UNIT-3

Date ___/___/___
Page ______

INTRODUCTION TO NETWORK STORAGE -

① JBOD, DAS, NAS, SAN & CAS evolution and comparison -

Refer Unit 1 Evolution of various storage technologies. CAS is the lastest technology after IP SAN. CAS is discussed in further section in this unit. CAS → Content-Addressable Storage.

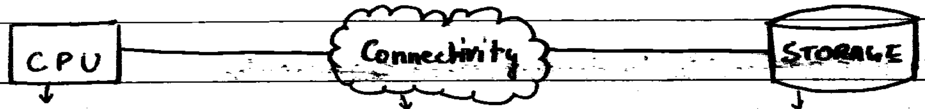
~~② Direct Storage~~

② Direct-Attached Storage (DAS) -

It is an architecture where storage connects directly to servers. Applications access data from DAS using block-level access protocols.

→ Applications of DAS - Internal HDD of a host, Tape libraries and directly connected external HDD packs.

→ Physical Elements of DAS -



CPU	Connectivity	Storage
Motherboard	Internal	Hard disk(s), CD-ROM drive,
Cluster group of processors	External	Optical drive, Removal media
Processor cards		Tape devices/Tape library, RAID/
Complete system		Intelligent array(s), Portable media devices

→ Types of DAS -

(1) Internal DAS - In this architecture, the storage device is internally connected to the host by a serial or parallel bus.

(2) External DAS - In this architecture, the server connects directly to the external storage devices

→ Connectivity -

Internal → serial or parallel bus, External → wires having SCSI or FC protocol

→ DAS Management -

Internal -

· Host provides disk partitioning (volume management), file system layout and data addressing.

DAS managed individually through the server and the OS.

IDE/ATA → Integrated Device Electronics / Advantaged technology attachment
SATA → Serial ATA
SAS → Serial Attached SCSI.

Date ___/___/___
Page ______

Extemal -
~~Array based management~~, Availability - multi-path I/O
Lower Total Cost of Ownership (TCO) for managing data and storage infrastructure

→ Security -
(1) Large scale DAS not very secure because of the distributed nature of the servers
(2) DAS hosted on Windows servers can be made secure by using group policies
(3) DAS does not provide security that is associated with a SAN or NAS configuration

→ Standards - (Others are IDE/ATA, SATA, SAS).
(i) SCSI → Small Computer System Interface.
(2) FC → Fibre Connectivity (usually point-to-point).

→ Limitations -
(1) DAS does not scale well (No scalability) ~~Limited no. of ports & hosts and~~ Limited bandwidth & distance
(2) Unused Resources cannot be easily reallocated.
(3) Data not accessible by diverse user ~~grou~~ groups.
(4) Allows only one user at a time.
(5) High administrative costs.
(6) DAS performance is effected by disk utilization, throughput, cache memory of storage device, virtual memory of host, RAID level configurations, ~~Storage~~ storage controller protocols and the efficiency of the bus.

→ Benefits -
(i) Low cost
(2) Configuration is simple and can be deployed easily and rapidly.
(3) Storage management tasks is easy for small and medium enterprises.
(4) Requires less hardware and software elements.
(5) Reliability
(6) Low Complexity
(7) Absence of storage interconnects and network latency.

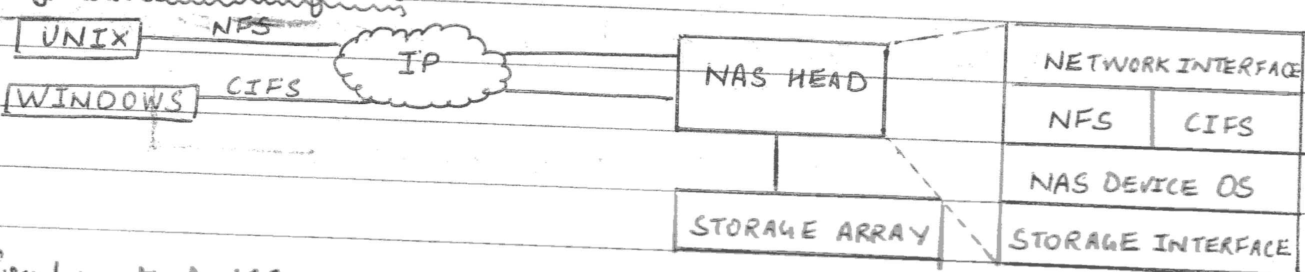
③ Network Attached Storage (NAS) -

It is an IP-based file-sharing device attached to a local area network. NAS provides the advantages of server consolidation by eliminating the need for multiple file servers. It provides storage consolidation through file-level data access and sharing.

→ Applications - NAS enables both UNIX and Microsoft Windows users to share the same data seamlessly.

It uses NFS for UNIX, CIFS for Windows and FTP and other protocols.

→ Physical Elements of NAS -



Components of NAS are -

1) NAS Head (CPU and Memory)
2) One or more network interface cards (NICs), which provide connectivity to the network. Eg - Gigabit Ethernet, Fast Ethernet
3) An optimized OS for managing NAS functionality
4) NFS and CIFS protocols for file sharing.
5) Industry standard storage protocols to connect and manage physical disk resources such as ATA, SCSI or FC.

