CMR
INSTITUTE OF
TECHNOLOGY

USN

Internal Assessment Test 3 – Dec 2020

Sub:	**Storage Area Networks**							**Sub Code:**	**18MCA554**
Date:	**14/12/2020**	**Duration:**	**90 min's**	**Max Marks:**	**50**	**Sem & Sec:**	**5 A & B**	**Branch:**	**MCA**

Note : Answer any full FIVE Questions

1	a) Explain the architecture of block level access and file level access with a neat diagram? b) What is Virtualization ? Write SNIA (Storage Networking Industry Association) storage virtualization taxonomy.	[4] [6]	CO3	L2

1a) Virtual storage is about providing logical storage to hosts and applications independent of physical resources. Virtualization can be implemented in both SAN and NAS storage environments.
In a SAN, virtualization is applied at the block level, whereas in NAS, it is applied at the file level.

i. Block-Level Storage Virtualization

- Block-level storage virtualization provides a translation layer in the SAN, between the hosts and the storage arrays, as shown in Figure 10-6.
- Instead of being directed to the LUNs on the individual storage arrays, the hosts are directed to the virtualized LUNs on the virtualization device.
- The virtualization device translates between the virtual LUNs and the physical LUNs on the individual arrays. This facilitates the use of arrays from different vendors simultaneously, without any interoperability issues.
- For a host, all the arrays appear like a single target device and LUNs can be distributed or even split across multiple arrays.
- Block-level storage virtualization extends storage volumes online, resolves application growth requirements, consolidates heterogeneous storage arrays, and enables transparent volume access. It also provides the advantage of non-disruptive data migration.
- In traditional SAN environments, LUN migration from one array to another was an offline event because the hosts needed to be updated to reflect the new array configuration.
- With a block-level virtualization solution in place, the virtualization engine handles the back-end migration of data, which enables LUNs to remain online and accessible while data is being migrated. No physical changes are required because the host still points to the same virtual targets on the virtualization device.

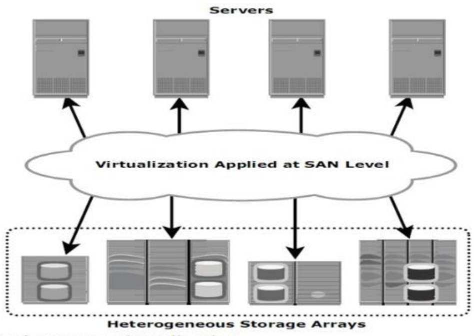

Servers
Virtualization Applied at SAN Level
Heterogeneous Storage Arrays

Figure 10-6: Block-level storage virtualization

ii. File-Level Virtualization

- File-level virtualization addresses the NAS challenges by eliminating the dependencies between the data accessed at the file level and the location where the files are physically stored.
- This provides opportunities to optimize storage utilization and server consolidation and to perform nondisruptive file migrations. Figure 10-7 illustrates a NAS environment before and after the implementation of file-level virtualization.
- Before virtualization, each NAS device or file server is physically and logically independent. Each host knows exactly where its file-level resources are located. Underutilized storage resources and capacity problems result because files are bound to a specific file server.
- It is necessary to move the files from one server to another because of performance reasons or when the file server fills up. Moving files across the environment is not easy and requires downtime for the file servers.
- Moreover, hosts and applications need to be reconfigured with the new path, making it difficult for storage administrators to improve storage efficiency while maintaining the required service level.
- **File-level virtualization simplifies file mobility.** It provides user or application independence from the location where the files are stored.
- File-level virtualization creates a logical pool of storage, enabling users to use a logical path, rather than a physical path, to access files.
- File-level virtualization facilitates the movement of file systems across the online file servers. This means that while the files are being moved, clients can access their files non-disruptively.

- Clients can also read their files from the old location and write them back to the new location without realizing that the physical location has changed.
- Multiple clients connected to multiple servers can perform online movement of their files to optimize utilization of their resources. A global namespace can be used to map the logical path of a file to the physical path names.

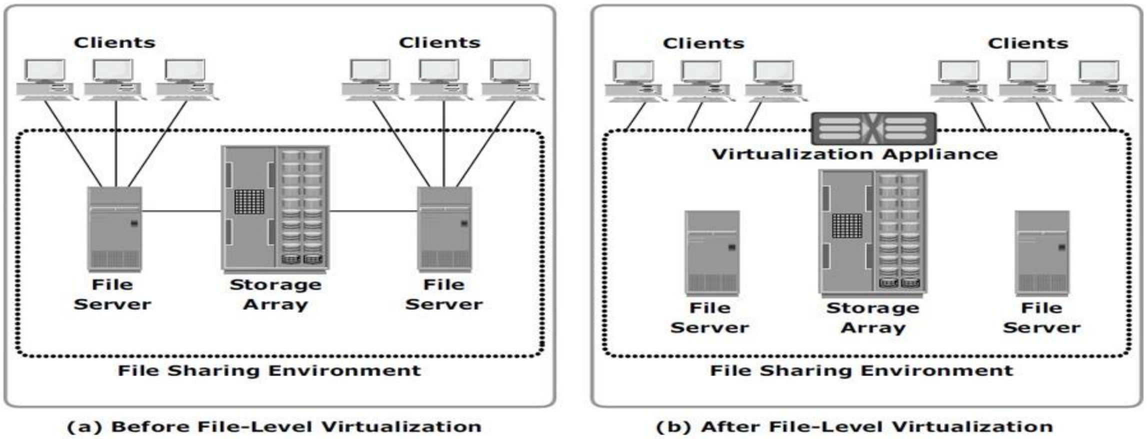


Figure 10-7: NAS device before and after file-level virtualization

1 b) Virtualization is the technique of masking or abstracting physical resources, which simplifies the infrastructure and accommodates the increasing pace of business and technological changes. It increases the utilization and capability of IT resources, such as servers, networks, or storage devices, beyond their physical limits. Virtualization simplifies resource management by pooling and sharing resources for maximum utilization and makes them appear as logical resources with enhanced capabilities.

The **SNIA (Storage Networking Industry Association) storage virtualization taxonomy** provides a systematic classification of storage virtualization, with three levels defining what, where, and how storage can be virtualized.

- The **first level** of the storage virtualization taxonomy addresses ―what is created. It specifies the types of virtualization: block virtualization, file virtualization, disk virtualization, tape virtualization, or any other device virtualization.
- The **second level** describes ―where the virtualization can take place. This requires a multilevel approach that characterizes virtualization at all three levels of the storage environment: server, storage network, and storage, as shown in Figure.
- The **third level** of the storage virtualization taxonomy specifies the network level virtualization methodology, in-band or out-of-band.

