SEE Question Papers

1. Numerical

R_s = 10.767 ms, the maximum IOPS will be
$1/(10.767*10^{-3})$ = 92.876 IOPS

6. Consider a disk I/O system in which an I/O request arrives at the rate of 80 IOPS. The disk service time is 6 ms.
   a. Compute the following: Utilization of I/O controller, Total response time, Average queue size, and Total time spent by a request in a queue.
   b. Compute the preceding parameter if the service time is halved.

Solution/Hint:

Arrival rate (a) = 80 IOPS, consequently, the arrival time
Ra = 1/a = 1/80 = 12.5 ms
Rs = 6 ms (given)

1. Utilization (U) = Rs/Ra
   = 6/12.5 = 0.48 or 48%
2. Response time (R) = Rs/(1-U)
   = 6/(1-0.48) = 11.5 ms
3. Average queue size = $U^2/(1-U)$
   = $(0.48)^2/(1-0.48)$
   = 0.44
4. Time spent by a request in a queue = UxR, or the total response time-service time = 0.48x11.5 = 5.52 ms

Now, if controller power is doubled, or the service time is halved; consequently Rs = 3 ms in this scenario.

1. Utilization (U) = Rs/Ra = 3/12.5 = 0.24 or 24 %
2. Response time (R) = Rs/(1-U) = 3/(1-0.24) = 3.9 ms
3. Average queue size = $U^2/(1-U) = (0.24)^2/(1-0.24)$ = 0.08
4. Time spent by a request in a queue = 0.24x3.9 = 0.936 ms

Consider a disk I/O system in which an I/O request arrives at a rate of 100 I/Os per second. The service time, R_S, is 8 ms. The following measures of disk performance can be computed using the relationships developed above — utilization of I/O controller (U), total response time (R), average queue size $[U^2 / (1 - U)]$ and total time spent by a request in a queue $(U \times R)$, as follows:

Arrival rate (a) = 100 I/O/s; consequently, the arrival time

$R_a = 1/a = 10$ ms

$R_S = 8$ ms (given)

1. Utilization (U) = $R_S / R_a = 8 / 10 = 0.8$ or 80%
2. Response time (R) = $R_S / (1 - U) = 8 / (1 - 0.8) = 40$ ms
3. Average queue size = $U^2 / (1 - U) = (0.8)^2 / (1 - 0.8) = 3.2$
4. Time spent by a request in a queue = $U \times R$, or the total response time-service time = 32 ms

Now, if controller power is doubled, the service time is halved; consequently, $R_S = 4$ ms in this scenario.

1. Utilization (U) = $4 / 10 = 0.4$ or 40%
2. Response time (R) = $4 / (1 - 0.4) = 6.67$ ms
3. Average queue size = $(0.4)^2 / (1 - 0.4) = 0.26$
4. Time spent by a request in a queue = $0.4 \times 6.67 = 2.67$ ms

2. Explain about RAID 4 & RAID 5

RAID 4

- RAID 4 stripes data for high performance and uses parity for improved fault tolerance. Data is striped across all disks except the parity disk in the array.
- Parity information is stored on a dedicated disk so that the data can be rebuilt if a drive fails. Striping is done at the block level.
- Unlike RAID 3, data disks in RAID 4 can be accessed independently so that specific data elements can be read or written on a single disk without the reading or writing of an entire stripe. RAID 4 provides good read throughput and reasonable write throughput.

