DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING
INTERNAL ASSESSMENT TEST – 3
STORAGE AREA NETWORKS (10IS765)
QUESTION PAPER & KEY

USN | | | | | | | | | |

CMRIT
CELEBRATING 25 YEARS
CMR INSTITUTE OF TECHNOLOGY, BENGALURU.
ACCREDITED WITH A+ GRADE BY NAAC

Improvement Test – Nov. 2017

Sub:	STORAGE AREA NETWORKS (SAN)				Sub Code:	10IS765	Branch:	ISE
Date:	20.11.2017	Duration:	90 min's	Max Marks: 50	Sem / Sec:	VII		OBE

	Answer any FIVE FULL Questions	MARKS	CO	RBT
1 (a)	Consider a disk I/O system, I/O request arrives at a rate of 100 I/Os per second. The service time Rs is 6ms. Compute the following measures of disk performance: a. Utilization of I/O controller, b. Total response time, c. Average Queue size, & d. Total time spent by request in a queue. Now, if controller power is doubled, the service time is halved; consequently, Rs = 3ms. In this scenario, determine the above measures.	[10]	CO1	L3
2 (a)	With a neat diagram, discuss the features of high end storage systems.	[10]	C02	L3
3 (a)	Explain the different FC ports with a neat diagram.	[10]	C03	L4
4 (a)	What is storage virtualization? Describe the types of storage virtualization in detail with diagram.	[10]	C04	L3
5 (a)	With neat diagram, explain the steps involved in BC backup and restore operation .	[10]	C05	L3
6 (a)	Describe SCSI - 3 architecture in detail with diagram.	[10]	C03	L4
7 (a)	What is storage array based remote replication? Differentiate between synchronous and asynchronous replication mode in it.	[10]	C06	L2
8(a)	What is storage array based local replication? List the various storage array based local replication technologies. Explain any one in detail.	[10]	C06	L3

1. Consider a disk I/O system, I/O request arrives at a rate of 100 I/Os per second. The service time Rs is 8ms. Compute the following measures of disk performance:

a. Utilization of I/O controller,

b. Total response time,

c. Average Queue size, &

d. Total time spent by request in a queue.

Now, if controller power is doubled, the service time is halved; consequently, Rs = 4ms. In this scenario, determine the above measures.

Solution:

- Arrival Rate (a)

R_a= 1/a = 1 sec / 100 = 1000ms/100

R_a = 10 ms

- Rs = 6 ms (Given)

a) Utilization (U) = Rs / Ra = 6 / 10 = 0.6 or 60%

b) Response Time (R) = Rs / (1-U) = 6 / (1 – 0.6) = 6 / 0.4 = 15 ms
c) Average Queue Size = U / (1-U) = 0.6 / (1-0.6) = 0.6 / 0.2 = 3
d) Total time spent by request in a queue = U * R = 0.6 * 15 ms = 9 ms

- Arrival Rate (a)

R_a= 1/a = 1 sec / 200 = 1000ms/200

R_a = 5 ms

- Rs =3 ms (Given)
  a) Utilization (U) = Rs / Ra = 3 / 5 = 0.6 or 60%
  b) Response Time (R) = Rs / (1-U) = 3 / (1 – 0.6) = 3 / 0.4 = 7.5 ms
  c) Average Queue Size = U / (1-U) = 0.6 / (1-0.6) = 0.6 / 0.4 = 1.5
  d) Total time spent by request in a queue = U * R = 0.6 * 7.5 ms = 4.5 ms

