MODULE 5 QB -PP

1) Explain the uses of Local Replicas.

Ans)

1) Alternative source for backup:

- The local replica contains an exact point-in-time (PIT) copy of the source data.
- Therefore can be used as a source to perform backup operations.
- This reduces the backup I/O workload on the production volumes.
- Another benefit of using local replica for backup is that it reduces the backup window to zero.

2) Fast recovery:

- If data loss or data corruption occurs on the source, a local replica might be used to recover the lost or corrupted data.
- If a complete failure of the source occurs some replication solution enables a replica to be used to restore data onto a different set of source devices, or production can be restarted on the replica.
- In either case, this method provides faster recovery and minimal RTO compared to traditional recovery from tape backups.

3) Decision-support activities, such as reporting or data warehousing:

- Running the reports using the data on the replicas greatly reduces the I/O burden placed on the production device.
- Local replicas are also used for data warehousing applications.
- The data warehouse application may be populated by the data on the replica.
- This avoids the impact on the production environment.

4) Data migration:

- Another use of local replica is data migration.
- Data migrations are performed for various reasons such as, migrating from a smaller capacity LUN to one of a larger capacity for newer versions of the application.

5) Testing platform:

- Local replicas are also used for testing new applications or upgrades.
- For example, an organization may use the replica to test the production application upgrade.
- If the test is successful, the upgrade may be implemented on the production environment.

2) Explain Local Replication Technology: Host based methods and Storage array based methods.

Ans)

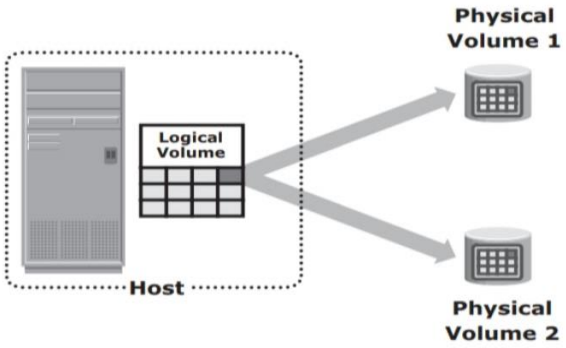
Host based method:

LVM-based replication and File system (FS) snapshot are two common methods of host-based local replication.

1) LVM-Based Replication:

- In LVM-based replication, the logical volume manager is responsible for creating and controlling the host-level logical volumes.
- An LVM has three components: physical volumes, volume groups and logical volumes.
  - A volume group is created by grouping one or more physical volumes.
  - Logical volumes are created within a given volume group.

- An application write to a logical volume is written to the two physical volumes by the LVM device driver. This is also known as LVM mirroring.



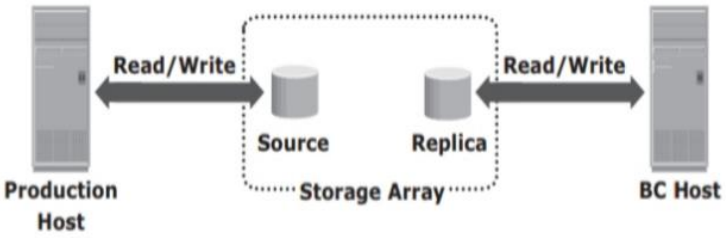
LVM-based mirroring

2) File System Snapshot:

- A file system (FS) snapshot is a pointer-based replica, which uses the Copy on First Write (CoFW) principle to create snapshots.
- When a snapshot is created, a bitmap and block map is 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 the snap creation.
- The blockmap is used to indicate the exact address from which the data is to be read when the data is accessed from the Snap FS.

Storage Array Based Local Replication:

- In this, the array-operating environment performs the local replication process.
- Required number of replica devices should be selected on the same array and then data should be replicated between the source-replica pairs.
- The source and target are in the same array and are accessed by different hosts.



1) Full Volume mirroring:

- Here, the target is attached to the source and established as a mirror of the source.
- Both the source and the target can be accessed for read and write operations by the production and business continuity hosts respectively.

2) Pointer-Based Full Volume Replication

- Similar to full-volume, this technology can provide full copies of the source data on the targets.
- It can be activated in either Copy on First Access (CoFA) mode or Full Copy mode.

3) Pointer-Based Virtual Replication:

- It is a technique used in data storage that allows multiple virtual copies of data to be created and managed efficiently without requiring the duplication of the entire dataset.
- This approach is commonly used in cloud storage environments where data needs to be accessed by multiple users simultaneously, or when data needs to be backed up.

3) Explain Remote replication technology: Storage Array based methods.

