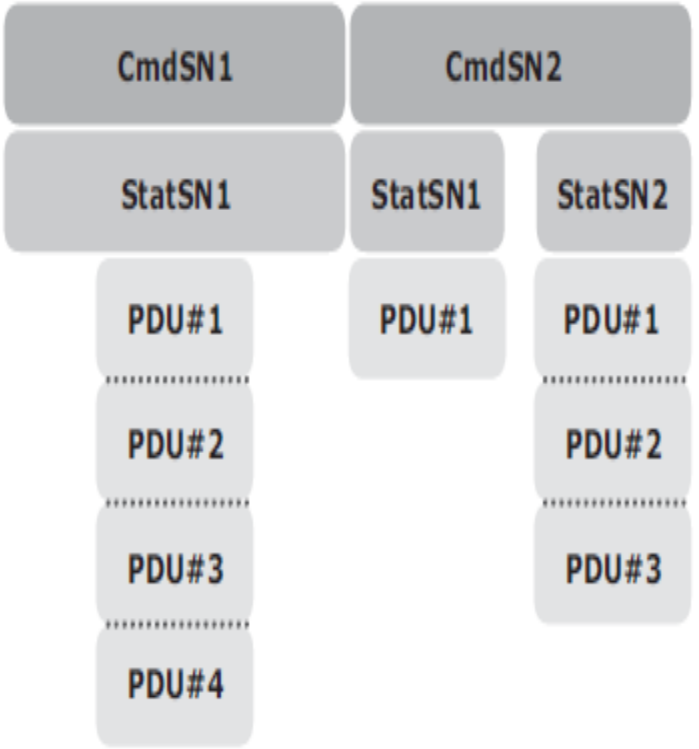


Figure 6-8: Command and status sequence number

iSCSI Command Sequencing

A target sends request-to-transfer (R2T) PDUs to the initiator when it is ready to accept data.

A data sequence number (DataSN) is used to ensure in-order delivery of data within the same command.

The DataSN and R2TSN are used to sequence data PDUs and R2Ts, respectively.

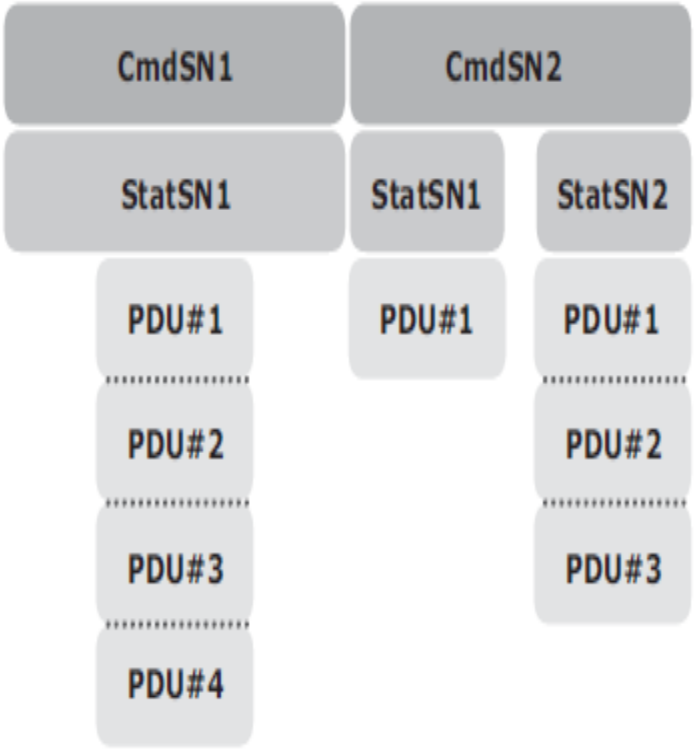


Figure 6-8: Command and status sequence number

UNIT 3- Storage Networking Technologies

IP SAN and FCoE(Chapter 6)

- iSCSI (Small Computer Systems Interface (iSCSI)
- FCIP (Fibre Channel over IP)
- FCoE(Fibre Channel Over Ethernet)

FCIP (Fibre Channel over IP)

FCIP is a tunneling protocol that enables distributed FC SAN islands to be interconnected over the existing IP-based networks.

The FCIP standard has rapidly gained acceptance as a manageable, cost effective way to blend the best of the two worlds: FC SAN and the proven

FCIP is extensively used in disaster recovery implementations in which data is duplicated to the storage located at a remote site.

FCIP Protocol Stack

Applications generate SCSI commands and data, which are processed by various layers of the protocol stack.

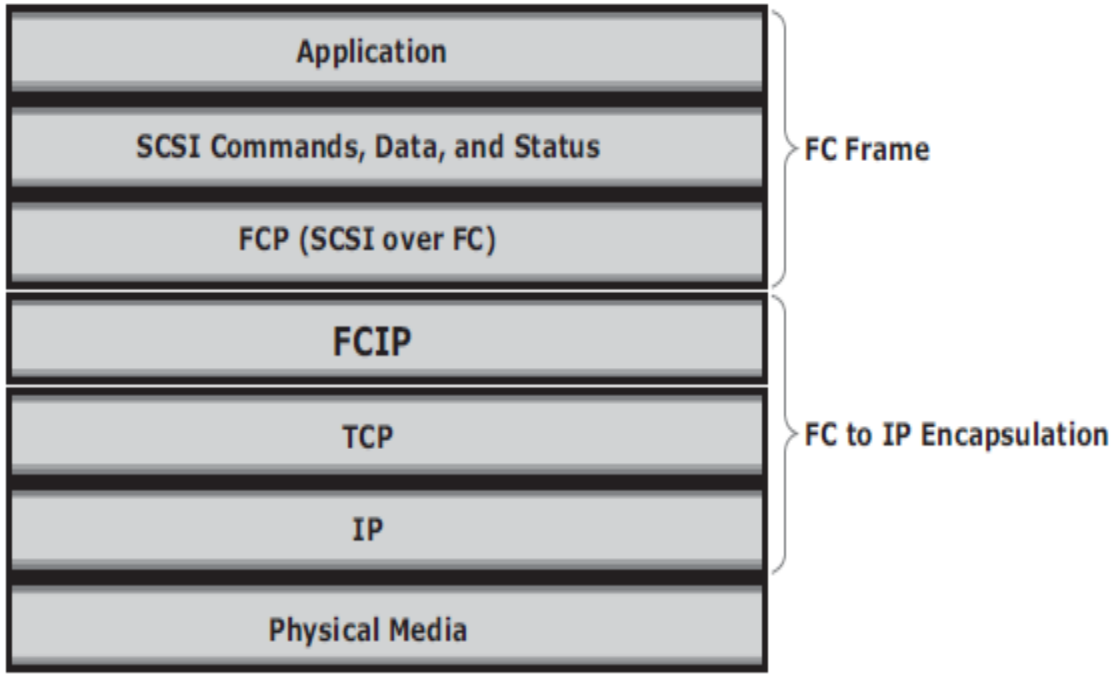


Figure 6-9: FCIP protocol stack

FCIP Protocol Stack

- The upper layer protocol SCSI includes the SCSI driver program that executes the read-and-write commands.
- Fibre Channel Protocol (FCP) layer, which is simply a Fibre Channel frame whose payload is SCSI.
- The FCP layer rides on top of the Fibre Channel transport layer.
- This enables the FC frames to run natively within a SAN fabric environment.

