

Week 06: Memory

Session 1: Introduction to Memory, Memory access method

Subject: CHMA - 20CS32P.

Class: III Semester.

Course: Computer Science and Engineering.

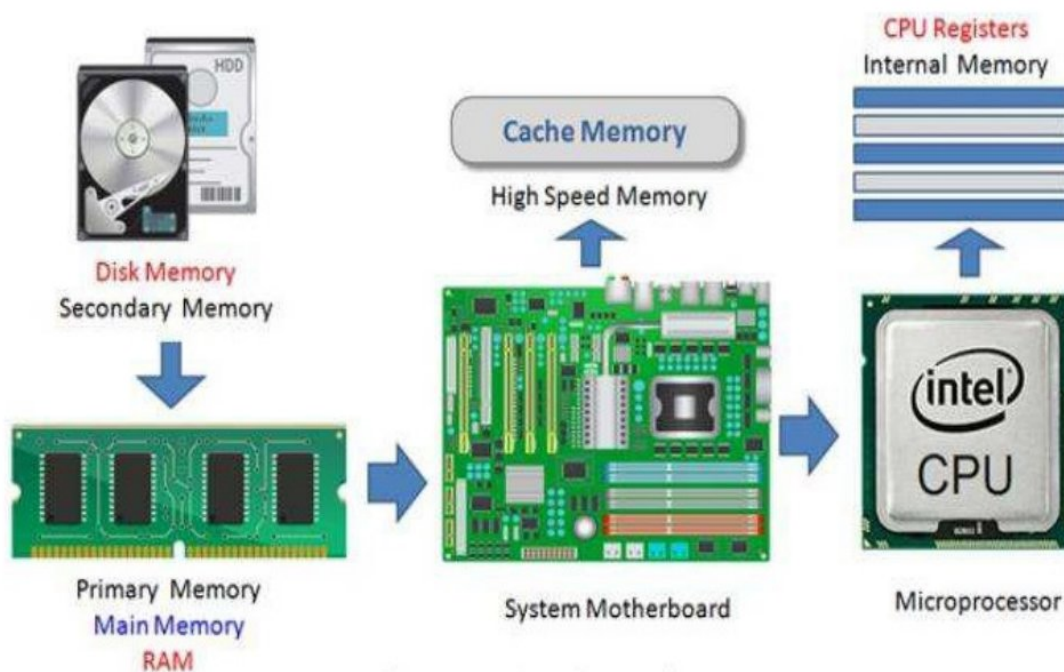
Memory

A memory is just like a human brain. It is used to store data and instructions. Computer memory is the storage space in the computer, where data is to be processed and instructions required for processing is stored. The memory is divided into large number of small parts called cells. Each location or cell has a unique address, which varies from zero to memory size minus one.

For example: If the computer has 64k words, then this memory unit has $64 \times 1024 = 65536$ memory locations. The address of these locations varies from 0 to 65535.

Memory is primarily of three types -

1. Cache Memory
2. Primary Memory/Main Memory
3. Secondary Memory



Memory Units:

Computer Memory Measurement Units

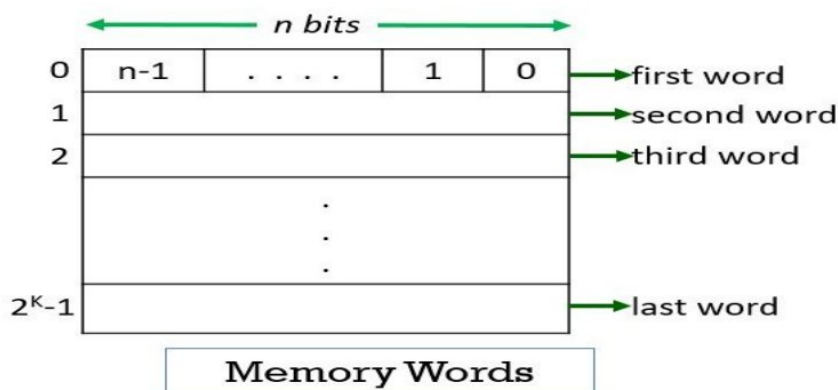
SYMBOL	FULL FORM	QUANTITY
1 BIT	BINARY DIGIT	1 CELL, BINARY 0 OR 1
4 BITS	NIBBLE	1/2 BYTE
8 BITS	BYTE	1 BYTE
1024 BYTE	KILOBYTE	1 KILOBYTE
1024 KILOBYTE	MEGABYTE	1 MEGABYTE
1024 MEGABYTE	GIGABYTE	1 GIGABYTE
1024 GIGABYTE	TERABYTE	1 TERABYTE
1024 TERABYTE	PETABYTE	1 PETABYTE
1024 PETABYTE	HEXABYTE	1 HEXABYTE
1024 HEXABYTE	ZEETABYTE	1 ZEETABYTE