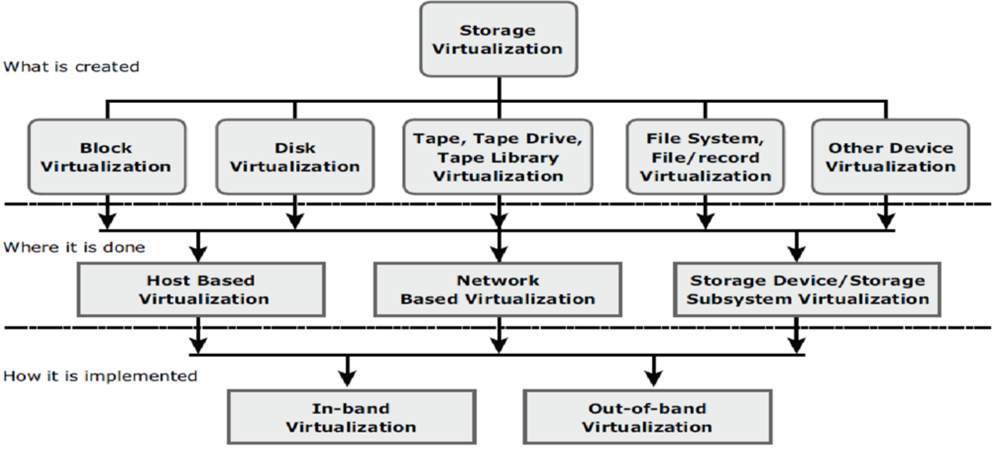
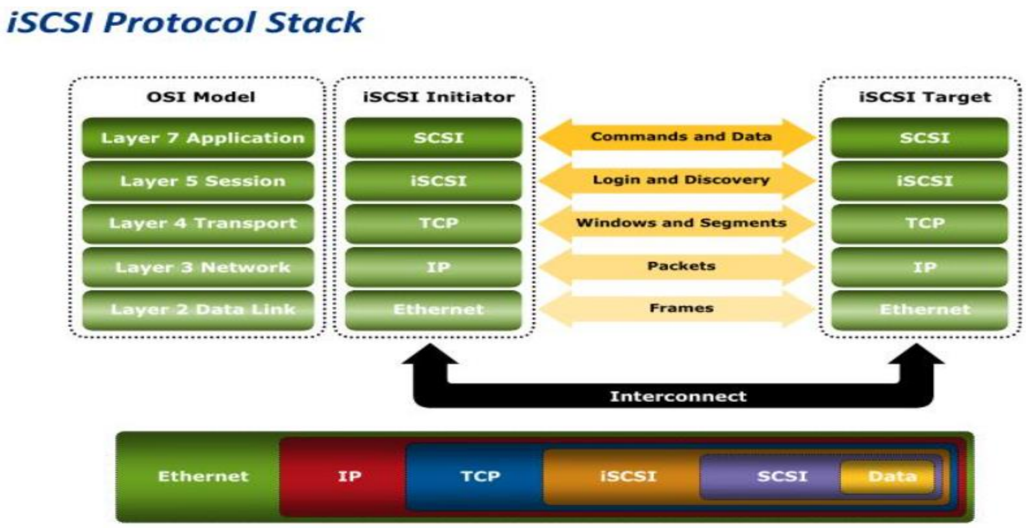

Storage Virtualization
What is created
Block Virtualization
Disk Virtualization
Tape, Tape Drive, Tape Library Virtualization
File System, File/record Virtualization
Other Device Virtualization
Where it is done
Host Based Virtualization
Network Based Virtualization
Storage Device/Storage Subsystem Virtualization
How it is implemented
In-band Virtualization
Out-of-band Virtualization

Fig : SNIA storage virtualization taxonomy

2	a) Explain ISCSI protocol stack? b) Explain different types of ISCSI discovery with a neat diagram?	[10]	CO3	L3

2a)

- encapsulates of SCSI commands for their delivery through a physical carrier.
- SCSI is the command protocol that works at the application layer of the OSI model.
- 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 initiators and targets.
- iSCSI is the session- layer protocol that initiates a reliable session between a device that recognizes SCSI commands and TCP/IP.
- The iSCSI session-layer interface is responsible for handling login, authentication, target discovery, and session management.



iSCSI Protocol Stack

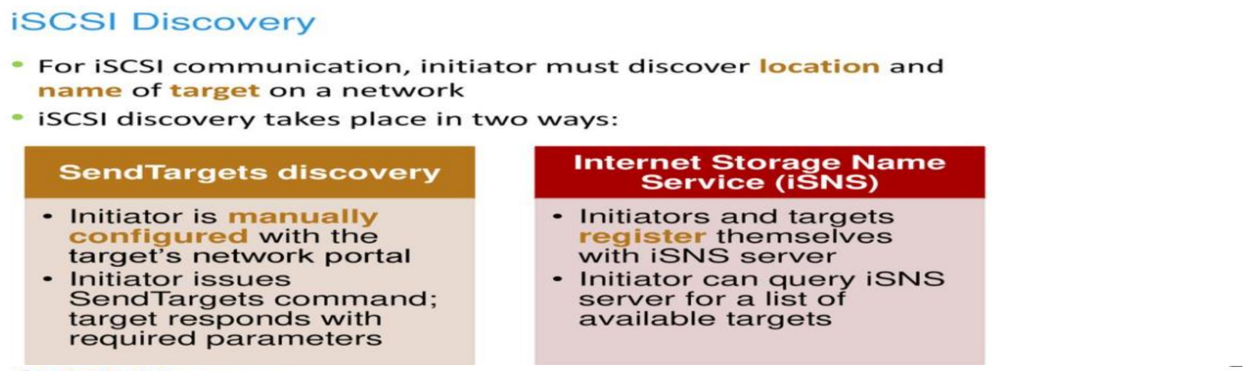
OSI Model	iSCSI Initiator		iSCSI Target
Layer 7 Application	SCSI	Commands and Data	SCSI
Layer 5 Session	iSCSI	Login and Discovery	iSCSI
Layer 4 Transport	TCP	Windows and Segments	TCP
Layer 3 Network	IP	Packets	IP
Layer 2 Data Link	Ethernet	Frames	Ethernet

Interconnect

Ethernet	IP	TCP	iSCSI	SCSI	Data

2b) An iSCSI session is established between an initiator and a target.

- ▶ Discovery of available targets to the initiator and the location of a specific target on a network
- ▶ Normal operation of iSCSI (transferring data between initiators and targets)



iSCSI Discovery

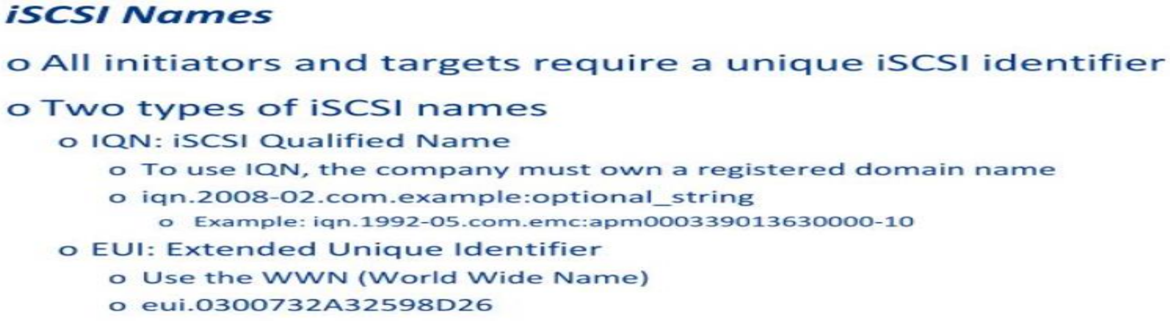
- For iSCSI communication, initiator must discover location and name of target on a network
- iSCSI discovery takes place in two ways:

SendTargets discovery

- Initiator is manually configured with the target's network portal
- Initiator issues SendTargets command; target responds with required parameters

Internet Storage Name Service (iSNS)

- Initiators and targets register themselves with iSNS server
- Initiator can query iSNS server for a list of available targets



iSCSI Names

- o All initiators and targets require a unique iSCSI identifier
- o Two types of iSCSI names
  - o IQN: iSCSI Qualified Name
    - o To use IQN, the company must own a registered domain name
    - o iqn.2008-02.com.example:optional_string
      - o Example: iqn.1992-05.com.emc:apm000339013630000-10
  - o EUI: Extended Unique Identifier
    - o Use the WWN (World Wide Name)
    - o eui.0300732A32598D26

3	a) Explain FCIP protocol stack? Explain FCIP encapsulation? b) Explain FC SAN topologies with the neat diagram?	[10]	CO3	L3

3a) Two primary protocols that leverage IP as the transport mechanism are iSCSI and Fibre Channel over IP (FCIP). FCIP is an IP-based storage networking technology that combines the advantages of Fibre Channel and IP. It creates virtual FC links that connect devices in a different fabric. FCIP uses a pair of bridges (FCIP gateways) communicating over TCP/IP as the transport protocol. FCIP is used to extend FC networks over distances and/or an existing IP-based infrastructure shown in fig. FCIP is extensively used in disaster-recovery implementations, where data is duplicated on disk or tape to an alternate site.