→ NAS implementations -

Integrated NAS - It has all the components of NAS, such as the NAS head and storage, in a single enclosure, or frame. This makes the integrated NAS a self-contained environment.

Simple management functions. Scalability is limited

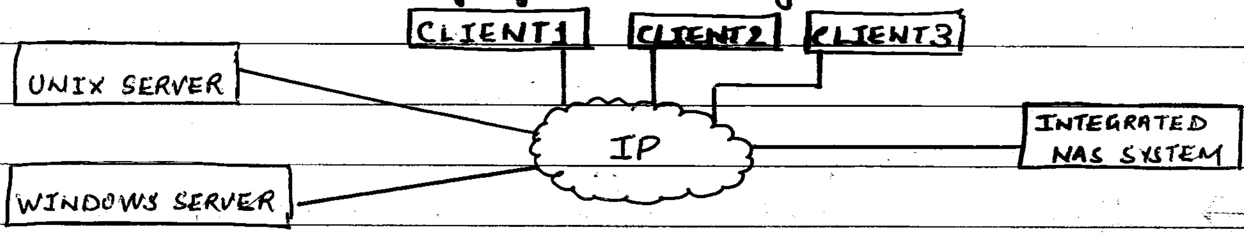
Gateway NAS - It consists of an independent NAS head and one or more storage arrays.

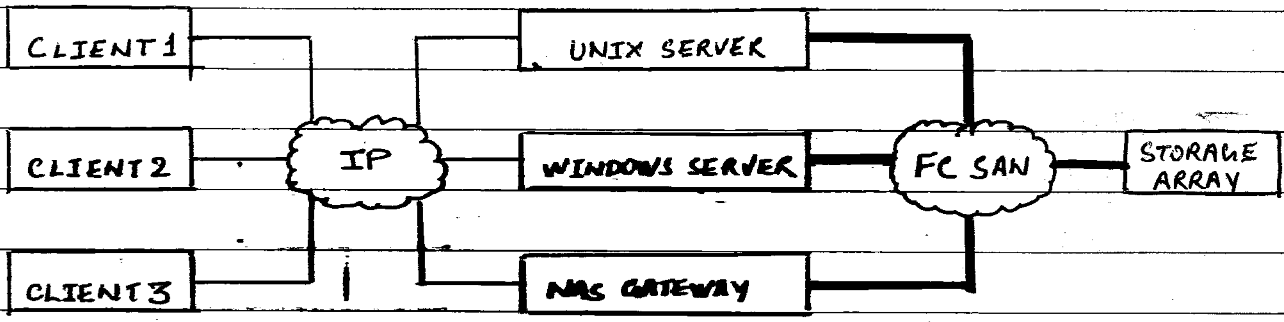
Complex management functions. Scalability as per required.

→ Connectivity -

Integrated NAS - Standard IP network

Gateway NAS - Front-end connectivity is done by IP Network but communication between the NAS and storage system is done by FC SAN.





→ NAS Management -

→ NAS File I/O -

NAS uses file-level access for all of its I/O operations. File I/O is a high-level request that specifies the file to be accessed, but does not specify its logical block address.

1) File Systems and Remote file sharing - A file system is a structured way of storing and organizing data files. Many file systems maintain a file access table to ~~sp~~ simplify the process of finding and accessing files.

2) Accessing a File System - A file system must be mounted before it can be used. When mounting a file system, the OS organizes files and directories in a tree-like structure and grants the user the privilege of accessing this structure.

Files are located at the leaf nodes, and directories and sub-directories are located at intermediate node.

3) File sharing - A user who creates the file (owner) determines the type of access to be given to other users (read, write, execute, append, delete, and list). and controls changes to the file.

File Transfer Protocol (FTP) is used in remote file sharing between the servers and ~~users~~ clients.

→ NAS I/O operations -

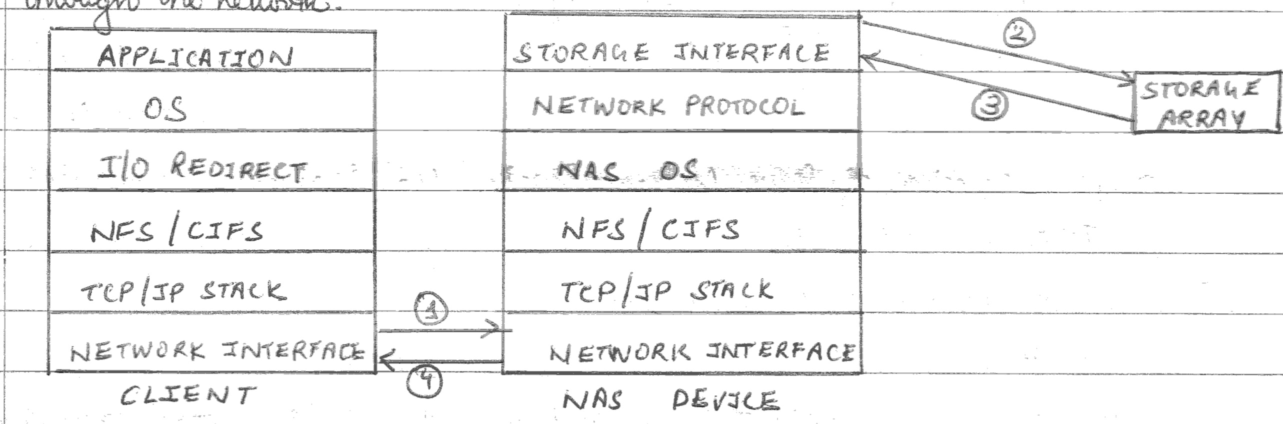
The NFS and CIFS protocols handle file I/O requests to a remote file system, which is managed by the NAS device. The process of NAS I/O is as follows -