Ans)

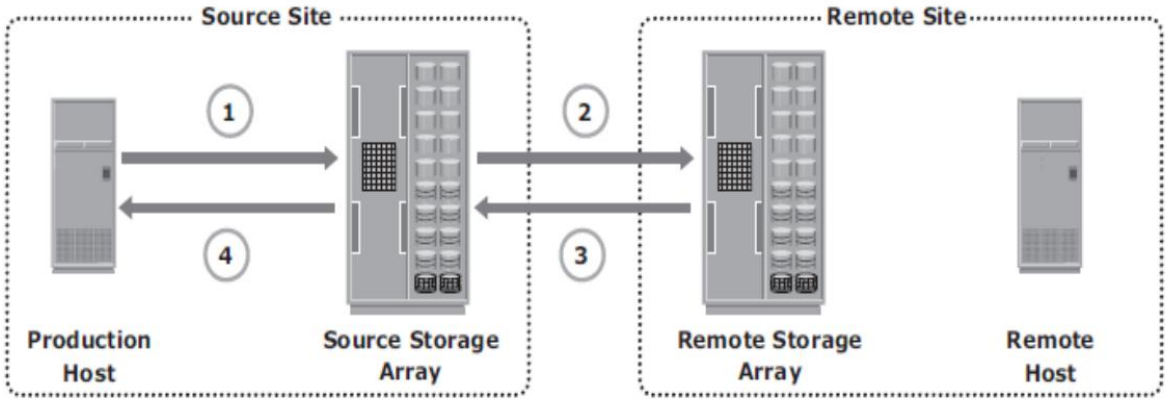
Storage Array based Remote Replication:

- Here, the array-operating environment and resources perform and manage data replication.
- A source and its replica device reside on different storage arrays.
- Data can be transmitted from the source storage array to the target storage array over a shared or a dedicated network.

Synchronous Replication Mode:

The following steps take place:

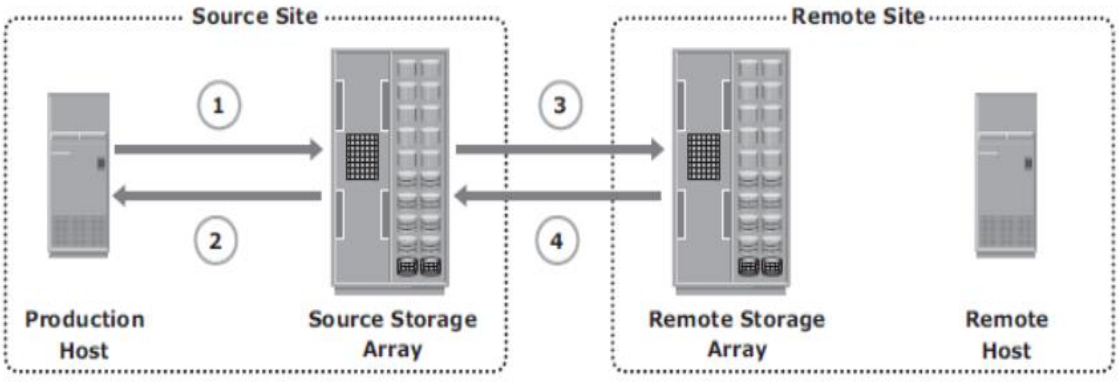
- The production host writes to the source storage array.
- Write is then transmitted to the remote storage array.
- Acknowledgement is sent to the source storage array by the remote storage array.
- Source storage array signals write-completion to the production host.



Asynchronous Replication Mode:

The following steps take place:

