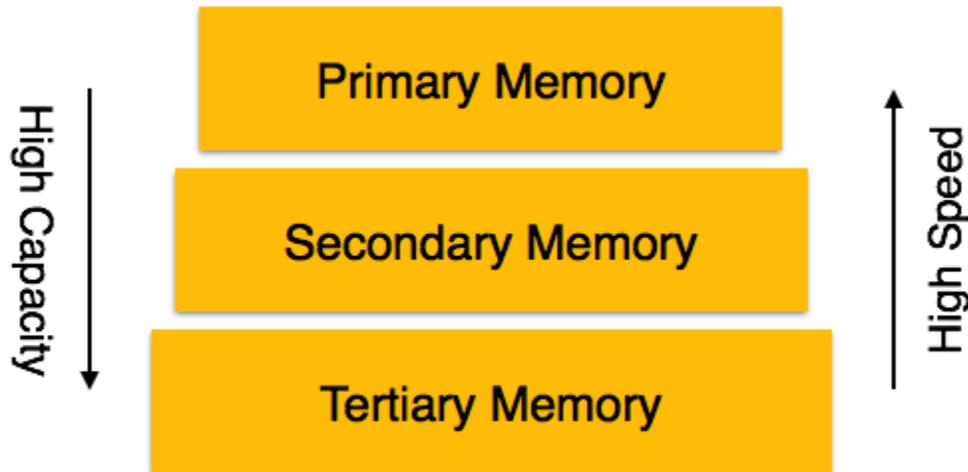# DBMS Storage System

Databases are stored in file formats, which contains records. At physical level, actual data is stored in electromagnetic format on some device capable of storing it for a longer amount of time. These storage devices can be broadly categorized in three types:



[*Image: Memory Types*]

- **Primary Storage:** The memory storage, which is directly accessible by the CPU, comes under this category. CPU's internal memory (registers), fast memory (cache) and main memory (RAM) are directly accessible to CPU as they all are placed on the motherboard or CPU chipset. This storage is typically very small, ultra fast and volatile. This storage needs continuous power supply in order to maintain its state, i.e. in case of power failure all data are lost.
- **Secondary Storage:** The need to store data for longer amount of time and to retain it even after the power supply is interrupted gave birth to secondary data storage. All memory devices, which are not part of CPU chipset or motherboard comes under this category. Broadly, magnetic disks, all optical disks (DVD, CD etc.), flash drives and magnetic tapes are not directly accessible by the CPU. Hard disk drives, which contain the operating system and generally not removed from the computers are, considered secondary storage and all other are called tertiary storage.
- **Tertiary Storage:** Third level in memory hierarchy is called tertiary storage. This is used to store huge amount of data. Because this storage is external to the computer system, it is the slowest in speed. These storage devices are mostly used to backup the entire system. Optical disk and magnetic tapes are widely used storage devices as tertiary storage.

## Memory Hierarchy

A computer system has well-defined hierarchy of memory. CPU has inbuilt registers, which saves data being operated on. Computer system has main memory, which is also directly accessible by CPU. Because the access time of main memory and CPU speed varies a lot, to minimize the loss cache memory is introduced. Cache memory contains most recently used data and data which may be referred by CPU in near future.

The memory with fastest access is the costliest one and is the very reason of hierarchy of memory system. Larger storage offers slow speed but can store huge amount of data compared to CPU registers or Cache memory and these are less expensive.

## Magnetic Disks

Hard disk drives are the most common secondary storage devices in present day computer systems. These are called magnetic disks because it uses the concept of magnetization to store information. Hard disks consist of metal disks coated with magnetizable material. These disks are placed vertically a spindle. A read/write head moves in

between the disks and is used to magnetize or de-magnetize the spot under it. Magnetized spot can be recognized as 0 (zero) or 1 (one).

Hard disks are formatted in a well-defined order to stored data efficiently. A hard disk plate has many concentric circles on it, called tracks. Every track is further divided into sectors. A sector on a hard disk typically stores 512 bytes of data.