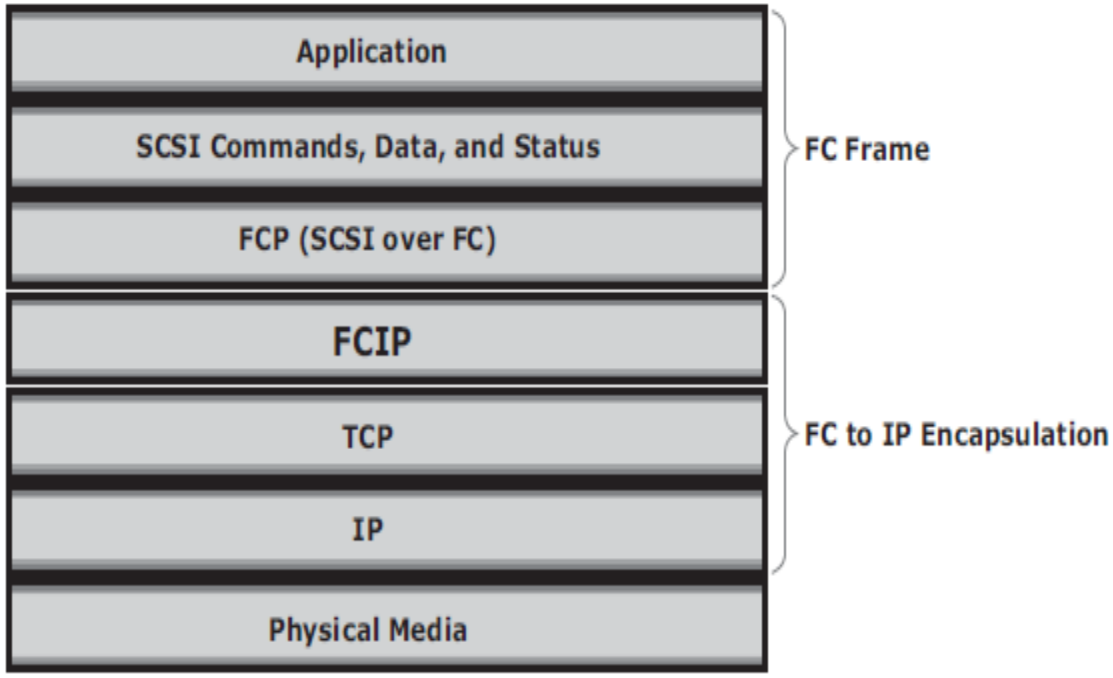


Figure 6-9: FCIP protocol stack

FCIP Protocol Stack

FC frames can be encapsulated into the IP packet and sent to a remote SAN over the IP.

The FCIP layer encapsulates the Fibre Channel frames onto the IP payload and passes them to the TCP layer (see Figure 6-10).

TCP and IP are used for transporting the encapsulated information across Ethernet, wireless, or other media that support the TCP/IP traffic.

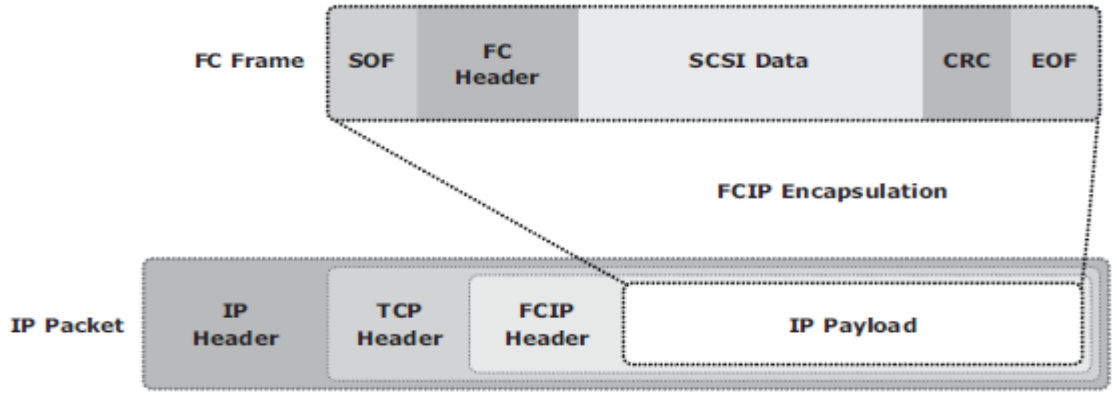


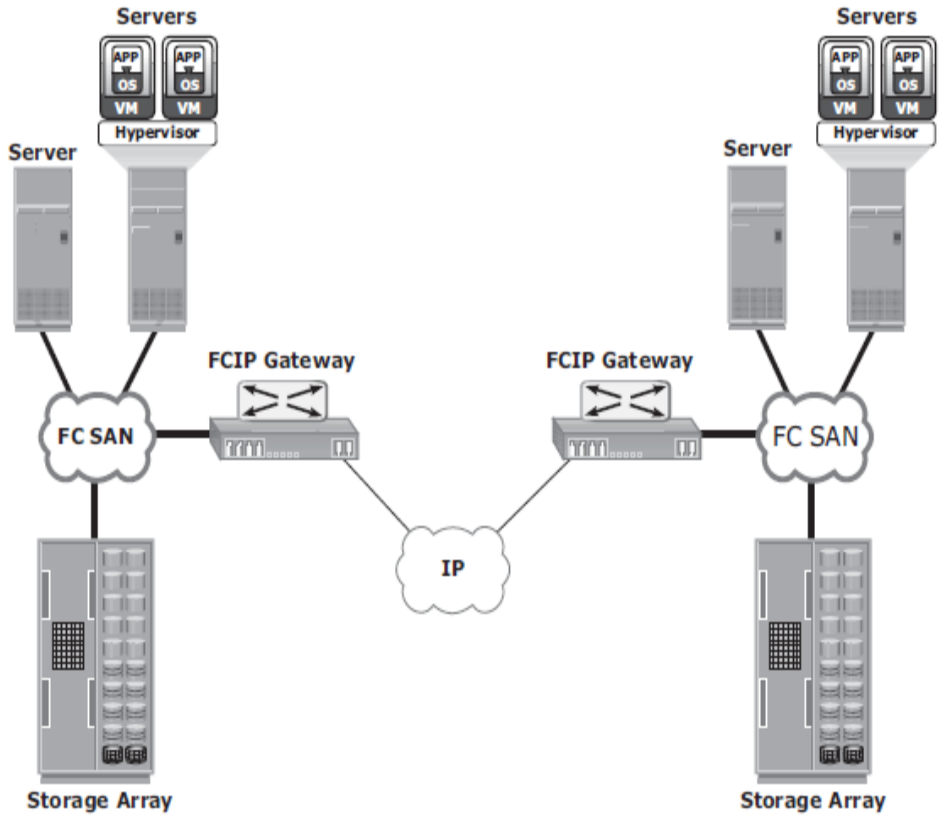
Figure 6-10: FCIP encapsulation

FCIP Topology

In an FCIP environment, an FCIP gateway is connected to each fabric via a standard FC connection.