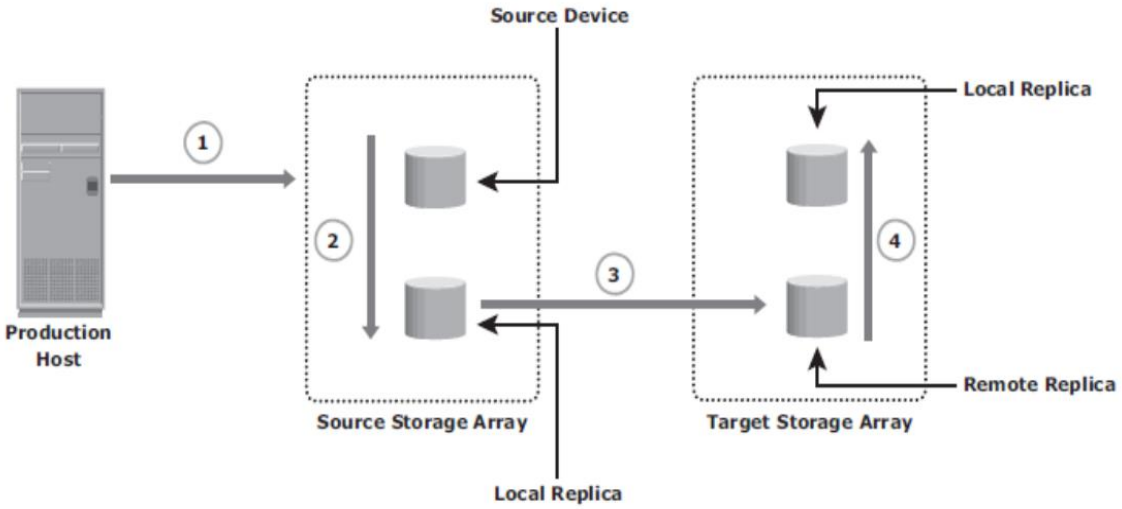
- The production host writes to the source storage array.
- The source array immediately acknowledges the production host.
- These writes are then transmitted to the target array.
- After the writes are received by the target array, it sends an acknowledgement to the source array.



Disk-Buffered Replication Mode:

The following steps take place:

- The production host writes data to the source device.
- A consistent PIT local replica of the source device is created.
- Data from the local replica in the source array is transmitted to its remote replica in the target array.
- Optionally, a local PIT replica of the remote device is created on the target array.



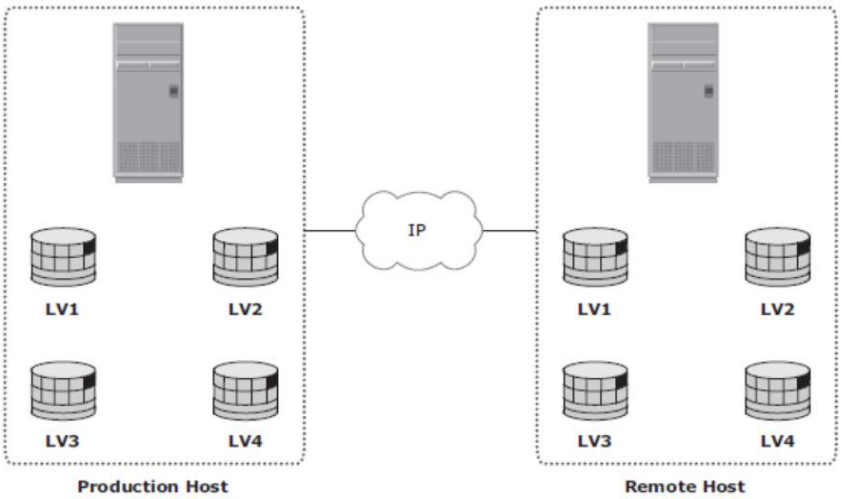
4) Explain Remote replication technology: Host-Based method.

Ans)

Host-Based method Remote Replication:

1) LVM Based Remote Replication:

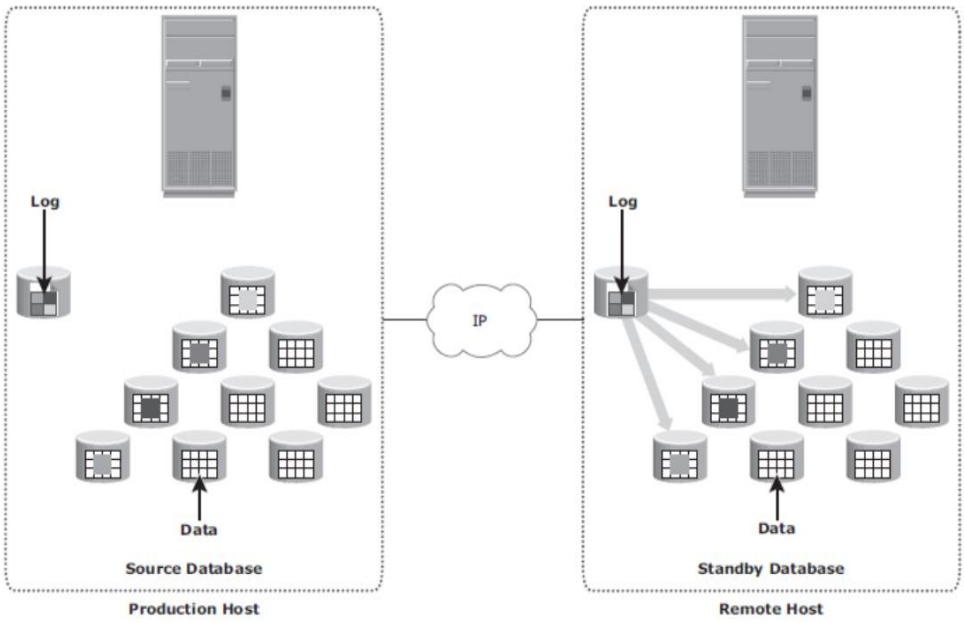
- LVM-based remote replication is performed and managed at the volume group level.
- Writes to the source volumes are transmitted to the remote host by the LVM.
- The LVM on the remote host receives the writes and commits them to the remote volume group.
- Prior to the start of replication, identical volume groups, logical volumes and file systems are created at the source and target sites.
- LVM-based remote replication supports both synchronous and asynchronous modes of replication.
- If a failure occurs at the source site, applications can be restarted on the remote host, using the data on the remote replicas.



