M.S. Ramaiah Institute of Technology
(Autonomous Institute, Affiliated to VTU)
Department of Computer Science and Engineering

Course Name: Storage Area Networks

Course Code: CSE734

Credits: 3:0:0

UNIT - 1

RAMAIAH
Institute of Technology

UNIT 1- Storage System

Introduction to Information Storage(Chapter 1)

- Information Storage
- Evolution of Storage Architecture,
- Data Centre Infrastructure
- Virtualization and Cloud Computing.

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

UNIT 1- Introduction to Information Storage

- Information Storage
- Evolution of Storage Architecture,
- Data Centre Infrastructure
- Virtualization and Cloud Computing.

Introduction

- Information is increasingly important in our daily lives.
- We access the Internet every day to perform searches, participate in social networking, send and receive e-mails, share pictures and videos, and scores of other applications.
- Equipped with a growing number of content-generating devices, more information is being created by individuals than by businesses.
- Information created by individuals gains value when shared with others.
- When created, information resides locally on devices such as cell phones, cameras, and laptops.
- To share this information, it needs to be uploaded via networks to data centers.

Introduction

- The importance, dependency, and volume of information for the business world also continue to grow at astounding rates.
- Some of the business applications that process information include airline reservations, telephone billing systems, e-commerce, ATMs, product designs, inventory management, e-mail archives, Web portals, patient records, credit cards, life sciences, and global capital markets.

Introduction

- The increasing criticality of information to the businesses has amplified the challenges in protecting and managing the data.
- The volume of data that business must manage has driven strategies to :
  - classify data according to its value
  - create rules for the treatment of this data

RAMAIAH
Institute of Technology

Introduction

- Data centers now view information storage as one of their core elements, along with applications, databases, operating systems, and networks.
- Storage technology continues to evolve with technical advancements offering increasingly higher levels of availability, security, scalability, performance, integrity, capacity, and manageability.

Introduction

- Let us understand:

The Evolution of information storage architecture from simple direct-attached models to complex networked topologies

Information lifecycle management (ILM) strategy, which aligns the information technology (IT) infrastructure with business priorities.

Information Storage-Data

- Data is a collection of raw facts from which conclusions may be drawn.

Examples of data:

- 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,
- account holder’s passbooks

RAMAIAH
Institute of Technology

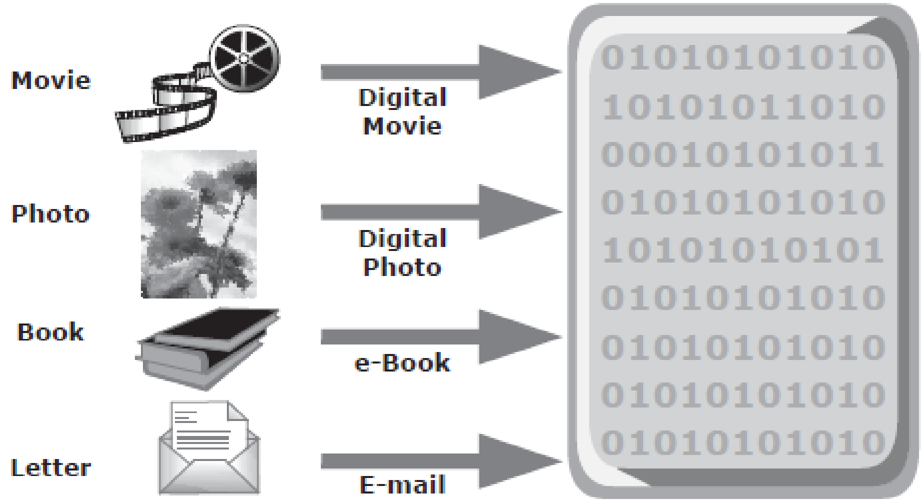
Information Storage-Data

Before the advent of computers, the methods adopted for data creation and sharing were limited to fewer forms, such as paper and fi lm.

Today, the same data can be converted into more convenient forms, such as an e-mail message,an e-book, a digital image, or a digital movie.

Information Storage-Data

This data can be generated using a computer and stored as strings of binary numbers (0s and 1s), as shown in Figure 1-1.



Data in this form is called *digital data* and is accessible by the user only after a computer processes it.

Information Storage-Data

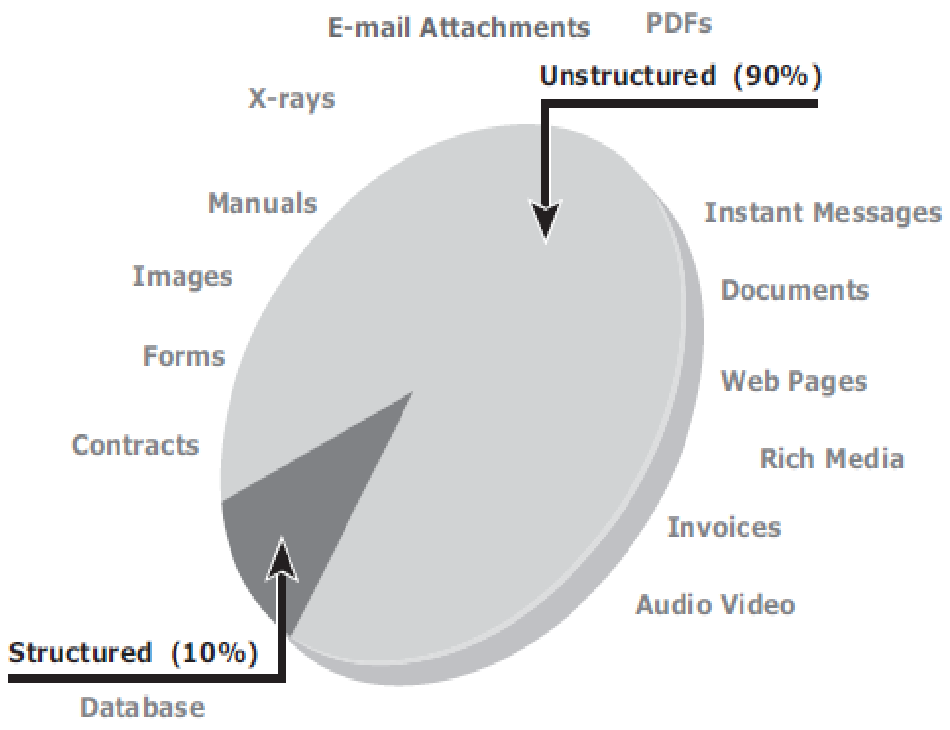
- The following is a list of some of the factors that have contributed to the growth of digital data:
- Increase in data processing capabilities: Modern-day computers provide a significant increase in processing and storage capabilities. This enables the conversion of various types of content and media from conventional forms to digital formats.
- Lower cost of digital storage: Technological advances and decrease in the cost of storage devices have provided low-cost solutions and encouraged the development of less expensive data storage devices. This cost benefit has increased the rate at which data is being generated and stored.
- Affordable and faster communication technology: The rate of sharing digital data is now much faster than traditional approaches. A handwritten letter may take a week to reach its destination, whereas it only takes a few seconds for an e-mail message to reach its recipient.

Information Storage-**Types of Data**

Data can be classified as structured or unstructured (see Figure 1-2) 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.

Structured data is typically stored using a database management system (DBMS).

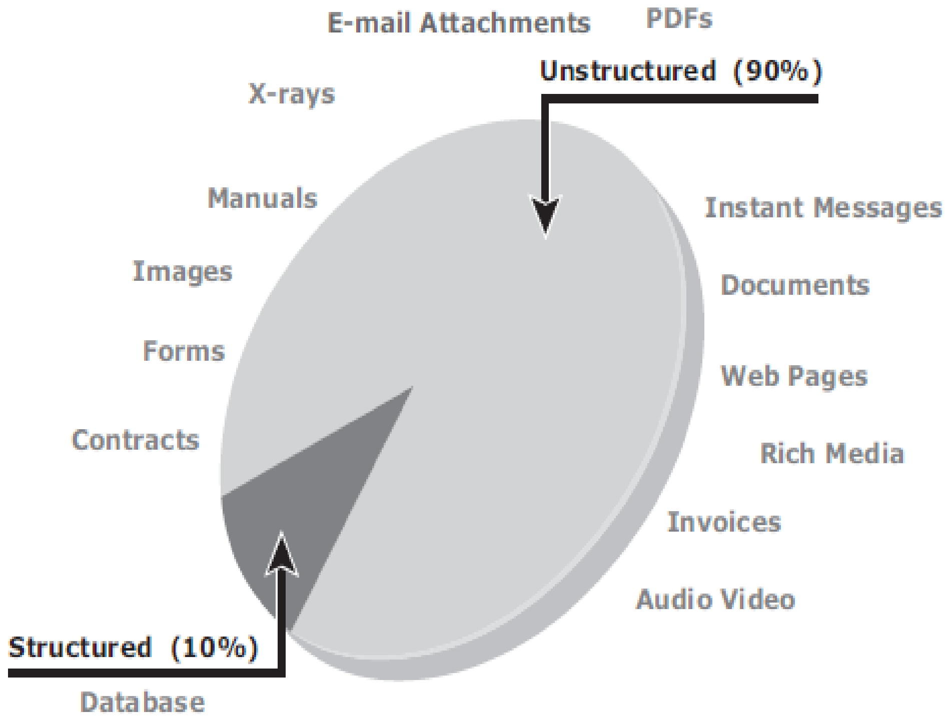


Information Storage-Types of Data

Unstructured Data - Elements are not stored in rows and columns, which makes it difficult to query and retrieve by applications.

For example, customer contacts that are stored in various forms such as sticky notes, e-mail messages,business cards, or even digital format files, such as .doc, .txt, and .pdf.

Due to its unstructured nature, it is difficult to retrieve this data using a traditional customer relationship management application.



Information Storage- **Big Data**

- Big data - refers to data sets whose sizes are beyond the capability of commonly used software tools to capture, store, manage, and process within acceptable time limits.
- It includes both structured and unstructured data generated by a variety of sources, including business application transactions, web pages, vidéos, images, e-mails, social media, and so on.
- These data sets typically require real-time capture or updates for analysis, predictive modeling, and decision making.

Information Storage- **Big Data**

The big data ecosystem consists of the following:

1. Devices that collect data from multiple locations and also generate new data about this data (metadata).

2. Data collectors who gather data from devices and users.

3. Data aggregators that compile the collected data to extract meaningful information.

4. Data users and buyers who benefit from the information collected and aggregated by others in the data value chain.

Information Storage -Information

- Data, whether structured or unstructured, does not fulfill any purpose for individuals or businesses unless it is presented in a meaningful form.
- Businesses need to analyze data for it to be of value.
- Information is the intelligence and knowledge derived from data.

Information Storage -Information

Example:

A retailer identifies customers' preferred products and brand names by analyzing their purchase patterns and maintaining an inventory of those products.

Information Storage -Information

- 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

Example : Job portal

job seekers post their résumés on various websites offering job search facilities.

These websites collect the résumés and post them on centrally accessible locations for prospective employers.

Information Storage -Information

Example : Job portal

Companies post available positions on job search sites.

Job-matching software matches keywords from resumes to keywords in job postings.

In this manner, the job search engine uses data and turns it into information for employers and job seekers.

Information Storage -Storage

In a computing environment, devices designed for storing data are termed storage devices or simply storage.

The type of storage used varies based on the type of data and the rate at which it is created and used.

Examples of storage devices:

Media card in a cell phone or digital camera, DVDs,CD-ROMs, and disk drives in personal computers

Information Storage -Storage

Example : Job portal

Companies post available positions on job search sites.

Job-matching software matches keywords from resumes to keywords in job postings.

In this manner, the job search engine uses data and turns it into information for employers and job seekers.

UNIT 1- Introduction to Information Storage

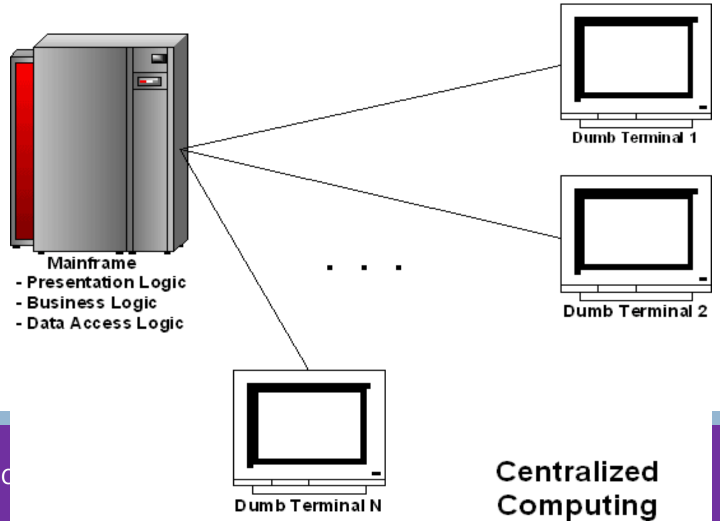
- Information Storage
- Evolution of Storage Architecture,
- Data Centre Infrastructure
- Virtualization and Cloud Computing.

Compare the Storage Area Network (SAN) architecture with the traditional server-centric architecture. Highlight differences in scalability, performance, and data sharing.

Evolution of Storage Architecture

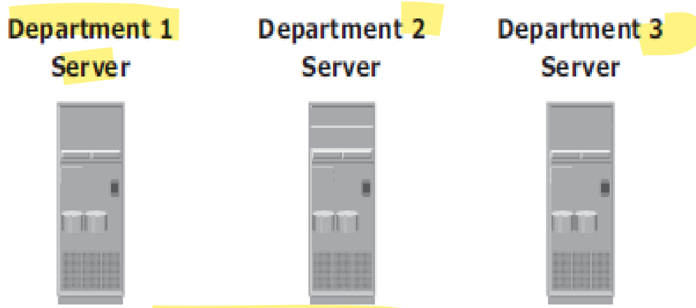
Historically, organizations had centralized computers (mainframes) and information storage devices (tape reels and disk packs) in their data center.

The evolution of open systems, their affordability, and ease of deployment made it possible for business units/departments to have their own servers and storage.



Evolution of Storage Architecture

- In earlier implementations of open systems, the storage was typically internal to the server.
- These storage devices could not be shared with any other servers.
- This approach is referred to as server-centric storage architecture
- Each server has a limited number of storage devices, and any administrative tasks, such as maintenance of the server or increasing storage capacity, might result in unavailability of information.

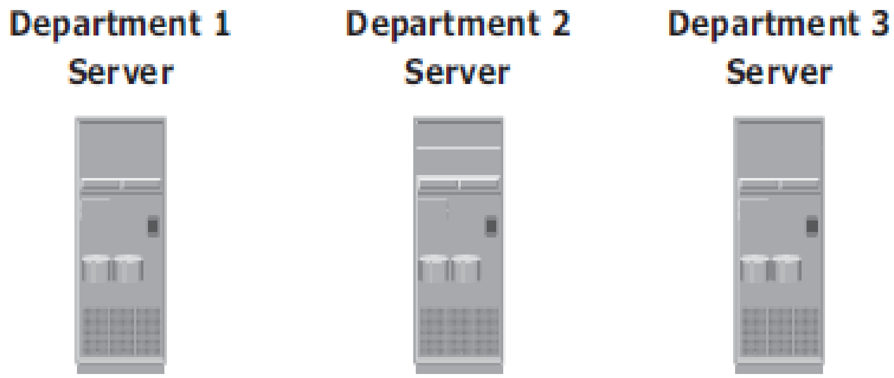


(a) Server-Centric Storage Architecture

Evolution of Storage Architecture

The proliferation of departmental servers in an enterprise resulted in:

- unprotected,
- unmanaged,
- fragmented islands of information
- increased capital
- operating expenses

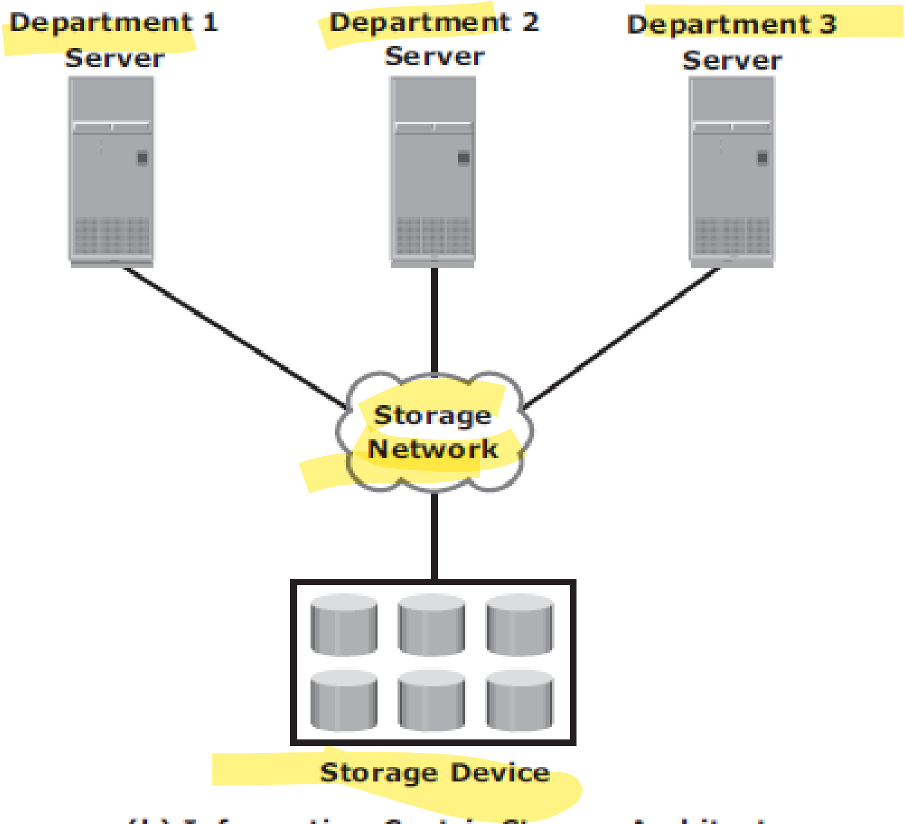


(a) Server-Centric Storage Architecture

Evolution of Storage Architecture

To overcome these challenges, storage evolved from server-centric to information- centric architecture

In this architecture, storage devices are managed centrally and independent of servers.

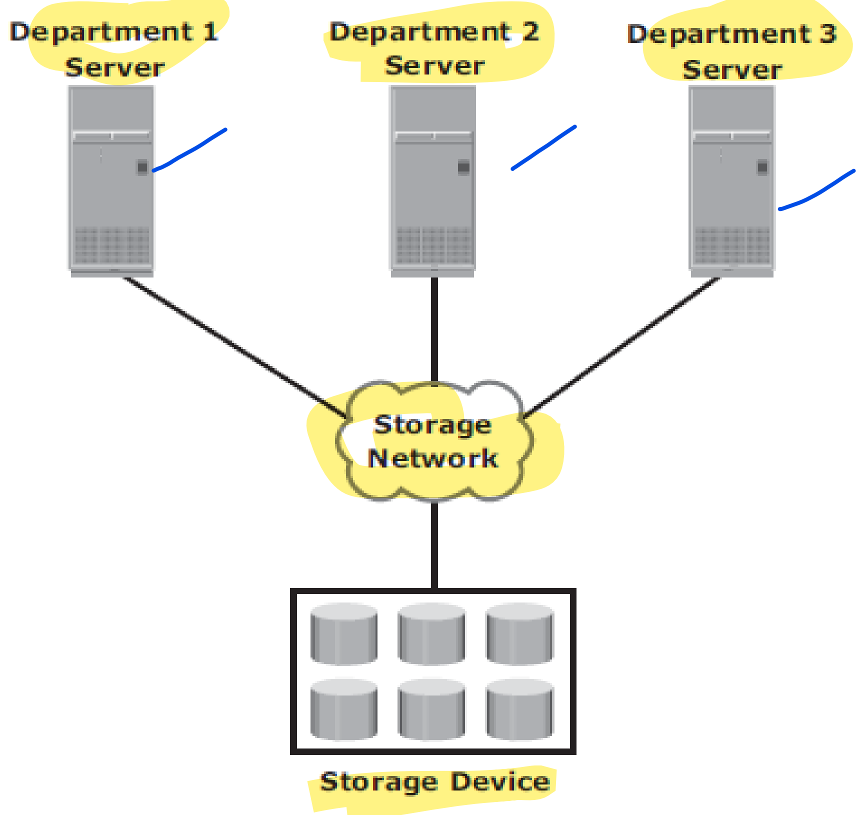


(b) Information-Centric Storage Architecture

RAMAIAH
Institute of Technology

Evolution of Storage Architecture

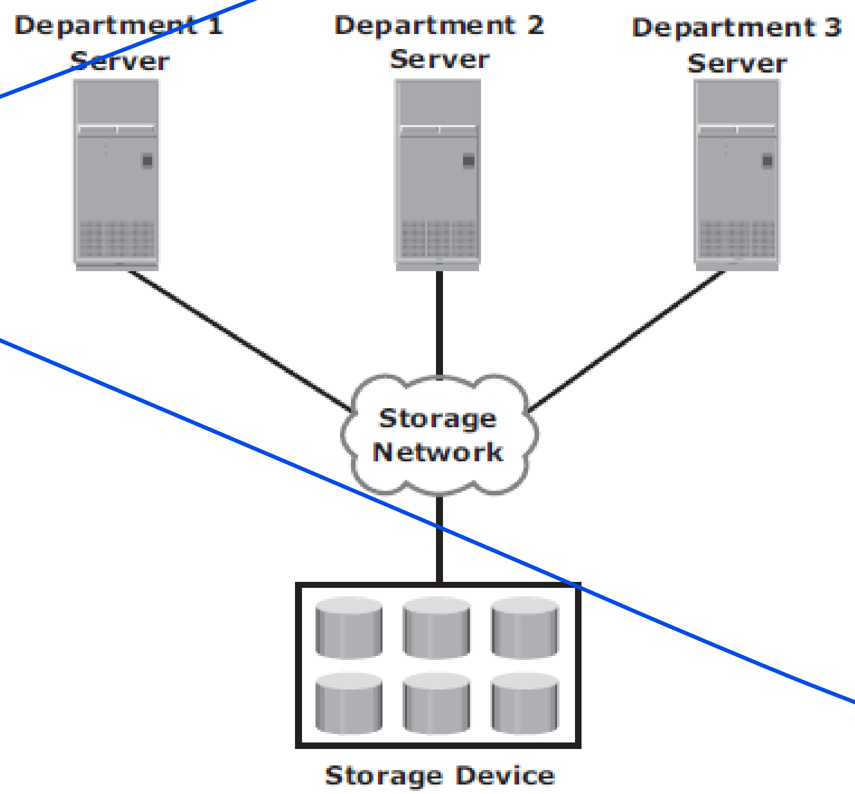
- These centrally-managed storage devices are shared with multiple servers.
- When a new server is deployed in the environment, storage is assigned from the same shared storage devices to that server.
- The capacity of shared storage can be increased dynamically by adding more storage devices without impacting information availability.
- In this architecture, information management is easier and cost-effective.



(b) Information-Centric Storage Architecture

Evolution of Storage Architecture

- These centrally-managed storage devices are shared with multiple servers.
- When a new server is deployed in the environment, storage is assigned from the same shared storage devices to that server.
- The capacity of shared storage can be increased dynamically by adding more storage devices without impacting information availability.
- In this architecture, information management is easier and cost-effective.



(b) Information-Centric Storage Architecture

UNIT 1- Introduction to Information Storage

- Information Storage
- Evolution of Storage Architecture,
- Data Centre Infrastructure
- Virtualization and Cloud Computing.

Data Centre Infrastructure

- Data centers manage large amounts of data.
- The data center infrastructure includes
  - Hardware components - computers, storage systems, network devices, and power backups;
  - Software components - applications, operating systems, and management software.
  - Environmental controls - air conditioning, fire suppression, and ventilation.
- Large organizations often maintain more than one data center to distribute data processing workloads and provide backup if a disaster occurs.

RAMAIAH Institute of Technology

List and explain the core elements and key characteristics of a data center.

A C M S S P D

Data Centre Infrastructure- Core Elements of a Data Center

Application: A computer program that provides the logic for computing operations

Database management system (DBMS): Provides a structured way to store data in logically organized tables that are interrelated

Host or compute: A computing platform (hardware, firmware, and software) that runs applications and databases

Network: A data path that facilitates communication among various networked devices

Storage: A device that stores data persistently for subsequent use

Data Centre Infrastructure- **Core Elements of a Data Center**

Figure shows an example of an online order transaction system that involves the five core elements of a data center and illustrates their functionality in a business process.

1. Application
2. Database management system (DBMS)
3. Host or computer
4. Network
5. Storage

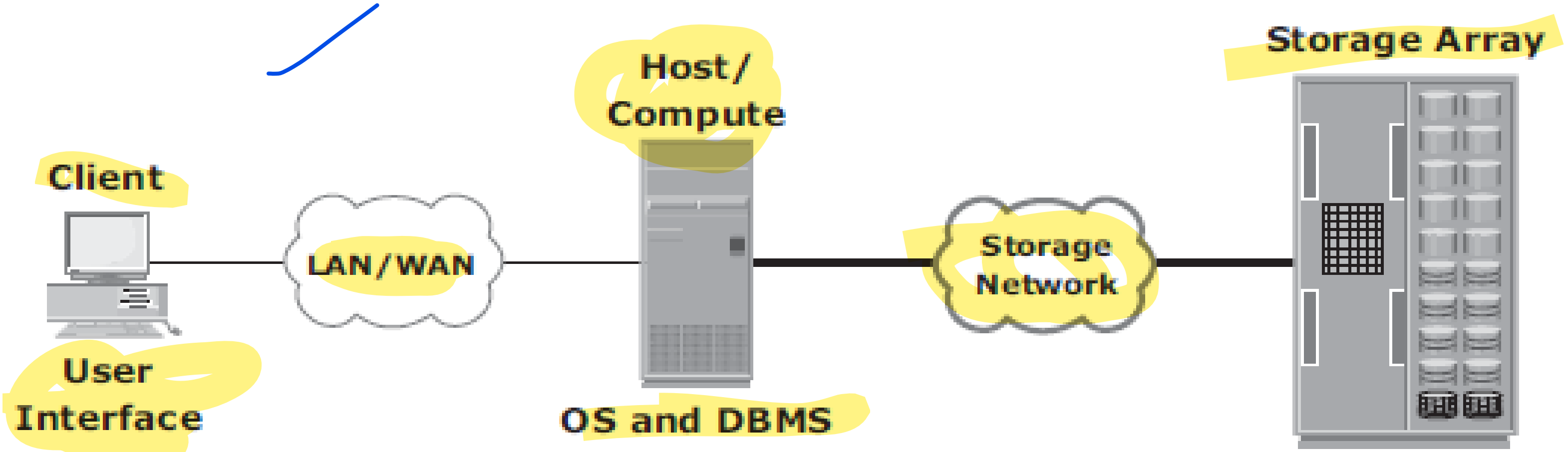


Figure 1-5: Example of an online order transaction system

Data Centre Infrastructure- **Core Elements of a Data Center**

- A customer places an order through a client machine connected over a LAN/WAN to a host running an order-processing application.
- The client accesses the DBMS on the host through the application to provide order-related information, such as the customer name, address, payment method, products ordered, and quantity ordered.

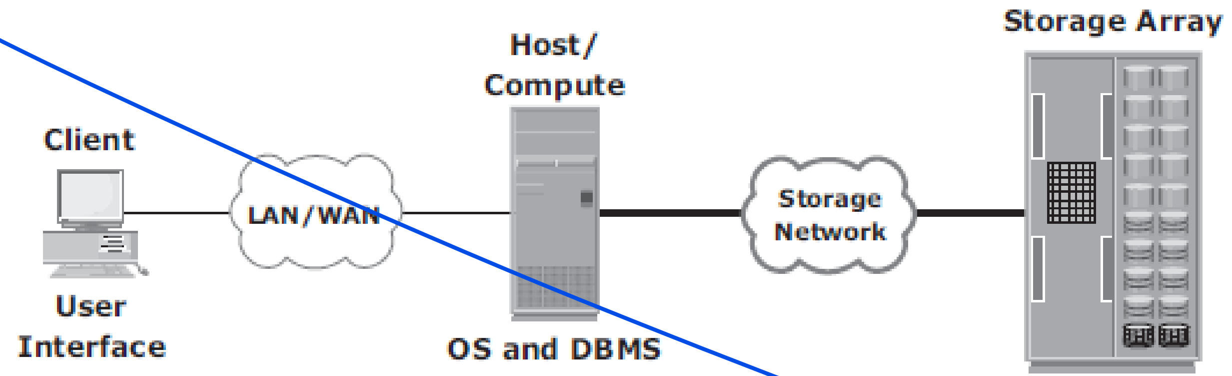


Figure 1-5: Example of an online order transaction system

Data Centre Infrastructure- **Core Elements of a Data Center**

- The DBMS uses the host operating system to write this data to the physical disks in the storage array.
- The storage networks provide the communication link between the host and the storage array and transports the request to read or write data between them.
- The storage array, after receiving the read or write request from the host, performs the necessary operations to store the data on physical disks.