The FCIP gateway at one end of the IP network encapsulates the FC frames into IP packets.

The gateway at the other end removes the IP wrapper and sends the FC data to the layer 2 fabric.

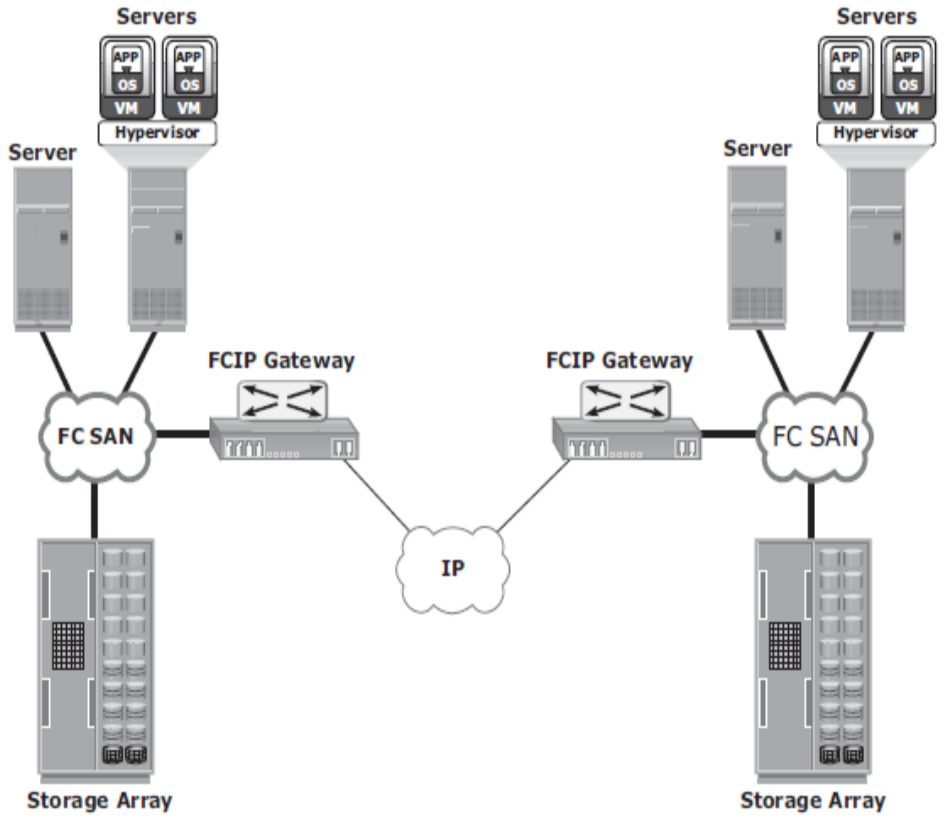


FCIP Topology

The fabric treats these gateways as layer 2 fabric switches.

An IP address is assigned to the port on the gateway, which is connected to an IP network.

After the IP connectivity is established, the nodes in the two independent fabrics can communicate with each other.



FCIP Performance and Security

When implementing storage solutions, parameters should always be taken into consideration

1. Performance
2. reliability
3. security

RAMAIAH Institute of Technology

UNIT 3- Storage Networking Technologies

IP SAN and FCoE(Chapter 6)

- iSCSI (Small Computer Systems Interface (iSCSI)
- FCIP (Fibre Channel over IP)
- FCoE(Fibre Channel Over Ethernet)

FCoE(Fibre Channel Over Ethernet)

- Data centers typically have multiple networks to handle various types of I/O Traffic —
- Example,
  - Ethernet network for TCP/IP communication
  - FC network for FC communication.
- TCP/IP -used for client-server communication, data backup, infrastructure management communication, and so on.
- FC -used for moving block-level data between storage and servers.
- To support multiple networks, servers in a data center are equipped with multiple redundant physical network interfaces — for example, multiple Ethernet and FC cards/adapters.

FCoE(Fibre Channel Over Ethernet)

- To enable the communication, different types of networking switches and physical cabling infrastructure are implemented in data centers.
- Fibre Channel over Ethernet (FCoE) protocol provides consolidation of LAN and SAN traffic over a single physical interface infrastructure.
- FCoE uses the Converged Enhanced Ethernet (CEE) link (10 Gigabit Ethernet) to send FC frames over Ethernet.

RAMAIAH
Institute of Technology

I/O Consolidation Using FCoE

- Figure 6-12 represents the infrastructure before FCoE deployment.
- Here, the storage resources are accessed using HBAs, and the IP network resources are accessed using NICs by the servers.
- Typically, in a data center, a server is configured with 2 to 4 NIC cards and redundant HBA cards.

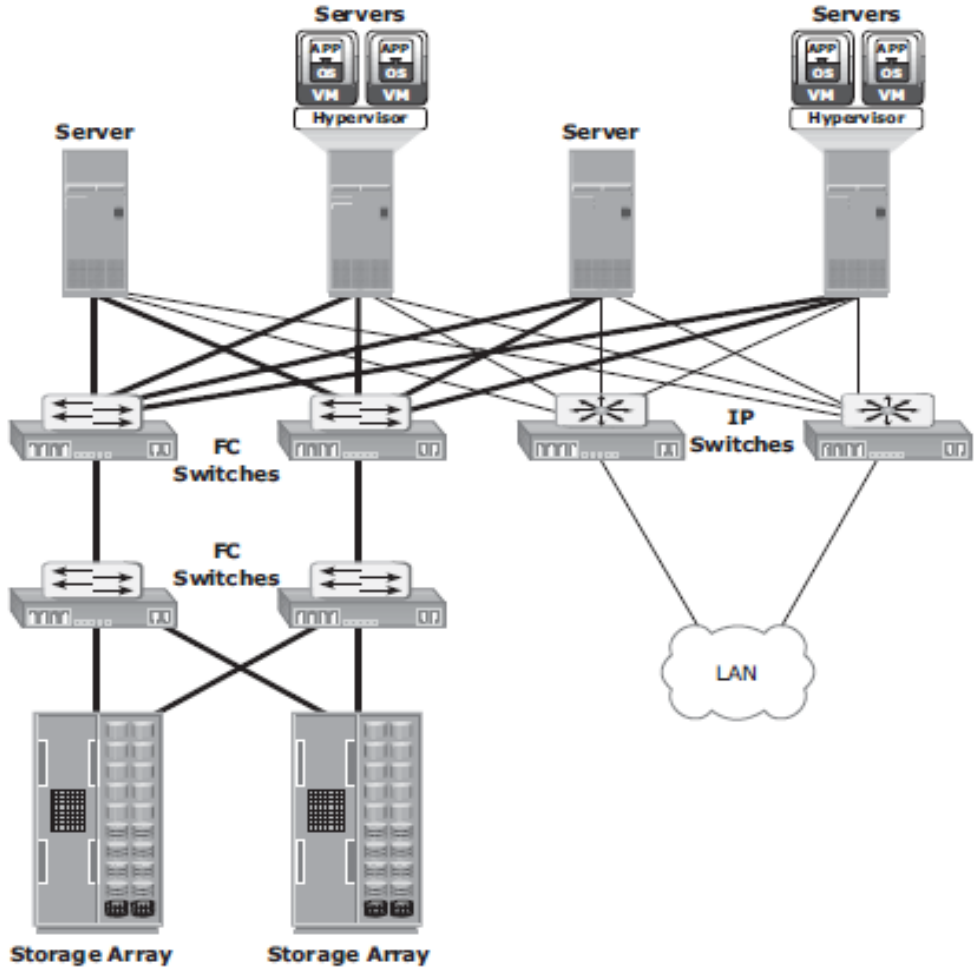


Figure 6-12: Infrastructure before using FCoE

RAMAIAH
Institute of Technology

I/O Consolidation Using FCoE

- If the data center has hundreds of servers, it would require a large number of adapters, cables, and switches.
- This leads to a complex environment, which is difficult to manage and scale.
- The cost of power, cooling, and floor space further adds to the challenge.

