File Organization

Radhika Rani Chintala

File Organization

Part-1

Agenda

- Introduction
- Storage Devices
- Secondary Storage devices
- Hardware Mechanism of Disk drive
- Tertiary Storage devices

Introduction

The collection of data that makes up a computerized database must be stored physically on some computer **storage medium**.

The DBMS software can then retrieve, update, and process this data as needed.

Computer storage media form a **Storage Hierarchy** that includes two main categories:

Primary storage

Secondary storage

Tertiary storage

Storage Devices

Primary Storage

- Direct access by CPU.
- Fast access to data
- Limited storage capacity.
- More expensive.
- Volatile
- Main memory & faster cache memory.

Secondary Storage

- No direct access to CPU
- Slow access to data
- Large capacity
- Less expensive
- Non-Volatile
- Online devices
- Magnetic disk & Flash memory
- Most databases are stored permanently on magnetic disks

Tertiary Storage

- No direct access to CPU
- Slower access to data
- Large capacity
- Less expensive than disks
- Non-Volatile
- Offline devices
- Magnetic tapes & Optical disks
- Used for backing up databases

Storage Hierarchy
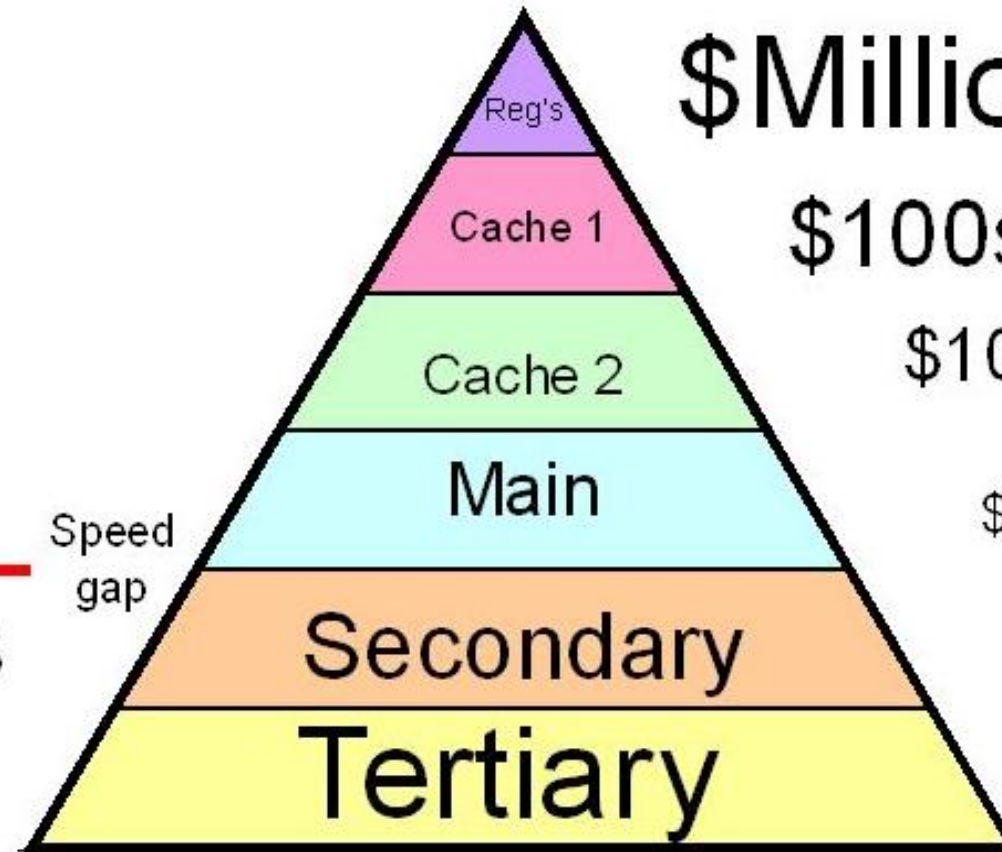
Secondary Storage Devices

