**How iSCSI Works**

**iSCSI (Internet Small Computer System Interface) allows SCSI commands to be encapsulated in TCP/IP packets, enabling storage devices to communicate over IP-based networks such as Ethernet. It creates a SAN (Storage Area Network) using standard networking hardware.**

**Workflow of iSCSI Communication**

1. **Initiator:**
   * **A client device (e.g., server) with an iSCSI initiator software or hardware.**
   * **Sends SCSI commands encapsulated in IP packets.**
2. **Network (IP Transport):**
   * **iSCSI commands are transmitted over a standard TCP/IP network.**
   * **Can use Ethernet, VLANs, or other IP-based technologies.**
3. **Target:**
   * **A storage system (e.g., NAS or dedicated SAN device) configured as the iSCSI target.**
   * **Receives, decapsulates, and executes the SCSI commands on the storage hardware.**
4. **Data Return:**
   * **The target sends the requested data or status back to the initiator using the same TCP/IP path.**

**iSCSI Workflow Diagram**

**A computer screen shot of a server

Description automatically generated**

**Key Features:**

* **Scalability: Easily expands storage by adding new targets.**
* **Low Cost: Uses standard networking hardware.**
* **Flexibility: Works over LAN, WAN, or even the internet.**

**TCP/IP (Transmission Control Protocol/Internet Protocol):**

* The foundational protocol suite of the internet and most networks.
* **TCP:** Ensures reliable data delivery with error checking and retransmission.
* **IP:** Handles addressing and routing data packets to their destination.

Both protocols are essential for enabling communication and storage access over networks.

Snapshots, backups, and replications are essential data protection and recovery techniques. Each serves different purposes, with unique characteristics. Here's a comparison:

**1. Snapshots**

**Definition:** A snapshot is a point-in-time representation of a dataset, capturing its state at that moment.

**Key Features:**

* **Quick Creation:** Snapshots are created almost instantly.
* **Space-Efficient:** Stores only the changes made after the snapshot (e.g., via copy-on-write or redirect-on-write).
* **Short-Term Use:** Ideal for testing, quick rollbacks, or recovering recently deleted files.
* **Dependencies:** Often rely on the underlying storage system.

**Use Cases:**

* Test environments.
* Quickly reverting to a known-good state after an update or configuration change.

**2. Backups**

**Definition:** A backup is a complete copy of data stored separately from the source, often used for long-term retention and disaster recovery.

**Key Features:**

* **Independent:** Stored on separate media (disk, tape, or cloud).
* **Comprehensive:** Includes all selected files, folders, or system data.
* **Long-Term Storage:** Retained for weeks, months, or years.
* **Time-Consuming:** Takes longer to create compared to snapshots.

**Use Cases:**

* Recovering data after accidental deletion or corruption.
* Disaster recovery from hardware failures or cyberattacks.

**3. Replications**

**Definition:** Replication involves copying data in real-time or near-real-time from one location to another for redundancy or high availability.

**Key Features:**

* **Continuous Process:** Updates the replica as changes are made to the source.
* **High Availability:** Ensures minimal downtime by providing a secondary active copy.
* **Synchronous or Asynchronous:**
  + **Synchronous:** Writes occur simultaneously on source and replica (ideal for local replication).
  + **Asynchronous:** Writes occur after the source updates (better for long-distance replication).

**Use Cases:**

* Business continuity during a failure.
* Disaster recovery with near-zero data loss (RPO).

**Comparison Table**

| **Feature** | **Snapshots** | **Backups** | **Replications** |
| --- | --- | --- | --- |
| **Purpose** | Point-in-time state | Data recovery & archiving | Redundancy & high availability |
| **Frequency** | Frequent, near-instant | Periodic (daily/weekly) | Continuous (real-time) |
| **Storage** | Shared with source | Separate media | Another device/location |
| **Retention** | Short-term | Long-term | Continuous updates |
| **Performance Impact** | Low (space-efficient) | Moderate (full copy) | High (real-time updates) |
| **Dependency** | Storage system | Independent | Storage & network |

**Choosing the Right Option**

* **Snapshots**: For fast, short-term recovery or testing.
* **Backups**: For long-term retention, historical data recovery, and compliance.
* **Replications**: For minimizing downtime and data loss in high-availability environments.

Snapshots, backups, and replications are essential data protection and recovery techniques. Each serves different purposes, with unique characteristics. Here's a comparison:

**1. Snapshots**

**Definition:** A snapshot is a point-in-time representation of a dataset, capturing its state at that moment.

**Key Features:**

* **Quick Creation:** Snapshots are created almost instantly.
* **Space-Efficient:** Stores only the changes made after the snapshot (e.g., via copy-on-write or redirect-on-write).
* **Short-Term Use:** Ideal for testing, quick rollbacks, or recovering recently deleted files.
* **Dependencies:** Often rely on the underlying storage system.

**Use Cases:**

* Test environments.
* Quickly reverting to a known-good state after an update or configuration change.