(1) The requestor packages an I/O request into TCP/IP and forwards it through the network stack. The NAS device receives this request from the network.

(2) The NAS device converts the I/O request into an appropriate physical storage request which is a block level I/O, and then performs the operation against the physical storage pool.

(3) When the data is returned from the physical storage pool, the NAS device processes and repackages the data into an appropriate file protocol response.

(4) The NAS device packages this response into TCP/IP again and forwards it to the client through the network.



Hosting and accessing files on NAS - (Stepwise process)

(1) Create storage array volumes
(2) Create NAS volumes
(3) Create NAS file systems
(4) Mount file systems
(5) Access the file systems.

→ Standards -

→ NAS file sharing protocols -

(1) NFS (Network file system) -

It is a client/ server protocol. It uses a machine-independent model to represent user data.

It uses RPC mechanisms over TCP protocol.

Mount point grant access to remote hierarchical file structures for local file system structures.

Access to the mount can be controlled by permissions.

Three versions of NFS are-

NFS version 2 (NFSv2) - Uses ~~UDP~~ UDP to provide stateless network connection.

Features → locking are handled outside the protocol.

NFS version 3 (NFSv3) - Uses TCP or UDP to provide stateless network connection.

Features → asynchronous writes, 64 bit file size.

NFS version 4 (NFSv4) - Use TCP to provide stateful network connection

Features → Enhanced security.

(2) CIFS (Common Internet File System) -

It is a client/server application protocol that enables client programs to make requests for files and services on remote computers over TCP/IP. It is a public or open, variation of Server Message Block (SMB) protocol. It is a stateful protocol.

To ensure data integrity :-

(i) It uses file and record locking

(ii) It runs over TCP

(iii) It supports fault tolerance and can automatically restore connections and reopen files that were open prior to interruption.

→ Security -

(1) Enhanced security due to the use of NFSv4 protocol.

(2) Provide use ~~authen~~ authentication and authorization.

→ Limitations -

(1) Low performance as compared to SAN.

(2) Consumes more bandwidth

(3) IP network might be congested.

(4) Factors that affect NAS performance are number of hops, authentication with the directory service, retransmission, overutilized routers and switches, file/directory lookup and metadata requests, Overutilized NAS devices and overutilized clients.

→ Benefits of NAS -

(1) Support comprehensive access to information.
(2) Improved efficiency
(3) Improved flexibility
(4) Centralized storage
~~(5)~~
(5) Simplified management
(6) Scalability
(7) High availability
(8) Security.

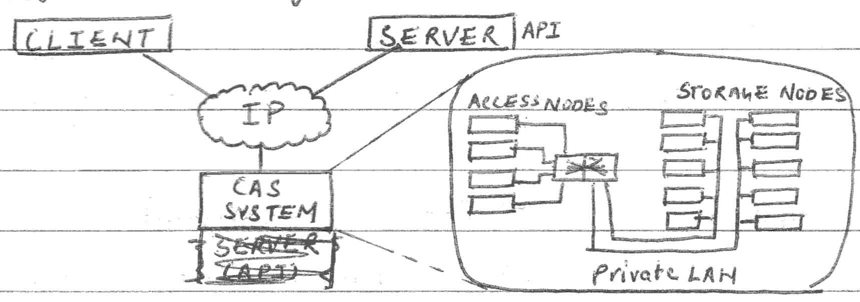
④ Content - Addressed Storage (CAS) -

It is an object-based system that has been purposely built for storing fixed content data. It is designed for secure online storage and retrieval of fixed content. The stored object is assigned a globally unique address known as content address (CA).

It provides an optimized and centrally managed storage solution that can support single-instance storage (SiS) to eliminate multiple copies of the same data.

→ Applications - Health care solutions (Storing Patient Studies), Finance solution (storing financial records)

→ Physical Elements of CAS -



CAS ARCHITECTURE

Physical elements of CAS are Client, Server and CAS based Storage. Logical elements of CAS are API, metadata and object-level access protocols.

A client accesses the CAS-based storage over a LAN through the server that runs the ~~server~~ CAS API which is responsible for performing functions that enable an application to store and retrieve the data.

CAS architecture is a redundant array of independent nodes (RAIN).