Magnetic Disk

- Stores large amounts of data.
- The device that holds the disks is referred to as a **hard disk drive**, or **HDD**.
- Basic unit of data is a **bit.**
- By magnetizing an area on a disk, it can represent a bit value of either 0 or 1.
- Bits are grouped into **bytes** (or **characters**) -- > 4 to 8 bits.
- **Capacity** of a disk = No. of bytes it can store
- Hard disks can hold from several hundred gigabytes up to a few terabytes.
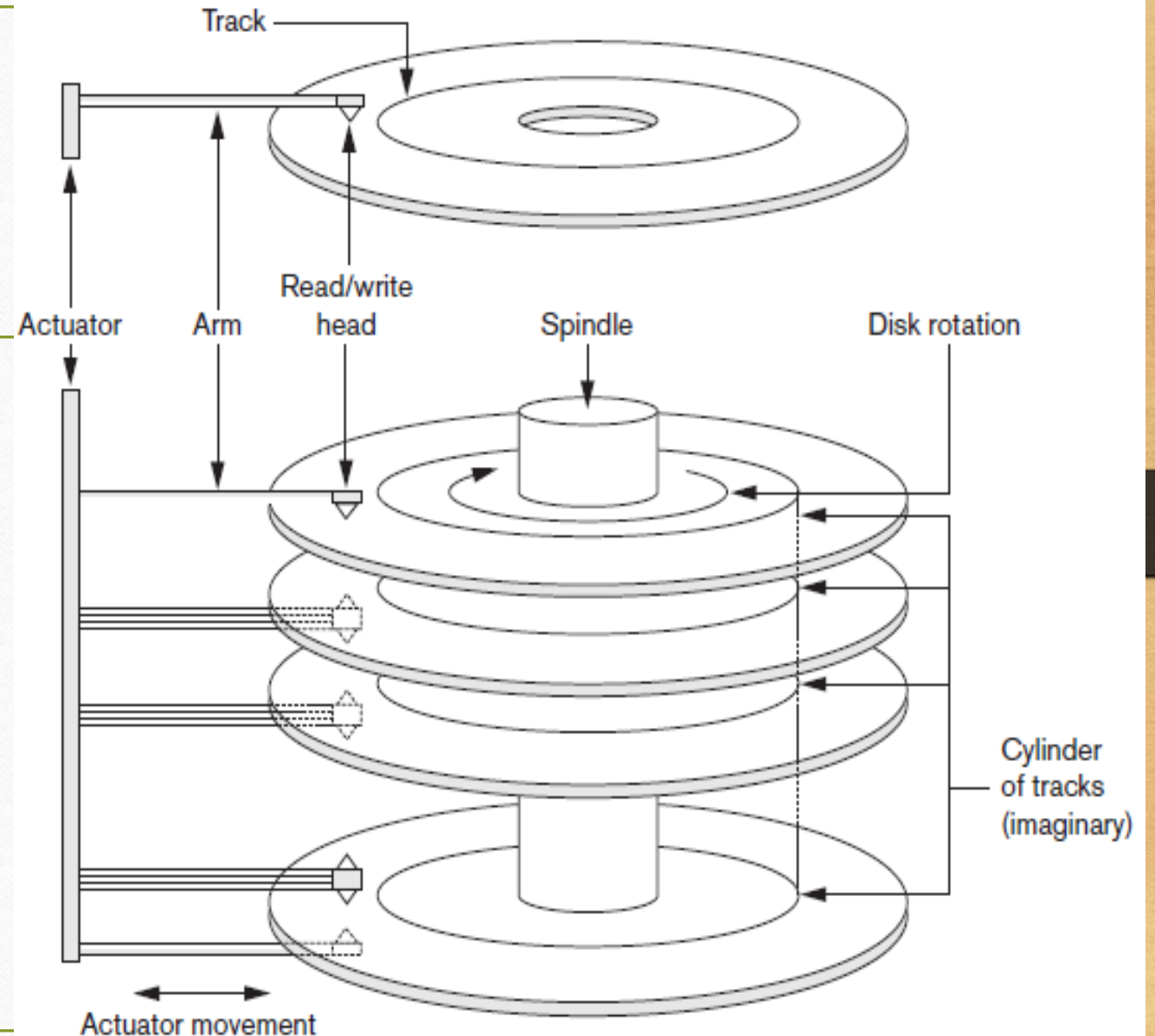
# Hardware description of Disk Devices
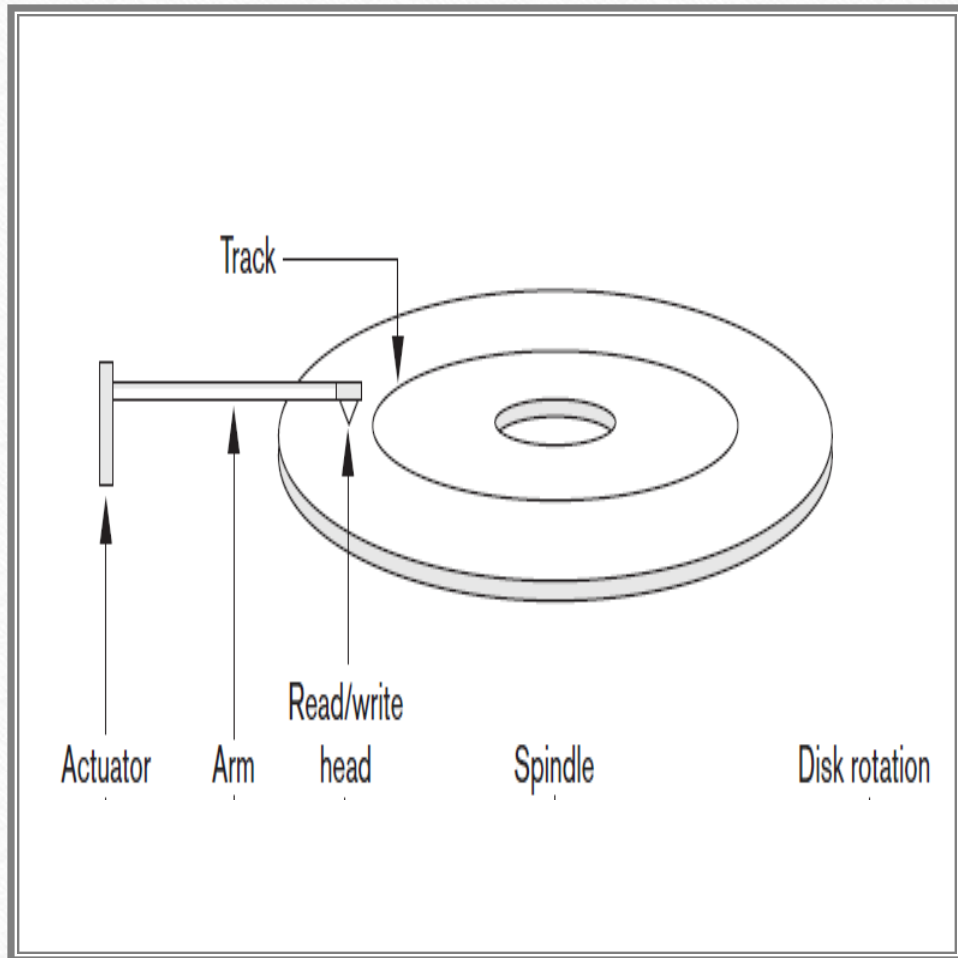
- Made of magnetic material
- Shaped as a thin circular disk
- Protected by a plastic or acrylic cover.
- Single-sided disk and Double-sided disk.
- 3.5" & 2.5" diameter.
- Disks are assembled into a **disk pack,** which may include many disks and therefore many surfaces.

