Multi-disciplinary Course

Information and Communication Technologies (ICT)

Semester-I

MDC-1: Information and Communication Technologies-I

Notes compiled by Dr. Yajnya Sapkota, Department of Physics, Dudhnoi College

---

Chapter-5
Introduction to Storage Systems: Overview of storage hierarchy and storage technologies

Storage systems are essential components of computer architecture, enabling the storage, retrieval, and management of data in various forms. These systems can be categorized into different levels based on speed, capacity, and cost.

Overview of Storage Hierarchy:

1. Registers (Level 1)

- **Speed**: Fastest (typically at the CPU level)
- **Cost**: Most expensive
- **Capacity**: Very small
- **Function**: Registers are the fastest form of storage located directly within the CPU. They are used to hold data and instructions that the CPU is currently processing.

2. Cache Memory (Level 2)

- **Speed**: Very fast
- **Cost**: Expensive
- **Capacity**: Small (usually in the range of kilobytes to megabytes)
- **Function**: Cache memory is used to store frequently accessed data or instructions for quick retrieval. It sits between the CPU and RAM, significantly improving the processing speed by reducing the access time to frequently used data.

3. Main Memory (RAM) (Level 3)

- **Speed**: Fast
- **Cost**: Moderate
- **Capacity**: Medium (usually gigabytes)
- **Function**: RAM is used for temporary storage of data that the CPU needs to access quickly. It is volatile memory, meaning data is lost when the system is powered down. It stores the operating system, running programs, and currently processed data.

4. Secondary Storage (Level 4)

- **Speed**: Slower than RAM
- **Cost**: Lower
- **Capacity**: Large (gigabytes to terabytes)

- **Function**: This level includes hard drives (HDDs), solid-state drives (SSDs), and optical drives (CD/DVD). Secondary storage provides long-term storage for the operating system, applications, and user data.

5. Tertiary Storage (Level 5)

- **Speed**: Slowest
- **Cost**: Very low
- **Capacity**: Very large (terabytes to petabytes)
- **Function**: Tertiary storage includes devices like tape drives, used for archival storage and backup purposes. Data stored here is infrequently accessed, and the retrieval process is slow but cost-effective for storing large volumes of data over long periods.

Characteristics of primary, secondary, and tertiary storage

1. Primary Storage

Definition: Primary storage, also known as **main memory**, is used for storing data that the CPU needs immediate access to during processing. It is volatile, meaning data is lost when the power is turned off.

Characteristics:

- **Speed**: **Very fast** (much faster than secondary and tertiary storage).
- **Capacity**: **Limited** (typically measured in gigabytes or less).
- **Volatility**: **Volatile**; data is lost when the system is powered off.
- **Cost**: **Expensive** per unit of storage.
- **Examples**:
    - **RAM (Random Access Memory)**: Used for temporary storage of data currently in use.
    - **Cache Memory**: High-speed memory used to store frequently accessed data close to the CPU.

Usage:

- **Active data storage**: Stores data that the CPU is actively processing.
- **Fast access for active programs**: Enables quick access to running applications and processes.

2. Secondary Storage

Definition: Secondary storage provides **long-term storage** for the operating system, applications, and user data. It is non-volatile, meaning data is retained even when the power is turned off.

Characteristics:

- **Speed**: **Slower** than primary storage but faster than tertiary storage.
- **Capacity**: **Large** (usually measured in gigabytes to terabytes).
- **Volatility**: **Non-volatile**; data is retained even when the system is powered off.
- **Cost**: **Moderate** compared to primary storage.
- **Examples**:
    - **Hard Disk Drives (HDDs)**: Magnetic storage devices that offer large capacity but slower speeds compared to SSDs.
    - **Solid-State Drives (SSDs)**: Faster than HDDs, using flash memory to store data.
    - **Optical Discs (CDs, DVDs)**: Used for storing data in a portable, removable format.

Usage:

- **Long-term data storage**: Stores operating systems, installed applications, user files, and documents.
- **Persistent storage**: Data is stored permanently unless deleted, ensuring persistence after the computer is shut down.

3. Tertiary Storage

Definition: Tertiary storage is used primarily for **archival and backup purposes**. It is slower than both primary and secondary storage and is typically used for large-scale data that is rarely accessed.

Characteristics:

- **Speed**: **Slowest** among all three types of storage.
- **Capacity**: **Very large** (typically measured in terabytes to petabytes).
- **Volatility**: **Non-volatile**; data is retained for long periods.
- **Cost**: **Very low** compared to primary and secondary storage.
- **Examples**:
    - **Magnetic Tape Drives**: Often used for backup and archival storage.
    - **Cloud Storage**: Remote storage solutions for large data storage and disaster recovery.
    - **Optical Storage (e.g., Blu-ray Discs, Archival-grade DVDs)**: Used for long-term storage of data, often in an archival or backup context.