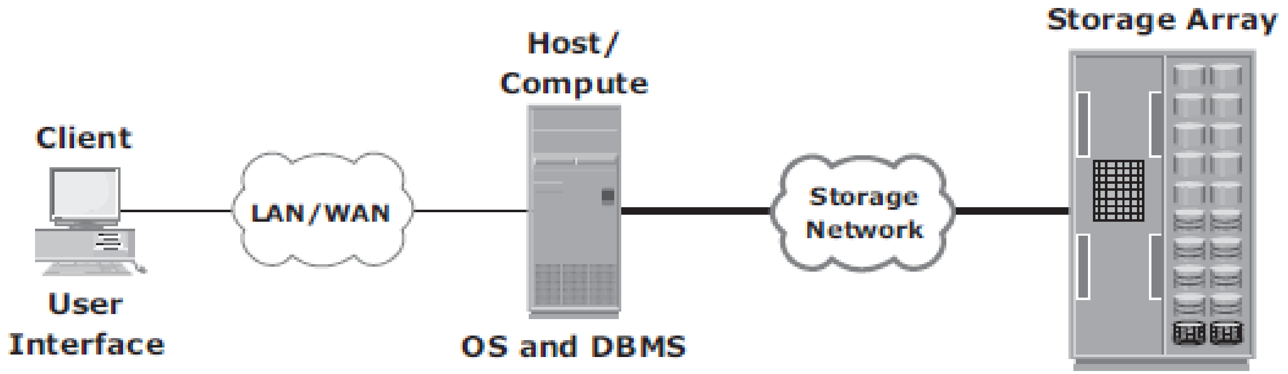


Figure 1-5: Example of an online order transaction system

A C M S S P D

Data Centre Infrastructure- Key Characteristics of a Data Center

Availability: A data center should ensure the availability of information when required. Unavailability of information could cost millions of dollars per hour to businesses, such as financial services, telecommunications, and e-commerce.

Security: Data centers must establish policies, procedures, and core element integration to prevent unauthorized access to information.

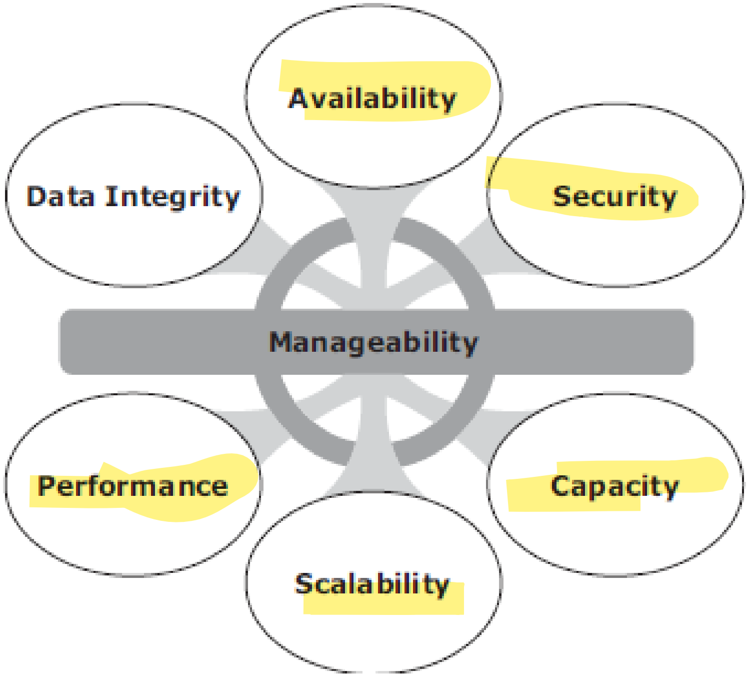


Figure 1-6: Key characteristics of a data center

Data Centre Infrastructure- **Key Characteristics of a Data Center**

Scalability: Business growth often requires deploying more servers, new applications, and additional databases. Data center resources should scale based on requirements, without interrupting business operations.

Performance: All the elements of the data center should provide optimal performance based on the required service levels.

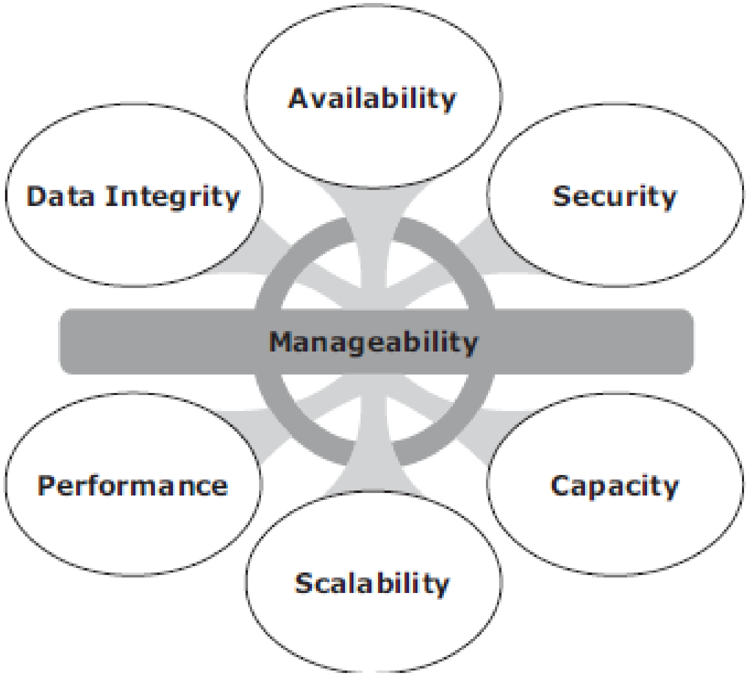


Figure 1-6: Key characteristics of a data center

Data Centre Infrastructure- **Key Characteristics of a Data Center**

Data integrity: Data integrity refers to mechanisms, such as error correction codes or parity bits, which ensure that data is stored and retrieved exactly as it was received.

Capacity: When capacity requirements increase, the data center must provide additional capacity without interrupting availability or with minimal disruption. Capacity may be managed by reallocating the existing resources or by adding new resources.

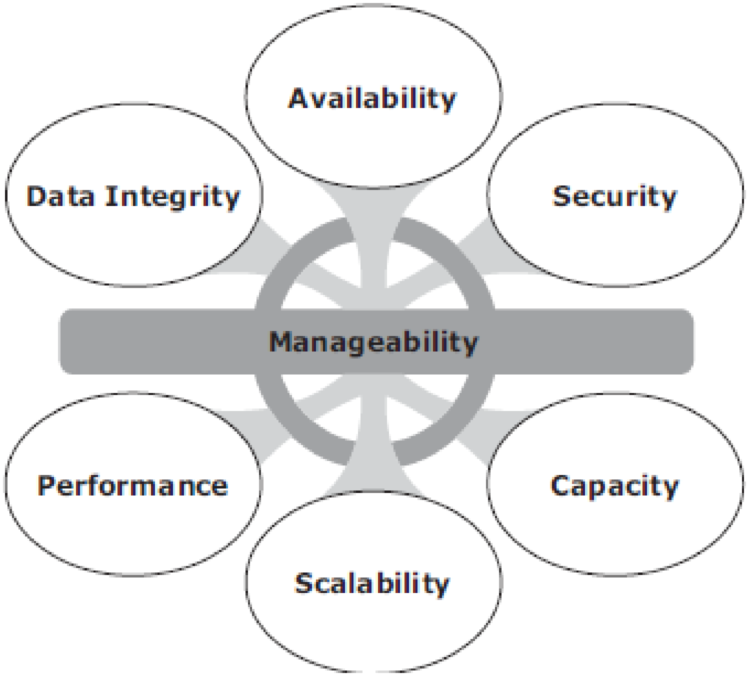


Figure 1-6: Key characteristics of a data center

RAMAIAH
Institute of Technology

Data Centre Infrastructure- **Key Characteristics of a Data Center**

Manageability: A data center should provide easy and integrated management of all its elements.

Manageability can be achieved through automation and reduction of human (manual) intervention in common tasks.

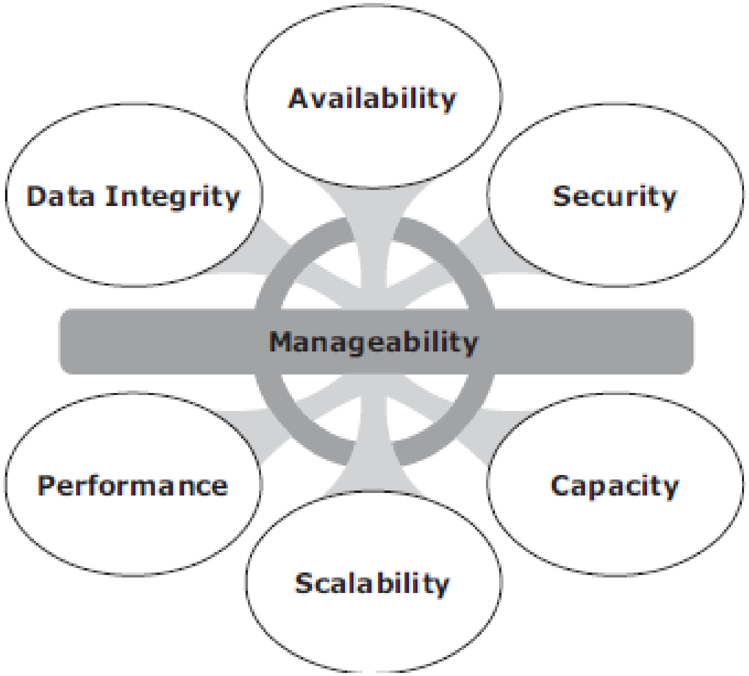


Figure 1-6: Key characteristics of a data center

Data Centre Infrastructure- **Managing a Data Center**

The key management activities include the following task:

Monitoring: It is a continuous process of gathering information on various elements and services running in a data center. The aspects of a data center that are monitored include security, performance, availability, and capacity.

Reporting: It is done periodically on resource performance, capacity, and utilization. Reporting tasks help to establish business justifications and chargeback of costs associated with data center operations.

Data Centre Infrastructure- **Managing a Data Center**

The key management activities include the following task:

Provisioning: It is a process of providing the hardware, software, and other resources required to run a data center.

Provisioning activities primarily include resources management to meet capacity, availability, performance, and security requirements.

UNIT 1- Introduction to Information Storage

- Information Storage
- Evolution of Storage Architecture,
- Data Centre Infrastructure
- Virtualization and Cloud Computing.

Virtualization and Cloud Computing

Virtualization is a technique of abstracting physical resources, such as computer,storage, and network, and making them appear as logical resources.

Examples of virtualization are virtual memory used on compute systems and partitioning of raw disks.

RAMAIAH
Institute of Technology

Virtualization and Cloud Computing

Virtualization enables pooling of physical resources and providing an aggregated view of the physical resource capabilities.

For example, storage virtualization enables multiple pooled storage devices to appear as a single large storage entity.

Similarly, by using compute virtualization, the CPU capacity of the pooled physical servers can be viewed as the aggregation of the power of all CPUs (in megahertz).

Virtualization also enables centralized management of pooled resources.

Virtualization and Cloud Computing

Virtual resources can be created and provisioned from the pooled physical resources.

For example, a virtual disk of a given capacity can be created from a storage pool or a virtual server with specific CPU power and memory can be configured from a compute pool.

These virtual resources share pooled physical resources, which improves the utilization of physical IT resources.

RAMAIAH
Institute of Technology

Virtualization and Cloud Computing

Based on business requirements, capacity can be added to or removed from the virtual resources without any disruption to applications or users.

Virtualization and Cloud Computing

Cloud computing, addresses these challenges efficiently.

Cloud computing enables individuals or businesses to use IT resources as a service over the network.

It provides highly scalable and flexible computing that enables provisioning of resources on demand.

Users can scale up or scale down the demand of computing resources, including storage capacity, with minimal management effort or service provider interaction.

Virtualization and Cloud Computing

Cloud computing, addresses these challenges efficiently.

Cloud computing enables individuals or businesses to use IT resources as a service over the network.

It provides highly scalable and flexible computing that enables provisioning of resources on demand.

Users can scale up or scale down the demand of computing resources, including storage capacity, with minimal management effort or service provider interaction.

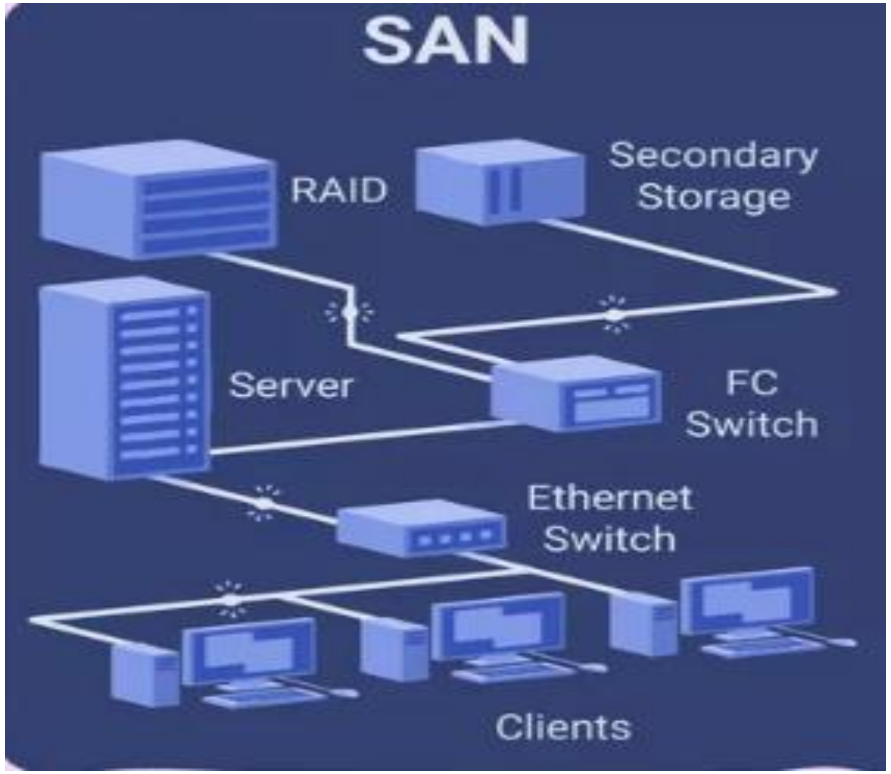
UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Introduction

SAN is a local network of several devices.



1. **Application**
2. **Database management system (DBMS)**
3. **Host or computer**
4. **Network**
5. **Storage**

Application

An application is a computer program that provides the logic for computing operations.

The application sends requests to the underlying operating system to perform read/write (R/W) operations on the storage devices.

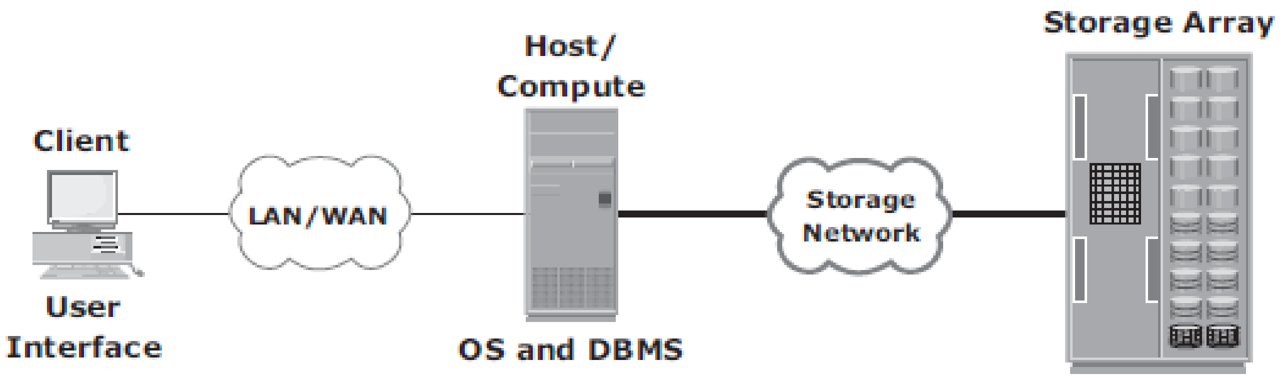


Figure 1-5: Example of an online order transaction system

Application

Applications can be layered on the database, which in turn uses the OS services to perform R/W operations on the storage devices.

Applications are categorized as

- business applications,
- Infrastructure management applications,
- data protection applications, and
- security applications.

Some examples of these applications are e-mail, enterprise resource planning (ERP), decision support system (DSS), resource management, backup, authentication and antivirus applications, and so on.

Database Management System (DBMS)

A database is a structured way to store data in logically organized tables that are interrelated.

A database helps to optimize the storage and retrieval of data.

A DBMS controls the creation, maintenance, and use of a database.

The DBMS processes an application’s request for data and instructs the operating system to transfer the appropriate data from the storage.

Host (Compute)

Users store and retrieve data through applications.

The computers on which these applications run are referred to as hosts or compute systems.

RAMAIAH Institute of Technology

Host (Compute)

- A host consists of CPU, memory, I/O devices, and a collection of software to perform computing operations.
- The CPU consists of four components: Arithmetic Logic Unit (ALU), control unit, registers, and L1 cache. There are two types of memory on a host, Random Access Memory (RAM) and Read-Only Memory (ROM). I/O devices enable communication with a host. Examples of I/O devices are keyboard, mouse, monitor, etc.
- This software includes the operating system, file system, logical volume manager, device drivers, and so on.

Host (Compute)

- Examples of physical hosts include desktop computers, servers or a cluster of servers, laptops, and mobile devices.
- Hosts can be physical or virtual machines.
- Various software components that are essential parts of a host system are:
  - **Operating System**
  - **Device Driver**
  - **Volume Manager**
  - **File System**
  - **Compute Virtualization**

Host (Compute) - Operating System

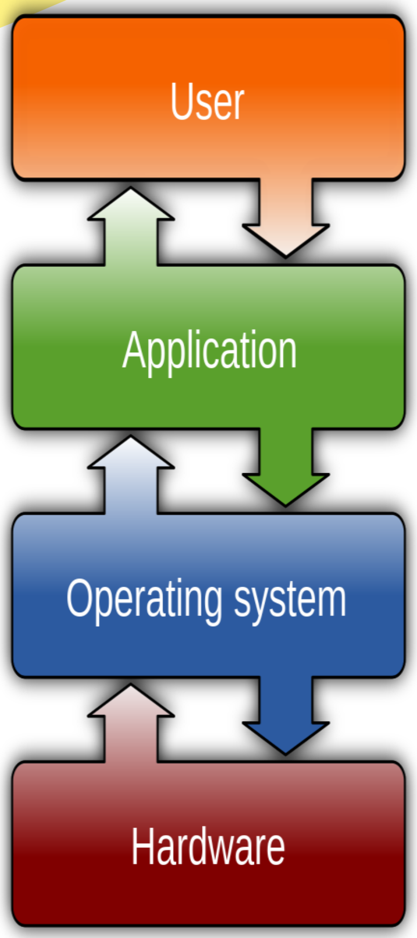
An operating system controls all aspects of computing.

It works between the application and the physical components of a compute system.

The operating system also monitors and responds to user actions and the environment.

It organizes and controls hardware components and manages the allocation of hardware resources.

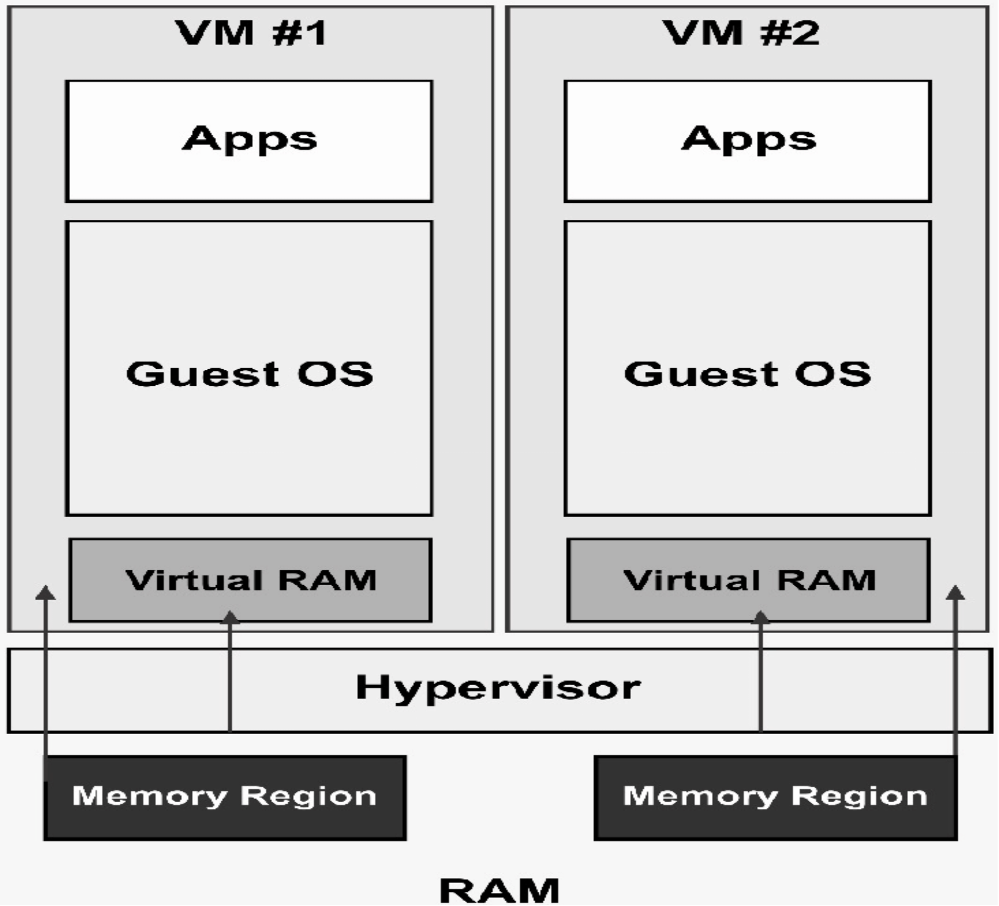
An operating system also performs basic storage management tasks while managing other underlying components, such as the file system, volume manager, and device drivers.



Host (Compute) - Operating System

In a virtualized compute environment, the virtualization layer works between the operating system and the hardware resources.

OS works as a guest and performs only the activities related to application interaction.



Host (Compute) - Operating System

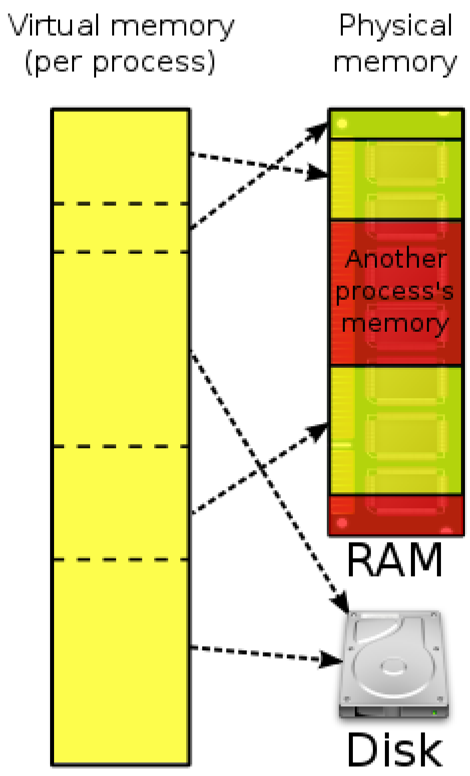
Memory Virtualization

Memory is an expensive component of a host.

It determines both the size and number of applications that can run on a host.

Memory virtualization is an operating system feature that virtualizes the physical memory (RAM) of a host.

It creates virtual memory with an address space larger than the physical memory space present in the compute system.



RAMAIAH Institute of Technology

Host (Compute) - Operating System

Memory Virtualization

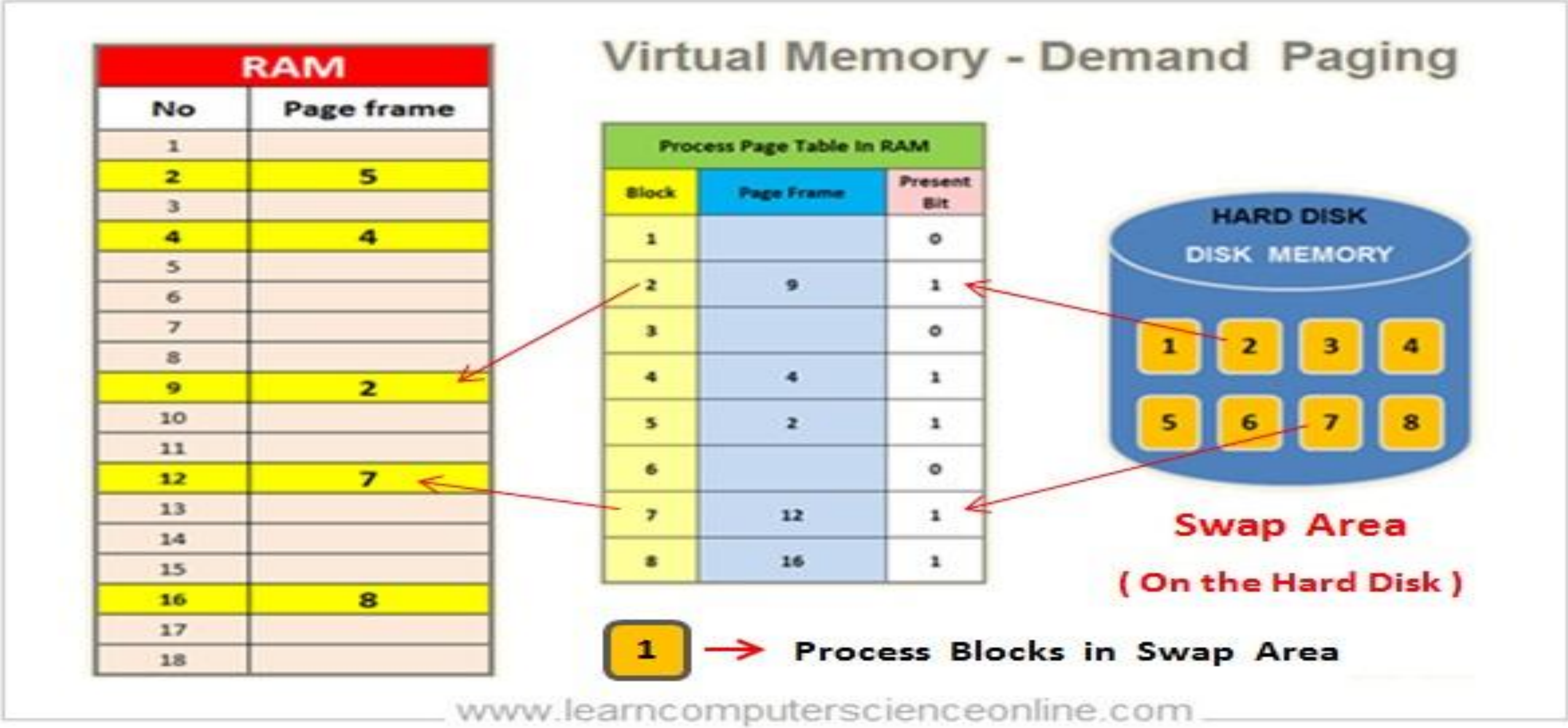
The operating system utility that manages the virtual memory is known as the virtual memory manager (VMM).

The VMM manages the virtual-to-physical memory mapping and fetches data from the disk storage when a process references a virtual address that points to data at the disk storage.

The space used by the VMM on the disk is known as a swap space.

A swap space (also known as page file or swap file) is a portion of the disk drive that appears to be physical memory to the operating system.

Virtual Memory - Demand Paging



RAMAIAH Institute of Technology

Host (Compute) - Operating System

Memory Virtualization

In a virtual memory implementation, the memory of a system is divided into contiguous blocks of fixed-size pages.

A process known as paging moves inactive physical memory pages onto the swap file and brings them back to the physical memory when required.

This enables efficient use of the available physical memory among different applications.