- **Disk Pack:** Which include multiple disks and thus many surfaces.

- **Tracks:** Information is stored on a disk surface in concentric circles of small width, where each circle is called a **track**.

- **Cylinder:** Tracks with the same diameter on the various surfaces are called a **cylinder.**

Track

Actuator   Arm   Read/write head   Spindle   Disk rotation

Cylinder of tracks (imaginary)

Actuator movement

- **Sectors:** Track usually contains a large amount of information, so it is divided into smaller blocks called sectors.

- **Disk Block:** The division of a track into equal-sized **disk blocks** (or **pages**) is set by the operating system during disk formatting (or initialization).

- Sectors subdivided or combined into blocks during initialization.
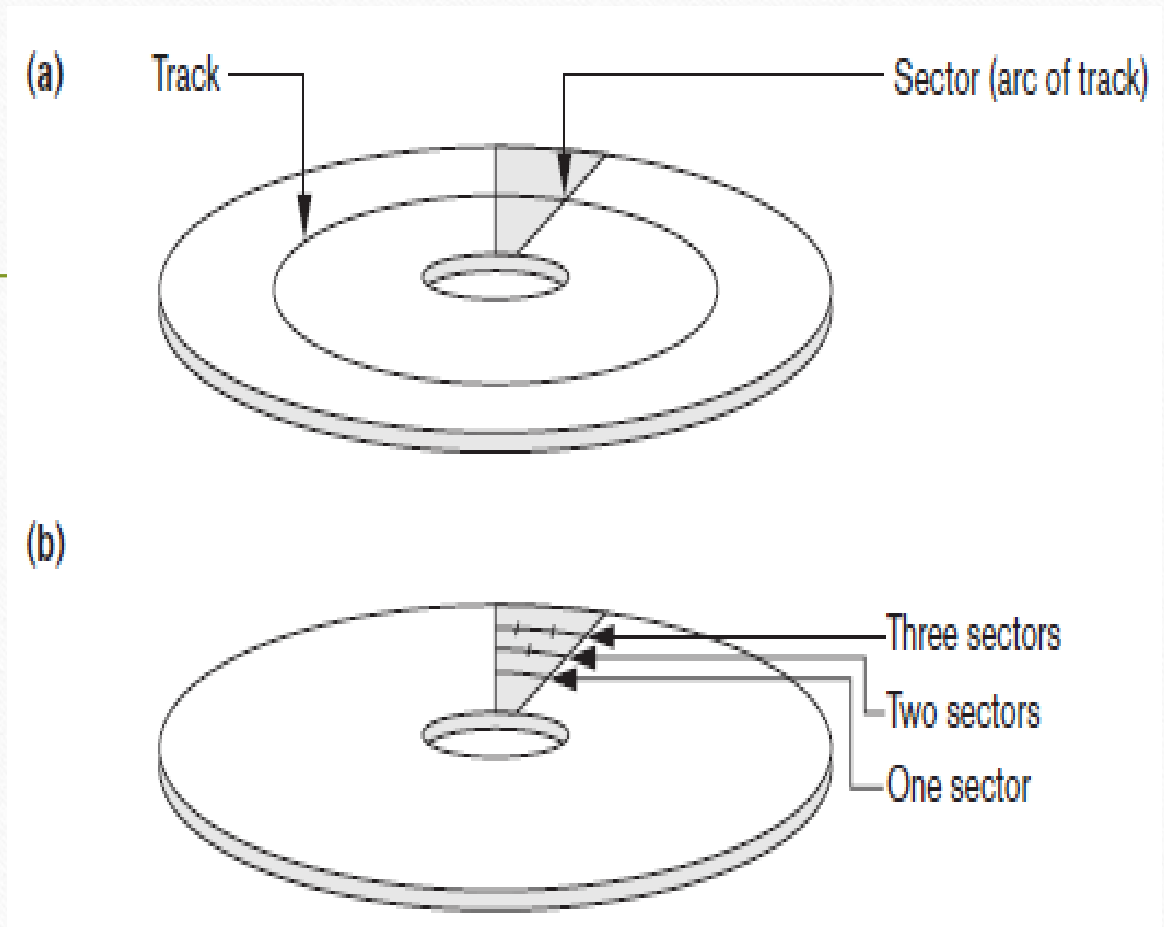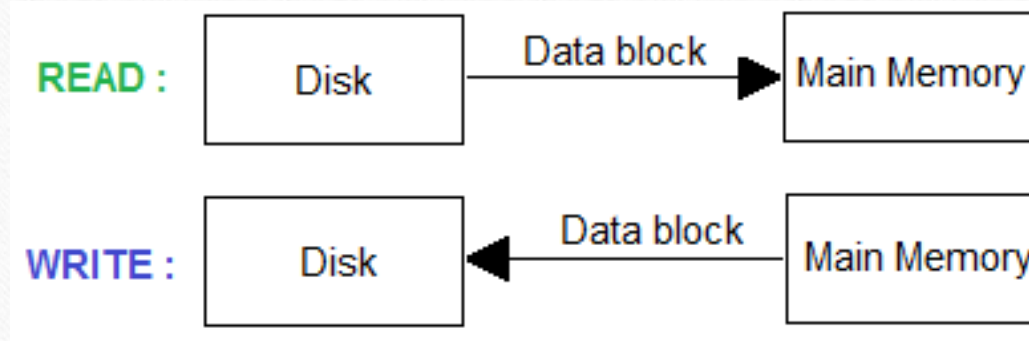
- **Data Transfer:** Data transfer b/w main memory and disk takes place in units of disk blocks.