Usage:

- **Data backup and archiving**: Used for storing data that is not frequently accessed but needs to be preserved for long-term purposes, such as backups and legal records.
- **Infrequent access**: Data stored on tertiary devices is rarely accessed but is kept for potential recovery or analysis in the future.

Summary Comparison of Primary, Secondary, and Tertiary Storage

Characteristic	Primary Storage	Secondary Storage	Tertiary Storage
Speed	Very fast (RAM, cache memory)	Slower than primary but fast (HDDs, SSDs)	Slowest (Magnetic Tape, Cloud Storage)
Capacity	Limited (GB)	Large (GB to TB)	Very large (TB to PB)
Volatility	Volatile (data lost on power off)	Non-volatile (retains data)	Non-volatile (retains data)
Cost	Expensive	Moderate (cost per GB)	Low cost
Examples	RAM, Cache memory	HDDs, SSDs, Optical Discs	Magnetic tape, Cloud storage
Usage	Active data processing	Long-term data storage	Archiving and backups

---

Introduction to cloud storage models (e.g., S3, Azure Blob Storage)

Cloud storage is a modern solution that allows users to store data on remote servers managed by providers like Amazon and Microsoft, rather than on local devices or physical servers. Two popular cloud storage services are Amazon S3 and Azure Blob Storage, each offering flexible, scalable, and cost-efficient ways to store large amounts of data.

Amazon S3 is a widely used cloud storage service that lets users store and retrieve data from anywhere on the internet. It is particularly suited for businesses and developers who need reliable and scalable storage. S3 offers various options based on how often data is accessed, allowing users to optimize costs by selecting storage options that fit their needs, whether it's for frequently accessed data or long-term archival storage. It is highly secure, with built-in encryption and access controls to ensure that data is protected.

Azure Blob Storage is a similar service provided by Microsoft through its Azure platform. It is designed for storing large amounts of unstructured data, such as videos, images, and documents. Like Amazon S3, Azure Blob Storage offers different pricing tiers to help users manage costs effectively, depending on how often they need to access the stored data. It also integrates well with other Microsoft services, making it an attractive option for businesses already using Microsoft technologies.

Both Amazon S3 and Azure Blob Storage are reliable, scalable solutions that offer businesses the flexibility to store vast amounts of data while ensuring security and easy access. The choice between the two depends largely on the specific needs of the organization, such as existing software environments and preferred service providers.

Data durability, availability, and scalability in cloud storage

In cloud storage, data durability, availability, and scalability are essential features that ensure data is stored safely, can be accessed when needed, and can grow with the needs of the business.

Data durability refers to how securely and reliably data is stored in the cloud. Cloud storage services aim for very high durability, meaning the risk of losing data is minimal, even if hardware fails. This is achieved through techniques like replicating data across multiple locations, ensuring that if one server or data center fails, the data is still accessible from another.

Data availability is about ensuring that data is accessible whenever it's needed. Cloud storage providers aim to keep data available around the clock, often offering service levels with guaranteed uptime. This means users can rely on cloud storage to access their files anytime, without worrying about downtime or service interruptions.

Scalability refers to the ability to easily increase or decrease the amount of storage based on changing needs. As businesses grow and data volumes increase, cloud storage can expand to accommodate that growth without the need for large upfront investments in physical hardware. Users can add more storage space as required, ensuring they only pay for what they use and can adjust quickly to meet their needs.

Together, these three features—durability, availability, and scalability—make cloud storage an attractive option for businesses, ensuring data is safe, always accessible, and capable of growing as the business does.

Emerging Trends in Storage: Flash-based storage technologies (e.g., SSD, NVMe)

Emerging trends in storage technology are shifting towards flash-based storage solutions, such as Solid-State Drives (SSDs) and Non-Volatile Memory Express (NVMe), due to their significant improvements in speed, efficiency, and reliability compared to traditional hard drives.

Solid-State Drives (SSDs) are becoming the go-to solution for both personal and enterprise storage needs. Unlike traditional Hard Disk Drives (HDDs), which rely on spinning disks and mechanical parts, SSDs use flash memory to store data, offering much faster read and write speeds. This results in quicker boot times, faster file transfers, and overall better performance for applications and services. SSDs are also more durable and reliable, as they have no moving parts, reducing the risk of mechanical failure.

Non-Volatile Memory Express (NVMe) is a newer protocol designed to further enhance the performance of SSDs by enabling faster data transfer speeds. NVMe drives are connected via the Peripheral Component Interconnect Express (PCIe) interface, which provides a much higher bandwidth than the older Serial ATA (Serial Advanced Technology Attachment) interface used by many SSDs. This allows NVMe drives to deliver even faster data access, making them ideal for high-performance computing tasks like gaming, video editing, and data-intensive applications.

The shift towards flash-based storage technologies like SSDs and NVMe is largely driven by the increasing demand for faster, more reliable, and energy-efficient storage solutions. As prices for these technologies continue to decrease and their capacity increases, they are expected to replace traditional storage systems across various industries, offering both consumers and businesses faster data access and improved system performance.
