Module 2

1 List and explain different FC connectivity options with a neat diagram.

FC Connectivity

The FC architecture supports three basic interconnectivity options:

1) Point-To-point,

2) Arbitrated Loop (Fc-AL),

3) FC Switched Fabric

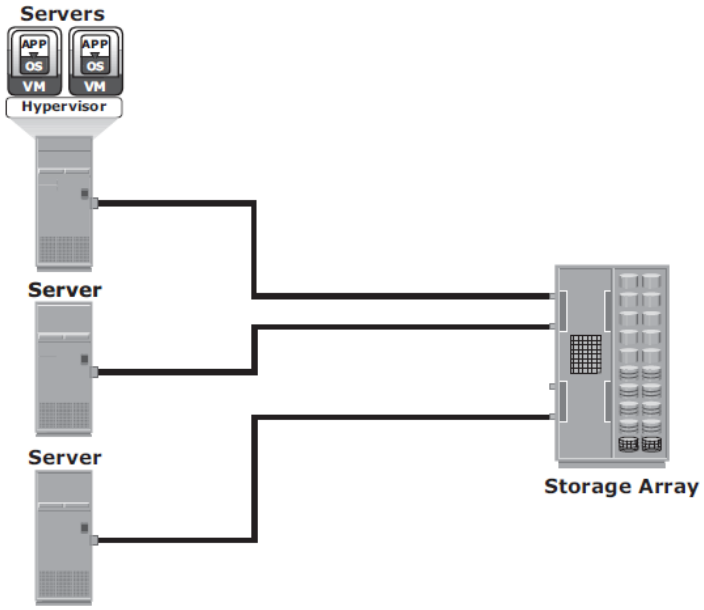
Point-to-Point

 Point-to-point is the simplest FC configuration – two devices are connected directly to each other, as shown in Fig 2.4.

 This configuration provides a dedicated connection for data transmission between nodes.

 The point-to-point configuration offers limited connectivity, as only two devices can communicate with each other at a given time.

 It cannot be scaled to accommodate a large number of network devices. Standard DAS uses point to- point connectivity.



Fibre Channel Arbitrated Loop

 In the FC-AL configuration, devices are attached to a shared loop, as shown in Fig 2.5.

 FC-AL has the characteristics of a token ring topology and a physical star topology.

 In FC-AL, each device contends with other devices to perform I/O operations. Devices on the loop must "arbitrate" to gain control of the loop.

 At any given time, only one device can perform I/O operations on the loop.

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

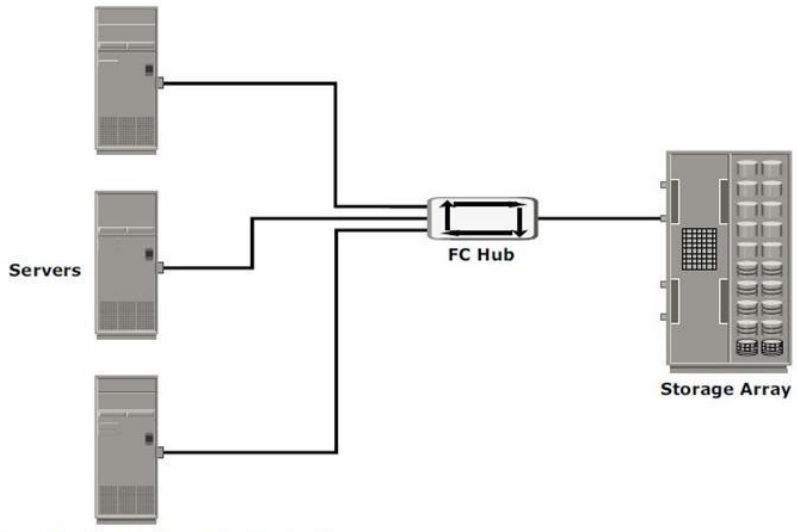
The FC-AL configuration has the following limitations in terms of scalability:

 FC-AL shares the bandwidth in the loop.

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

 FC-AL uses 8-bit addressing. It can support up to 127 devices on a loop.

 Adding or removing a device results in loop re-initialization, which can cause a momentary pause in loop traffic.



Fibre Channel Switched Fabric (FC-SW)

 FC-SW provides dedicated data path and scalability.

 The addition and removal of a device does not affect the on-going traffic between other devices.

 FC-SW is 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.

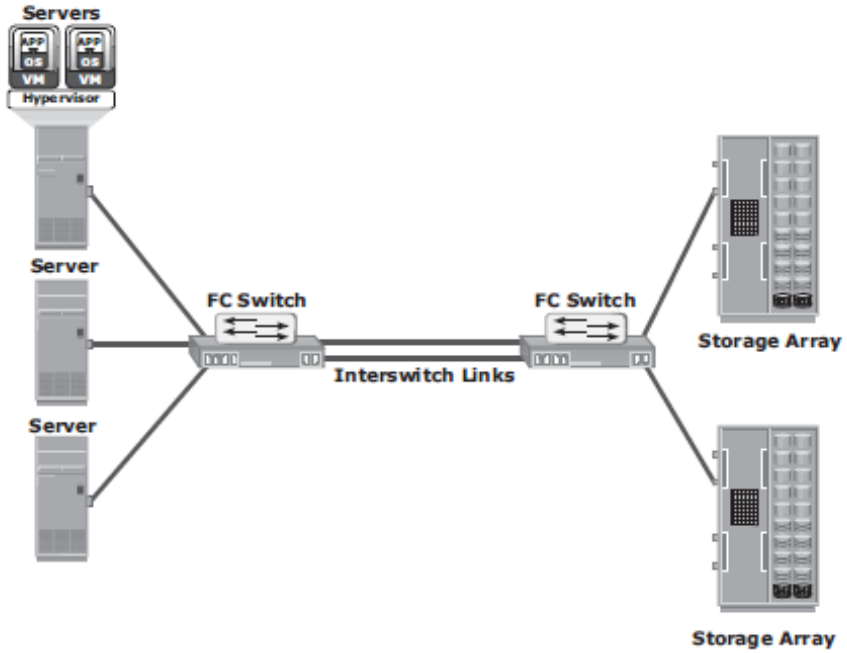
 Each switch in a fabric contains a a unique domain identifi er, which is part of the fabric's addressing scheme.

 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.

