Competency 02

2.4 Examines PC memory system to identify different types of memory and their main characteristics.

Time: 6 periods

Learning Outcomes

- Briefly explains the memory hierarchy with a suitable diagram.
- Describes the need for different types of memory and their characteristics.
- Briefly explains the volatile and non- volatile memory.
- Lists volatile and non-volatile memories in computer.
- Describes the characteristics of memory in terms of performance, location, capacity, access method, cost, physical type and physical arrangement of data (bits into words)
- Lists and briefly explains the types and characteristics of ROMs
- Compares and contrasts each type of memory in terms of access time, cost / MB, capacity (typical amount used)

Memory hierarchy

In the Computer System Design, Memory Hierarchy is an enhancement to organize the memory such that it can minimize the access time. The Memory Hierarchy was developed based on a program behavior known as locality of references.

This Memory Hierarchy Design is divided into 2 main types:

1. **External Memory or Secondary Memory –**

Comprising of Magnetic Disk, Optical Disk, and Magnetic Tape i.e. peripheral storage devices which are accessible by the processor via I/O Module.

2. **Internal Memory or Primary Memory –**

Comprising of Main Memory, Cache Memory & CPU registers. This is directly accessible by the processor.

Characteristics of Memory Hierarchy Design

Capacity:

It is the global volume of information the memory can store. As we move from top to bottom in the Hierarchy, the capacity increases.

Access Time:

It is the time interval between the read/write request and the availability of the data. As we move from top to bottom in the Hierarchy, the access time increases.

Performance:

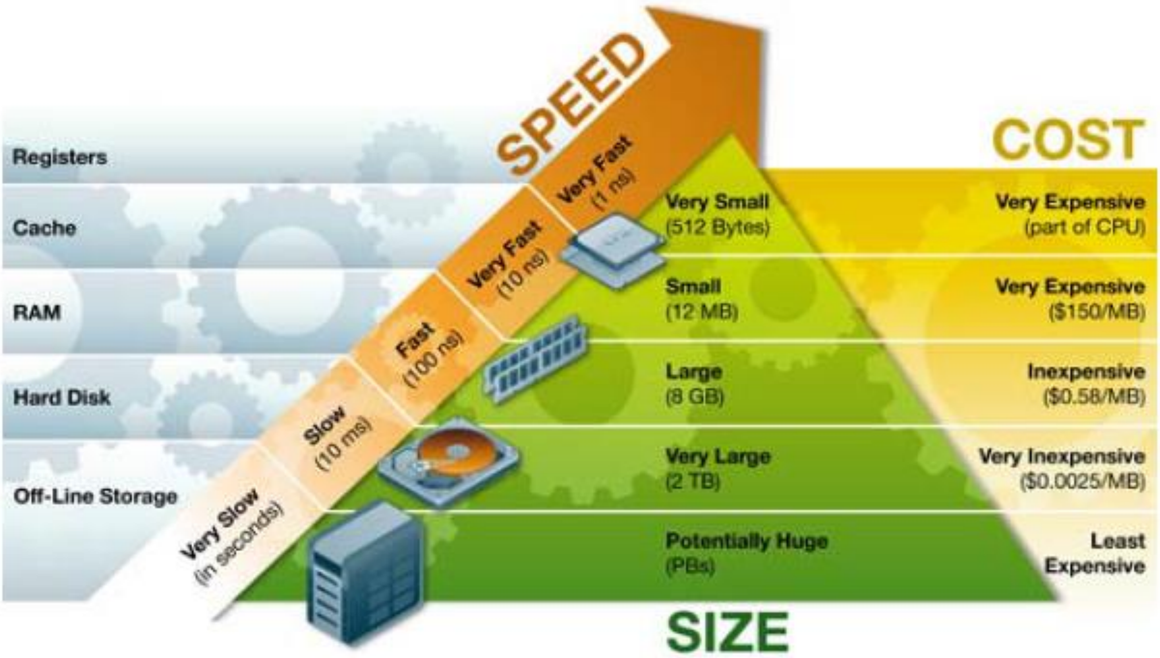
Earlier when the computer system was designed without Memory Hierarchy design, the speed gap increases between the CPU registers and Main Memory due to large difference in access time. This results in lower performance of the system and thus, enhancement was required. This enhancement was made in the form of Memory Hierarchy Design because of which the performance of the system increases. One of the most significant ways to increase system performance is minimizing how far down the memory hierarchy one has to go to manipulate data.