RAID 5

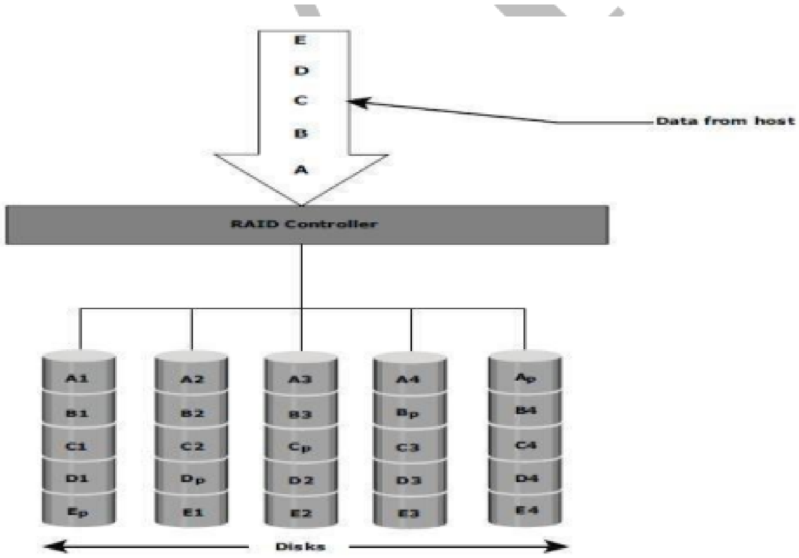
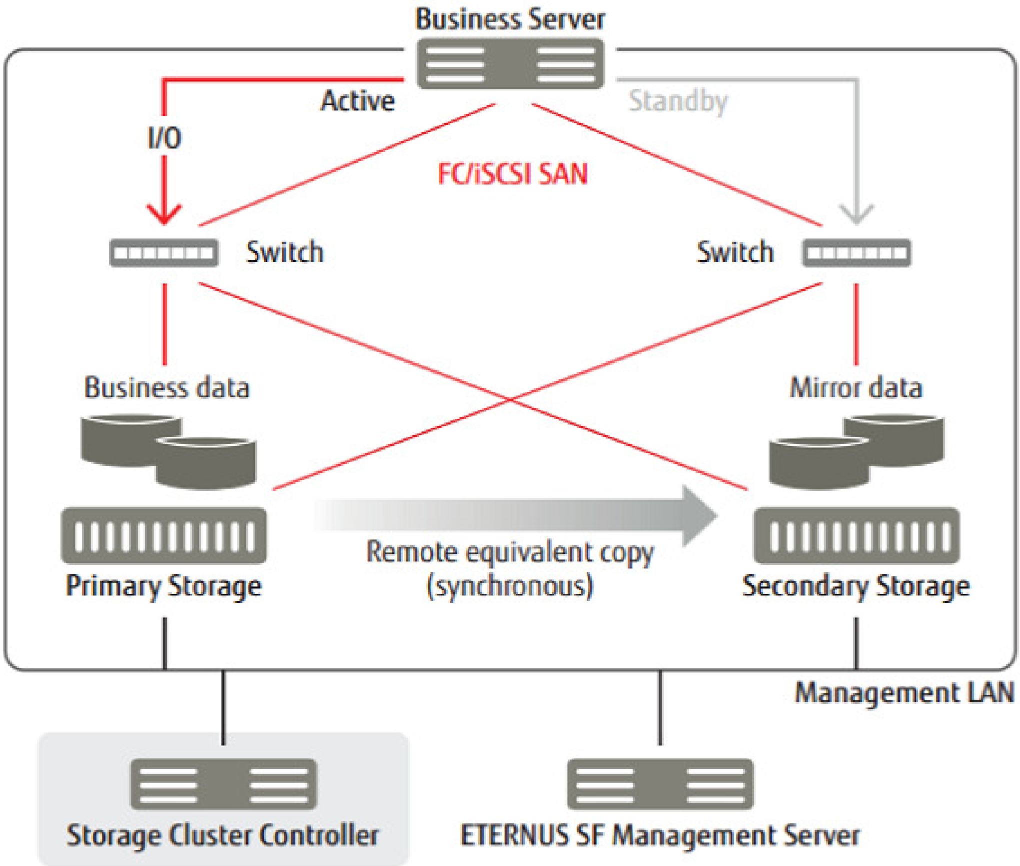