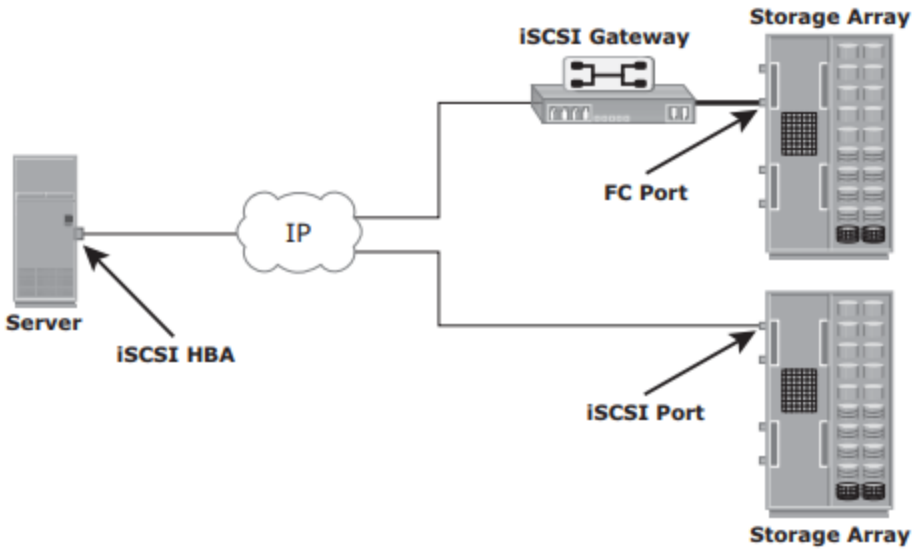


2 With diagram explain ISCSI implementation.

 iSCSI is an IP based protocol that establishes and manages connections between host and storage over IP, as shown in Fig 2.21.

 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



Components of iSCSI

 An initiator (host), target (storage or iSCSI gateway), and an IP-based network are the key iSCSI components.

 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.

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

iSCSI Host Connectivity

The three iSCSI host connectivity options are:

- A standard NIC with software iSCSI initiator,
- a TCP offload engine (TOE) NIC with software iSCSI initiator,
- an iSCSI HBA

 The function of the iSCSI initiator is to route the SCSI commands over an IP network.

 A standard NIC with a software iSCSI initiator is the simplest and least expensive connectivity option. It is easy to implement because most servers come with at least one, and in many cases two, embedded NICs. It requires only a software initiator for iSCSI functionality. 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. TOE NIC helps reduce 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.

 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. The functionality of an iSCSI HBA is similar to the functionality of an FC HBA.

iSCSI Topologies

 Two topologies of iSCSI implementations are native and bridged.

 Native topology does not have FC components.

 The initiators may be either directly attached to targets or connected through the IP network.

 Bridged topology enables the coexistence of FC with IP by providing iSCSI-to-FC bridging functionality.

 For example, the initiators can exist in an IP environment while the storage remains in an FC environment.

Native iSCSI Connectivity

 FC components are not required for iSCSI connectivity if an iSCSI-enabled array is deployed.

 In Fig 2.22(a), the array has one or more iSCSI ports configured with an IP address and is connected to a standard Ethernet switch.

 After an initiator is logged on to the network, it can access the available LUNs on the storage array.

 A single array port can service multiple hosts or initiators as long as the array port can handle the amount of storage traffic that the hosts generate.

Bridged iSCSI Connectivity

 A bridged iSCSI implementation includes FC components in its configuration.

 Fig 2.22(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.

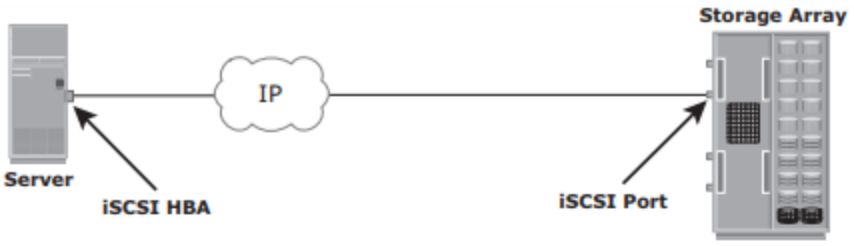
 The bridge devices contain both FC and Ethernet ports to facilitate the communication between the FC and IP environments.

 In a bridged iSCSI implementation, the iSCSI initiator is configured with the gateway's IP address as its target destination.

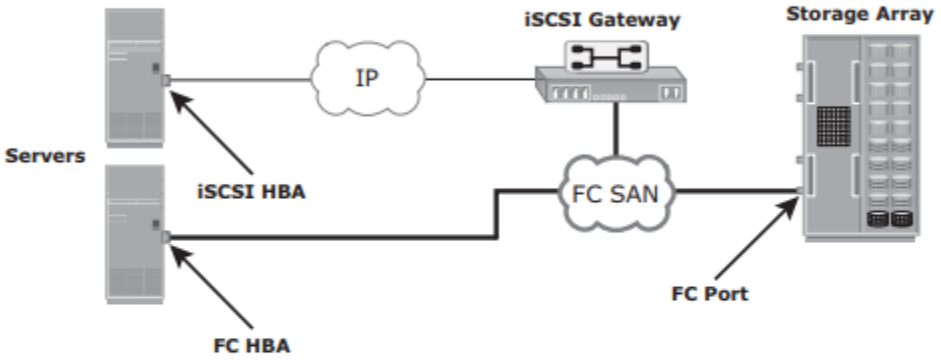
 On the other side, the gateway is configured as an FC initiator to the storage array.

Combining FC and Native iSCSI Connectivity:

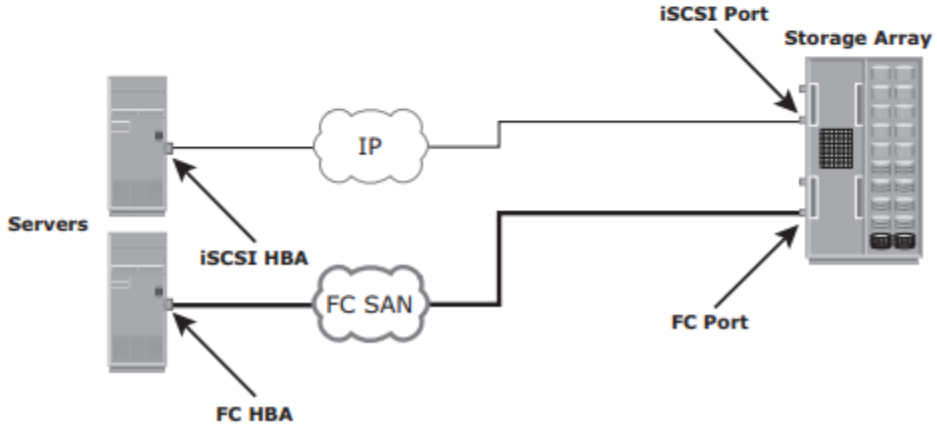
The most common topology is a combination of FC and native iSCSI. Typically, a storage array comes with both FC and iSCSI ports that enable iSCSI and FC connectivity in the same environment, as shown in Fig 2.22(c).



(a) Native iSCSI Connectivity



(b) Bridged iSCSI Connectivity



(c) Combining FC and Native iSCSI Connectivity

3 What is NAS? Explain NAS implementation in detail

What is NAS?

