DHANALAKSHMI SRINIVASAN ENGINEERING COLLEGE
(AUTONOMOUS)
(Approved by AICTE & Affiliated to Anna University, Chennai)
Accredited with 'A' Grade by NAAC, Accredited by TCS
Accredited by NBA with BME, ECE & EEE
PERAMBALUR - 621 212. Tamil Nadu.
website : www.dsengg.ac.in

U20CS701 / INFORMATION STORAGE MANAGEMENT

Part A & B

UNIT I- STORAGE SYSTEMS

Part A

1. **Define Information Storage and Management?**

   Information Storage and Management (ISM) refers to the discipline and practices involved in effectively managing and storing digital information within an organization. It encompasses the strategies, technologies, and processes used to capture, store, organize, protect, and retrieve data throughout its lifecycle.

2. **What is Information Storage?**

   Organizations process data to derive the information required for their day-to-day operations. Storage is a repository that enables users to persistently store and retrieve this digital data.

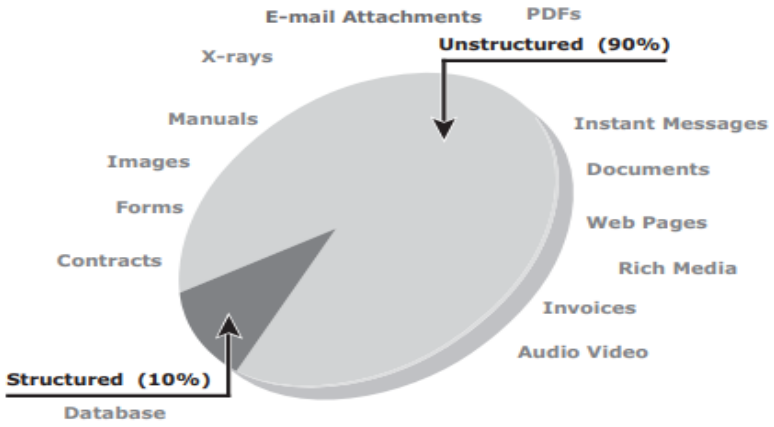
3. **Define data and information. (NOV/DEC 2017) (NOV/DEC 2015)**

   Data is a collection of raw facts. Handwritten letters, a printed book, a family photograph, a movie on video tape, printed and duly signed copies of mortgage papers, a bank's ledgers, and an account holder's passbooks are all examples of data.

   Information is the intelligence and knowledge derived from data. Effective data analysis not only extends its benefits to existing businesses, but also creates the potential for new business opportunities by using the information in creative ways.

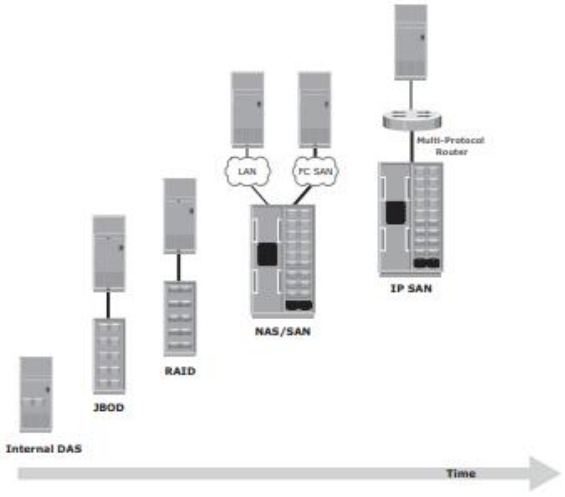
4. **List some of the factors that have contributed to the growth of digital data.**
   - **a.** Increase in data-processing capabilities
   - **b.** Lower cost of digital storage
   - **c.** Affordable and faster communication technology
   - **d.** Proliferation of applications and smart devices

5. **Recall types of data.**
   - Data can be classified as structured or unstructured based on how it is stored and managed.
   - Structured data is organized in rows and columns in a rigidly defined format so that applications can retrieve and process it efficiently.
   - Data is unstructured if its elements cannot be stored in rows and columns, which makes it difficult to query and retrieve by applications.



6. State the evolution of Storage Technology and Architecture.

- Redundant Array of Independent Disks (RAID)
- Direct-attached storage (DAS)
- Storage area network (SAN)
- Network-attached storage (NAS)
- Internet Protocol SAN (IP-SAN)



7. What are the limitations of data center

- Companies hiring the resources from the data center providers will not have complete control locally.
- This is due to the fact that the human resources and hardware lies at the remote location.
- The usage and quality of data center services vary based on internet connectivity at the customer premises.
- There are limitations on security features offered by data center service providers.
- Some companies charge technical support fees from customers.
- In case of troubleshooting and issues, customers have to rely on support staff from data center providers. Hence resolution of the same depends on skills and knowledge of the support staff.

8. List the key characteristics of data center.

- Location
- Physical Security
- Network Redundancy
- Financial Stability
- Scalability
- Disaster Recovery Plan

- Think to the Future
- 24/7 Support and Monitoring
- Operational Performance
- Price

9. Point out the benefits of information centric storage architecture.

Data becomes independent from location, application, storage, and means of transportation, enabling in-network caching and replication. The expected benefits are improved efficiency, better scalability with respect to information/bandwidth demand and better robustness in challenging communication scenarios.

