CMR

INSTITUTE OF

TECHNOLOGY

Internal Assessment Test 1 – Sept. 2016 Scheme and solution

Sub:	STORAGE AREA NETWORKS						Code:	10IS765
Date:	8/ 09 / 2016	Duration:	90 mins	Max Marks:	50	Sem: VII A & B	Branch:	ISE/CSE

Note: ***Answer any five questions.***

1 Explain FC connectivity with related diagrams (6m)

The FC architecture supports three basic interconnectivity options: point-topoint, arbitrated loop (FC-AL), and fabric connect.

Point-to-Point

Point-to-point is the simplest FC configuration — two devices are connected directly to each other, as shown in Figure 6-6. This configuration provides a dedicated connection for data transmission between nodes. However, the point-to-point configuration offers limited connectivity, as only two devices can communicate with each other at a given time. Moreover, it cannot be scaled to accommodate a large number of network devices. Standard DAS usess pointto-point connectivity.

Fibre Channel Arbitrated Loop

In the FC-AL configuration, devices are attached to a shared loop, as shown in Figure 6-7. FC-AL has the characteristics of a token ring topology and a physical star topology. In FC-AL, each device contends with other devices to perform I/O operations. Devices on the loop must "arbitrate" to gain control of the loop. At any given time, only one device can perform I/O operations on the loop.As a loop configuration, FC-AL can be implemented without any interconnectingdevices by directly connecting one device to another in a ring through cables

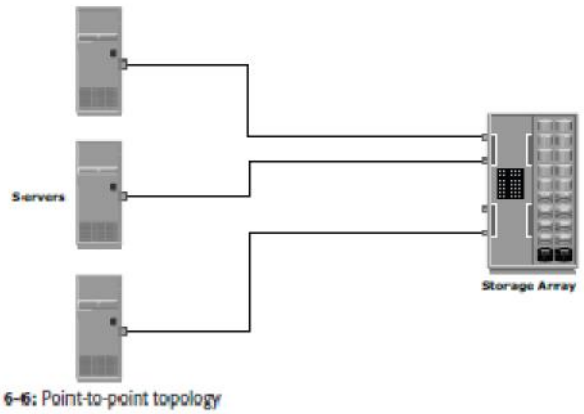


Figure 6-6: Point-to-point topology

However, FC-AL implementations may also use hubs whereby the arbitrated loop is physically connected in a star topology. The FC-AL configuration has the following limitations in terms of scalability:

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

an FC-AL topology.

■■ FC-AL uses 8-bit addressing. It can support up to 127 devices on a loop.
■■Adding or removing a device results in loop re-initialization, which can cause a momentary pause in loop traffic.

FC-AL Transmission

When a node in the FC-AL topology attempts to transmit data, the node sends an *arbitration (ARB)* frame to each node on the loop. If two nodes simultaneously attempt to gain control of the loop, the node with the highest priority is allowed to communicate with another node. This priority is determined on the basis of Arbitrated Loop Physical Address (AL-PA) and Loop ID,

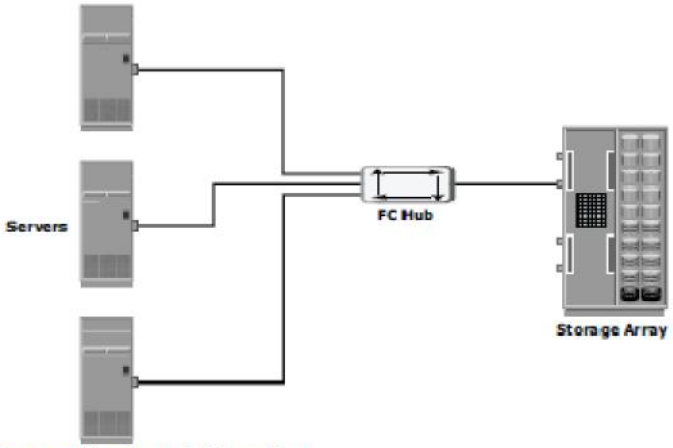
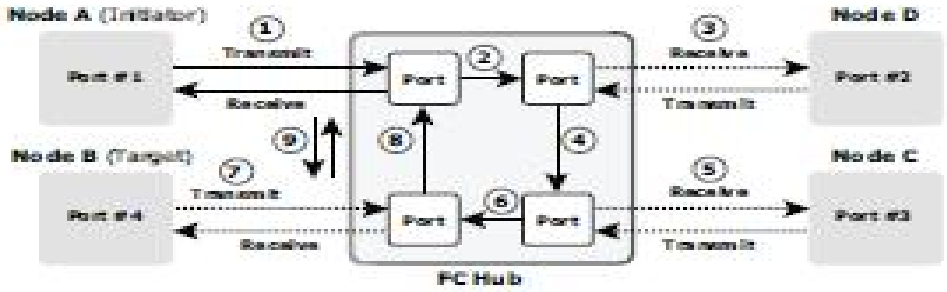


Figure 6-7: Fibre Channel arbitrated loop

When the initiator node receives the ARB request it sent, it gains control of the loop. The initiator then transmits data to the node with which it has established a virtual connection. Figure 6-8 illustrates the process of data transmission in an FC-AL configuration.



Node A want to communicate with Node B

① High priority initiator, Node A inserts the ARB frame in the loop.
② ARB frame is passed to the next node (Node D) in the loop.
③ Node D receives high priority ARB, therefore remains idle.
④ ARB is forwarded to next node (Node C) in the loop.
⑤ Node C receives high priority ARB, therefore remains idle.
⑥ ARB is forwarded to next node (Node B) in the loop.
⑦ Node B receives high priority ARB, therefore remains idle and
⑧ ARB is forwarded to next node (Node A) in the loop.
⑨ Node A receives ARB back; now it gains control of the loop and can start communicating with target Node B.