Memory locations and address space

Memory locations and addresses determine how the computer's memory is organized so that the user can efficiently store or retrieve information from the computer. The computer's memory is made of a silicon chip which has millions of storage cells, where each storage cell is capable to store a bit of information whose value is either 0 or 1.

But the fact is, computer memory holds instructions and data. And a single bit is very small to hold this information so bits are rarely used individually. As a solution to this, the bits are grouped in fixed sizes of 'n' bits. The memory of the computer is organized in such a way that the group of these 'n' bits can be stored and retrieved easily by the computer in a single operation.

The group of 'n' bits is termed as word where 'n' is termed as the word length. The word length of the computer has evolved from 8, 16, 24, 32 to 64 bits. General-purpose computers nowadays have 32 to 64 bits. The group of 8 bits is called a byte.



Now, whenever we want to store any instruction or data may it be of a byte or a word we have to access a memory location. To access the memory location either we must know the memory location by its unique name or it is required to provide a unique address to each memory location.

The memory locations are addressed from 0 to $2K-1$ i.e. a memory has $2K$ addressable locations. And thus the address space of the computer has $2K$ addresses. Let us try some suitable values for K .

Memory Access Method

Data need to be accessed from the memory for various purposes.

There are several methods to access memory as listed below:

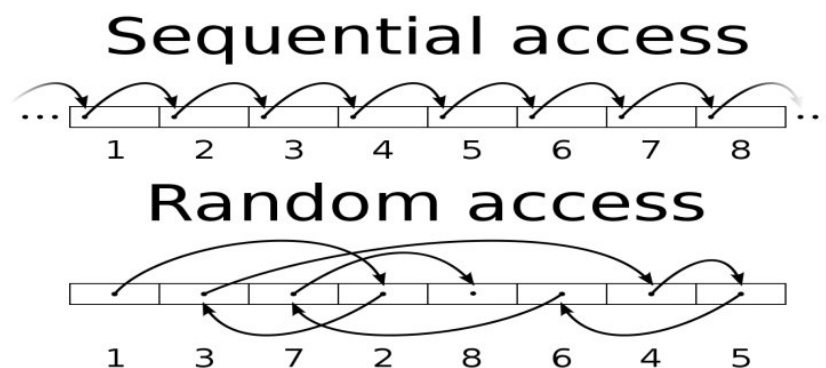
1. Sequential access
2. Direct access
3. Random access
4. Associative access

Sequential Access Method

In sequential memory access method, the memory is accessed in linear sequential way. The time to access data in this type of method depends on the location of the data.

Example of sequential access; Magnetic disks, magnetic tapes, optical memory use sequential access method.

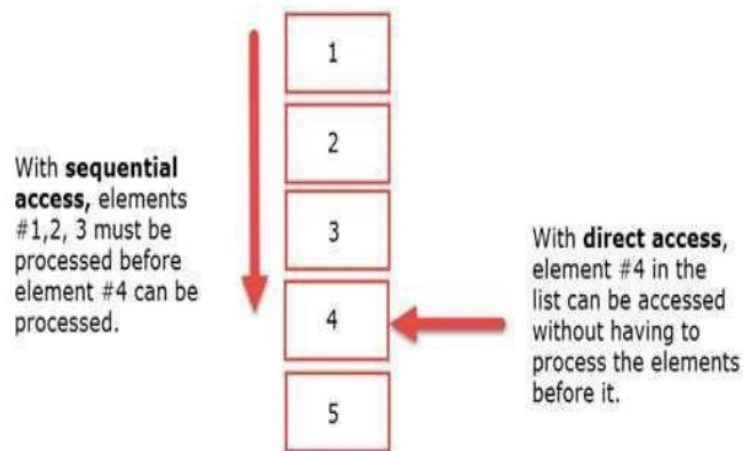
Random Access Method



In random access method, data from any location of the memory can be accessed randomly. The access to any location is not related with its physical location and is independent of other locations. There is a separate access mechanism for each location.