**2. Backups**

**Definition:** A backup is a complete copy of data stored separately from the source, often used for long-term retention and disaster recovery.

**Key Features:**

* **Independent:** Stored on separate media (disk, tape, or cloud).
* **Comprehensive:** Includes all selected files, folders, or system data.
* **Long-Term Storage:** Retained for weeks, months, or years.
* **Time-Consuming:** Takes longer to create compared to snapshots.

**Use Cases:**

* Recovering data after accidental deletion or corruption.
* Disaster recovery from hardware failures or cyberattacks.

**3. Replications**

**Definition::** Replication involves copying data in real-time or near-real-time from one location to another for redundancy or high availability.

**Key Features:**

* **Continuous Process:** Updates the replica as changes are made to the source.
* **High Availability:** Ensures minimal downtime by providing a secondary active copy.
* **Synchronous or Asynchronous:**
  + **Synchronous:** Writes occur simultaneously on source and replica (ideal for local replication).
  + **Asynchronous:** Writes occur after the source updates (better for long-distance replication).

**Use Cases:**

* Business continuity during a failure.
* Disaster recovery with near-zero data loss (RPO).

**Comparison Table**

| **Feature** | **Snapshots** | **Backups** | **Replications** |
| --- | --- | --- | --- |
| **Purpose** | Point-in-time state | Data recovery & archiving | Redundancy & high availability |
| **Frequency** | Frequent, near-instant | Periodic (daily/weekly) | Continuous (real-time) |
| **Storage** | Shared with source | Separate media | Another device/location |
| **Retention** | Short-term | Long-term | Continuous updates |
| **Performance Impact** | Low (space-efficient) | Moderate (full copy) | High (real-time updates) |
| **Dependency** | Storage system | Independent | Storage & network |

**Choosing the Right Option**

* **Snapshots**: For fast, short-term recovery or testing.
* **Backups**: For long-term retention, historical data recovery, and compliance.
* **Replications**: For minimizing downtime and data loss in high-availability environments.

A Storage Area Network (SAN) is a specialized high-speed network that provides block-level storage to servers. It allows multiple servers to access shared storage devices, such as disk arrays and tape libraries. Here’s an overview of the internal hardware details of a SAN, accompanied by an explanation of its key components:

**Internal Hardware Components of SAN**

1. **Host Bus Adapter (HBA)**
   * A hardware component installed on servers, enabling communication with the SAN.
   * Converts server data into a format suitable for storage networks (e.g., Fibre Channel or iSCSI).
2. **SAN Switches**
   * Central devices in the SAN network that connect servers (HBAs) to storage devices.
   * Use high-speed protocols like Fibre Channel or Ethernet.
   * Enable zoning, isolating traffic for security and performance.
3. **Storage Controllers**
   * Found within storage arrays, they manage read/write operations between servers and the physical storage devices.
   * Provide redundancy and caching for better performance and reliability.
4. **Storage Arrays**
   * Hardware systems housing multiple disk drives or SSDs.
   * Provide storage for SAN and include RAID controllers for data redundancy.
5. **Interconnects (Cables)**
   * High-speed connections (e.g., Fibre Channel, Ethernet) between servers, switches, and storage arrays.
   * Maintain data integrity and low latency during transmission.
6. **Management Software**
   * Monitors and controls the SAN hardware.
   * Includes tools for provisioning storage, setting up zoning, and ensuring optimal performance.

**SAN Internal Hardware Architecture**

Here’s a conceptual diagram:

1. **Servers (Hosts)**
   * Run applications needing storage.
2. **HBAs**
   * Installed in each server for SAN connectivity.
3. **SAN Switches**
   * Provide connectivity between servers and storage arrays.
4. **Storage Arrays**
   * Contain multiple drives for data storage.

**Direct-Attached Storage (DAS)** and **Network-Attached Storage (NAS)** are two different approaches to data storage. Here's a comparison:

| **Aspect** | **DAS (Direct-Attached Storage)** | **NAS (Network-Attached Storage)** |
| --- | --- | --- |
| **Connection** | Directly attached to a single server or computer (e.g., via USB, SATA, SAS). | Connected to a network, accessible by multiple devices. |
| **Access** | Accessible only by the attached device. | Accessible over a network by multiple devices. |
| **Usage** | Ideal for single-user environments or specific servers. | Ideal for file sharing and multi-user environments. |
| **Protocol** | No network protocols; uses local file systems (e.g., NTFS, ext4). | Uses network protocols like SMB/CIFS, NFS, or FTP. |
| **Scalability** | Limited scalability, dependent on server hardware. | Highly scalable by adding more storage units to the network. |
| **Performance** | High performance for direct workloads. | May have lower performance due to network latency. |
| **Management** | Managed locally through the attached device. | Centralized management for multiple users. |
| **Cost** | Generally lower upfront cost. | Higher cost due to networking and shared access capabilities. |
| **Backup/Recovery** | Backup requires external tools or solutions. | Often comes with built-in backup and recovery features. |

**Key Use Cases**:

* **DAS**: Single server applications, direct data access for high performance, small-scale setups.
* **NAS**: Centralized file sharing, collaborative work environments, home or small business storage.