Figure 6–8: Data transmission in FC-AL

Fibre Channel Switched Fabric

Unlike a loop configuration, a Fibre Channel switched fabric (FC-SW) network provides interconnected devices, dedicated bandwidth, and scalability. The addition or removal of a device in a switched fabric is minimally disruptive; it does not affect the ongoing traffic between other devices. FC-SW is also referred to as *fabric connect*. A fabric is a logical space in which all nodes communicate with one another in a network. This virtual space can be created with a switch or a network of switches. Each switch in a fabric contains
a unique domain identifier, which is part of the fabric's addressing scheme. In FC-SW, nodes do not share a loop; instead, data is transferred through a dedicated path between the nodes. Each port in a fabric has a unique 24-bit fibre channel address for communication. Figure 6-9 shows an example of FC-SW.
A fabric topology can be described by the number of tiers it contains. The number of tiers in a fabric is based on the number of switches traversed between two points that are farthest from each other. However, note that this number is based on the infrastructure constructed by the fabric topology; it disregards how the storage and server are connected across the switches.

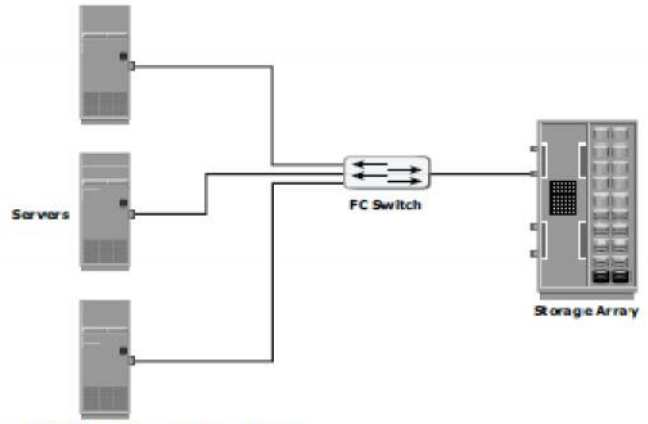


Figure 6–9: Fibre Channel switched fabric

When the number of tiers in a fabric increases, the distance that a fabric management message must travel to reach each switch in the fabric also increases. The increase in the distance also increases the time taken to propagate and complete a fabric reconfiguration event, such as the addition of a new switch, or a zone set propagation event (detailed later in this chapter). Figure 6-10 illustrates two-tier and three-tier fabric architecture.

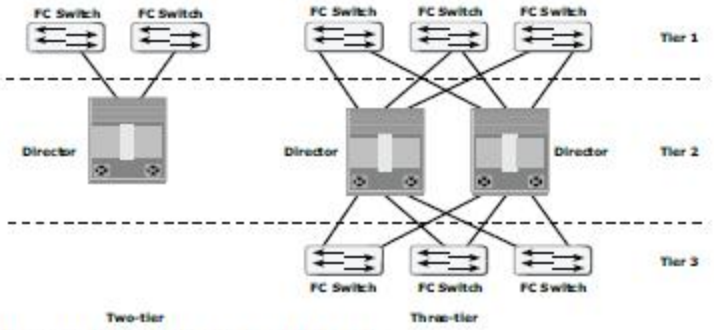


Figure 6-10: Tiered structure of FC-SW topology

FC-SW Transmission

FC-SW uses switches that are intelligent devices. They can switch data traffic from an initiator node to a target node directly through switch ports. Frames are routed between source and destination by the fabric.

As shown in Figure 6-11, if node B wants to communicate with node D, Nodes should individually login first and then transmit data via the FC-SW. This link is considered a dedicated connection between the initiator and the target.

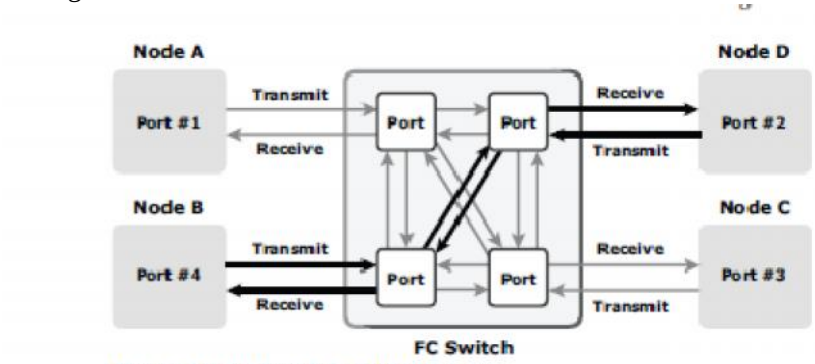


Figure 6-11: Data tansmission in FC-SW topology

2.a.Explain SCSI-3 architecture.(5m)

SCSI-3 Architecture

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:

■■ **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.

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

■■ **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.