- NAS is an IP based dedicated, high-performance file sharing and storage device.
- Enables NAS clients to share files over an IP network.
- Uses network and file-sharing protocols to provide access to the file data.
- Ex: Common Internet File System (CIFS) and Network File System (NFS).
- Enables both UNIX and Microsoft Windows users to share the same data seamlessly.

 NAS device uses its own operating system and integrated hardware and software components to meet specific file-service needs.

 Its operating system is optimized for file I/O which performs better than a general-purpose server.

 A NAS device can serve more clients than general-purpose servers and provide the benefit of server consolidation.

Components of NAS

 NAS device has two key components (as shown in Fig 2.33): NAS head and storage.

 In some NAS implementations, the storage could be external to the NAS device and shared with other hosts.

 NAS head includes the following components:

 CPU and memory

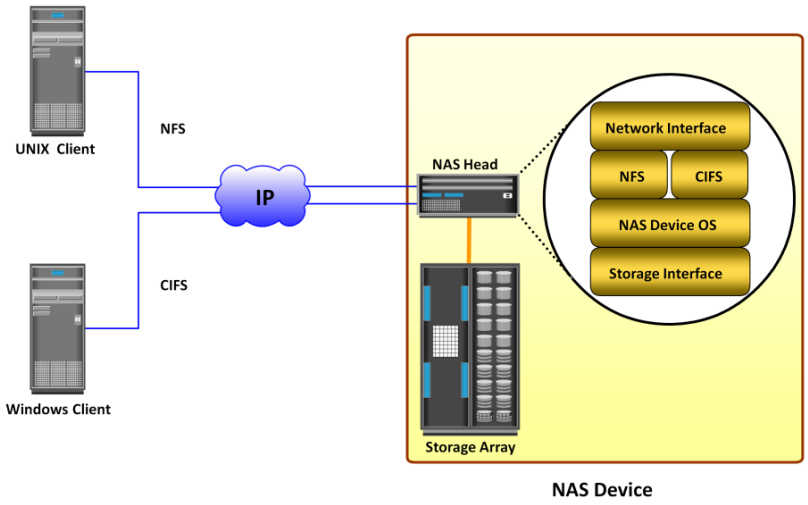
 One or more network interface cards (NICs), which provide connectivity to the client network.

 An optimized operating system for managing the NAS functionality. It translates file-level requests into block-storage requests and further converts the data supplied at the block level to file data

 NFS, CIFS, and other protocols for file sharing

 Industry-standard storage protocols and ports to connect and manage physical disk resources

 The NAS environment includes clients accessing a NAS device over an IP network using filesharing protocols.



4 List the key features of Content Addressed Storage (CAS). Illustrate with a neat block diagram the unified storage for CAS system.

CAS is an object-based system that has been built for storing fixed content data. It is designed for secure online storage and retrieval of fixed content.

CAS stores user data and its attributes as separate objects. The stored object is assigned a globally unique address known as a content address (CA). This address is derived from the object's binary representation. CAS provides an optimized and centrally managed storage solution that can support single-instance storage (SiS) to eliminate multiple copies of the same data.

This chapter describes fixed content and archives, traditional solutions deployed for archives and their limitations, the features and benefits of CAS, CAS architecture, storage and retrieval in a CAS environment, and examples of CAS solutions.

Features refer page no. 187 in text-book

Unified Storage

- Unified storage consolidates block, file, and object access into one storage solution.
- It supports multiple protocols, such as CIFS, NFS, iSCSI, FC, FCoE, REST (representational state transfer), and SOAP (simple object access protocol).

Components of Unified Storage

A unified storage system consists of the following key components:

- storage controller,
- NAS head,
- OSD node,
- storage.

Fig 2.41 illustrates the block diagram of a unified storage platform.

The storage controller or storage processor provides block-level access to application servers through iSCSI, FC, or FCoE protocols. It contains the corresponding front-end ports for direct block access. The storage controller is also responsible for managing the back-end storage pool in the storage system.

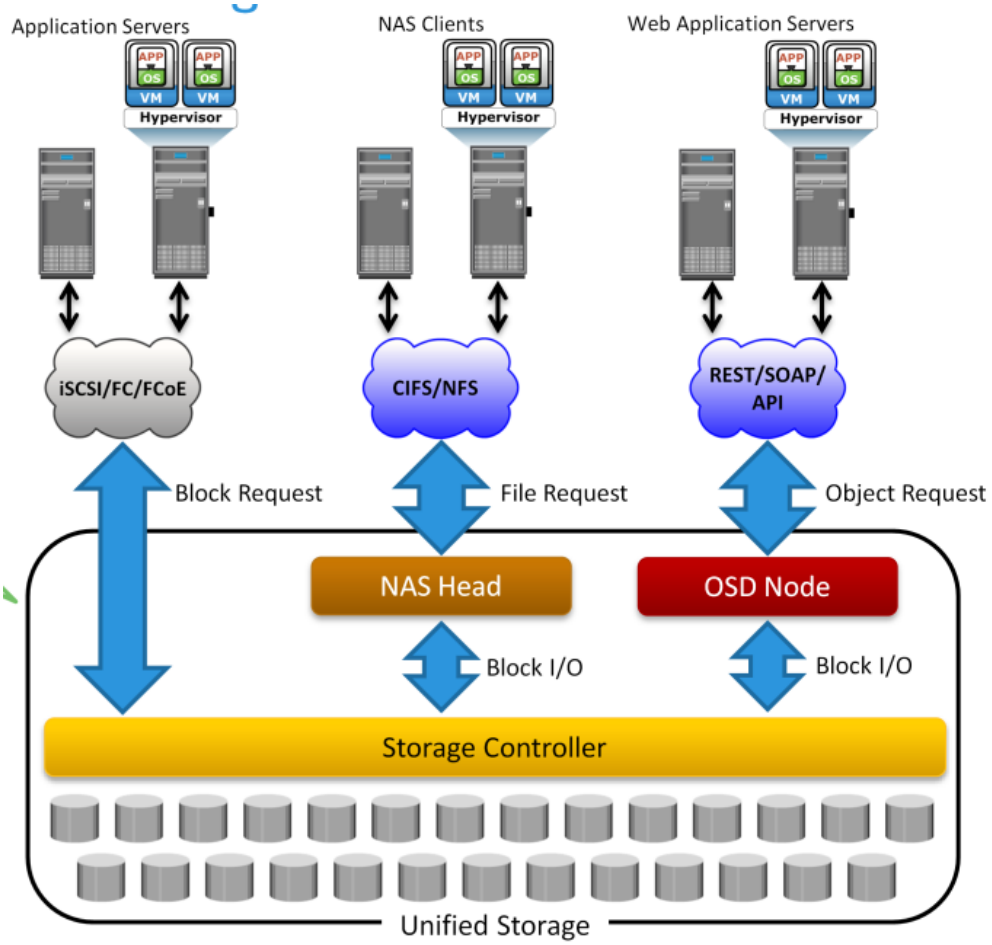
 The controller configures LUNs and presents them to application servers, NAS heads, and OSD nodes. The LUNs presented to the application server appear as local physical disks. A file system is configured on these LUNs and is made available to applications for storing data.