**Redundancy in Storage** refers to the implementation of additional resources or mechanisms to ensure **data availability and protection against failures**. It provides a fail-safe to minimize the risk of data loss or downtime by duplicating critical components or data within a storage system.

**Key Concepts of Storage Redundancy:**

1. **Data Redundancy**:
   * Storing copies of data across multiple locations (disks, systems, or sites).
   * Commonly achieved using RAID, replication, or snapshots.
2. **Hardware Redundancy**:
   * Duplicating critical hardware components (e.g., power supplies, network interfaces, storage controllers).
   * Ensures the system remains operational if one component fails.
3. **Path Redundancy**:
   * Using multiple communication paths (e.g., Fibre Channel or Ethernet links) between servers and storage systems.
   * Provides high availability and prevents single points of failure.

**Techniques for Redundancy:**

1. **RAID (Redundant Array of Independent Disks)**:
   * Combines multiple physical disks into logical units for redundancy and performance.
   * Common RAID levels:
     + RAID 1: Mirroring (exact copy of data on another disk).
     + RAID 5: Striping with parity (data spread across disks with parity for recovery).
     + RAID 6: Striping with dual parity (tolerates two disk failures).
     + RAID 10: Combination of striping and mirroring.
2. **Data Replication**:
   * Copies data to a different location, either within the same storage system or to a remote site.
   * Types:
     + **Synchronous Replication**: Immediate copy, ensuring consistency but requiring low latency.
     + **Asynchronous Replication**: Copy occurs with a delay, suitable for distant sites.
3. **Snapshots**:
   * Point-in-time copies of data stored within the same system.
   * Used for quick recovery without requiring full duplication.
4. **Storage Cluster Redundancy**:
   * Groups of storage systems work together to replicate data and provide failover capabilities.
   * Often used in hyper-converged or distributed storage solutions.

**Advantages of Storage Redundancy:**

* **Data Availability**: Ensures continuous access to data even during failures.
* **Fault Tolerance**: Protects against hardware or software malfunctions.
* **Disaster Recovery**: Provides mechanisms for quick recovery in case of catastrophic failures.
* **Operational Continuity**: Reduces downtime, maintaining productivity.

**Challenges:**

* **Cost**: Additional hardware, software, and storage increase expenses.
* **Complexity**: Managing redundant systems and configurations can be challenging.
* **Performance Overhead**: Some redundancy methods (e.g., RAID 6) may slightly affect write performance.

**Use Case Example:**

* **Enterprise Databases**: Use RAID 10 for fast read/write with high redundancy.
* **Virtualized Environments**: Employ storage clusters with replication for high availability.
* **Backup Systems**: Implement snapshot-based redundancy for quick recovery points.

Redundancy ensures that storage systems remain reliable and resilient in the face of failures, safeguarding critical data and operations.

LUN (Logical Unit Number) storage is a key concept in storage area networks (SANs) and other block-level storage systems. It represents a logical unit of storage that is created from a physical or virtual disk and is presented to a host system. The host system interacts with the LUN as if it were a physical disk, enabling efficient data management and storage allocation. Here's a breakdown of the concept:

**Key Components of LUN Storage**

1. **Logical Abstraction**:  
   A LUN is a logical representation of a storage unit. It abstracts the physical characteristics of the underlying storage medium, allowing for flexibility and scalability.
2. **SAN Environment**:  
   In a SAN, storage devices (like disk arrays) are connected to servers. LUNs are used to carve out specific portions of storage from these devices and allocate them to servers.
3. **Block-Level Storage**:  
   LUNs provide block-level storage access, meaning data is stored and retrieved in fixed-sized blocks. This is ideal for databases, virtualization, and high-performance applications.
4. **Provisioning**:  
   Storage administrators can define LUNs with specific sizes and characteristics. They can then assign these LUNs to one or more hosts depending on the requirements.
5. **Multipathing and Access**:  
   LUNs often support multipathing, enabling multiple routes for data access to ensure redundancy and reliability.
6. **Types of LUNs**:
   * **Thick-Provisioned LUNs**: Allocates the entire storage space upfront.
   * **Thin-Provisioned LUNs**: Dynamically allocates storage as needed to optimize utilization.

**Advantages of LUN Storage**

* **Resource Optimization**: Allows fine-grained allocation of storage resources.
* **Scalability**: Easy to resize or allocate additional storage as needed.
* **Flexibility**: Hosts can access specific portions of storage without interference.
* **Performance**: Block-level access provides high-speed data transfer.
* **Security and Isolation**: Specific LUNs can be isolated to certain hosts to ensure data security.