10. Define Big Data Analytics.

Big data analytics is the use of advanced analytic techniques against very large, diverse big data sets that include structured, semi-structured and unstructured data, from different sources, and in different sizes from terabytes to zettabytes.

11. Illustrate the structure of information centric storage architecture

The challenges of storingand managing unstructured data.Discuss the benefits of information-centric storage architecture over server-centric storage architectureWhat are the attributes of big data? Research and prepare presentation on big data analytics. Research how businesses use their information assets to derive competitive advantage and new business opportunities. Research and prepare a presentation on personal data management.

12. Classify the various types of cloud storage

- Google Drive,
- OneDrive,
- Dropbox, Box and
- Amazon aren't your only cloud storage options.

13. Distinguish between IaaS and PaaS in cloud

Basis	IAAS	PAAS
Stands for	Infrastructure as a service.	Platform as a service.
Uses	IAAS is used by network architects.	PAAS is used by developers.
Access	IAAS gives access to the resources like virtual machines and virtual storage.	PAAS gives access to run time environment to deployment and development tools for application.
Model	It is a service model that provides virtualized computing resources over the internet.	It is a cloud computing model that delivers tools that are used for the development of applications.
Technical understanding.	It requires technical knowledge.	Some knowledge is required for the basic setup.
Enterprise services.	AWS virtual private cloud.	Microsoft Azure.

14. Give the characteristics of file system

- Arranging the files and folders hierarchically.
- Report generation.
- Notes.
- Status.

- Assigning documents for processing in a queue.
- Add or edit metadata of files.
- Create, modify, delete, or manage other file operations.
- Simple interact to access and manage files.

15. Show the schematic operation of SCSI

- It must be configured for a specific computer.
- It has limited system BIOS support.
- Its variations (speeds, connectors) can be bewildering.
- There is no common software interface.
- It's fast -- up to 160 megabytes per second (MBps).
- It's reliable.
- It allows you to put multiple devices on one bus.
- It works on most computer systems.

16. Classify Disk Service time.

The list of various disks scheduling algorithm is given below. Each algorithm is carrying some advantages and disadvantages. The limitation of each algorithm leads to the evolution of a new algorithm.

- FCFS scheduling algorithm
- SSTF (shortest seek time first) algorithm
- SCAN scheduling
- C-SCAN scheduling
- LOOK Scheduling
- C-LOOK scheduling

17. Compose a rotational latency

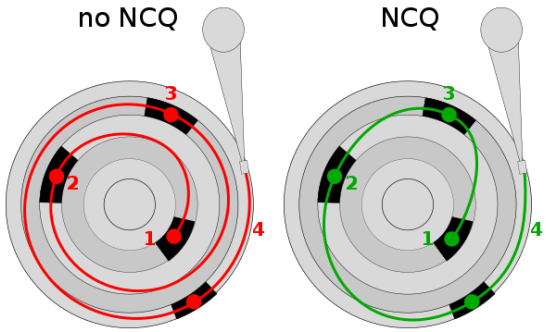
Rotational latency (in milliseconds) describes the time required to position a specific sector under the read–write head. Average latency is typically given as the time it takes the drive to perform half a rotation of the platter, and is directly dependent on its RPM rating.

18. Assess the limitation in terms of performance and management in storage device.

- Over or under exaggerating recent performances
- An unstructured process in place
- Lack of rewards and/or recognition
- No annual performance evaluation
- Inconsistent feedback, goals and rewards

19. Develop the disk command queuing with suitable diagram

(NCQ) is a technology designed to increase performance of SATA hard disks on computers. It can do this under certain situations by allowing the individual hard disk to internally optimize the order in which it executes the read and write requests it received. This can reduce the amount of unnecessary drive head movement and resulting in better performance (and slightly decreased wear of the drive) for workloads where multiple read/write requests are outstanding at the same time.



20. Give a reason why interface protocols enable between host and storage.

- A protocol enables communication between the host and storage.
- Protocols are implemented using interface devices (or controllers) at both source and destination.
- The popular interface protocols used for host to storage communications are:

i. Integrated Device Electronics/Advanced Technology Attachment (IDE/ATA)
ii. Small Computer System Interface (SCSI), i
ii. Fibre Channel (FC) iv. Internet Protocol (IP)

Part B

1. **(i)Analyze the various types of storage system .(7)**
   (ii)Explain in detail about virtualization and cloud computing. (6)
2. **Describe in detail about Big Data ecosystem. (13)**
3. **Analyze the evolution of storage architecture with neat diagram. (13)**
4. **(i)Demonstrate the working principle of Platter.(7)**
   (ii) Demonstrate the working principle of Read/Write head.(6)
5. **(i)How does nested RAID used in storage system. (7)**
   (ii)Identify the features of flash drive. (6)
6. **(i)Explain the advantages of virtualized data center over a classic data center. (7)**
   (ii)Which component constitutes the disk service time. (3)
   (iii) Which components contribute the largest percentage of the disk service time in random I/O operations? (3)
7. **Differentiate between types of RAID levels. (13)**
8. **(i)Classify the two methods of RAID implementation. (7)**
   (ii)Examine the components of the RAID array. (6)
9. **Generalize the role of RAID techniques. (13)**
10. **(i)Examine the applications of IOPS and RAID configurations. (7)**
    (ii)Describe the Hot Spares. (6) .
11. **Explain the key challenges in managing information. (13)**
12. **Discuss the process of data recovery in case of a drive failure in RAID 5. (13)**
13. **Describe the impact of random and sequential I/Os in different RAID configurations. (13)**
14. **Summarize the benefits of using RAID 3 in backup application. (13)**
15. **Discuss: (i) Data transfer rate. (6)**
    (ii) Seek time. (7)
16. **What is structured and unstructured data? Examine The challenges of storing and managing unstructured data. (13)**
17. **(i) Describe the disk partition and concatenation. (6)**
    (ii) Explain the process of mapping user files to disk storage. (7)