Example of random access; Semiconductor memories like RAM, ROM use random access method.

Direct Access Method



Direct access method can be seen as combination of sequential access method and random access method.

Magnetic hard disks contain many rotating storage tracks. Here each track has its own read or writes head and the tracks can be accessed randomly. But access within each track is sequential.

Example of direct access: Memory devices such as magnetic hard disks.

Associative Access Method

Associative access method is a special type of random access method. It enables comparison of desired bit locations within a word for a specific match and to do this for all words simultaneously. Thus based on portion of word's content, word is retrieved rather than its address.

Example of associative access: Cache memory uses associative access method.

Week 06: Memory

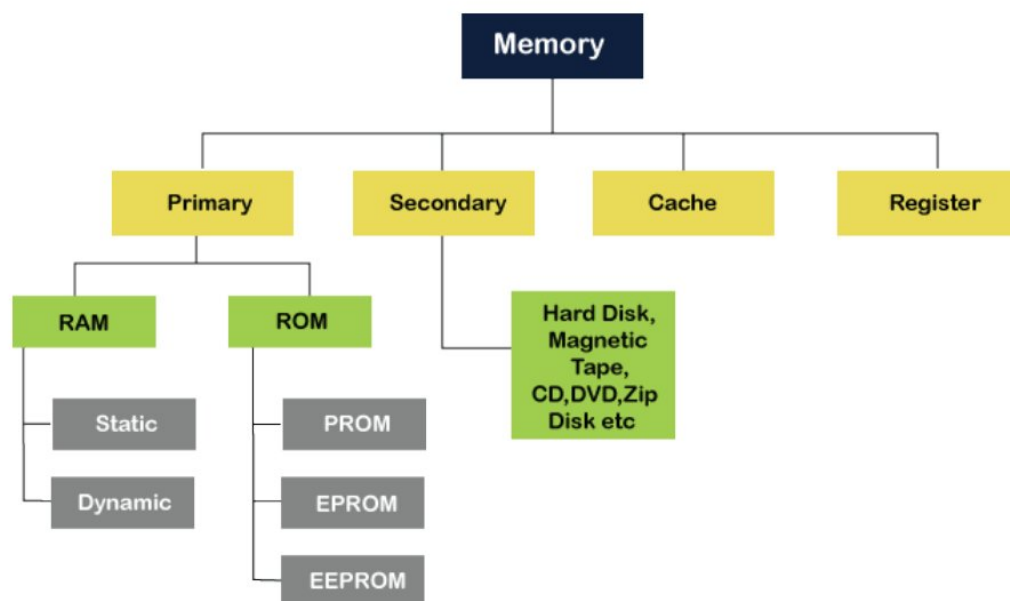
Session 2: Memory classification, Primary Memory, RAM and ROM

Subject: CHMA - 20CS32P.

Class: III Semester.

Course: Computer Science and Engineering.

Memory Classification



Primary Memory (Main Memory)

Primary memory is also known as the computer system's main memory that communicates directly within the CPU, Auxiliary memory and the Cache memory.

Main memory is used to keep programs or data when the processor is active to use them. When a program or data is activated to execute, the processor first loads instructions or programs from secondary memory into main memory, and then the processor starts execution. Accessing or executing of data from primary memory is faster because it has a cache or register memory that provides faster response, and it is located closer to the CPU. The primary memory is volatile, which means the data in memory can be lost if it is not saved when a power failure occurs. It is costlier than secondary memory, and the main memory capacity is limited as compared to secondary memory.