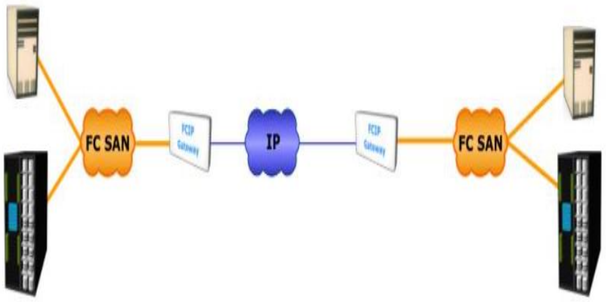


Fig : FCIP Network

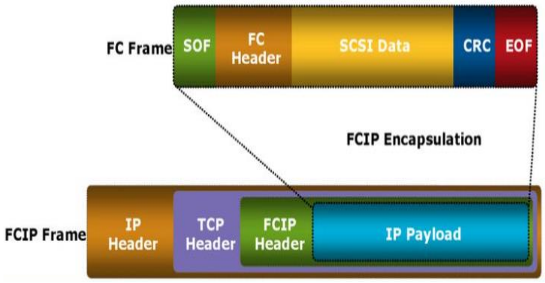


Fig : FCIP frame encapsulation

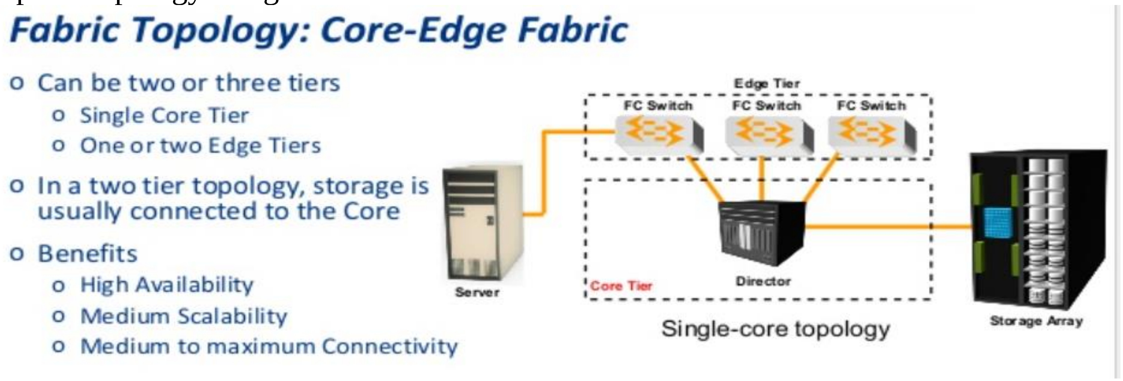
In FCIP, the FC frames are encapsulated onto the IP payload, FCIP does not manipulate FC frames.

FCIP frame

- Encapsulation of C frames onto IP packet that could cause the IP packet to be **fragmented.**
- When an IP packet is fragmented, the required parts of the header must be **copied** by all fragments.
- When a TCP packet is **segmented,** normal TC operations are responsible for receiving and re-sequencing the data prior t passing it on to the FC processing portion of the device.

3b)

Fabric design follows standard topologies to connect devices. Core-edge fabric is one of the popular topology designs



Fabric Topology: Core-Edge Fabric

- Can be two or three tiers
  - Single Core Tier
  - One or two Edge Tiers
- In a two tier topology, storage is usually connected to the Core
- Benefits
  - High Availability
  - Medium Scalability
  - Medium to maximum Connectivity

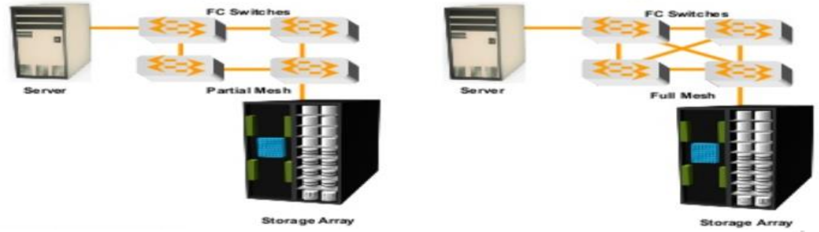
1.The core-edge fabric provides one-hop storage access to all storage in the system. Because each tier's switch is used for either storage or hosts, one can easily identify which resources are approaching their capacity, making it easier to develop a set of rules for scaling and apportioning.

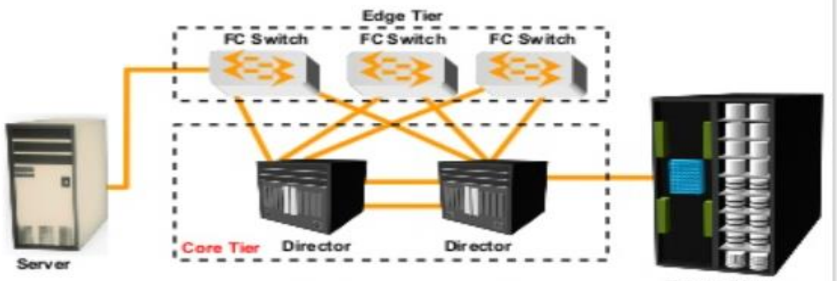
2. A well-defined, easily reproducible building-block approach makes rolling out new fabrics easier. Core-edge fabrics can be scaled to larger environments by linking core switches, adding more core switches, or adding more edge switches.

3. This method can be used to extend the existing simple core-edge model or to expand the fabric into a compound or complex core-edge model.

Fabric Topology: Mesh

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





Dual-core topology

4	a) Explain the key Benefits of Content Addressed Storage b) Define the following: i) C-Clip iii) BLOB iii) C-Clip Descriptor File (CDF)?	[6] [4]	CO4	L2

4a) 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.

Benefits of CAS

1. Content authenticity: The genuineness of stored content is achieved by generating a unique content address and automating the process for stored objects. Content authenticity is assured because the address assigned to each piece of fixed content is as unique as a fingerprint. Every time an object is read, CAS uses a hashing algorithm to recalculate the object's content address as a validation step and compares the result to its original content address.

2. Content integrity: Refers to the assurance that the stored content has not been altered. Use of hashing algorithm for content authenticity also ensures content integrity in CAS. If the fixed content is altered, CAS assigns a new address to the altered content, rather than overwrite the original fixed content, providing an audit trail and maintaining the fixed content in its original state.

3. Location independence: CAS uses a unique identifier that applications can leverage to retrieve data rather than a centralized directory, path names, or URLs. Using a content address to access fixed content makes the physical location of the data irrelevant to the application requesting the data. Therefore, the location from which the data is accessed is transparent to the application. This yields complete content mobility to applications across locations.