Cost per bit:

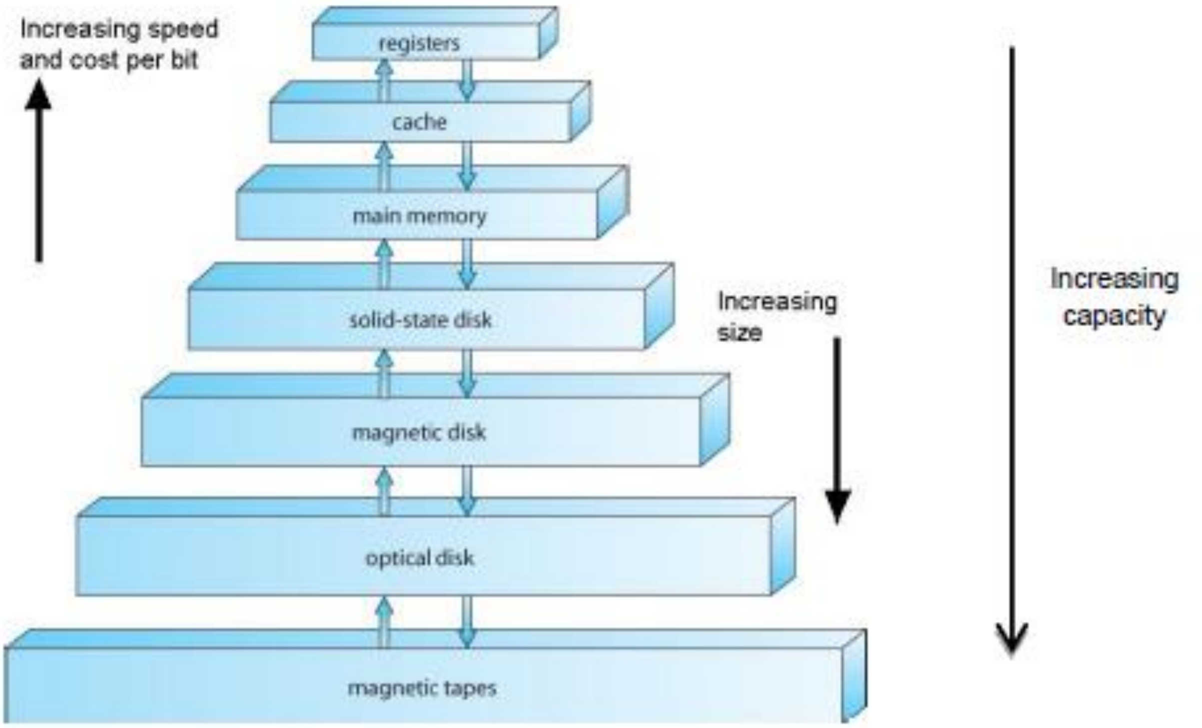
As we move from bottom to top in the Hierarchy, the cost per bit increases i.e. Internal Memory is costlier than External Memory.

Physical size/ density of data

When we move from top to bottom in the Hierarchy, the physical size of memory is increases.



Note: as of 2015



Different types of memory and their characteristics

There are two types of memory

1. Volatile memory
2. Non-volatile memory

The main difference between volatile memory and non-volatile memory is that volatile memory is not stored permanently memory and in non-volatile memory, memory is stored permanently.

Basic comparison between volatile and non-volatile memory

Basis	**Volatile Memory**	**Non Volatile Memory**
Definition	Data is not stored in memory as soon as power is gone.	Data remains stored in the memory even if power is done
Temporary	Temporary memory	Not temporary memory but is permanent memory
Performance	Fast	Slow that Volatile memory
Example	RAM	ROM
Storage	Primary	Secondary

Volatile memory

Volatile memory is a computer storage that only holds the data while the device is powered.