2. With a neat diagram, discuss the features of high end storage systems.
Solution:

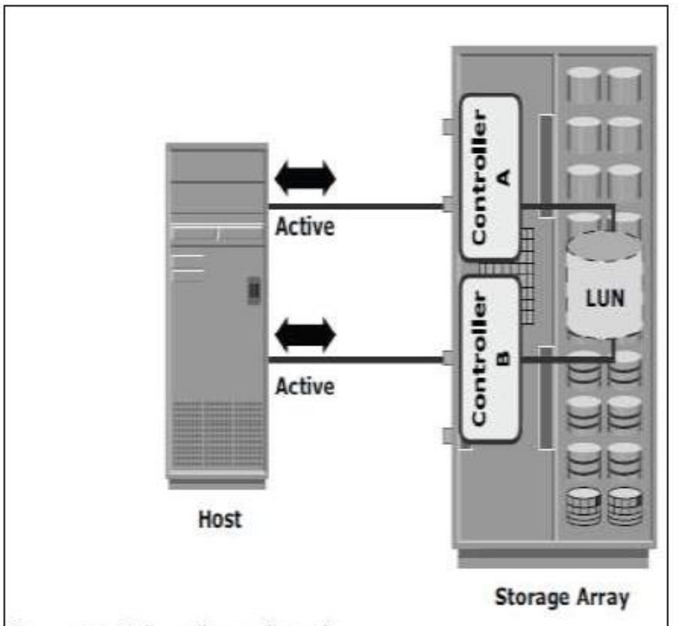


Figure 4-7: Active-active configuration

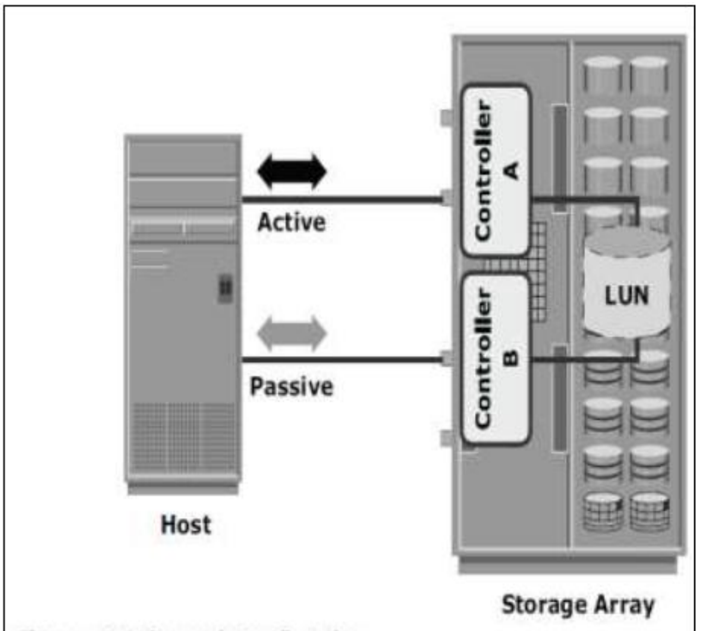


Figure 4-8: Active-passive configuration

High-end storage systems, referred to as ***active-active arrays***, are aimed at large enterprises for centralizing corporate data. These arrays are designed with a large number of controllers and cache memory. An active-active array implies that the host can perform I/Os to its LUNs across any of the available paths (see Figure 4-7).

To address the enterprise storage needs, these arrays provide the following capabilities:
1. Large storage capacity
2. Large amounts of cache to service host I/Os optimally
3. Fault tolerance architecture to improve data availability
4. Connectivity to mainframe computers and open systems hosts
5. Availability of multiple front-end ports and interface protocols to serve a large number of hosts
6. Availability of multiple back-end Fibre Channel or SCSI RAID controllers to manage disk processing
7. Scalability to support increased connectivity, performance, and storage capacity requirements
8. Ability to handle large amounts of concurrent I/Os from a number of servers and applications
9. Support for array-based local and remote replication

3. Explain the different FC ports with a neat diagram.

Solution:

Fibre Channel Ports

Ports are the basic building blocks of an FC network. Ports on the switch can be one of the following types:

1. N_port: An end point in the fabric. This port is also known as the *node port*. Typically, it is a host port (HBA) or a storage array port that is connected to a switch in a switched fabric.

2. NL_port: A node port that supports the arbitrated loop topology. This port is also known as the *node loop port*.

3. E_port: An FC port that forms the connection between two FC switches. This port is also known as the *expansion port*. The E_port on an FC switch connects to the E_port of another FC switch in the fabric through a link, which is called an *Inter-Switch Link (ISL)*. ISLs are used to transfer host-to-storage data as well as the fabric management traffic from one switch to another. ISL is also one of the scaling mechanisms in SAN connectivity.

4. F_port: A port on a switch that connects an N_port. It is also known as a *fabric port* and cannot participate in FC-AL.

5. FL_port: A fabric port that participates in FC-AL. This port is connected to the NL_ports on an FC-AL loop. A FL_port also connects a loop to a switch in a switched fabric. As a result, all NL_ports in the loop can participate in FC-SW. This configuration is referred to as a *public loop*. In contrast, an arbitrated loop without any switches is referred to as a *private loop*. A private loop contains nodes with NL_ports, and does not contain FL_port.

6. G_port: A generic port that can operate as an E_port or an F_port and determines its functionality automatically during initialization.

Figure 6-12 shows various FC ports located in the fabric.