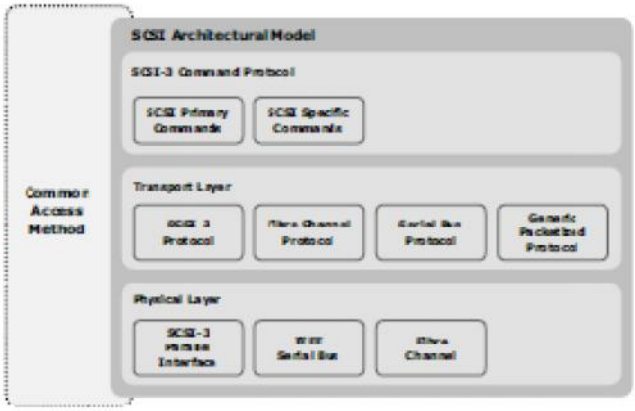
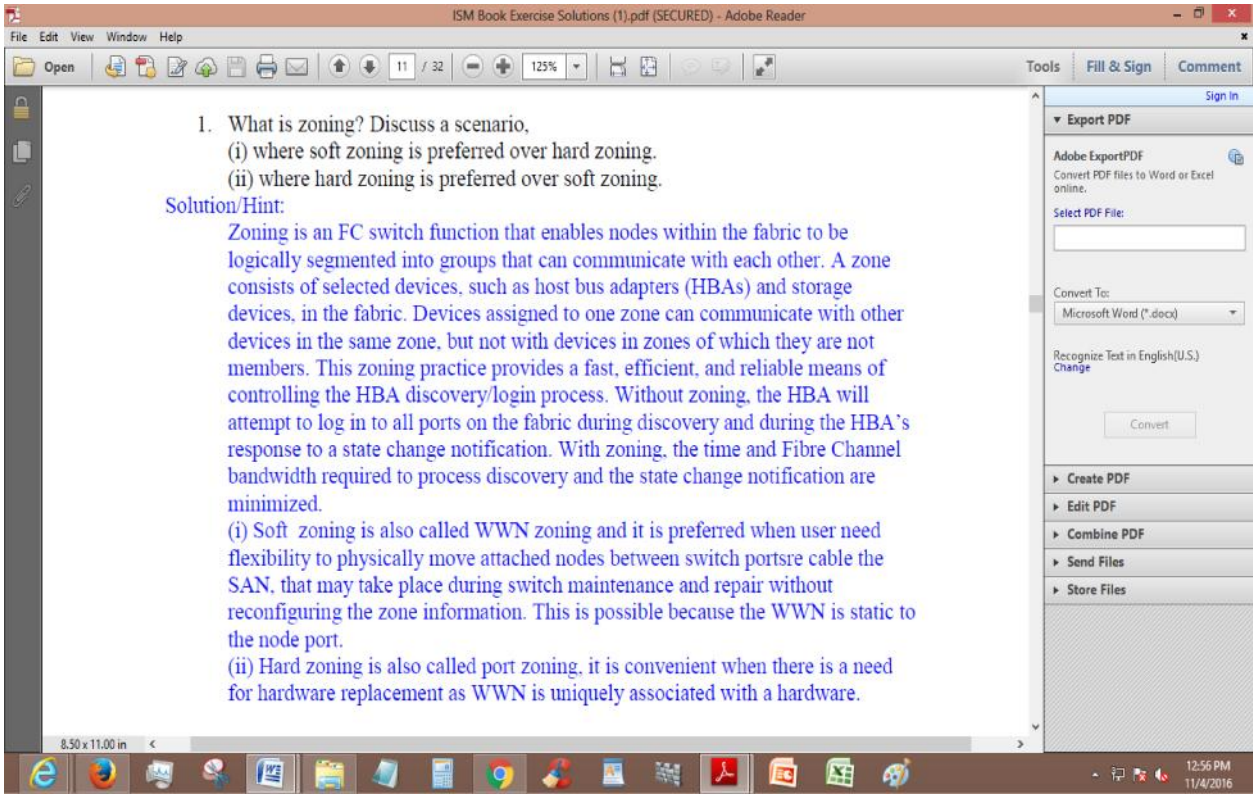


Figure 5-3: SCSI-3 standards architecture

b.What is zoning? Discuss a scenario where soft zoning is preferred & where hard zoning is preferred.(5M)



1. What is zoning? Discuss a scenario,
   (i) where soft zoning is preferred over hard zoning.
   (ii) where hard zoning is preferred over soft zoning.

Solution/Hint:

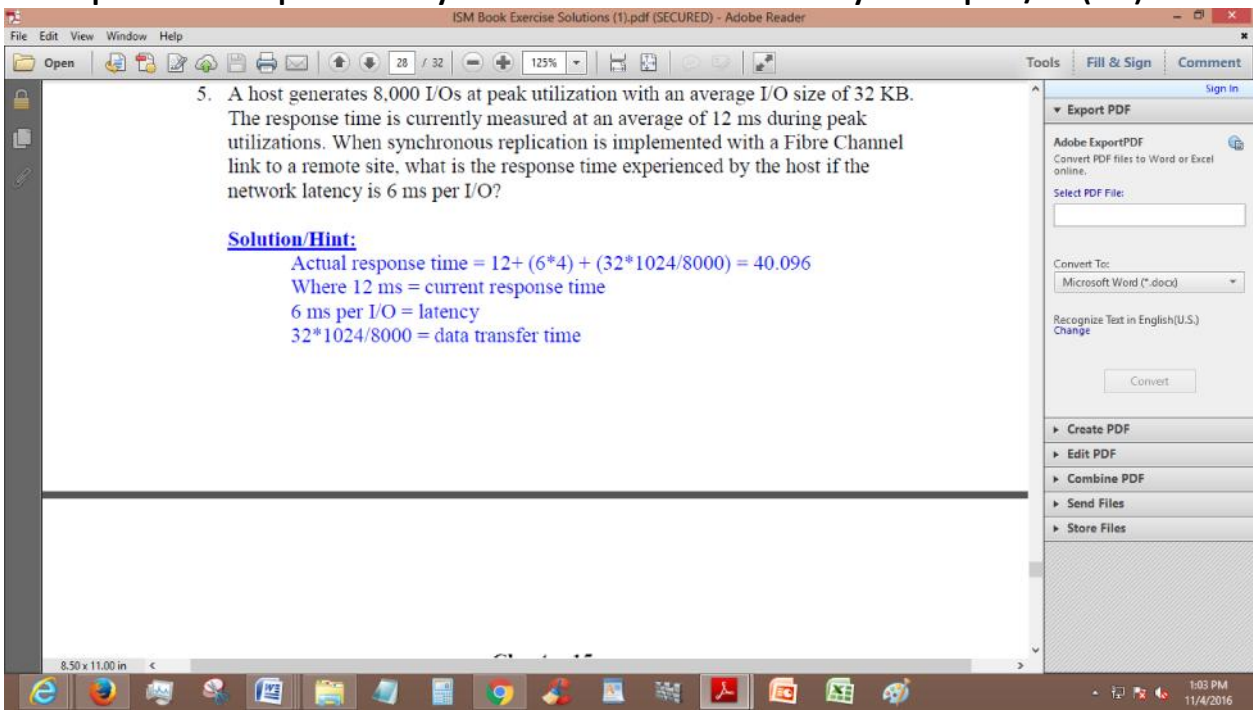
Zoning is an FC switch function that enables nodes within the fabric to be logically segmented into groups that can communicate with each other. A zone consists of selected devices, such as host bus adapters (HBAs) and storage devices, in the fabric. Devices assigned to one zone can communicate with other devices in the same zone, but not with devices in zones of which they are not members. This zoning practice provides a fast, efficient, and reliable means of controlling the HBA discovery/login process. Without zoning, the HBA will attempt to log in to all ports on the fabric during discovery and during the HBA's response to a state change notification. With zoning, the time and Fibre Channel bandwidth required to process discovery and the state change notification are minimized.