The primary memory is further divided into two parts:

1. RAM (Random Access Memory)
2. ROM (Read Only Memory)

Random Access Memory (RAM)



Random Access Memory (RAM) is one of the faster types of main memory accessed directly by the CPU. It is the hardware in a computer device to temporarily store data, programs or program results. It is used to read/write data in memory until the machine is working. It is volatile, which means if a power failure occurs or the computer is turned off, the information stored in RAM will be lost. All data stored in computer memory can be read or accessed randomly at any time.

There are two types of RAM:

- SRAM
- DRAM

SRAM: Static Random-Access Memory is a type of RAM used to store static data in the memory. It means to store data in SRAM remains active as long as the computer system has a power supply. However, data is lost in SRAM when power failures have occurred.

Characteristics of Static Ram

1. It does not require refreshing.
2. It is faster than DRAM
3. It is expensive.
4. High power consumption
5. Longer life
6. Large size
7. Uses as a cache memory

DRAM: Dynamic Random-Access Memory is a type of RAM that is used for the dynamic storage of data in RAM. In DRAM, each cell carries one-bit information. The cell is made up of two parts: a capacitor and a transistor. The size of the capacitor and the transistor is so small, requiring millions of them to store on a single chip. Hence, a DRAM chip can hold more data than an SRAM chip of the same size. However, the capacitor needs to be continuously refreshed to retain information because DRAM is volatile. If the power is switched off, the data store in memory is lost.

Characteristics of DRAM

1. It requires continuously refreshed to retain the data.
2. It is slower than SRAM
3. It holds a large amount of data
4. It is the combination of capacitor and transistor
5. It is less expensive as compared to SRAM
6. Less power consumption

Read-Only Memory (ROM)



ROM is a memory device or storage medium that is used to permanently store information inside a chip. It is a read-only memory that can only read stored information, data or programs, but we cannot write or modify anything.

A ROM contains some important instructions or program data that are required to start or boot a computer. It is a non-volatile memory; it means that the stored information cannot be lost even when the power is turned off or the system is shut down.

Types of ROM

There are five types of Read Only Memory:

MROM (Masked Read Only Memory): MROM is the oldest type of read-only memory whose program or data is pre-configured by the integrated circuit manufacture at the time of manufacturing. Therefore, a program or instruction stored within the MROM chip cannot be changed by the user.

PROM (Programmable Read Only Memory): It is a type of digital read-only memory, in which the user can write any type of information or program only once. It means it is the empty PROM chip in which the user can write the desired content or program only once using the special PROM programmer or PROM burner device; after that, the data or instruction cannot be changed or erased.

EPROM (Erasable and Programmable Read Only Memory): It is the type of read only memory in which stored data can be erased and re-programmed only once in the EPROM memory. It is a non-volatile memo chip that holds data when there is no power supply and can also store data for a minimum of 10 to 20 years. In EPROM, if we want to erase any stored data and re-programmed it, first, we need to pass the ultraviolet light for 40 minutes to erase the data; after that, the data is re-created in EPROM.

EEPROM (Electrically Erasable and Programmable Read Only Memory): The EEROM is an electrically erasable and programmable read only memory used to erase stored data using a high voltage electrical charge and re-programmed it. It is also a non-volatile memory whose data cannot be erased or lost; even the power is turned off. In EEPROM, the stored data can be erased and reprogrammed up to 10 thousand times, and the data erase one byte at a time.

Flash ROM: Flash memory is a non-volatile storage memory chip that can be written or programmed in small units called Block or Sector. Flash Memory is an EEPROM form of computer memory, and the contents or data cannot be lost when the power source is turned off. It is also used to transfer data between the computer and digital devices.

Week 06: Memory

Session 3: Secondary Memory, Cache Memory and Memory Modules

Subject: CHMA - 20CS32P.

Class: III Semester.

Course: Computer Science and Engineering.

Secondary Memory (Auxiliary memory)

Secondary memory or Auxiliary memory is a permanent storage space to hold a large amount of data. Secondary memory is also known as external memory that representing the various storage media (hard drives, USB, CDs, flash drives and DVDs) on which the computer data and program can be saved on a long term basis. However, it is cheaper and slower than the main memory. Unlike primary memory, secondary memory cannot be accessed directly by the CPU. Instead of that, secondary memory data is first loaded into the RAM (Random Access Memory) and then sent to the processor to read and update the data. Secondary memory devices also include magnetic disks like hard disk and floppy disks, an optical disk such as CDs and CDROMs, and magnetic tapes.