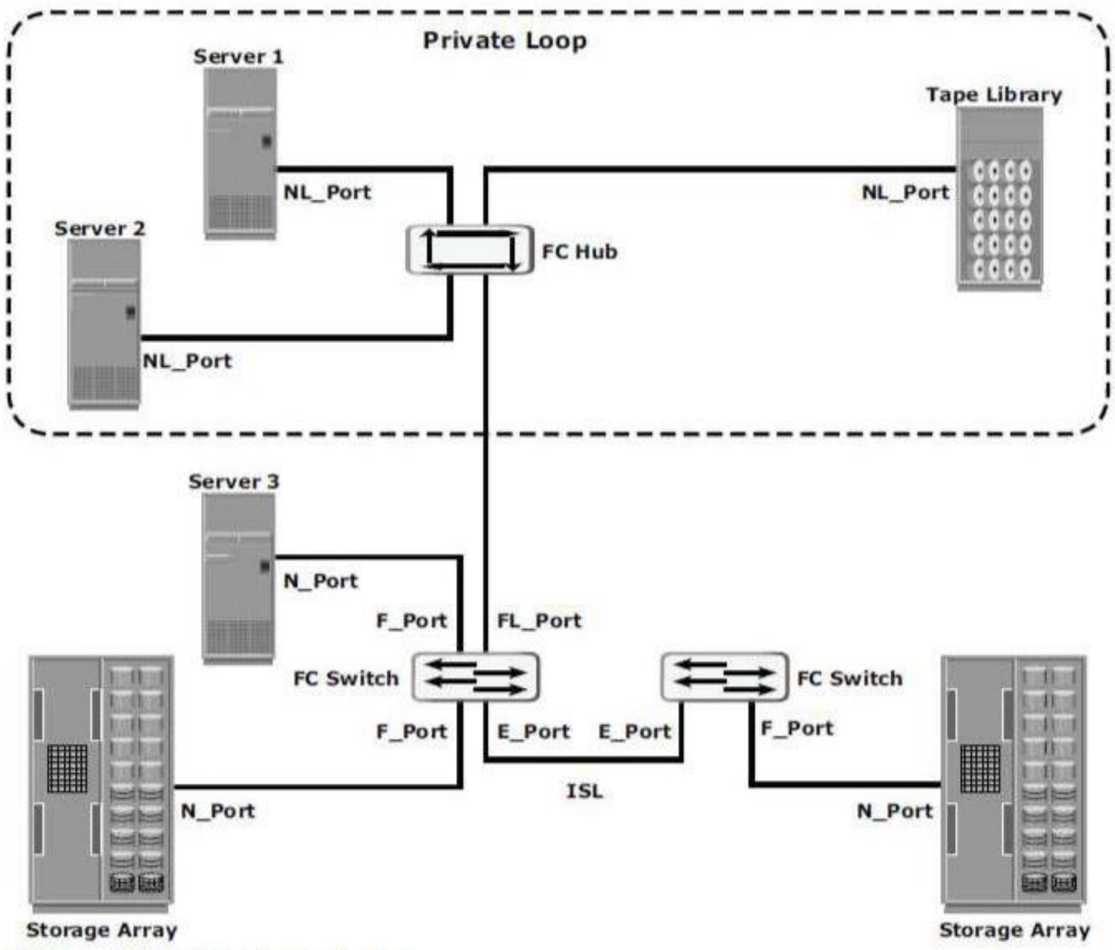


Figure 6-12: Fibre channel ports

4. What is storage virtualization? Describe the types of storage virtualization in detail with diagram.

Solution:
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.

10.5.1 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.

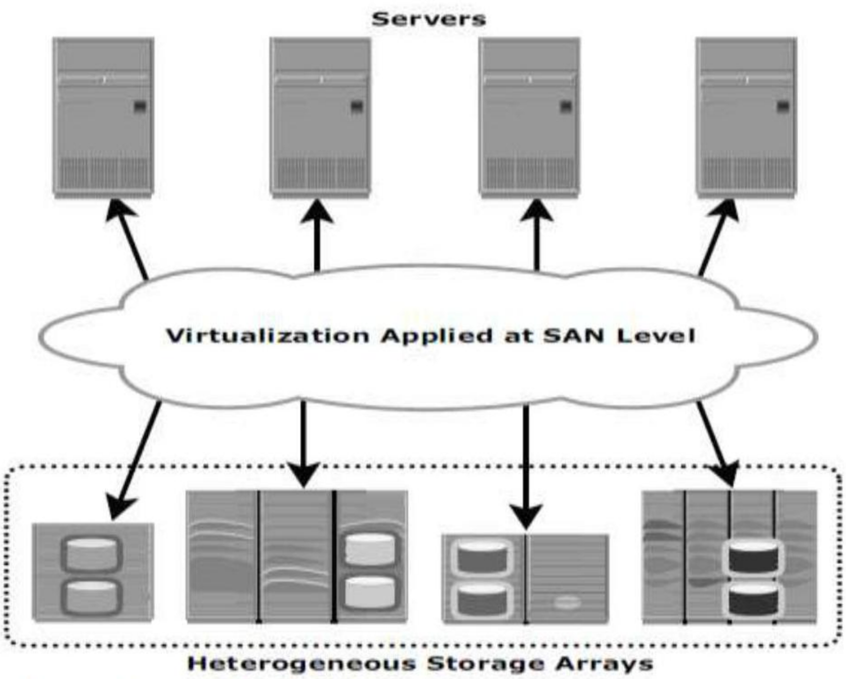


Figure 10-6: Block-level storage virtualization

10.5.2 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.