UNIT II- STORAGE NETWORKING TECHNOLOGIES

Part A

1. **Define DAS.**

Direct-attached storage (DAS) is a type of storage that is attached directly to a computer without going through a network. The storage might be connected internally or externally. Only the host computer can access the data directly.

2. **What is meant by SCSI?**

A small computer systems interface (SCSI) is a standard interface for connecting peripheral devices to a PC. Depending on the standard, generally it can connect up to 16 peripheral devices using a single bus including one host adapter. SCSI is used to increase performance, deliver faster data transfer transmission and provide larger expansion for devices such as CD-ROM drives, scanners, DVD drives and CD writers. SCSI is also frequently used with RAID, servers, high-performance PCs and storage area networks SCSI has a controller in charge of transferring data between the devices and the SCSI bus.

3. **Express the idea of disk drive interfaces.**

The interface of an internal hard disk drive defines the way it connects to the rest of the host system, usually the computer's motherboard, and interacts with it. Hard disk drives are accessed over one of a number of bus types, including parallel ATA (PATA, also called IDE or EIDE; described before the introduction of SATA as ATA), Serial ATA (SATA), SCSI, Serial Attached SCSI (SAS), and Fibre Channel.

4. **What do you know about 'fiber channel'?**

Fibre Channel is a high-speed data transfer protocol designed for enabling storage area networks (SANs). Given that SANs are intended to connect blocks of data across storage devices to servers, Fibre Channel Protocol is optimized to transmit in-order, lossless raw block data with low latency.

5. **Predict the functions of interface protocol.**

The TCP/IP stack dynamically creates and activates an internal shared memory (ISM) interface for specific physical network ID (PNetID) when both of the following conditions are true:

- The first SMC-D capable IPAQENET, IPAQENET6, IPAQIDIO, or IPAQIDIO6 interface is started for that PNetID.
-
- An ISM device with the same PNetID value is found

6. **Differentiate between FC SAN and IP SAN.**

What is SAN?

The storage area network (SAN) is a high-speed network dedicated to storage operations. It is usually independent of the computer local area network (LAN). A SAN connects a host and a storage device and provides a dedicated communication channel for any host and any storage device on the SAN. SAN separates storage devices from servers to share storage resources at the server level. SANs bring channel and network technologies into storage environments, providing a new networked storage solution

that meets the requirements of throughput, availability, reliability, scalability, and manageability simultaneously.

What is IP SAN?

In brief, IP-SAN (IP storage) uses IP channels instead of Fibre Channel to connect servers and storage devices. In addition to iSCSI, standards such as FCIP and iFCP are being developed. iSCSI is the fastest growing network and has become a powerful representative of IP storage.

Like Fibre Channel, IP storage is interchangeable, but unlike Fibre Channel, IP networks are mature and do not have interoperability problems, which is the biggest headache for Fibre Channel SANs. IP has been widely recognized in the IT industry, and there are many network management software and service products available.

7. Formulate the modal dispersion.

Modal dispersion is a distortion mechanism occurring in multimode fibers and other waveguides, in which the signal is spread in time because the propagation velocity of the optical signal is not the same for all modes. Modal dispersion is a distortion mechanism occurring in multimode fibers and other waveguides, in which the signal is spread in time because the propagation velocity of the optical signal is not the same for all modes.

8. Show the multi-mode fiber and single mode fiber.

The major difference single-mode and multimode optical fiber is that in single-mode optical fiber light ray propagates only through a single path. On the contrary, in multimode optical fiber several light rays propagate through the waveguide at the same time.

9. Define point-to-point connectivity.

A Point to Point Connection is a private data connection securely connecting two or more locations for private data services. A point to point connection is a closed network data transport service which does not traverse the public Internet and is inherently secure with no data encryption needed.

10. List the FC-AL configuration limitations in terms of scalability.

- Node Ports
- Cables and Connectors
- Interconnect Devices
- SAN Management Software
- Fibre Channel Protocol (FCP)

11. Infer the advantage of FCP.

- FCP has high performance
- provides good backup and restoration and simplified consolidation
- also offers congestion-free data flow, Gigabit bandwidth, compatibility with multiple topologies and protocols, flow control, and self-management
- It is providing high-speed data transfer
- It is cost-efficient
- High-speed data can be transferred over a distance of 10km
- supports several fault-tolerant features

- FCP having good bandwidth and speed

12. Compare and contrast FC-4 layer and FC-2 layer.

❖ Fibre Channel FC-4 is the fourth layer of the Fibre Channel (FC) protocol stack. It provides a standard set of services, known as application protocols, to upper layer protocols such as SCSI (Small Computer System Interface) and IP (Internet Protocol). ❖ These services allow FC-4 to carry various data payloads, including block data, file data, and video data, over a Fibre Channel network. The goal of FC-4 is to provide a flexible, scalable, and high-performance data transport solution that supports a variety of storage and networking applications.	❖ Fibre Channel FC-2 refers to the network layer of the Fibre Channel architecture. It is responsible for providing routing and switching of data frames between devices in a Fibre Channel network. ❖ The FC-2 layer performs functions such as address mapping, routing, and switching of data frames between devices. It uses unique device addresses to identify devices in the network and routes data frames to the correct destination. FC-2 layer also provides network-level flow control, which regulates the flow of data frames to prevent congestion in the network. This helps to ensure that data frames are transmitted efficiently and effectively in the Fibre Channel network.