(i) Soft zoning is also called WWN zoning and it is preferred when user need flexibility to physically move attached nodes between switch portsre cable the SAN, that may take place during switch maintenance and repair without reconfiguring the zone information. This is possible because the WWN is static to the node port.

(ii) Hard zoning is also called port zoning, it is convenient when there is a need for hardware replacement as WWN is uniquely associated with a hardware.

3. a.A host generates 8000 I/Os at peak utilization with an average I/O size of 32KB.The response time is currently measured at an average of 12 ms during peak utilizations. When

synchronous replication is implemented with a fiber channel link to its remote site, what is the response time experienced by the host if the network latency is 6ms per I/O? (5m)



b.Write about Monitoring Components in storage infrastructure in detail (5m)

Hosts, networks, and storage are components within the storage environment that should be monitored for accessibility, capacity, performance, and security.

Hosts

Mission-critical application hosts should be monitored continuously. The accessibility of a host depends on the status of the hardware components and software processes running on it. For example, an application crash due to host hardware failure can cause instant unavailability of the data to the user Servers are used in a cluster to ensure high availability. In a server virtualization environment, multiple virtual machines share a pool of resources. These resources are dynamically reallocated, which ensures application accessibility and ease of management. File system utilization of hosts also needs to be monitored. Monitoring helps in estimating the file system's growth rate and helps in predicting when it will reach 100 percent. Accordingly, the administrator can extend (manually or automatically) the file system's space proactively to prevent a failure resulting from a file system being full. New provisioning technologies even enable the allocation of storage on demand as the need arises. Alternatively, system administrators can enforce a quota for users, provisioning a fixed amount of space for their files. For example, a quota could be specified at a user level, restricting the maximum space to 10 GB per user, or at a file level that restricts a file to a maximum of 100 MB. Server performance mainly depends on I/O profile, utilization of CPU and memory. For example, if a server running an application is experiencing 80 percent of CPU utilization continuously, this suggests that the server may be running out of processing power, which can lead to degraded performance and slower response time. Administrators can take several actions to correct the problem, such as upgrading or adding more processors, shifting the workload to different servers, and restricting the number of simultaneous client access. In a virtualized environment, CPU and memory may be allocated dynamically from another physical server or from the same server.

Memory utilization is measured by the amount of free memory available. Databases, applications, and file systems utilize the server's physical memory(RAM) for data manipulation. Insufficient memory leads to excessive swapping and paging on the disk, which in turn affects response time to the applications.

Security monitoring on servers involves tracking of login failures and execution of unauthorized applications or software processes. Proactive measures against unauthorized access to the servers are based on the threat

identified. For example, an administrator can block access to an unauthorized user if multiple login failures are logged.

Storage Network

The storage network needs to be monitored to ensure proper communication between the server and the storage array. Uninterrupted access to data over the storage network depends on the accessibility of the physical and logical components in the storage network. The physical components of a storage network include elements such as switches, ports, cables, GBICs, and power supplies. The logical components include constructs, such as zones and fabrics. Any failure in the physical or logical components may cause data unavailability. For example errors in zoning such as specifying the wrong WWN of a port results in failure to access that port, which potentially prevents access from a host to its storage. Capacity monitoring in a storage network involves monitoring the availability of ports on a switch, the number of available ports in the entire fabric, the utilization of the inter-switch links, individual ports, and each interconnect device in the fabric. Capacity monitoring provides all required inputs for future planning and optimization of the fabric with additional interconnect devices. Monitoring the performance of a storage network is useful in assessing individual component performance and helps to identify network bottlenecks. For example, monitoring port performance is done by measuring receive ortransmit link utilization metrics, which indicate how busy the switch port is, based on expected maximum throughput. Heavily used ports can cause queuing delays on the server.

For IP networks, monitoring performance includes monitoring network latency,packet loss, bandwidth utilization for I/O, network errors, and collisions. Storage network security monitoring provides information for any unauthorized change to the configuration of the fabric—for example, changes to the zone policies that can affect data security. Login failures and unauthorized access to switches for performing administrative changes should be logged and monitored continuously.

Storage

The accessibility of the storage array should be monitored for its hardware components and various processes. Storage arrays configured with redundant components do not affect accessibility in the event of an individual component failure, but failure of any process can disrupt or compromise business continuity operations. For example, the failure of a replication task affects disaster recovery capabilities. Some storage arrays also provide the capability to send a message to the vendor's support center in the event of hardware or process failures, referred to as a *call home*. Capacity monitoring of a storage array enables the administrator to respond to storage needs as they occur. Information about fan-in or fan-out ratios and the availability of front-end ports is useful when a new server is given access to the storage array. A storage array can be monitored by a number of performance metrics, such as utilization rates of the various storage array components, I/O response time, and cache utilization. A high utilization rate of storage array components may lead to performance degradation.A storage array is usually a shared resource, which may be exposed to security breaches. Monitoring security helps to track unauthorized configuration of the storage array or corruption of data and ensures that only authorized users are allowed to access it.

4.Explain local replication technologies in detail(10m)

Host-based and storage-based replications are the two major technologies adopted for local replication. File system replication and LVM-based replication are examples of host-based local replication technology. Storage array–based replication can be implemented with distinct solutions namely, full-volume mirroring, pointerbased full-volume replication, and pointer-based virtual replication.

Host-Based Local Replication

In host-based replication, logical volume managers (LVMs) or the file systems perform the local replication process. LVM-based replication and file system (FS) snapshot are examples of host-based local replication.