Eg: Register, Cache memory, RAM

Cache memory

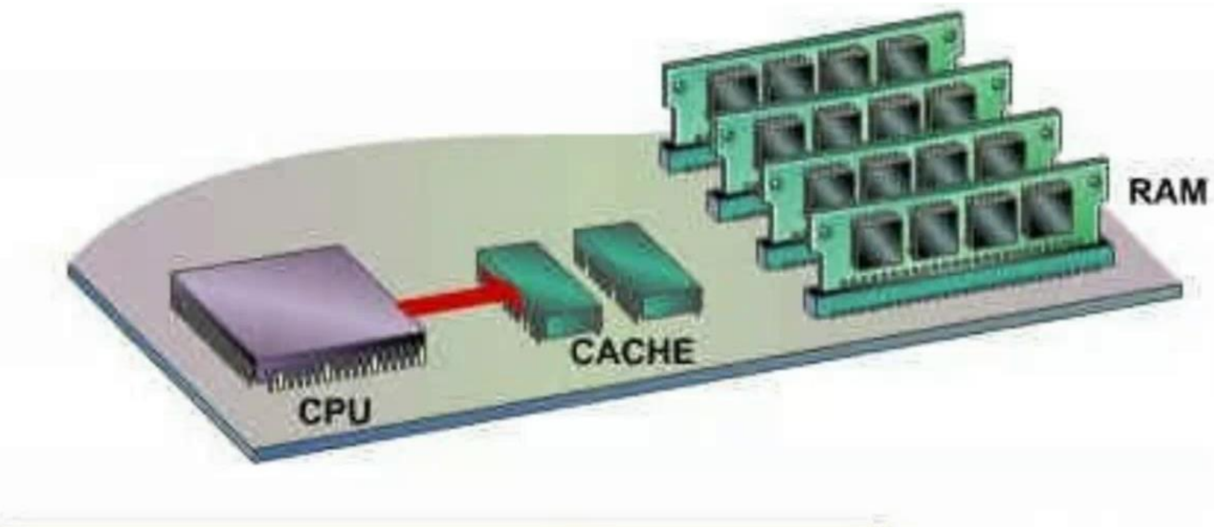
The cache memory is used to store program instructions that are frequently accessed by software during operation.

Types of cache memories

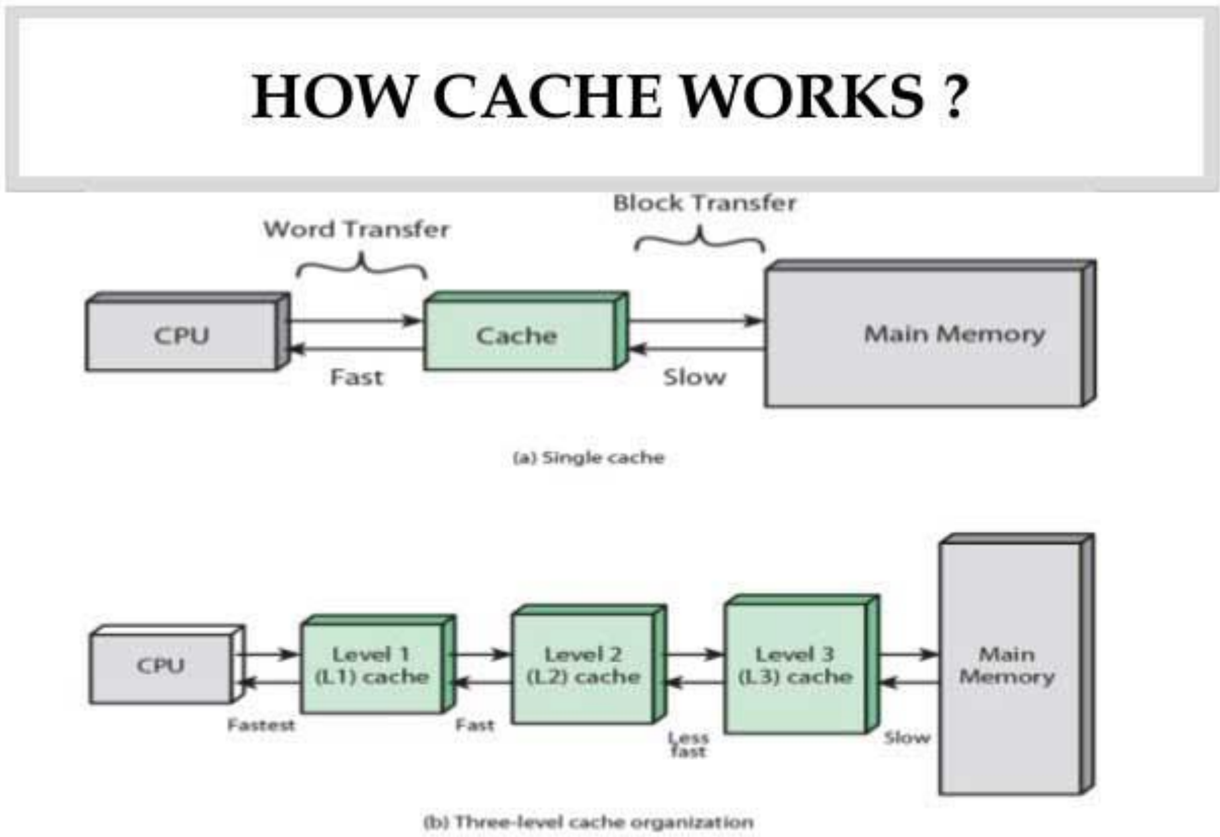
• **Level 1 (L1) cache** is extremely fast but relatively small, and is usually embedded in the processor chip (CPU).