Fig 1.18: RAID 5

- RAID 5 is a versatile RAID implementation.
- It is similar to RAID 4 because it uses striping. The drives (strips) are also independently accessible.
- The difference between RAID 4 and RAID 5 is the parity location. In RAID 4, parity is written to a dedicated drive, creating a write bottleneck for the parity disk.
- In RAID 5, parity is distributed across all disks. The distribution of parity in RAID 5 overcomes the Write bottleneck. Below Figure illustrates the RAID 5 implementation.
- RAID 5 is good for random, read-intensive I/O applications and preferred for messaging, data mining, medium-performance media serving, and relational database management system (RDBMS) implementations, in which database administrators (DBAs) optimize data access

3. Explain Remote Mirroring with neat diagram



Remote Mirroring is to provide data accessibility protection for an application/s using physically separate locations. A remote mirror volume is a mirror volume with a source in another cluster. You can use remote mirrors for offsite backup, for data transfer to remote facilities, and for load and latency balancing for large websites. By mirroring the cluster's root volume and all other volumes in the cluster, you can create an entire mirrored cluster that keeps in sync with the source cluster.

Synchronous remote mirroring

Synchronous remote mirroring does this by neither completing nor acknowledging the local write until the remote write is completed and acknowledged. Additional writes can't occur until each preceding write has been completed and acknowledged.

Asynchronous remote mirroring

local writes are completed and acknowledged before the remote writes. Asynchronous remote mirroring is a "store-and-forward" technique that reduces I/Os and wait delays, allowing remote writes to fall behind the local writes. This means the RPO for lost data can range from seconds to minutes, and even hours in some cases.

4. Explain with a neat diagram different layers of Fiber Channel Protocol Stack.

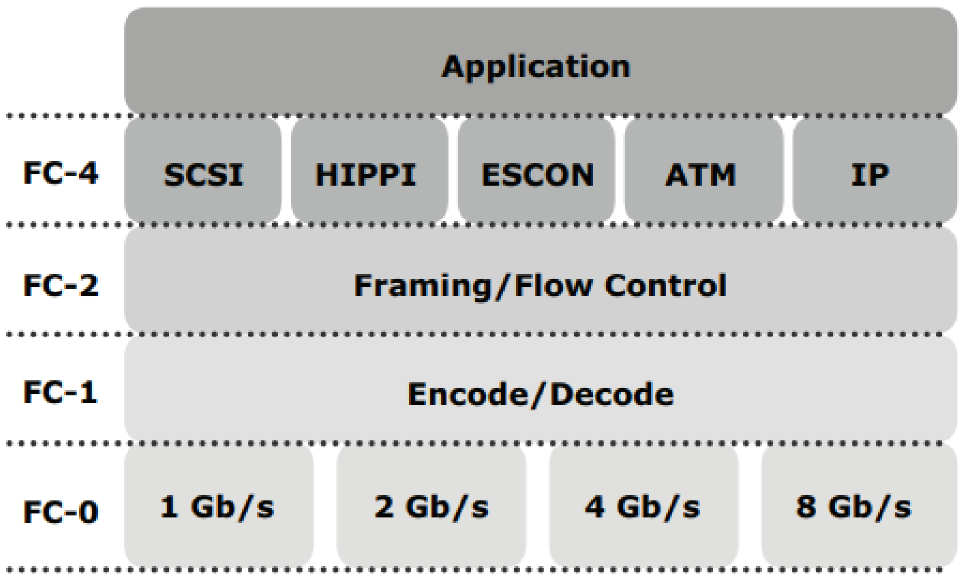


Figure 6-13: Fibre channel protocol stack

FCP defines the communication protocol in five layers:

FC-4 Upper Layer Protocol

FC-4 is the uppermost layer in the FCP stack. This layer defines the application interfaces and the way Upper Layer Protocols (ULPs) are mapped to the lower FC layers. The FC standard defines several protocols that can operate on the FC-4 layer (see Figure 6-7). Some of the protocols include SCSI, HIPPI Framing Protocol, Enterprise Storage Connectivity (ESCON), ATM, and IP.

FC-2 Transport Layer

The FC-2 is the transport layer that contains the payload, addresses of the source and destination ports, and link control information. The FC-2 layer provides Fibre Channel addressing, structure, and organization of data (frames, sequences, and exchanges). It also defines fabric services, classes of service, flow control, and routing.