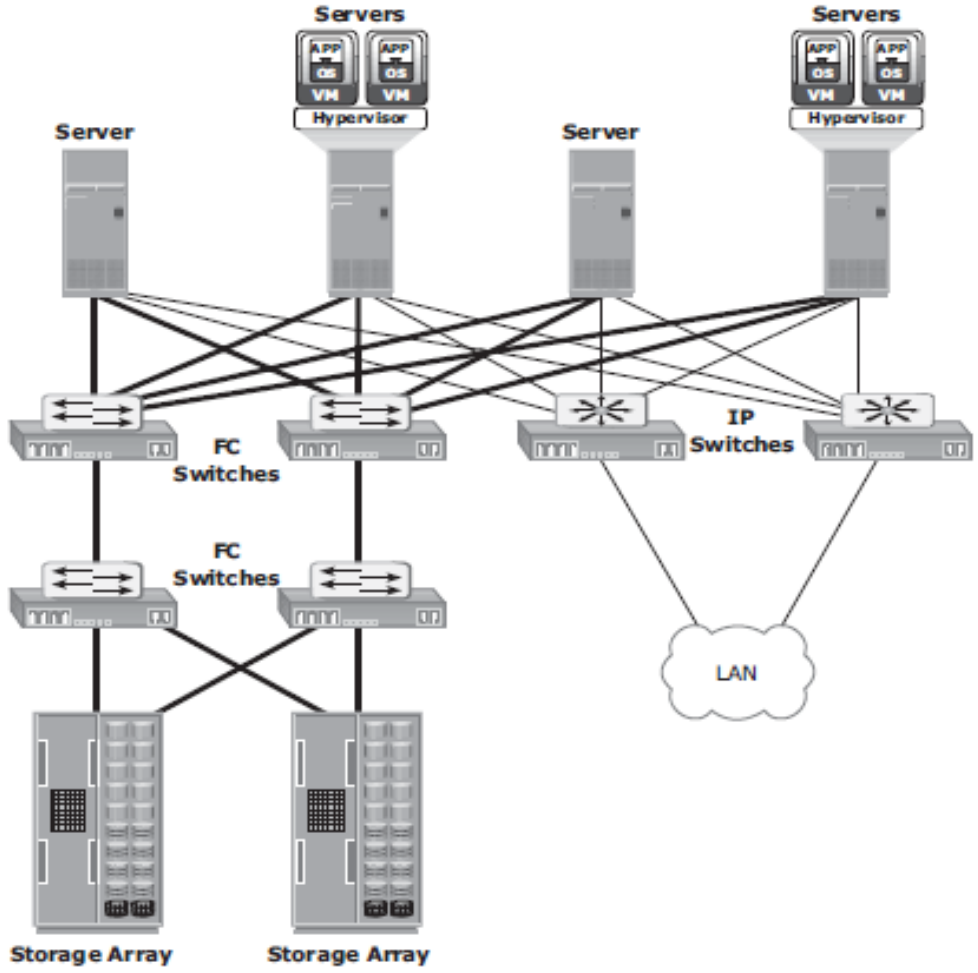


Figure 6-12: Infrastructure before using FCoE

RAMAIAH Institute of Technology

I/O Consolidation Using FCoE

- Figure 6-13 shows the I/O consolidation with FCoE using FCoE switches and Converged Network Adapters (CNAs).
- A CNA replaces both HBAs and NICs in the server and consolidates both the IP and FC traffic.
- This reduces the requirement of multiple network adapters at the server to connect to different networks.
- This reduces the requirement of adapters, cables, and switches.
- This also considerably reduces the cost and management overhead.