- **Hardware Block Address** = (cylinder number, track number, block number)

- **Logical block address(LBA)** = number between 0 and n (assuming the total capacity of the disk is n + 1 blocks).

- **Buffer:** a contiguous reserved area in main memory that holds one disk block.

- **Cluster**(several contiguous blocks) may be transferred as a unit. In this case, the buffer size is adjusted to match the number of bytes in the cluster.
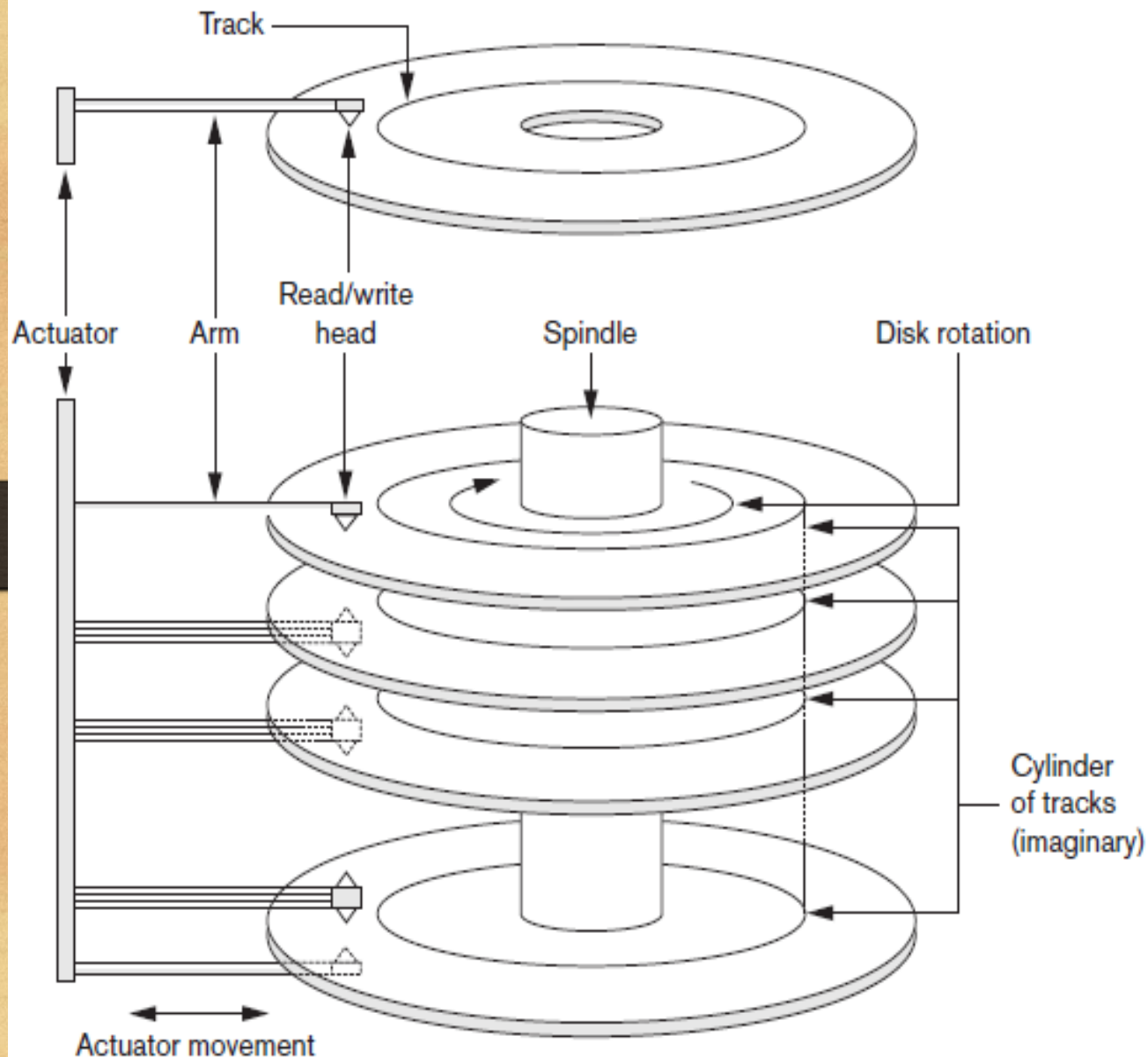
Hardware Mechanism of Disk drive

Disk Drive

Read/Write head

Mechanical Arm

Actuator

Disk Types → Fixed-head disks → Movable-head disks

Interfacing Disk Drives to Computer Systems

Disk Controller : Controls the disk drive and interfaces it to the computer system.