# RAID

Exponential growth in technology evolved the concept of larger secondary storage medium. To mitigate the requirement RAID is introduced. RAID stands for Redundant Array of Independent Disks, which is a technology to connect multiple secondary storage devices and make use of them as a single storage media.

RAID consists an array of disk in which multiple disks are connected together to achieve different goals. RAID levels define the use of disk arrays.

- **RAID 0:** In this level a striped array of disks is implemented. The data is broken down into blocks and all blocks are distributed among all disks. Each disk receives a block of data to write/read in parallel. This enhances the speed and performance of storage device. There is no parity and backup in Level 0.

[Image: RAID 0]

- **RAID 1:** This level uses mirroring techniques. When data is sent to RAID controller it sends a copy of data to all disks in array. RAID level 1 is also called mirroring and provides 100% redundancy in case of failure.
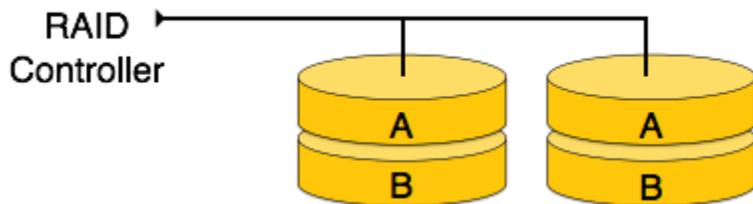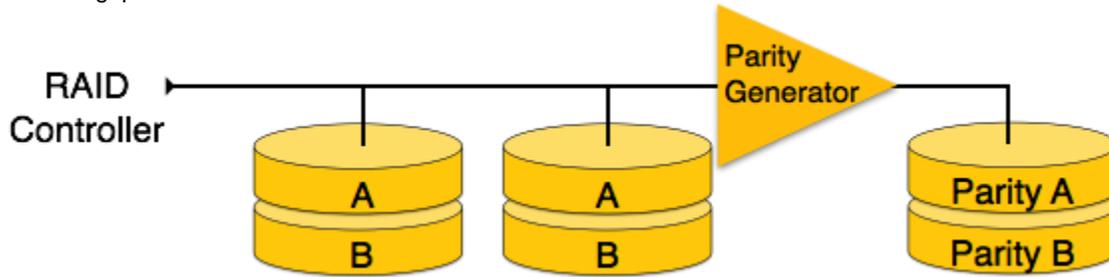
[Image: RAID 1]

- **RAID 2:** This level records the Error Correction Code using Hamming distance for its data striped on different disks. Like level 0, each data bit in a word is recorded on a separate disk and ECC codes of the data words are stored on different set disks. Because of its complex structure and high cost, RAID 2 is not commercially available.
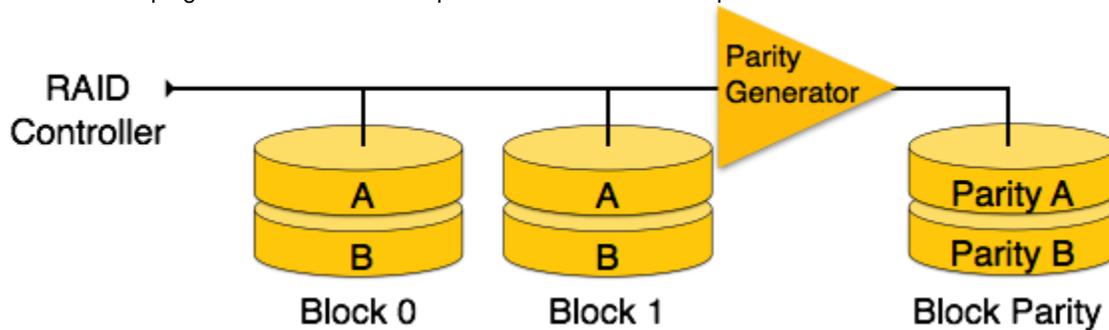
[Image: RAID 2]

