**CCS367-STORAGE TECHNOLOGIES**

**UNIT-1**

**PART-A**

**1. What is Information Lifecycle Management (ILM)?**

**ILM** is a policy-based approach to managing data through its lifecycle — from creation to deletion. It ensures data is stored, protected, and accessed efficiently based on its current value.

**Stages:**

* Creation & capture
* Classification
* Storage
* Usage/access
* Archival
* Deletion/destruction

**2. What are the Core Elements of Data Center Infrastructure?**

* **Servers** – Perform compute operations.
* **Storage systems** – Hold structured and unstructured data.
* **Networking** – Switches, routers, and cabling for connectivity.
* **Power** – Backup generators, UPS systems.
* **Cooling systems** – Maintain optimal temperatures.
* **Security systems** – Firewalls, access control, surveillance.
* **Management software** – For monitoring and automation.

**3. What are the Key Characteristics of Data Center Elements?**

* **Reliability** – High availability and fault tolerance.
* **Scalability** – Ability to grow with demand.
* **Performance** – Fast computing and low-latency storage/networking.
* **Security** – Physical and cyber protections.
* **Efficiency** – Power and space optimization.
* **Manageability** – Centralized monitoring and control.

**4. What is the Evolution of Computing Platforms?**

1. **Mainframe computing** – Centralized, large systems.
2. **Client-server** – Distributed, with dedicated servers and user PCs.
3. **Web-based computing** – Services accessible over the internet.
4. **Virtualization** – Hardware abstraction to run multiple systems on one.
5. **Cloud computing** – On-demand resources over the internet.
6. **Edge computing & IoT** – Data processing closer to the source.
7. **AI/ML-driven platforms** – Intelligent, adaptive systems.

**5. What are the Cloud Services?**

* **IaaS (Infrastructure as a Service)** – Virtual servers, storage, networks (e.g., AWS EC2).
* **PaaS (Platform as a Service)** – Development tools and environments (e.g., Google App Engine).
* **SaaS (Software as a Service)** – Applications delivered online (e.g., Microsoft 365).
* **FaaS (Function as a Service)** – Serverless computing (e.g., AWS Lambda).

**6. What are the Cloud Deployment Models?**

* **Public Cloud** – Services offered over the internet (e.g., AWS, Azure).
* **Private Cloud** – Dedicated cloud environment for one organization.
* **Hybrid Cloud** – Mix of public and private cloud.
* **Community Cloud** – Shared infrastructure for specific user groups.

**7. Characteristics of Third Platform Infrastructure**

* Built on **Cloud**, **Big Data**, **Mobile**, and **Social** technologies.
* **Highly scalable**
* **Flexible and elastic**
* **API-driven**
* **Real-time analytics**
* **DevOps and CI/CD ready**

**8. What is the Building Block of a Data Center?**

The **server** is the foundational building block — often housed in **racks**. Other building blocks include:

* **Storage arrays**
* **Network switches**
* **Power and cooling units**
* **Management systems**

**9. What is Big Data Analysis and Give the Details?**

**Big Data Analysis** involves examining large and complex data sets to uncover hidden patterns, correlations, trends, and insights.

**Key Aspects:**

* **Volume** – Massive amounts of data.
* **Velocity** – High speed of data generation.
* **Variety** – Different types (text, images, logs).
* **Veracity** – Ensuring data accuracy and quality.
* **Value** – Turning raw data into actionable insights.

**Tools:** Hadoop, Spark, NoSQL databases, machine learning models.

**10. What is Digital Data and Its Types?**

**Digital data** is information stored using binary format (0s and 1s).

**Types:**

* **Structured** – Databases, spreadsheets.
* **Unstructured** – Emails, videos, images.
* **Semi-structured** – XML, JSON, log files.
* **Metadata** – Data about data (e.g., file size, creation date).

**UNIT-2**

**PART-A**

**1. What is an Intelligent Storage System?**

An **Intelligent Storage System** is a storage array that includes advanced features for data management, performance optimization, and high availability. It uses embedded software and hardware to automate tasks and improve efficiency.