Seek time	• Time taken by the disk controller to position the read/write head on the correct track.
Rotational delay or latency	• Rotational time taken to position the beginning of the desired block under the read/write head.
Block transfer time	• Time taken to transfer the data
Bulk Transfer	• Transfer several consecutive blocks on the same track or cylinder.

Total Block Transfer Time Calculation

Total time = Seek time + Rotational delay + Block transfer time

Bulk transfer rate = Seek time + Rotational delay + (n * Block transfer time)

Tertiary Storage Devices

Magnetic Tapes

- Sequential access devices .
- high-capacity.
- Data is stored on reels. Similar to audiotapes or videotapes.
- A tape drive is required
- A read/write head is used to read or write data blocks on tape.
- Blocks may be larger than those for disks
- Slow access devices
- Used for Backing up the database

Storage Devices

Magnetic Disks

Hardware Mechanism of Disk drive

Magnetic Tapes

Summary of Part-1
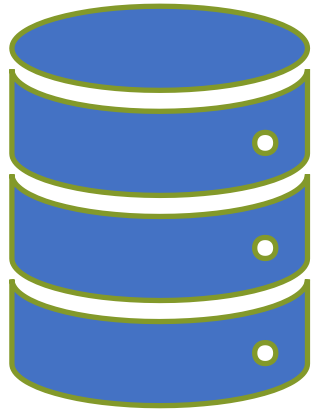
File Organization

Part-2

File Organization

Part-2

Agenda

- Buffer Management

- Various ways of formatting and storing file records on disk.

- Various types of operations that are typically applied to file records.

- Three primary methods for organizing file records on disk:
  - unordered records
  - ordered records
  - hashed records.

Buffer Management

Buffering of Blocks

Multiple Buffers: When several blocks need to be transferred from disk to main memory, several buffers can be reserved in main memory to speed up the transfer.
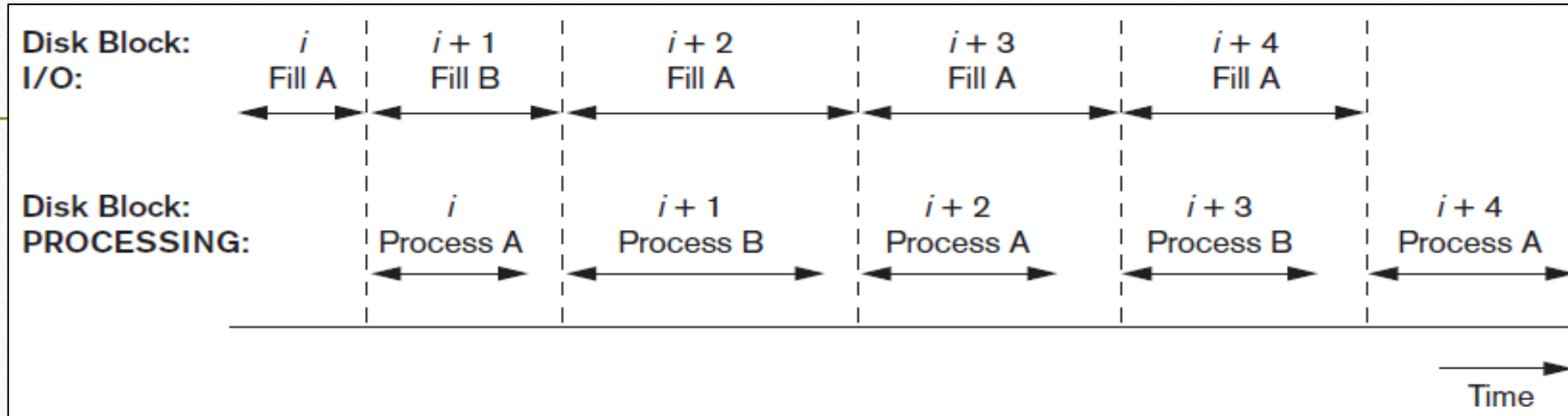
Disk I/O Processor: Controller that takes care of I/O operations.

Parallel processing of Buffers: While one buffer is being read or written, CPU can process data in the other buffer. CPU processing and Disk I/O processing can be done parallelly.

Double Buffering

Disk Block: I/O:	i Fill A	$i + 1$ Fill B	$i + 2$ Fill A	$i + 3$ Fill A	$i + 4$ Fill A

Disk Block: PROCESSING:		i Process A	$i + 1$ Process B	$i + 2$ Process A	$i + 3$ Process B	$i + 4$ Process A

Time

- Reading and processing can proceed in parallel.