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

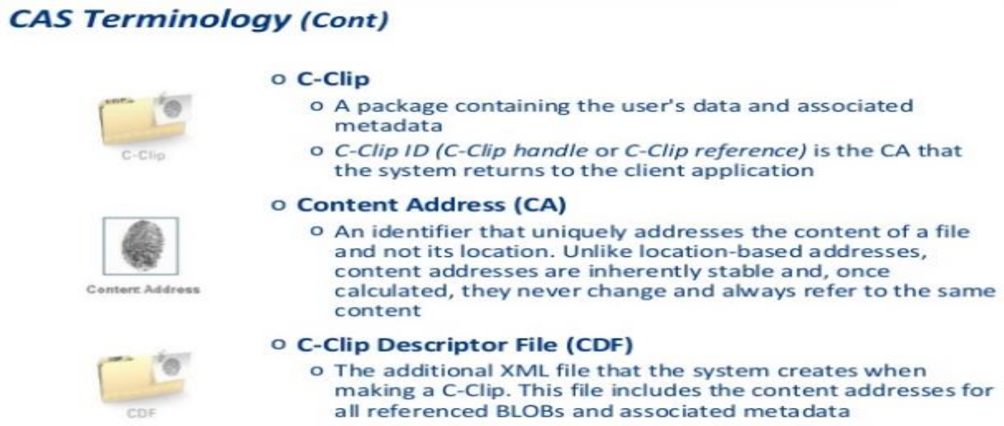
4. Retention enforcement : Protecting and retaining data objects is a core requirement of an archive system. CAS creates two immutable components: a data object and a meta-object for every object stored.

The metaobject stores object's attributes and data handling policies. For systems that support object-retention capabilities, the retention policies are enforced until the policies expire. Record-level protection and disposition

5. Technology independence: The CAS system interface is impervious to technology changes. As long as the application server is able to map the original content address the data remains accessible and ensure compatibility across platforms.

6. Fast record retrieval: CAS maintains all content on disks that provide subsecond —time to first byte (200 ms–400 ms) in a single cluster.

4b)



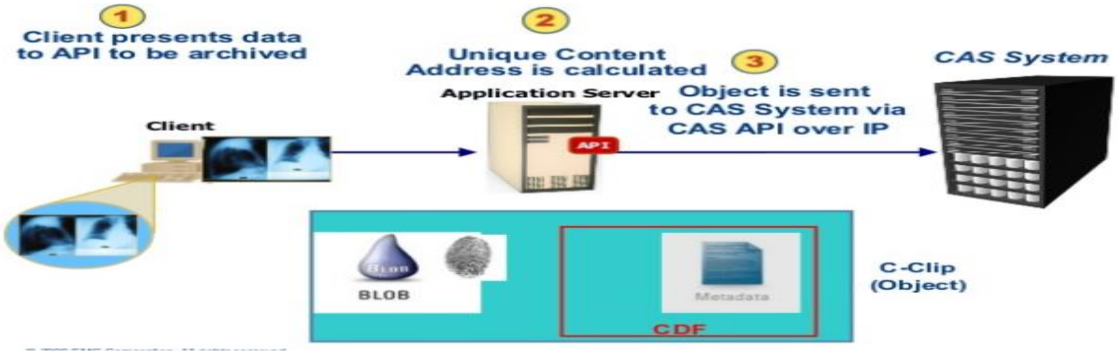
CAS Terminology (Cont)

- C-Clip
  - A package containing the user's data and associated metadata
  - *C-Clip ID (C-Clip handle* or *C-Clip reference)* is the CA that the system returns to the client application
- Content Address (CA)
  - An identifier that uniquely addresses the content of a file and not its location. Unlike location-based addresses, content addresses are inherently stable and, once calculated, they never change and always refer to the same content
- C-Clip Descriptor File (CDF)
  - The additional XML file that the system creates when making a C-Clip. This file includes the content addresses for all referenced BLOBs and associated metadata

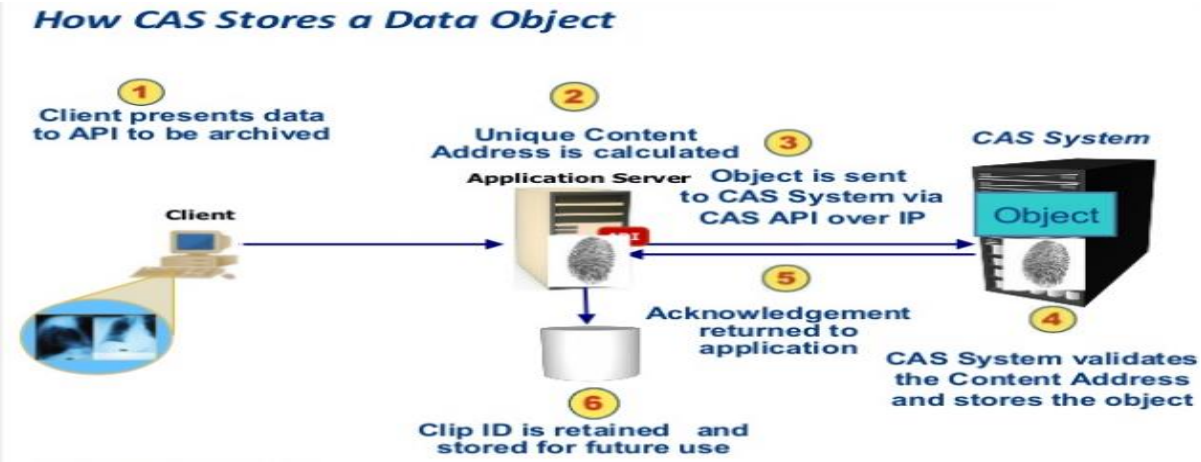
5	a) How CAS stores and retrieves the data object? Explain with an example as Health Care Solution? b) Explain back operation and restore operation with a neat diagram	[4] [6]	CO4 CO3	L3

5a)

How CAS Stores a Data Object

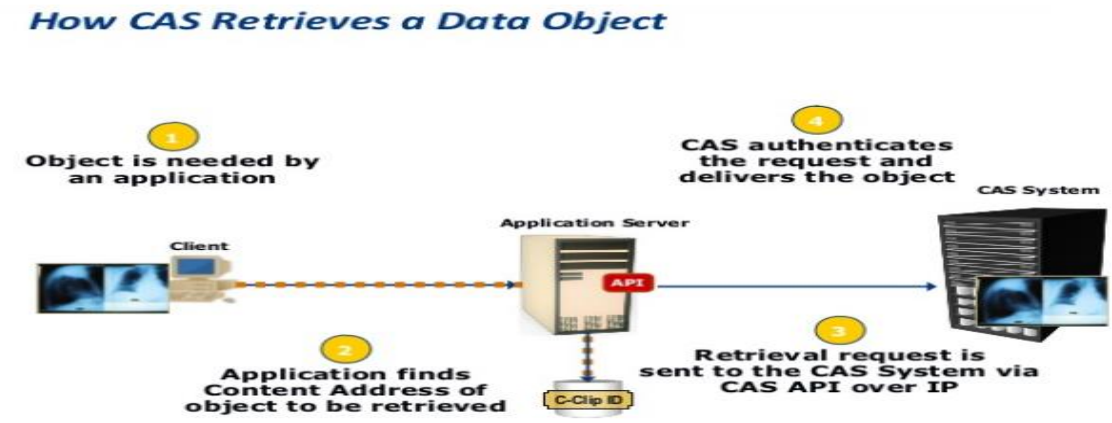


1
Client presents data to API to be archived
2
Unique Content Address is calculated
3
Object is sent to CAS System via CAS API over IP
CAS System
Client
Application Server
API
BLOB
Metadata
CDF
C-Clip (Object)