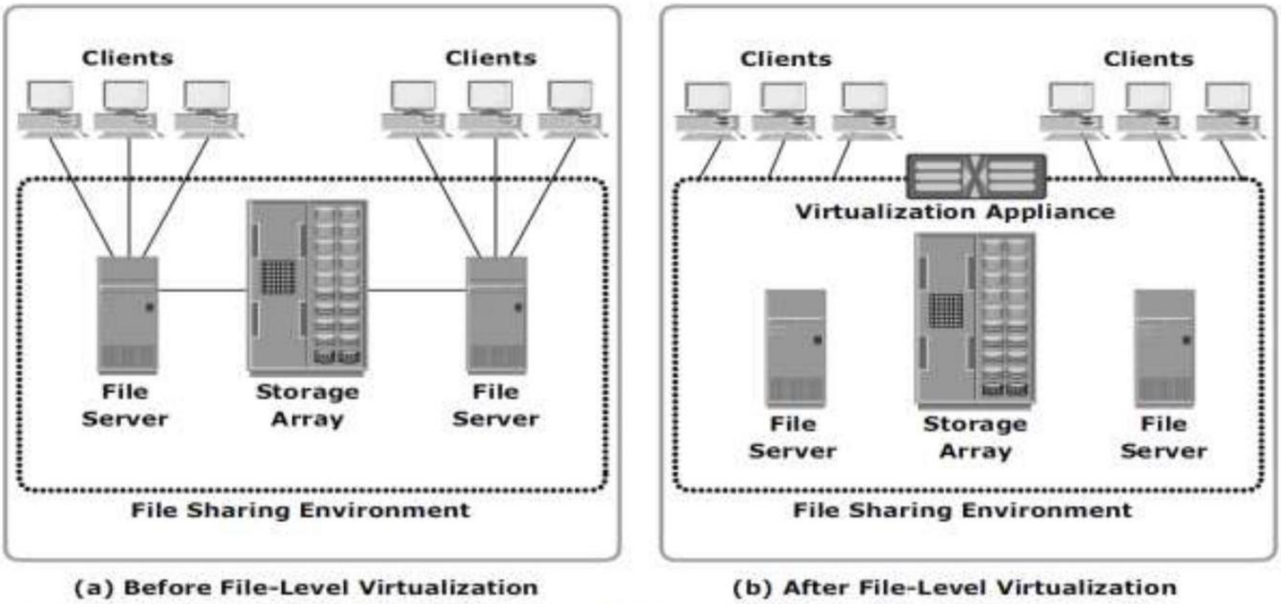


Figure 10-7: NAS device before and after file-level virtualization

5. With neat diagram, explain the steps involved in BC backup and restore operation .
Solution:

When a backup process is initiated, significant network communication takes place between the different components of a backup infrastructure. The backup server initiates the backup process for different clients based on the backup schedule configured for them. For example, the backup process for a group of clients may be scheduled to start at 3:00 am every day.
The backup server coordinates the backup process with all the components in a backup configuration (see Figure 12-5). The backup server maintains the information about backup clients to be contacted and storage nodes to be used in a backup operation. The backup server retrieves the backup-related information from the backup catalog and, based on this information, instructs the storage node to load the appropriate backup media into the backup devices. Simultaneously, it instructs the backup clients to start scanning the data, package it, and send it over the network to the assigned storage node. The storage node, in turn, sends metadata to the backup server to keep it updated about the media being used in the backup process. The backup server continuously updates the backup catalog with this information.