2) Host Based Log Shipping:

- Database replication via log shipping is a host-based replication technology supported by most databases.
- Transactions to the source database are captured in logs, which are periodically transmitted by the source host to the remote host.
- The remote host receives the logs and applies them to the remote database.

- All DBMS's switch log files at preconfigured time intervals or when a log file is full.
- This process ensures that the standby database is consistent up to the last committed log.
- RPO at the remote site is finite and depends on the size of the log and the frequency.



5) Explain Risk Triads: Assets, Threats and Vulnerability.

Ans) Risk triad defines risk in terms of assets, threats and vulnerabilities.

1) Assets:

- Information is one of the most important assets for any organization.
- Other assets include hardware, software and other infrastructure components required to access the information.
- To protect these assets, organizations must develop a set of parameters to ensure the availability of the resources to authorized users and trusted networks.
- Security methods have two objectives.
  - The first objective is to ensure that the network is easily accessible to authorized users.
  - The second objective is to make it difficult for potential attackers to access and compromise the system.
- The security method must ensure that updates to the operating system and other software are installed regularly.

2) Security Threats:

- Threats are the potential attacks that can be carried out on an IT infrastructure.
- Attacks can be classified as active or passive.

  1) **Passive attacks** are attempts to gain unauthorized access into the system.
  They pose threats to confidentiality of information.

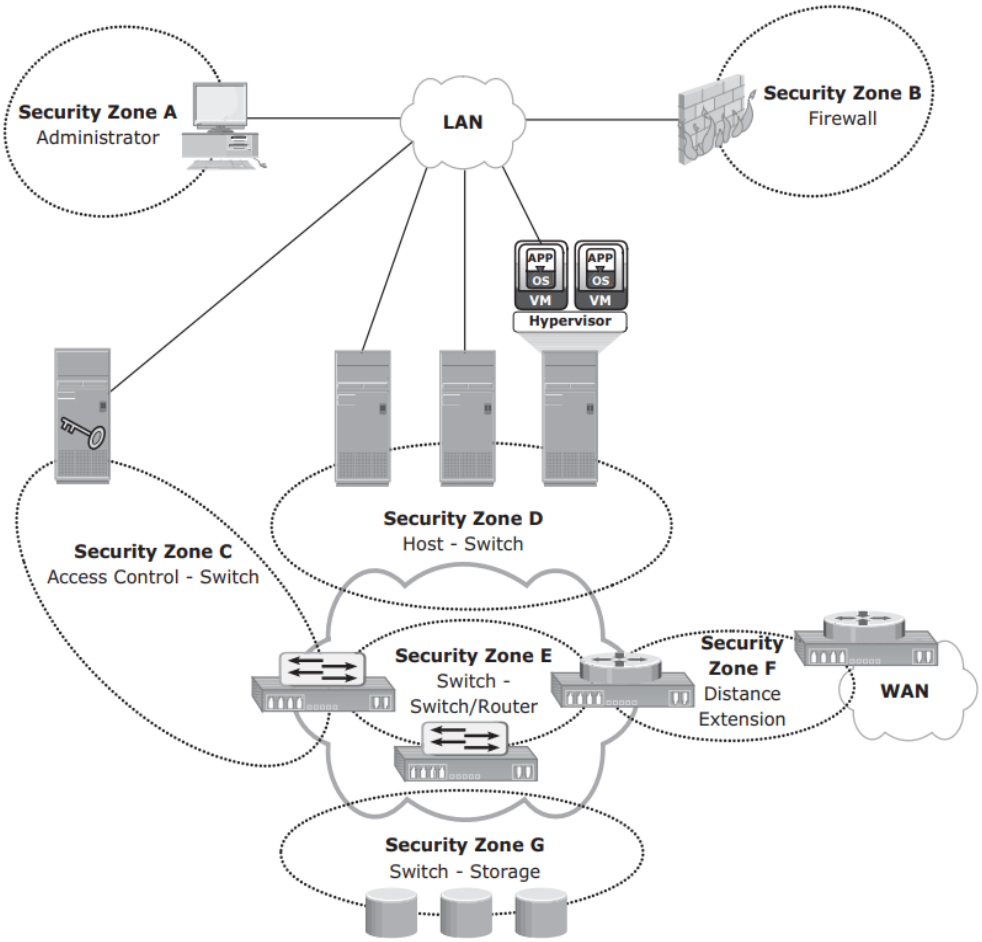
  2) **Active attacks** include data modification, Denial of Service (DoS) and repudiation attacks.