**Use Cases**

* **Virtualization**: Assigning storage to virtual machines in a hypervisor environment.
* **Databases**: Storing and managing large volumes of structured data.
* **Backups**: Dedication of LUNs for backup operations to ensure data integrity.
* **High-Performance Computing**: Supporting workloads requiring rapid and frequent data access.

LUN storage is foundational to modern enterprise storage systems, providing the flexibility and efficiency needed for diverse workloads.

Thick and Thin LUNs refer to different provisioning methods for allocating storage space in a storage system. These approaches determine how storage is reserved and used within a storage array or system. Here's a detailed explanation of both:

**Thick-Provisioned LUN**

**Definition:**

A **Thick-Provisioned LUN** allocates the full amount of storage space specified during its creation upfront, regardless of whether the host system or application uses it immediately.

**Key Characteristics:**

1. **Fixed Allocation**:  
   The entire requested size is reserved on the physical storage from the start. For example, if you create a 500 GB LUN, 500 GB is immediately set aside, even if only 100 GB is initially used.
2. **Guaranteed Space**:  
   Ensures that the allocated space will always be available for the application or host, which is useful for predictable workloads.
3. **Performance**:  
   Since the storage space is preallocated, there is no overhead related to dynamically assigning additional space, which can lead to better performance consistency.
4. **Storage Utilization**:  
   Can lead to unused space, especially if the host or application does not consume the allocated storage, resulting in less efficient utilization.
5. **Use Cases**:
   * Mission-critical applications where predictable performance is required.
   * Workloads with a fixed, known storage requirement.

**Thin-Provisioned LUN**

A **Thin-Provisioned LUN** allocates storage space dynamically as it is needed, rather than reserving it upfront.

**Key Characteristics:**

1. **Dynamic Allocation**:  
   Space is allocated on demand when data is written. For example, creating a 500 GB thin LUN does not consume 500 GB immediately; space is assigned incrementally as data grows.
2. **Efficient Utilization**:  
   Thin provisioning allows for over-commitment, where the total virtual storage assigned to hosts or applications exceeds the physical storage capacity.
3. **Overhead**:  
   Writing new data to a thin LUN may introduce performance overhead, as the system needs to allocate space dynamically.
4. **Risk of Overcommitment**:  
   If over-commitment is not managed properly and the physical storage becomes full, there is a risk of application failure.
5. **Use Cases**:
   * Environments with variable or unpredictable storage growth.
   * Test and development systems where full storage usage is not guaranteed.
   * Virtualized environments to optimize shared resources.

**Comparison of Thick and Thin LUNs**

| **Feature** | **Thick LUN** | **Thin LUN** |
| --- | --- | --- |
| **Allocation Method** | Fixed upfront | Dynamic on demand |
| **Space Utilization** | Less efficient | More efficient |
| **Performance** | Consistent and predictable | May have slight overhead |
| **Risk** | No risk of over-commitment | Over-commitment risk if not managed |
| **Storage Management** | Simpler | Requires monitoring and management |
| **Typical Use Cases** | Critical applications, databases | Virtualized environments, testing |

**Choosing Between Thick and Thin LUNs**

* Use **Thick LUNs** for workloads requiring guaranteed performance and predictable storage needs.
* Use **Thin LUNs** to maximize storage efficiency, especially in environments with variable or unknown storage requirements.

By understanding the trade-offs, organizations can optimize their storage infrastructure for both performance and efficiency.

**50 TOP SAN – Storage Area Networks Questions and Answers**

[**SAN Interview Questions and Answers**](http://saninterviewquestionspdf.blogspot.com/2017/01/50-top-san-storage-area-networks.html)**List**

**1. What is a**[**SAN**](http://saninterviewquestionspdf.blogspot.com/2017/01/50-top-san-storage-area-networks.html)**?**  
SAN is short for Storage Area Network. It is a high-speed network of storage elements, similar in form and function to a LAN that establishes direct and indirect connections between multiple servers and multiple storage elements. The SAN is an extension of the server’s storage bus  
  
**2. What does a**[**SAN**](http://saninterviewquestionspdf.blogspot.com/2017/01/50-top-san-storage-area-networks.html)**do?**  
SANs create connectivity. SANs offer a method of attaching storage that improves data reliability, availability and performance  
SAN overcomes traditional network bottlenecks by connecting in three ways:  
· Server-to-storage (direct attached storage)  
· Server-to-server (network attached storage)  
· Storage-to-storage (SAN Attached Storage)  
  