- The CPU can start processing a block once its transfer to main memory is completed; at the same time, the disk I/O processor can be reading and transferring the next block into a different buffer.
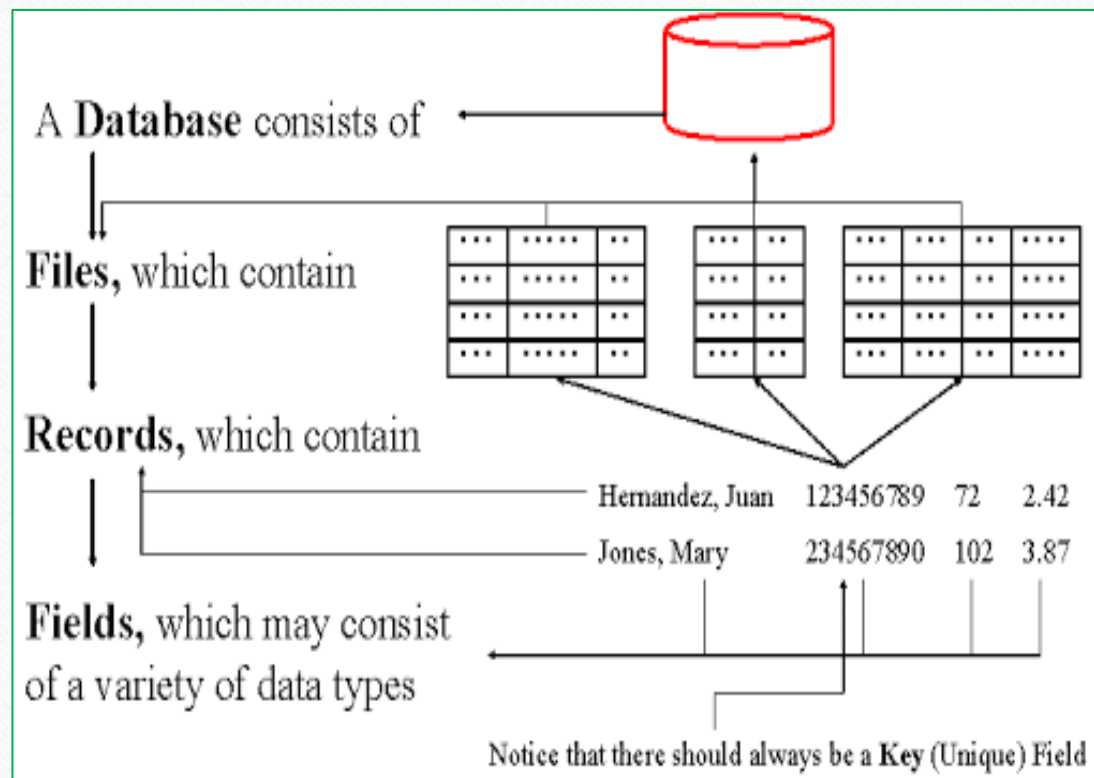
Placing File Records on Disk

Files, Fixed-Length Records, and Variable-Length Records

A **Database** consists of

Files, which contain

Records, which contain

Hernandez, Juan   123456789   72   2.42

Jones, Mary   234567890   102   3.87

Fields, which may consist of a variety of data types

Notice that there should always be a **Key** (Unique) Field

- **File:** a sequence of records. All records in a file are of the same record type.

- **Fixed-length records:** Every record in the file has exactly the same size (in bytes).

- **Variable-length records: D**ifferent records in the file have different sizes.
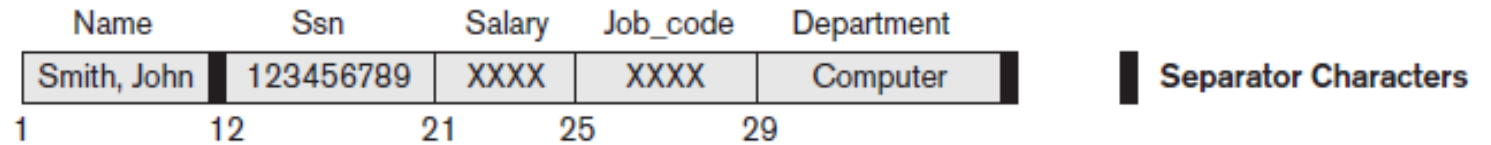
Three Storage Formats

- Fixed Length Record

- Variable Length Record with Separator

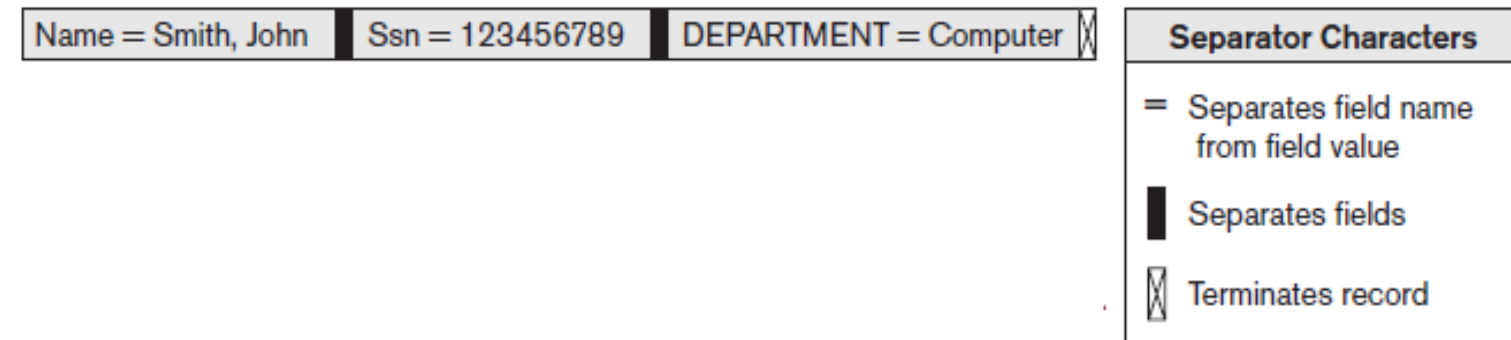- Variable Length Record with multiple Separators