RAMAIAH Institute of Technology

Host (Compute) - Operating System

Memory Virtualization

The operating system typically moves the least used pages into the swap file so that enough RAM is available for processes that are more active.

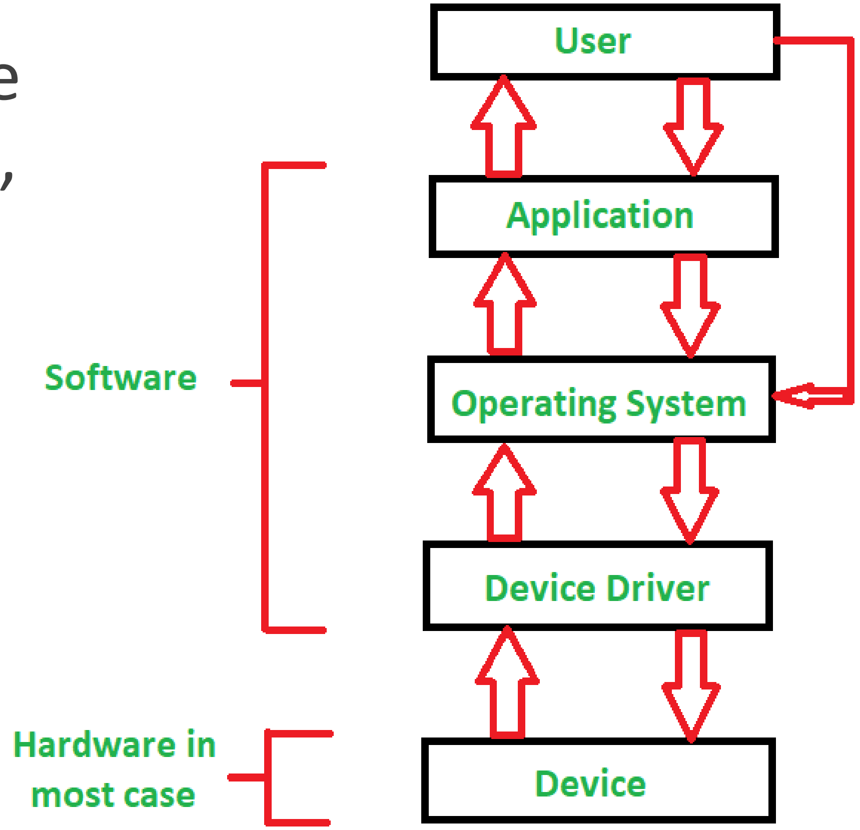
Access to swap file pages is slower than access to physical memory pages because swap file pages are allocated on the disk drive, which is slower than physical memory.

Host (Compute) - Device Driver

A device driver is special software that permits the operating system to interact with a specific device, such as a printer, a mouse, or a disk drive.

A device driver enables the operating system to recognize the device and to access and control devices.

Device drivers are hardware-dependent and operating-system-specific.



Host (Compute)-Volume Manager

- In the early days, disk drives appeared to the operating system as a number of continuous disk blocks.
- The entire disk drive would be allocated to the file system or other data entity used by the operating system or application.
- The disadvantage was lack of flexibility.
  - When a disk drive ran out of space, there was no easy way to extend the file system's size.
- Also, as the storage capacity of the disk drive increased, allocating the entire disk drive for the file system often resulted in underutilization of storage capacity.

Justify the role of Logical Volume Managers (LVMs) in providing dynamic file system expansion and efficient storage management.

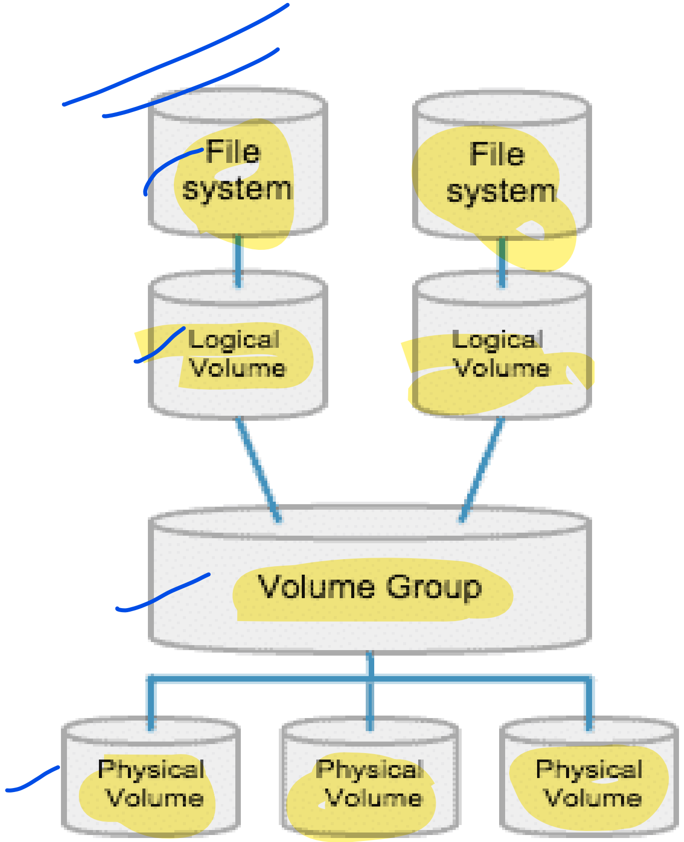
RAMAIAH
Institute of Technology

Host (Compute)-Volume Manager

The evolution of Logical Volume Managers (LVMs) enabled dynamic extension of file system capacity and efficient storage management.

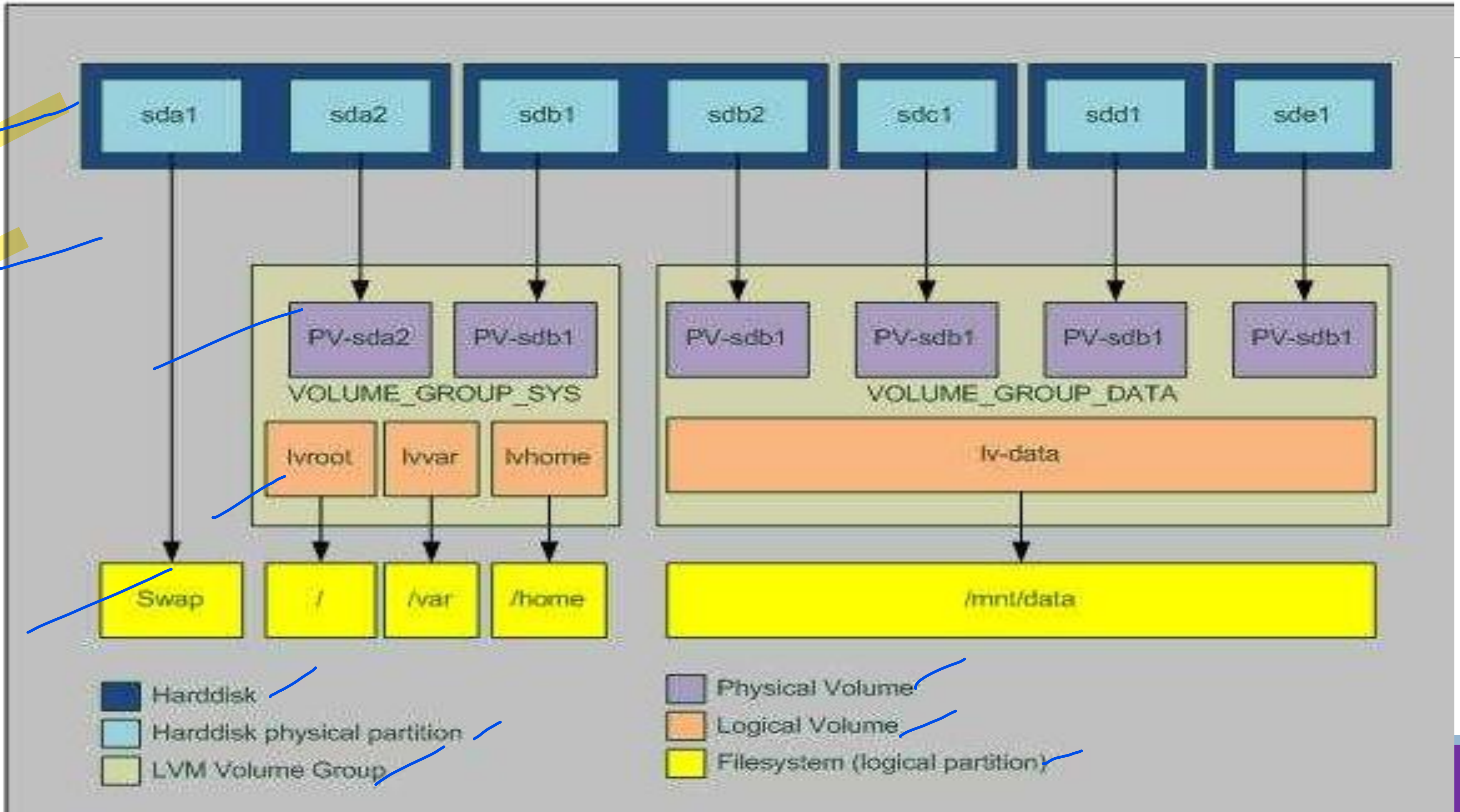
The LVM is software that runs on the compute system and manages logical and physical storage.

LVM is an intermediate layer between the file system and the physical disk.



RAMAIAH
Institute of Technology

Host (Compute)-Volume Manager

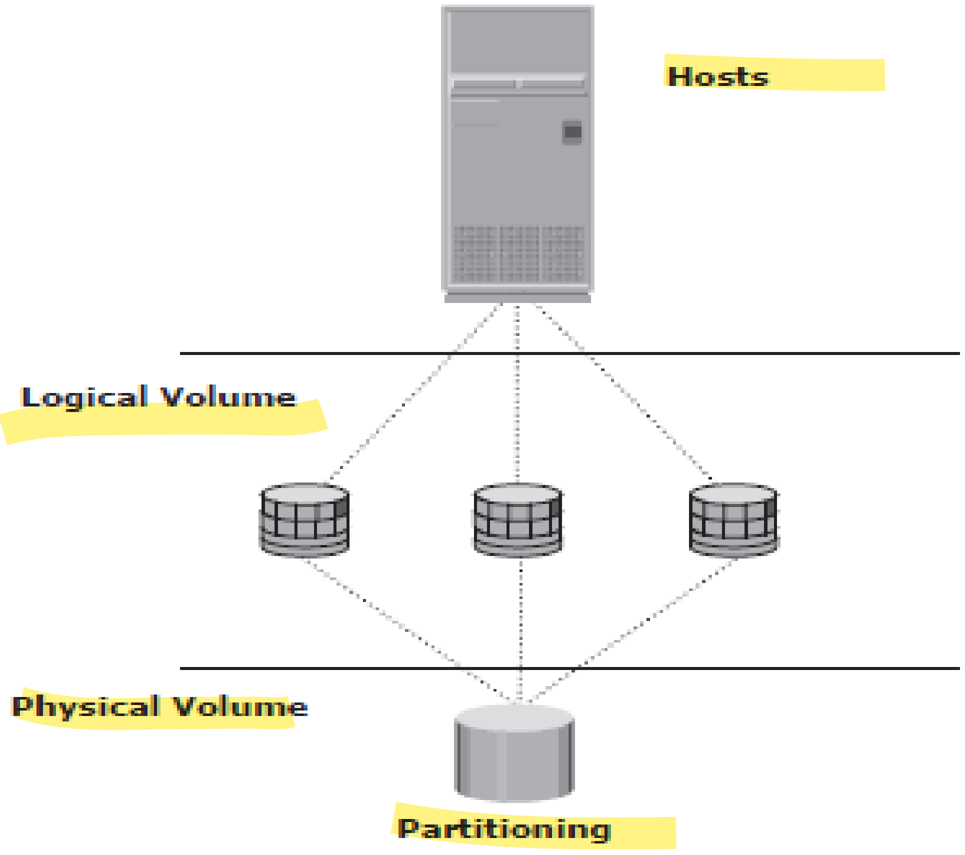


Host (Compute)-Volume Manager

Partitioning - It can partition a larger-capacity disk into virtual logical containers called logical volumes (LVs)

The partitions are created from groups of contiguous cylinders when the hard disk is initially set up on the host.

The host's file system accesses the logical volumes without any knowledge of partitioning and physical structure of the disk.

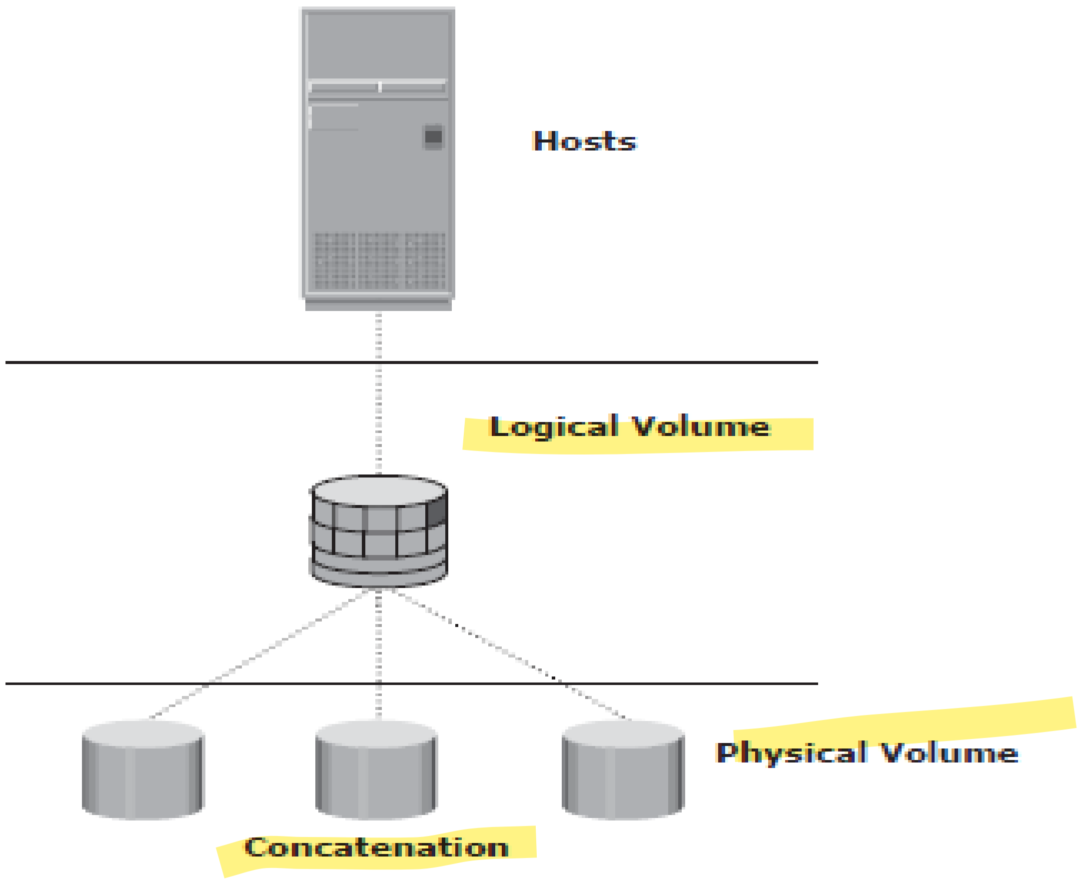


RAMAIAH
Institute of Technology

Host (Compute)-Volume Manager

Concatenation -It can aggregate several smaller disks to form a larger virtual volume.

These volumes are then presented to applications.



Host (Compute) - File System

A file is a collection of related records or data stored as a unit with a name.

A file system is a hierarchical structure of files.

A file system enables easy access to data files residing within a disk drive, a disk partition, or a logical volume.

A file system consists of logical structures and software routines that control access to files.

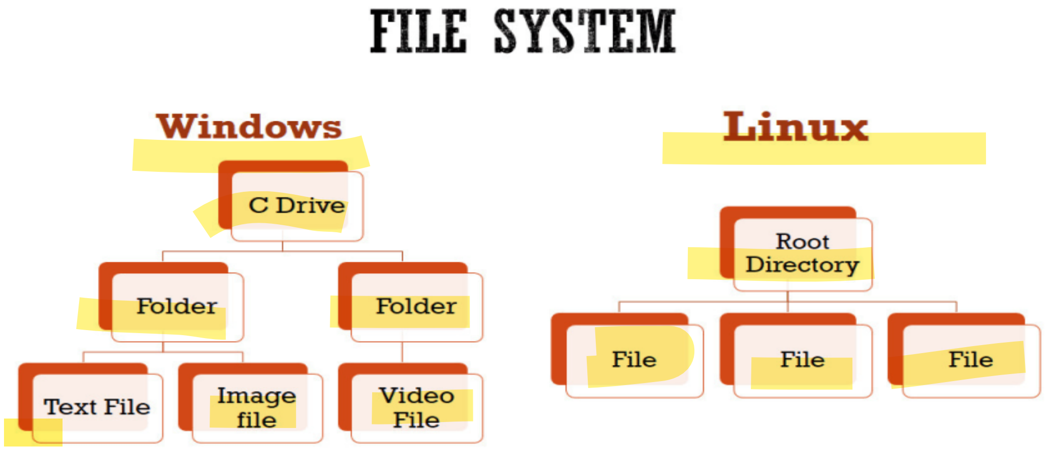
Host (Compute) - File System

A file is a collection of related records or data stored as a unit with a name.

A file system is a hierarchical structure of files.

A file system enables easy access to data files residing within a disk drive, a disk partition, or a logical volume.

A file system consists of logical structures and software routines that control access to files.



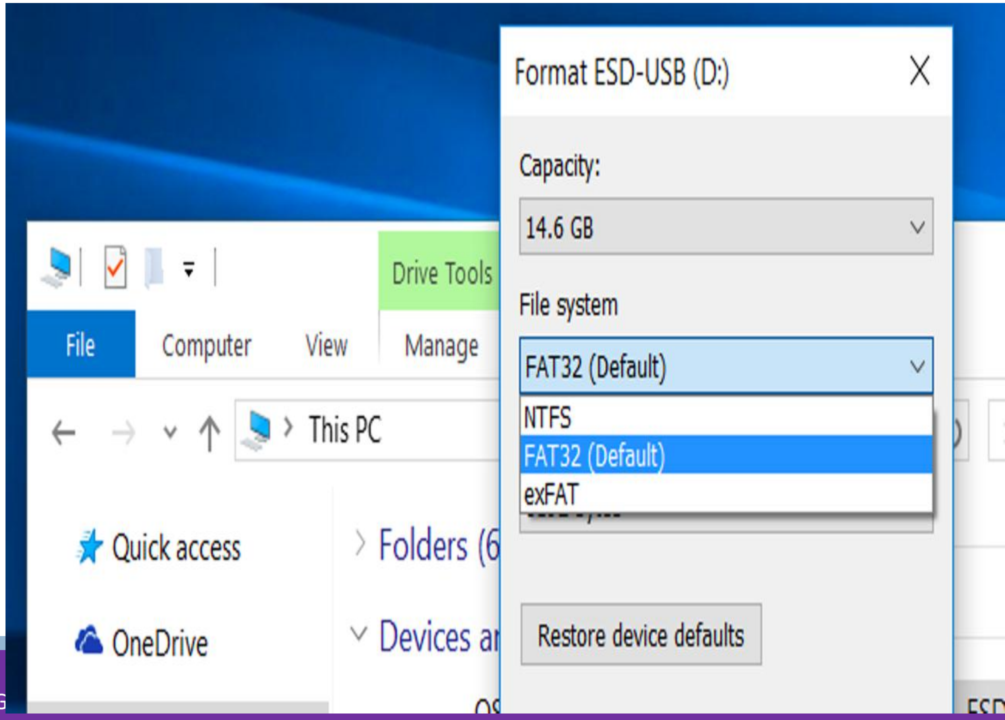
Host (Compute) - File System

A file system organizes data in a structured hierarchical manner via the use of directories, which are containers for storing pointers to multiple files.

All file systems maintain a pointer map to the directories, subdirectories, and files that are part of the file system.

Examples of common file systems are:

- FAT 32 (File Allocation Table) for Microsoft Windows
- NT File System (NTFS) for Microsoft Windows
- UNIX File System (UFS) for UNIX
- Extended File System (EXT2/3) for Linux



Explain the process of mapping user files to disk storage with the help of a neat diagram. Describe how files are translated from user-level structures to physical disk locations.

Host (Compute) - File System

The following list shows the process of mapping user files to the disk storage subsystem with an LVM

1. Files are created and managed by users and applications.

2. These files reside in the file systems.

3. The file systems are mapped to file system blocks.

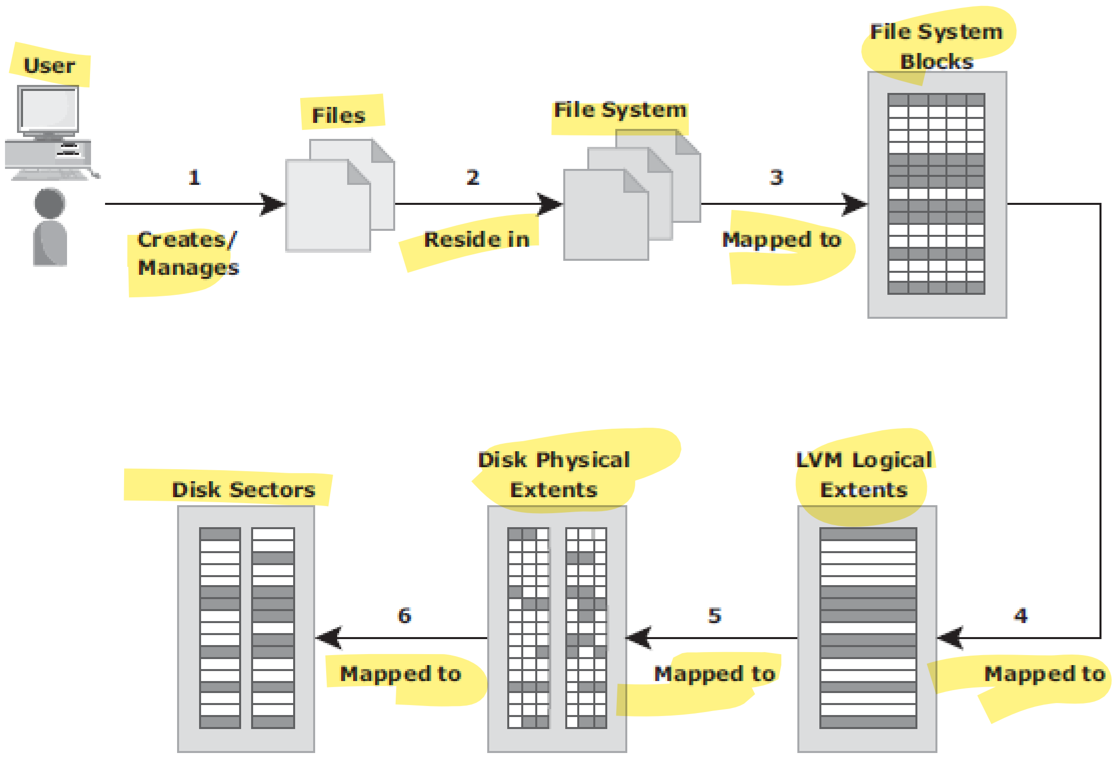


Figure 2-2: Process of mapping user files to disk storage

Host (Compute) - File System

4. The file system blocks are mapped to logical extents of a logical volume.

5. These logical extents in turn are mapped to the disk physical extents either by the operating system or by the LVM.

6. These physical extents are mapped to the disk sectors in a storage subsystem.

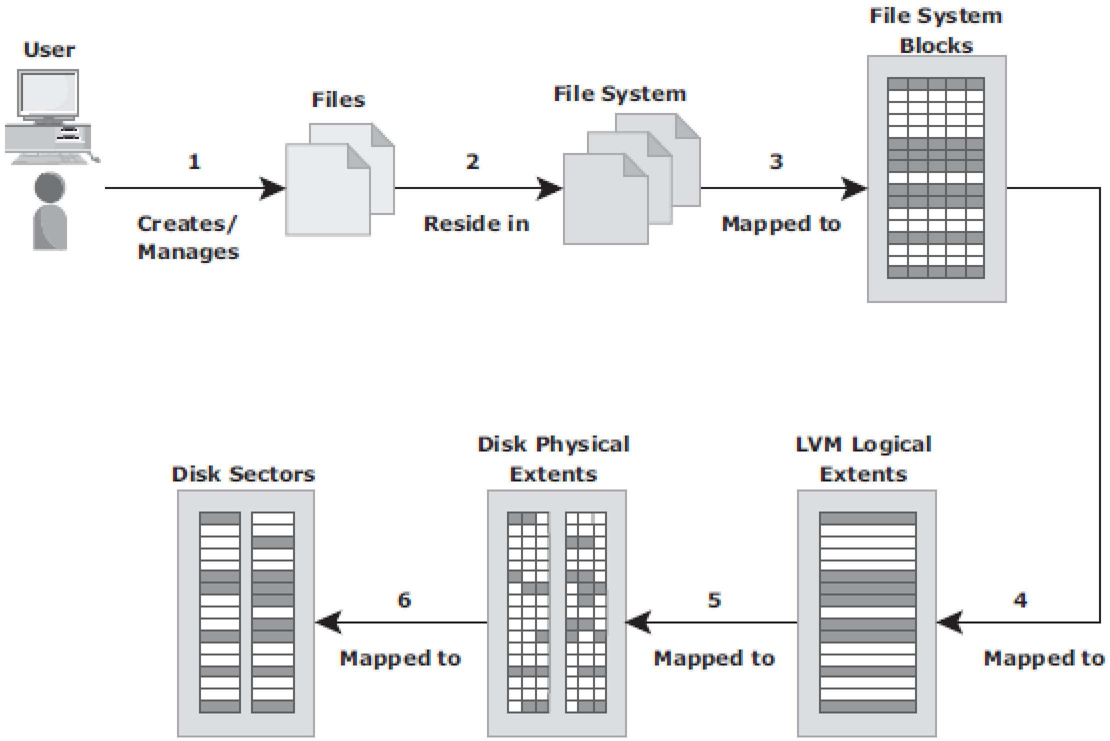


Figure 2-2: Process of mapping user files to disk storage

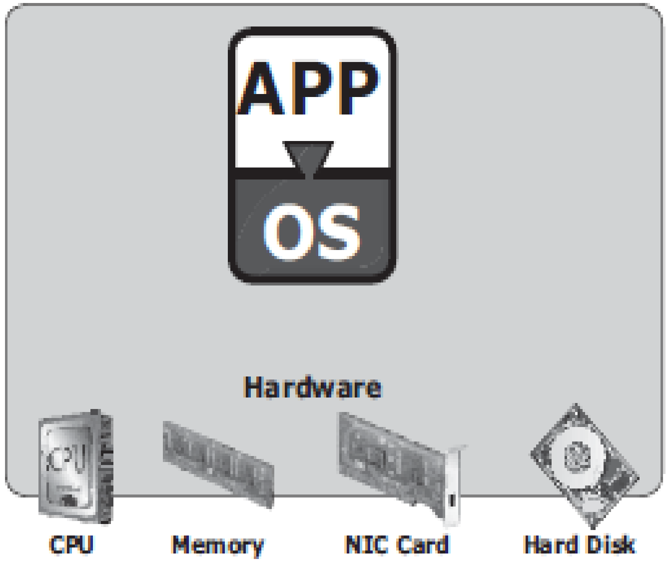
Host (Compute) - Compute Virtualization

- Compute virtualization is a technique for masking or abstracting the physical hardware from the operating system.
- It enables multiple operating systems to run concurrently on single or clustered physical machines.
- This technique enables creating portable virtual compute systems called virtual machines (VMs).
- A virtual machine is a logical entity but appears like a physical host to the operating system, with its own CPU, memory, network controller, and disks.

Host (Compute) - Compute Virtualization

The servers are limited to serve only one application at a time, as shown in Figure.

This causes organizations to purchase new physical machines for every application they deploy, resulting in expensive and inflexible infrastructure.

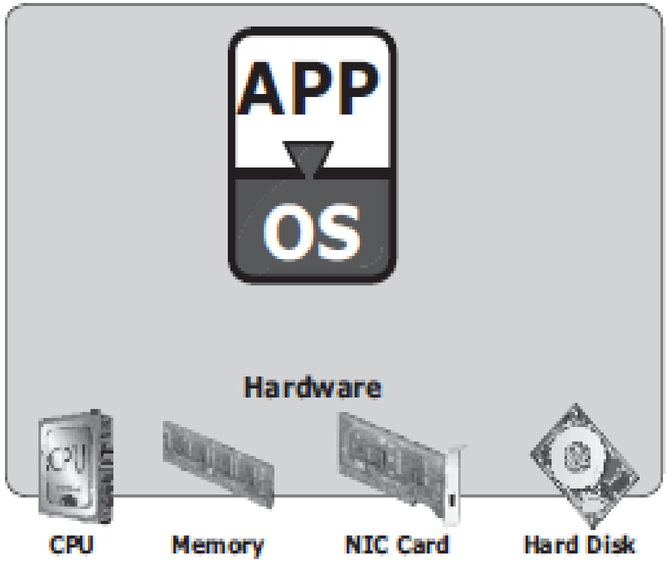


(a) Before Compute Virtualization

Host (Compute) - Compute Virtualization

Many applications do not take full advantage of the hardware capabilities available to them.