**3. Name some of the**[**SAN**](http://saninterviewquestionspdf.blogspot.com/2017/01/50-top-san-storage-area-networks.html)**topologies and Explain each of them ?**  
Point-to-point, arbitrated loop, and switched fabric topologies  
a) Point-to-Point  
A point-to-point connection is the simplest topology. It is used when there are exactly two nodes and future expansion is not predicted. There is no sharing of the media, which allows the devices to use the total bandwidth of the link. A simple link initialization is needed before communications can begin.  
b) Arbitrary Loop  
Our second topology is Fiber Channel Arbitrated Loop (FC-AL). FC-AL is more useful for storage applications. It is a loop of up to 126 nodes (NL\_Ports) that is managed as a shared bus. Traffic flows in one direction, carrying data frames and primitives around the loop with a total bandwidth of 400 MBps (or 200 MBps for a loop based on 2 Gbps technology).  
c) Switched Fabric Loop  
It applies to switches and directors that support the FC-SW standard, that is, it is not limited to switches as its name suggests. A Fibre Channel fabric is one or more fabric switches in a single, sometimes extended, configuration. Switched fabrics provide full bandwidth per port compared to the shared bandwidth per port in arbitrated loop  
Implementations.  
  
**4. What’s the need for separate network for storage why LAN cannot be used?**  
LAN hardware and operating systems are geared to user traffic, and LANs are tuned for a fast user response to messaging requests. With a SAN, the storage units can be secured separately from the servers and totally apart from the user network enhancing storage access in data blocks (bulk data transfers), advantageous for server-less backups.  
  
**5. What is FCP?**  
The Fibre Channel Protocol (FCP) is the interface protocol of SCSI on Fibre Channel. It is a gigabit speed network technology primarily used for Storage Networking. Fibre Channel is standardized in the T11 Technical Committee of the InterNational Committee for Information Technology Standards (INCITS), an American National Standard Institute (ANSI) accredited standards committee. It started for use primarily in the supercomputer field, but has become the standard connection type for storage area networks in enterprise storage. Despite its name, Fibre Channel signaling can run on both twisted-pair copper wire and fiber optic cables.  
  
**6.What is iSCSI ?**  
Internet SCSI (iSCSI) is a transport protocol that carries SCSI commands from an initiator to a target. It is a data storage networking protocol that transports standard Small Computer System Interface (SCSI) requests over the standard Transmission Control Protocol/Internet Protocol (TCP/IP) networking technology.  
iSCSI enables the implementation of IP-based storage area networks (SANs), enabling customers to use the same networking technologies — for both storage and data networks. As it uses TCP/IP, iSCSI is also well suited to run over almost any physical network. By eliminating the need for a second network technology just for storage, iSCSI has the potential to lower the costs of deploying networked storage.  
  
**7.What is FCIP ?**  
Fibre Channel over IP (FCIP) is also known as Fibre Channel tunneling or storage tunneling. It is a method to allow the transmission of Fibre Channel  
information to be tunnelled through the IP network. FCIP encapsulates Fibre Channel block data and subsequently transports it over a TCP socket. TCP/IP services are utilized to establish connectivity between remote SANs. Any congestion control and management, as well as data error and data loss recovery, is handled by TCP/IP services, and does not affect FC fabric services. The major point with FCIP is that is does not replace FC with IP, it simply allows deployments of FC fabrics using IP tunneling  
  
**8. What is iFCP**  
Internet Fibre Channel Protocol (iFCP) is a mechanism for transmitting data to and from Fibre Channel storage devices in a SAN, or on the Internet using TCP/IP. iFCP gives the ability to incorporate already existing SCSI and Fibre Channel networks into the Internet. iFCP is able to be used in tandem with existing Fibre Channel protocols, such as FCIP, or it can replace them. Whereas FCIP is a tunneled solution, iFCP is an FCP routed solution.iFCP is a gateway-to-gateway protocol, and does not simply encapsulate FC block data. Gateway devices are used as the medium between the FC initiators and targets. As these gateways can either replace or be used in tandem with existing FC fabrics, iFCP could be used to help migration from a Fibre  Channel SAN to an IP SAN, or allow a combination of both  
  
**9. What is FICON address ?**  
FICON generates the 24-bit FC port address field in yet another way. When communication is required from the FICON channel port to the FICON CU port,  
the FICON channel (using FC-SB-2 and FC-FS protocol information) will provide both the address of its port, the source port address identifier (S\_ID), and the address of the CU port, the destination port address identifier (D\_ID) when the communication is from the channel N\_Port to the CU N\_Port.  
  
**10. What is zoning?**  
Fabric management service that can be used to create logical subsets of devices within a SAN. This enables portioning of resources for management and access control purpose.

**11. What are the two major classification of zoning?**  
Two types of zoning are  
a) Software Zoning  
b) Hardware Zoning  
  
**12. What are different levels of zoning?**  
a) Port Level zoning  
b) WWN Level zoning  
c) Device Level zoning  
d) Protocol Level zoning  
e) LUN Level zoning  
  
**13. What is FICON ?**  
FICON is a protocol that uses Fibre Channel as its physical medium. FICON channels are capable of data rates up to 200 MBps full duplex, they extend the channel distance (up to 100 km), increase the number of control unit images per link, increase the number of device addresses per control unit link, and retain the topology and switch management characteristics of ESCON.  
  