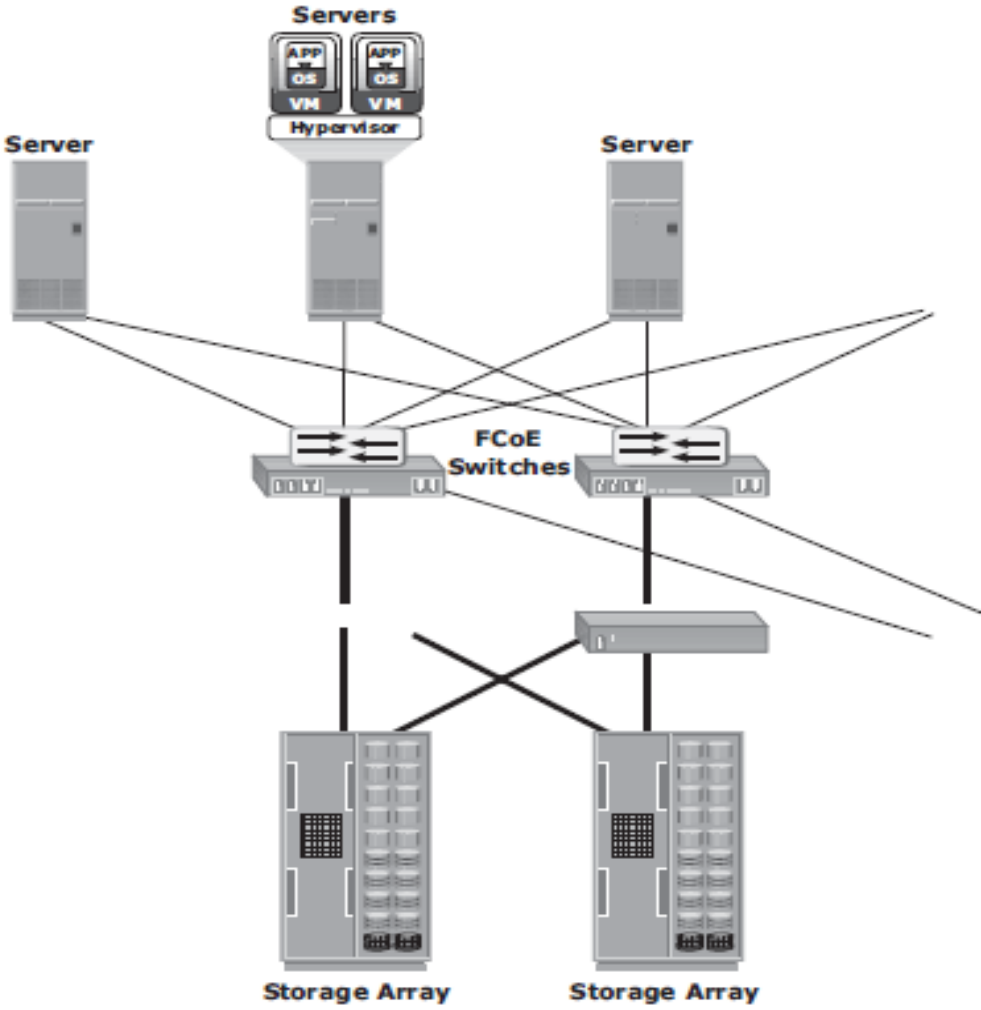


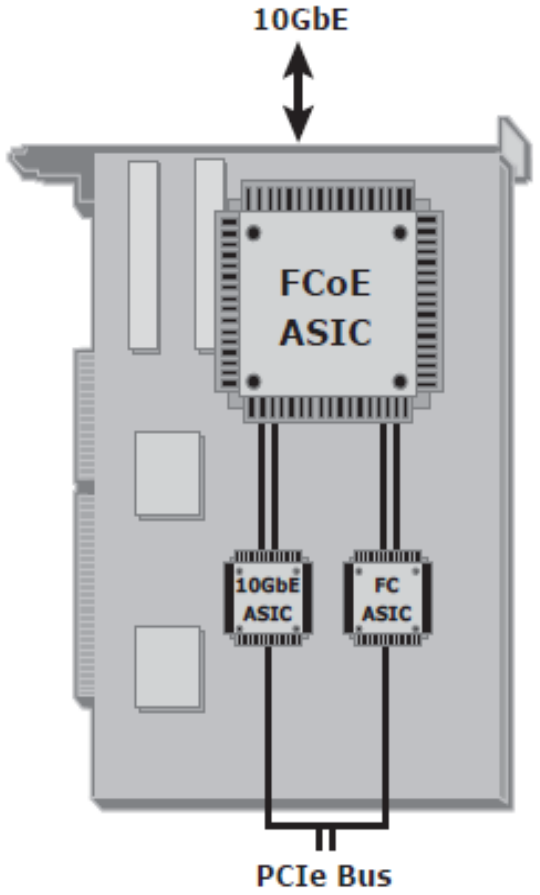
Figure 6-13: Infrastructure after using FCoE

Converged Network Adapter

A CNA provides the functionality of both a standard NIC and an FC HBA in a single adapter and consolidates both types of traffic.

CNA eliminates the need to deploy separate adapters and cables for FC and Ethernet communications, thereby reducing the required number of server slots and switch ports.

CAN offloads the FCoE protocol processing task from the server, thereby freeing the server CPU resources for application processing.

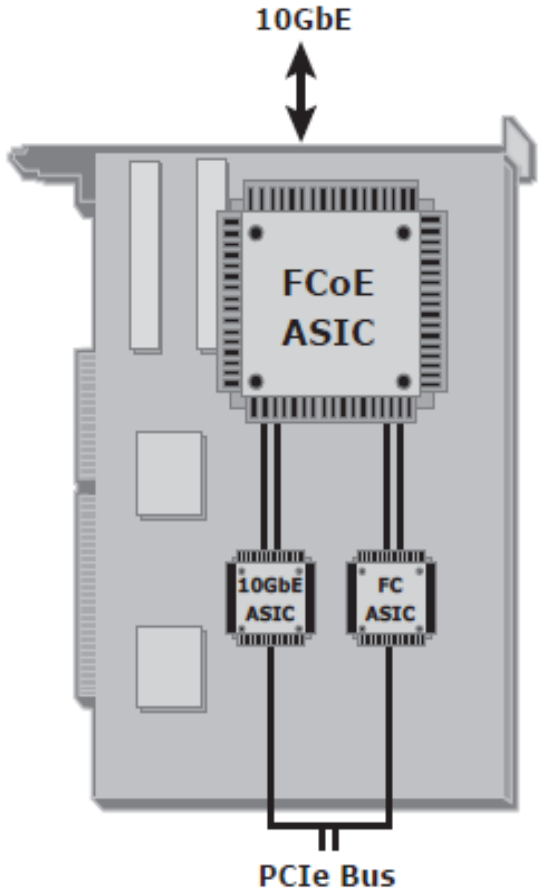


Converged Network Adapter

As shown in Figure 6-14, a CAN contains separate modules for 10 Gigabit Ethernet, Fibre Channel, and FCoE Application Specific Integrated Circuits (ASICs).

The FCoE ASIC encapsulates FC frames into Ethernet frames.

One end of this ASIC is connected to 10GbE and FC ASICs for server connectivity, while the other end provides a 10GbE interface to connect to an FCoE switch.



Cables

Two FCoE cabling:

1. Copper based Twinax
2. standard fiber optical cables.

A Twinax cable is composed of two pairs of copper cables covered with a shielded casing.

The Twinax cable can transmit data at the speed of 10 Gbps over shorter distances up to 10 meters.

Twinax cables require less power and are less expensive than fiber optic cables.

The Small Form Factor Pluggable Plus (SFP+) connector is the primary connector used for FCoE links and can be used with both optical and copper cables.

FCoE Switches

- An FCoE switch has both functionalities.
  - Ethernet switch
  - Fibre Channel switch
- The FCoE switch has a
  - Fibre Channel Forwarder (FCF)
  - Ethernet Bridge
  - set of Ethernet ports
  - optional FC ports

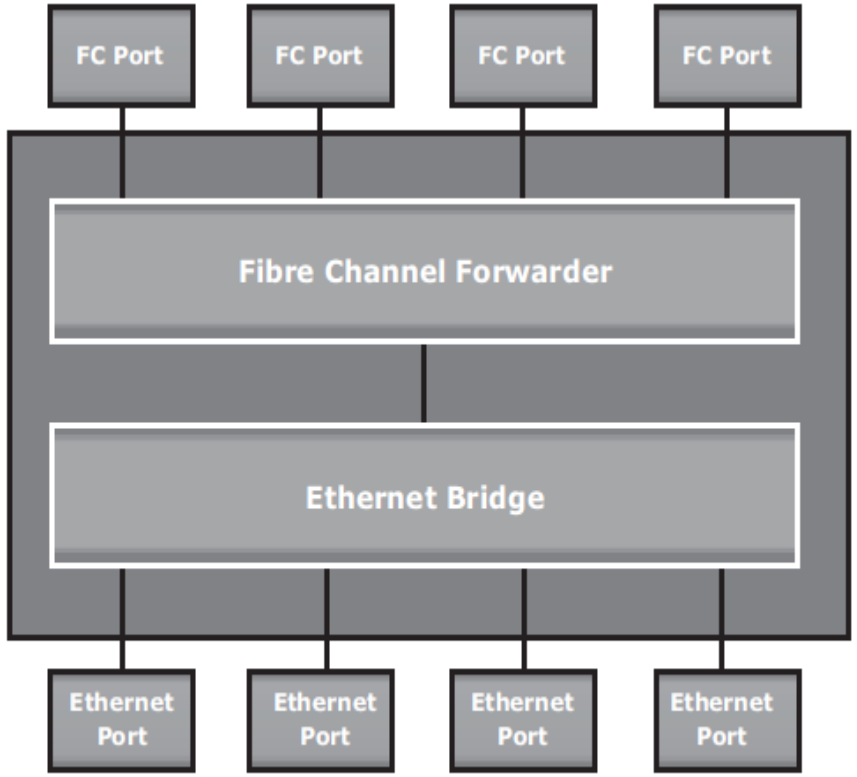


Figure 6-15: FCoE switch generic architecture

FCoE Switches

Fibre Channel Forwarder (FCF)

- The function of the FCF is
  - to encapsulate the FC frames received from the FC port into the FCoE frames and
  - to de-encapsulate the FCoE frames received from the Ethernet Bridge to the FC frames.