13. Develop the 24 bit FC address of N port.

The FC address is 24 bits long and contains the following 8-bit fields: Domain_ID—A domain represents a switch and all N_Ports connected to the switch. For more information about N_Ports, see "Port modes." A domain ID, which is in the range of 1 to 239, uniquely identifies an FC switch. For more information about N_Ports, see "Port modes." A domain ID, which is in the range of 1 to 239, uniquely identifies an FC switch.

14. Point out the types of Zoning.

- Port zones
- Mixed zones
- Worldwide name zones

15. Assess the term ‘mesh topology’.

Mesh topology is a type of networking in which all the computers are inter-connected to each other. In Mesh Topology, the connections between devices take place randomly. The connected nodes can be computers, switches, hubs, or any other devices. In this topology setup, even if one of the connections goes down, it allows other nodes to be distributed. This type of topology is very expensive and does not have any hierarchy, interdependency, and uniform pattern between nodes. The connections of the mesh topology are not easier to establish.

16. Illustrate the fiber channel protocol stack.

It is easier to understand a communication protocol by viewing it as a structure of independent layers.

FCP defines the communication protocol in five layers: FC-0 through FC-4 (except FC-3 layer, which is not implemented).

17. Demonstrate the switched fabric port.

Node ports, which are on a host or storage system. F_Ports. Fabric ports, which are on a switch and provide connections between a node and the network. E_Ports. Expansion ports, which are on a switch and connect to other switches, sometimes by means of inter-switch links (ISLs).

18. Show the Single mode fiber.

Single-mode fiber is a common type of optical fiber that is used to transmit over longer distances. It is one of the two optical fiber types, the other being multi-mode fiber. A single-mode fiber is a single glass fiber strand used to transmit a single mode or ray of light

19. What is Compute System?

A management information system (MIS) is a computer system of hardware and software that acts as the foundation for an organization's operations. An MIS collects data from various online systems to support management decision-making, analyses the information, and reports data.

20. Express the virtual memory manager

The Virtual Memory Manager (VMM) manages the memory requests made by the system and its applications.

Virtual-memory segments are partitioned in units called *pages*; each page is either located in real physical memory (RAM) or stored on disk until it is needed. AIX® uses virtual memory to address more memory than is physically available in the system. The management of memory pages in RAM or on disk is handled by the VMM.

21. Define swap space

A computer has a sufficient amount of physical memory but most of the time we need more so we swap some memory on disk. Swap space is a space on a hard disk that is a substitute for physical memory. It is used as virtual memory which contains process memory images. Whenever our computer runs short of physical memory it uses its virtual memory and stores information in memory on disk

22. Assess the paging

Paging is a memory management scheme that eliminates the need for contiguous allocation of physical memory. The process of retrieving processes in the form of pages from the secondary storage into the main memory is known as paging. The basic purpose of paging is to separate each procedure into pages. Additionally, frames will be used to split the main memory.This scheme permits the physical address space of a process to be non – contiguous.

Part B

1. Examine Direct Attached Storage in detail. (13)
2. Give the detail about
i) Command and status sequence number.(7)
ii) FCIP(6)
3. Describe the FC SAN evolution in detail.(13)
4. With a neat diagram explain components of FC SAN (13)
5. Demonstrate the FC connectivity with neat diagram (13)
6. Summarize:
i) The advantages of FCP (6)
ii) Fiber channel switched fabric (7)
7. With a neat diagram explain the data transmission in Fiber channel switched fabric.(13)
8. Examine the Fiber channel Architecture in detail. (13)
9. Explain the Fabric Services (7)
ii) Compare the Class 1, Class 2 and Class 3 services of Fabric channel. (6)
10. Summarize the Switched Fabric Login types (13)
11. Describe FC SAN Topologies. (13)
12. (i) Explain the block level storage virtualization. (7)
(ii)Write a note on Virtual SAN. (6)
13. (i)Discuss about component Core Edge Topologies. (7)
(ii)Discuss benefits and limitations of Core Edge Fabric. (6)
14. Analyze the types of Zoning. (13)
15. (i)Demonstrate Fabric services provided by FC switches. (7)
(ii)Explain the structure and organization of FC data (6)
16.Summarize:
(i)Fiber channel arbitrated loop(6)
(ii)SAN Management software.(7)
17.Write short notes on
(i)Fiber channel addressing (7)
(ii)FC frame(6)

UNIT III- ADVANCED STORAGE NETWORKING AND VIRTUALIZATION

Part A

1. Define.iSCSI

ISCSI, which stands for Internet Small Computer System Interface, works on top of the Transport Control Protocol (TCP) and allows the SCSI command to be sent end-to-end over local-area networks (LANs), wide-area networks (WANs) or the internet

2. Tabulate internet Storage Name.

- Hard Drive Disks
- Floppy Disks
- Tapes
- Compact Discs (CDs)
- DVD and Blu-ray Discs
- USB Flash Drives
- Secure Digital Cards (SD Card)s
- Solid-State Drives (SSDs)
- Cloud Storage

3. Show the FCIT protocol stack.

A protocol stack refers to a group of protocols that are runnning concurrently that are employed for the implementation of network protocol suite. The protocols in a stack determine the interconnectivity rules for a layered network model such as in the OSI or TCP/IP models. To become a stack the protocols must be interoperable being able to connect both vertically between the layers of the network and horizontally between the end-points of each transmission segment.