**Key Features:**

* Data tiering
* Automated storage provisioning
* Caching for performance
* RAID support
* Remote replication

**2. What is Cache?**

**Cache** is a small, high-speed memory (usually DRAM or flash) used to temporarily store frequently accessed data, reducing access time and improving performance.

**Types:**

* **Read cache** – Stores recently read data.
* **Write cache** – Temporarily holds data before writing to disk.

**3. What is a Physical Disk?**

A **physical disk** is the actual hardware device where data is stored. Examples include:

* **HDDs (Hard Disk Drives)** – Magnetic-based storage.
* **SSDs (Solid-State Drives)** – Flash-based, no moving parts.

**4. What is a Platter?**

A **platter** is a circular disk inside an HDD that stores data magnetically. Each platter has:

* A top and bottom surface
* Read/write heads
* Spins at high RPM (e.g., 7200 RPM)

Multiple platters can exist in one HDD.

**5. What is Logical Block Addressing (LBA)?**

**LBA** is a method of identifying data locations on a disk using **sequential block numbers** rather than cylinder/head/sector. It simplifies how systems interact with storage devices.

**6. What is Seek Time?**

**Seek time** is the time it takes for the HDD’s read/write head to move to the track where the data is stored.

* Measured in milliseconds (ms)
* Lower is better (e.g., 4ms to 9ms)

**7. What is RAID? And Its Types?**

**RAID (Redundant Array of Independent Disks)** combines multiple physical disks into one logical unit for **redundancy**, **performance**, or both.

**Common RAID Types:**

* **RAID 0** – Striping (performance, no redundancy)
* **RAID 1** – Mirroring (redundancy, low capacity)
* **RAID 5** – Striping + parity (balanced)
* **RAID 6** – Striping + double parity (extra fault tolerance)
* **RAID 10** – Mirroring + striping (high performance + redundancy)

**8. What are the Components of an Intelligent Storage System?**

* **Front-end** – Interfaces with host systems (e.g., Fibre Channel, iSCSI).
* **Cache** – Temporary high-speed memory.
* **Storage processors** – Control and manage I/O operations.
* **Back-end** – Interfaces with the physical disks.
* **Physical disks** – Where data is stored persistently.

**9. What is Scale-Up and Scale-Out Storage Architecture?**

* **Scale-Up** – Add more capacity to a single storage system (e.g., add more drives to a server).
* **Scale-Out** – Add more nodes to increase both capacity and performance (e.g., distributed systems like NetApp, Dell EMC Isilon).

**10. What is Mirroring?**

**Mirroring** is a data protection technique that creates an exact copy of data on two or more disks. It’s used in RAID 1 and RAID 10 setups.

* Offers **high redundancy**
* Great for **critical data**

**UNIT-3**

**PART-A**

**1. What is FCoE Architecture?**

**FCoE (Fibre Channel over Ethernet)** is a storage protocol that encapsulates Fibre Channel frames over Ethernet networks.

**Architecture Highlights:**

* Uses **Converged Network Adapters (CNAs)** to handle both Ethernet and FC traffic.
* Reduces need for separate storage and network cabling.
* Works over **Data Center Bridging (DCB)** Ethernet — a lossless Ethernet extension.

**Goal:** Combine storage and IP network traffic into a single network fabric.

**2. What is Virtualization?**

**Virtualization** is the process of creating a virtual version of physical resources — like servers, storage, or networks.

**Types:**

* Server virtualization
* Storage virtualization
* Network virtualization
* Desktop virtualization

**Benefits:** Improved resource utilization, flexibility, scalability, and isolation.

**3. What is Virtualization in FC SAN Environment?**

In **FC SAN (Fibre Channel Storage Area Network)**:

* **Storage virtualization** abstracts physical storage into logical volumes.
* Allows pooling and dynamic allocation of storage.
* Enables features like thin provisioning, replication, and snapshots.