LVM-Based Replication

In LVM-based replication, logical volume manager is responsible for creating and controlling the host-level logical volume. An 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, as shown in Figure 13-4. 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.

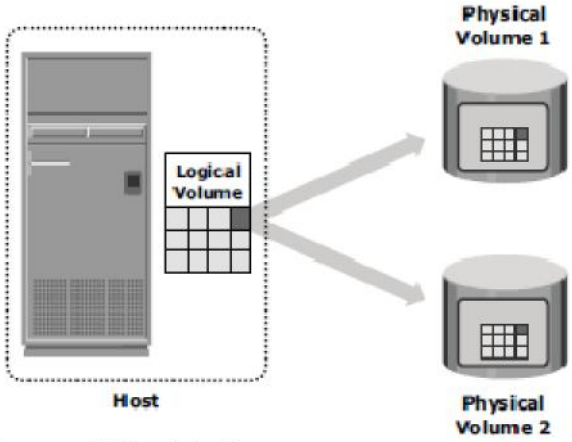


Figure 13-4: LVM-based mirroring

Advantages of LVM-Based Replication

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

Limitations of LVM-Based Replication

As every write generated by an application translates into two writes on the disk, an additional burden is placed on the host CPU. This can degrade application performance. Presenting an LVM-based local replica to a second host is usually not possible because the replica will still be part of the volume group, which is usually accessed by one host at any given time. Tracking changes to the mirrors and performing incremental synchronization operations is also a challenge as all LVMs do not support incremental resynchronization. If the devices are already protected by some level of RAID on the array, then the additional protection provided by mirroring is unnecessary.

This solution does not scale to provide replicas of federated databases and applications. Both the replica and the source are stored within the same volume group. Therefore, the replica itself may become unavailable if there is an error in the volume group. If the server fails, both source and replica are unavailable until the server is brought back online.

File System Snapshot

File system (FS) snapshot is a pointer-based replica that requires a fraction of the space used by the original FS. This snapshot can be implemented by either FS itself or by LVM. It uses Copy on First Write (CoFW) principle.

When the snapshot is created, a bitmap and a blockmap 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 blockmap is usedto 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 blockmap and data is read from that address. Reads from the production FS work as normal.

Storage Array–Based Replication

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 13-5 shows storage array based local replication, where source and target are in the same array and accessed by different hosts.

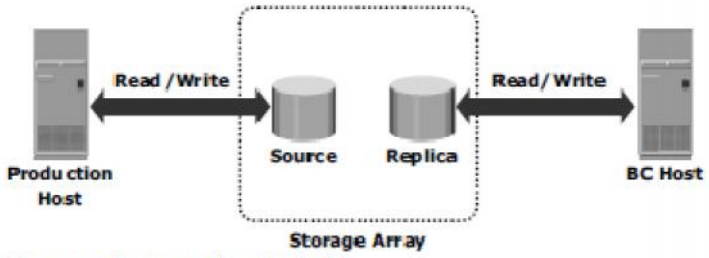


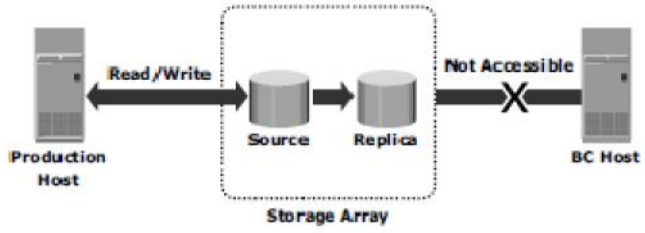
Figure 13-5: Storage array-based replication

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.

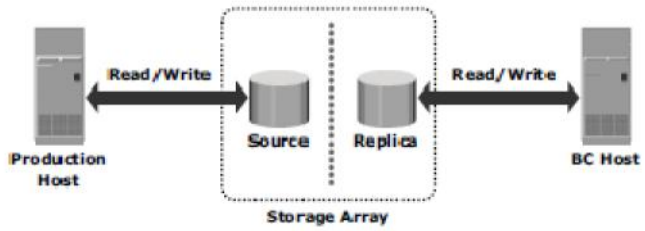
Full-Volume Mirroring

In *full-volume mirroring*, the target is attached to the source and established as a mirror of the source (Figure 13-6 [a]). 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. Figure 13-6 (b) shows full-volume mirroring when the target is detached from the source. Notice that both the source and the target can be accessed for read and write operations by the production hosts

.



(a) Full volume mirroring with source attached to replica



(b) Full volume mirroring with source detached from replica

Figure 13-6: Full-volume mirroring

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. *Pointer-Based, Full-Volume Replication*

An alternative to full-volume mirroring is *pointer-based full-volume replication*. Like full-volume mirroring, this technology can provide full copies of the source data on the targets. Unlike full-volume mirroring, the target is made immediately available at the activation of the replication session. Hence, one need not wait for data synchronization to, and detachment of, the target in order to access it. The time of activation defines the PIT copy of source. Pointer-based, full-volume replication can be activated in either Copy on First Access (CoFA) mode or Full Copy mode. In either case, at the time of activation, a protection bitmap is created for all data on the source devices. Pointers are initialized to map the (currently) empty data blocks on the target to the corresponding original data blocks on the source. The granularity can range from 512 byte blocks to 64 KB blocks or higher. Data is then copied from the source to the target, based on the mode of activation. In CoFA, after the replication session is initiated, data is copied from the source to the target when the following occurs:

A write operation is issued to a specific address on the ■■ source for the first time (see Figure 13-7).

■■A read or write operation is issued to a specific address on the target for the first time (see Figure 13-8 and Figure 13-9) .

When a write is issued to the source for the first time after session activation, original data at that address is copied to the target. After this operation, the new data is updated on the source. This ensures that original data at the point-in-time of activation is preserved on the target. This is illustrated in Figure 13-7.