4. Define CAS .

CAS stands for Confirmation of Acceptance for Studies. International students who want to qualify for a Tier 4 (General) student visa to study in the UK need the CAS.Important details about the student's course of study are included in the CAS, including the start and end dates, the cost of the course, and the degree of study. The CAS also verifies that the student has taken the offer of a spot in the course and has satisfied the university's or college's admission requirements.

5. Classify the Send Targets discovery.

You can discover iSCSI target portals by using the send targets method. You can discover iSCSI target portals by using the Internet storage name service (iSNS) method. You can connect to a specific portal on a discovered target by using the iscsiadm command.

6. Differentiate between Hierarchical file systems and Flat address space.

- Hierarchical file system organizes data in the form of files/directories
  - Limits the number of files that can be stored
- OSD uses flat address space that enables storing large number of objects
  - Enables the OSD to meet the scale-out storage requirement of third platform

7. Give the components OSD

- OSD agent.
- OSD manager.
- Image repository.
- Network boot listener.
- Storage device.
- NAMPUser.
- Projects.
- OSD wizard.

8. List the Data storage process in OSD

- The application server presents the data to be stored to the OSD node.
- The OSD node divides the data into two parts: user data and metadata.
- The OSD node generates the object ID using a specialized algorithm.
- The algorithm is executed against the contents of the user data to derive an ID unique to this data.

9. Discuss about the benefits of object based storage

- Object storage can automatically replicate data and store it across multiple devices and geographical locations.
- This can help protect against outages, safeguard against data loss, and help support disaster recovery strategies.
- Object storage is best used for large amounts of unstructured data, especially when durability, unlimited storage, scalability, and complex metadata management are relevant factors for overall performance.

10.Analyze the purpose of CAS

Content addressed storage or CAS is an object-based system that has been made for storing and managing fixed content data. It is designed for secure online storage and recovery of fixed content.

11.List the key features of CAS

- Content Authenticity: ...
- Content Integrity: ...
- Location Independence: ...
- Single Instance Storage: ...
- Retention Enforcement: ...
- Record Level Protection and Disposition: ...
- Technology Independence: ...
- Fast Record Retrieval

12. Classify the major functions in CAS

When storing data, a CAS system assigns a content address to each object. The content address is a unique identifier that is calculated based on the content itself, providing a digital fingerprint that ensures the data's authenticity and uniqueness. Applications that need to access data in a CAS system must use the content addresses to find and retrieve the desired objects. In CAS, data is stored on disk, rather than on tape, which streamlines the process of searching for archival data.

13. Create discovery using iSNS.

Before you use the iSNS discovery, you must do the following tasks:

- Configure the system to use iSCSI.
- Configure the system to use your iSNS server. See .
- Verify that the system nodes registered their target portals with the iSNS server. If the node iSCSI addresses are visible to the iSNS server, this registration occurs automatically.

About this task

The iSNS method is configured from the Discovery tab of the iSCSI Initiator control panel.

Procedure

1. Click the Add button that is associated with the list of iSNS servers.
2. Type the IP address or DNS host name of your iSNS server.

14. Illustrate FCoE Frame Structure

FCoE transports Fibre Channel directly over Ethernet while being independent of the Ethernet forwarding scheme. The FCoE protocol specification replaces the FC0 and FC1 layers of the Fibre Channel stack with Ethernet. By retaining the native Fibre Channel constructs, FCoE was meant to integrate with existing Fibre Channel networks and management software.

Data centers used Ethernet for TCP/IP networks and Fibre Channel for storage area networks (SANs). With FCoE, Fibre Channel becomes another network protocol running on Ethernet, alongside traditional Internet Protocol (IP) traffic. FCoE operates directly above Ethernet in the network protocol stack, in contrast to iSCSI which runs on top of TCP and IP. As a consequence, FCoE is not routable at the IP layer, and will not work across routed IP networks.

15. Show the differences between FCoE Enabling Technologies.

Many enterprise data centers rely on Ethernet for their LAN and data traffic, and on Fibre Channel (FC) networks for their storage infrastructure. With increased adoption of 10GbE in the data center, the availability of Fibre Channel over Ethernet (FCoE), and new lossless 10GbE technologies, it is now possible to consolidate FC SAN storage data flows with LAN and data traffic onto the same unified Ethernet cable. Network convergence will enable enterprises to preserve their existing investments in FC storage, reduce data center costs and complexity, and simplify network management.