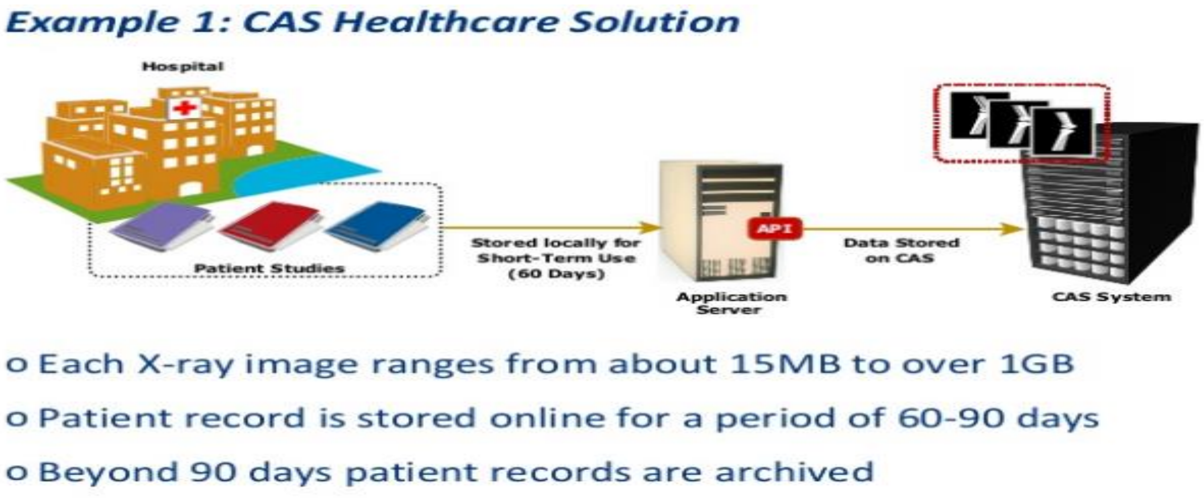
How CAS Stores a Data Object

1
Client presents data to API to be archived
2
Unique Content Address is calculated
3
Object is sent to CAS System via CAS API over IP
CAS System
Application Server
Client
Object
API
5
Acknowledgement returned to application
4
CAS System validates the Content Address and stores the object
6
Clip ID is retained and stored for future use



How CAS Retrieves a Data Object

1
Object is needed by an application
4
CAS authenticates the request and delivers the object
CAS System
Application Server
Client
API
2
Application finds Content Address of object to be retrieved
C-Clip ID
3
Retrieval request is sent to the CAS System via CAS API over IP



Example 1: CAS Healthcare Solution

- Each X-ray image ranges from about 15MB to over 1GB
- Patient record is stored online for a period of 60-90 days
- Beyond 90 days patient records are archived

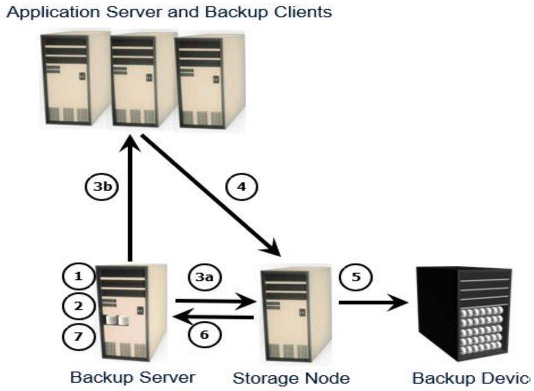
5b) **i) Back up Operation**

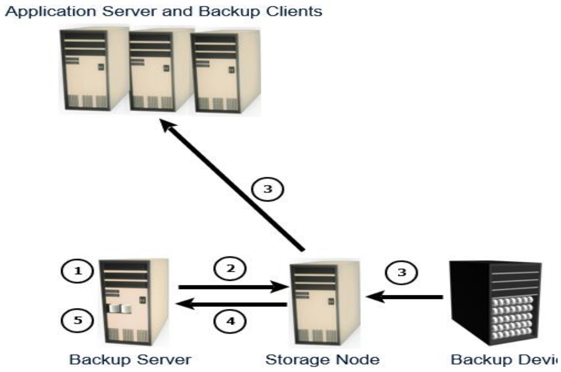
Step-1: Start of scheduled backup process Backup server retrieves backup related information from backup catalog.

Step-2: Backup server instructs backup clients to send its metadata to the backup server and data to be backed up

to storage node.

Step-3a: Backup server instructs storage node to load backup media in backup device
Step-3b: Backup server instructs backup clients to send its metadata to the backup server and data to be backed up
to storage node.
Step-4: Backup clients send data to storage node.
Step-5: Storage node sends data to backup device Storage node sends media information to backup server
Step-6: Backup server update catalog and records the status.





ii) Restore Operation

Step-1: Backup server scans backup catalog to identify data to be restore and the client that will receive data.
Step-2: Backup server instructs storage node to load backup media in backup device.
Step-3: Data is then read and send to backup client.
Step-4: Storage node sends restore metadata to backup server.
Step-5: Backup server updates catalog.

6	a) Explain BC planning life cycle? b) Explain the following BC terminology: i) disaster recovery, disaster restart ii) RPO & RTO iii)) hot & cold site	[4] [6]	CO4	L2

6a) BC planning must follow a disciplined approach like any other planning process. Organizations today dedicate specialized resources to develop and maintain BC plans. From the conceptualization to the realization of the BC plan, a lifecycle of activities can be defined for the BC process. The BC planning lifecycle includes five stages

Several activities are performed at each stage of the BC planning lifecycle, including the following key activities:

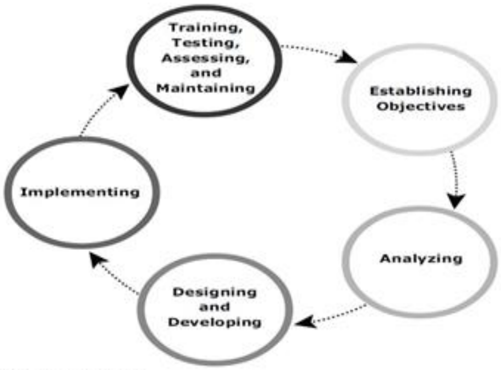
1. Establishing objectives

- Determine BC requirements
- Estimate the scope and budget to achieve requirements.
- Select a BC team by considering subject matter experts from all areas of the business, whether internal or external.
- Create BC policies.

2. Analyzing

- Collect information on data profiles, business processes, infrastructure support, dependencies, and frequency of using business infrastructure.

- Identify critical business needs and assign recovery priorities.
- Create a risk analysis for critical areas and mitigation strategies.
- Conduct a Business Impact Analysis (BIA).
- Create a cost and benefit analysis based on the consequences of data unavailability.
- Evaluate options.



BC planning lifecycle

3. Designing and developing

- Define the team structure and assign individual roles and responsibilities. For example, different teams are formed for activities such as emergency response, damage assessment, and infrastructure and application recovery.
- Design data protection strategies and develop infrastructure.
- Develop contingency scenarios.
- Develop emergency response procedures.
- Detail recovery and restart procedures.

