

# Data storage & Indexing

Page No.:

Date: / /

## Physical storage media

- i) storage media are classified by the speed with which data can be accessed by the cost per unit of data to buy the medium & by the medium's reliability.
- ii) Data can be loss on power failure or system crash & physical failure of the storage device.
- iii) volatile storage  
Data can be loss when power is switched off.

## Non volatile storage

→ content persist when power is switched off.

- It includes secondary & tertiary storage as well as battery backup main memory.

### 1) Cache

- It is a fastest & most costly form of storage.
- It is relatively small, its use is managed by the computer system hardware.

### 2) Main memory

- The storage medium used for data that are available to be operated on is main memory.
- ii) Accessing time is fast, 10s to 100s of nanoseconds  
so 1 nanosecond =  $10^{-9}$  seconds.
- iii) It is generally too small to store the entire database.
- iv) contents of main memory are usually lost if a power failure or system crash occurs.

### 3) Flash memory

- i) In flash memory data survives if power failure.
- ii) In that data can be written at location only once but location can be erased & written to again.

- iii) It can support only limited number of write-erase cycles.
- iv) Erasing of memory has to be done to an entire bank of memory.
- v) Reads are roughly as fast as main memory but writes are slow, erase is slower.
- vi) Cost per unit of storage roughly similar to main memory.
- vii) It is widely used in embedded devices such as digital cameras.
- viii) It is a type of EEPROM (Electrical Erasable Programmable Read only Memory)

#### 4) Magnetic disk

- i) In magnetic disk data is stored on a spinning disk & read / written magnetically.
- ii) Data must be moved from disk to main memory for access & written back for storage.
- iii) Access time is slower than main memory.
- iv) We can read data on disk in any order unlike magnetic tape.
- v) Capacity of disk is upto 400GB.
- vi) It can survive power failure & system crashes, disk failure can destroy data but is rare.

#### 5) Optical storage

- i) It is a non volatile, data is read optically from spinning disk using a laser.
- ii) CD-ROM & DVD
- iii) Write one, read many (WORM) optical disks used for archival storage (CD-R, DVD-R, DVD+R)
- iv) Multiple write versions also available CD-RW, DVD-RW, DVD+RW & DVD-RAM
- v) It can read & write are slower than with magnetic disk.



## 6) Tape storage

- i) It is a non volatile, used primarily for backup & for archival data.
- ii) Access time is slower than disk.
- iii) Very high capacity available
- iv)

## Storage Hierarchy

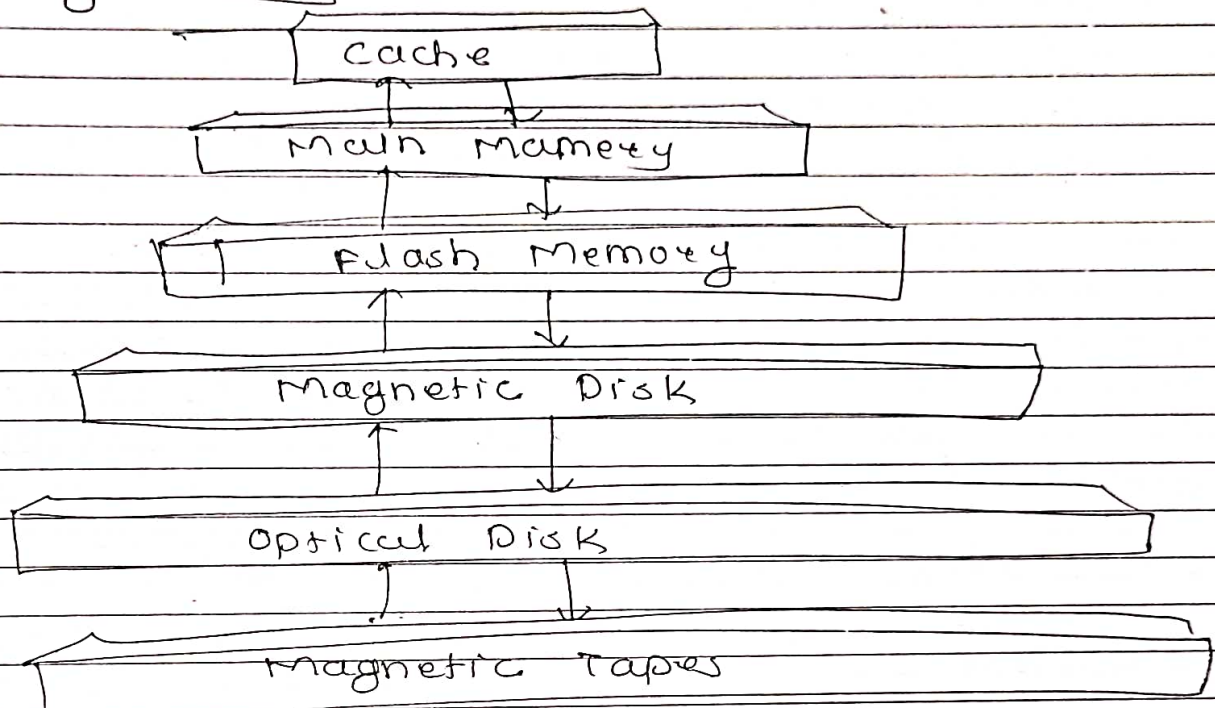


fig → Storage Hierarchy

### 1) Primary storage →

It is fastest but volatile. In that cache & main memory included.

### 2) Secondary storage

- It is the next level in hierarchy
- It is non-volatile
- Fast Access time.
- It is also called as online storage
- E.g → flash memory & magnetic disk

### 8) Tertiary storage

i) It is lowest level in hierarchy

ii) It is non volatile

iii) slow Access time

iv) It is also called as off line storage

v) e.g → Magnetic tape, optical storage

## File Organization

### i) File

Each file is a collection of records. & records is a sequence of fields.

- ii) A DB is mapped into a no. of different files that are maintained by the underlying OS. These files resides permanently on disk.
- iii) File is a logically sequence of records, these records are mapped onto disk blocks.
- iv) Each file is also logically partitioned into fixed length & storage units called blocks, which are the units of both storage allocation & data transfer.

v)