3) Vulnerabilities:

- The paths that provide access to information are often vulnerable to potential attacks.
- Each of the paths may contain various access points, which provide different levels of access to the storage resources.
- Attack surface, attack vector and work factor are the three important factors to be considered.
  - An **Attack surface** refers to the various entry points that an attacker can use to launch an attack.
  - An **Attack vector** is a step or a series of steps necessary to complete an attack.
  - **Work factor** refers to the amount of time and effort required to exploit an attack vector.

6) Explain FC San security architecture.

Ans)

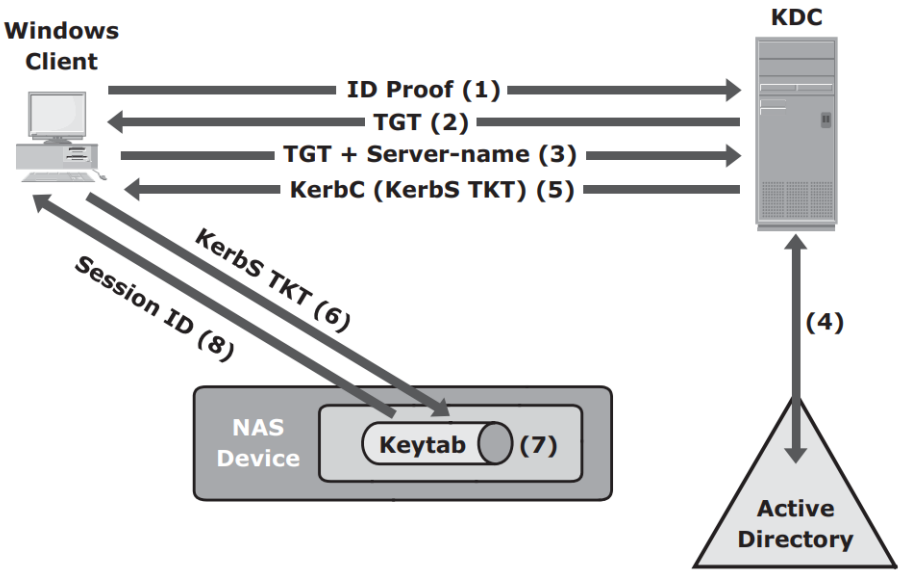


Security Zones	Protection Strategies
Zone A	• Restrict management LAN access to authorized users. • Implement VPN tunneling for secure remote access to the management LAN. • Use two-factor authentication for network access.
Zone B	• Block inappropriate traffic by filtering out addresses that should not be allowed on your LAN. • Block ports that are not in use.
Zone C	Authenticate users/administrators of FC switches using Remote Authentication Dial In User Service (RADIUS), DH-CHAP (Diffie-Hellman Challenge Handshake Authentication Protocol), and so on.
Zone D	• Restrict Fabric access to legitimate hosts by implementing ACLs. • Implementing a secure zoning method, such as port zoning (also known as hard zoning)
Zone E	Protect traffic on fabric by • using E-Port authentication • encrypting the traffic in transit • implementing FC switch controls and port controls.
Zone F	Implement encryption for in-flight data • FC-SP for long-distance FC extension • IP-Sec for SAN extension via FCIP
Zone G	Protect the storage arrays on your SAN via • WWPN-based LUN masking • S_ID locking: masking based on source FC address.

7) Explain Kerberos - network authentication protocol.

Ans) Kerberos is a network authentication protocol, which is designed to provide strong authentication for client/server applications by using secret-key cryptography.

- It uses cryptography so that a client and server can prove their identity to each other across an insecure network connection.
- In Kerberos, authentications occur between clients and servers.