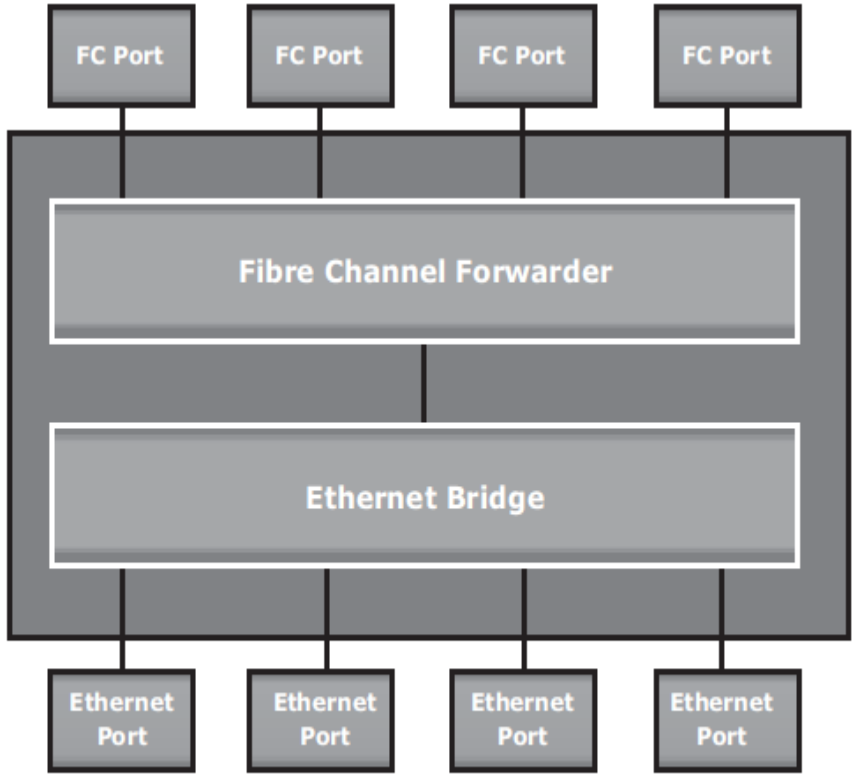


Figure 6-15: FCoE switch generic architecture

FCoE Switches

- Upon receiving the incoming traffic,
  - the FCoE switch inspects the Ethertype of the incoming frames and uses that to determine the destination.
- If the Ethertype of the frame is FCoE the switch recognizes that the frame contains an FC payload and forwards it to the FCF.
- If the Ethertype is not FCoE, the switch handles the traffic as usual Ethernet traffic and forwards it over the Ethernet ports.

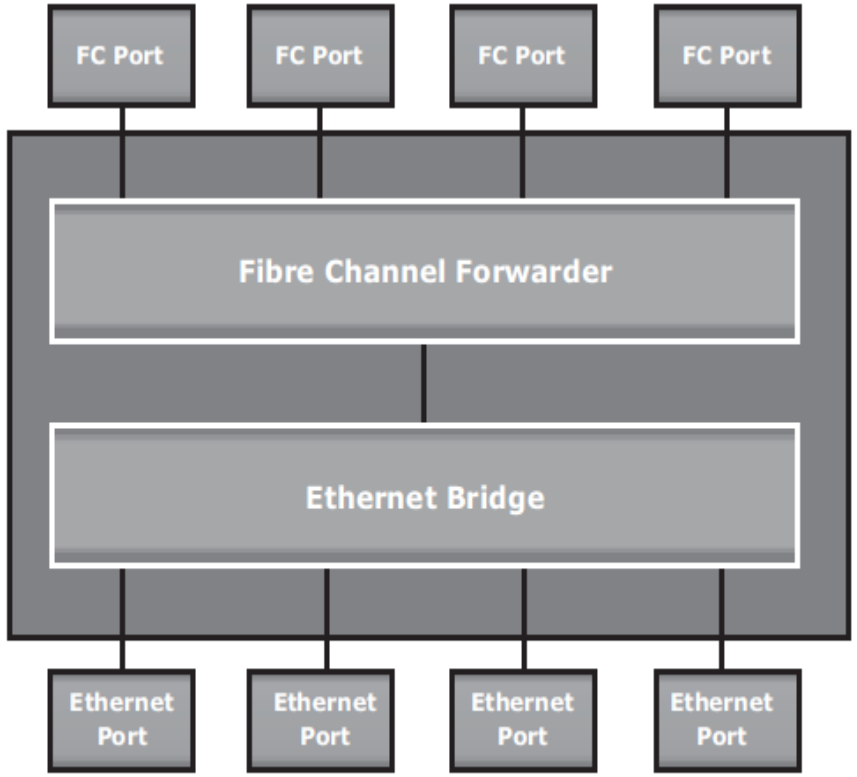


Figure 6-15: FCoE switch generic architecture

FCoE Frame Structure

Figure 6-16 shows the FCoE frame structure.

An FCoE frame is an Ethernet frame that contains an FCoE Protocol Data Unit.

- **48-bits - destination MAC address**
- **48-bits - source MAC address.**
- **32-bit IEEE 802.1Q tag supports the creation of multiple virtual networks (VLANs) across a single physical infrastructure.**