• **Level 2 (L2) cache** is often more capacity than L1. It may be located on the CPU or on a separate.

• **Level 3 (L3) cache** is typically specialized memory that works to improve the performance of L1 and L2. It can be significantly slower than L1 or L2, but is usually double the speed of RAM.



When the CPU needs data, it looks first in cache memory.



(a) Single cache

(b) Three-level cache organization

RAM (Random Access Memory)

- RAM is the **main memory /a primary storage** of the computer.
- RAM is directly accessible to processor. It holds instructions to execute, data being data to process and information to output.
- RAM Provides **random access** (any memory cell can be accessed directly).
- RAM is **volatile** (everything is lost as power is cutoff).
- RAM is fast and expensive (relative to the **secondary storage**)
- Two technologies are used in RAM,
    - Static RAM (SRAM)
    - Dynamic RAM (DRAM)

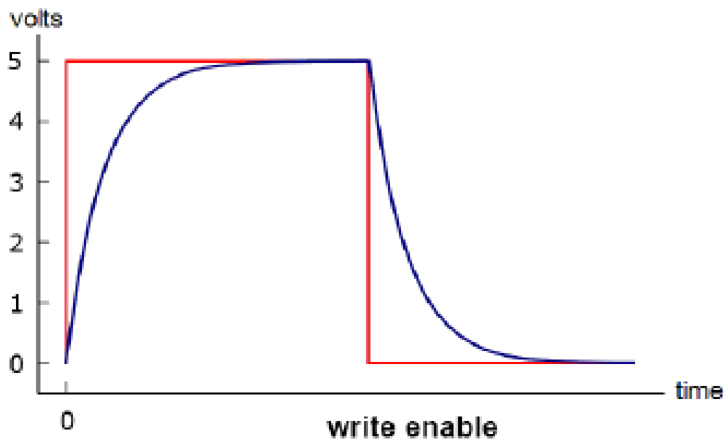
Static RAM

- In static RAM, each memory cell is implemented with a flip-flop and stores on bit.
- Very fast
- Low density (low capacity per area of silicon)
- Consumes more power and generates more heat
- Expensive

Dynamic RAM

- In Dynamic RAM, each memory cell is implemented with a transistor and a capacitor and stores one bit.
- Consumes less power and generate less heat
- High density
- Inexpensive , so widely used
- Not as fast as SRAM

- **Problem** of DRAM,
    - Capacitor looses charge with time
    - Stored value 1 may change to 0 with time