(a) A fixed-length record with six fields and size of 71 bytes.
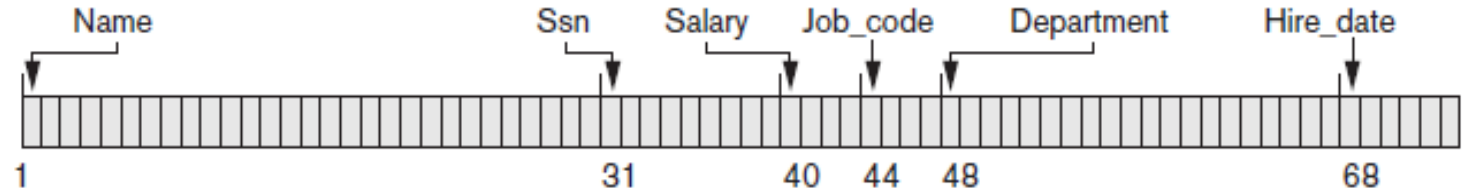
(b) A record with two variable-length fields and three fixed-length fields

(c) A variable-field record with three types of separator characters

Separator Characters

= Separates field name from field value

▮ Separates fields

⊠ Terminates record

Record Blocking

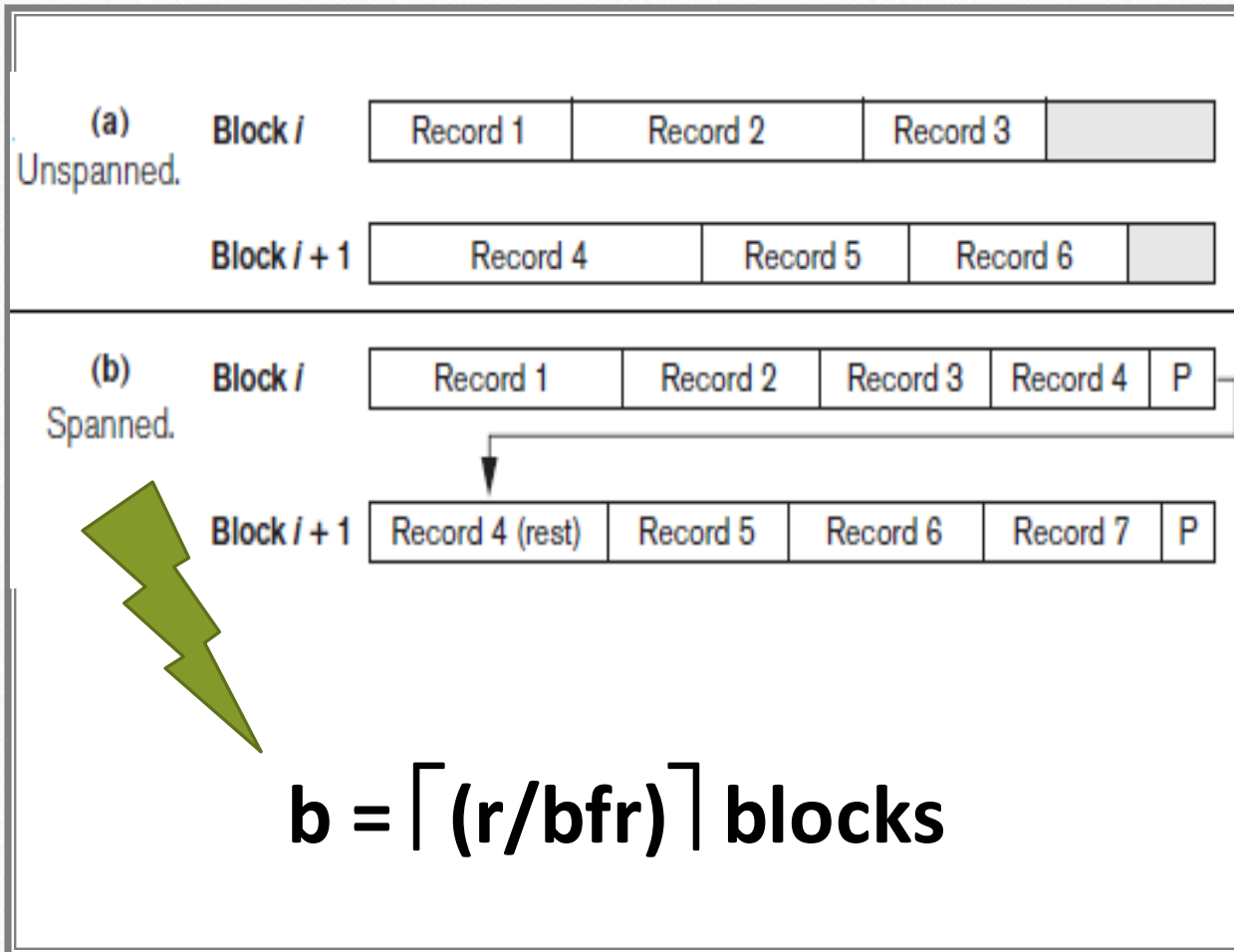- When the **block size > record size**, each block will contain numerous records

- When the **record size > block size**, each record occupies multiple blocks.