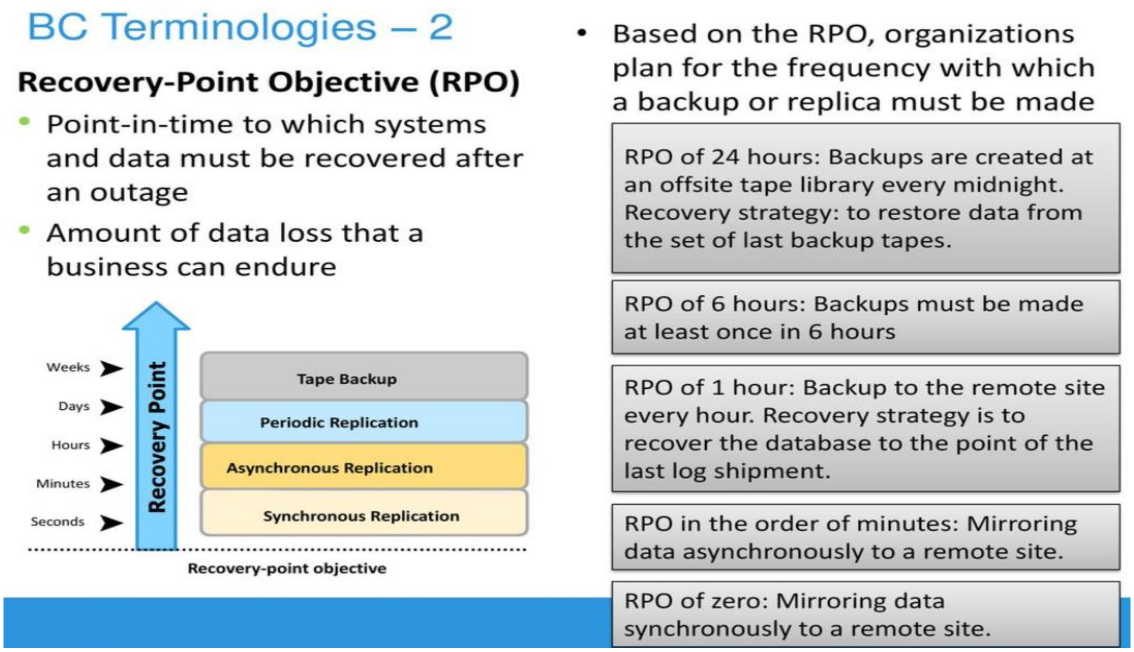
4. Implementing

- Implement risk management and mitigation procedures that include backup, replication, and management of resources.
- Prepare the disaster recovery sites that can be utilized if a disaster affects the primary data center.
- Implement redundancy for every resource in a data center to avoid single points of failure.

5. Training, testing, assessing, and maintaining

- Train the employees who are responsible for backup and replication of business-critical data on a regular basis or whenever there is a modification in the BC plan.
- Train employees on emergency response procedures when disasters are declared and recovery procedures based on contingency scenarios.
- Test the BC plan regularly to evaluate its performance and identify its limitations.
- Assess the performance reports and identify limitations.
- Update the BC plans and recovery/restart procedures to reflect regular changes within the data center.

6b)



BC Terminologies – 2

Recovery-Point Objective (RPO)

- Point-in-time to which systems and data must be recovered after an outage
- Amount of data loss that a business can endure

Weeks
Days
Hours
Minutes
Seconds

Recovery Point

Tape Backup

Periodic Replication

Asynchronous Replication

Synchronous Replication

Recovery-point objective

- Based on the RPO, organizations plan for the frequency with which a backup or replica must be made

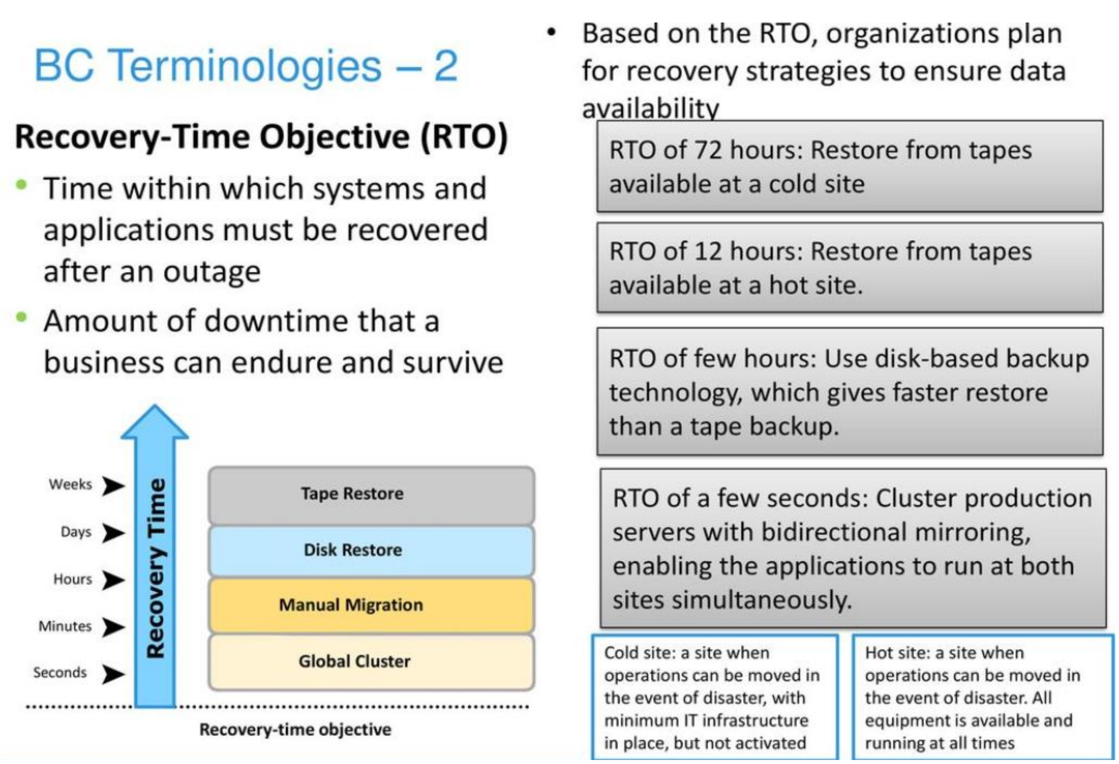
RPO of 24 hours: Backups are created at an offsite tape library every midnight. Recovery strategy: to restore data from the set of last backup tapes.

RPO of 6 hours: Backups must be made at least once in 6 hours

RPO of 1 hour: Backup to the remote site every hour. Recovery strategy is to recover the database to the point of the last log shipment.

RPO in the order of minutes: Mirroring data asynchronously to a remote site.

RPO of zero: Mirroring data synchronously to a remote site.



BC Terminologies – 2

Recovery-Time Objective (RTO)

- Time within which systems and applications must be recovered after an outage
- Amount of downtime that a business can endure and survive

Weeks
Days
Hours
Minutes
Seconds

Recovery Time

Tape Restore

Disk Restore

Manual Migration

Global Cluster

Recovery-time objective

- Based on the RTO, organizations plan for recovery strategies to ensure data availability

RTO of 72 hours: Restore from tapes available at a cold site

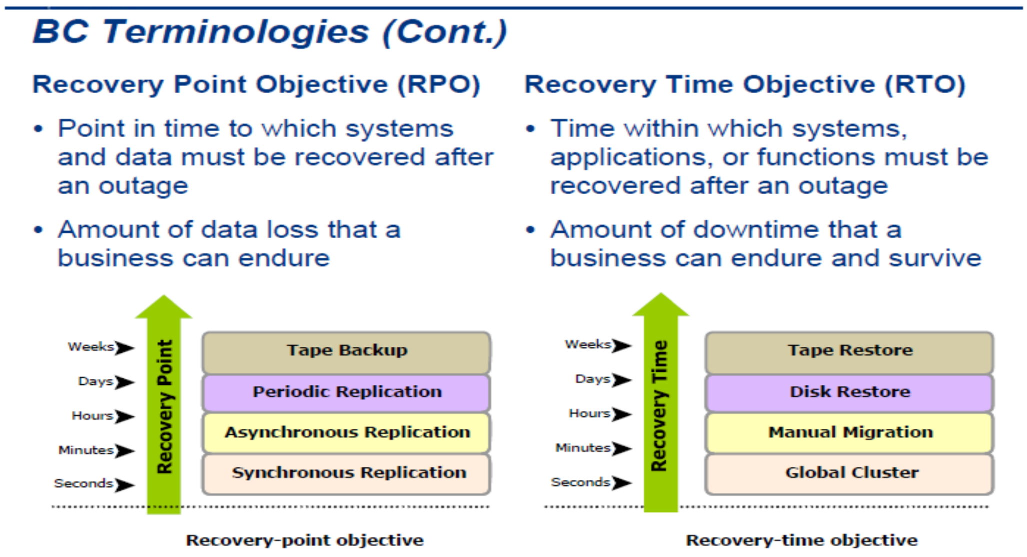
RTO of 12 hours: Restore from tapes available at a hot site.