**14. What is FSPF ?**  
FSPF keeps track of the links on all switches in the fabric and associates a cost with each link. The cost is always calculated as being directly proportional to the number of hops. The protocol computes paths from a switch to all other switches in the fabric by adding the cost of all links traversed by the path, and choosing the path that minimizes the cost.  
  
**15. How FSPF works**  
The collection of link states (including cost) of all switches in a fabric constitutes the topology database (or link state database). The topology database is kept in all switches in the fabric, and they are maintained and synchronized to each other. There is an initial database synchronization, and an update mechanism.  
82 Introduction to Storage Area Networks .The initial database synchronization is used when a switch is initialized, or when an ISL comes up. The update mechanism is used when there is a link state change. This ensures consistency among all switches in the fabric.  
  
**16. What is Network Attached Storage (NAS) ?**  
Network Attached Storage (NAS) is basically a LAN-attached file server that serves files using a network protocol such as Network File System (NFS). NAS is a term used to refer to storage elements that connect to a network and provide file access services to computer systems. A NAS storage element consists of an engine that implements the file services (using access protocols such as NFS or CIFS), and one or more devices, on which data is stored. NAS elements may be attached to any type of network. From a SAN perspective, a SAN-attached NAS engine is treated just like any other server, but a NAS does not provide any of the activities that a server in a server-centric system typically provides, such as e-mail, authentication, or file management.  
  
**17. How is Fiber Channel Different from iSCSI?**  
Fibre Channel and iSCSI each have a distinct place in the IT infrastructure as SAN alternatives to DAS. Fibre Channel generally provides high performance and high availability for business-critical applications, usually in the corporate data center. In contrast, iSCSI is generally used to provide SANs for business applications in smaller regional or departmental data centers.  
  
**18. What is Frames?**  
Fibre Channel places a restriction on the length of the data field of a frame at 528 transmission words, which is 2112 bytes. (See Table 3-2 on page 52.) Larger amounts of data must be transmitted in several frames. This larger unit that consists of multiple frames is called a sequence. An entire transaction between two ports is made up of sequences administered by an even larger unit called an exchange.  
A frame consists of the following elements:  
\_ SOF delimiter  
\_ Frame header  
\_ Optional headers and payload (data field)  
\_ CRC field  
\_ EOF delimiter  
  
**19. What is Loop address ?**  
An NL\_Port, like an N\_Port, has a 24-bit port address. If no switch connection exists, the two upper bytes of this port address are zeroes (x’00 00’) and referred to as a private loop. The devices on the loop have no connection with the outside world. If the loop is attached to a fabric and an NL\_Port supports a fabric login, the upper two bytes are assigned a positive value by the switch. We call this mode a public loop.  
  
**20. What is LUN?**  
LUN unique number that is assigned to each storage device or partition of the storage that the storage can support.

**21. What is LUN Masking?**  
A method used to create an exclusive storage area and access control. And this can be achieved by storage device control program.  
  
**22. What is WWN?**  
WWN is a 64bit address that is hard coded into a fiber channel HBA and this is used to identify individual port (N\_Port or F\_Port) in the fabric.  
  
**23. What is metaLUN?**  
A metaLUN is a type of LUN whose maximum capacity can be the combined capacities of all the LUNs that compose it. The metaLUN feature lets you dynamically expand the capacity of a single LUN (base          LUN) into a larger unit called a metaLUN. You do this by adding LUNs to the base LUN. You can also add LUNs to a metaLUN to further increase its capacity. Like a LUN, a metaLUN can belong to Storage Group, and can participate in Snap View, Mirror View and SAN copy sessions. MetaLUNs are supported only on CX-Series storage systems. A metaLUN may include multiple sets of LUNs and each set of LUNs is called a component. The LUNs within a component are striped together and are independent of other LUNs in the metaLUN.  
  
**24. EMC control center 5.1**  
The EMC Control Center Web Console uses [Port : 10799] data stored in the Repository to monitor your storage-attached network and manageControlCenter alerts remotely through a Web browser  
  
**25. What Is Emc Power path?**  
EMC PowerPath is a server-resident software solution that enhances performance and information availability. It integrates multiple path I/O capabilities, automatic load balancing, and path failover      functions into one comprehensive package for use on open server platforms connected to Symmetrix enterprise storage systems. PowerPath enables you to do more work in a shorter time so you can serve             more customers, run more applications, and exploit more business opportunities.  
  
**26. What is a HBA?**  
Host bus adapters (HBAs) are needed to connect the server (host) to the storage.  
  
**27. What is SAN fabric?**  
SAN fabric is a hardware device that connects workstations and servers to storage devices in a SAN network. It uses the Fibre Channel switching technology to connect a server to a storage device. The SAN fabric offers a high-speed dedicated network including high availability features, very low latency, and high throughput.  
  
**28. What is zoning?**  
Fabric management service that can be used to create logical subsets of devices within a SAN. This enables portioning of resources for management and access control purpose.  
  