→ CAS management -

→ Object storage and retrieval in CAS -

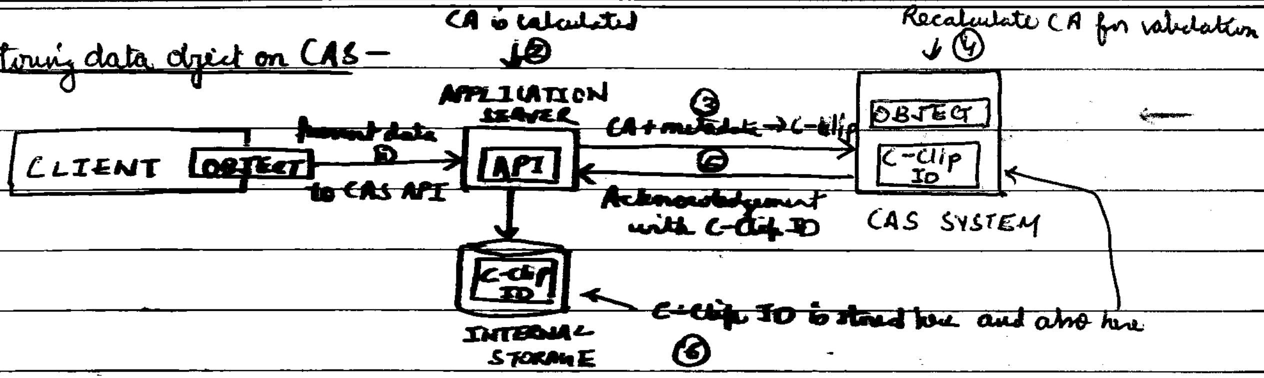
Some terminologies -

C-Clip - A virtual package that contains data and its associated CDF.

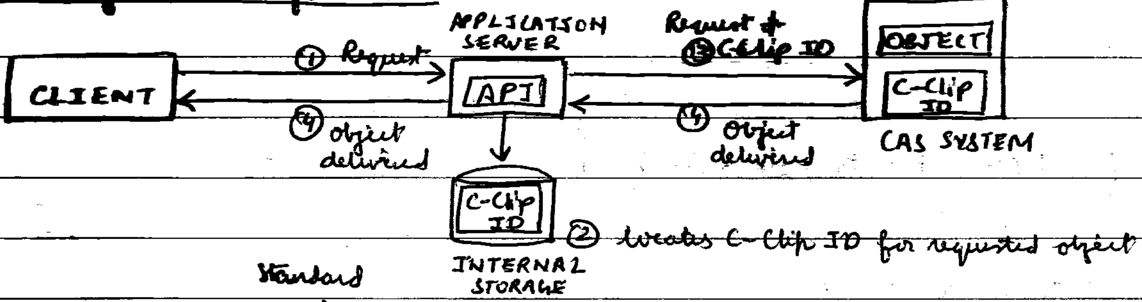
C-Clip Descriptor file (CDF) - XML file that a system creates while making a C-Clip

myCOMPANION

Storing data object on CAS -



Data object retrieval from CAS -



Standard

→ **Connectivity** - Use IP protocol (LAN for inside the CAS system).

→ **Security** - Data security is provided by constantly communication between the access nodes and the storage nodes.

→ **Limitations** -

(1) Can be slower than SAN, NAS or DAS

(2) Application integration

(3) Initial cost of ownership is higher even though TCO is significantly lower.

→ **Benefits** -

(1) Content authenticity
(2) Content Integrity
(3) Location Independence
(4) Single-instance Storage (SiS).
(5) Retention enforcement
(6) Record-level protection and disposition
(7) Technology Independence
(8) Fast record retrieval.

(5) Storage Area Network (SAN) -

It is a high speed, dedicated network of servers and shared storage devices. Traditionally connected over Fibre Channel (FC) networks, a SAN forms a single storage pool and facilitates data centralization and consolidation.

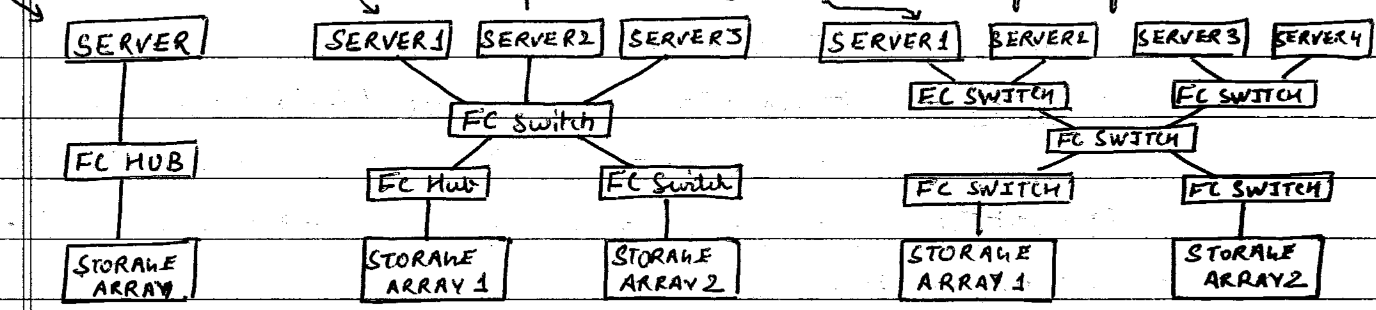
→ **Fibre Channel (FC)** -

It is a high speed network technology that runs on high-speed optical fibres (preferred for front-end SAN connectivity) and serial cables (preferred for back-end disk connectivity).