16. Point out the storage m virtualization taxonomy

The first classification discriminates against the service or entity that is being emulated

- Virtualization is mainly used to emulate execution environments, storage and networks
- Execution virtualization is the oldest, most popular •Two major categories: Process level, System level
- Process level techniques - on top of existing OS which has full control of the hardware
- System level- Directly on hardware and require minimum support from existing OS

17. Generalize the storage virtualization configuration

NAS systems, managing SANs can be a time-consuming process. Consolidating a number of block storage systems under a single management interface that often shields users from the tedious steps of LUN configuration, for example, can be a significant timesaver. An early version of block-based virtualization was IBM's SAN Volume Controller, now called IBM Spectrum Virtualize. The software runs on an appliance or storage array and creates a single pool of storage by virtualizing LUNs attached to servers connected to storage controllers. Spectrum Virtualize also enables customers to tier block data to public cloud storage.

18.Define Storage Virtualization

As we know that, there has been a strong link between the physical host and the locally installed storage devices. However, that paradigm has been changing drastically, almost local storage is no longer needed. As the technology progressing, more advanced storage devices are coming to the market that provide more functionality, and obsolete the local storage.

Storage virtualization is a major component for storage servers, in the form of functional RAID levels and controllers. Operating systems and applications with device can access the disks directly by themselves for writing. The controllers configure the local storage in RAID groups and present the storage to the operating system depending upon the configuration. However, the storage is abstracted and the controller is determining how to write the data or retrieve the requested data for the operating system.

19. Give Send Targets discovery,

The three target discovery methods available in Oracle Solaris are as follows:

- sendtargets discovery method
- Internet Storage Name Service (iSNS) discovery method
- Static discovery method

20. Show about iSCSI Qualified Name

iSCSI uses a unique name to identify an iSCSI node, either target or initiator. This name that is used as a way to universally identify the node. The iSCSI qualified name format is defined in RFC3720. iqn specifies the use of the iSCSI qualified name as the authority

21.Define Status sequence number

You can define restrictions and fixed sequences for the status change. For this reason, status sequences are defined in Customizing. Using the status sequence, you can also indicate which statuses are not allowed, and which are mandatory.

For each status, you define whether it is one of the following:

- a mandatory status
- a recommended status
- an optional status
- a prohibited status

21. Evaluate Data sequence number

The application program can also assign sequence numbers when sending expedited data-flow-control requests (such as a Signal request). For these requests, VTAM uses the current setting of the SEQNO field of the SEND RPL instead of generating a sequence number. This allows the application program to use the SEQNO field as another data field (for example, to relate a Signal request to a particular bracket). VTAM generates sequence numbers for all session-control requests whether sent by VTAM (such as UNBIND), or by the application program (such as RQR).

Part B

1.Explain the components of iSCSI (13)
2.Describe the iSCSI topology in detail.(13)
3. (i)Illustrate iSCSI protocol stack (7)
(ii)Illustrate iSCSI PDU encapsulated in an IP packet. (6)
4. (i)iDemonstrate iSCSI names. (7)
(ii)What is iSCSI session? (6)
5. (i) Explain iSCSI command sequencing. (7)
(ii)Why FCIP might require high network bandwidth. (6)
6. Describe in detail about
(i)FCIP Protocol stack (7)
(ii)FCIP topology (6)
7.Classify the iSCSI Session in detail (13)
8.Explain File systems and Network file sharing (13)
9.Describe in detail about NFS (13)
10.Summarize the File Level Virtualization (13)
11. Explain object based storage architecture. (13)
12.Summarize the process of retrieval from OSD .(13)
13.(i)Analyze the benefits of Object based storage (7)
(ii)Describe Common used case for object based storage. (6)
14.Explain the key features of CAS (13)
15.(i)Explain the CAS use cases with example (6)
(ii)Describe the components of unified storage (7)
16. Show the object storage and retrieval process in an OSD system. (13)
17.Explain which object storage is better choice over SAN and NAS(13)

UNIT IV- BUSINESS CONTINUITY

Part A

1. Define Information availability.

Information availability (IA) refers to the ability of the infrastructure to function according to business expectations during its specified time of operation. Information availability ensures that people (employees, customers, suppliers, and partners) can access information whenever they need it.

2. Summarize the average cost of downtime per hour.

In the IT industry, downtime is typically calculated at about $5,600 per minute. Depending on the company's size, the full range of its lost revenue spans from $145,000 to $450,000 per hour. In the auto industry, downtime costs rise to about $50,000 per minute, which translates to about $3 million per hour

3. What is MTTR?

MTTR (mean time to recovery or mean time to restore) is the average time it takes to recover from a product or system failure. As we touched on earlier, the MTTR formula is the total unplanned maintenance time divided by the total number of repairs (failures). MTTR is most commonly represented in hours. Keep in mind, MTTR assumes tasks are performed sequentially and by trained maintenance personnel.

4. Classify the RPO.

A Recovery Point Objective (RPO)? Recovery point objective (RPO) is the maximum acceptable amount of data loss after an unplanned data-loss incident, expressed as an amount of time. Describes the acceptable amount of data loss over a given timeRecovery Point Objective describes a point in time at which a disaster occurs 120 seconds.

5. Develop a QC planning.

A Quality Management Plan (QMP) is a defined set of processes and methods that accompany the development of a product and ensure that it meets certain quality standards. It comes as a document that outlines the policies, procedures, and objectives for achieving quality in a project. A QMP is designed to ensure that the project team is aware of the quality requirements and that they are met throughout the project. It also helps to identify and manage potential quality risks.