**29. What are the two major classification of zoning?**  
Two types of zoning are  
a) Software Zoning  
b) Hardware Zoning  
  
**30. What are different levels of zoning?**  
a) Port Level zoning  
b) WWN Level zoning  
c) Device Level zoning  
d) Protocol Level zoning  
e) LUN Level zoning

**31. What is storage virtualization?**  
Storage virtualization is amalgamation of multiple n/w storage devices into single storage unit.  
  
**32. What is virtualization?**  
A technique of hiding the physical characteristics of computer resources from the way in which other system application or end user interact with those resources. Aggregation, spanning or concatenation of the combined multiple resources into larger resource pools.  
  
**33. What is Multipath I/O?**  
Fault tolerant technique where, there is more than one physical path between the CPU in the computer systems and its main storage devices through the buses, controllers, switches and other bridge devices connecting them.  
  
**34. What are the 3 prominent characteristics of SAS Protocol?**  
a) Native Command Queuing (NCQ)  
b) Port Multiplier  
c) Port Selector  
  
**35. What is the purpose of disk array?**  
Probability of unavailability of data stored on the disk array due to single point failure is totally eliminated.  
  
**36. What is disk array?**  
Set of high performance storage disks that can store several terabytes of data. Single disk array can support multiple points of connection to the network.  
  
**37. What are the advantages of RAID?**  
**“Redundant Array of Inexpensive Disks”**  
Depending on how we configure the array, we can have the  
- data mirrored [RAID 1] (duplicate copies on separate drives)  
- striped [RAID 0] (interleaved across several drives), or  
- parity protected [RAID 5](extra data written to identify errors).  
These can be used in combination to deliver the balance of performance and reliability that the user requires.  
  
**38. How is a SAN managed?**  
There are many management software’s used for managing SAN's to name a few Santricity  
- IBM Tivoli Storage Manager.  
- CA Unicenter.  
- Veritas Volumemanger.  
  
**39. Which one is the Default ID for SCSI HBA?**  
Generally the default ID for SCSI HBA is 7.  
SCSI- Small Computer System Interface  
HBA - Host Bus Adaptor  
  
**40. What is the highest and lowest priority of SCSI?**  
There are 16 different ID’s which can be assigned to SCSI device 7, 6, 5, 4, 3, 2, 1, 0, 15, 14, 13, 12, 11, 10, 9, 8.  
Highest priority of SCSI is ID 7 and lowest ID is 8.

**41. what is SRDF ?**  
RDF (Symmetrix Remote Data Facility) is a family of EMC products that facilitates the data replication from one Symmetrix storage array to another through a Storage Area Network or IP network.SRDF logically pairs a device or a group of devices from each array and replicates data from one to the other synchronously or asynchronously. An established pair of devices can be split, so that separate hosts can access the same data independently (maybe for backup), and then resynchronized.  
In synchronous mode (SRDF/S), the primary array waits until the secondary array has acknowledged each write before the next write is accepted, ensuring that the replicated copy of the data is always as current as the primary. However, the latency due to propagation increases significantly with distance.  
  
Asynchronous SRDF (SRDF/A) transfers changes to the secondary array in units called delta sets, which are transferred at defined intervals. Although the remote copy of the data will never be as      current as the primary copy, this method can replicate data over considerable distances and with reduced bandwidth requirements and minimal impact on host performance. Other forms of SRDF exist to integrate with clustered environments and to manage multiple SRDF pairs where replication of multiple devices must be consistent (such as with the data files and log files of a database application).  
  
**42. Define RAID? Which one you feel is good choice?**  
RAID (Redundant array of Independent Disks) is a technology to achieve redundancy with faster I/O. There are Many Levels of RAID to meet different needs of the customer which are: R0, R1, R3, R4, R5, R10, R6. Generally customer chooses R5 to achieve better redundancy and speed and it is cost effective.  
  
R0 – Striped set without parity/[Non-Redundant Array].  
  Provides improved performance and additional storage but no fault tolerance. Any disk failure destroys the array, which becomes more likely with more disks in the array. A single disk failure destroys the entire array because when data is written to a RAID 0 drive, the data is broken into fragments. The number of fragments is dictated by the number of disks in the drive. The fragments are written to their respective disks simultaneously on the same sector. This allows smaller sections of the entire chunk of data to be read off the drive in parallel, giving this type of arrangement huge bandwidth.  
RAID 0 does not implement error checking so any error is unrecoverable. More disks in the array means higher bandwidth, but greater risk of data loss  
  
R1 - Mirrored set without parity.  
 Provides fault tolerance from disk errors and failure of all but one of the drives. Increased read performance occurs when using a multi-threaded operating system that supports split seeks, very small performance reduction when writing. Array continues to operate so long as at least one drive is functioning. Using RAID 1 with a separate controller for each disk is sometimes called duplexing.  
  