RTO of few hours: Use disk-based backup technology, which gives faster restore than a tape backup.

RTO of a few seconds: Cluster production servers with bidirectional mirroring, enabling the applications to run at both sites simultaneously.

Cold site: a site when operations can be moved in the event of disaster, with minimum IT infrastructure in place, but not activated

Hot site: a site when operations can be moved in the event of disaster. All equipment is available and running at all times



BC Terminologies (Cont.)

Recovery Point Objective (RPO)

- Point in time to which systems and data must be recovered after an outage
- Amount of data loss that a business can endure

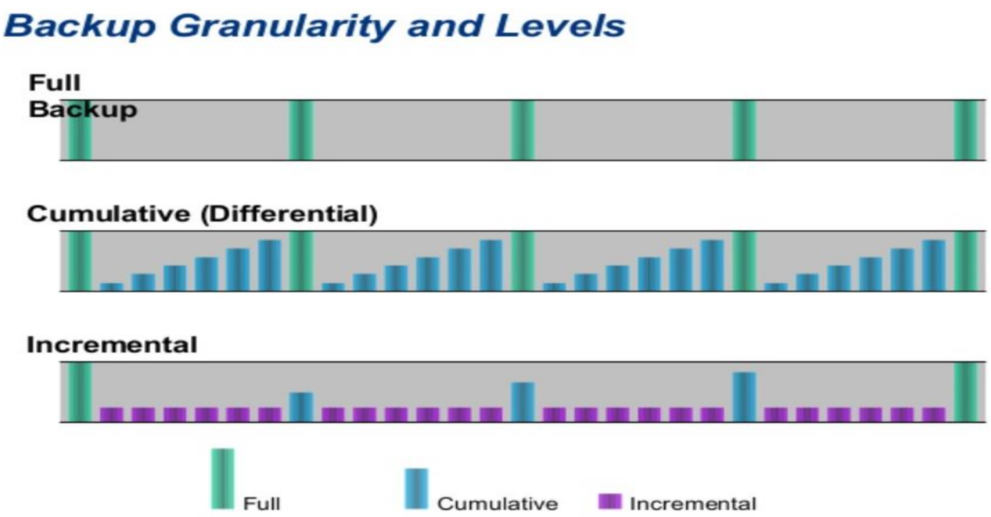
Recovery Time Objective (RTO)

- Time within which systems, applications, or functions must be recovered after an outage
- Amount of downtime that a business can endure and survive

7	a) Explain different level of granularity in Backup with a neat diagram b) Discuss different backup Topologies	[4] [6]	CO4	L3

7a) Backup granularity depends on business needs and required RTO/RPO

- **Full backup** is a backup of the complete data on the production volumes at a certain point in time. A full backup copy is created by copying the data on the production volumes to a secondary storage device.
- **Incremental backup** copies the data that has changed since the last full or incremental backup, whichever has occurred more recently. This is much faster (because the volume of data backed up is restricted to changed data), but it takes longer to restore.
- **Cumulative (or differential) backup** copies the data that has changed since the last full backup. This method takes longer than incremental backup but is faster to restore.



Backup Granularity and Levels

7b) Backup Topologies

Three basic topologies are used in a backup environment: direct attached backup, LAN based backup, and SAN based backup. A mixed topology is also used by combining LAN based and SAN based topologies.

1. In a direct-attached backup, a backup device is attached directly to the client. Only the metadata is sent to the backup server through the LAN. This configuration frees the LAN from backup traffic. The example shown in Figure 12-7 depicts use of a backup device that is not shared. As the environment grows, however, there will be a need for central management of all backup devices and to share the resources to optimize costs.

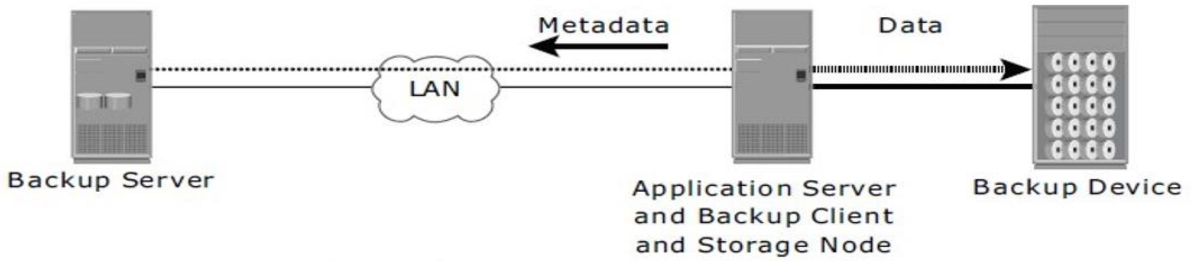


Figure 12-7: Direct-attached backup topology

2. In LAN-based backup, all servers are connected to the LAN and all storage devices are directly attached to the storage node (see Figure 12-8). The data to be backed up is transferred from the backup client (source), to the backup device (destination) over the LAN, which may affect network performance. Streaming across the LAN also affects network performance of all systems connected to the same segment as the backup server. Network resources are severely constrained when multiple clients access and share the same tape library unit (TLU).

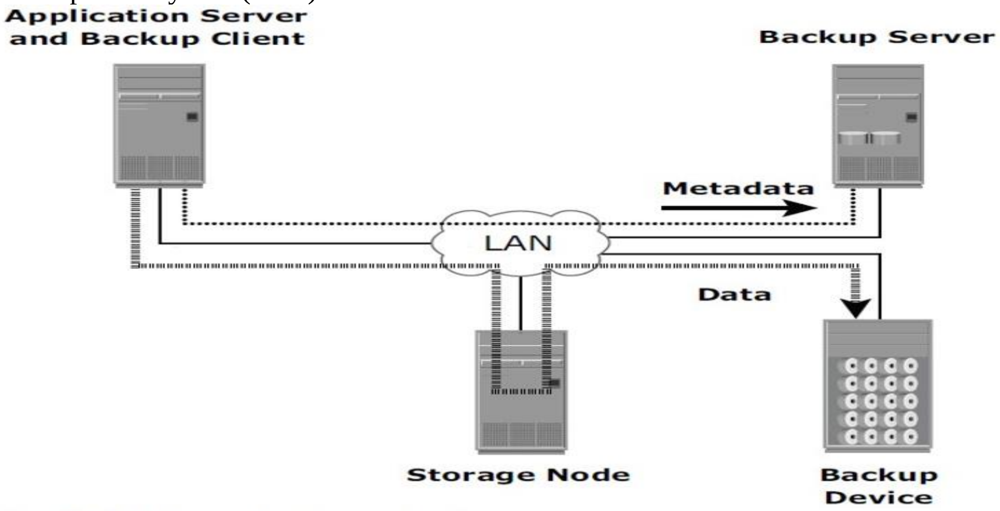


Figure 12-8: LAN-based backup topology

3. The SAN-based backup is also known as the LAN-free backup. Figure 12-9 illustrates a SAN-based backup. The SAN- based backup topology is the most appropriate solution when a backup device needs to be shared among the clients. In this case the backup device and clients are attached to the SAN.