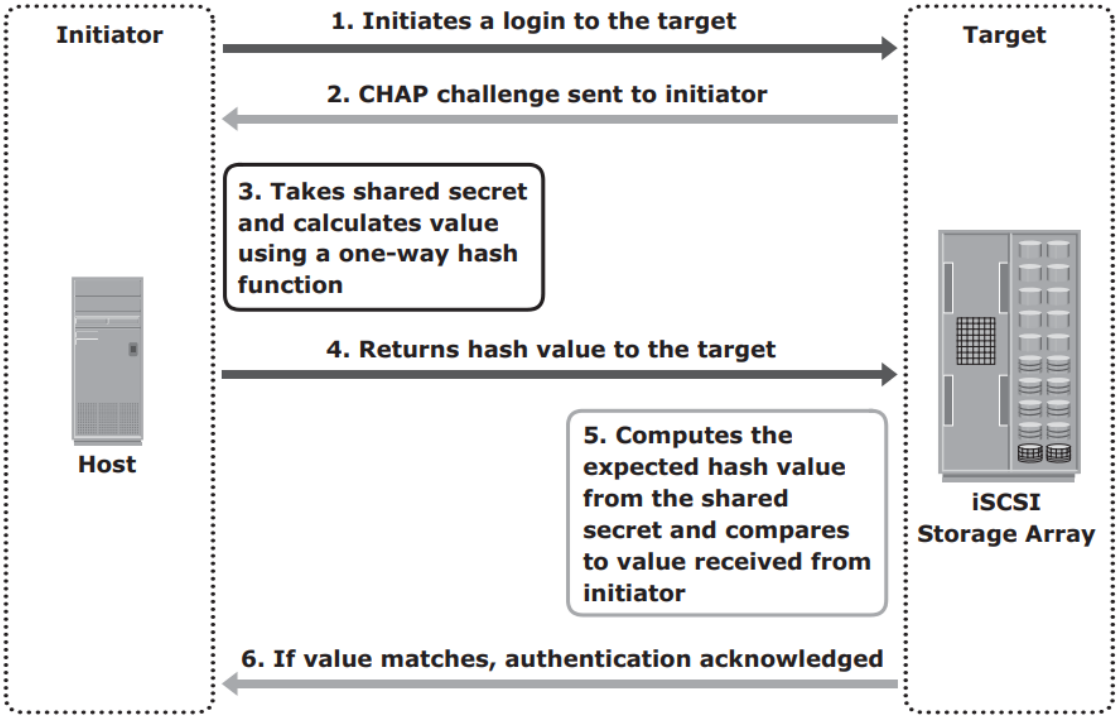
- The user logs on to the workstation in the Active Directory domain using an ID and a password.
- The client computer sends a request to the AS running on the KDC for a Kerberos ticket.
- The KDC verifies the user's login information from Active Directory.
- The KDC responds with an encrypted Ticket Granting Ticket (TGT) and an encrypted session key.
- When the client requests a service from a server, it sends a request, consisting of the previously generated TGT, encrypted with the session key and the resource information to the KDC.
- The KDC checks the permissions in Active Directory and ensures that the user is authorized to use that service.
- The KDC returns a service ticket to the client.
- The client then sends the service ticket to the server that houses the required resources.
- The server, in this case the NAS device, decrypts the server portion of the ticket and stores the information in a key tab file.
- As long as the client's Kerberos ticket is valid, this authorization process does not need to be repeated.
- The server automatically allows the client to access the appropriate resources.
- A client-server session is now established.

8) Explain Authentication mechanism in IP SAN.

Ans) The Challenge-Handshake Authentication Protocol (CHAP) is a basic authentication mechanism that has been widely adopted by network devices and hosts.

- CHAP provides a method for initiators and targets to authenticate each other by using a secret code or password.
- CHAP secrets are usually random secrets of 12 to 128 characters.
- The secret is never exchanged directly over the communication channel.
- Rather, a one-way hash function converts it into a hash value, which is then exchanged.
- A hash function, using the MD5 algorithm, transforms data in such a way that the result is unique and cannot be changed back to its original form.
- If the initiator requires reverse CHAP authentication, the initiator authenticates the target in same way.

- The CHAP secret must be configured on the initiator and the target.
- Both the initiator and target store a CHAP entry that consists of a node's name and its associated secret.
- The same steps are executed in a two-way CHAP authentication scenario.