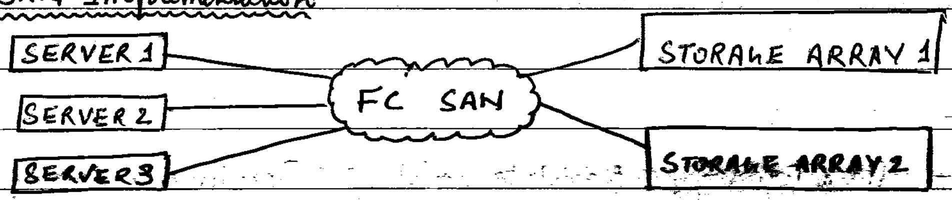
→ Evolution of SAN -

SAN Islands FC Arbitrated loop (FC-AL) → Interconnected SANs FC Switched fabric

Enterprise SANs FC switched ~~flash~~ fabric



→ SAN Implementation -

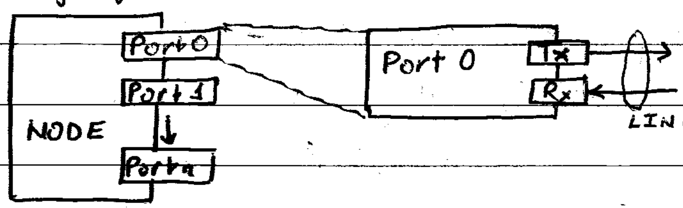


→ Components of SAN - (Physical Elements)

Three basic components - Servers, Network Infrastructure and storage. These components can be further broken down into five key elements -

(1) Node Ports -

Each node requires one or more ports to provide a physical interface for communicating with other nodes. These ports are integral components of an HBA (Host Bus Adaptors) and the storage front-end adapters.



(2) Cabling -

SAN implementation uses optical fibre cabling. Copper can be used for shorter distances for back-end connectivity, as it provides a better signal-to-noise ratio for distances up to 30 metres.

Optical fibre cables carry data in form of light. There are two types of optical

Date ___/___/___
Page ______

cables, ~~&~~ multimode fibre (MMF) → carry multiple beams of light and Single-mode fibre (SMF) → carry a single beam of light.

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

Connectors – Standard Connector (SC) → 1 Gb/s transmission speed,
Lucent connector (LC) → 4 Gb/s transmission speed,

Straight Tip (ST) connector has a plug and a socket that is locked with a half twisted bayonet lock.

Small Form-factor Pluggable (SFP) is an optical transceiver used in optical communication. (10 Gb/s data rates)

(3) Interconnecting Devices –

Hub ~~Pass~~ → Bus topology, Broadcast (No longer used), Bandwidth sharing

Switch → Star topology, point-to-point, No Bandwidth sharing

Directors → Larger than switches and are deployed for data center implementation. The function of directors is similar to that of FC switches, but directors have high port count and fault tolerance capability.

(4) Storage arrays –

It provides high availability, high ~~a~~ and redundancy, improved performance, business continuity and multiple host connectivity.

(5) SAN Management software –

It manages the interfaces between hosts, interconnect devices and storage arrays. It provides Key management functions – mapping of storage devices, switches, and servers, monitoring and generating alerts for discovered devices, and logical partitioning of the SAN, called zoning.

→ FC Connectivity –

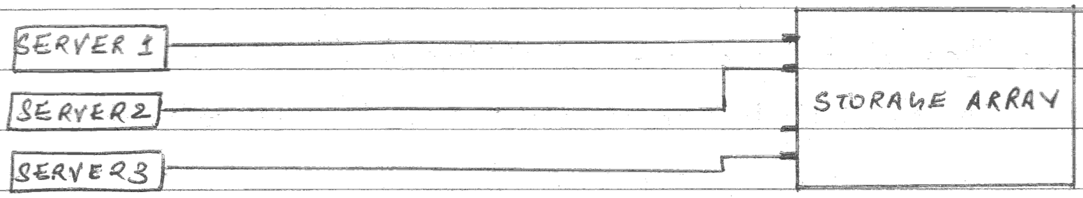
Three basic interconnectivity options –

(1) Point-to-Point –

Simplest FC configuration – two devices are connected directly to each other. Offers limited connectivity and no scalability.

Date ___/___/___

Page ______



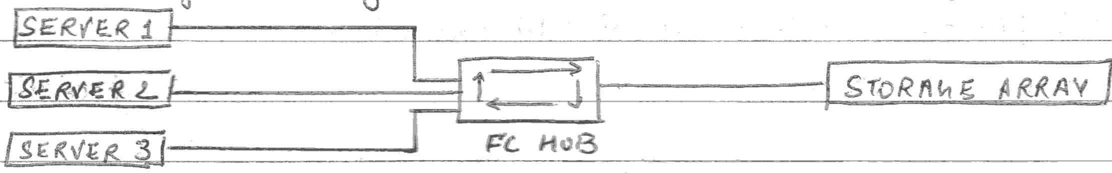
(2) <u>Fibre Channel Arbitrated Loop</u> (FC-AL) -

Device are attached to a shared loop. It has characteristics of a token ring topology and a physical star topology.