Consequently, resources such as processors, memory, and storage remain underutilized.

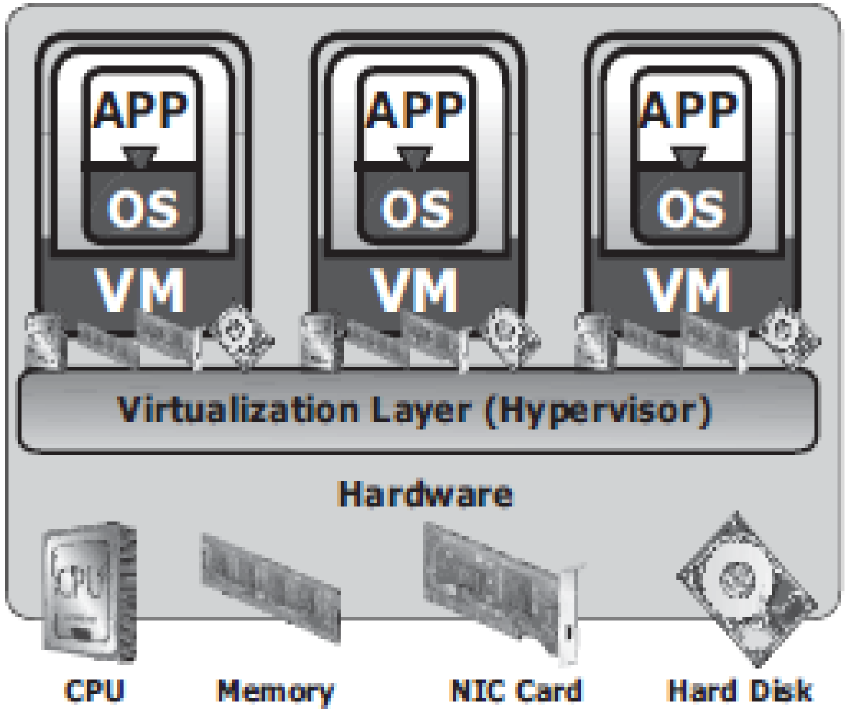


(a) Before Compute Virtualization

Host (Compute) - Compute Virtualization

Compute virtualization enables users to overcome these challenges (see Figure)by allowing multiple operating systems and applications to run on a single physical machine.

This technique significantly improves server utilization and provides server consolidation.



(b) After Compute Virtualization

RAMAIAH Institute of Technology

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Connectivity

Connectivity refers to the interconnection

- between hosts
- between a host and peripheral devices, such as printers or storage devices.
- between the host and the storage device.

Connectivity and communication between host and storage are enabled using physical components and interface protocols.

Connectivity - Physical Components of Connectivity

Connectivity refers to the interconnection between hosts or between a host and peripheral devices, such as printers or storage devices.

Connectivity between the host and the storage device.

Connectivity and communication between host and storage are enabled using physical components and interface protocols.

Connectivity - Physical Components of Connectivity

The physical components of connectivity are the hardware elements that connect the host to storage.

Three physical components of connectivity between the host and storage are the

1. host interface device
2. port
3. cable

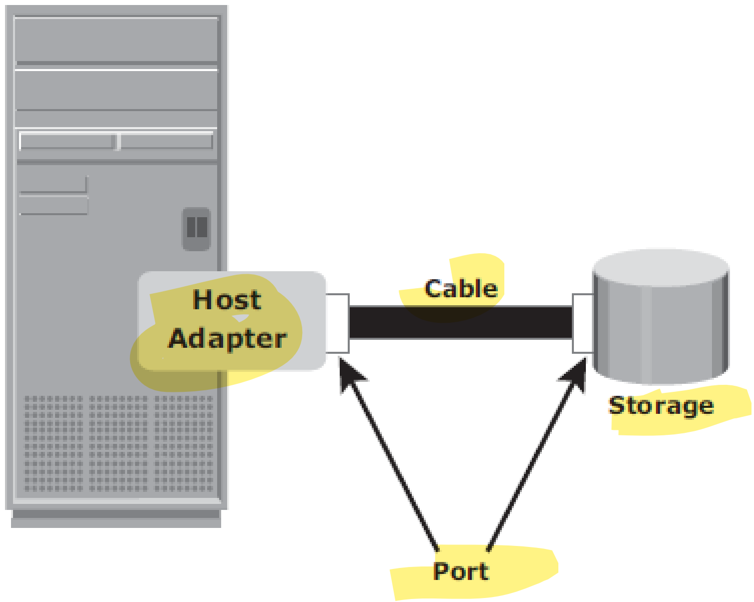
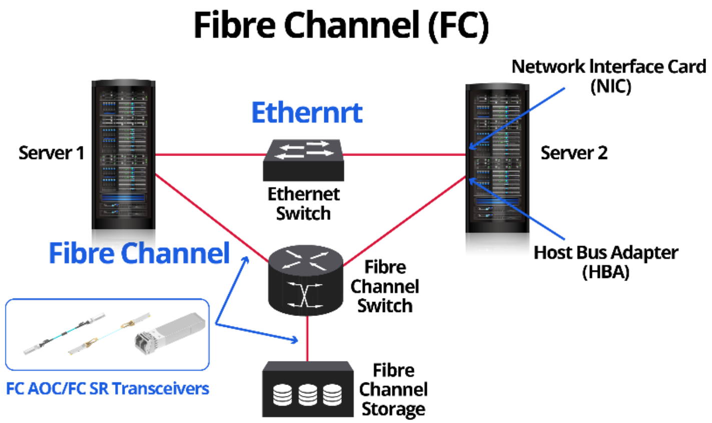


Figure 2-4: Physical components of connectivity

Connectivity - Physical Components of Connectivity

Host interface device or host adapter: connects a host to other hosts and storage devices.



Examples of host interface devices are

- host bus adapter (HBA)
- network interface card (NIC).

Host bus adaptor is an application-specific integrated circuit (ASIC) board that performs I/O interface functions between the host and storage, relieving the CPU from additional I/O processing workload.

A host typically contains multiple HBAs.

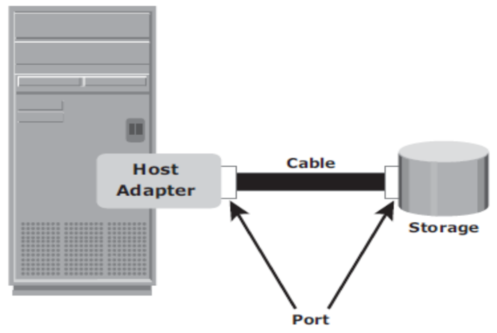


Figure 2-4: Physical components of connectivity

RAMAIAH
Institute of Technology

Connectivity - Physical Components of Connectivity

A port is a specialized outlet that enables connectivity between the host and external devices.

An HBA may contain one or more ports to connect the host to the storage device.

Cables connect hosts to internal or external devices using copper or fiber optic media.

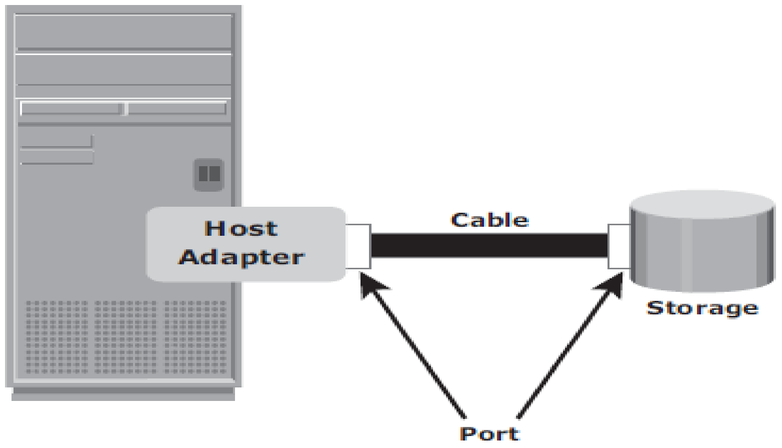


Figure 2-4: Physical components of connectivity

RAMAIAH Institute of Technology

Identify and describe the popular interface protocols used for host-to-storage communication

IDE/ATA
SCSI
FibreChannel (FC)
IP

Connectivity - Interface Protocols

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

1. Integrated Device Electronics/Advanced Technology Attachment (IDE/ATA),
2. Small Computer System Interface (SCSI),
3. Fibre Channel (FC)
4. Internet Protocol (IP).

Connectivity - Interface Protocols - IDE/ATA and Serial ATA

- IDE/ATA is a popular interface protocol standard used for connecting storage devices, such as disk drives and CD-ROM drives.
- This protocol supports parallel transmission and therefore is also known as Parallel ATA (PATA) or simply ATA.
- IDE/ATA has a variety of standards and names.
- The Ultra DMA/133 version of ATA supports a throughput of 133 MB per second.

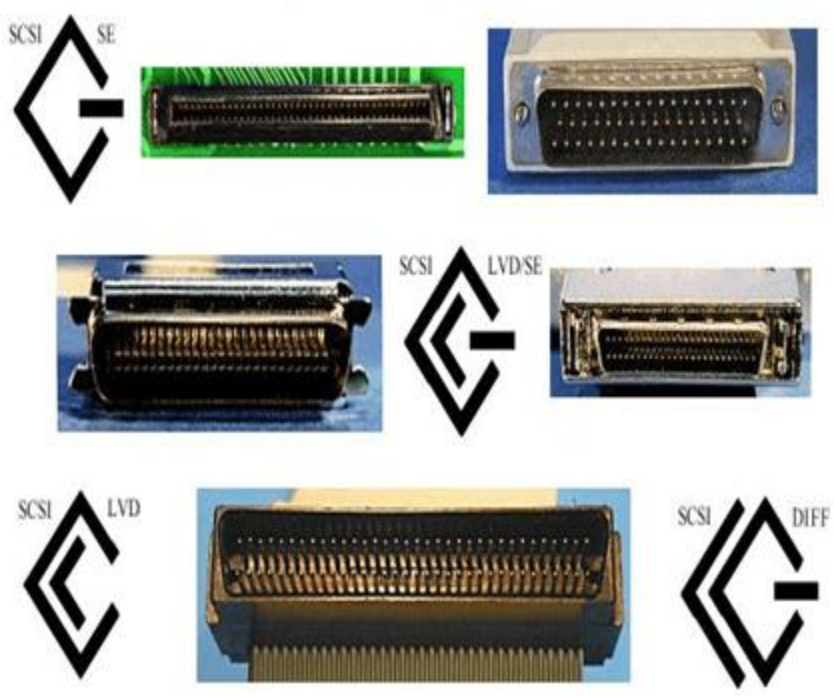
RAMAIAH
Institute of Technology

Connectivity - Interface Protocols - IDE/ATA and Serial ATA

- In a master-slave configuration, an ATA interface supports two storage devices per connector.
- The serial version of this protocol supports single bit serial transmission and is known as Serial ATA (SATA).
- High performance and low cost SATA has largely replaced PATA in newer systems.
- SATA revision 3.0 provides a data transfer rate up to 6 Gb/s.

Connectivity - Interface Protocols - SCSI and Serial SCSI

- **Small Computer Systems Interface (SCSI)** has emerged as a preferred connectivity protocol in high-end computers.
- This protocol supports parallel transmission and offers improved performance, scalability, and compatibility compared to ATA.
- However, the high cost associated with SCSI limits its popularity among home or personal desktop users.



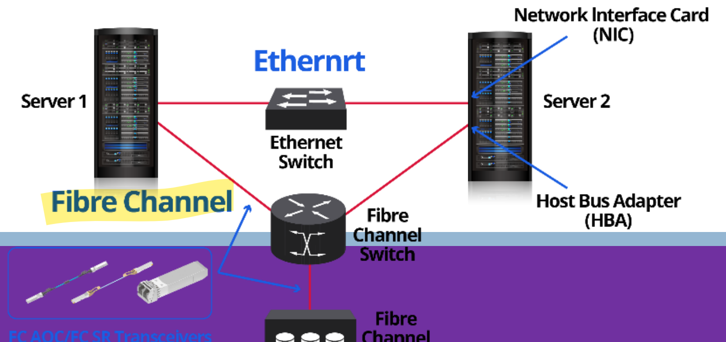
Connectivity - Interface Protocols - SCSI and Serial SCSI

- SCSI supports up to 16 devices on a single bus and provides data transfer rates up to 640 MB/s (for the Ultra-640 version).
- Serial attached SCSI (SAS) is a point-to-point serial protocol that provides an alternative to parallel SCSI.
- A newer version of serial SCSI (SAS 2.0) supports a data transfer rate up to 6 Gb/s.

Connectivity - Interface Protocols - Fibre Channel

- Fibre Channel is a widely used protocol for high-speed communication to the storage device.
- The Fibre Channel interface provides gigabit network speed.
- It provides a serial data transmission that operates over copper wire and optical fiber.
- The latest version of the FC interface (16FC) allows transmission of data up to 16 Gb/s.

Fibre Channel (FC)



Connectivity - Interface Protocols - Internet Protocol (IP)

- IP is a network protocol that has been traditionally used for host-to-host traffic.
- With the emergence of new technologies, an IP network has become a viable option for host-to-storage communication.
- IP offers several advantages in terms of cost and maturity and enables organizations to leverage their existing IP-based network.
- iSCSI and FCIP protocols are common examples that leverage IP for host-to-storage communication.

RAMAIAH
Institute of Technology

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Storage

Storage is a core component in a data center.

A storage device uses magnetic,optic, or solid state media.

Disks, tapes, and diskettes use magnetic media,whereas CD/DVD uses optical media for storage.

Removable Flash memory or Flash drives are examples of solid state media.

Storage

Tapes were the most popular storage option for backups because of their low cost.

Tapes have various limitations in terms of performance and management, as listed here:

- Data is stored on the tape linearly along the length of the tape. Search and retrieval of data are done sequentially, and it invariably takes several seconds to access the data. As a result, random data access is slow and time-consuming. This limits tapes as a viable option for applications that require real-time, rapid access to data.

Storage

- In a shared computing environment, data stored on tape cannot be accessed by multiple applications simultaneously, restricting its use to one application at a time.
- On a tape drive, the read/write head touches the tape surface, so the tape degrades or wears out after repeated use.
- The storage and retrieval requirements of data from the tape and the overhead associated with managing the tape media are significant.

Storage

Optical disc storage is popular in small, single-user computing environments.

It is frequently used by individuals to store photos or as a backup medium on personal or laptop computers.

It is also used as a distribution medium for small applications, such as games, or as a means to transfer small amounts of data from one computer system to another.

Optical discs have limited capacity and speed, which limit the use of optical media as a business data storage solution.

Storage

- The capability to write once and read many (WORM) is one advantage of optical disc storage.
- A CD-ROM is an example of a WORM device.
- Optical discs can be used as a low-cost alternative for long-term storage of relatively small amounts of fixed content that do not change after it is created.
- Collections of optical discs in an array, called a jukebox, are still used as a fixed-content storage solution.
- Other forms of optical discs include CD-RW, Blu-ray disc, and other variations of DVD.

RAMAIAH Institute of Technology

Storage

- Disk drives are the most popular storage medium used in modern computers for storing and accessing data for performance-intensive, online applications.
- Disks support rapid access to random data locations. This means that data can be written or retrieved quickly for a large number of simultaneous users or applications.
- Disks have a large capacity.
- Disk storage arrays are configured with multiple disks to provide increased capacity and enhanced performance.

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

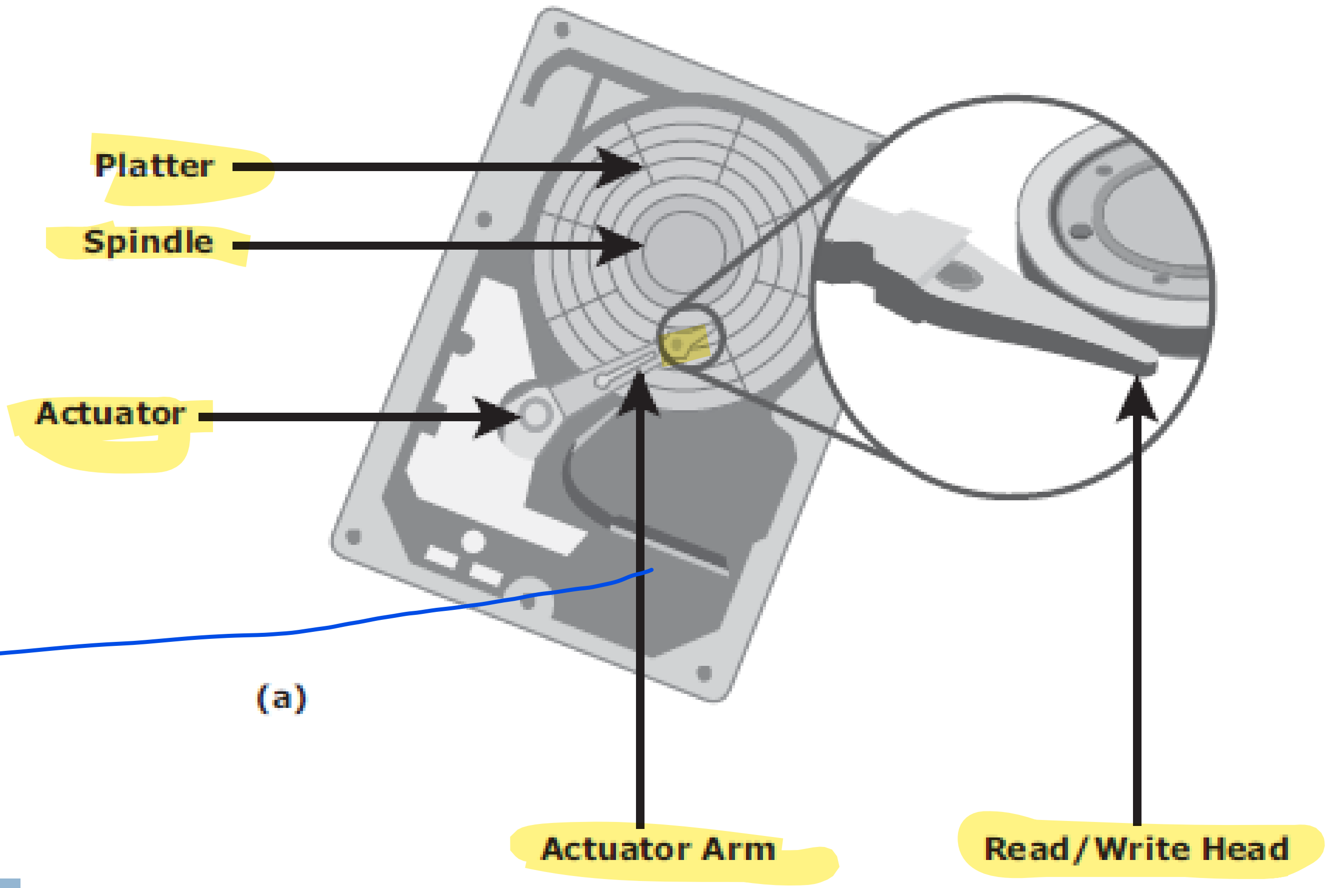
- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Describe the various components of a Hard Disk Drive (HDD) and explain the function of
each component in storing and retrieving data.

Disk Drive Components

The key components of a hard disk drive are

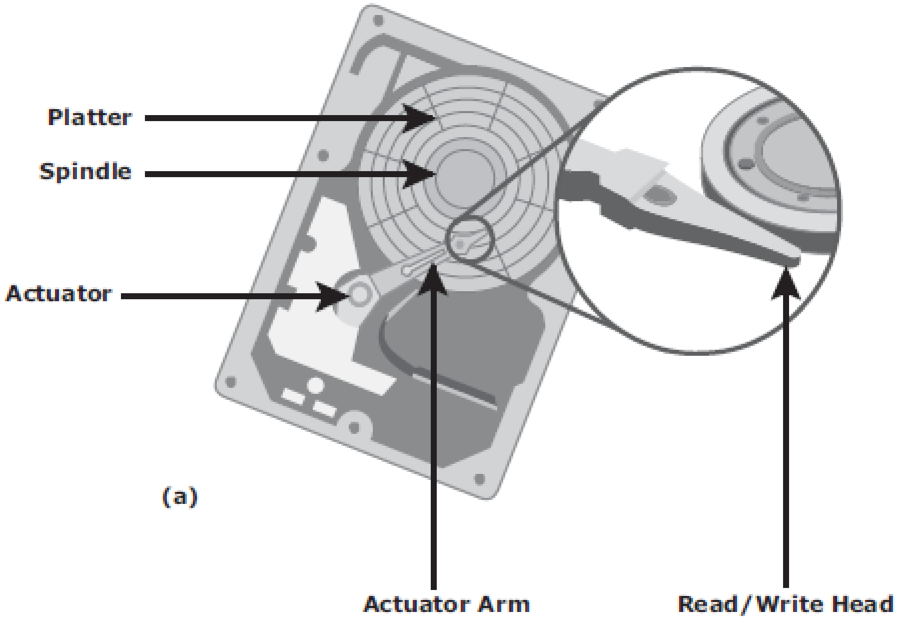
1. platter
2. spindle
3. read-write head
4. actuator arm assembly
5. controller board



(a)

Disk Drive Components

- I/O operations in a HDD are performed by rapidly moving the arm across the rotating flat platters coated with magnetic particles.
- Data is transferred between the disk controller and magnetic platters through the read-write (R/W) head which is attached to the arm.
- Data can be recorded and erased on magnetic platters any number of times.



(a)

RAMAIAH
Institute of Technology

Disk Drive Components

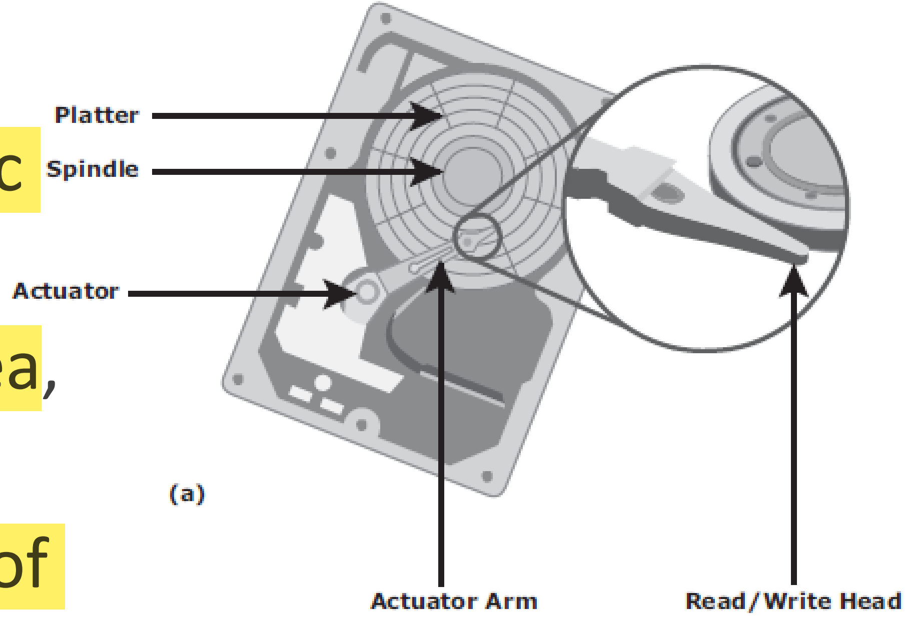
1.Platter

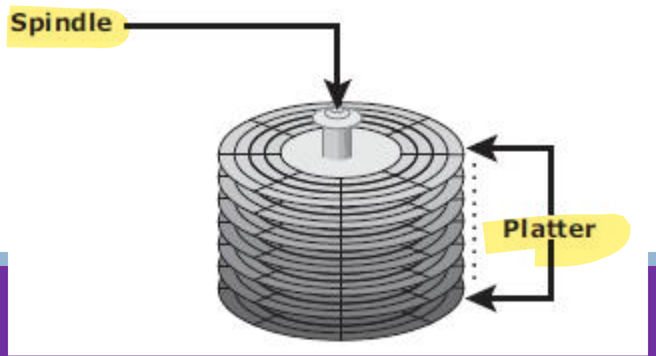
A platter is a rigid, round disk coated with magnetic material on both surfaces (top and bottom).

The data is encoded by polarizing the magnetic area, or domains, of the disk surface.

Data can be written to or read from both surfaces of the platter.

The number of platters and the storage capacity of each platter determine the total capacity of the drive.





RAMAIAH Institute of Technology

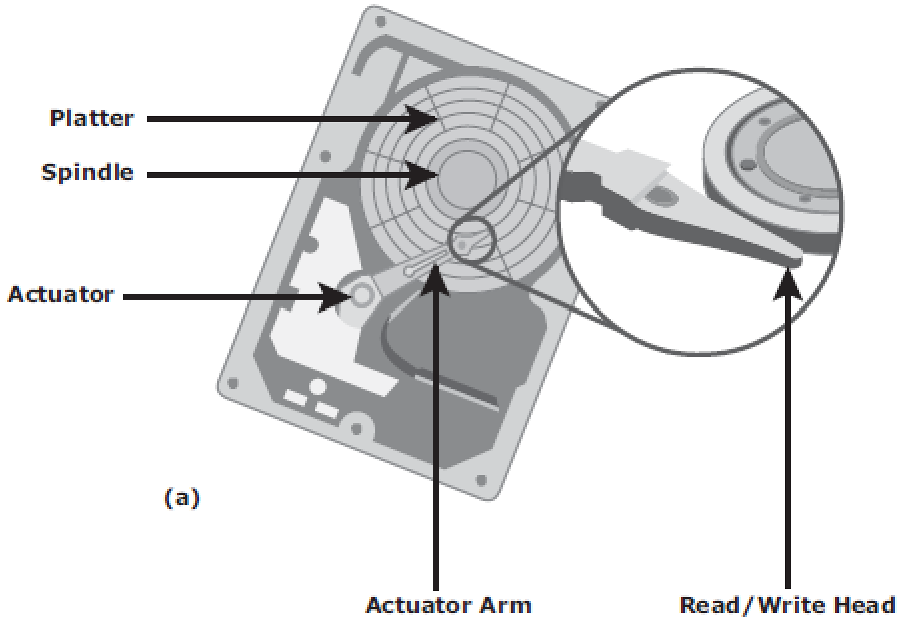
Disk Drive Components

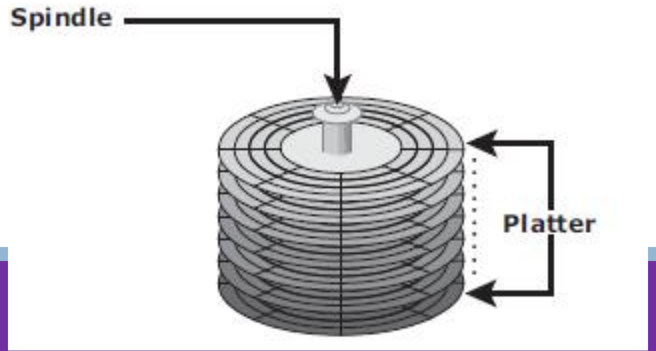
1.Platter

A typical HDD consists of one or more flat circular disks called platters.

The data is recorded on these platters in binary codes (0s and 1s).

The set of rotating platters is sealed in a case, called the Head Disk Assembly (HDA).