 A NAS head is a dedicated file server that provides file access to NAS clients. The NAS head is connected to the storage via the storage controller typically using a FC or FCoE connection. The system typically has two or more NAS heads for redundancy.

 The LUNs presented to the NAS head appear as physical disks. The NAS head configures the file systems on these disks, creates a NFS, CIFS, or mixed share, and exports the share to the NAS clients.

 The OSD node also accesses the storage through the storage controller using a FC or FCoE connection.

 The LUNs assigned to the OSD node appear as physical disks. These disks are configured by the OSD nodes, enabling them to store the data from the web application servers.



Data Access from Unified Storage

 In a unified storage system, block, file, and object requests to the storage travel through different I/O paths. Fig 2.41 also illustrates the different I/O paths for block, file, and object access.

 Block I/O request: The application servers are connected to an FC, iSCSI, or FCoE port on the storage controller. The server sends a block request and the storage processor (SP) processes the I/O and responds to the application server.

 File I/O request: The NAS clients send a file request to the NAS head using NFS or CIFS protocol. The NAS head receives the request, converts it into a block request, and forwards it to the storage controller. Upon receiving the block data, the NAS head again converts the block request back to the file request and sends back it to the clients.

 Object I/O request: The web application servers send an object request, typically using REST or SOAP protocols, to the OSD node. The OSD node receives the request, converts it into a block request, and sends it to the disk through the storage controller. The controller in turn processes the block request and responds back to the OSD node, which in turn provides the requested object to the web application server.

5 Explain block-level storage virtualization with neat diagram. Explain VSAN in brief

SAN based virtualization and VSAN technology

There are two network-based virtualization techniques in a SAN environment:

 block-level storage virtualization

 virtual SAN (VSAN).

Block-level Storage Virtualization

 Block-level storage virtualization aggregates block storage devices (LUNs) and enables provisioning of virtual storage volumes, independent of the underlying physical storage.

 A virtualization layer, which exists at the SAN, abstracts the identity of physical storage devices and creates a storage pool from heterogeneous storage devices.

 Virtual volumes are created from the storage pool and assigned to the hosts.

 Instead of being directed to the LUNs on the individual storage arrays, the hosts are directed to the virtual volumes provided by the virtualization layer.

 For hosts and storage arrays, the virtualization layer appears as the target and initiator devices, respectively.

 The virtualization layer maps the virtual volumes to the LUNs on the individual arrays.

 The hosts remain unaware of the mapping operation and access the virtual volumes as if they were accessing the physical storage attached to them.

 Typically, the virtualization layer is managed via a dedicated virtualization appliance to which the hosts and the storage arrays are connected.

 Fig 2.19 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 the storage network to the mapped LUNs.

 Depending on the capabilities of the virtualization appliance, the architecture may allow for more complex mapping between array LUNs and virtual volumes.

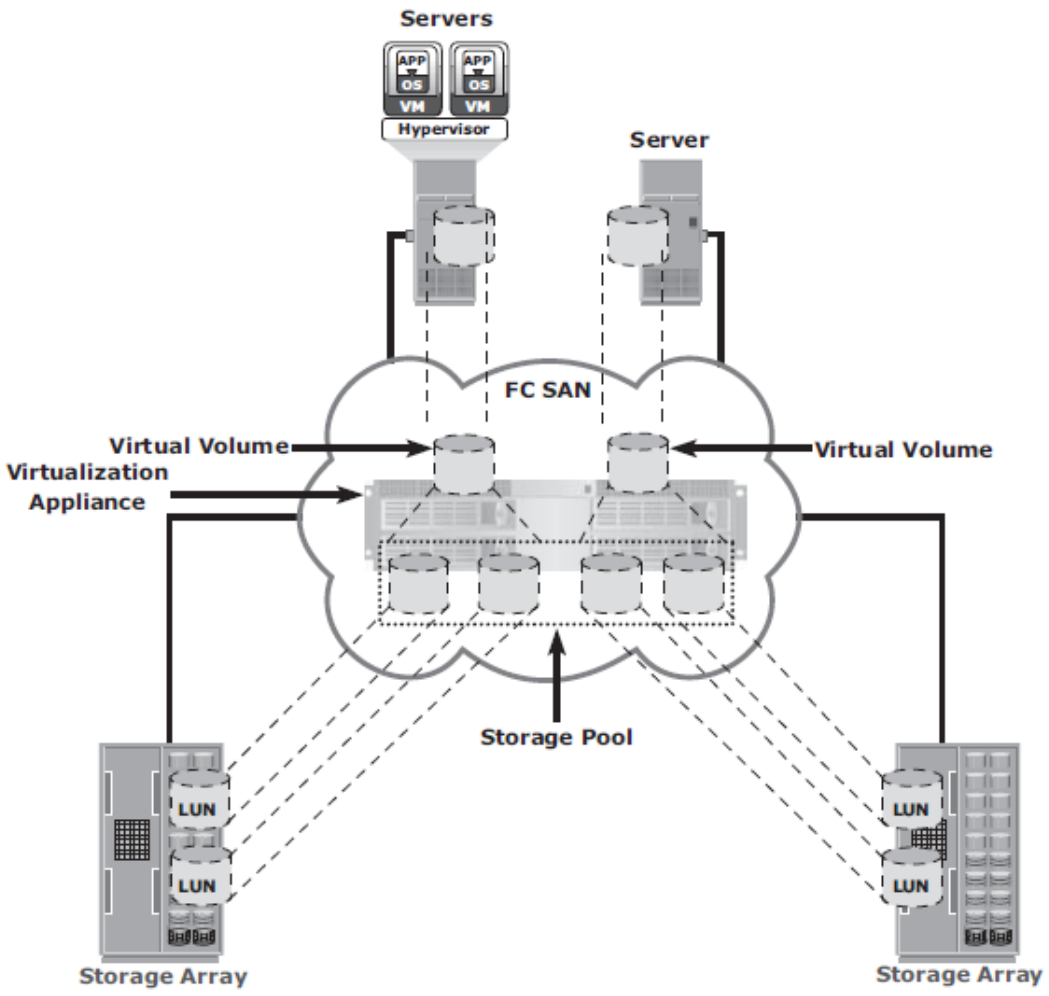


Fig 2.19 Block-level storage virtualization

 Block-level storage virtualization also provides the advantage of nondisruptive data migration.

 In a traditional SAN environment, LUN migration from one array to another is an offline event because the hosts needed to be updated to reflect the new array configuration.

 In other instances, host CPU cycles were required to migrate data from one array to the other, especially in a multivendor environment.