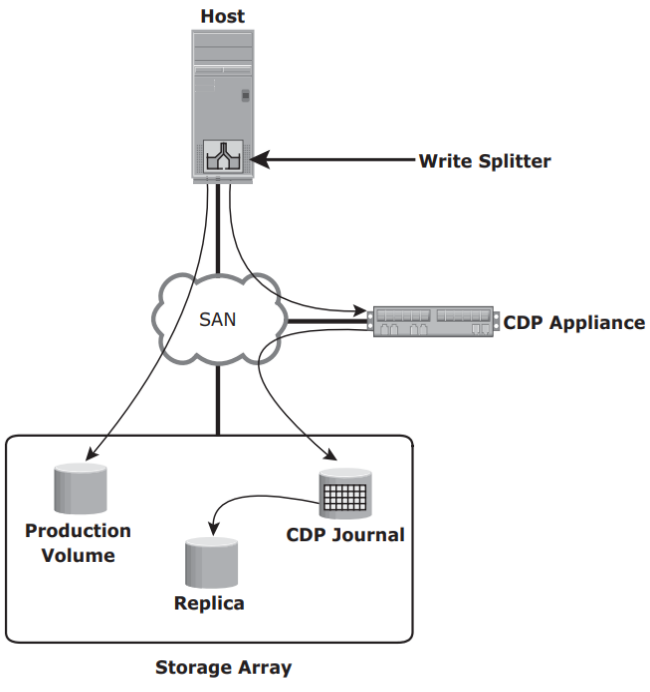


9) Explain Network-Based Local Replication.

Ans) In network-based replication, the replication occurs at the network layer between the hosts and storage arrays.

Continuous data protection:

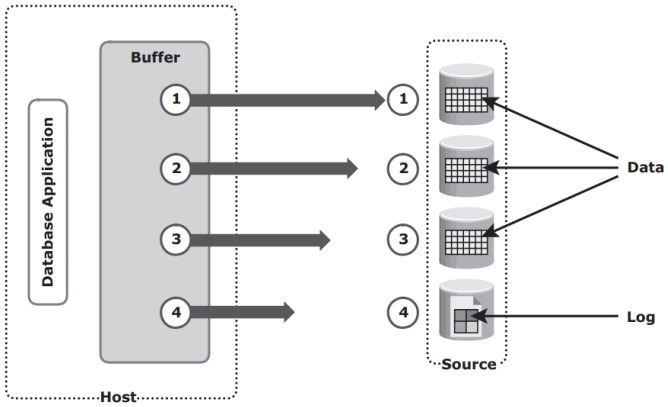
- Continuous data protection (CDP) is a technology used for network-based local and remote replications.
- In CDP, data changes are continuously captured and stored in a separate location from the primary storage.
- With CDP, recovery from data corruption poses no problem because it allows going back to a PIT image prior to the data corruption incident.
- CDP uses a journal volume to store all data changes on the primary storage.
- CDP is implemented using CDP appliance and write splitters.
- CDP appliance is an intelligent hardware platform that runs the CDP software.
- CDP appliance manages local and remote data replications.
- Write splitters intercept writes to the production volume from the host and split each write into two copies.



10) Explain Consistency of a Replicated Database.

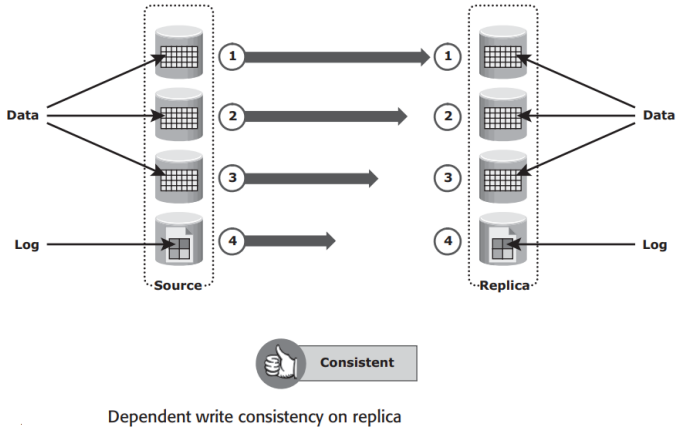
Ans)

- A database may be spread over numerous files, file systems and devices.
- All of these must be replicated consistently to ensure that the replica is restorable and restartable.
- Replication is performed with the database offline or online.
- If the database is offline during the creation of the replica, it is not available for I/O operations.
- If the database is online, it is available for I/O operations and transactions to the database update the data continuously.
- When a database is replicated while it is online, changes made to the database at this time must be applied to the replica to make it consistent.
- Many applications and DBMS use a dependent write I/O principle to maintain consistency.
- Databases need writes to happen in a specific order for a transaction to be considered complete.
- These writes will be recorded on the various devices or file systems.
- Figure below shows the process of flushing the buffer from the host to the source.



Dependent write consistency on sources

- When the replica is created, all the writes to the source devices must be captured on the replica devices to ensure data consistency.
- Figure below shows the process of replication from the source to the replica.