**Helps simplify storage management and boost utilization.**

**4. What is Block-Based Storage System?**

A **block-based storage system** stores data in fixed-size blocks (e.g., 512 bytes, 4 KB). Each block has its own address.

**Common in:** SANs  
**Use cases:** Databases, virtual machines  
**Protocols:** iSCSI, FC, FCoE

**5. What is File-Based Storage System?**

A **file-based storage system** stores data as files in a directory hierarchy.

**Common in:** NAS (Network Attached Storage)  
**Protocols:** NFS (Linux), SMB/CIFS (Windows)  
**Use cases:** File sharing, home directories

**6. What is Object Storage System?**

**Object storage** manages data as objects (which include the data, metadata, and a unique ID).

**Benefits:**

* Highly scalable
* Great for unstructured data (videos, images)
* Supports HTTP/REST API access

**Used in:** Cloud storage (e.g., Amazon S3)

**7. Types of SDN (Software Defined Networking):**

* **Open SDN:** Based on **OpenFlow** protocol, controller decides data flow.
* **API-Based SDN:** Uses APIs to configure network devices (e.g., Cisco ACI).
* **Overlay SDN:** Builds virtual networks over physical ones (e.g., VXLAN).
* **Hybrid SDN:** Mix of traditional and SDN elements.

**8. Important Advantages of SDN:**

* **Centralized control**
* **Dynamic network provisioning**
* **Faster deployment of services**
* **Improved scalability**
* **Reduced operational costs**
* **Vendor neutrality**

**10. What is Zoning?**

**Zoning** is a security mechanism in FC SANs that controls which devices can communicate.

**Types:**

* **Soft zoning:** Controlled via software (Name-based).
* **Hard zoning:** Enforced by the switch hardware (Port-based).

**Purpose:** Isolate devices, enhance security and performance.

**11. What is Switch Aggregation?**

**Switch Aggregation** is the process of combining multiple physical switches or switch ports to increase bandwidth and reliability.

* Also known as **link aggregation** or **port-channeling**.
* Improves **throughput** and **fault tolerance**.
* Common in **Ethernet** and **SAN environments**.

**UNIT-4**

**PART-A**

**1. What is Backup Architecture?**

**Backup architecture** is the structured design for protecting data by creating copies (backups) that can be restored in case of data loss or corruption.

**Components:**

* **Backup sources** – Systems or files to back up
* **Backup server/software** – Manages and schedules backups
* **Backup targets** – Storage devices (disk, tape, cloud)
* **Retention policy** – Defines how long backups are kept
* **Recovery strategy** – Steps for restoring data

**2. Characteristics of Replication**

**Replication** is the process of copying data from one location to another in real time or near real time.

**Key Characteristics:**

* **Real-time or scheduled**
* **Synchronous or asynchronous**
* **Source and target consistency**
* **Continuous availability**
* **Supports disaster recovery**

**3. What is Disaster Recovery as a Service (DRaaS)?**

**DRaaS** is a cloud-based service that provides full system backup, replication, and failover to ensure business operations continue during a disaster.

**Includes:**

* Virtual machine replication
* Automated failover/failback
* Testing without downtime
* Managed recovery plans

**4. What is Business Continuity?**

**Business Continuity (BC)** is the ability of an organization to continue operations during and after a disruption (e.g., natural disaster, cyberattack).

**Includes:**

* Risk assessment
* Disaster recovery planning
* Backup strategies
* Communication plans

**5. What is MTBF and MTTR?**

* **MTBF (Mean Time Between Failures):** The average time between hardware or system failures. Higher = more reliable.
* **MTTR (Mean Time To Repair):** The average time to fix a failure. Lower = faster recovery.

**6. What is Data Migration?**

**Data migration** is the process of moving data from one system/storage location to another.

**Common Scenarios:**

* Hardware upgrades
* Cloud adoption
* System consolidations
* Storage type changes (e.g., HDD to SSD)

**7. What is Backup?**

**Backup** is the process of creating copies of data so it can be recovered in case of loss, corruption, or failure.

**Types:**

* **Full** – Entire data set
* **Incremental** – Only changed data since last backup
* **Differential** – Changed data since last full backup

**8. What is Cloud-Based Backup?**

Cloud-based backup is a method where data is copied and stored in a **remote cloud storage** location.