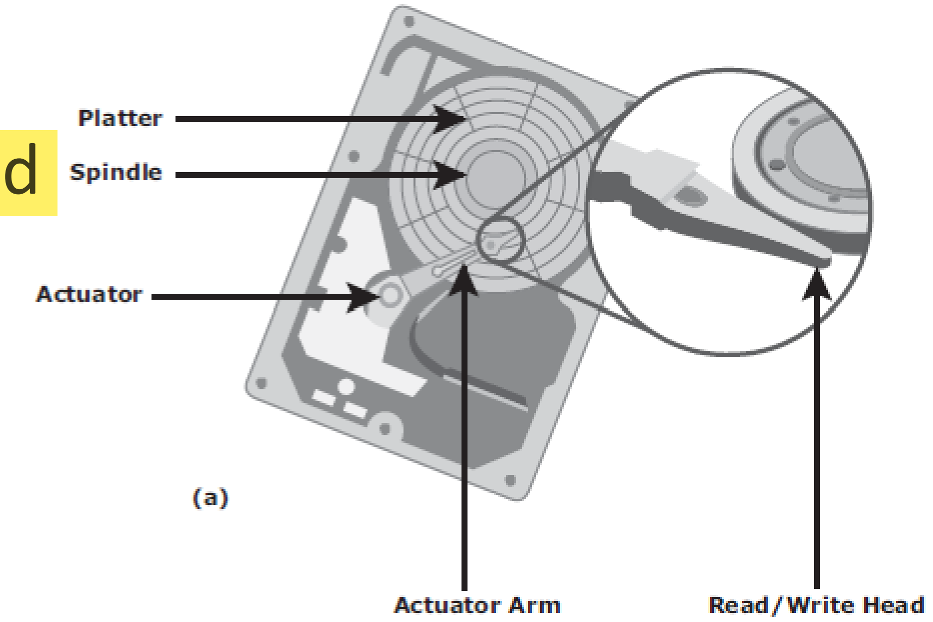


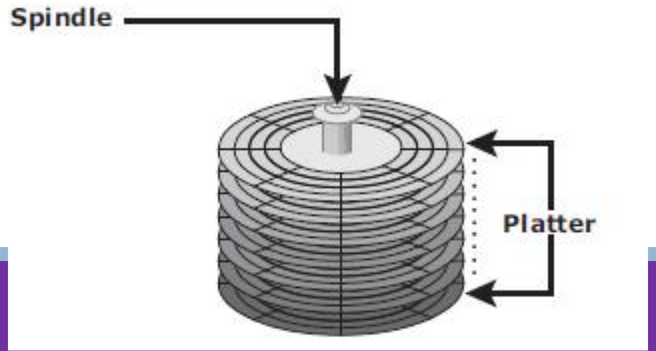
Disk Drive Components

2.Spindle

A spindle connects all the platters and is connected to a motor.

The motor of the spindle rotates with a constant speed.





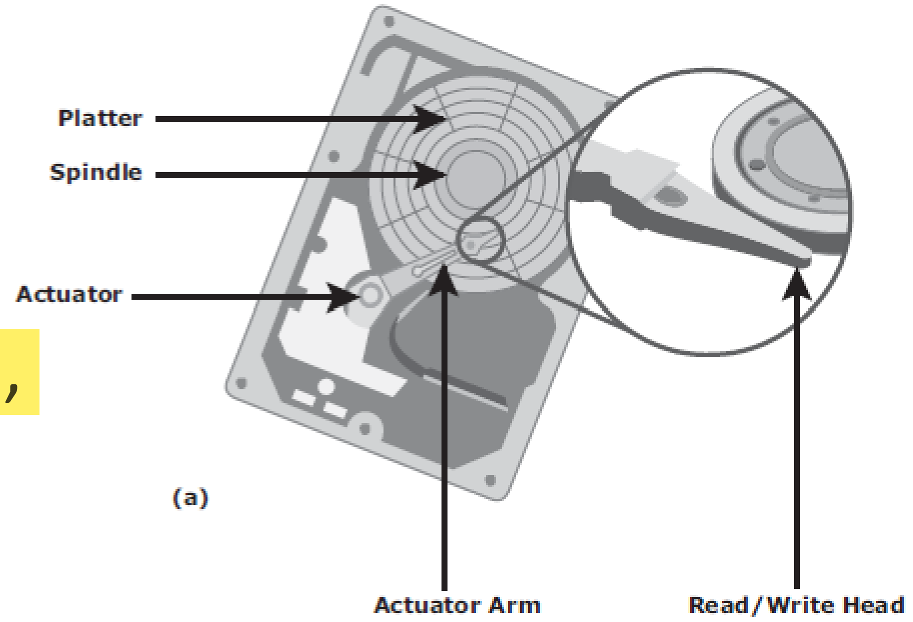
Disk Drive Components

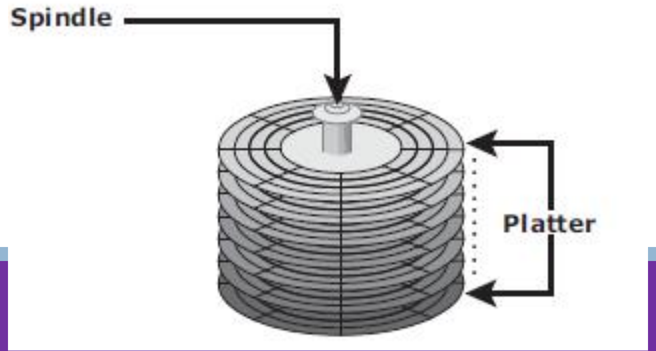
2.Spindle

The disk platter spins at a speed of several thousands of revolutions per minute (rpm).

Common spindle speeds are 5,400 rpm, 7,200 rpm, 10,000 rpm, and 15,000 rpm.

The speed of the platter is increasing with improvements in technology, although the extent to which it can be improved is limited.





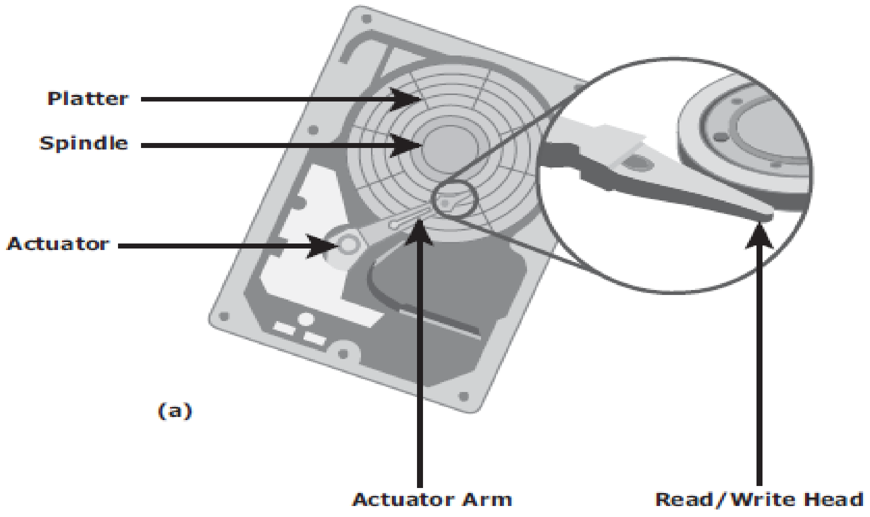
RAMAIAH
Institute of Technology

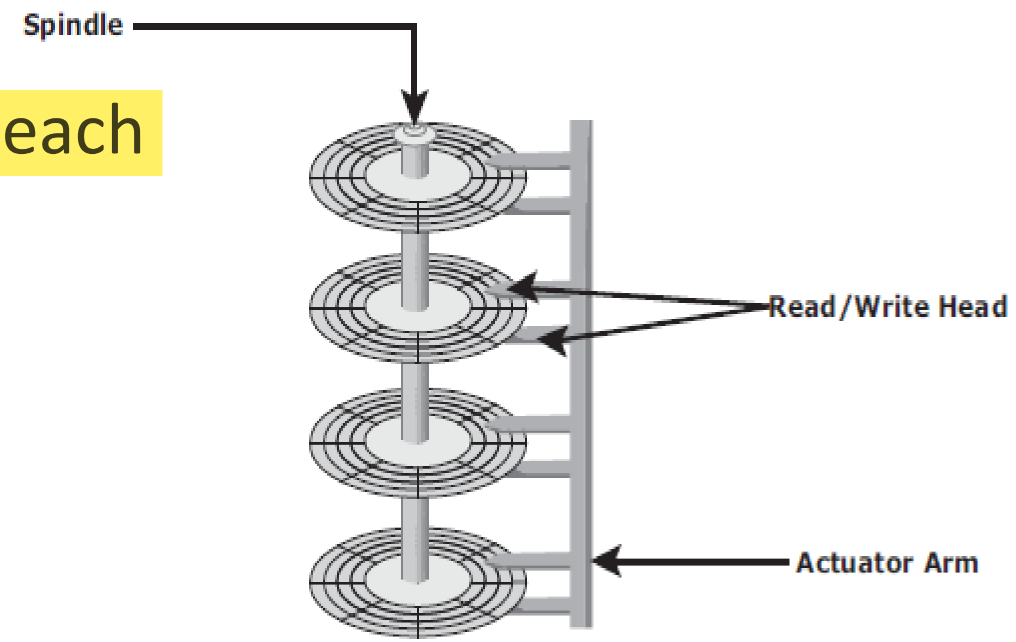
Disk Drive Components

3.Read/Write Head

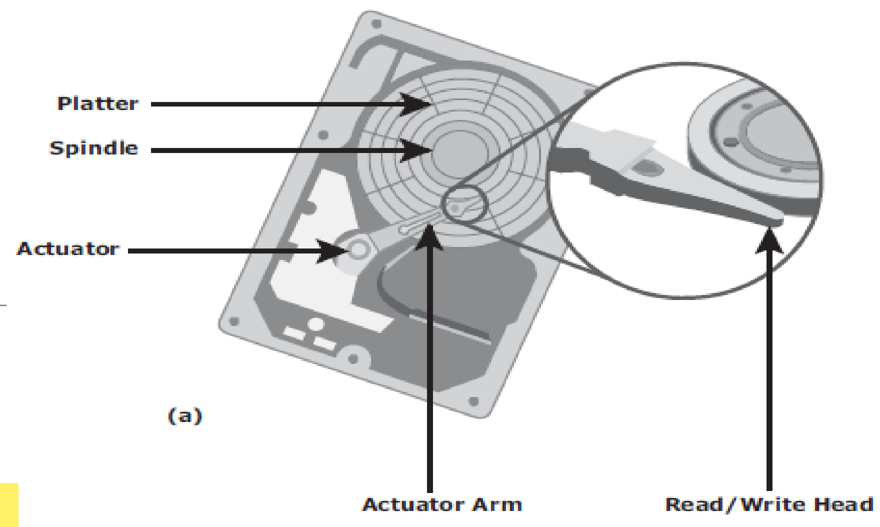
Read/Write (R/W) heads, read and write data from or to platters.

Drives have two R/W heads per platter, one for each surface of the platter.





Disk Drive Components



When writing data-

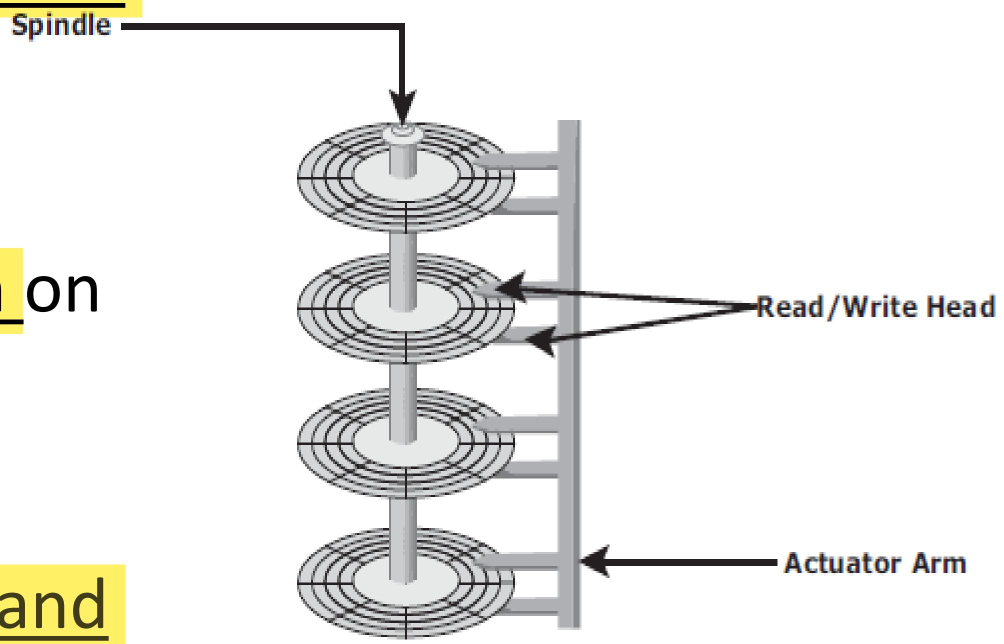
- The R/W head changes the magnetic polarization on the surface of the platter

While reading data -

- The R/W head detects the magnetic polarization on the surface of the platter.

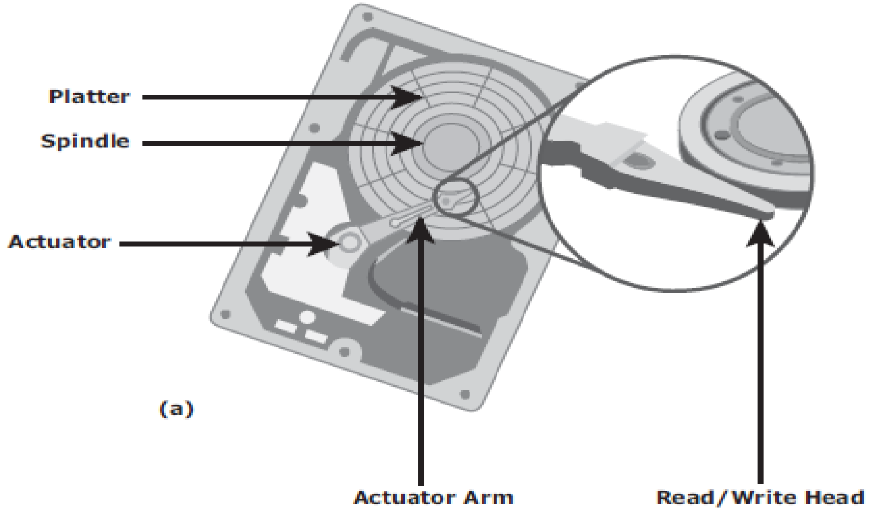
During reads and writes-

The R/W head senses the magnetic polarization and never touches the surface of the platter.



Ramaiah Institute of Technology

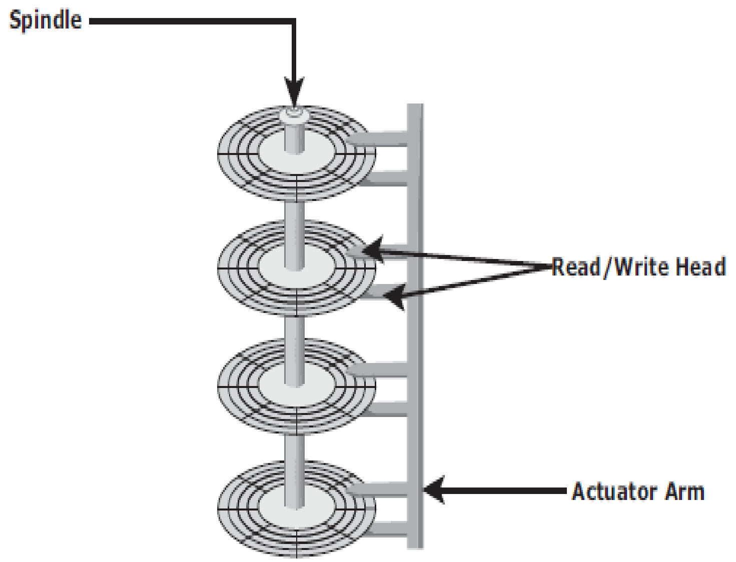
Disk Drive Components



Head flying height

When the spindle is rotating, there is a microscopic air gap maintained between the R/W heads and the platters.

This air gap is removed when the spindle stops rotating and the R/W head rests on a special area on the platter near the spindle. This area is called the landing



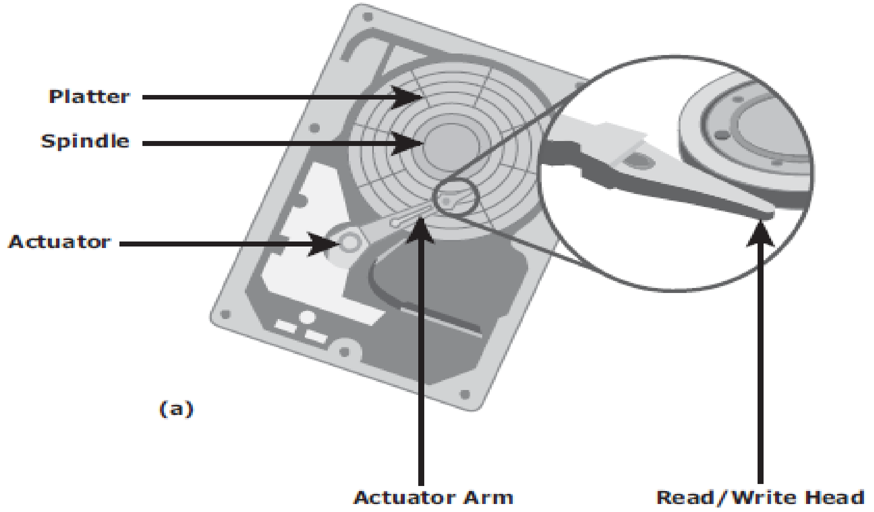
Disk Drive Components

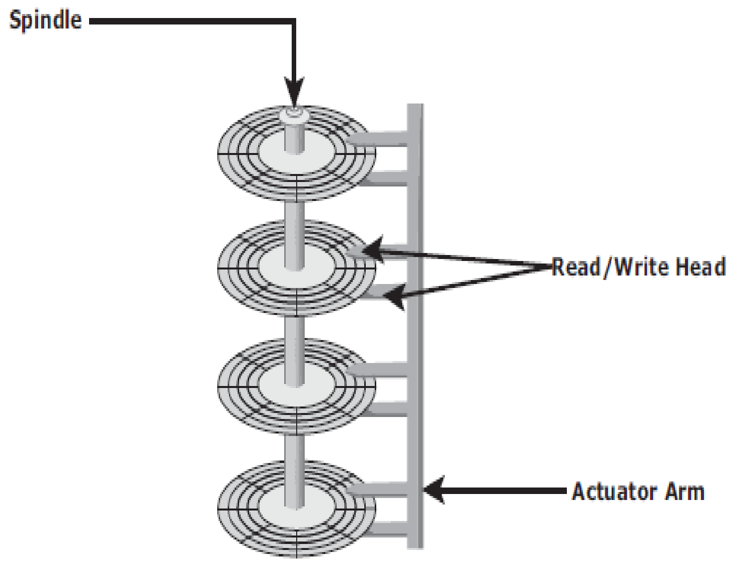
Head flying height

When the spindle is rotating, there is a microscopic air gap maintained between the R/W heads and the platters.

This air gap is removed when the spindle stops rotating and the R/W head rests on a special area on the platter near the spindle. This area is called the landing zone.

The landing zone is coated with a lubricant to reduce friction between the head and the platter.





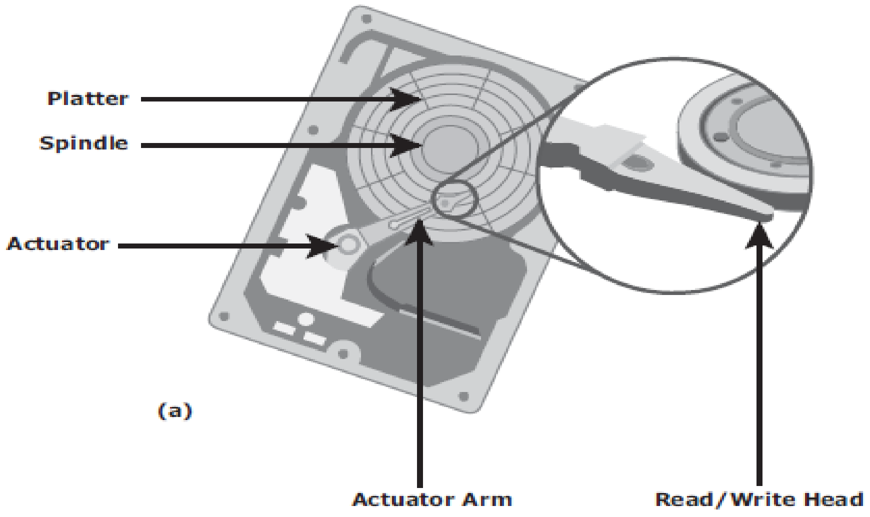
Disk Drive Components

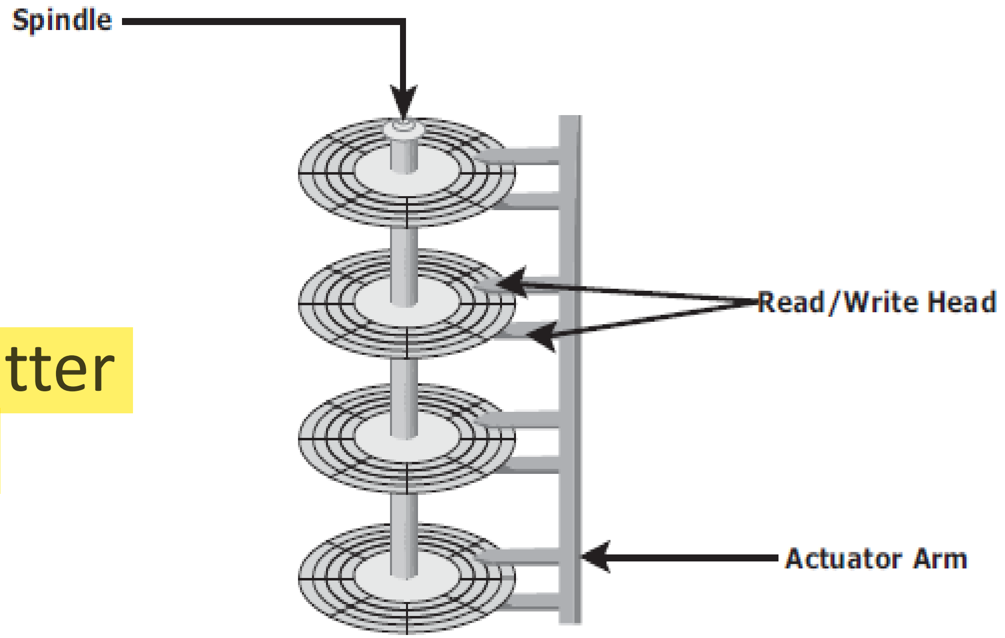
The logic on the disk drive ensures that heads are moved to the landing zone before they touch the surface.

If the drive malfunctions and the R/W head accidentally touches the surface of the platter outside the landing zone, a head crash occurs.

In a head crash, the magnetic coating on the platter is scratched and may cause damage to the R/W head.

A head crash generally results in data loss.



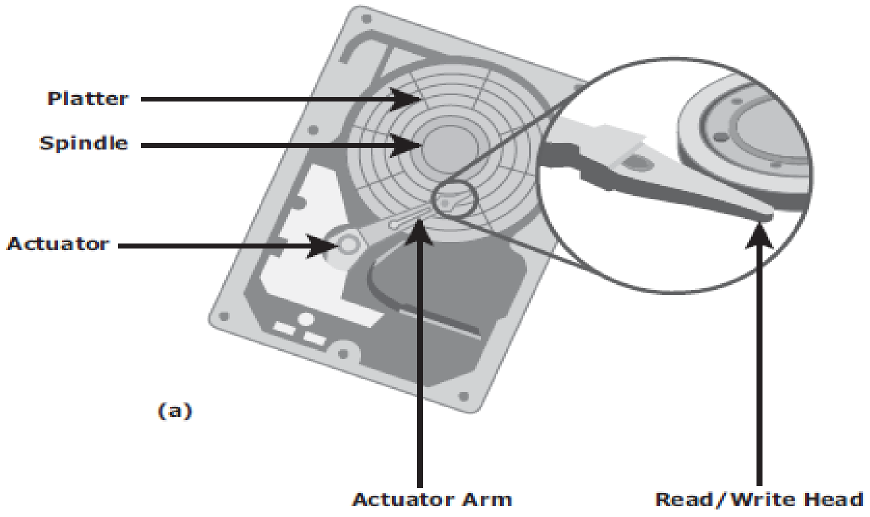


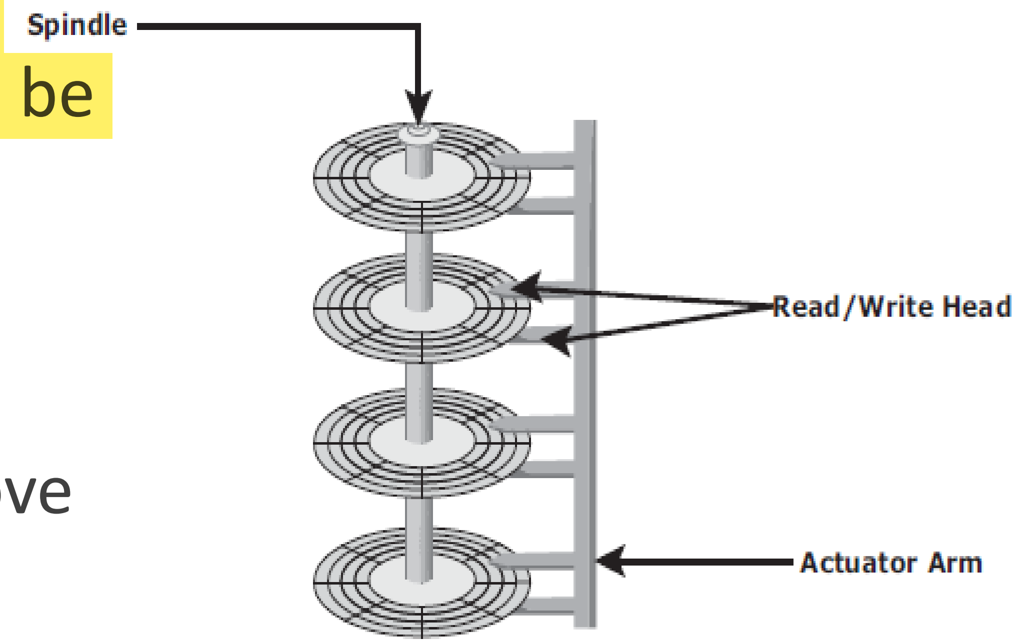
Disk Drive Components

4.Actuator Arm Assembly

- R/W heads are mounted on the actuator arm assembly, which positions the R/W head at the location on the platter where the data needs to be written or read.

- The R/W heads for all platters on a drive are attached to one actuator arm assembly and move across the platters simultaneously.





RAMAIAH
Institute of Technology

Disk Drive Components

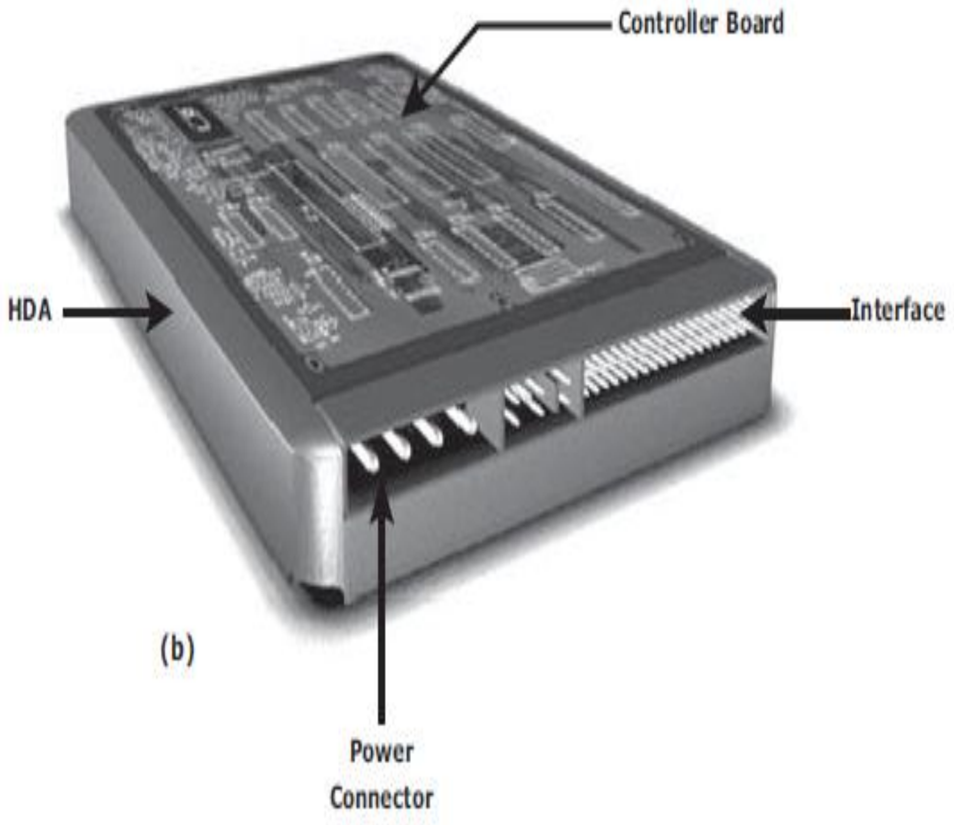
5. Drive Controller Board

The controller is a printed circuit board, mounted at the bottom of a disk drive.