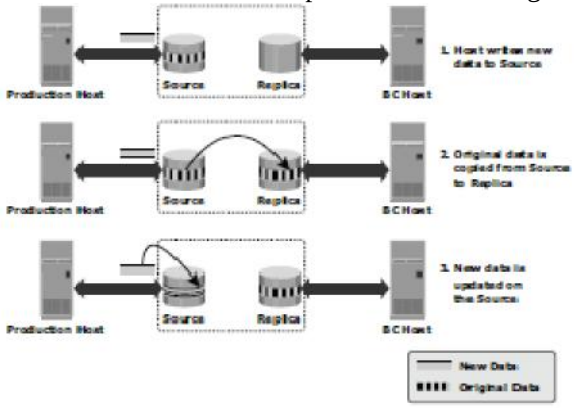


Figure 13-7: Copy on first access (CoFA) – write to source

When a read is issued to the target for the first time after session activation, the original data is copied from the source to the target and is made available to the host. This is illustrated in Figure 13-8.

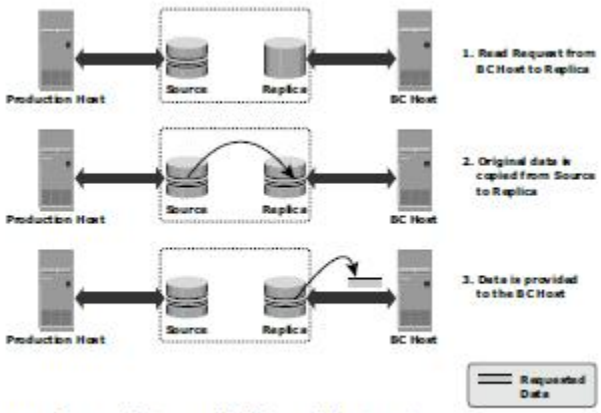


Figure 13-8: Copy on first access (CoFA) – read from target

When a write is issued to the target for the first time after session activation, the original data is copied from the source to the target. After this, the new data is updated on the target. This is illustrated in Figure 13-9.
In all cases, the protection bit for that block is reset to indicate that the original data has been copied over to the target. The pointer to the source data can now be discarded. Subsequent writes to the same data block on the source, and reads or writes to the same data blocks on the target, do not trigger a copy operation (and hence are termed Copy on First Access).

If the replication session is terminated, then the target device only has the data that was accessed until the termination, not the entire contents of the source at the point-in-time. In this case, the data on the target cannot be used for a restore, as it is not a full replica of the source.

In Full Copy mode, all data from the source is copied to the target in the background. Data is copied regardless of access. If access to a block that has not yet been copied is required, this block is preferentially copied to the target. In a complete cycle of the Full Copy mode, all data from the source is copied to the target. If the replication session is terminated now, the target will contain all the original data from the source at the point-in-time of activation. This makes the target a viable copy for recovery, restore, or other business continuity operations.

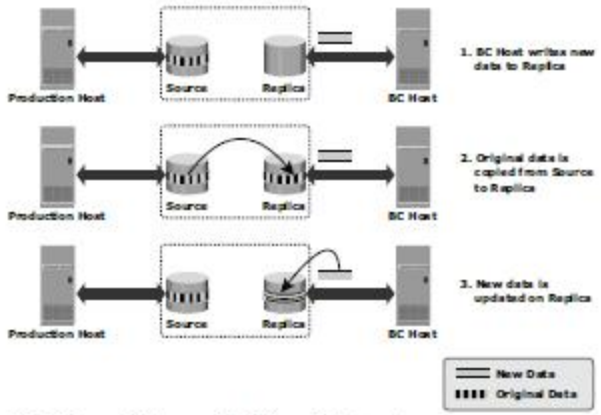


Figure 13-9: Copy on first access (CoFA) – write to target

The key difference between pointer-based, Full Copy mode and full-volume mirroring is that the target is immediately accessible on session activation in Full Copy mode. In contrast, one has to wait for synchronization and detachment to access the target in full-volume mirroring. Both the full-volume mirroring and pointer-based full-volume replication technologies require the target devices to be at least as large as the source devices. In addition, full-volume mirroring and pointer-based full-volume replication in Full Copy mode can provide incremental resynchronization or restore capability.

Pointer-Based Virtual Replication

In *pointer-based virtual replication*, at the time of session activation, the target contains pointers to the location of data on the source. The target does not contain data, at any time. Hence, the target is known as a

virtual replica. Similar to pointer-based full-volume replication, a protection bitmap is created for all data on the source device, and the target is immediately accessible. Granularity can range from 512 byte blocks to 64 KB blocks or greater. When a write is issued to the source for the first time after session activation, original data at that address is copied to a predefined area in the array. This area is generally termed the *save location*. The pointer in the target is updated to point to this data address in the save location. After this, the new write is updated on the source. This process is illustrated in Figure 13-10.

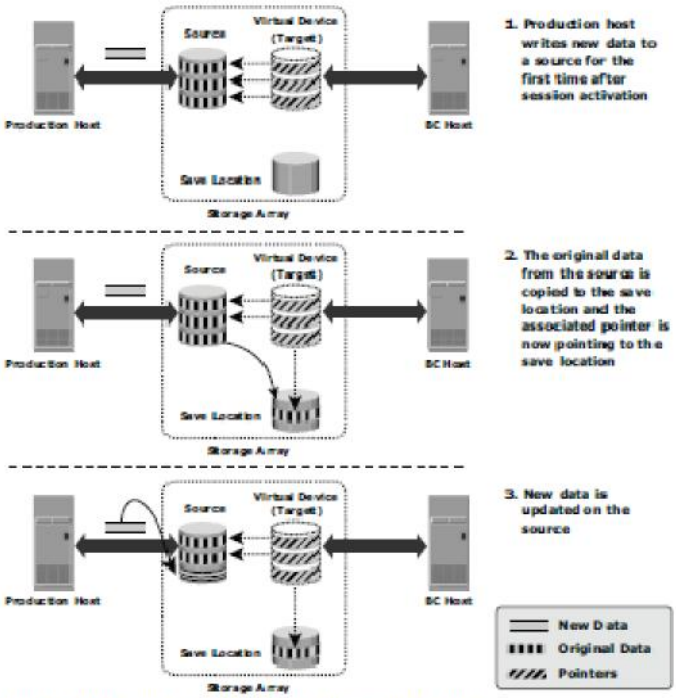


Figure 13–10: Pointer-based virtual replication – write to source

When a write is issued to the target for the first time after session activation, original data is copied from the source to the save location and similarly the pointer is updated to data in save location. Another copy of the original data is created in the save location before the new write is updated on the save location. This process is illustrated in Figure 13-11.