 With a block-level virtualization as a solution, the virtualization layer handles the back-end migration of data, which enables LUNs to remain online and accessible while data is migrating.

 No physical changes are required because the host still points to the same virtual targets on the virtualization layer.

 Previously, 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 (Fig 2.20). This enables the federation of block-storage resources both within and across data centers. The virtual volumes are created from the federated storage resources.

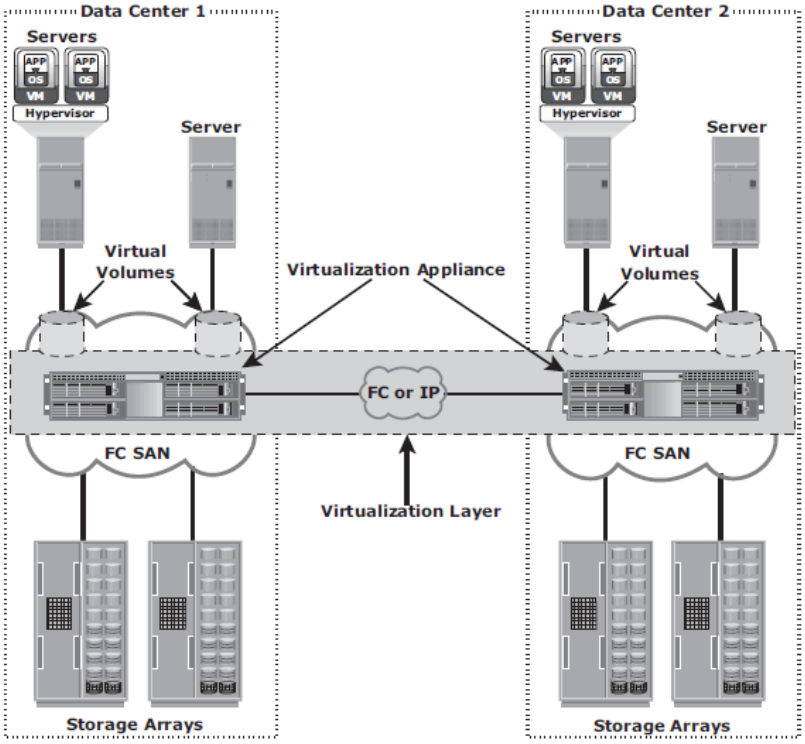


Fig 2.20 Federation of block storage across data centers

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.

 VSANs improve SAN security, scalability, availability, and manageability.

 VSANs facilitate an easy, flexible, and less expensive way to manage networks.

 Configuring VSANs is easier and quicker compared to building separate physical FC SANs for various node groups.

 To regroup nodes, an administrator simply changes the VSAN configurations without moving nodes and recabling.

6 What is FCoE? Explain the components of FCoE with neat diagram.

FCoE (Fiber Channel over Ethernet)

 Data centers typically have multiple networks to handle various types of I/O traffic — for example, an Ethernet network for TCP/IP communication and an FC network for FC communication.

 TCP/IP is typically used for client-server communication, data backup, infrastructure management communication, and so on.

 FC is typically 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. In addition, to enable the communication, different types of networking switches and physical cabling infrastructure are implemented in data centers.

 The need for two different kinds of physical network infrastructure increases the overall cost and complexity of data center operation.

 Fibre Channel over Ethernet (FCoE) protocol provides consolidation of LAN and SAN traffic over a single physical interface infrastructure.

 FCoE helps organizations address the challenges of having multiple discrete network infrastructures.

 FCoE uses the Converged Enhanced Ethernet (CEE) link (10 Gigabit Ethernet) to send FC frames over Ethernet.