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

Devices on the loop must "arbitrate" to gain control of the loop.

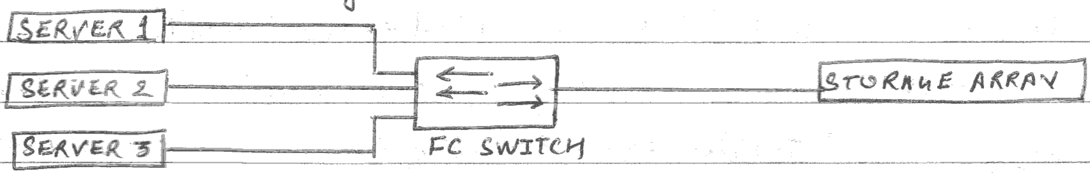
<u>Limited scalability</u> due to bandwidth sharing, uses 8-bit addressing (127 devices), and adding or removing a device results in loop re-initialization.



(3) <u>Fibre Channel Switched Fabric</u> (FC-SW or fabric connect) -

It provides interconnected devices, dedicated bandwidth and scalability.

A <u>fabric</u> is a logical space in which all nodes communicate with one another in a network which is created using a switch or a network of switches. Each switch in a fabric contains a <u>unique domain identifier</u>, which is part of the fabric's addressing scheme.



→ <u>Fibre Channel Ports</u> -

(1) ~~Node~~ <u>N-port</u> (Node Port) - An end point in the fabric

(2) <u>NL-port</u> (Node Loop Port) → Node port that supports the arbitrated loop topology.

(3) <u>E-port</u> (Expansion Port) → FC port that forms the connection between two FC switches.

(4) <u>F-port</u> (Fabric Port) → A port on a switch that connects an N-port.

(5) <u>FL-port</u> ~~&~~ → A fabric port that participates in FC-AL.

(6) <u>G-port</u> → A generic port that can operate as an E-port or an F-port.

Date ___/___/___

Page ________

→ Fibre Channel Architecture - (Standards)

The FC architecture represents true channel/network integration with standard interconnecting devices.

Channel technologies provide high levels of performance with low protocol overheads.

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

(1) Sustained transmission bandwidth over long distances.

(2) Support for a large number of addressable devices over a network (15 million).

(3) Exhibits the characteristics of channel transport and provides speed upto 8.5 Gb/s

→ Fibre Channel Protocol Stack -

	APPLICATION				
FC-4	SCSI	HIPPI	ESCON	ATM	IP
FC-2	FRAMING / FLOW CONTROL				
FC-1	ENCODE / DECODE				
FC-0	1Gb/s	2Gb/s	4Gb/s	8Gb/s	

FC-4 (Upper layer protocol) - Defines the application interfaces and the way Upper layer protocols (ULP) are mapped to the lower FC layers.

Some of the protocols are SCSI, High Performance Parallel Interface (HIPPI) framing protocol, Enterprise Storage Connectivity (ESCON), ATM and IP.

FC-2 (Transport layer) - Contains the payload, addresses of the source and destination ports, and link control information.

It also defines fabric services, classes of service, flow control and routing.

FC-1 (Transmission Protocol) - Includes serial encoding and decoding values, special characters used, and error control.

FC-0 (Physical Interface) - defines the physical interface, media, and transmission of raw bits.

→ Fibre Channel addressing -

An FC address is dynamically assigned when a port logs onto the fabric.

24 bit FC address of N_port -

DOMAIN ID	AREA ID	PORT ID	
8 bits (239) 17 addresses are reserved	8-bits (256)	8-bits (256)	= 15,663,104 address

Area-ID is used to identify a group of F-ports.

24 bit FC address of NL-port -

LOOP ID / UNUSED	AL_PA ID
LOOP ID in case of public loop (16-bits) Unused in case of private loop (16-bits)	8-bits (127) 1 address is reserved

Public loop → when an arbitrated loop is connected to a fabric through an FL-port.

World Wide Names (WWN) - Similar to MAC address

Each device in the FC environment is assigned a 64 bit unique identifier called the WWN. It is a static name for each device on an FC network.

Two types - World Wide Node Name (WWNN) and World Wide Port Name (WWPN)

→ FC Frame -

SOF 4 BYTES	FRAME HEADER 24 BYTES	DATA FIELD 0 - 2112 BYTES	CRC 4 BYTES	EOF 4 BYTES
Start of frame			Cyclic Redundancy Check	End of frame

~~delimiter~~ delimiters → SOF; EOF ← delimiters

ROUTING CONTROL (R_CTL)	DESTINATION ID (D_ID)	
CLASS SPECIFIC CONTROL (CS_CTL)	SOURCE ID (S_ID)	
TYPE	FRAME CONTROL (F_CTL)	
SEQUENCE ID (SEQ_ID)	DATA FIELD CONTROL (DF_CTL)	(SEQ_CNT) SEQUENCE COUNT
ORIGINATING EXCHANGE ID (OX_ID)		RESPONDER EXCHANGE ID (RX_ID)
← 1 BYTE →	← 1 BYTE →	← 2 BYTES →

OFFSET

Type - describes the Upper layer protocol is to be carried on the frame

Routing Control - denotes whether the frame is a link control frame or a data frame.