When reads are issued to the target, unchanged data blocks since session activation are read from the source. Original data blocks that have changed are read from the save location. Pointer-based virtual replication uses CoFW technology. Subsequent writes to the same data block on the source or the target do not trigger a copy operation.

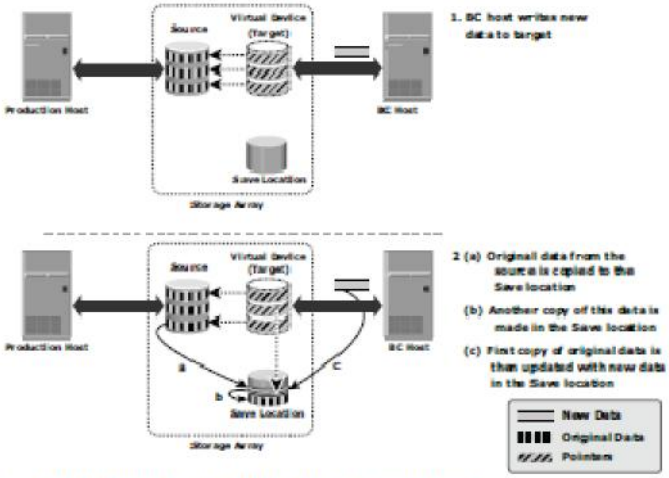


Figure 13-11: Pointer-based virtual replication – write to target

Data on the target is a combined view of unchanged data on the source and data on the save location. Unavailability of the source device invalidates the data on the target. As the target only contains pointers to data, the physical capacity required for the target is a fraction of the source device. The capacity required for the save location depends on the amount of expected data change.

5.Discuss various modes of replication in three site replication in detail(10m)

Three-Site Replication

In synchronous and asynchronous replication, under normal conditions the workload is running at the source site. Operations at the source site will not be disrupted by any failure to the target site or to the network used for replication. The replication process resumes as soon as the link or target site issues are resolved. The source site continues to operate without any remote protection. If failure occurs at the source site during this time, RPO will be extended. In synchronous replication, source and target sites are usually within 200 KM (125 miles) of each other. Hence, in the event of a regional disaster, both the source and the target sites could become unavailable. This will lead to extended RPO and RTO because the last known good copy of data would have to come from another source, such as offsite tape library. A regional disaster will not affect the target site in asynchronous replication,as the sites are typically several hundred or several thousand kilometers apart. If the source site fails, production can be shifted to the target site, but there will be no remote protection until the failure is resolved. *Three-site replication* is used to mitigate the risks identified in two-site replication. In a three-site replication, data from the source site is replicated to two remote data centers. Replication can be synchronous to one of the two data centers, providing a zero-RPO solution. It can be asynchronous or disk buffered to the otherremote data center, providing a finite RPO. Three-site remote replication can be implemented as a cascade/multi-hop or a triangle/multi-target solution.

Three-Site Replication—Cascade/Multi-hop

In the *cascade/multi-hop* form of replication, data flows from the source to the intermediate storage array, known as a *bunker*, in the first hop and then from abunker to a storage array at a remote site in the second hop. Replication between the source and the bunker occurs synchronously, but replication between the bunker and the remote site can be achieved in two ways: disk-buffered mode or asynchronous mode.

Synchronous + Asynchronous

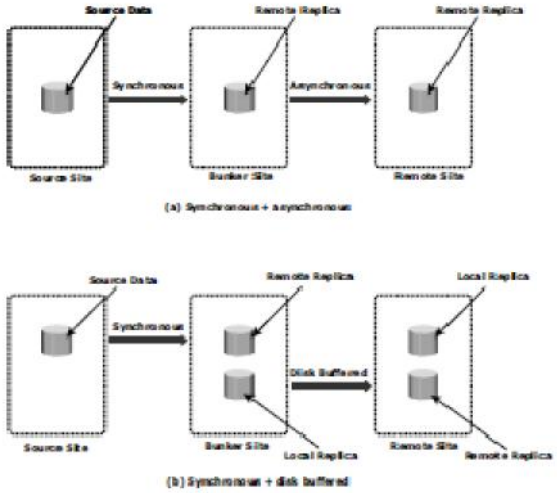
This method employs a combination of synchronous and asynchronous remote replication technologies. Synchronous replication occurs between the source and the bunker. Asynchronous replication occurs between the bunker and the remote site. The remote replica in the bunker acts as the source for the asynchronous replication to create a remote replica at the remote site. Figure 14-10(a) illustrates the synchronous + asynchronous method.RPO at the remote site is usually on the order of minutes in this implementation.

In this method, a minimum of three storage devices are required (including the source) to replicate one storage device. The devices containing a synchronous remote replica at the bunker and the asynchronous replica at the remote are the other two devices. If there is a disaster at the source, operations are failed over to the bunker site with zero or near-zero data loss. But unlike the synchronous two-site situation, there is still remote protection at the third site. The RPO between the bunker and third site could be on the order of minutes.If there is a disaster at the bunker site or if there is a network link failure between the source and bunker sites, the source site will continue to operate as normal but without any remote replication. This situation is very similar to two-site replication when a failure/disaster occurs at the target site. The updates to the remote site cannot occur due to the failure in the bunker site. Hence, the data at the remote site keeps falling behind, but the advantage here is that if the source fails during this time, operations can be resumed at the remote site. RPO at the remote site depends on the time difference between the bunker site failure and source site failure.