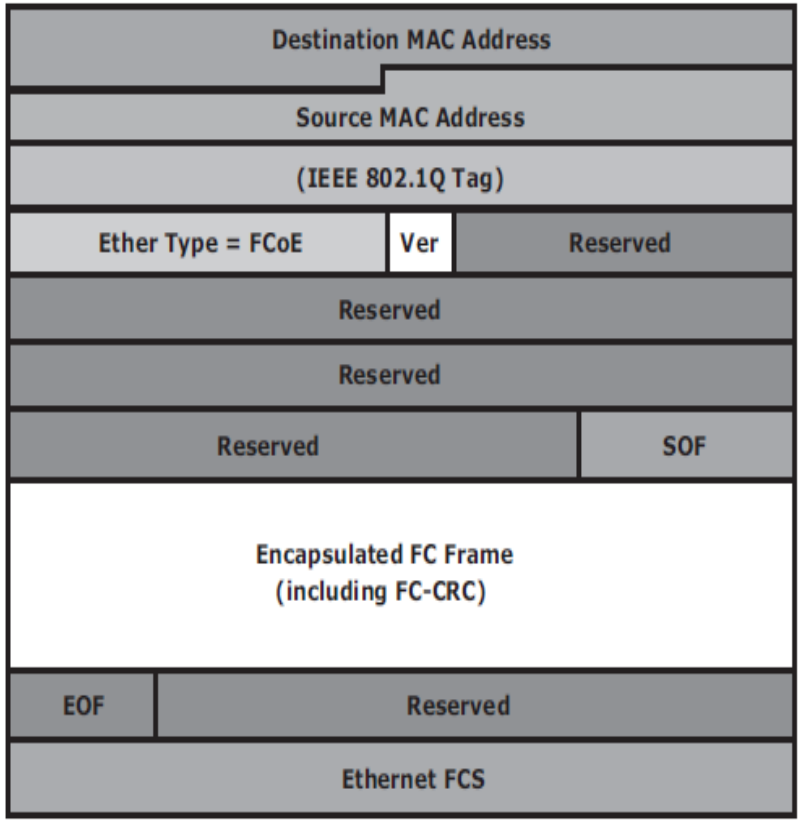


Figure 6-16: FCoE frame structure

FCoE Frame Structure

- **16 bits - FCoE has its own Ethertype**
- **4-bit - version field.**
- **100-bits are reserved**
- **8-bit Start of Frame**

Actual FC frame - The encapsulated Fibre Channel frame consists of the original 24-byte FC header and the data being transported (including the Fibre Channel CRC).

- **8-bit End of Frame delimiter**
- **24 reserved bits**
- **32-bits dedicated to the Frame Check Sequence (FCS) function that provides error detection for the Ethernet frame.**

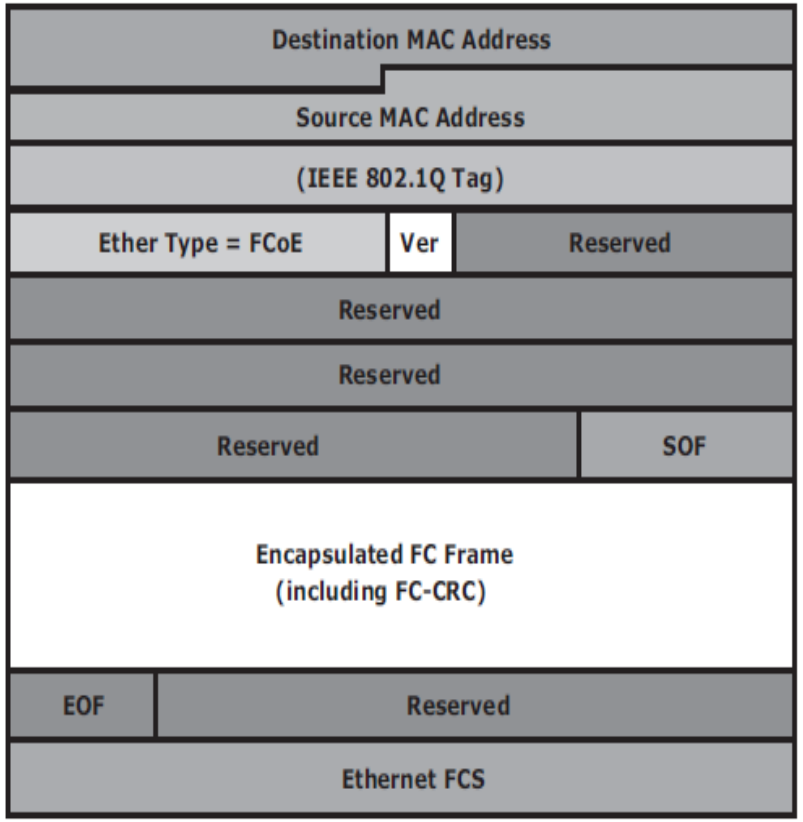


Figure 6-16: FCoE frame structure

FCoE Frame Mapping

The encapsulation of the Fibre Channel frame occurs through the mapping of the FC frames onto Ethernet, as shown in Figure 6-17.

Fibre Channel and traditional networks have stacks of layers where each layer in the stack represents a set of functionalities.

The FC stack consists of five layers:

- FC-0 through FC-4.

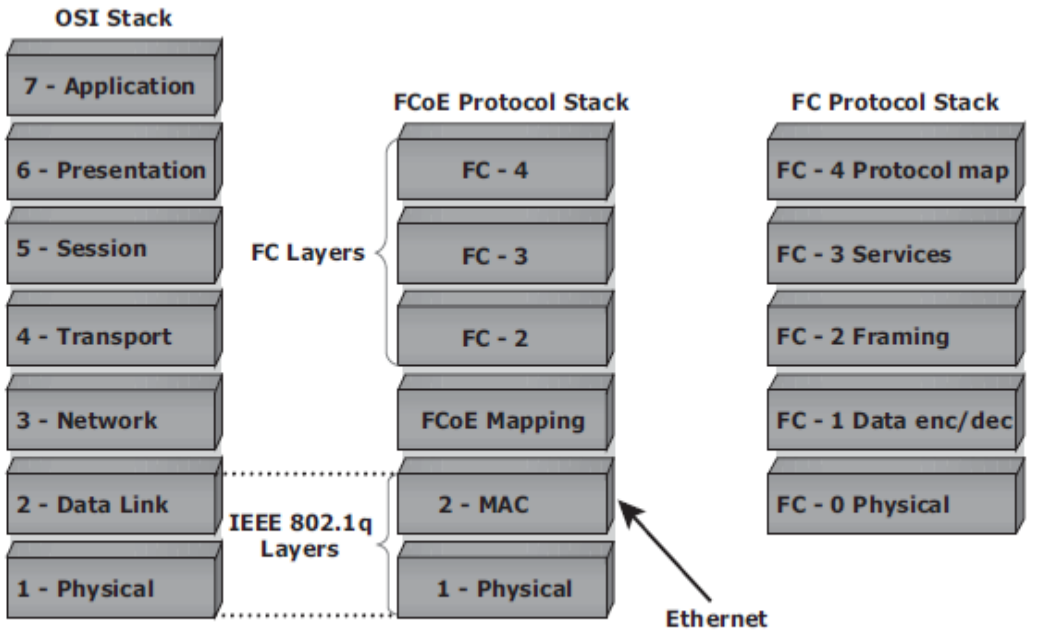


Figure 6-17: FCoE frame mapping

FCoE Frame Mapping

- Ethernet is typically considered as a set of protocols that operates at the physical and data link layers in the seven layer OSI stack.
- The FCoE protocol specification replaces the FC-0 and FC-1 layers of the FC stack with Ethernet.
- This provides the capability to carry the FC-2 to the FC-4 layer over the Ethernet layer.