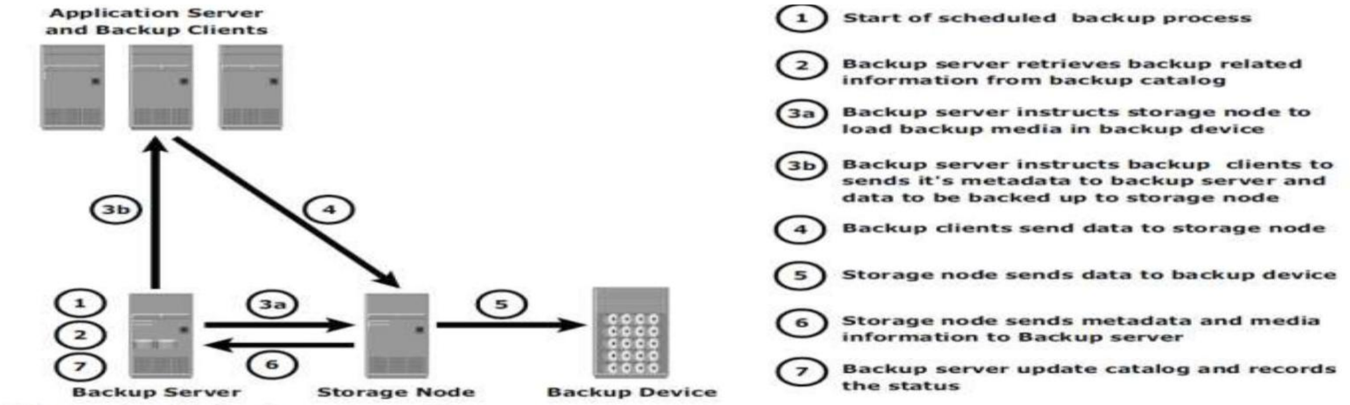


Figure 12-5: Backup operation

After the data is backed up, it can be restored when required. A restore process must be manually initiated. Some backup software has a separate application for restore operations. These restore applications are accessible only to the administrators.
Figure 12-6 depicts a restore process.