It consists of a microprocessor, internal memory, circuitry, and firmware.

The firmware controls the power to the spindle motor and the speed of the motor.

It also manages the communication between the drive and the host.



Disk Drive Components - Physical Disk Structure

- Data on the disk is recorded on tracks, which are concentric rings on the platter around the spindle.
- The tracks are numbered, starting from zero, from the outer edge of the platter.
- The number of tracks per inch (TPI) on the platter (or the track density) measures how tightly the tracks are packed on a platter.

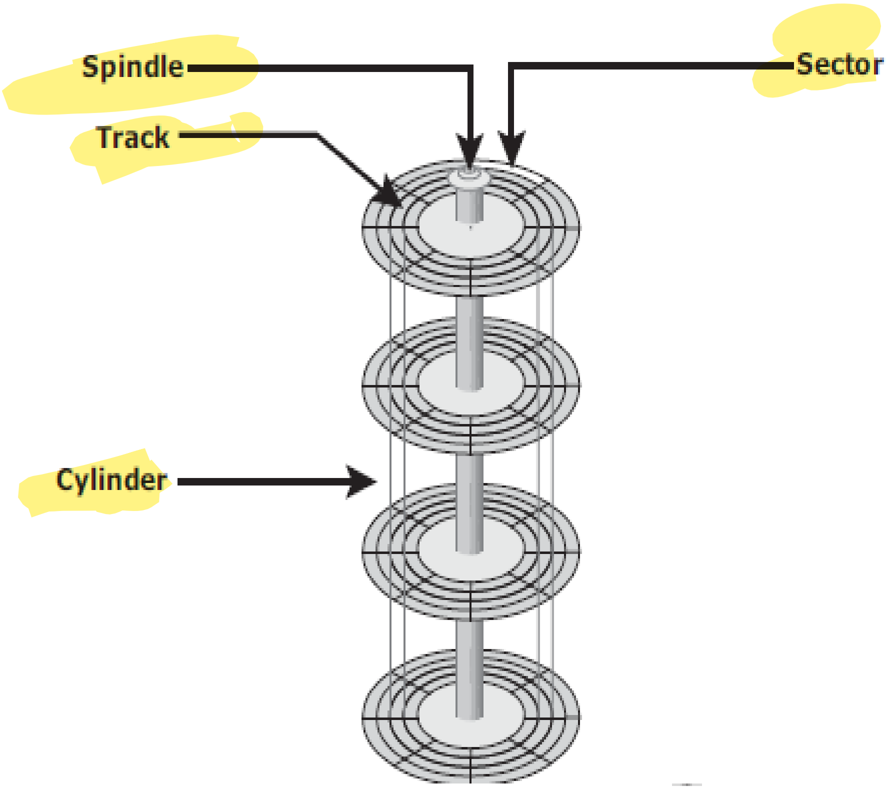


Figure 2-8: Disk structure: sectors, tracks, and cylinders

RAMAIAH
Institute of Technology

Disk Drive Components - Physical Disk Structure

- Each track is divided into smaller units called sectors.
- A sector is the smallest, individually addressable unit of storage.
- A sector holds 512 bytes of user data
- The number of sectors per track varies according to the drive type.

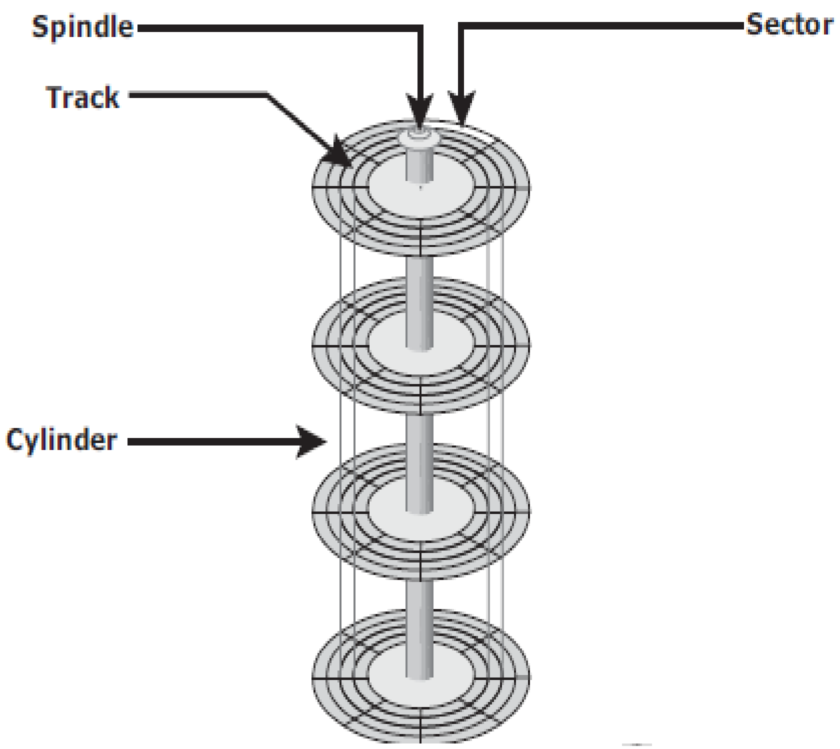


Figure 2-8: Disk structure: sectors, tracks, and cylinders

Disk Drive Components - Physical Disk Structure

A cylinder is a set of identical tracks on both surfaces of each drive platter.

The location of R/W heads is referred to by the cylinder number, not by the track number.

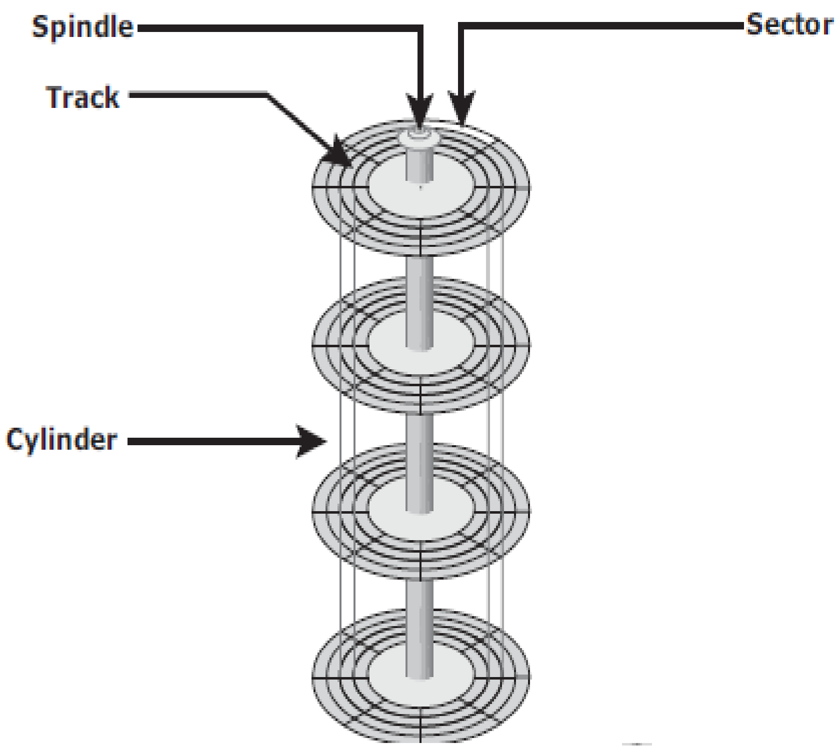


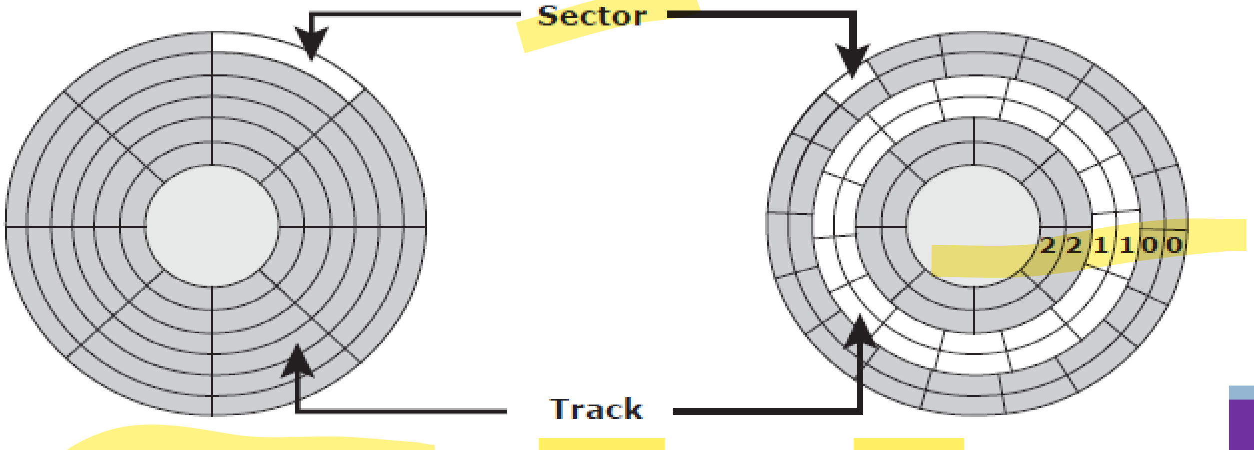
Figure 2-8: Disk structure: sectors, tracks, and cylinders

Explain the concepts of Zone Bit Recording (ZBR) and Native Command Queueing (NCQ) used in disk storage. Describe the advantages of adopting these techniques.

Disk Drive Components - Zoned Bit Recording

Platters are made of concentric tracks; the outer tracks can hold more data than the inner tracks because the outer tracks are physically longer than the inner tracks.

On older disk drives, the outer tracks had the same number of sectors as the inner tracks, so data density was low on the outer tracks. This was an inefficient use of the available space, as shown in Figure (a).



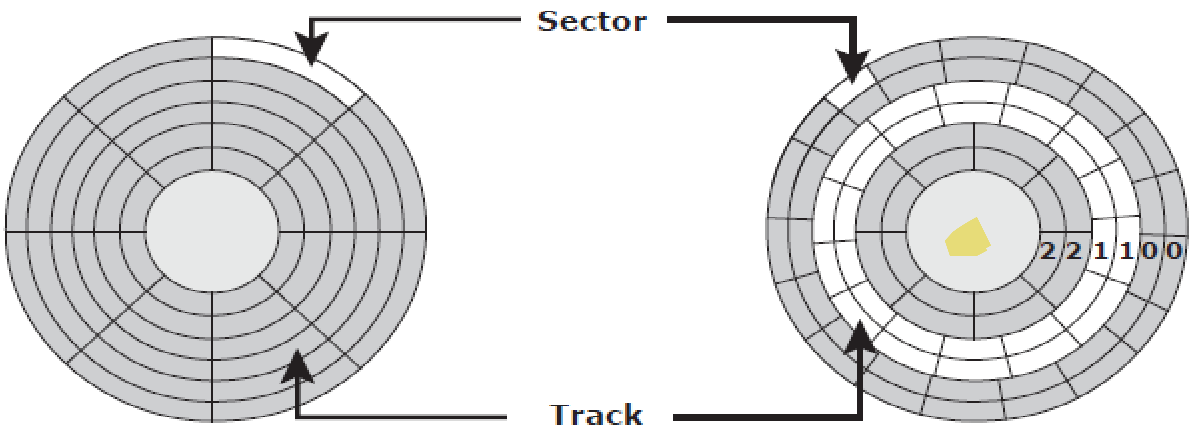
(a) Platter Without Zones (b) Platter With Zones

Disk Drive Components - Zoned Bit Recording

Zoned bit recording uses the disk efficiently.

As shown in Figure (b), this mechanism groups tracks into zones based on their distance from the center of the disk.

The zones are numbered, with the outermost zone being zone 0.



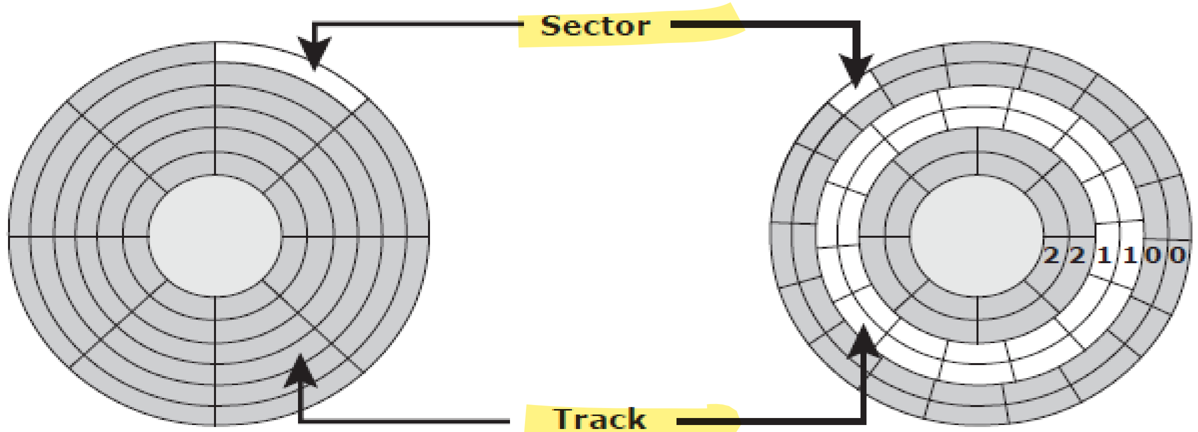
(a) Platter Without Zones

(b) Platter With Zones

Disk Drive Components - Zoned Bit Recording

An appropriate number of sectors per track are assigned to each zone, so a zone near the center of the platter has fewer sectors per track than a zone on the outer edge.

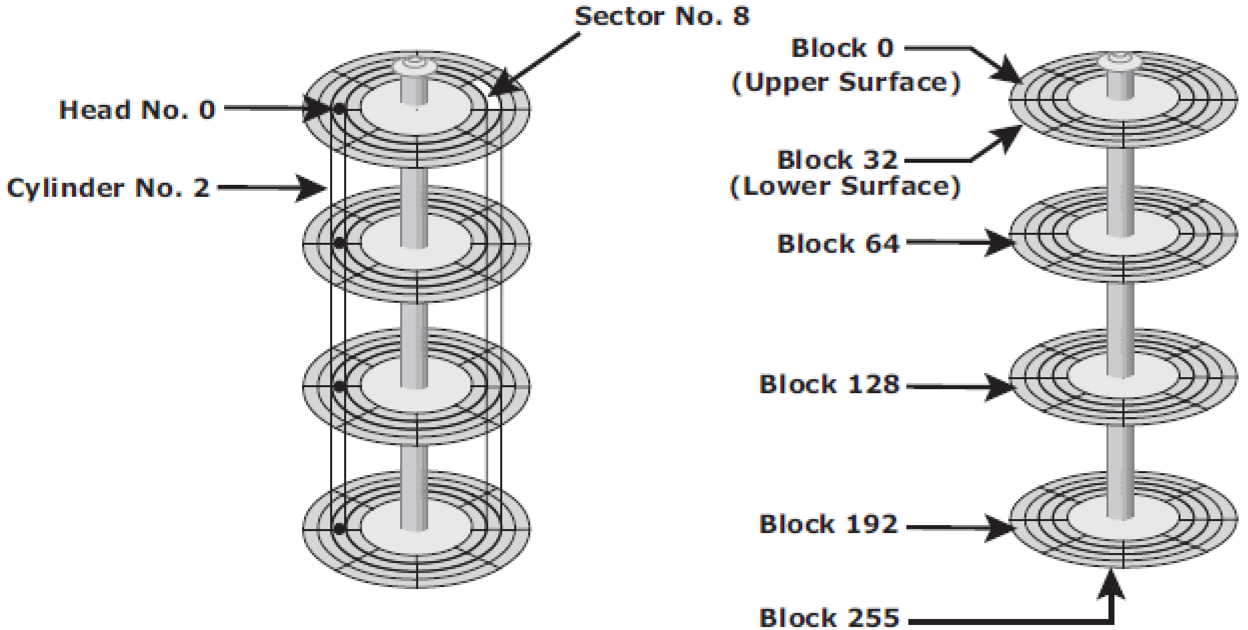
However, tracks within a particular zone have the same number of sectors.



(a) Platter Without Zones (b) Platter With Zones

Disk Drive Components - Logical Block Addressing

Drives used physical addresses consisting of the cylinder, head, and sector (CHS) number to refer to specific locations on the disk, as shown in Figure (a)



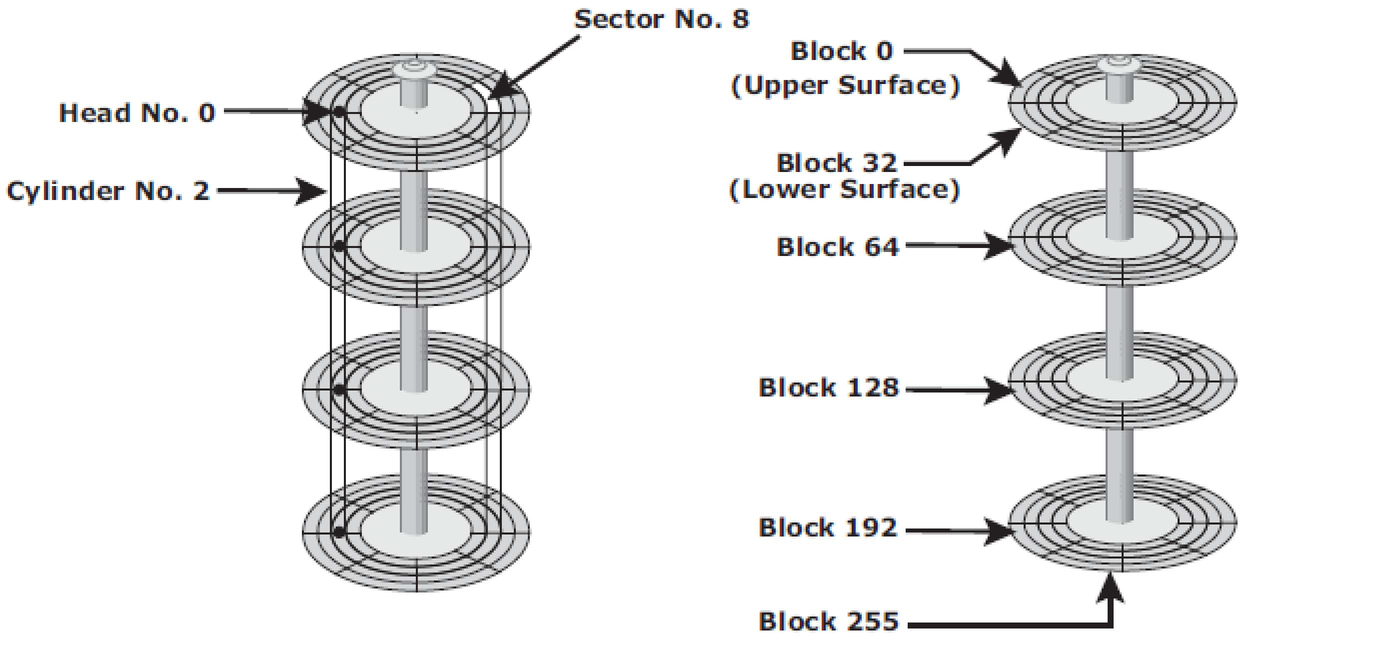
(a) Physical Address = CHS

(b) Logical Block Address = Block#

Disk Drive Components - Logical Block Addressing

Logical block addressing (LBA), as shown in Figure (b),

It simplifies addressing by using a linear address to access physical blocks of data.



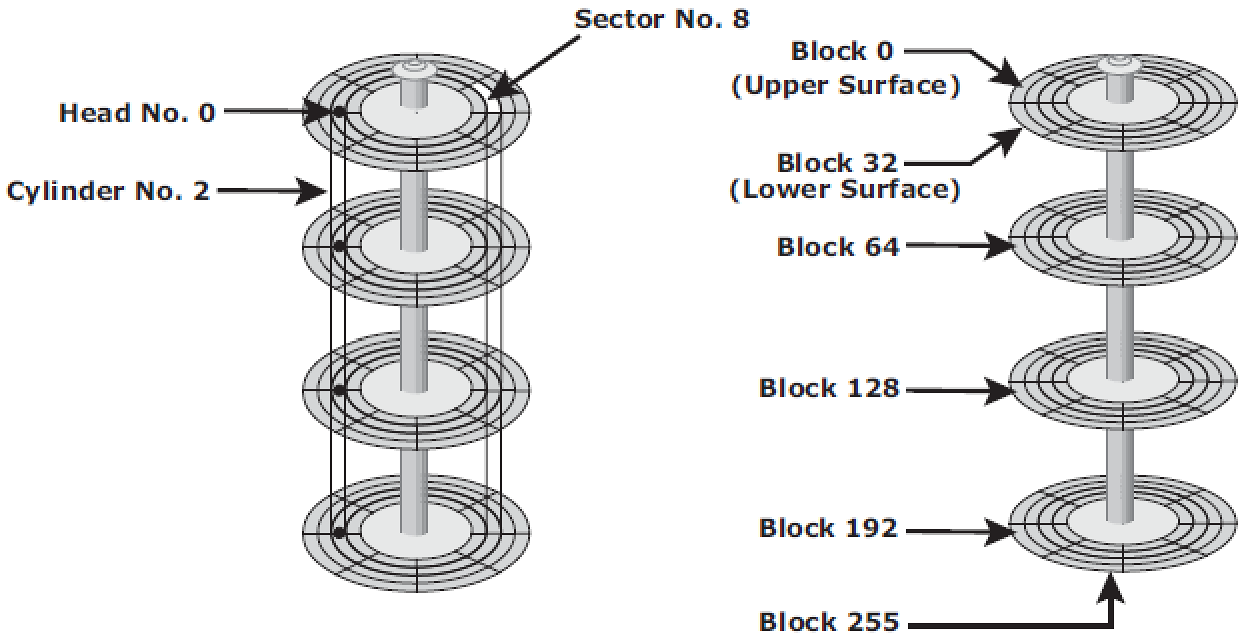
(a) Physical Address = CHS

(b) Logical Block Address = Block#

Disk Drive Components - Logical Block Addressing

The disk controller translates LBA to a CHS address, and the host needs to know only the size of the disk drive in terms of the number of blocks.

The logical blocks are mapped to physical sectors on a 1:1 basis.



(a) Physical Address = CHS

(b) Logical Block Address = Block#

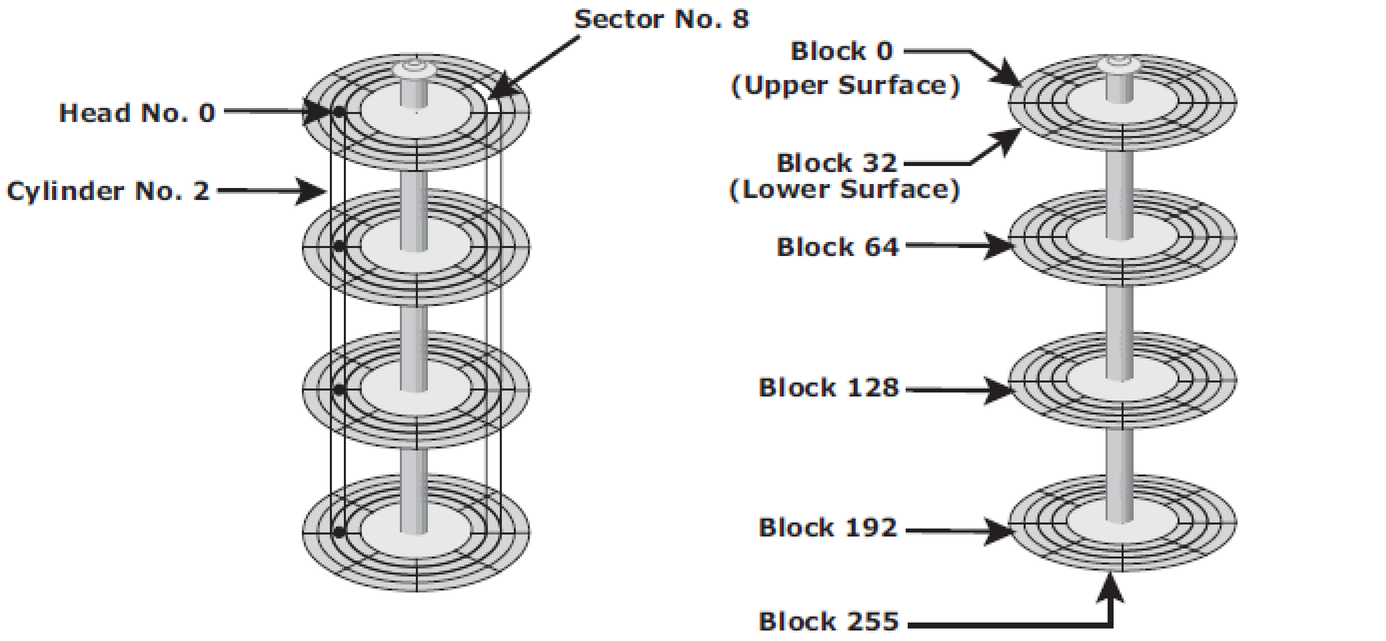
Disk Drive Components - Logical Block Addressing

In Figure (b), the drive shows eight sectors per track, eight heads, and four cylinders.

This means a total of 8 × 8 × 4 = 256 blocks, so the block number ranges from 0 to 255.

Each block has its own unique address.

Assuming that the sector holds 512 bytes, a 500 GB drive with a formatted capacity of 465.7 GB has in excess of 976,000,000 blocks.



(a) Physical Address = CHS (b) Logical Block Address = Block#

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

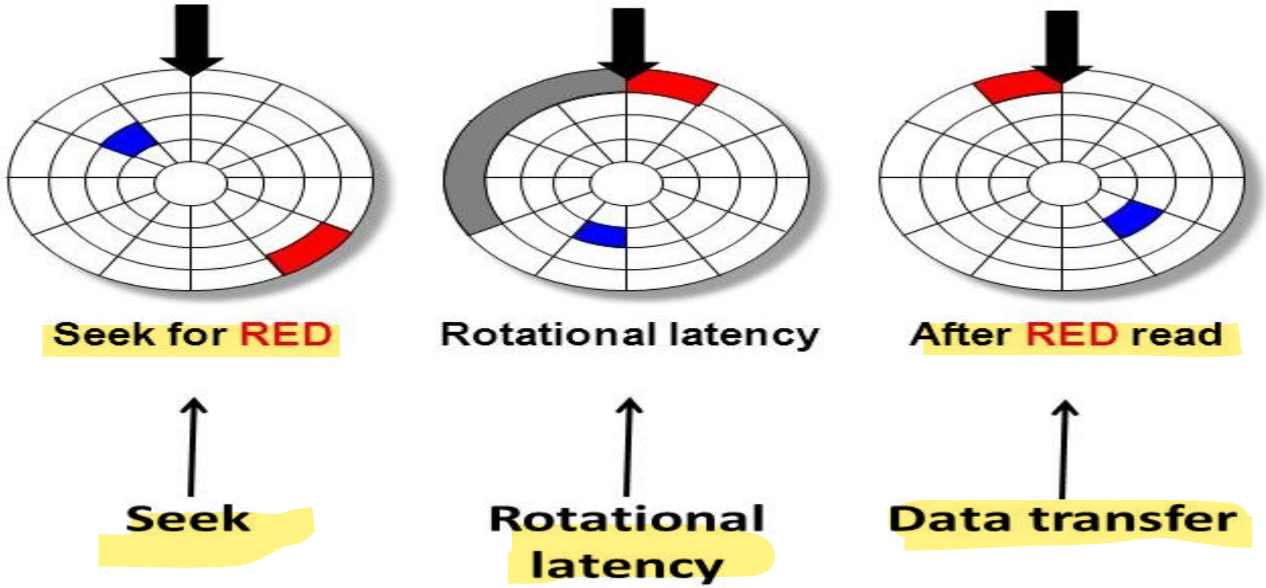
Examine the factors that affect the Hard disk performance? And also analyze the IOPS requirement of an application, how it affects the average response time?