6. Give single point of failure.

A single point of failure (SPOF) is a system component which, upon failure, renders an entire system unavailable or unreliable. When you design a highly available deployment, you identify potential SPOFs and investigate how these SPOFs can be mitigated

7. Analyze multi pathing software.

Multipathing, also called SAN multipathing or I/O multipathing, is the establishment of multiple physical routes between a server and the storage device that supports it. In storage networking, the physical path between a server and the storage device that supports it can sometimes fail.

8. Express the task which include in VIA

Storage Management is defined as it refers to the management of the data storage equipment's that are used to store the user/computer generated data. Hence it is a tool or set of processes used by an administrator to keep your data and storage equipment's safe. Storage management is a process for users to optimize the use of storage devices and to protect the integrity of data for any media on which it resides

and the category of storage management generally contain the different type of subcategories covering aspects such as security.

9. Distinguish the local replication and remote replication.

Local replication enables you to perform recovery operations in the event of data loss and also provides support for backup. Whereas in storage system-based remote replication, the replication is carried out between storage systems

10. List the stages of BC planning.

Four Steps to Developing an Effective Business Continuity Plan

- Identify threats or risks. ...
- Conduct a business impact analysis. ...
- Adopt controls for prevention and mitigation. ...
- Test, exercise and improve your plan routinely.

11. Differentiate between MTBF and MTTR

Mean time before failure vs mean time to repair

MTBF and MTTR are related as different steps in a larger process. MTBF measures the time between failures for devices that need to be repaired, MTTR is simply the time that it takes to repair those failed devices. In other words, MTBF measures the reliability of a device, whereas MTTR measures the efficiency of it's repairs.

12. List the goal of business continuity

The main goal of business continuity planning is to support key company activities during a crisis. Planning ensures a company can run with limited resources or restricted access to buildings. Continuity planning also aims to minimize revenue or reputation losses.

13. Design information availability metrics.

Three of the most important and commonly cited metrics are latency, I/O operations per second (IOPS), and throughput. In addition to these three, queue length and I/O splitting can also provide valuable insights into storage performance. Application availability is determined based on application-specific key performance indicators (KPIs) such as overall or timed application uptime and downtime, number of completed transactions, responsiveness, reliability, and other relevant metrics.

14. Interpret disaster recovery.

Disaster recovery relies upon the replication of data and computer processing in an off-premises location not affected by the disaster. When servers go down because of a natural disaster, equipment failure or cyber attack, a business needs to recover lost data from a second location where the data is backed up. Ideally, an organization can transfer its computer processing to that remote location as well in order to continue operations.

15. Tabulate Backup purpose

Backup storage is the supplementary, off-device storage of files and application data. It is both hardware- and software-based and allows for the secure maintenance, management, and retrieval of important backup data. There are mainly three types of backup: full, differential, and incremental.

16. Identify measuring information availability.

- Information availability (IA) refers to the ability of the infrastructure to function according to business expectations during its specified time of operation.

- Information availability ensures that people (employees, customers, suppliers, and partners) can access information whenever they need it.
- Information availability can be defined with the help of reliability, accessibility and timeliness

17. Interpret the strategies to meet RPO and RTO targets

The main difference is in their purposes – being focused on time, RTO is focused on downtime of services, applications, and processes, helping define resources to be allocated to business continuity; while RPO, being focused on amount of data, has as its sole purpose to define backup frequency.

18. Show established Objectives in BC planning

The main goal of business continuity planning is to support key company activities during a crisis. Planning ensures a company can run with limited resources or restricted access to buildings. Continuity planning also aims to minimize revenue or reputation losses.

19. Analyze Authentication with a directory service such as Active Directory

The client requests an authentication ticket from the AD server. The AD server returns the ticket to the client. The client sends this ticket to the Endpoint Server. The Server then returns an acknowledgment of authentication to the client. The main service in Active Directory is Domain Services (AD DS), which stores directory information and handles the interaction of the user with the domain. AD DS verifies access when a user signs into a device or attempts to connect to a server over a network.

20. Give Causes of latency

The main cause of latency is distance. The longer the distance between the browser making the request and the server responding to that request, the more time it'll take the requested data to travel there and back.Latency is a synonym for delay. In telecommunications, low latency is associated with a positive user experience (UX) while high latency is associated with poor UX. In computer networking, latency is an expression of how much time it takes for a data packet to travel from one designated point to another.

Part B

1. (i)Discuss about process of information unavailability (7)
(ii)Summarize measuring information availability (6)
2. Illustrate recovery-point objective and recovery-time objective (13)
3. Demonstrate BC planning life cycle in detail.(13)
4. Integrate resolving single points of failure .(13)
5. (i)Evaluate business impact analysis. (6)
(ii)Evaluate set of task that includes in business impact analysis. (7)
6.Describe power path features in detail.(13)
7.(i)Draw and explain the dynamic load balancing policy.(8)
(ii)Explain IO operation without power path. (5)
8. Describe path failure without power path. (13)
9.Explain path fail over with power path for an active-active array (13)
10.Write short notes on Path failover with power path for an active-passive array.(13)
11.Explain the automatic path failover in detail.(13)
12.Express the ways in which one or more copies of data maintained in business operations .(13)
13. Classify the RTO recovery strategies to ensure data availability. (13)
14. Point out how to calculate the IA with an example. (13)
15.Expain in detail about Dynamic Load Balancing.(13)
16.Illustrate in detail about BC Terminology.(13)
17.Give notes on
(i)Data Access from Unified Storage.
(ii)Components of Unified Storage

UNIT V- REPLICATION

Part A

1. **What is local replication?**
   Replication is **the process of creating an exact copy of data**. Creating one or more replicas of the production data is one of the ways to provide Business Continuity (BC). These replicas can be used for recovery and restart operations in the event of data loss.