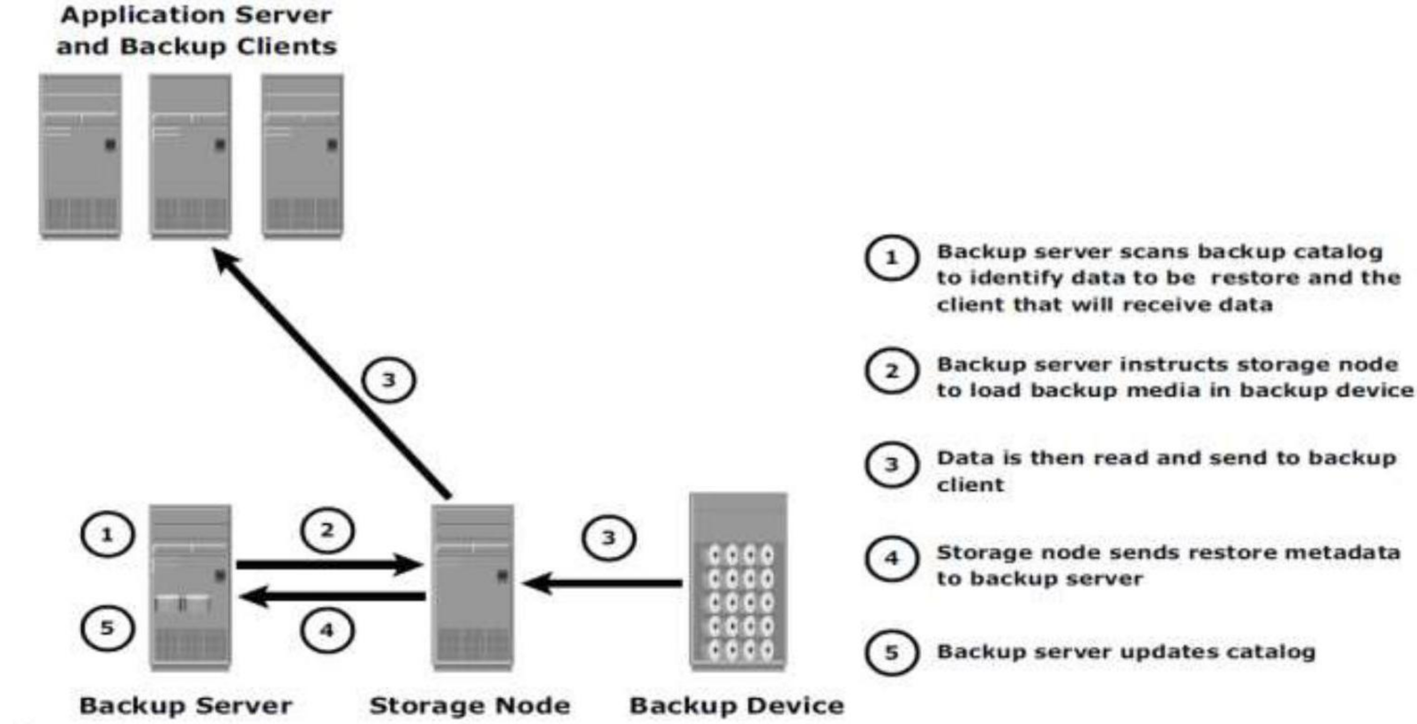


Figure 12-6: Restore operation

Upon receiving a restore request, an administrator opens the restore application to view the list of clients that have been backed up. While selecting the client for which a restore request has been made, the administrator also needs to identify the client that will receive the restored data. Data can be restored on the same client for whom the restore request has been made or on any other client.
The administrator then selects the data to be restored and the specified point in time to which the data has to be restored based on the RPO. Note that because all of this information comes from the backup catalog, the restore application must also communicate to the backup server.
The administrator first selects the data to be restored and initiates the restore process. The backup server, using the appropriate storage node, then identifies the backup media that needs to be mounted on the backup devices. Data is then read and sent to the client that has been identified to receive the restored data.

6. Describe SCSI - 3 architecture in detail with diagram.
Solution:

The SCSI-3 architecture defines and categorizes various SCSI-3 standards and requirements for SCSI-3 implementations. The SCSI-3 architecture was approved and published as the standard X.3.270-1996 by the ANSI. This architecture helps developers, hardware designers, and users to understand and effectively utilize SCSI.

The three major components of a SCSI architectural model are as follows:

1. SCSI-3 command protocol: This consists of primary commands that are common to all devices as well as device-specific commands that are unique to a given class of devices.

2. Transport layer protocols: These are a standard set of rules by which devices communicate and share information.

3. Physical layer interconnects: These are interface details such as electrical signaling methods and data transfer modes.

Common access methods are the ANSI software interfaces for SCSI devices.

Figure 5-3 shows the SCSI-3 standards architecture with interrelated groups of other standards within SCSI-3.

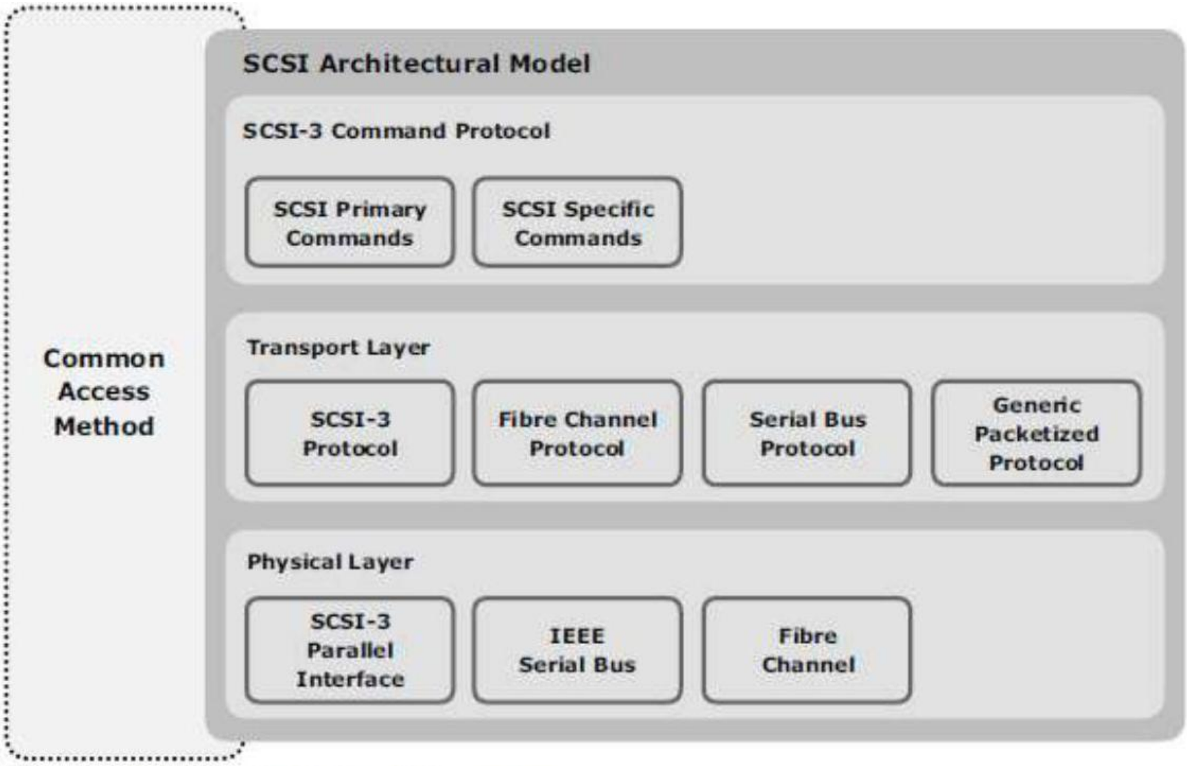


Figure 5-3: SCSI-3 standards architecture

7. What is storage array based remote replication? Differentiate between synchronous and asynchronous replication mode in it.