Class specific control - Specifies link speeds for class 1 & class 4 data transmission.

Data Field control - Indicates any optional headers present at the beginning or not.

Frame control - Contain control information related to frame content

Date ___/___/___

Page ______

→ **Structure and Organization of FC data** -

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

Exchange operation - Exchange of sequences.

Sequence - Contiguous set of frames

Frame - fundamental unit of data transfer.

→ **Flow Control** -

It defines the pace of the flow of data frames during data transmission. FC technology uses two flow-control mechanism -

1) Buffer-to-Buffer Credit (BB-Credit) -

It is used for hardware-based flow control. It controls the maximum number of frames. The transmitting port maintains a count of free receiver buffers and continues to send frames if the count is greater than zero.

2) End-to-End Credit (EE-Credit) - Similar to BB-Credit

When an initiator and a target establish themselves as node communicating with each other, they exchange the EE_Credit parameters (part of Port login).

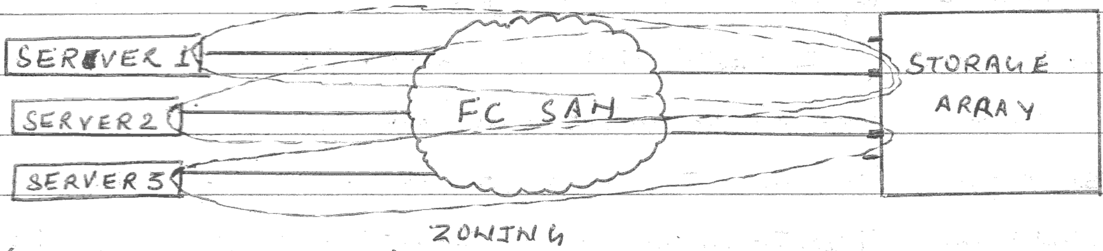
It affects the flow control for class 1 and class 2 traffic only.

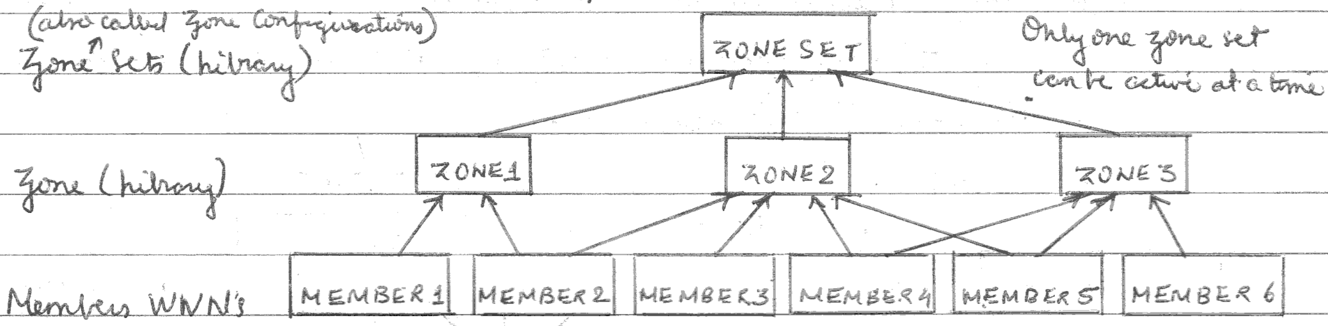
→ **Class of Service** -

	CLASS 1	CLASS 2	CLASS 3
Communication Type	Dedicated Connection	Nondedicated connection	Non dedicated connection
Flow Control	EE-credit	EE_credit & BB_credit	BB-credit
Frame Delivery	In order Delivery	Order not guaranteed	Order not guaranteed
Frame Acknowledgement	Acknowledged	Acknowledged	Not Acknowledged
Multiplexing	No	Yes	Yes
Bandwidth Utilization	Poor	Moderate	High

Zoning -

It is a FC switch function that enables nodes within the fabric to be logically segmented into groups that can communicate with each other.





Types of Zoning - Three types -

(1) Port Zoning - It uses the FC addresses of the physical ports to define zones. It is also called Hard zoning. Although this method is secure, it requires updating of zoning configuration information in the event of fabric reconfiguration.

(2) WWN Zoning - It uses World Wide Names to define zones. It is also called soft zoning. It is flexible because does not require ~~zoning~~ updating of zoning configuration information in the event of fabric reconfiguration.

(3) Mixed Zoning - It combines the quality of both WWN zoning and port zoning. Using mixed zoning enables a specific port to be tied to the WWN of a node.

Fabric Channel Login Types -

Fabric services define three login types -

(1) Fabric Login (FLOGI) - between N_port and F_port.