2. **Give the operation in replication environment?**
   Data replication refers to the process of creating numerous copies of your data and making them available across a diverse set of locations in a network to ensure high data availability and accessibility. It takes into consideration all of the Data Sources present in an organization's distributed infrastructure.

3. **Differentiate local replication and remote replication**
   Replication can be classified into two major categories: local and remote. Local replication refers to replicating data within the same array or the same data center. Remote replication refers to replicating data at a remote site. This chapter provides details about various local replication technologies, along with restore and restart considerations. This chapter also details local replication in a virtualized environment.

4. **Show the uses of local replicas.**
   - These replica copies are used for **restore and restart operations if data loss occurs**.
   - These replicas can also be assigned to other hosts to perform various business operations, such as backup, reporting, and testing.

5. **Define replicas consistency.**
   The states of the replicas of an application should remain to be identical at the end of the processing of each request. Replica consistency is necessary to mask a fault in some replicas.

6. **What is host-based local replication ?**

   - The Host-Based Replication is a process of duplicating data from one site to the target place and works with any storage type such as NAS, DAS, and SAN.
   - A file system filter driver (the replication agent) is installed onto the server to be replicated.
   - The agent operates on the I/O storage system's traffic and replicates it to its destination. Its software can be installed within an OS or a virtual machine, locally.

7. **Define the LBM based mirroring.**

   - In LVM-based replication, the 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.

8. **Express full-volume mirroring.**
   By using volume mirroring, **a volume can have two physical copies**. Each volume copy can belong to a different pool, and each copy has the same virtual capacity as the volume. In the management GUI, an asterisk (*) indicates the primary copy of the mirrored volume.
9. **Define pointer-based full-volume replication .**
   Pointer-based virtual replication (also referred as storage system-based snapshot) is **a space optimal solution when compared to full volume replica**. At the time of replication session activation, the target (snapshot) contains pointers to the location of the data on the source.
10. **How the condition occurs after the replication session initiate in CoFA.**
    Pointer-based, full-volume replication can be activated in either Copy on First Access(CoFA) mode or Full Copy mode.
    In CoFA, after the replication session is initiated, the data is copied from the sourceto the target only when the following condition occurs:
    •A write I/O is issued to a specifi c address on the source for the first time.
    •A read or write I/O is issued to a specifi c address on the target for the first time.
    •When a write is issued to the source for the fi rst time after replication sessionactivation, the original data at that address is copied to the target. After thisoperation, the new data is updated on the source.
11. **List the various storage array based replication technologies.**
    - Storage array-based local replication, where source and target are in the same array and accessed by different hosts.
    - Storage array-based local replication can be further categorized as full-volume mirroring, pointer-based full-volume replication, and pointer-based virtual replication.
12. **Point out array-based local replication technologies**
    Array-based replication has the advantage of being universal, at least to the data on that array. It's simply replicating bits of data as they change from one array to another, so any host that connects to the array can be replicated regardless of the OS that it's running. It also doesn't consume any of the host resources, so they can remain dedicated to application performance.

Part B

1. **Explain the common terms used to represent various entities and operation in a replication environment.(13)**
2. **i) Discuss about fast recovery and data migration.(6)**
   ii) Give the various purposes of one or more local replicas of the source may be created. . (7)
3. **Compare and contrast the local replication technologies (13)**
4. **(i) Summarize the advantages of LBM\-based replication (6)**
   (ii) Describe about the limitation of LBM-based replication (7)
5. **i)List the uses of Local replica.(8)**
   (ii)Write shot notes on Dependent write consistency on sources (5)
6. **i)Illustrate to production FS in detail (8)**
   ii)Write short notes on Continuous Data Protection.(6)
7. **i)Illustrate Continuous data protection — local replication.(7)**
   ii)Examine Tracking changes in detail.(6).
8. **Write short notes on Restore and Restart Considerations (13)**
9. **(i)What is Snap View Snapshot (7)**
   (ii)Identify Snap View Clone (6)
10. **. Generalize the functions of**
    (i) Time Finder/Clone (7)
    (ii) Time Finder/Snap (6)
11. **(i)Identify the VM Snapshot captures.(7)**
    (ii) Describe about the VM Clone (6)
12. **Explain Tracking Changes to Source and Replica.(13)**
13. **Analyze the administrator configures six pointer-based virtual replicas of a source LUN and creates eight full-volume replicas of the same LUN. The administrator then creates four pointer-based virtual replicas for each full-volume replica that was created. How many usable replicas are now available?. (13)**
14. **Explain the Flushing the file system buffer.(13)**
15. **illustrates the write operations to the production fi le system.(13)**
16. **Describe Synchronous replication in detail.(13)**
17. **Explain in detail about Network-Based Remote Replication.(13)**