# Answers

## 1. What is Storage Virtualization? Explain storage Virtualization at the Network level.

Storage virtualization is the process of presenting a logical view of the physical storage resources to a host. This logical storage appears and behaves as physical storage directly connected to the host.

### Storage Virtualization at Network level

Storage virtualization at the network is implemented using either the in-band or the out-of-band methodology. In an out-of-band implementation, the virtualized environment configuration is stored external to the data path.

The configuration is stored on the virtualization appliance configured external to the storage network that carries the data. This configuration is also called split-path because the control and data paths are split. This configuration enables the environment to process data at a network speed with only minimal latency added for translation of the virtual configuration to the physical storage.

## 2. Describe SNIA Storage Virtualization Taxonomy?

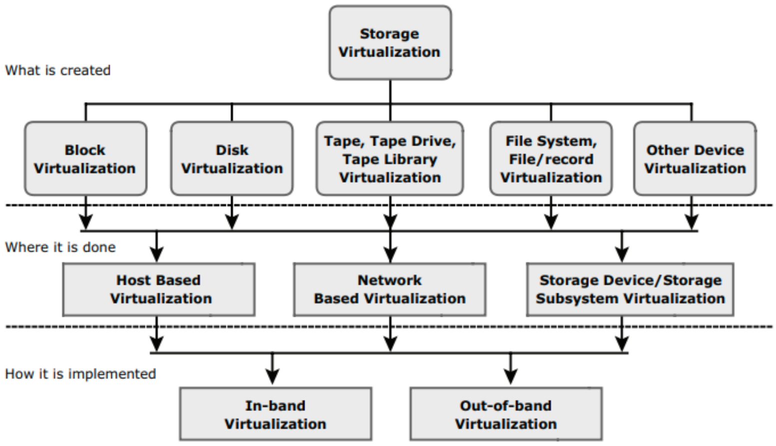

**Figure 10-3:** SNIA storage virtualization taxonomy

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

- **First Level:** The first level of the storage virtualization taxonomy addresses “what” is created. It specifies the types of virtualization: block virtualization, file virtualization, disk virtualization, tape virtualization, or any other device virtualization.
- **Second Level:** The second level describes “where” the virtualization can take place. This requires a multilevel approach that characterizes virtualization at all three levels of the storage environment: server, storage network, and storage. An effective virtualization strategy distributes intelligence across all three levels while centralizing the management and control functions.
- **Third Level:** The third level of the storage virtualization taxonomy specifies the network level virtualization methodology, in-band or out-of-band.

# 3. Explain the different types of Storage Virtualization:
# a. Block level b. File level

## Block Level

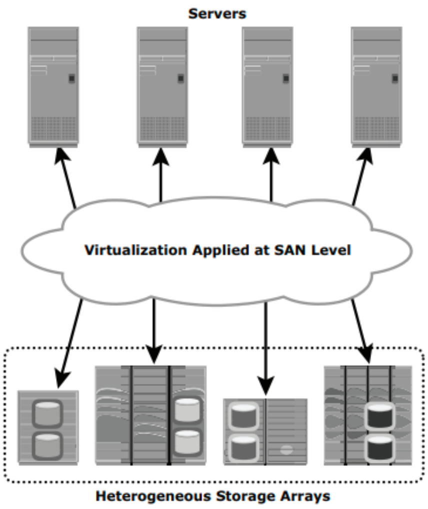

**Figure 10-6:** Block-level storage virtualization

- Block-level storage virtualization provides a translation layer in the SAN, between the hosts and the storage 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. 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.

# File Level

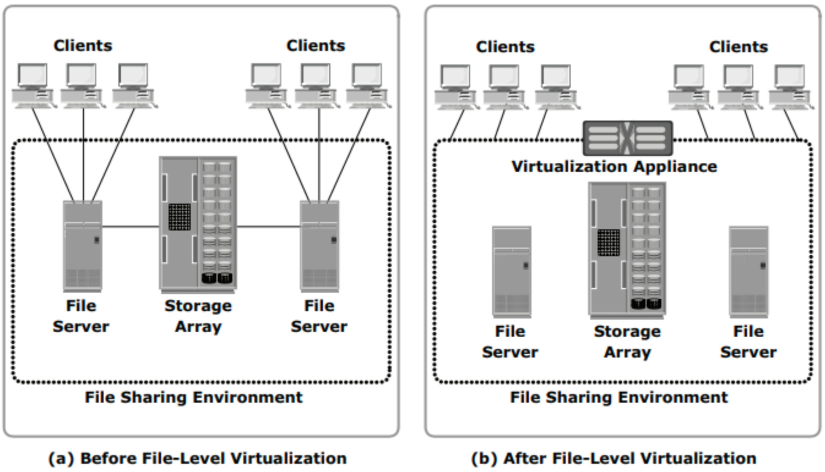


**Figure 10-7:** NAS device before and after file-level virtualization

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 non-disruptive file migrations.