Disk Drive Performance

- A disk drive is an electromechanical device that governs the overall performance of the storage system environment.
- The various factors that affect the performance of disk drives
- Disk Service Time
  - Seek Time
  - Rotational Latency
  - Data Transfer Rate
- Disk I/O Controller Utilization

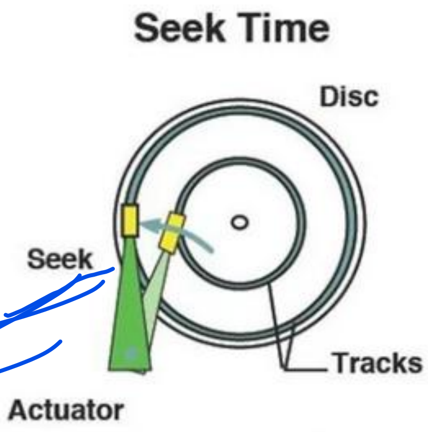
Disk Drive Performance -Disk Service Time

- Disk service time is the time taken by a disk to complete an I/O request.
- Components that contribute to the service time on a disk drive are
  - Seek time
  - Rotational latency
  - Data transfer rate



RAMAIAH
Institute of Technology

Disk Drive Performance -Disk Service Time



Seek time

The seek time (also called access time) describes the time taken to position the R/W heads across the platter with a radial movement (moving along the radius of the platter).

Disk vendors publish the following seek time specifications:

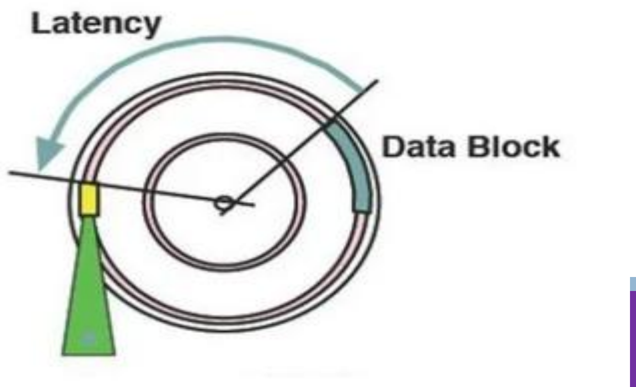
- **Full Stroke:** The time taken by the R/W head to move across the entire width of the disk, from the innermost track to the outermost track.
- **Average:** The average time taken by the R/W head to move from one random track to another, normally listed as the time for one-third of a full stroke.
- **Track-to-Track:** The time taken by the R/W head to move between adjacent tracks.

RAMAIAH Institute of Technology

Disk Drive Performance -Disk Service Time

Rotational Latency

- To access data, the actuator arm moves the R/W head over the platter to a particular track while the platter spins to position the requested sector under the R/W head.
- The time taken by the platter to rotate and position the data under the R/W head is called rotational latency.
- This latency depends on the rotation speed of the spindle and is measured in milliseconds.

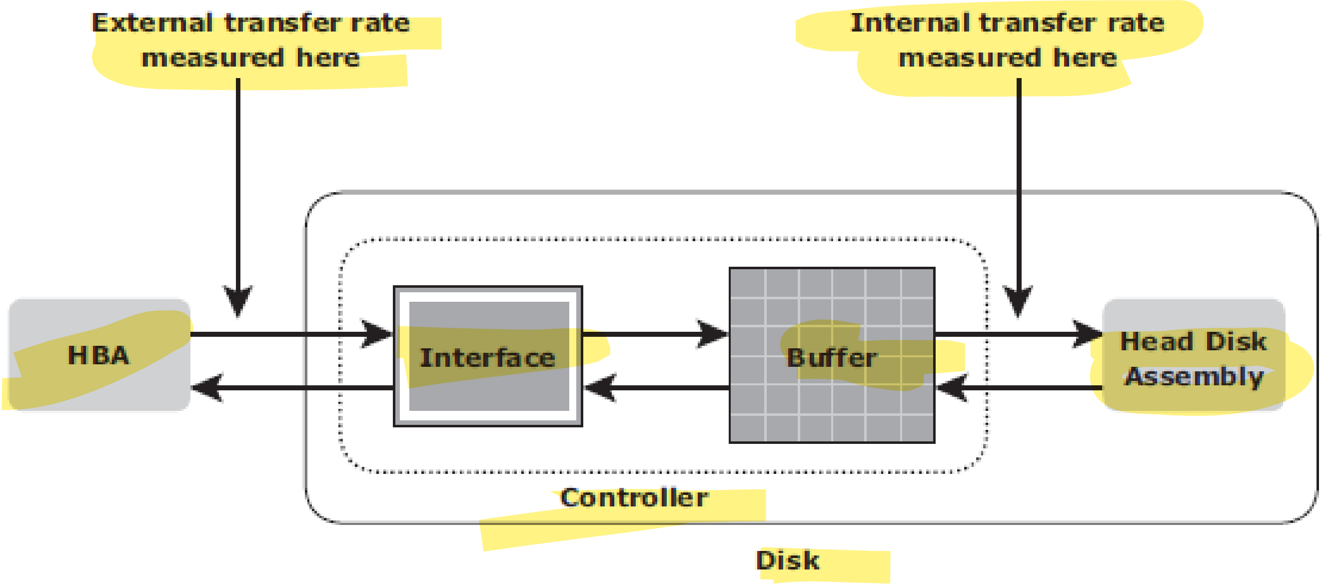


RAMAIAH
Institute of Technology

Disk Drive Performance -Disk Service Time

Data Transfer Rate

- The data transfer rate (also called transfer rate) refers to the average amount of data per unit time that the drive can deliver to the HBA.

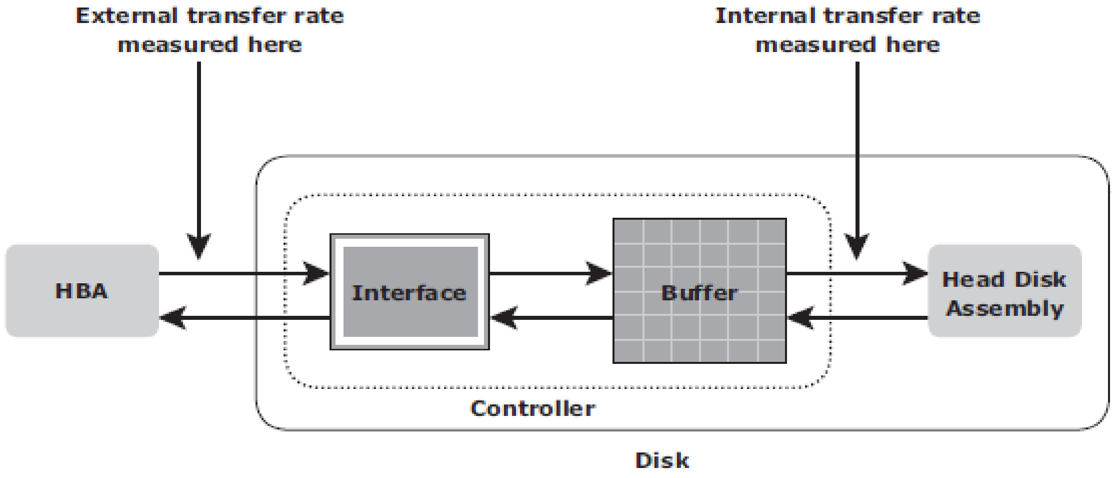


RAMAIAH
Institute of Technology

Disk Drive Performance -Disk Service Time

Data Transfer Rate

- In a read operation, the data first moves from disk platters to R/W heads; then it moves to the drive's internal buffer. Finally, data moves from the buffer through the interface to the host HBA.

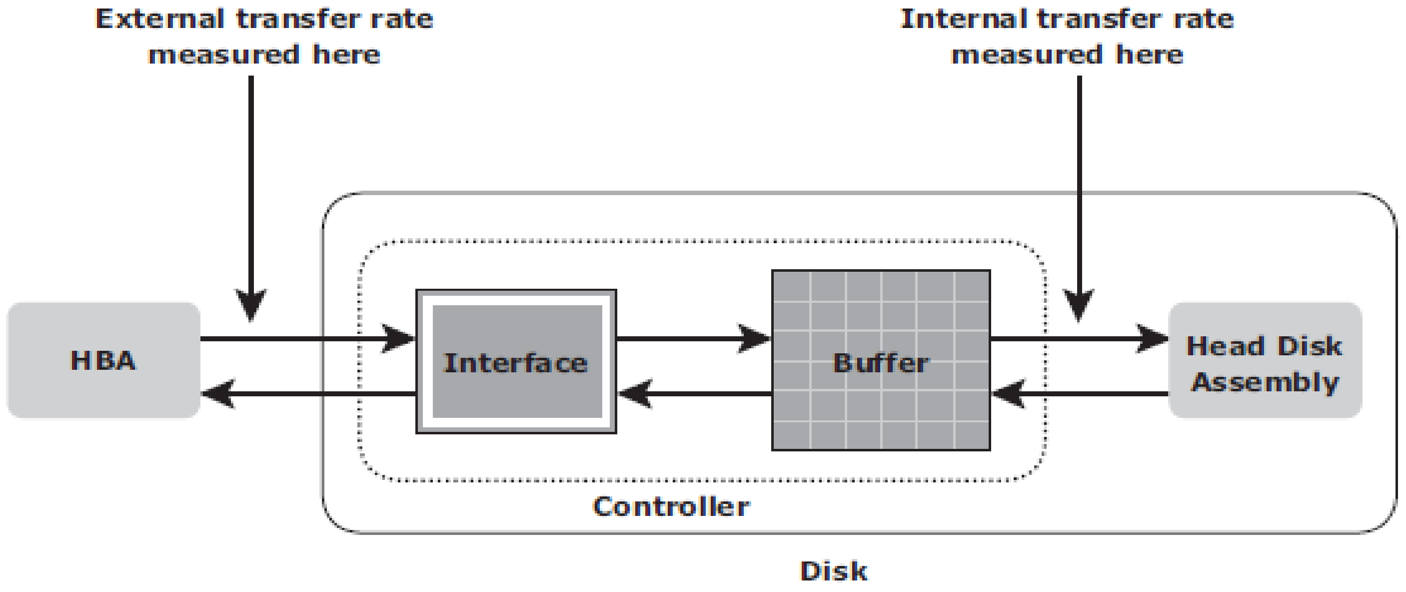


RAMAIAH Institute of Technology

Disk Drive Performance -Disk Service Time

Data Transfer Rate

- In a write operation, the data moves from the HBA to the internal buffer of the disk drive through the drive's interface. The data then moves from the buffer to the R/W heads. Finally, it moves from the R/W heads to the platters.

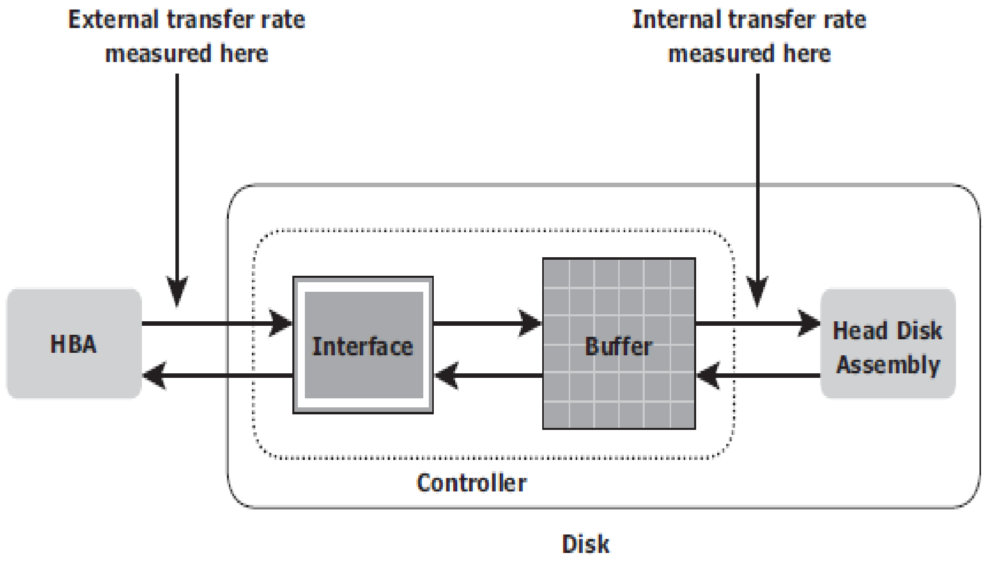


RAMAIAH
Institute of Technology

Disk Drive Performance -Disk Service Time

Data Transfer Rate

- The data transfer rates during the R/W operations are measured in terms of internal and external transfer rates
- Internal transfer rate is the speed at which data moves from a platter's surface to the internal buffer (cache) of the disk.
- The internal transfer rate takes into account fa
and rotational latency.

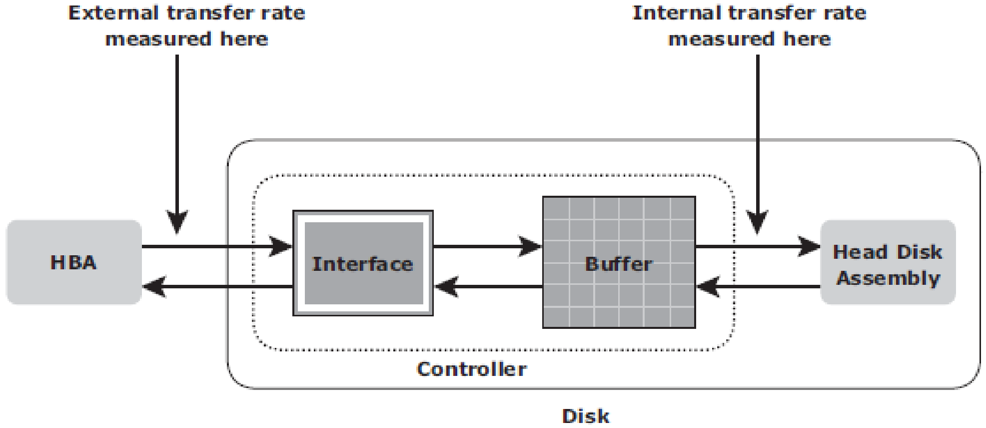


RAMAIAH
Institute of Technology

Disk Drive Performance -Disk Service Time

Data Transfer Rate

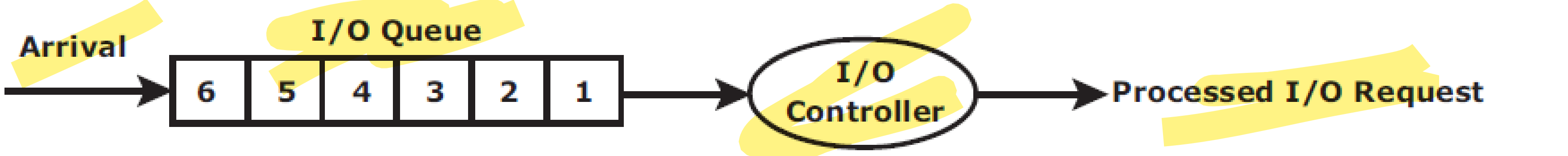
- The data transfer rates during the R/W operations are measured in terms of internal and external transfer rates
- External transfer rate is the rate at which data can move through the interface to the HBA
- The external transfer rate is generally the advertised speed of the interface, such as 133 MB/s for ATA.



Disk Drive Performance - Disk I/O Controller Utilization

- Utilization of a disk I/O controller has a significant impact on the I/O response time.
- Consider that a disk can be viewed as a black box consisting of two elements:
  - Queue: The location where an I/O request waits before it is processed by the I/O controller
  - Disk I/O Controller: Processes I/Os waiting in the queue one by one

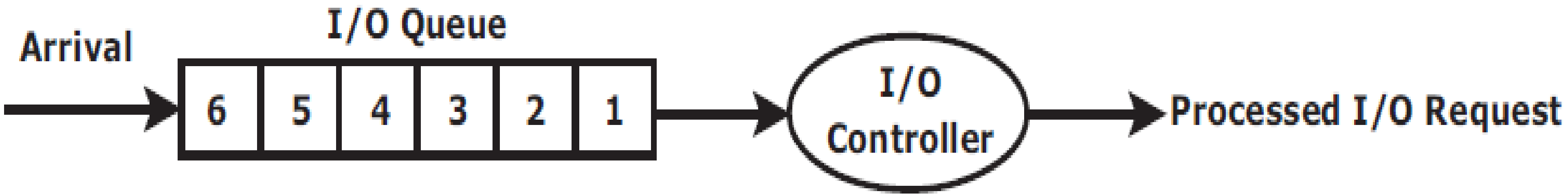
Requests are held in the I/O queue, and the I/O controller processes them one by one



RAMAIAH
Institute of Technology

Disk Drive Performance - Disk I/O Controller Utilization

- The I/O requests arrive at the controller at the rate generated by the application. This rate is also called the arrival rate.
- The I/O arrival rate, the queue length, and the time taken by the I/O controller to process each request determines the I/O response time. If the controller is busy or heavily utilized, the queue size will be large and the response time will be high.

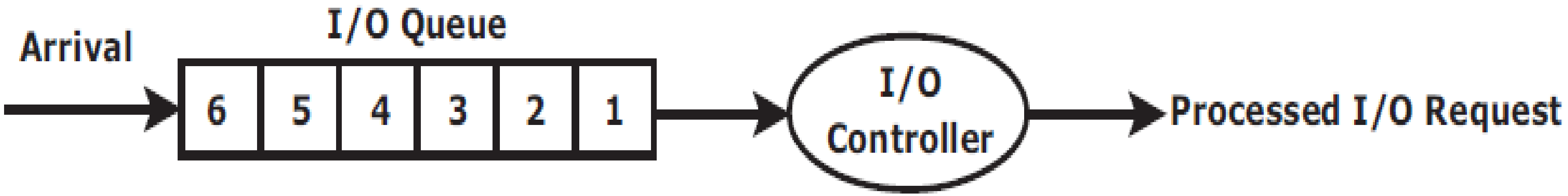


Disk Drive Performance - Disk I/O Controller Utilization

- Based on the fundamental laws of disk drive performance, the relationship between controller utilization and average response time is given as

Average response time (T_R) = Service time (T_S) / (1 – Utilization)

where T_S is the time taken by the controller to serve an I/O.



RAMAIAH
Institute of Technology

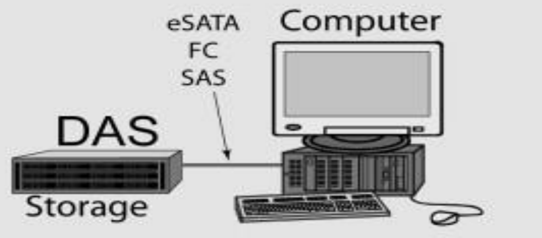
UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Host Access to Data

- The key components of infrastructure are the operating system (or file system), connectivity, and storage.
- The storage device can be internal or external to the host.
- host controller card accesses the storage devices using
  - IDE/ATA and SCSI for accessing data from internal storage.
  - FC and iSCSI protocols for accessing data from external storage
- External storage devices can be connected to the host directly or through the storage network.
- When the storage is connected directly to the host, it is referred as **direct-attached storage (DAS)**



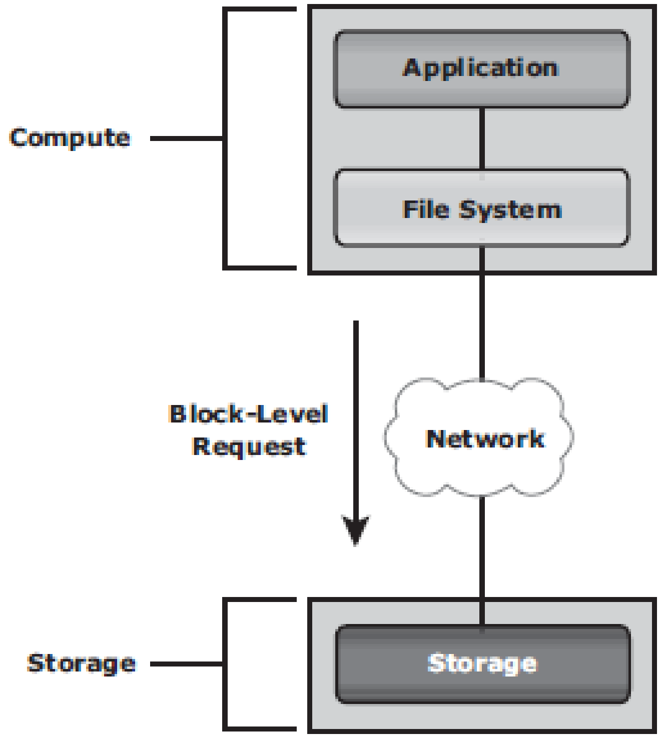
Direct Attached Storage (DAS)

Host Access to Data

- Data can be accessed over a network in one of the following ways:
    - block level
    - file level or object level.
- The application requests data from the file system (or operating system) by specifying the filename and location.
- The file system maps the file attributes to the logical block address of the data and sends the request to the storage device.
- The storage device converts the logical block address (LBA) to a cylinder-head-sector (CHS) address and fetches the data.

Host Access to Data

- In a block-level access, the file system is created on a host, and data is accessed on a network at the block level, as shown in Figure.
- In this case, raw disks or logical volumes are assigned to the host for creating the file system.

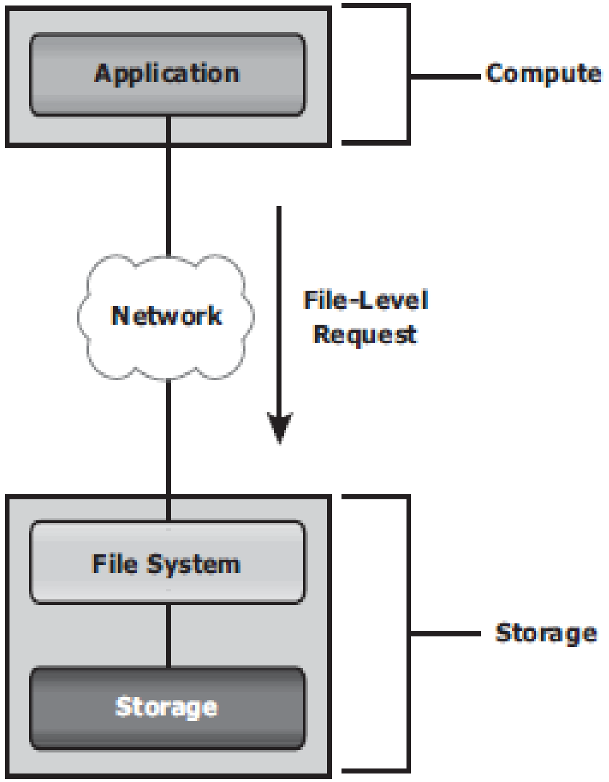


(a) Block-Level Access

RAMAIAH Institute of Technology

Host Access to Data

- In a file-level access, the file system is created on a separate file server or at the storage side, and the file-level request is sent over a network, as shown in Figure
- Object-level access is an intelligent evolution, whereby data is accessed over a network in terms of self-contained objects with a unique object identifier.



(b) File-Level Access

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

RAMAIAH Institute of Technology

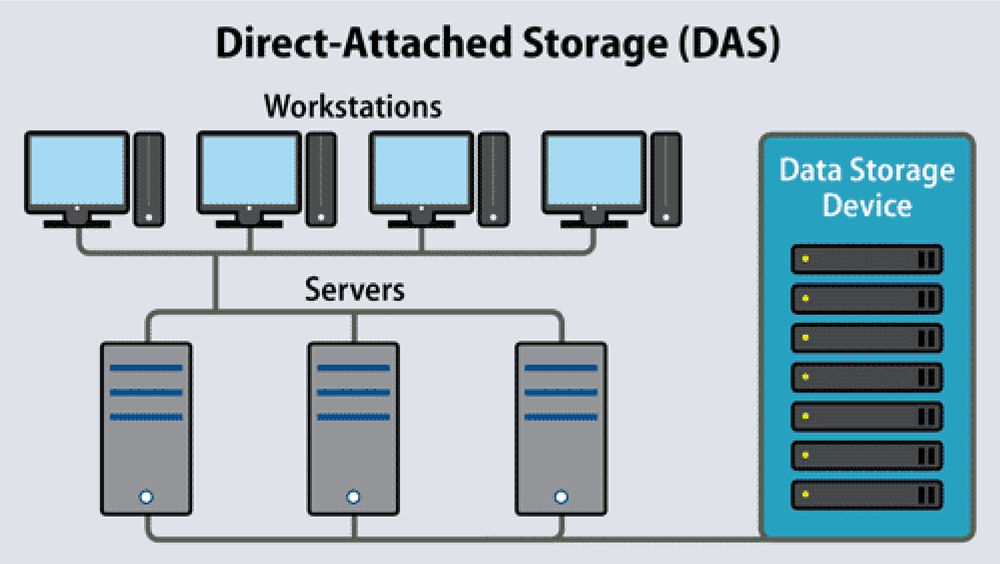
Explain how a host accesses storage using Direct-Attached Storage (DAS).
Discuss its benefits and limitations.

Direct-Attached Storage

- DAS is an architecture in which storage is connected directly to the hosts.
- DAS has remained suitable for localized data access in a small environment, such as personal computing and workgroups.

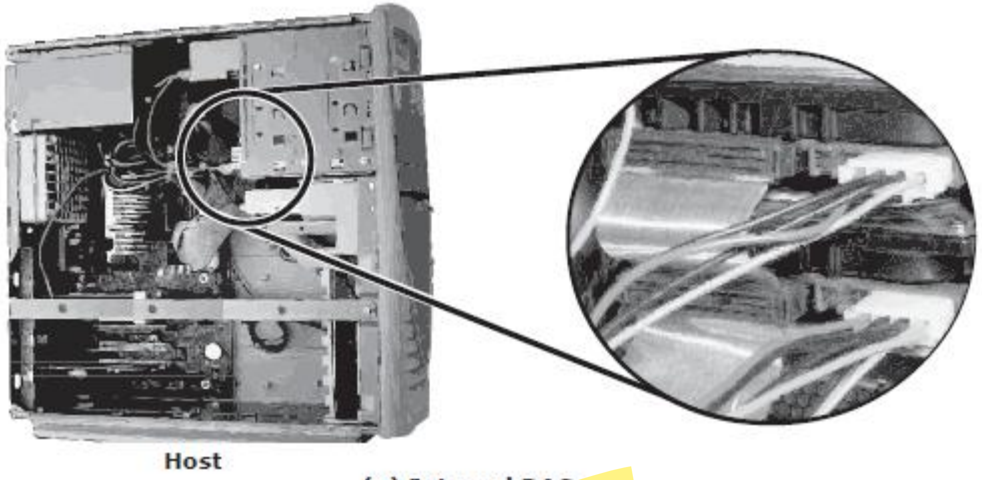
Examples:

- The internal disk drive of a host
- Directly-connected external storage array

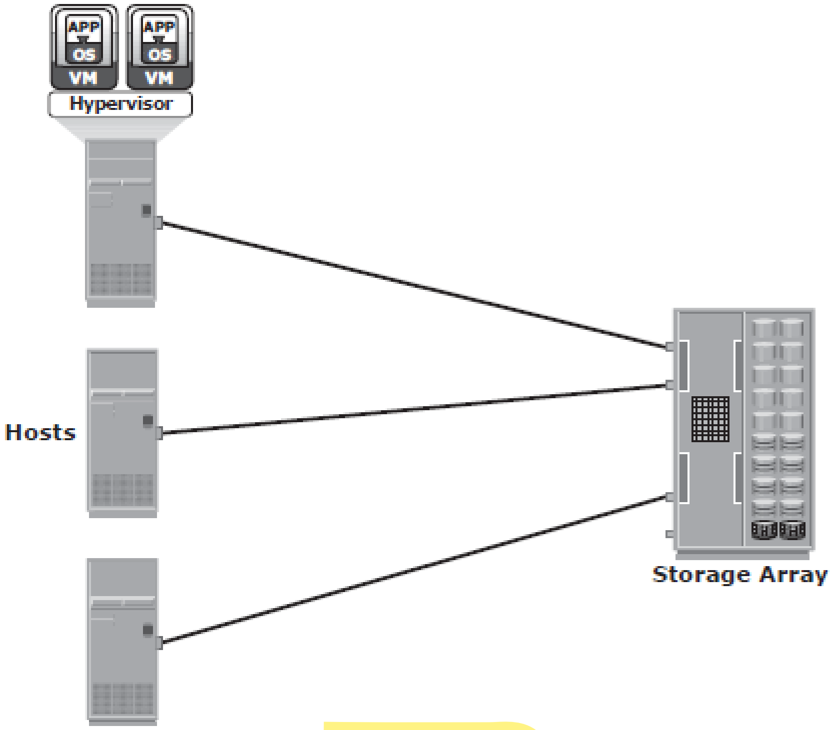


Direct-Attached Storage

- DAS is classified as internal or external, based on the location of the storage device with respect to the host.