Features of Secondary Memory

- Its speed is slower than the primary/ main memory.
- Store data cannot be lost due to non-volatile nature.
- It can store large collections of different types, such as audio, video, pictures, text, software, etc.
- All the stored data in a secondary memory cannot be lost because it is a permanent storage area; even the power is turned off.
- It has various optical and magnetic memories to store data.

Types of Secondary Memory

The following are the types of secondary memory devices:

Hard Disk



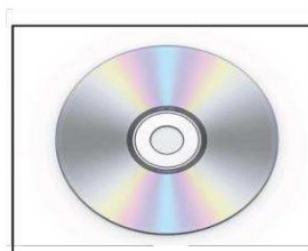
A hard disk is a computer's permanent storage device. It is a non-volatile disk that permanently stores data, programs, and files, and cannot lose store data when the computer's power source is switched off. Typically, it is located internally on computer's motherboard that stores and retrieves data using one or more rigid fast rotating disk platters inside an air-sealed casing. It is a large storage device, found on every computer or laptop for permanently storing installed software, music, text documentation, videos, operating system, and data until the user did not delete.

Floppy Disk



A floppy disk is a secondary storage system that consisting of thin, flexible magnetic coating disks for holding electronic data such as computer files. It is also known as Floppy Diskette that comes in three sizes like 8 inches, 5.5 inches and 3.5 inches. The stored data of a floppy disk can be accessed through the floppy disk drive. Furthermore, it is the only way through a new program installed on a computer or backup of the information. However, it is the oldest type of portable storage device, which can store data up to 1.44 MB. Since most programs were larger, that required multiple floppy diskettes to store large amounts of data. Therefore, it is not used due to very low memory storage.

CD (Compact Disc):

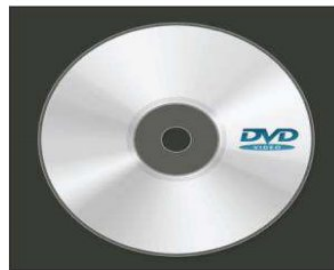


A CD is an optical disk storage device, stands for Compact Disc. It is a storage device used to store various data types like audio, videos, files, OS, Back-Up file, and any other information useful to a computer. The CD has a width of 1.2 mm and 12 cm in height, which can store approximately 783 MB of data size. It uses laser light to read and write data from the CDs.

Types of CDs

1. **CD-ROM (Compact Disc Read Only Memory):** It is mainly used for bulk size mass like audio CDs, software and computer games at the time of manufacture. Users can only read data, text, music, videos from the disc, but they cannot modify or burnt it.
2. **CD-R (Compact Disc Recordable):** The type of Compact Disc used to write once by the user; after that, it cannot be modified or erased.
3. **CD-RW (Compact Disc Rewritable):** It is a rewritable CD disc, often used to write or delete the stored data.

DVD Drive/Disc:



DVD is an optical disc storage device, stands for Digital Video Display or Digital Versatile Disc. It has the same size as a CD but can store a larger amount of data than a compact disc. It was developed in 1995 by Sony, Panasonic, Toshiba and Philips four electronics companies. DVD drives are divided into three types, such as DVD ROM (Read Only Memory), DVD R (Recordable) and DVD RW (Rewritable or Erasable). It can store multiple data formats like audio, videos, images, software, operating system, etc. The storing capacity of data in DVD is 4.7 GB to 17 GB.

Blu-Ray Disc (BD):



Blu-Ray is an Optical disc storage device used to store a large amount of data or high definition of video recording and playing other media files. It uses laser technology to read the

stored data of the Blu-ray Disk. It can store more data at a greater density as compared to CD/DVD.

For example, compact discs allow us to store 700 MB of data, and in DVDs, it provides up to 8 GB of storage capacity, while Blu-ray Discs provide 25 GB/50GB of space to store data.