Before virtualization, each NAS device or file server is physically and logically independent. Each host knows exactly where its file-level resources are located. 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 require downtime for the file servers. 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.

# 4. Explain different forms of Storage Virtualization:
# a. Memory Virtualization   b. Server Virtualization

## Memory Virtualization

- Virtual memory makes an application appear as if it has its own contiguous logical memory independent of the existing physical memory resources. With technological advancements, memory technology has changed and the cost of memory has decreased.
- In a virtual memory implementation, a memory address space is divided into contiguous blocks of fixed-size pages. A process known as *paging* saves inactive memory pages onto the disk and brings them back to physical memory when required. This enables efficient use of available physical memory among different processes.
- The space used by VMMs on the disk is known as a swap file. The operating system typically moves the least used data into the swap file so that RAM will be available for processes that are more active.

## Server Virtualization

- Server virtualization enables multiple operating systems and applications to run simultaneously on different virtual machines created on the same physical server.
- Virtual machines provide a layer of abstraction between the operating system and the underlying hardware. Within a physical server, any number of virtual servers can be established; depending on hardware capabilities.
- Each virtual server seems like a physical machine to the operating system, although all virtual servers share the same underlying physical hardware in an isolated manner.

- Individual virtual servers can be restarted, upgraded, or even crashed, without affecting the other virtual servers on the same physical machine.

# 5. Differentiate between Symmetric and Asymmetric Virtualization

## Symmetric Virtualization

Symmetric Storage virtualization In symmetric storage virtualization the data and control flow goes down the same path. This means that the abstraction from physical to logical storage necessary for virtualization must take place within the data flow. As a result, the metadata controller is positioned precisely in the data flow between server and storage devices, which is why symmetric virtualization is also called in-band virtualization.

### Advantages

- The application servers can easily be provided with data access both on the block and file levels, regardless of the underlying physical storage devices.
- The administrator has complete control over which storage resources are available to which servers at a central point. This increases security and eases the administration.

## Asymmetric Virtualization

In contrast to symmetric virtualization, in asymmetric virtualization, the data flow is separated from the control flow. This is achieved by moving all mapping operations from logical to physical drives to a metadata controller outside the data path.

The metadata controller now only has to look after the administrative and control tasks of virtualization, the flow of data takes place directly from the application servers to the storage devices. As a result, this approach is also called out-band virtualization.

### Advantages

- Complete control of storage resources by absolutely centralized management on the metadata controller.

- Maximum throughput between servers and storage devices by the separation of the control flow from the data flow, thus avoiding additional devices in the data path.
- In comparison to the development and administration of a fully functional volume manager on every server, the porting of the agent software is associated with a low cost.

# 6. What is SAN? Explain SAN and its Evolution.

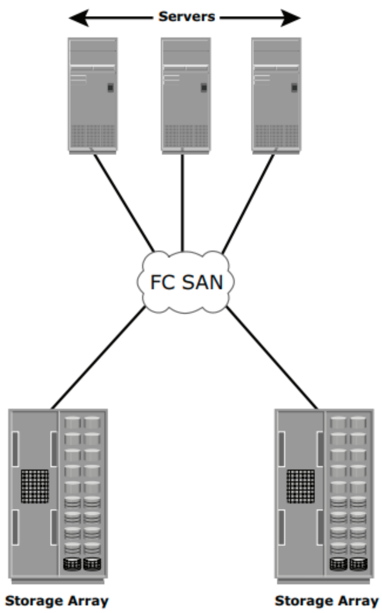

**Figure 6-1:** SAN implementation

## SAN

- A storage area network (SAN) carries data between servers (also known as hosts) and storage devices through fiber channel switches.
- A SAN enables storage consolidation and allows storage to be shared across multiple servers. It enables organizations to connect geographically dispersed servers and storage.
- A SAN provides the physical communication infrastructure and enables secure and robust communication between host and storage devices.

## Evolution of SAN

In its earliest implementation, the SAN was a simple grouping of hosts and the associated storage that was connected to a network using a hub as a connectivity device. This configuration of a SAN is known as a Fibre Channel Arbitrated Loop (FC-AL). The use of hubs resulted in isolated FC-AL SAN islands because hubs provide limited connectivity and bandwidth.

The inherent limitations associated with hubs gave way to high-performance FC switches. The switched fabric topologies improved connectivity and performance, which enabled SANs to be highly scalable.