Solution:

Remote replication is the process of creating replicas of information assets at remote sites (locations). Remote replicas help organizations mitigate the risks associated with regionally driven outages resulting from natural or human-made disasters. Similar to local replicas, they can also be used for other business operations. The infrastructure on which information assets are stored at the primary site is called the *source*. The infrastructure on which the replica is stored at the remote site is referred to as the *target*. Hosts that access the source or target are referred to as *source hosts* or *target hosts,* respectively.

Modes of Remote Replication

The two basic modes of remote replication are synchronous and asynchronous. In *synchronous remote replication*, writes must be committed to the source and the target, prior to acknowledging "write complete" to the host. Additional writes on the source cannot occur until each preceding write has been completed and acknowledged. This ensures that data is identical on the source and the replica at all times. Further writes are transmitted to the remote site exactly in the order in which they are received at the source. Hence, write

ordering is maintained. In the event of a failure of the source site, synchronous remote replication provides zero or near-zero RPO, as well as the lowest RTO. However, application response time is increased with any synchronous remote replication. The degree of the impact on the response time depends on the distance between sites, available bandwidth, and the network connectivity infrastructure. The distances over which synchronous replication can be deployed depend on the application's ability to tolerate extension in response time. Typically, it is deployed for distances less than 200 KM (125 miles) between the two sites. In *asynchronous remote replication*, a write is committed to the source and immediately acknowledged to the host. Data is buffered at the source and transmitted to the remote site later. This eliminates the impact to the application's response time. Data at the remote site will be behind the source by at least the size of the buffer. Hence, asynchronous remote replication provides a finite (nonzero) RPO disaster recovery solution. RPO depends on the size of the buffer, available network bandwidth, and the write workload to the source.
There is no impact on application response time, as the writes are acknowledged immediately to the source host. This enables deployment of asynchronous replication over extended distances. Asynchronous remote replication can be deployed over distances ranging from several hundred to several thousand kilometers between two sites.

Figure**:** Synchronous replication

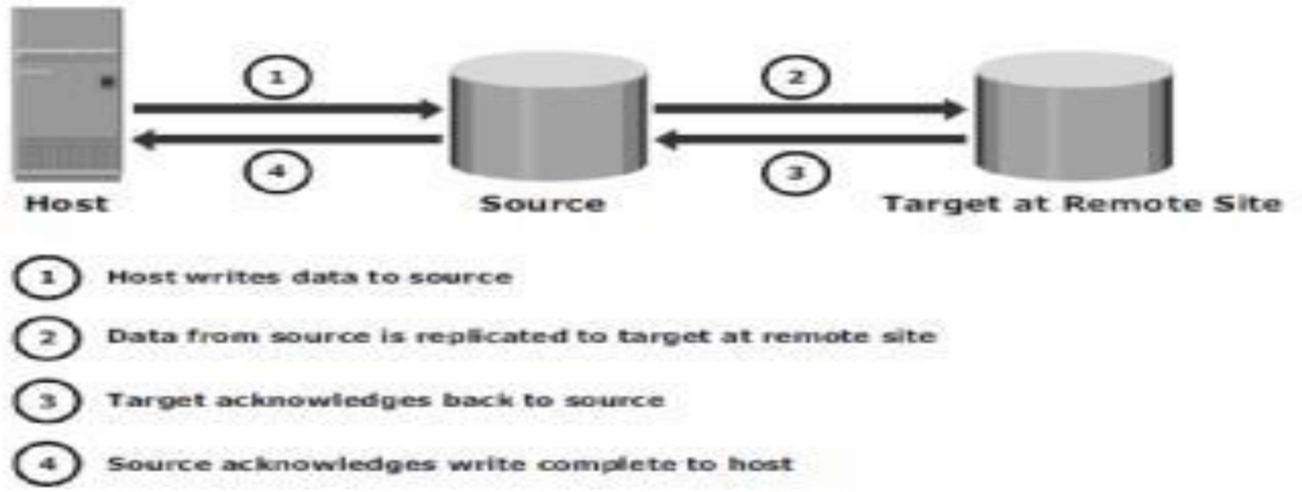
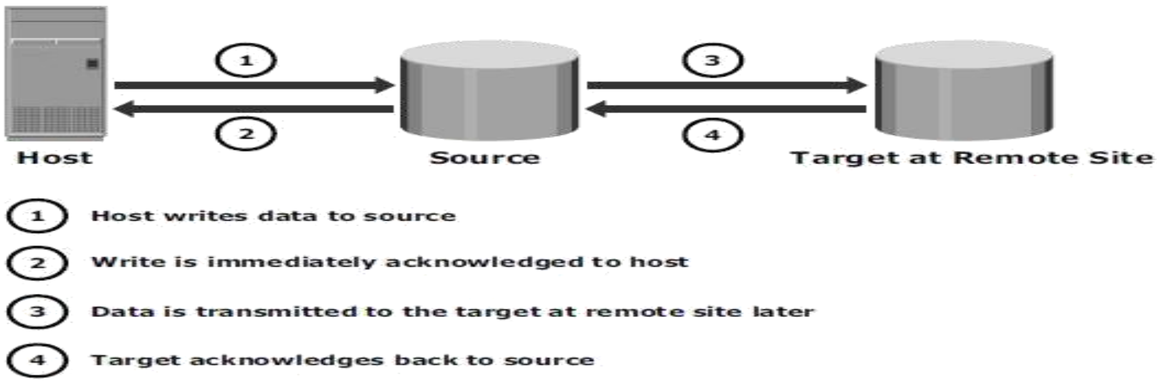