- For a file of fixed-length records of size R bytes, with $B >= R$, we can fit

bfr (blocking factor) = $\lfloor B/R \rfloor$ records per block

Unused space in each block = B − (bfr * R) bytes

(a) Unspanned.

| Block *i* | Record 1 | Record 2 | Record 3 | |
| Block *i* + 1 | Record 4 | Record 5 | Record 6 | |

(b) Spanned.

| Block *i* | Record 1 | Record 2 | Record 3 | Record 4 | P |
| Block *i* + 1 | Record 4 (rest) | Record 5 | Record 6 | Record 7 | P |

$$b = \lceil (r/bfr) \rceil \text{ blocks}$$

Spanned & Unspanned Records

- **Unspanned Records:**
  - A record is found in one and only one block.
  - Records do not span across block boundaries.
  - Used with fixed-length records having B > R

- **Spanned Records:**
  - Records are allowed to span across block boundaries.
  - Used with variable-length records having R > B

Allocating File blocks on Disk

| Contiguous Allocation	Linked Allocation	Indexed Allocation
• The file blocks are allocated to consecutive disk blocks. • Reading a file is very fast. • But expanding a file is difficult.	• Each file block contains a pointer to the next file block. • Easy to expand a file • But reading a file is slow	• One or more **index blocks** contain pointers to the actual file blocks.

It is also common to use combinations of these techniques.

Operations on Files
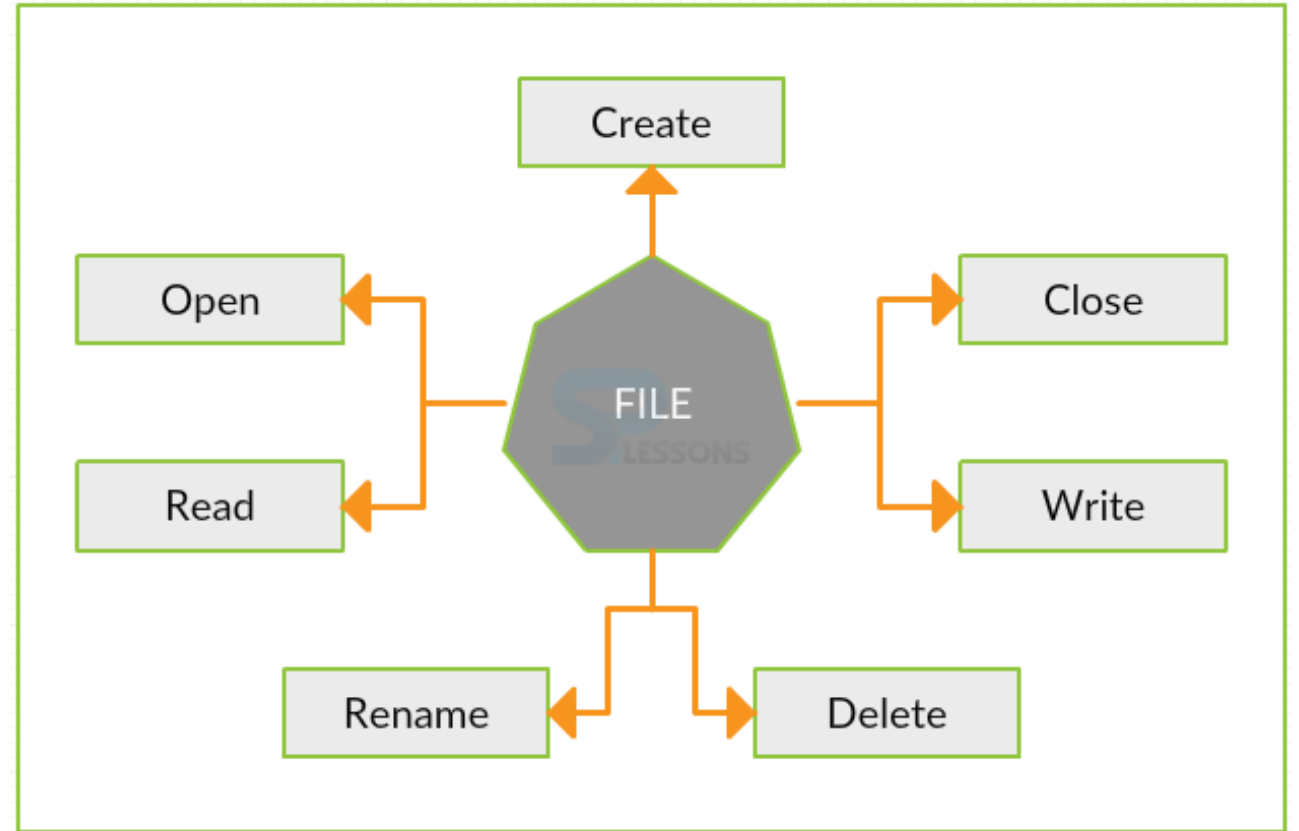
File Operations

- **Open**
- **Reset**
- **Find (or Locate)**
- **Read (or Get)**
- **FindNext**
- **Delete**
- **Modify**

- **Insert**
- **Close**
- **Scan**
- **FindAll**
- **Find (or Locate) n**
- **FindOrdered**
- **Reorganize**

Files of Unordered Records

and

Ordered Records

Files of Unordered Records and Ordered Records

Files of Unordered Records

- Also called as **Heap Files or Pile Files.**

- Records are stored in the same order in which they are created.

- **Insert operation:** *Fast*
- **Search (or Update) operation:** *Slow*
- **Delete operation:** *Slow*
- Deleting a record creates a hole in the page.

Files of Ordered Records

- Also called as **Sorted Files.**

- **Ordering Field**: Records are sorted on the values of one or more fields.

- **Insert operation:** *Poor*
- **Search (or Update) operation:** *Fast*
- **Delete operation:** *Fast*

Summary of Part-2

Buffer Management

Placing File Records on Disk

Operations on Files

Files of Unordered and Ordered Records