(a) Internal DAS

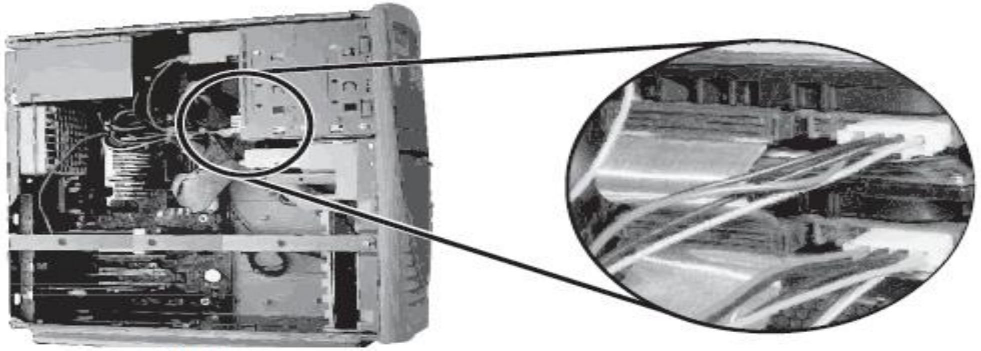


(b) External DAS

Direct-Attached Storage

Internal DAS architectures

- The storage device is internally connected to the host by a serial or parallel bus
- The physical bus has distance limitations and can be sustained only over a shorter distance for highspeed connectivity.
- Most internal buses can support only a limited number of devices, and they occupy a large amount of space inside the host, making maintenance of other components difficult.



Host

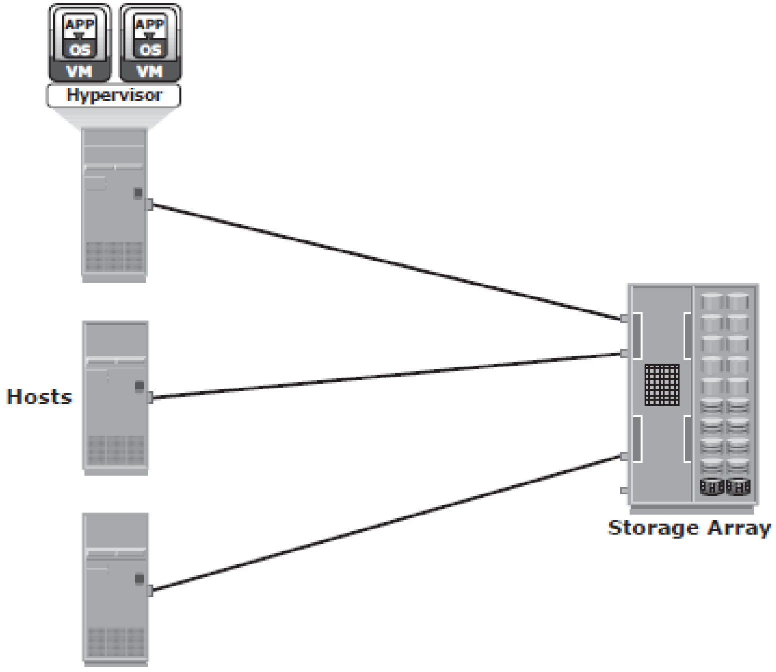
(a) Internal DAS

RAMAIAH
Institute of Technology

Direct-Attached Storage

External DAS architectures

- The host connects directly to the external storage device, and data is accessed at the block level .
- Communication between the host and the storage device takes place over a SCSI or FC protocol.
- External DAS overcomes the distance and device count limitations and provides centralized management of storage devices.



(b) External DAS

List any 4 DAS Benefits and Limitations

RAMAIAH
Institute of Technology

Direct-Attached Storage

DAS Benefits

- DAS requires a relatively lower initial investment than storage networking architectures.
- The DAS configuration is simple and can be deployed easily and rapidly.
- The setup is managed using host-based tools, such as the host OS, which makes storage management tasks easy for small environments.
- Because DAS has a simple architecture, it requires fewer management tasks and less hardware and software elements to set up and operate.

RAMAIAH
Institute of Technology

Direct-Attached Storage

DAS Limitations

- DAS does not scale well. A storage array has a limited number of ports, which restricts the number of hosts that can directly connect to the storage.
- When capacities are reached, the service availability may be compromised. DAS does not make optimal use of resources due to its limited capability to share front-end ports.
- In DAS environments, unused resources cannot be easily reallocated,
- resulting in islands of over-utilized and under-utilized storage pools.

RAMAIAH
Institute of Technology

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Storage Design Based on Application

- Determining storage requirements for an application begins with determining the required storage capacity.
- This is easily estimated by the size and number of file systems and database components used by applications.
- The I/O size, I/O characteristics, and the number of I/Os generated by the application at peak workload are other factors that affect disk performance, I/O response time, and design of storage systems.
- The I/O block size depends on the file system and the database on which the application is built.

Storage Design Based on Application

- Block size in a database environment is controlled by the underlying database engine and the environment variables.
- The disk service time (T_S) for an I/O is a key measure of disk performance; T_S, along with disk utilization rate (U), determines the I/O response time for an application.
- The total disk service time (T_S) is the sum of the seek time (T), rotational latency (L), and internal transfer time (X)

$$T_S = T + L + X$$

RAMAIAH
Institute of Technology

Storage Design Based on Application

Example with the following specifications provided for a disk:

- The average seek time is 5 ms in a random I/O environment; T = 5 ms.
- Disk rotation speed of 15,000 revolutions per minute or 250 revolutions per second

  Rotational latency (L) - one-half of the time taken for a full rotation

  L = 0.5/250 rps expressed in ms.

- 40 MB/s internal data transfer rate, from which the internal transfer time (X) is derived based on the block size of the I/O

  Example : I/O with a block size of 32 KB;

  X = 32 KB/40 MB.

Time taken by the I/O controller to serve an I/O of block size 32 KB is

$T_S = T + L + X$

$T_S = 5\text{ ms} + (0.5/250) + 32\text{ KB}/40\text{ MB}$

$T_S = 7.8\text{ ms}.$

Storage Design Based on Application

Example with the following specifications provided for a disk:

Maximum number of I/Os serviced per second(IOPS) is $(1/T_S) = 1/(7.8 \times 10^{-3}) = 128$ IOPS.

Table lists the maximum IOPS that can be serviced for different block sizes using the previous disk specifications.

Table 2-1: IOPS Performed by Disk Drive

BLOCK SIZE	$T_S = T + L + X$	IOPS = $1/T_S$
4 KB	5 ms + (0.5/250 rps) + 4 K/40 MB = 5 + 2 + 0.1 = 7.1	140
8 KB	5 ms + (0.5/250 rps) + 8 K/40 MB = 5 + 2 + 0.2 = 7.2	139
16 KB	5 ms + (0.5/250 rps) + 16 K/40 MB = 5 + 2 + 0.4 = 7.4	135
32 KB	5 ms + (0.5/250 rps) + 32 K/40 MB = 5 + 2 + 0.8 = 7.8	128
64 KB	5 ms + (0.5/250 rps) + 64 K/40 MB = 5 + 2 + 1.6 = 8.6	116

The IOPS ranging from 116 to 140 for different block sizes represents the IOPS that can be achieved at potentially high levels of utilization

RAMAIAH Institute of Technology

Storage Design Based on Application

Example with the following specifications provided for a disk:

The application response time - R, increases with an increase in disk controller utilization.

R for an I/O with a block size of 32 KB at 96 percent disk controller utilization is

response time (T_R) = Service time (T_S) / (1 – Utilization)

$R = T_S/(1 – U)$

$= 7.8/(1 – 0.96)$

$= 195$ ms

Table 2-1: IOPS Performed by Disk Drive

BLOCK SIZE	$T_S = T + L + X$	$IOPS = 1/T_S$
4 KB	5 ms + (0.5/250 rps) + 4 K/40 MB = 5 + 2 + 0.1 = 7.1	140
8 KB	5 ms + (0.5/250 rps) + 8 K/40 MB = 5 + 2 + 0.2 = 7.2	139
16 KB	5 ms + (0.5/250 rps) + 16 K/40 MB = 5 + 2 + 0.4 = 7.4	135
32 KB	5 ms + (0.5/250 rps) + 32 K/40 MB = 5 + 2 + 0.8 = 7.8	128
64 KB	5 ms + (0.5/250 rps) + 64 K/40 MB = 5 + 2 + 1.6 = 8.6	116

Storage Design Based on Application

The total number of disks required (D_R) for an application is computed as follows:

$$D_R = \text{Max}\ (D_C, D_I)$$

D_C is the number of disks required to meet the capacity,

D_I is the number of disks required to meet the application IOPS requirement.

Storage Design Based on Application

Consider an example in which the capacity requirement for an application is 1.46 TB. The number of IOPS generated by the application at peak workload is estimated at 9,000 IOPS. The vendor specifies that a 146-GB, 15,000-rpm drive is capable of doing a maximum 180 IOPS.

$D_C = 1.46$ TB/146 GB

$D_C = 10$ disks.

Storage Design Based on Application

Consider an example in which the capacity requirement for an application is 1.46 TB. The number of IOPS generated by the application at peak workload is estimated at 9,000 IOPS. The vendor specifies that a 146-GB, 15,000-rpm drive is capable of doing a maximum 180 IOPS.

To meet the application IOPS requirements, the number of disks required is

$D_I = 9{,}000/180$

$D_I = 50.$

Storage Design Based on Application

Consider an example in which the capacity requirement for an application is 1.46 TB. The number of IOPS generated by the application at peak workload is estimated at 9,000 IOPS. The vendor specifies that a 146-GB, 15,000-rpm drive is capable of doing a maximum 180 IOPS.

However, if the application is response-time sensitive, the number of IOPS a disk drive can perform should be calculated based on 70-percent disk utilization.

The number of IOPS a disk can perform at 70 percent utilization is

= 180 * 0.7

= 126 IOPS.

Storage Design Based on Application

Consider an example in which the capacity requirement for an application is 1.46 TB. The number of IOPS generated by the application at peak workload is estimated at 9,000 IOPS. The vendor specifies that a 146-GB, 15,000-rpm drive is capable of doing a maximum 180 IOPS.

The number of IOPS a disk can perform at 70 percent utilization is

= 126 IOPS.

The number of disks required to meet the application IOPS requirement will be D_I = 9,000/126

D_I = 72.

RAMAIAH Institute of Technology

Storage Design Based on Application

Consider an example in which the capacity requirement for an application is 1.46 TB. The number of IOPS generated by the application at peak workload is estimated at 9,000 IOPS. The vendor specifies that a 146-GB, 15,000-rpm drive is capable of doing a maximum 180 IOPS.

As a result, the number of disks required to meet the application requirements will be $D_R = \text{Max}(D_C, D_I)$

$= \text{Max}(10, 72)$

$D_R = 72$ disks.

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

RAMAIAH
Institute of Technology

Explain the concepts of Zone Bit Recording (ZBR) and Native Command Queueing (NCQ) used in disk storage. Describe the advantages of adopting these techniques.

Disk **Native Command Queuing** (NCQ)

Command queuing

- Is a technique implemented on modern disk drives that determines the execution order of received I/Os and reduces unnecessary drive-head movements to improve disk performance.
- When an I/O is received for execution at the disk controller, the command queuing algorithms assign a tag that defines a sequence in which the commands should be executed.
- With command queuing, commands are executed based on the organization of data on the disk, regardless of the order in which the commands are received.

Disk **Native Command Queuing** (NCQ)

Command queuing

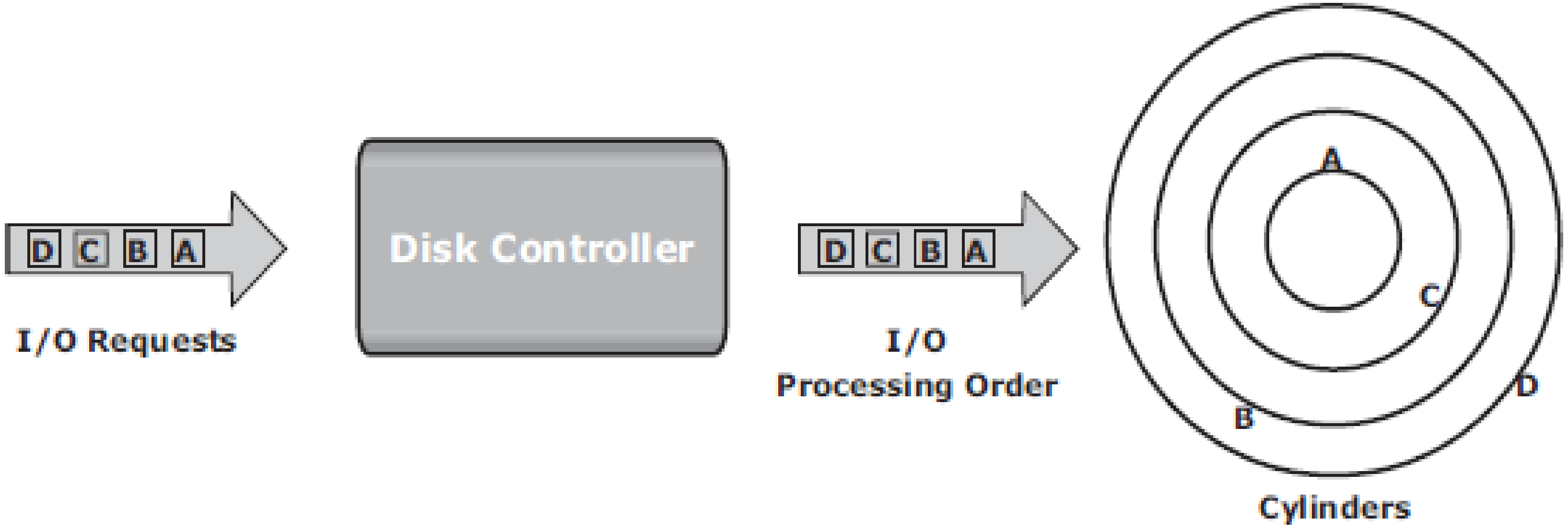
1. Seek time optimization
2. Access Time Optimization

RAMAIAH
Institute of Technology

Disk **Native Command Queuing** (NCQ)

Command queuing - Seek time optimization.

- Commands are executed based on optimizing read/write head movements, which might result in the reordering of commands.
- Without seek time optimization, the commands are executed in the order they are received.



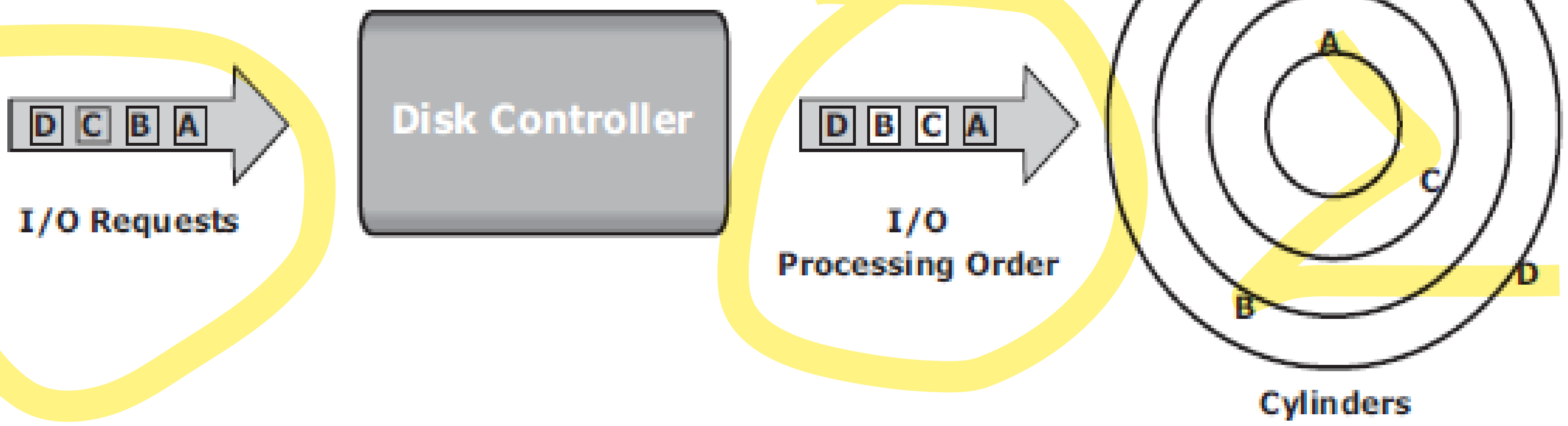
(a) Without Optimization

Ramaiah Institute of Technology

Disk **Native Command Queuing** (NCQ)

Command queuing - Seek time optimization.

- The radial movement required by the head to execute C immediately after A is less than what would be required to execute B.
- With seek time optimization, the command execution sequence would be A, C, B, and D,



(b) With Seek Time Optimization

Disk **Native Command Queuing** (NCQ)

Command queuing - Access Time Optimization

- With this algorithm, commands are executed based on the combination of seek time optimization and an analysis of rotational latency for optimal performance.
- Command queuing is also implemented on modern storage array controllers, which might further supplement the command queuing implemented on the disk drive.

UNIT 1- Storage System

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.

Introduction to Flash Drives

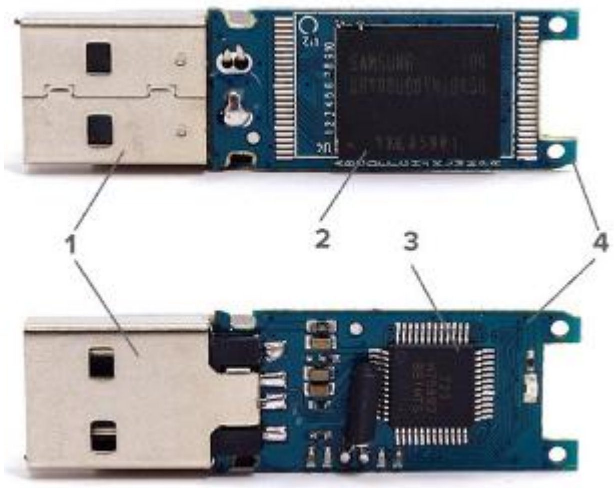
Flash drives(Solid state drives (SSDs))

- Flash drives use semiconductor-based solid state memory (flash memory) to store and retrieve data.
- Unlike conventional mechanical disk drives, flash drives contain no moving parts; therefore, they do not have seek and rotational latencies.
- Flash drives deliver a high number of IOPS with very low response times.
- Flash drives consume less power, compared to mechanical drives.
- Flash drives are especially suited for applications with small block size and random-read workloads that require consistently low (less than 1 millisecond) response times.

Introduction to Flash Drives

Components of Flash Drives

1. Controller-The controller manages the functioning of the drive
2. I/O interface -The I/O interface provides power and data access.
3. Mass storage (collection of memory chips) -Mass storage is an array of nonvolatile NAND (negated AND) memory chips used for storing data.
4. Cache -Cache serves as a temporary space or buffer for data transaction and operations.



1. USB-A Connector
2. NAND Memory Chip
3. Controller Chip
4. PCB (Printed Circuit Board)

Introduction to Flash Drives

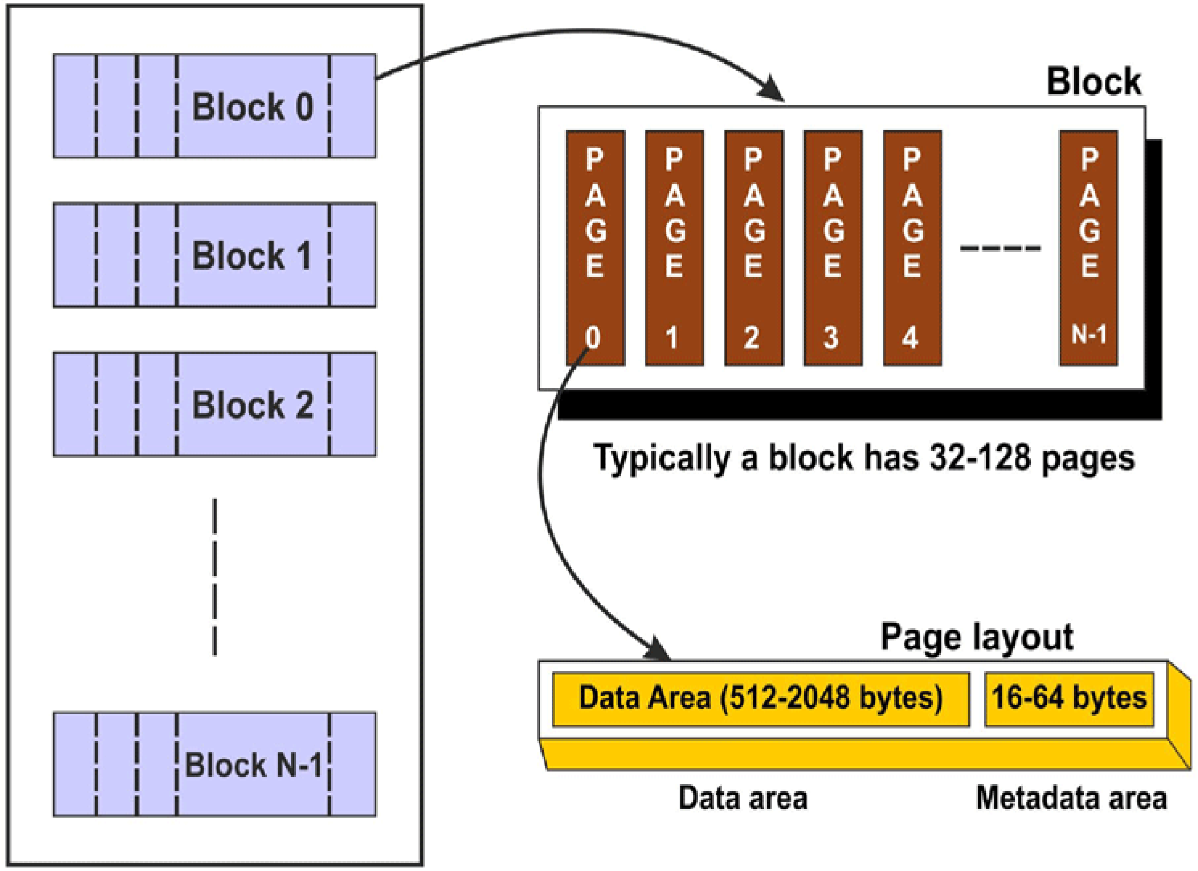
Architecture of Flash Drives

- A flash drive uses multiple parallel I/O channels (from its drive controller to the flash memory chips) for data access.
- Generally, the larger the number of flash memory chips and channels, the higher the drive’s internal bandwidth,and ultimately the higher the drive’s performance.
- Flash drives typically have eight to 24 channels.

Introduction to Flash Drives

Architecture of Flash Drives

- Memory chips in flash drives are logically organized in blocks and pages.
- A page is the smallest object that can be read or written on a flash drive.
- Pages are grouped together into blocks.
- A block may have 32, 64, or 128 pages.
- Pages do not have a standard size; typical page sizes are 4 KB,8 KB, and 16 KB.
-



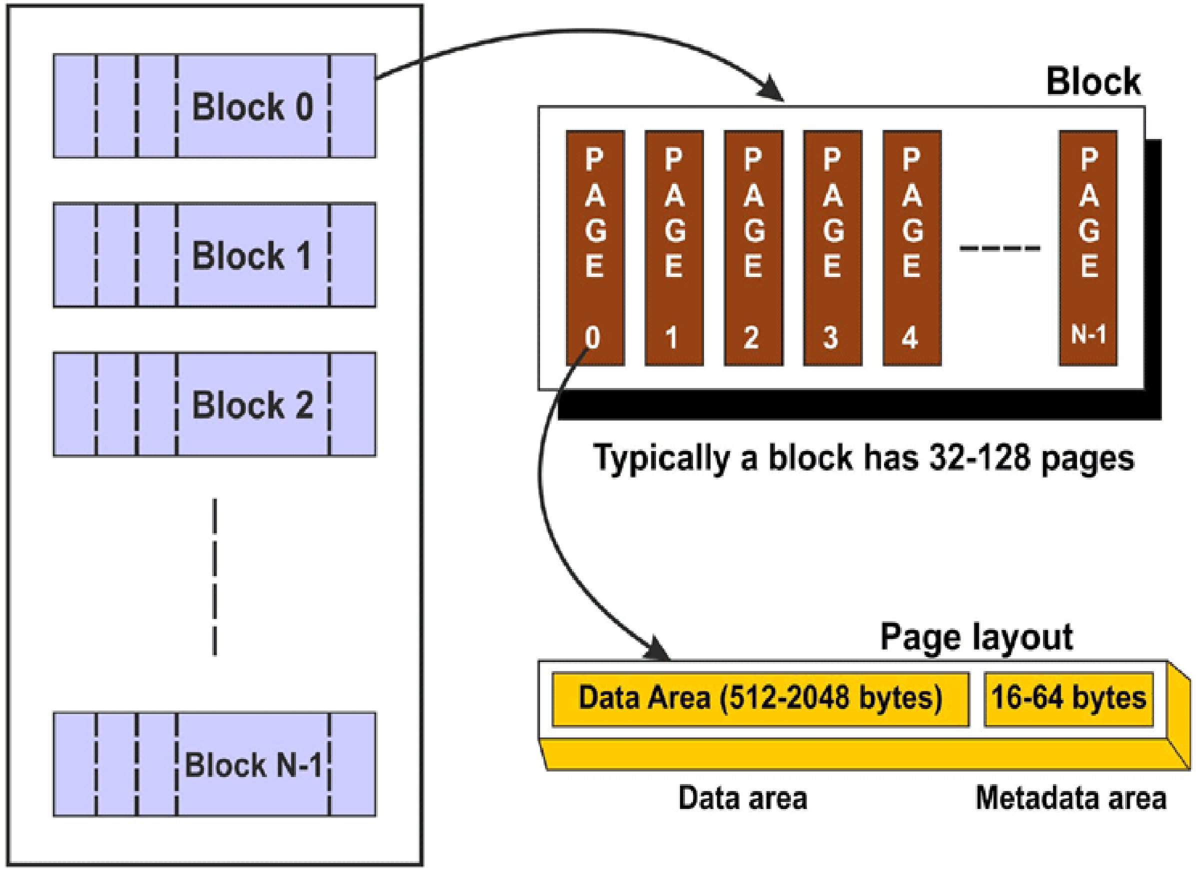
Introduction to Flash Drives

Architecture of Flash Drives

- For example, a 4-KB page would span across eight 512-byte data blocks with
- consecutive addresses.

In flash drives,

- ❑ a read operation can happen at the page level
- ❑ a write or an erase operation happens only at the block level.



RAMAIAH
Institute of Technology

Introduction to Flash Drives

Features of Enterprise Flash Drives

- NAND flash memory technology:
  - NAND memory technology is well suited for accessing random data.
  - A NAND device uses bad block tracking and error-correcting code (ECC) to maintain data integrity and provide the fastest write speeds.

Introduction to Flash Drives

Features of Enterprise Flash Drives

- Single-Level Cell (SLC)-based flash:
- NAND technology is available in two different cell designs.
  - A multi-level cell (MLC) stores more than one bit per cell by virtue of its capability to register multiple states, versus a single-level cell that can store only 1 bit.
  - SLC is the preferred technology for enterprise data applications due to its performance and longevity.

Introduction to Flash Drives

Features of Enterprise Flash Drives

- Single-Level Cell (SLC)-based flash:
  - SLC read speeds are typically rated at twice those of MLC devices, and write speeds are up to four times higher.
  - SLC devices typically have 10 times higher write erase cycles, compared to MLC designs.
  - SLC flash memory has higher reliability because it stores only 1 bit per cell. Hence, the likelihood for error is reduced.

RAMAIAH Institute of Technology

Introduction to Flash Drives

Features of Enterprise Flash Drives

- Write leveling technique:
- An important element of maximizing a flash drive’s useful life is ensuring that the individual memory cells experience uniform use.
- This means that data that is frequently updated is written to different locations to avoid rewriting the same cells.
- In EFDs, the device is designed to ensure that with any new write operation, the youngest block is used.

RAMAIAH Institute of Technology

UNIT 1- Storage System

Introduction to Information Storage(Chapter 1)

- Information Storage
- Evolution of Storage Architecture,
- Data Centre Infrastructure
- Virtualization and Cloud Computing.

Data Centre Environment(Chapter 2)

- Application
- Database Management System (DBMS)
- Host
- Connectivity
- Storage
- Disk Drive Components
- Disk Drive Performance
- Host Access to Data
- Direct-Attached Storage
- Storage Design Based on Application
- Disk Native Command Queuing,
- Introduction to Flash Drives.


Thank
You