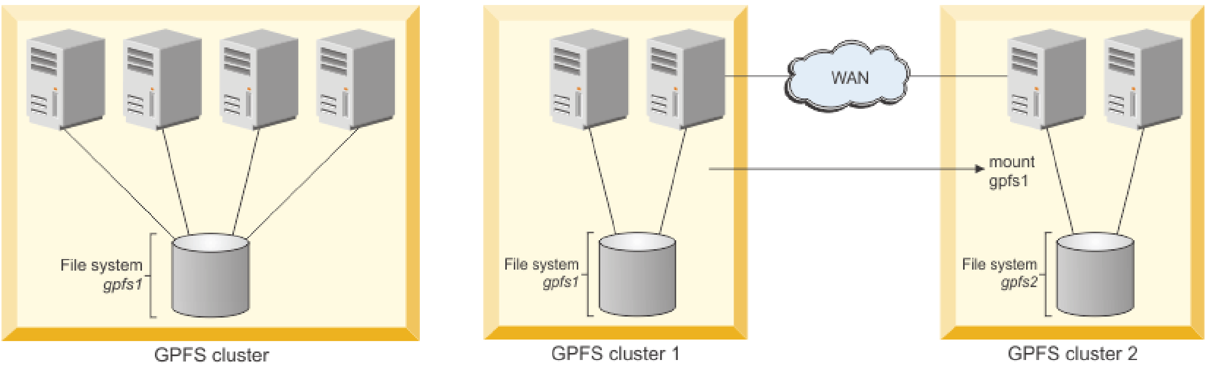
FC-1 Transmission Protocol

This layer defines the transmission protocol that includes serial encoding and decoding rules, special characters used, and error control. At the transmitter node, an 8-bit character is encoded into a 10-bit transmissions character. This character is then transmitted to the receiver node. At the receiver node, the 10-bit character is passed to the FC-1 layer, which decodes the 10-bit character into the original 8-bit character.

FC-0 Physical Interface

FC-0 is the lowest layer in the FCP stack. This layer defines the physical interface, media, and transmission of raw bits. The FC-0 specification includes cables, connectors, and optical and electrical parameters for a variety of data rates. The FC transmission can use both electrical and optical media.

5. What is GPFS? Explain with neat diagram



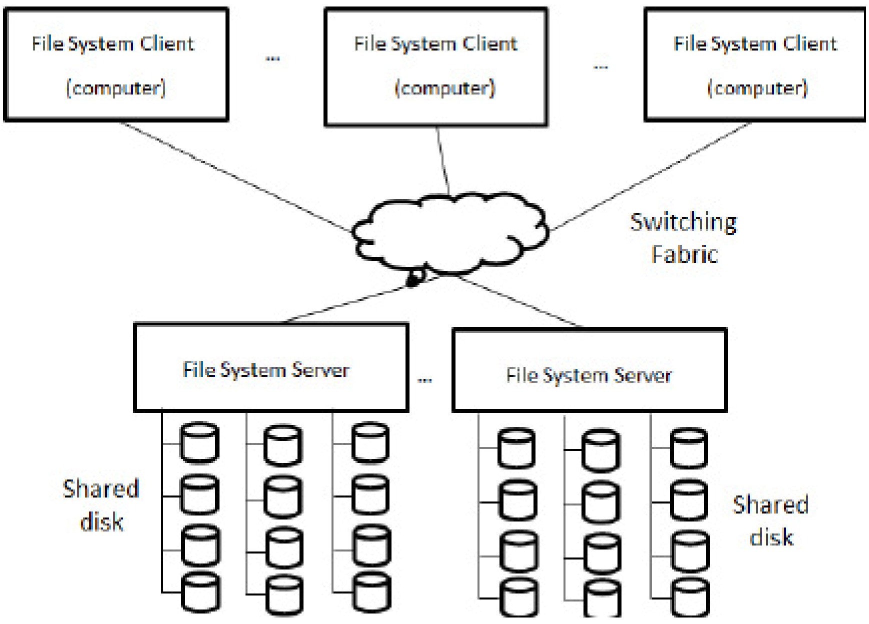
GPFS clusters can contain a mix of all supported node types including Linux, AIX, and Windows Server. These nodes can all be attached to a common set of SAN storage or through a mix of SAN and network attached nodes. Nodes can all be in a single cluster, or data can be shared across multiple clusters. A cluster can be contained in a single data center or spread across geographical locations.