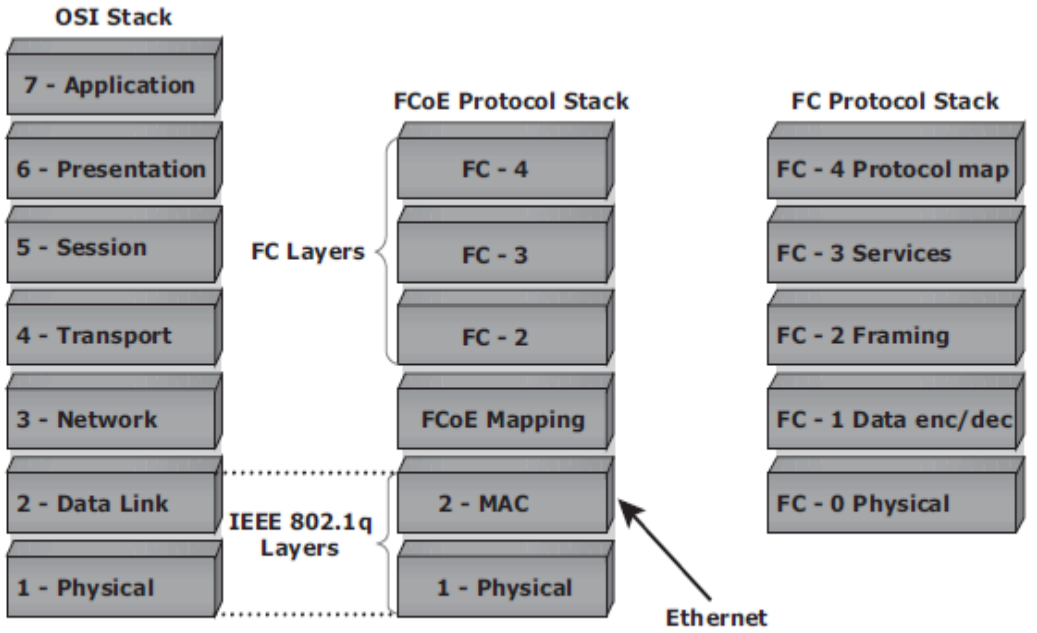


Figure 6-17: FCoE frame mapping

FCoE Enabling Technologies

- Conventional Ethernet is lossy in nature, which means that frames might be dropped or lost during transmission.
- Converged Enhanced Ethernet (CEE)/ lossless Ethernet, provides a new specification to the existing Ethernet standard that eliminates the lossy nature of Ethernet.
- This makes 10 Gb Ethernet a viable storage networking option, similar to FC.
- Lossless Ethernet requires certain functionalities.
- These functionalities are defined and maintained by the data center bridging (DCB) task group, which is a part of the IEEE 802.1 working group, and they are:
  - **Priority-Based Flow Control (PFC)**
  - **Enhanced Transmission Selection (ETS)**
  - **Congestion Notification (CN)**

FCoE Enabling Technologies

Priority-Based Flow Control (PFC)

- PFC provides a link level flow control mechanism.
- Physical Ethernet link is divided into eight virtual links and allows a PAUSE for a single virtual link without affecting the traffic for the others.
- PFC enables the pause mechanism based on user priorities or classes of service.
  - Enabling the pause based on priority allows creating lossless links for traffic, such as FCoE traffic.
- This PAUSE mechanism is typically implemented for FCoE while regular TCP/IP traffic continues to drop frames.

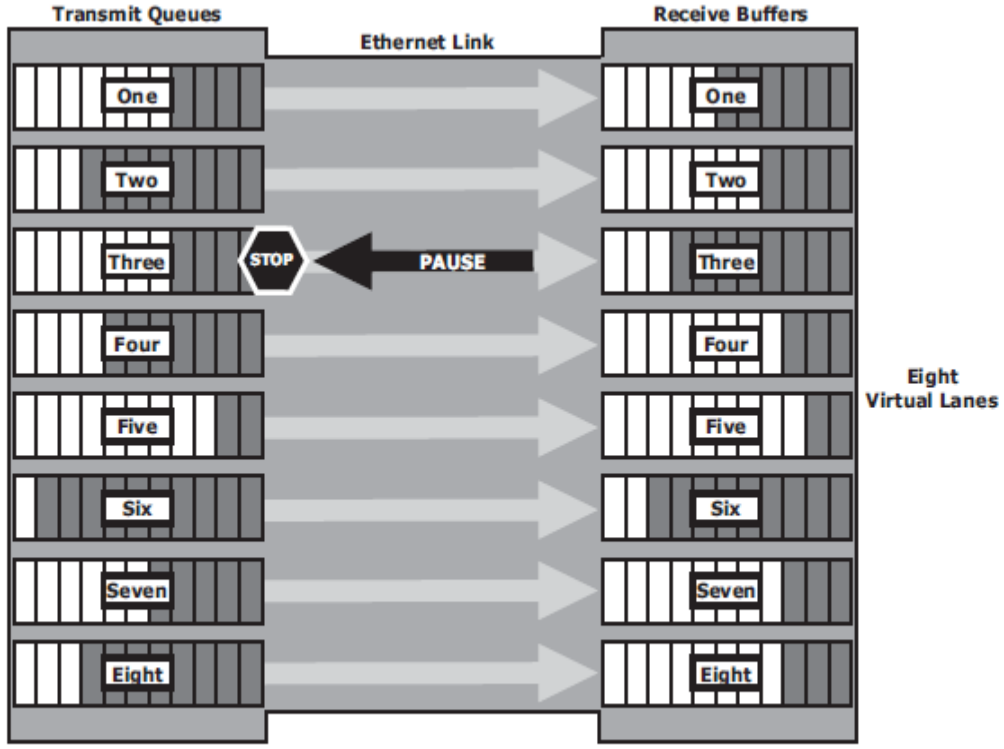


Figure 6-18: Priority-based flow control

FCoE Enabling Technologies

Enhanced Transmission Selection (ETS)

- Enhanced transmission selection provides a common management framework or the assignment of bandwidth to different traffic classes, such as LAN, SAN,and Inter Process Communication (IPC).
- When a particular class of traffic does not use its allocated bandwidth, ETS enables other traffic classes to use the available bandwidth.

FCoE Enabling Technologies

Congestion Notification (CN)

Congestion notification provides end-to-end congestion management for protocols, such as FCoE, that do not have built-in congestion control mechanisms.

Link level congestion notification provides a mechanism for detecting congestion and notifying the source to move the traffic flow away from the congested links.

Link level congestion notification enables a switch to send a signal to other ports that need to stop or slow down their transmissions.

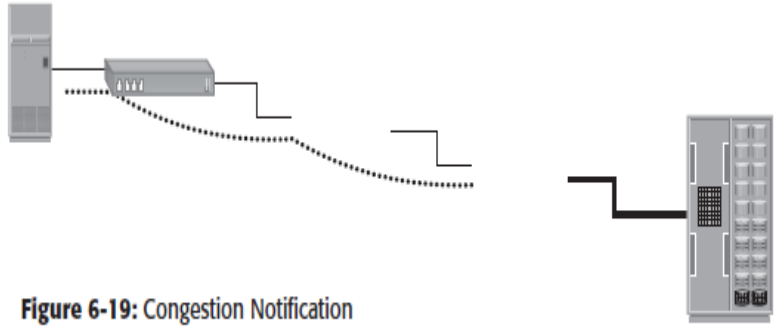


Figure 6-19: Congestion Notification

FCoE Enabling Technologies

The process of congestion notification and its management is shown in Figure 6-19, which represents the communication between the nodes A (sender) and B (receiver).

If congestion at the receiving end occurs, the algorithm running on the switch generates a congestion notification message to the sending node (Node A).

In response to the CN message, the sending end limits the rate of data transfer.

Figure 6-19: Congestion Notification


Thank
you