R3 - Striped set with dedicated parity/Bit interleaved parity.  
 This mechanism provides an improved performance and fault tolerance similar to RAID 5, but with a dedicated parity disk rather than rotated parity stripes. The single parity disk is a bottle-neck for writing since every write requires updating the parity data. One minor benefit is the dedicated parity disk allows the parity drive to fail and operation will continue without parity or performance penalty.  
  
R4 - Block level parity.  
 Identical to RAID 3, but does block-level striping instead of byte-level striping. In this setup, files can be distributed between multiple disks. Each disk operates independently which allows I/O requests to be performed in parallel, though data transfer speeds can suffer due to the type of parity. The error detection is achieved through dedicated parity and is stored in a separate, single disk unit.  
  
R5 - Striped set with distributed parity.  
 Distributed parity requires all drives but one to be present to operate; drive failure requires replacement, but the array is not destroyed by a single drive failure. Upon drive failure, any subsequent reads can be calculated from the distributed parity such that the drive failure is masked from the end user. The array will have data loss in the event of a second drive failure and is vulnerable until the data that was on the failed drive is rebuilt onto a replacement drive.  
  
R6 - Striped set with dual distributed Parity.  
 Provides fault tolerance from two drive failures; array continues to operate with up to two failed drives. This makes larger RAID groups more practical, especially for high availability systems. This becomes increasingly important because large-capacity drives lengthen the time needed to recover from the failure of a single drive. Single parity RAID levels are vulnerable to data loss until the failed drive is rebuilt: the larger the drive, the longer the rebuild will take. Dual parity gives time to rebuild the array without the data being at risk if one drive, but no more, fails before the rebuild is complete.  
  
**43. What is the different between mirroring, Routing and multipathing?**  
Redundancy Functions Relationships Role Mirroring Generates 2 ios to 2 storage targets Creates 2 copies of data Routing Determined by switches independent of SCSI Recreates n/w route after a Failure Multipathing Two initiator to one target Selects the LUN initiator pair to use.  
  
**44.  Briefly list the advantages of SAN?**  
SANs fully exploit high-performance, high connectivity network technologies  
SANs expand easily to keep pace with fast growing storage needs  
SANs allow any server to access any data  
SANs help centralize management of storage resources  
SANs reduce total cost of ownership (TCO).  
  
iSCSI fundamentals:-  
iSCSI is a protocol defined by the Internet Engineering Task Force (IETF) which enables SCSI commands to be encapsulated in TCP/IP traffic, thus allowing access to remote storage over low cost IP networks.  
  
  
**45. What advantages would using an iSCSI Storage Area Network (SAN) give to your organization over using Direct Attached Storage (DAS) or a Fibre Channel SAN?**  
iSCSI is cost effective, allowing use of low cost Ethernet rather than expensive Fibre architecture.  
· Traditionally expensive SCSI controllers and SCSI disks no longer need to be used in each server, reducing overall cost.  
· Many iSCSI arrays enable the use of cheaper SATA disks without losing hardware RAID functionality.  
· The iSCSI storage protocol is endorsed by Microsoft, IBM and Cisco, therefore it is an industry standard.  
· Administrative/Maintenance costs are reduced.  
· Increased utilisation of storage resources.  
· Expansion of storage space without downtime.  
· Easy server upgrades without the need for data migration.  
· Improved data backup/redundancy.  
  
**46. How many minimum drives are required to create R5 (RAID 5) ?**  
You need to have at least 3 disk drives to create R5.  
  
**47.  What are the advantages of SAN?**  
Massively extended scalability  
Greatly enhanced device connectivity  
Storage consolidation  
LAN-free backup  
Server-less (active-fabric) backup  
Server clustering  
Heterogeneous data sharing  
Disaster recovery - Remote mirroring  
While answering people do NOT portray clearly what they mean & what advantages each of them have, which are cost effective & which are to be used for the client's requirements.  
  
**48. What is the difference b/w SAN and NAS?**  
The basic difference between SAN and NAS, SAN is Fabric based and NAS is Ethernet based.  
SAN - Storage Area Network  
It accesses data on block level and produces space to host in form of disk.  
NAS - Network attached Storage  
It accesses data on file level and produces space to host in form of shared network folder.  
  
**49. What is a typical storage area network consists of - if we consider it for implementation in a small business setup?**  
 If we consider any small business following are essentials components of SAN  
- Fabric Switch  
- FC Controllers  
- JBOD's  
  
**50. Can you briefly explain each of these Storage area components?**  
Fabric Switch: It's a device which interconnects multiple network devices .There are switches starting from 16 port to 32 ports which connect 16 or 32 machine nodes etc. vendors who manufacture these kind of switches are Brocade, McData.  
FC Controllers: These are Data transfer media they will sit on PCI slots of Server; you can configure Arrays and volumes on it.  
JBOD: Just Bunch of Disks is Storage Box, it consists of Enclosure where set of hard-drives are hosted in many combinations such SCSI drives, SAS, FC, SATA.