You can configure a GPFS cluster in which some nodes have a direct attachment to the disks and others access the disks through other GPFS nodes. This configuration is often used in large clusters or to provide a cost-effective, high-performance solution.

When a GPFS node is providing access to a disk for another GPFS node, it is called an NSD Server. The GPFS node accessing the data through an NSD server is called a GPFS client.

GPFS allows you to share data across multiple GPFS clusters. Once a file system is mounted in another GPFS cluster all access to the data is the same as if you were in the host cluster. You can connect multiple clusters within the same data center or across long distances over a WAN. In a multicluster configuration, each cluster can be placed in a separate administrative group simplifying administration or provide a common view of data across multiple organizations.

6. Explain Shared Disk File System with neat diagram.



A shared disk file system, also known as clustered file system or SAN file system, is an enterprise storage file system which can be shared (concurrently accessed for reading and writing) by multiple computers. Such devices are usually clustered servers, which connect to the underlying block device over an external storage device. Such a device is commonly a storage area network (SAN).Shared disk file systems are necessary because with regular file systems, if multiple instances (servers) were to attempt concurrent access to the same physical device, the data would rapidly become corrupt, because there is nothing to prevent two devices from performing a modification of the same part of the file system at the same time. Conventional file locking is no aid in this, as file locking operates above the file system level; file locking can protect files against concurrent access but offers no protection when server / operating system manipulates the file system directly.

Directly means the server manipulates blocks of storage device. The files in the file system consist of these blocks.