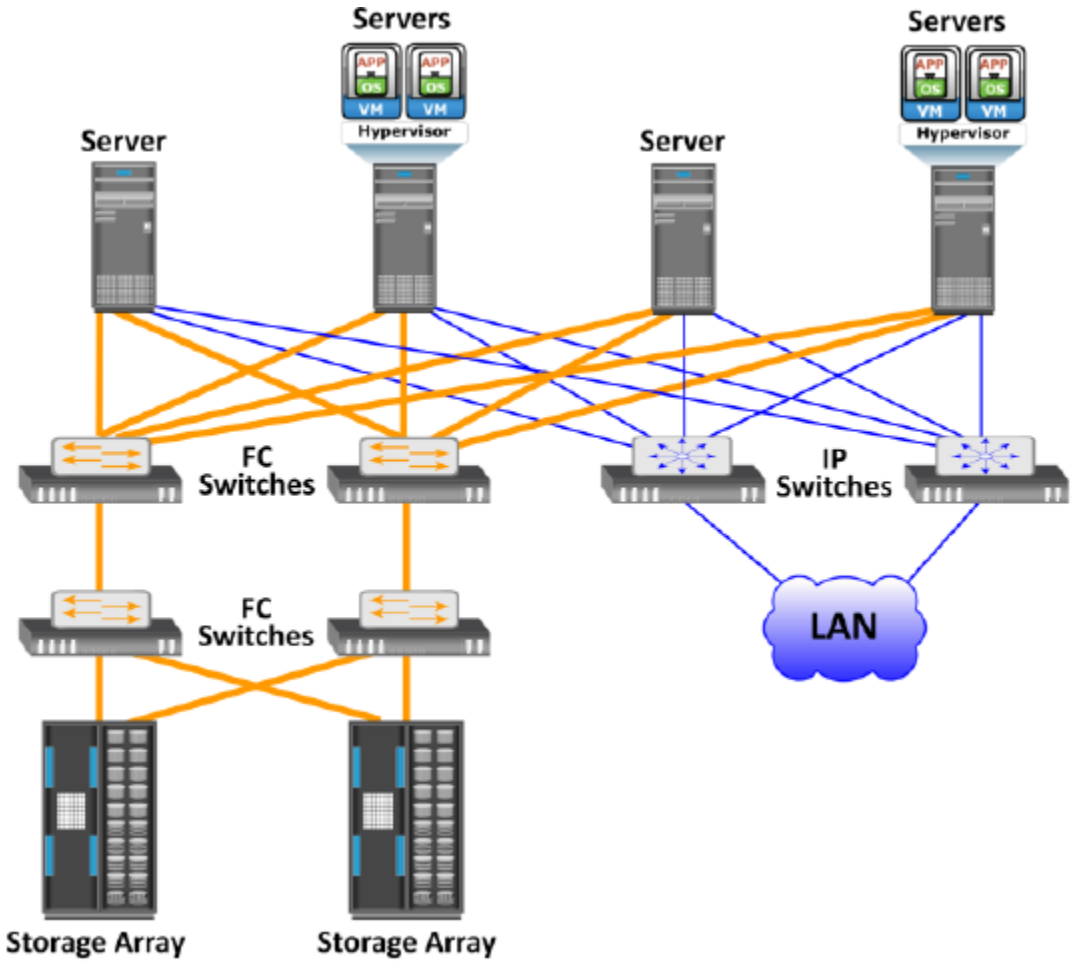


Fig 2.30 Before using FCOE

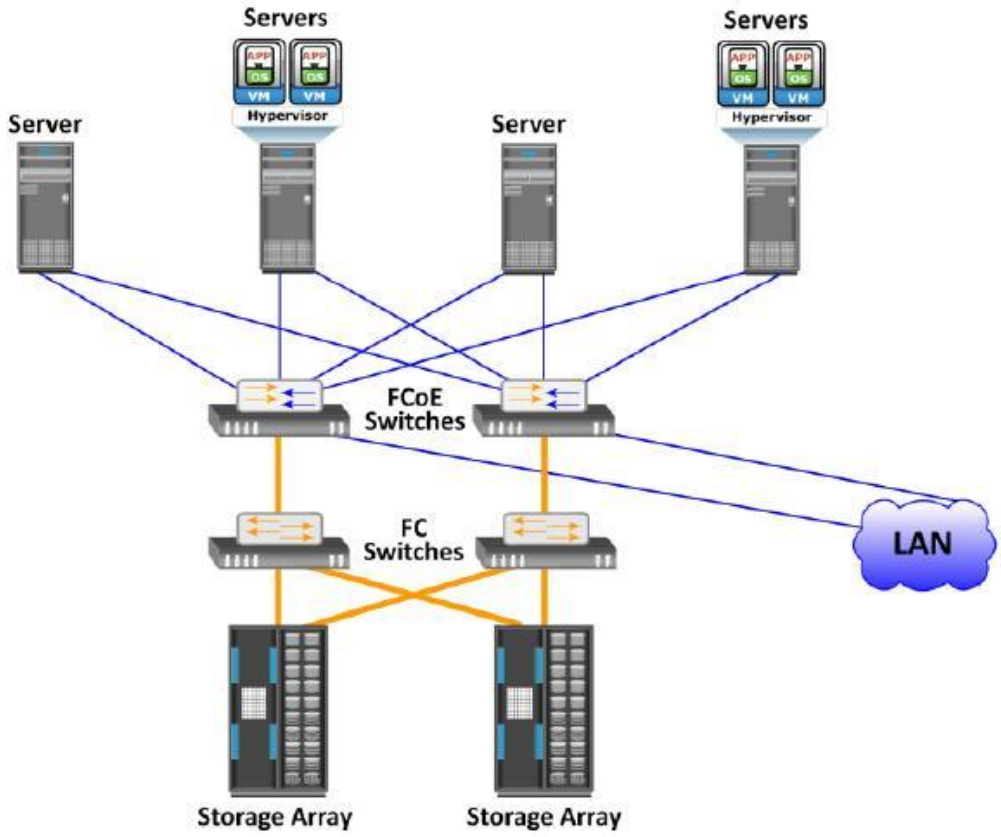


Fig 2.30 After using FCOE

The key components of FCOE are:

- Converged Network Adaptors (CNA)
- Cables
- FCOE Switches

Converged Network Adaptors (CNA)

- 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.
- As shown in Fig 2.31, a CNA contains separate modules for 10 Gigabit Ethernet, Fiber 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.

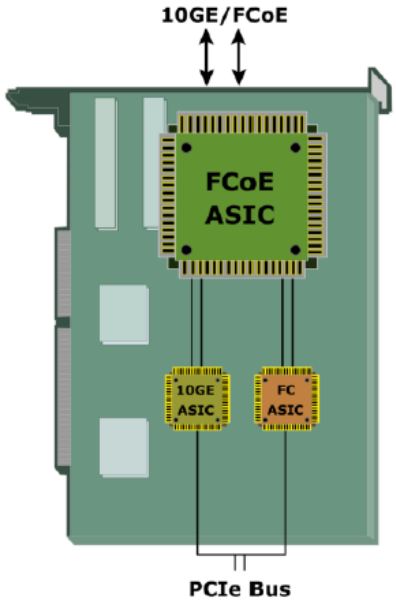


Fig 2.31 Converged Network Adapter

Cables

 There are two options available for 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 fi ber 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 Ethernet switch and Fibre Channel switch functionalities.

 As shown in Fig 2.32, FCoE switch consists of:

1. Fibre Channel Forwarder (FCF),

2. Ethernet Bridge,

3. set of Ethernet ports

4. optional FC ports

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

 Upon receiving the incoming traffic, the FCoE switch inspects the Ethertype (used to indicate which protocol is encapsulated in the payload of an Ethernet frame) 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. From there, the FC is extracted from the FCoE frame and transmitted to FC SAN over the FC ports.

 If the Ethertype is not FCoE, the switch handles the traffic as usual Ethernet traffic and forwards it over the Ethernet ports.

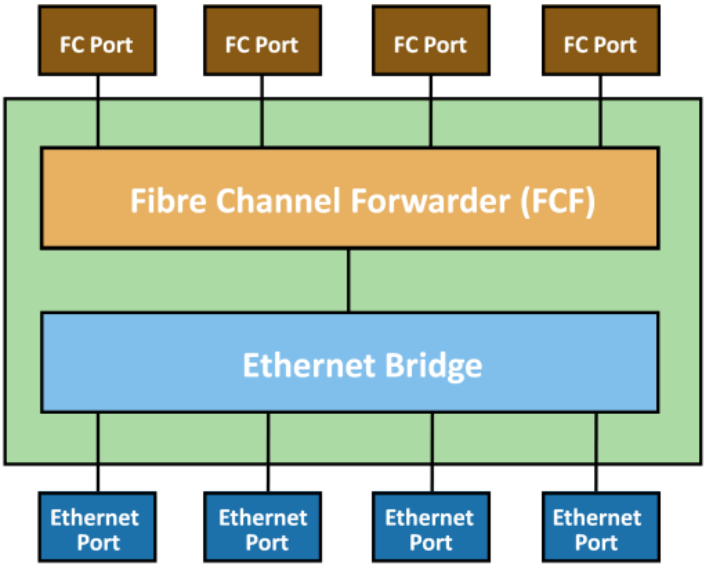


Fig 2.32 FCoE switch generic architecture

7 What is NAS? Explain the benefits of NAS

Page no. 159 from text-book

8 Explain with neat diagram the components of Fiber Channels (FC) storage Area Networks.

Components of SAN

 Components of FC SAN infrastructure are:

1) Node Ports,

2) Cabling,

3) Connectors,

4) Interconnecting Devices (Such As Fc Switches Or Hubs),

5) San Management Software.

Node Ports

 In fibre channel, devices such as hosts, storage and tape libraries are all referred to as Nodes.

 Each node is a source or destination of information for one or more nodes.

 Each node requires one or more ports to provide a physical interface for communicating with other nodes.