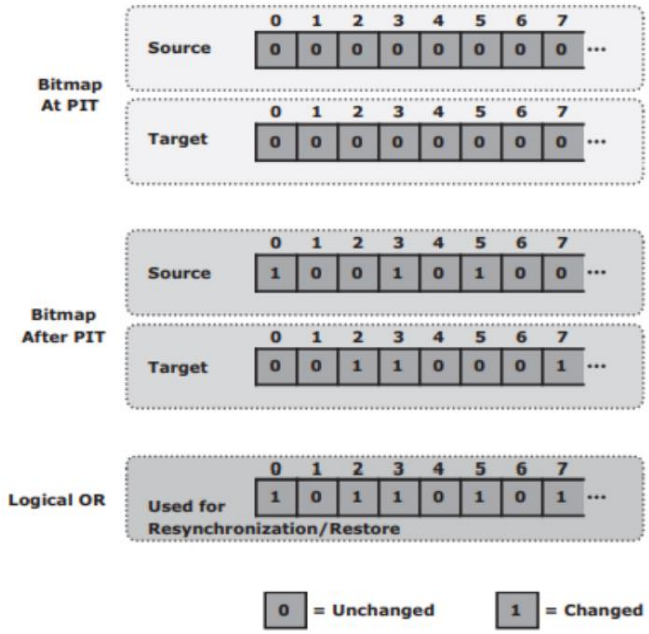
Dependent write consistency on replica

11) Explain Tracking Changes to Source and Replica.

Ans)

- Updates can occur on the source device after the creation of PIT local replicas.
- If the primary purpose of local replication is to have a viable PIT copy for data recovery or restore operations, then the replica devices should not be modified.
- To enable incremental resynchronization or restore operations, changes to both the source and replica devices after the PIT should be tracked.

- Resynchronization is enabled using bitmaps, where each bit represents a block of data.
- The data block sizes can range from 512 bytes to 64 KB or greater.
- For example, if the block size is 32 KB, then a 1-GB device would require 32,768 bits.
- The bits in the source and target bitmaps are all set to 0 (zero) when the replica is created.
- Any changes to the source or replica are then flagged by setting the appropriate bits to 1 in the bitmap.
- When resynchronization or restore is required, a logical OR operation between the source bitmap and the target bitmap is performed.
- If resynchronization is required, changes to the replica are overwritten with the corresponding blocks from the source.
- In this example, that would be blocks labelled 2, 3, and 7 on the replica.
- If a restore is required, changes to the source are overwritten with the corresponding blocks from the replica.
- In this example, that would be blocks labeled 0, 3 and 5 on the source.



12) List and explain the basic security goals of information security framework.

Ans) The basic security goals of information security framework are:

1) Confidentiality
2) Integrity
3) Availability
4) Accountability service

1) Confidentiality:

- Provides the required secrecy of information and ensures that only authorized users have access to data.
- This requires authentication of users who need to access information.

2) Integrity:

- Ensures that the information is unaltered.
- Ensuring integrity requires detection of and protection against unauthorized alteration or deletion of information.
- Ensuring integrity stipulates measures such as error detection and correction for both data and systems.

3) Availability:

- This ensures that authorized users have reliable and timely access to systems, data and applications residing on these systems.
- Availability requires protection against unauthorized deletion of data and Denial of service.
- Availability also implies that sufficient resources are available to provide a service.

4) Accountability service:

- Refers to accounting for all the events and operations that take place in the data center infrastructure.
- It maintains a log of events that can be audited or traced later for the purpose of security.

13) Explain Security threats in a backup, replication and archive environment.

Ans)

- Backup, replication and archive is the third domain that needs to be secured against an attack.
- A backup involves copying the data from a storage array to backup media, such as tapes or disks.
- Securing backup is complex and is based on the backup software that accesses the storage arrays.
- It also depends on the configuration of the storage environments at the primary and secondary sites, especially with remote backup solutions performed directly on a remote tape device or using array-based remote replication.
- Organizations must ensure that the disaster recovery (DR) site maintains the same level of security for the backed up data.
- Protecting the backup, replication and archive infrastructure requires addressing several threats, including spoofing the legitimate identity of a DR site, tampering with data, network snooping, DoS attacks and media theft.
- Such threats represent potential violations of integrity, confidentiality and availability.
- The physical threat of a backup tape being lost, stolen, or misplaced, especially if the tapes contain highly confidential information, is another type of threat.
- Backup-to-tape applications are vulnerable to severe security implications if they do not encrypt data while backing it up.