A *regional disaster* in three-site cascade/multihop replication is very similar to a source site failure in two-site asynchronous replication. Operations will failover to the remote site with an RPO on the order of minutes. There is no remote protection until the regional disaster is resolved. Local replication technologies could be used at the remote site during this time.If a disaster occurs at the remote site, or if the network links between the bunker and the remote site fail, the source site continues to work as normal with disaster recovery protection provided at the bunker site.

Synchronous + Disk Buffered

This method employs a combination of local and remote replication technologies. Synchronous replication occurs between the source and the bunker: A consistent PIT local replica is created at the bunker. Data is transmitted from the local replica at the bunker to the remote replica at the remote site. Optionally, a local replica can be created at the remote site after data is received from the bunker. Figure 14-10(b) illustrates the synchronous + disk buffered method.In this method, a minimum of four storage devices are required (includingthe source) to replicate one storage device. The other three devices are the synchronous remote replica at the bunker, a consistent PIT local replica at the bunker, and the replica at the remote site. RPO at the remote site is usually in the order of hours in this implementation. For example, if a local replica is created at 10:00 am at the bunker and it takes an hour to transmit this data to the remote site, changes made to the remote replica at the bunker since 10:00 am are tracked. Hence only one hour's worth of data has to be resynchronized between the bunker and the remote site during the next cycle. RPO in this case will also be two hours, similar to disk-buffered replication.



igure 14-10: Three-site replication

The process of creating the consistent PIT copy at the bunker and incrementally updating the remote replica and the local replica at the remote site occurs continuously in a cycle. This process can be automated and controlled from the source.

Three-Site Replication—Triangle/Multi-target

In the *three-site triangle/multi-target replication,* data at the source storage array is concurrently replicated to two different arrays. The source-to-bunker site (target 1) replication is synchronous, with a near-zero RPO. The source-toremote site (target 2) replication is asynchronous, with an RPO of minutes. The distance between the source and the remote site could be thousands of miles. This configuration does not depend on the bunker site for updating data on the remote site, because data is asynchronously copied to the remote site directly from the source. The key benefit of three-site triangle/multi-target replication is the ability to failover to either of the two remote sites in the case of source site failure, withdisaster recovery (asynchronous) protection between them. Resynchronization between the two surviving target sites is incremental. Disaster recovery protection is always available in the event of any one site failure. During normal operations all three sites are available and the workload is at the source site. At any given instant, the data at the bunker and the source is identical. The data at the remote site is behind the data at the source and the bunker. The replication network links between the bunker and remote sites will be in place but not in use. Thus, during normal operations there is no data movement between the bunker and remote arrays. The difference in the data between the bunker and remote sites is tracked, so that in the event of a source site disaster, operations can be resumed at the bunker or the remote sites with incremental resynchronization between the sites

6.Describe storage infrastructure management activities(10M)

All the management tasks in a storage infrastructure can be broadly categorized into availability management, capacity management, performance management, security management, and reporting.

Availability management

The critical task in availability management is establishing a proper guideline for all configurations to ensure availability based on service levels. For example, when a server is deployed to support a critical business function, the highest availability standard is usually required. This is generally accomplished by deploying two or more HBAs, multipathing software with path failover capability, and server clustering. The server must be connected to the storage array using at least two independent fabrics and switches that have built-in redundancy. Storage devices with RAID protection are made available to the server using at least two front-end ports. In addition, these storage arrays should have built-in redundancy for various components, support backup, and local and remote replication. Virtualization technologies have significantly improved the availability management task. With virtualization in place resources can be dynamically added or removed to maintain the availability.

Capacity management

The goal of *capacity management* is to ensure adequate availability of resources for all services based on their service level requirements. Capacity management provides capacity analysis, comparing allocated storage to forecasted storage on a regular basis. It also provides trend analysis of actual utilization of allocated storage and rate of consumption, which must be rationalized against storage acquisition and deployment timetables.

Storage provisioning is an example of capacity management. It involves activities such as device configuration and LUN masking on the storage array and zoning configuration on the SAN and HBA components. Capacity management also takes into account the future needs of resources, and setting up monitors and analytics to gather such information.

Performance management

Performance management ensures the optimal operational efficiency of all components. Performance analysis is an important activity that helps to identify the performance of storage infrastructure components. This analysis provides the information — whether a component is meeting expected performance levels.

Several performance management activities are initiated for the deployment of an application or server in the existing storage infrastructure. Every component must be validated for adequate performance capabilities as defined by the service levels. For example, to optimize expected performance levels, activities on the server such as the volume configuration, designing the database, application layout configuration of multiple HBAs, and intelligent multipathing software must be fine-tuned. The performance management tasks on a SAN include designing sufficient ISLs in a multi-switch fabric with adequate bandwidth to support the required performance levels. The storage array configuration tasks include selecting the appropriate RAID type and LUN layout, front-end and back-end ports, and LUN accessibility (LUN masking) while considering the end-to-end performance.

Security Management

Security management prevents unauthorized access and configuration of storage infrastructure components. For example, while deploying an application or a server, the security management tasks include managing user accounts and access policies, that authorizes users to perform role-based activities. The security management tasks in the SAN environment include configuration of zoning to restrict an HBA's unauthorized access to the specific storage array ports. LUN masking prevents data corruption on the storage array by restricting host access to a defined set of logical devices.

Reporting

It is difficult for businesses to keep track of the resources they have in their data centers, for example, the number of storage arrays, the array vendors, how the storage arrays are being used, and by which applications. Reporting on a storage infrastructure involves keeping track and gathering information from various components/processes. This information is compiled to generate reports for trend analysis, capacity planning, chargeback, performance, and to illustrate the basic configuration of storage infrastructure components. Capacity planning reports also contain current and historic information about utilization of storage, file system, database tablespace, and ports. Configuration or asset management reports include details about device allocation, local or remote replicas, and fabric configuration; and list all equipment, with details such as their value, purchase date, lease status, and maintenance records. Chargeback reports contain information about the allocation or utilization of storage infrastructure components by various departments or user groups. Performance reports provide details about the performance of various storage infrastructure components.