 A port operates in full-duplex data transmission mode with a transmit (Tx) link and a receive (Rx) link (see Fig 2.1).

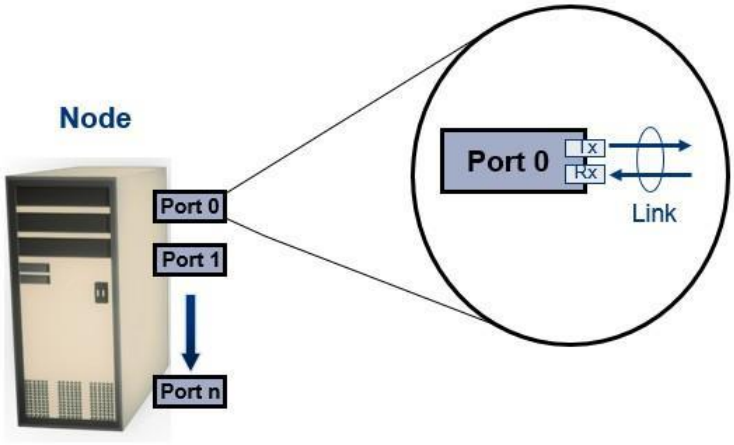


Fig 2.1: Nodes, Ports, links

Cabling

 SAN implementations use optical fiber cabling.

 Copper can be used for shorter distances for back-end connectivity

 Optical fiber cables carry data in the form of light.

 There are two types of optical cables :Multi-Mode And Single-Mode.

1) Multi-mode fiber (MMF) cable carries multiple beams of light projected at different angles simultaneously onto the core of the cable (see Fig 2.2 (a)).

 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.

 MMFs are generally used within data centers for shorter distance runs

2) Single-mode fiber (SMF) carries a single ray of light projected at the center of the core (see Fig 2.2 (b)).

 In an SMF transmission, a single light beam travels in a straight line through the core of the fiber.

 The small core and the single light wave limits modal dispersion. Among all types of fibre cables, single-mode provides minimum signal attenuation over maximum distance (up to 10 km).

 A single-mode cable is used for long-distance cable runs, limited only by the power of the laser at the transmitter and sensitivity of the receiver.

 SMFs are used for longer distances.

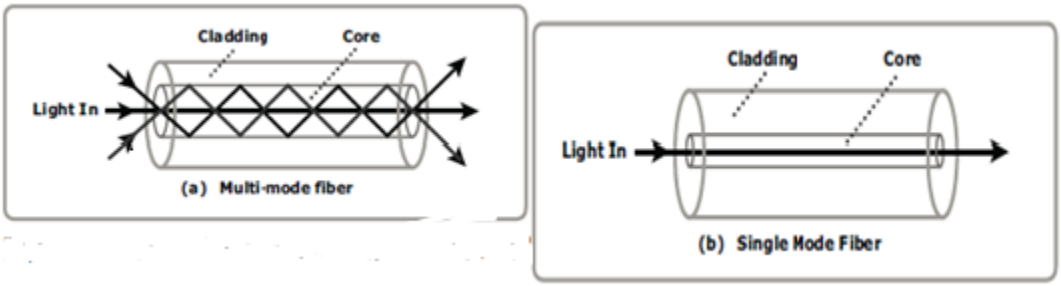


Fig 2.2: Multimode fiber and single-mode fiber

Connectors

 They are attached at the end of the cable to enable swift connection and disconnection of the cable to and from a port.

 A Standard connector (SC) (see Fig 2.3 (a)) and a Lucent connector (LC) (see Fig 2.3 (b)) are two commonly used connectors for fiber optic cables.

 An SC is used for data transmission speeds up to 1 Gb/s, whereas an LC is used for speeds up to 4 Gb/s.

 Figure 2.3 depicts a Lucent connector and a Standard connector.

 A Straight Tip (ST) is a fiber optic connector with a plug and a socket that is locked with a half-twisted bayonet lock (see Fig 2.3 (c)).

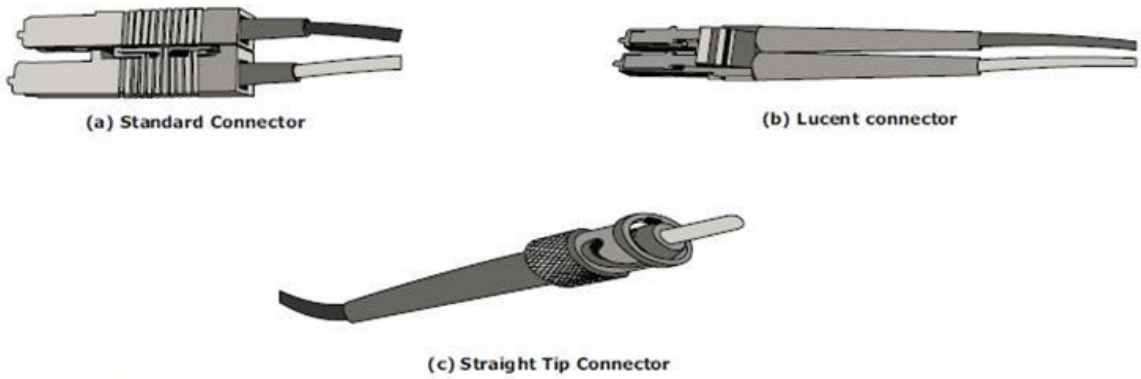


Fig 2.3: SC,LC, and ST connectors

Interconnect Devices

The commonly used interconnecting devices in SAN are

1) Hubs,

2) Switches,

3) Directors

 Hubs are used as communication devices in FC-AL implementations. Hubs physically connect nodes in a logical loop or a physical star topology.

 All the nodes must share the bandwidth because data travels through all the connection points. Because of availability of low cost and high performance switches, hubs are no longer used in SANs.

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

 Switches are available with:

 Fixed port count

 Modular design : port count is increased by installing additional port cards to open slots.

 Directors are larger than switches and are deployed for data center implementations.

 The function of directors is similar to that of FC switches, but directors have higher port count and fault tolerance capabilities.

 Port card or blade has multiple ports for connecting nodes and other FC switches

SAN Management Software

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

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

 It provides key management functions, including mapping of storage devices, switches, and servers, monitoring and generating alerts for discovered devices, and logical partitioning of the SAN, called zoning

9 What is zoning? Explain its types.

Zoning

 Zoning is an FC switch function that enables nodes within the fabric to be logically segmented into groups that can communicate with each other (see Fig 2.11).

 Whenever a change takes place in the name server database, the fabric controller sends a Registered State Change Notifi cation (RSCN) to all the nodes impacted by the change.

 If zoning is not configured, the fabric controller sends an RSCN to all the nodes in the fabric. Involving the nodes that are not impacted by the change results in increased fabric management traffic.