7. Explain Out Band Management which goes through different Interfaces.

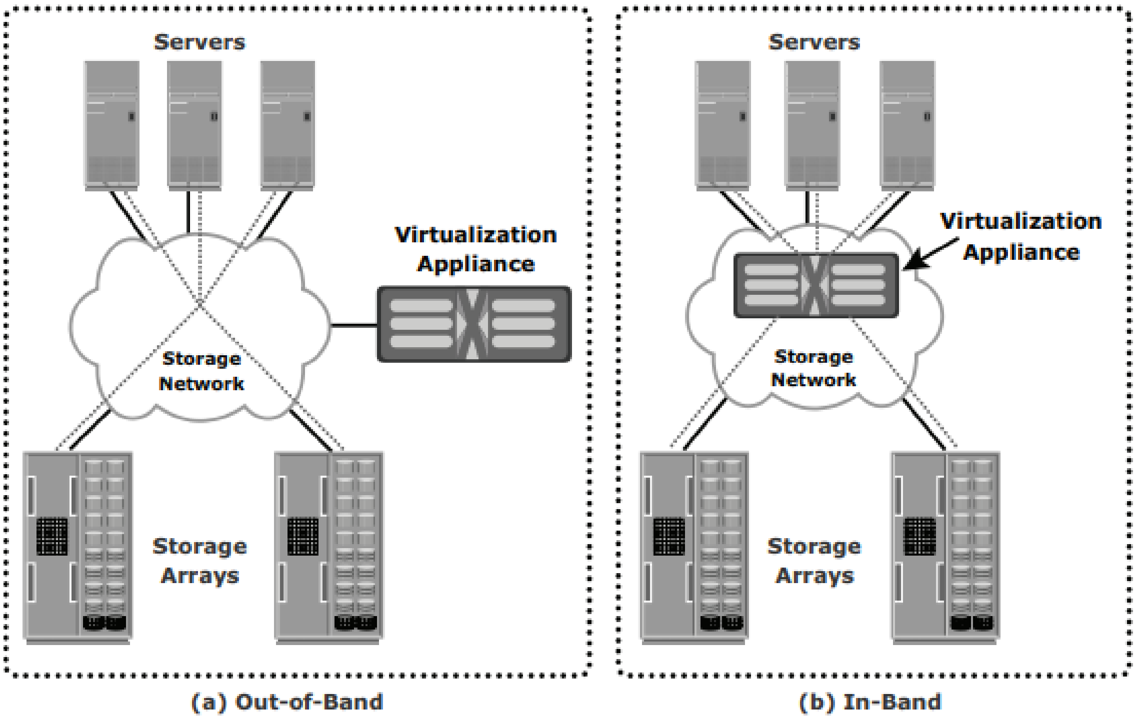


Figure 10-5: Storage virtualization configuration

Out Band

In systems management, out-of-band management involves the use of management interfaces (or serial ports) for managing and networking equipment. Out-of-band (OOB) is a networking term which refers to having a separate channel of communication which does not travel over the usual data stream.

Out-of-band management allows the network operator to establish trust boundaries in accessing the management function to apply it to network resources. It also can be used to ensure management connectivity (including the ability to determine the status of any network component) independent of the status of other in-band network components.

In computing, one form of out-of-band management is sometimes called lights-out management (LOM) and involves the use of a dedicated management channel for device maintenance. It allows a system administrator to monitor and manage servers and other network-attached equipment by remote control regardless of whether the machine is powered on or whether an OS is installed or functional.

By contrast, in-band management through VNC or SSH is based on in-band connectivity (the usual network channel). It typically requires software that must be installed on the remote system being managed and only works after the operating system has been booted and networking is brought up. It does not allow management of remote network components independently of the current status of other network components. A classic example of this limitation is when a sysadmin attempts to reconfigure the network on a remote machine only to find themselves locked out and unable to fix the problem without physically going to the machine. Despite these limitations, in-band solutions are still common because they are simpler and much lower-cost.

Both in-band and out-of-band management are usually done through a network connection, but an out-of-band management card can use a physically separated network connector if preferred. A remote management card usually has at least a partially independent power supply and can switch the main machine on and off through the network. Because a special device is required for each machine, out-of-bandwidth management can be much more expensive.

Out-of-band

In an out-of-band implementation, the data flow is separated from the control flow. This is achieved by separating the data and meta-data (data about the data) into different places. Out-of-band virtualization involves moving all mapping and locking tables to a separate server (the meta-data controller) that contains the meta-data of the files. In an out-of-band solution the servers request authorization to data from the meta-data controller, which grants it, handles locking, and so on. Once they are authorized, servers access the data directly without any meta-data controller intervention. Once a client has obtained access to a file, all I/O will go directly over the SAN to the storage devices. For many operations, the meta-data controller does not even intervene. Separating the flow of control and data in this manner allows the I/O to use the full bandwidth that a SAN provides, while control could go over a separate network or routes in the SAN that are isolated for this purpose.

Other advantages include:

- Releasing the customer from a particular vendor’s storage
- Providing storage management for the SAN

- Offering excellent scalability
- Offloading host processing
- Supporting storage management from multiple vendors
- Integrating well with storage management software
- Supporting multiple heterogeneous hosts
- Relatively low overhead in the data path

In-band

When we implement an in-band virtual storage network, both data and control flow over the same path. Levels of abstraction exist in the data path, and storage can be pooled under the control of a domain manager. In general, in-band solutions are perceived to be simpler to implement, especially because they do not require special software to be installed in servers (other than conventional multi-pathing software). In-band solutions can also provide caching and advanced functions within the storage network. This can help to improve the performance of existing disk systems and can extend their useful life, and reduce the cost of new storage capacity by enabling the use of lower function, lower cost disk systems, without the loss of performance.

Other advantages include:

- Ability to offload function from the host
- Providing storage management for the SAN
- Performing performance optimizations in the data path
- Supporting host systems not in a cluster
- Supporting multiple heterogeneous hosts
- Releasing the customer from a particular vendor’s storage
- Integrating with storage to create a better management picture
- Offering excellent scalability

The IBM TotalStorage SAN Volume Controller is an example of an in-band solution.