Today, Internet Protocol (IP) has become an option to interconnect geographically separated SANs. Two popular protocols that extend block-level access to applications over IP are iSCSI and Fibre Channel over IP (FCIP).

# 7. Explain the different components of SAN

## Node Port

In fiber channels, devices such as hosts, storage, and tape libraries are all referred to as nodes. Each node is a source or destination of information for one or more nodes. Each node requires one or more ports to provide a physical interface for communicating with other nodes. These ports are integral components of an HBA and the storage front-end adapters.

## Cabling

SAN implementations use optical fiber cabling. Copper can be used for shorter distances for back-end connectivity, as it provides a better signal-to-noise ratio for distances up to 30 meters. Optical fiber cables carry data in the form of light. There are two types of optical cables, multi-mode, and single-mode.

- Multi-mode fiber (MMF) cable carries multiple beams of light projected at different angles simultaneously onto the core of the cable.
- Single-mode fiber (SMF) carries a single ray of light projected at the center of the core.

MMFs are generally used within data centers for shorter distance runs, while SMFs are used for longer distances.

## Interconnect Devices

Hubs, switches, and directors are the interconnect devices commonly used in SAN. Hubs are used as communication devices in FC-AL implementations. Hubs physically connect nodes in a logical loop or a physical star topology. Because of the availability of low-cost and high-performance switches, hubs are no longer used in SANs.

## Storage Arrays

- The fundamental purpose of a SAN is to provide host access to storage resources. The large storage capacities offered by modern storage arrays have been exploited in SAN environments for storage consolidation and centralization.
- SAN implementations complement the standard features of storage arrays by providing high availability and redundancy, improved performance, business continuity, and multiple host connectivity.

## SAN Management Software

SAN management software manages the interfaces between hosts, interconnect devices, and storage arrays. The software provides a view of the SAN environment and enables the management of various resources from one central console.

# 8. Explain three basic interconnectivity options of fiber channel Architecture.

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

## Point-to-Point

- Point-to-point is the simplest FC configuration, in which two devices are connected directly to each other.
- 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.

## Fibre Channel Arbitrated Loop

In the FC-AL configuration, devices are attached to a shared loop. FC-AL has the characteristics of token ring topology and 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.

## 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.

# 9. Explain General-purpose versus NAS devices?

- A NAS device is optimized for file-serving functions such as storing, retrieving and accessing files for applications and clients.
- A general-purpose server can be used to host any application, as it runs a generic operating system. Unlike a general-purpose server, a NAS device is dedicated to file-serving.
- It has a real-time operating system dedicated to file serving by using open-standard protocols. Some NAS vendors support features such as native clustering for high availability.

# 10. Describe the different benefits of NAS?

- **Improved flexibility:** Compatible for clients on both UNIX and Windows platforms using industry-standard protocols. NAS is flexible and can serve requests from different types of clients from the same source.

- **Centralized storage:** Centralizes data storage to minimize data duplication on client workstations, simplify data management, and ensure greater data protection.
- **Simplified management:** Provides a centralized console that makes it possible to manage file systems efficiently.
- **Scalability:** Scales well in accordance with different utilization profiles and types of business applications because of the high performance and low-latency design.
- **High availability:** Offers efficient replication and recovery options, enabling high data availability. NAS uses redundant networking components that provide maximum connectivity options. A NAS device can use clustering technology for failover.
- **Security:** Ensures security, user authentication, and file locking in conjunction with industry-standard security schemas.

# 11. Explain NAS file I/O in detail

- NAS uses file-level access for all of its I/O operations.
- File I/O is a high-level request that specifies the file to be accessed but does not specify its logical block address. For example, a file I/O request from a client may specify reading 256 bytes from byte number 1152 onward in a specific file.
- Unlike block I/O, there is no disk volume or disk sector information in a file I/O request.
- The NAS operating system keeps track of the location of files on the disk volume and converts client file I/O into block-level I/O to retrieve data.
- The NAS operating system issues a block I/O request to fulfill the file read and write requests that it receives. The retrieved data is again converted to file level I/O for applications and clients.

# 12. What are the components of NAS explain in detail.

A NAS device has the following components:

- NAS head (CPU and Memory)
- One or more network interface cards (NICs), which provide connectivity to the network. Examples of NICs include Gigabit Ethernet, Fast Ethernet, ATM, and Fiber Distributed Data Interface (FDDI).
- An optimized operating system for managing NAS functionality
- NFS and CIFS protocols for file sharing
- Industry-standard storage protocols to connect and manage physical disk resources, such as ATA, SCSI, or FC

# 13. Explain NFS and CIFS file sharing protocols of NAS

# 14. Discuss the advantage of FC-SW over FC-AL.