 Zoning helps to limit the number of RSCNs in a fabric. In the presence of zoning, a fabric sends the RSCN to only those nodes in a zone where the change has occurred.

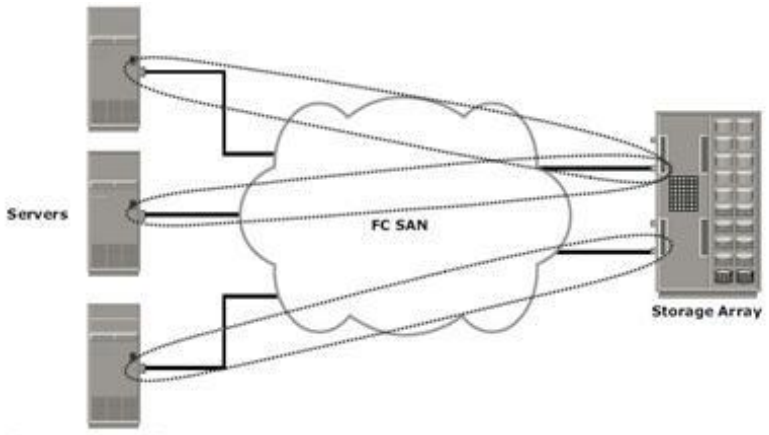


Fig 2.11 Zoning

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

 A zone set is a set of zones and a zone is a set of members.

 A member may be in multiple zones. Members, zones, and zone sets form the hierarchy defined in the zoning process (see Fig 2.12).

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

 Zones comprise a set of members that have access to one another. A port or a node can be a member of multiple zones.

 Zone sets comprise a group of zones that can be activated or deactivated as a single entity in a fabric. Only one zone set per fabric can be active at a time.

 Zone sets are also referred to as zone configurations.

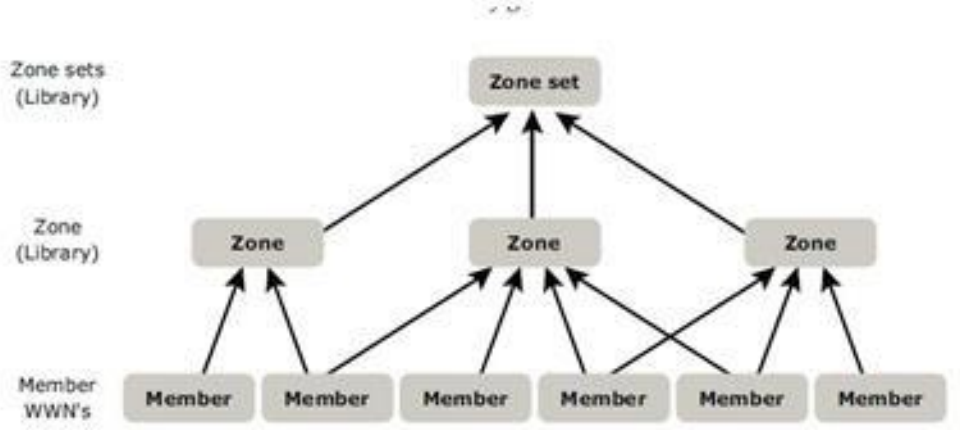


Fig 2.12: Members,Zones, and Zone sets

Types of Zoning

Zoning can be categorized into three types:

1) Port zoning

2) WWN zoning

3) Mixed zoning

Port zoning:

 It uses the FC addresses of the physical ports to define zones.

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

 The FC address is dynamically assigned when the port logs on to the fabric. Therefore, any change in the fabric configuration affects zoning.

 Port zoning is also called hard zoning.

 Although this method is secure, it requires updating of zoning configuration information in the event of fabric reconfiguration.

WWN zoning:

 It uses World Wide Names to define zones.

 WWN zoning is also referred to as soft zoning.

 A major advantage of WWN zoning is its flexibility.

 It allows the SAN to be recabled without reconfiguring the zone information. This is possible because the WWN is static to the node port.

Mixed zoning:

 It combines the qualities of both WWN zoning and port zoning.

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

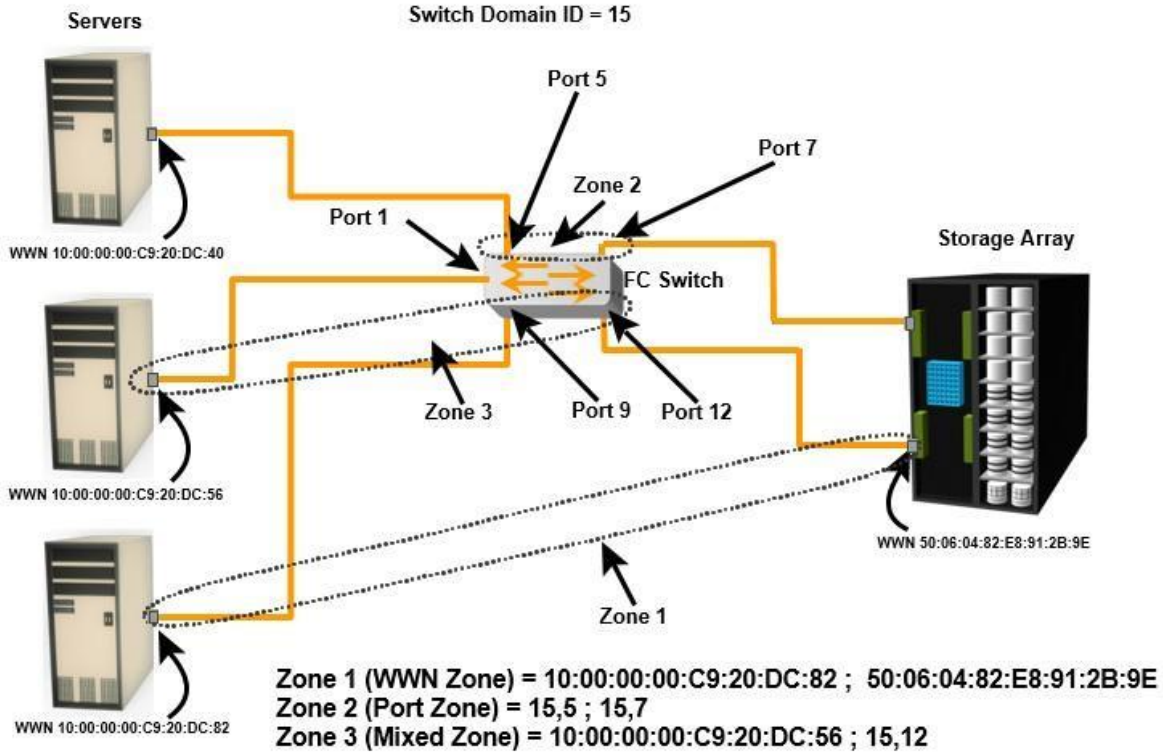


Fig 2.14: Types of Zoning

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