**Benefits:**

* Offsite protection
* Scalability
* Automation
* Disaster recovery support

**Examples:** Acronis, Veeam, AWS Backup

**9. What is Storage-Based Replication?**

Storage-based replication is handled **at the storage system level**, without needing the host to participate.

**Features:**

* Faster
* Offloads processing from servers
* Supports synchronous/asynchronous modes
* Used in enterprise SAN environments

**10. What is Network-Based Replication?**

Network-based replication uses the **network layer** and typically operates through software installed on hosts or via network appliances.

**Features:**

* Doesn’t depend on specific storage vendors
* Can work across heterogeneous environments
* May be slightly slower than storage-based replication

**UNIT-5**

**PART-A**

**1. What is Storage Infrastructure Management Process?**

It’s the end-to-end process of managing physical and virtual storage resources to ensure **availability, performance, security**, and **efficiency**.

**Steps Involved:**

* **Provisioning** – Allocating storage to applications/servers.
* **Monitoring** – Tracking performance, health, and capacity.
* **Reporting** – Usage analytics, trends, and compliance reports.
* **Maintenance** – Updates, patching, and hardware management.
* **Troubleshooting** – Detecting and resolving issues.

**2. What Are Security Controls to Protect a Storage Infrastructure?**

Security controls ensure that data is protected from unauthorized access, corruption, or loss.

**Types of Controls:**

* **Access control** – Role-based access, authentication (e.g., LDAP).
* **Encryption** – For data at rest and in transit.
* **Data masking** – Hiding sensitive data.
* **Firewalls and VLANs** – Network segmentation.
* **Logging and auditing** – Track access and changes.
* **Backup and replication** – To prevent data loss.

**3. What Are the GRC Tools?**

**GRC (Governance, Risk, and Compliance) tools** help manage regulatory requirements, risks, and organizational policies.

**Popular GRC Tools:**

* **RSA Archer**
* **MetricStream**
* **LogicManager**
* **SAP GRC**
* **IBM OpenPages**
* **ServiceNow GRC**

**4. What Are the Challenges of GRC?**

* **Complex regulations** – Constantly evolving laws and standards.
* **Data silos** – Lack of centralized data visibility.
* **Manual processes** – Prone to error and inefficiency.
* **Integration issues** – Difficulty aligning IT systems with GRC tools.
* **Lack of real-time risk visibility** – Delayed response to threats.

**5. What is Automation?**

**Automation** in IT is using software tools/scripts to perform tasks **without manual intervention**.

**In storage & security:**

* Automating backup, monitoring, provisioning.
* Automating patching or security alerts.
* Reduces errors, saves time, and improves consistency.

**6. What is Risk Triad?**

The **Risk Triad** in cybersecurity typically refers to the **CIA Triad**:

* **Confidentiality** – Prevent unauthorized access.
* **Integrity** – Ensure data is accurate and untampered.
* **Availability** – Keep systems and data accessible when needed.

It’s a foundational model for assessing and securing IT systems.

**7. What is a Threat?**

A **threat** is any potential event or action that can cause harm to data, systems, or networks.

**Examples:**

* Malware
* Hackers
* Insider threats
* Natural disasters
* Phishing attacks

**8. What is a Vulnerability?**

A **vulnerability** is a **weakness** in a system that could be exploited by a threat to gain unauthorized access or cause harm.

**Examples:**

* Unpatched software
* Weak passwords
* Misconfigured servers
* Open ports

**9. What is Data Encryption?**

**Data encryption** is the process of converting readable data into a coded format (**ciphertext**) to prevent unauthorized access.

**Types:**

* **Symmetric encryption** – Same key for encryption/decryption (e.g., AES)
* **Asymmetric encryption** – Public/private key pair (e.g., RSA)

**Used for:**

* Data at rest (disks, databases)
* Data in transit (emails, HTTPS)