**Database System Concepts**

**Notes**

1. **Physical Storage**

To understand how databases store data, it is necessary to understand the physical constraints of our storage. The physical characteristics of storage devices play a major role in the way data are stored. For example: Disk access takes tens of milliseconds, whereas memory access takes a tenth of a microsecond.

Storage media are classified by the speed with which data can be accessed, the cost per unit of data to buy the medium and the reliability of the medium. Typically, we can divide media into the following:

1. **Cache:** Fastest, most costly, volatile, usually managed by computer hardware. We will not concern ourselves with managing this in our database however some implementors do.
2. **Main Memory:** Storage medium used for data that are available to be operated on. Can be up to several gigabytes but entire database is not stored here. Volatile.
3. **Flash Memory:** Non-volatile aka data is retained upon power loss. Mainly two types -> NAND and NOR Flash. Of these, NAND Flash is better and is widely used. Cheaper than Main Memory. Solid-State Drives, USBs etcetera use Flash Memory.
4. **Magnetic-disk storage:** Primary medium for long term storage in 2009. System moves data from here to main memory for operation. Non-volatile but sometimes fails and destroys data, but such failures are rare.
5. **Optical Storage:** Old stuff; Compact Disks (CDs), Digital Video Disks (DVDs), Blu-ray DVDs. Optical Disks are also used in read-only cases (CD-ROM, DVD-ROM). There are also record-once variants (CD-R, DVD-R) These are called write-once, read-many (WORM) disks. Can also have write-many variants (CD-RW, DVD-RW)
6. **Tape Storage:** Archaic stuff, long tapes that are read sequentially and hold enormous amounts of data. Archival in nature.

And therefore, we have a hierarchy of storages, based on the cost-capacity-speed metric, with cache at top for highest speed, cost and lowest capacity and moving all the way down to tape storage.

Cache and Main Memory are called primary storage. Magnetic Disks are referred to as secondary or online storage and the lowest level is referred to as tertiary storage or offline storage.

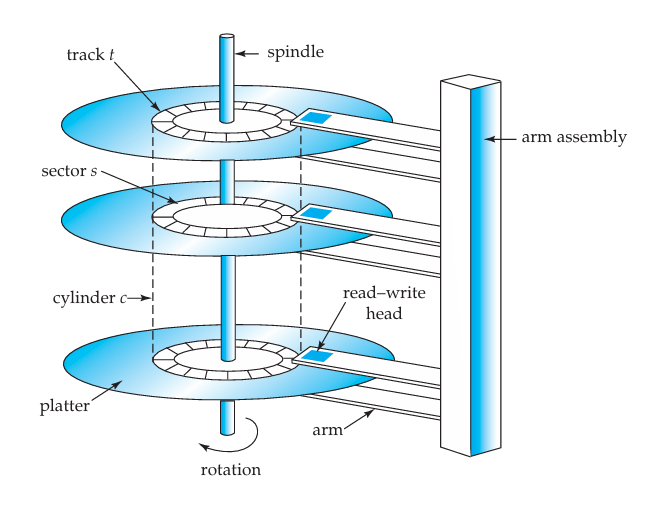
A second sorting is enforced on the basis of volatility with cache and main memory being the only volatile ones.

1. **Magnetic Disk and Flash Storage**
2. **Physical Characteristics of Disks:**

A physical disk has platters. Each platter has a flat, circular shape. Its 2 surfaces are covered with magnetic material and information is recorded on the surfaces.

A disk surface is logically divided into tracks, which are subdivided into sectors.

A sector is the smallest unit of information that can be read from or written to the disk. Sector sizes are typically 512 bytes. There are about 50,000 to 100,000 tracks per platter and 1 to 5 platters per disk. The inner tracks (closer to the spindle) are of smaller length and the outer tracks contain more sectors than the inner tracks; typical numbers are around 500 to 1000 sectors per track in the inner tracks, and around 1000 to 2000sectors per track in the outer tracks.



The read–write head stores information on a sector magnetically as reversals of the direction of magnetization of the magnetic material.

A disk controller interfaces between the computer system and the actual hardware of the disk drive. It accepts high-level commands to read or write a sector, and initiates actions, such as moving the disk arm to the right track and actually reading or writing the data.

Another interesting task that disk controllers perform is remapping of bad sectors. If the controller detects that a sector is damaged when the disk is initially formatted, or when an attempt is made to write the sector, it can logically map the sector to a different physical location

1. **Performance Measures of Disks**

Access time is the time from when a read or write request is issued to when data transfer begins. The time for repositioning the arm is called the seek time, and it increases with the distance that the arm must move.

Once the head has reached the desired track, the time spent waiting for the sector to be accessed to appear under the head is called the rotational latency time. The access time is then the sum of the seek time and the latency, and ranges from 8 to 20 milliseconds.

The mean time to failure of a disk (or of any other system) is the amount of time that, on average, we can expect the system to run continuously without any failure. According to vendors’ claims, the mean time to failure of disks today ranges from 500,000 to 1,200,000 hours— about 57 to 136 years.

1. **Optimization of Disk-Block Access**

A block is a logical unit consisting of a fixed number of contiguous sectors. Block sizes range from 512 bytes to several kilobytes.

A sequence of requests for blocks from disk may be classified as a sequential access pattern or a random-access pattern. In a sequential access pattern, successive requests are for successive block numbers, which are on the same track, or on adjacent tracks.

In a random-access pattern, successive requests are for blocks that are randomly located on disk. Each such request would require a seek.