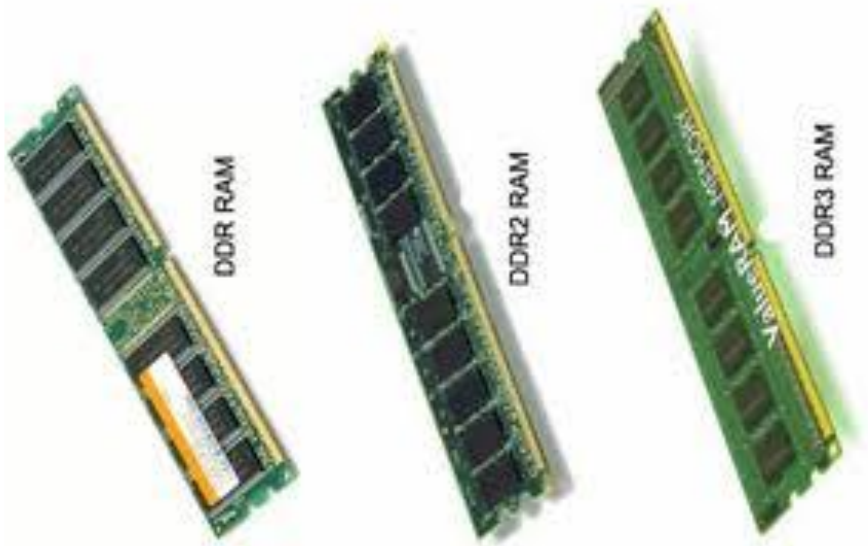


- **Solution**
    - Refresh the charge periodically
    - While refreshing, cannot read/write

Synchronous dynamic random access memory (SDRAM)

SDRAM is dynamic random access memory (DRAM) with an interface synchronous with the system bus carrying data between the CPU and the memory controller hub. SDRAM has a rapidly responding synchronous interface, which is in sync with the system bus. SDRAM waits for the clock signal before it responds to control inputs.

SDRAM preceded DDR.DDR stands for Double Data Rate. DDR RAM enables two data transfer per clock cycle. DDR, DDR2 and DDR3 are different versions based on the same DDR technology. The design of all three RAMs is based on Synchronous Dynamic Random Access Memory. DDR is also known as double pumped, dual-pumped and double transition process.



There are three significant characteristics differentiating SDRAM and DDR:

- The main difference is the amount of data transmitted with each cycle, not the speed.
- SDRAM sends signals once per clock cycle. DDR transfers data twice per clock cycle. (Both SDRAM and DDR use the same frequencies.)
- SDRAM uses one edge of the clock. DDR uses both edges of the clock.

Non-volatile memory

This is a type of computer memory that has the capability to hold saved data even if the power is turned off.

Eg: ROM, Hard disk etc

ROM (Read Only Memory)

ROM remains its contents even when the computer is turned off. ROM stores essential programs such as the program that boots the computer.

Types of ROM

MROM (Mask ROM)

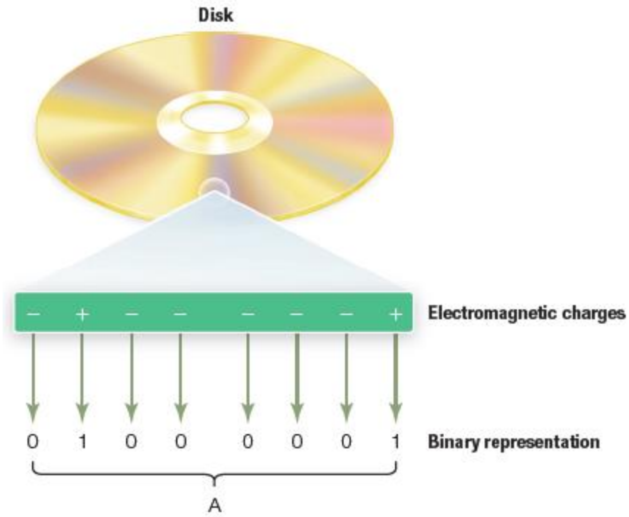
PROM (Programmable ROM)

EPROM (Erasable PROM)