- **RAID 3:** This level also stripes the data onto multiple disks in array. The parity bit generated for data word is stored on a different disk. This technique makes it to overcome single disk failure and a single disk failure does not impact the throughput.
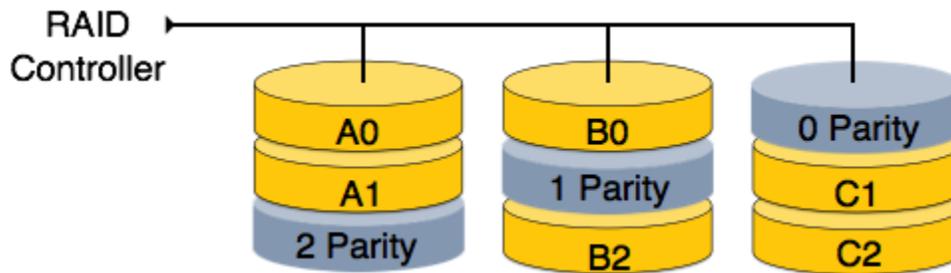


[*Image: RAID 3*]

- **RAID 4:** In this level an entire block of data is written onto data disks and then the parity is generated and stored on a different disk. The prime difference between level 3 and 4 is, level 3 uses byte level striping whereas level 4 uses block level striping. Both level 3 and 4 requires at least 3 disks to implement RAID.
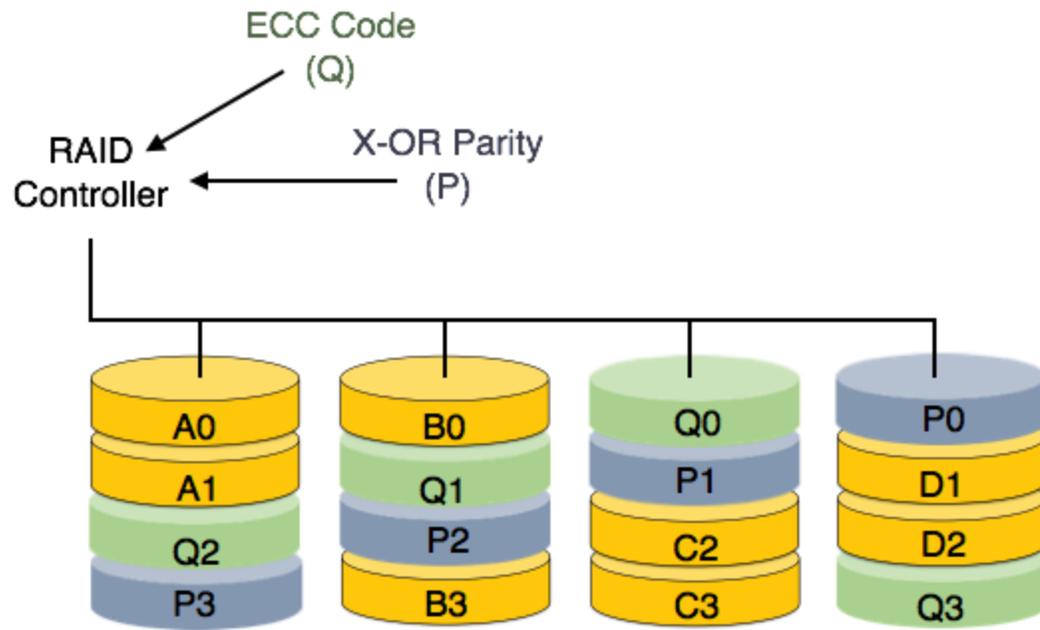


[*Image: RAID 4*]

- **RAID 5:** This level also writes whole data blocks on to different disks but the parity generated for data block stripe is not stored on a different dedicated disk, but is distributed among all the data disks.



[*Image: RAID 5*]

- **RAID 6:** This level is an extension of level 5. In this level two independent parities are generated and stored in distributed fashion among disks. Two parities provide additional fault tolerance. This level requires at least 4 disk drives to be implemented.

ECC Code
(Q)

RAID
Controller

X-OR Parity
(P)

| A0 | B0 | Q0 | P0 |
| A1 | Q1 | P1 | D1 |
| Q2 | P2 | C2 | D2 |
| P3 | B3 | C3 | Q3 |

[*Image: RAID 6*]

Source:

http://www.tutorialspoint.com/dbms/dbms_storage_system.htm