In this example, clients read the data from the mail servers in the SAN and write to the SAN attached backup device. The backup data traffic is restricted to the SAN, and backup metadata is transported over the LAN. However, the volume of metadata is insignificant when compared to production data. LAN performance is not degraded in this configuration.

4. The mixed topology uses both the LAN-based and SAN-based topologies, as shown in Figure 12-10. This topology might be implemented for several reasons, including cost, server location, reduction in administrative overhead, and performance considerations.

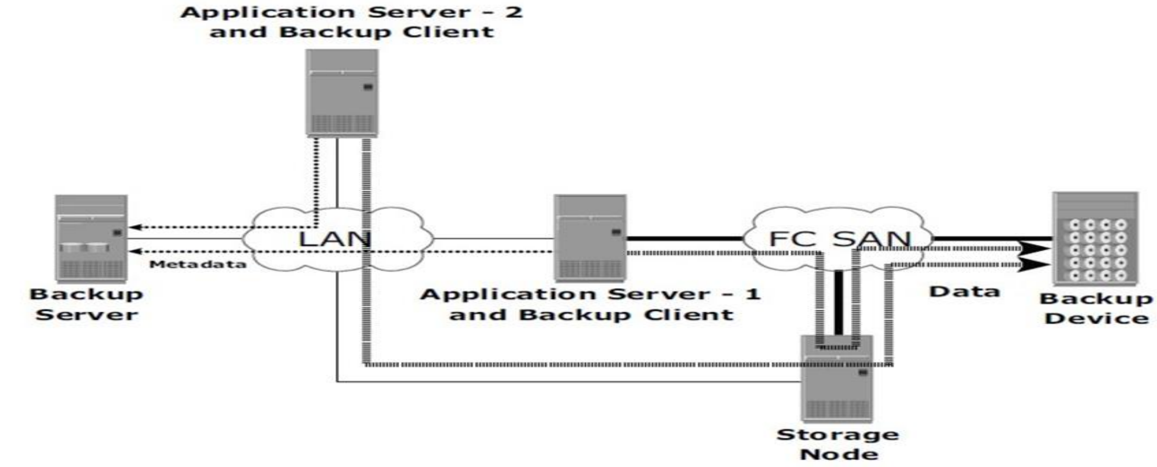


Figure 12-10: Mixed backup topology

8	a) Explain local Replication technology using Host based methods	[10]	CO4	L2

Sol : Replication is the process of creating an exact copy of data. Creating one or more replicas of the production data is one of the ways to provide Business Continuity (BC). These replicas can be used for recovery and restart operations in the event of data loss. Local Replication technologies can be classified based on where the replication is performed.

i) Host based Local Replication
ii) Array based Local Replication

Host-Based Local Replication

In this,all the replication is performed by using the CPU resources of the host using software that is running on the host.There are two common methods of host-based local replication.

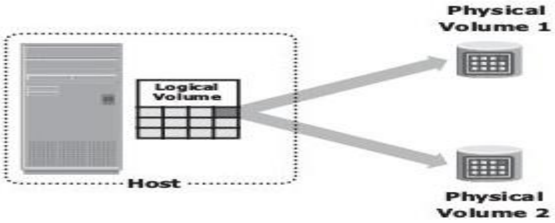
- LVM-based replication
- File System Snapshot (FS)

i)**LVM-based replication**

LVM (logical volume manager) is responsible for creating and controlling the host-level logical volume.

LVM has three components: physical volumes (physical disk), volume groups, and logical volumes.

- ➢ A *volume group* is created by grouping together one or more physical volumes. *Logical volumes* are created within a given volume group.
- ➢ A volume group can have multiple logical volumes. In LVM-based replication, each logical partition in a logical volume is mapped to two physical partitions on two different physical volumes.
- ➢ An application write to a logical partition is written to the two physical partitions by the LVM device driver. This is also known as LVM mirroring.
- ➢ Mirrors can be split and the data contained therein can be independently accessed. LVM mirrors can be added or removed dynamically.



Advantage - LVM-based replication is not dependent on a vendor-specific storage system. LVM is part of the operating system, and no additional license is required to deploy LVM mirroring.

Limitation : Every write generated by an application translates into two writes on the disk, and thus, an additional burden is placed on the host CPU.This can degrade application performance.

ii) File system (FS) snapshot

File system (FS) snapshot is a pointer-based replica that requires a fraction of the space used by the original FS.

- It uses Copy on First Write (CoFW) principle. When the snapshot is created, a bitmap and a block map are created in the metadata of the Snap FS.
- The bitmap is used to keep track of blocks that are changed on the production FS after creation of the snap. The block map is used to indicate the exact address from which data is to be read when the data is accessed from the Snap FS.
- Immediately after creation of the snapshot all reads from the snapshot will actually be served by reading the production FS.
- To read from the Snap FS, the bitmap is consulted. If the bit is 0, then the read is directed to the production FS. If the bit is 1, then the block address is obtained from the block map and data is read from that address.

For example an update occurs to the 3rd block in the Prod FS which has a value of ‘Data c’. The new value should be ‘Data C’ The snapshot application will hold the IO to the Prod FS and will first copy the old data ‘Data c’ to an available data block on the Snap FS. The bitmap and block map of location 3 is changed in the snap metadata. The bitmap of 3 is changed to 1 indicating that this block has changed on the Prod FS. The block map of 3 is changed to indicate the block number where the data is actually written in Snap FS, (in this case block 2). Once this is done the IO to the Prod FS will be allowed to complete. Any new writes to the 3th block on the Prod. In a similar manner if an IO is issued the 4th block on the Prod FS to change the value from ‘Data d’ to ‘Data D’

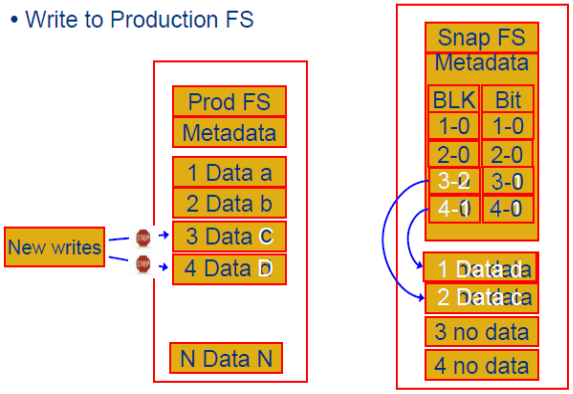


Fig-1: Write to Source

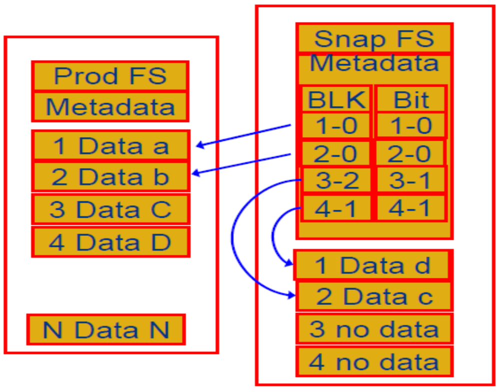


Fig-2: Read from production

The bitmap of the Snap FS indicates that the data in blocks 1 and 2 have not changed since the creation of the snap hence the data should actually be read from the Prod FS. The bitmap for blocks 3 and 4 indicate that the data on the Prod FS has changed since the creation of the snap and that the block map should be consulted. The block map for block 3 indicates that the data is located in the 2nd data block of the Snap FS, while the block map for block 4 indicates that the data is located in the 1st data block.