FLOGI frame contains WWNN & WWPN parameters send by N_port & F_port send Accept (ACC) frame to N_port. N_port sends the frame at FC address FFFFFE.

(2) Port Login (PLOGI) - between N_port and another N_port to establish a session.

(3) Process Login (PRLI) - between N_port and another N_port to exchange the information about the FC type in use, the SCSI initiator ~~and~~ or the target.

Date ___/___/___

Page ______

FC topologies -

(1) Core-Edge Fabric -

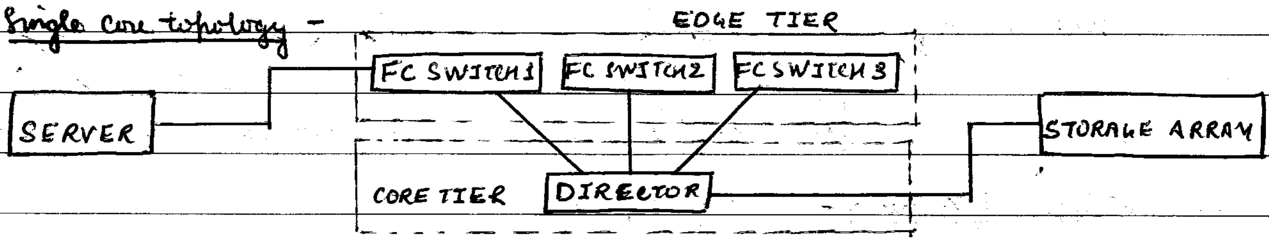
In this topology, there are two types of switch tiers in the fabric.

Edge tier - Comprises switches and offers an inexpensive approach to adding more hosts in a fabric. The ~~tier~~ tier at the edge fans out from the tier at the core. The nodes on the edge can communicate with each other.

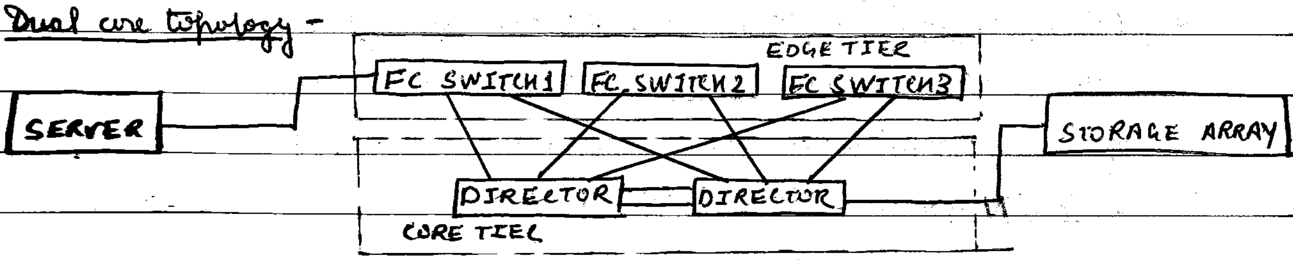
~~(a)~~ Core tier - Comprises enterprise directors that ensure high fabric availability. Additionally all tier traffic has to either traverse through or terminate at this tier.

This topology increases connectivity within the SAN while conserving overall port utilization. Two variations are -

Single core topology -



Dual core topology -



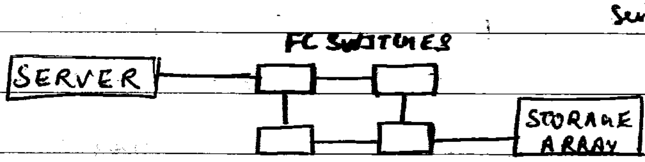
Benefits - ~~(1)~~ Provides one hop storage access to all storage in the system.
(2) provides easier calculation of ISL loading and traffic patterns
(3) Scalable to larger environments

Limitations - (1) Many hop leads to transmission delay
(2) Performance is affected by increase in ISLs and domain count in the fabric

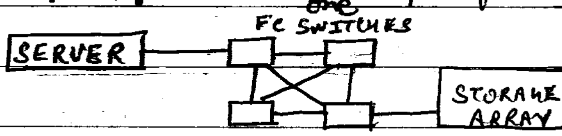
(2) Mesh Topology -

Each switch is directly connected to other switches by using ISLs. It enhanced connectivity. Two types - Partial Mesh and Full Mesh

Several ~~each~~ hop is required — ~~several~~ One hop is required



PARTIAL MESH



FULL MESH

Date ___/___/___
Page ______

→ Benefits of SAN -
(1) High Bandwidth
(2) SCSI Extension
(3) Resource Consolidation
(4) Scalability
(5) Secure Access.

→ Security - Isolation and filtering features are present in SAN.

→ Management -
(1) Infrastructure protection.
(2) Fabric Management
(3) Storage Allocation
(4) Capacity Tracking
(5) Performance Management

→ Applications -
Used in disk arrays, tape libraries and optical Jukeboxes to enhance storage devices.

→ Limitations -
(1) Very expensive
(2) Management of SAN is difficult
(3) Complexity of SAN