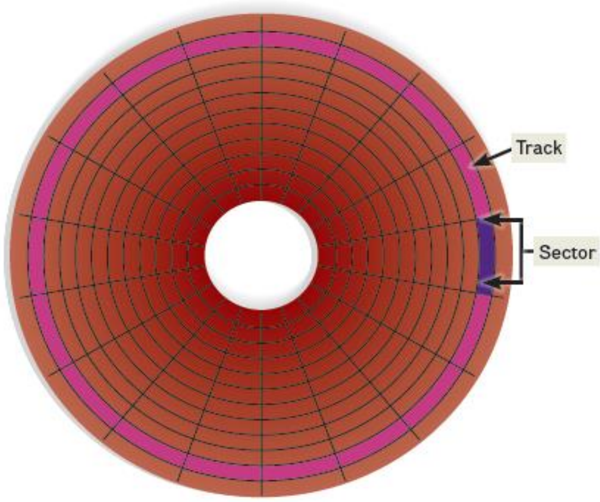
EEPROM (Electrically Erasable PROM)

Hard disk

Hard disks save files by altering the magnetic charges of the disk's surface to represent 1s and 0s. Hard disks retrieve data and programs by reading these charges from the magnetic disk. Characters are represented by positive (+) and negative (-) charges using the ASCII, EBCDIC, or Unicode binary codes. For example, the letter A would require a series of 8 charges. Density refers to how tightly these charges can be packed next to one another on the disk.



How charges on a disk surface store the letter A

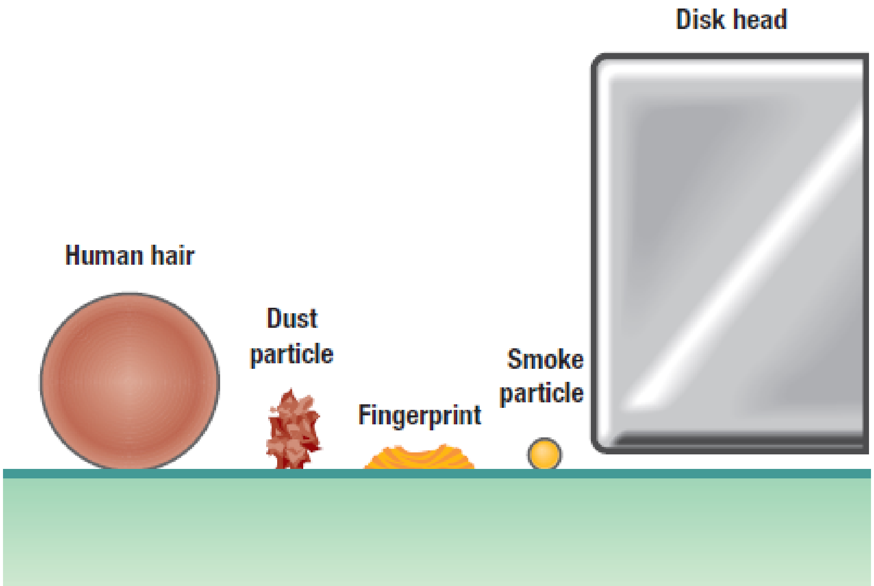


Tracks and sectors

Hard disks use inflexible metallic platters that are stacked one on top of another. Hard disks store and organize files using tracks, sectors, and cylinders. Tracks are rings of concentric circles on the platter. Each track is divided into invisible wedge-shaped sections called sectors. A cylinder runs through each track of a stack of platters.

Cylinders are necessary to differentiate files stored on the same track and sector of different platters. When a hard disk is formatted, tracks, sectors, and cylinders are assigned.

Hard disks are sensitive instruments. Their read/write heads ride on a cushion of air about 0.000001 inch thick. It is so thin that a smoke particle, fingerprint, dust, or human hair could cause what is known as a head crash.



A head crash occurs when a read-write head makes contact with the hard disk's surface or with particles on its surface. A head crash is a disaster for a hard disk. The disk surface is scratched, and some or all of the data is destroyed. At one time, head crashes were commonplace. Now, fortunately, they are rare.

There are two basic types of hard disks: internal and external.

References

Teachers' Guide 2017

T. O'Leary, L. O'Leary and D. O'Leary, Computing Essentials 2017: Making IT work for you, 26th Edition, pp. 116 - 128, McGraw-Hill, 2017

https://www.geeksforgeeks.org/memory-hierarchy-design-and-its-characteristics/

http://www.differencebetween.info/difference-between-ddr-ddr2-and-ddr3-ram