Figure**:** Asynchronous replication



8. What is storage array based local replication? List the various storage array based local replication technologies. Explain any one in detail.
Solution:

In *storage array-based local replication*, the array operating environment performs the local replication process. The host resources such as CPU and memory are not used in the replication process. Consequently, the host is not burdened by the replication operations. The replica can be accessed by an alternate host for any business operations. In this replication, the required number of replica devices should be selected on the same array and then data is replicated between source-replica pairs. A database could be laid out over multiple physical volumes and in that case all the devices must be replicated for a consistent PIT copy of the database.

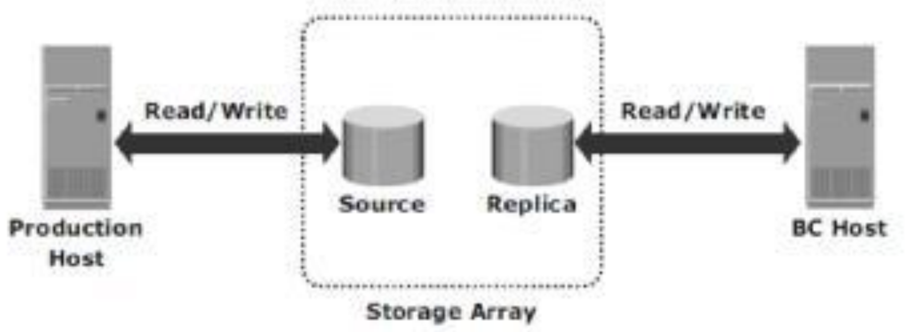


Figure**:** Storage array-based replication

Storage array-based local replication can be further categorized as full-volume mirroring, pointer-based full-volume replication, and pointer-based virtual replication. Replica devices are also referred as target devices, accessible by business continuity host. The different techniques are:

- ***Full-Volume Mirroring***
- ***Pointer-Based, Full-Volume Replication***
- ***Pointer-Based Virtual Replication***

Full-Volume Mirroring

In *full-volume mirroring*, the target is attached to the source and established as a mirror of the source. Existing data on the source is copied to the target. New updates to the source are also updated on the target. After all the data is copied and both the source and the target contain identical data, the target can be considered a mirror of the source. While the target is attached to the source and the synchronization is taking place, the target remains unavailable to any other host. However, the production host can access the source. After synchronization is complete, the target can be detached from the source and is made available for BC operations. Notice that both the source and the target can be accessed for read and write operations by the production hosts. After the split from the source, the target becomes a PIT copy of the source. The point-in-time of a replica is determined by the time when the source is detached from the target. For example, if the time of detachment is 4:00 pm, the PIT for the target is 4:00 pm. After detachment, changes made to both source and replica can be tracked at some predefined granularity. This enables incremental resynchronization (source to target) or incremental restore (target to source). The granularity of the data change can range from 512 byte blocks to 64 KB blocks. Changes are typically tracked using bitmaps, with one bit assigned for each block. If any updates occur to a particular block, the whole block is marked as changed, regardless of the size of the actual update. However, for resynchronization (or restore), only the changed blocks have to be copied, eliminating the need for a full synchronization (or restore) operation. This method reduces the time required for these operations considerably. In full-volume mirroring, the target is inaccessible for the duration of the synchronization process, until detachment from the source. For large databases, this can take a long time. Figure**:** Full-volume mirroring