Pen Drive



A pen drive is a portable device used to permanently store data and is also known as a USB flash drive. It is commonly used to store and transfer the data connected to a computer using a USB port. It does not have any moveable part to store the data; it uses an integrated circuit chip that stores the data. It allows the users to store and transfer data like audio, videos, images, etc. from one computer to any USB pen drive. The storing capacity of pen drives from 64 MB to 128 GB or more.

Cache Memory

It is a small-sized chip-based computer memory that lies between the CPU and the main memory. It is a faster, high performance and temporary memory to enhance the performance of the CPU. It stores all the data and instructions that are often used by computer CPUs.

Cache memory is sometimes called CPU (central processing unit) memory because it is typically integrated directly into the CPU chip or placed on a separate chip that has a separate bus interconnect with the CPU. Therefore, it is more accessible to the processor, and able to increase efficiency, because it's physically close to the processor.

Different cache levels are:

- **L1 Cache:** The L1 cache is also known as the onboard, internal, or primary cache. It is built with the help of the CPU. Its speed is very high, and the size of the L1 cache varies from 8 KB to 128 KB.
- **L2 Cache:** It is also known as external or secondary cache, which requires fast access time to store temporary data. It is built into a separate chip in a motherboard, not built into the CPU like the L1 level. The size of the L2 cache may be 128 KB to 1 MB.
- **L3 Cache:** L3 cache levels are generally used with high performance and capacity of the computer. It is built into a motherboard. Its speed is very slow, and the maximum size up to 8 MB.

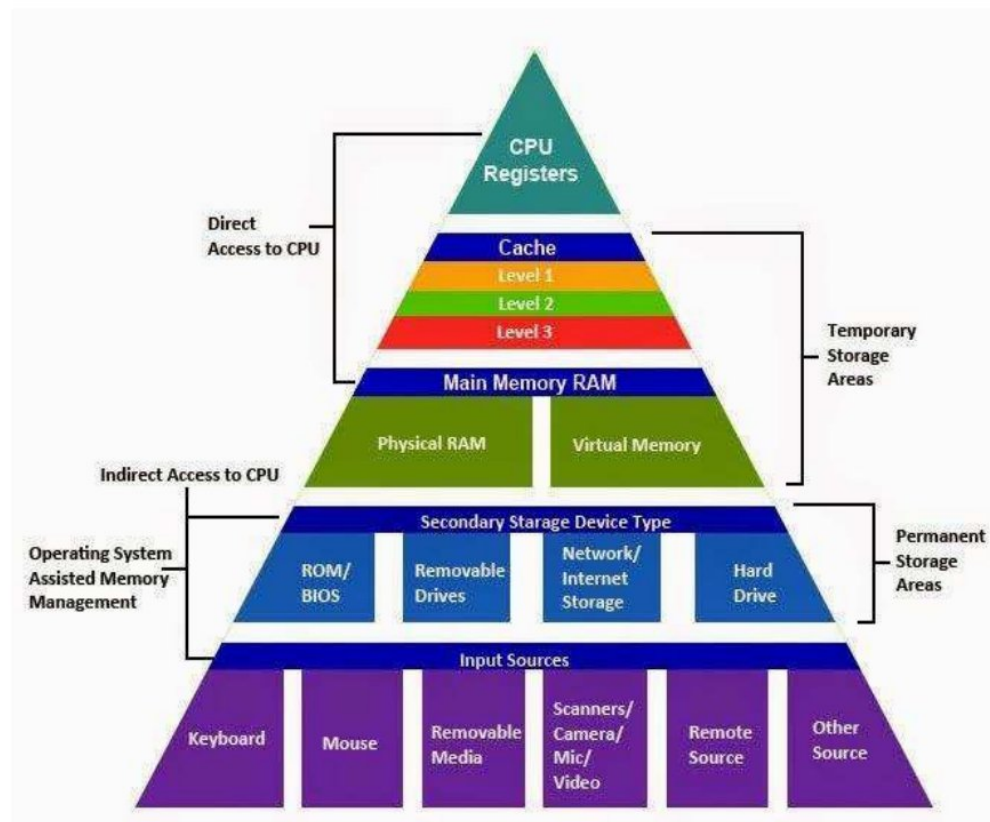
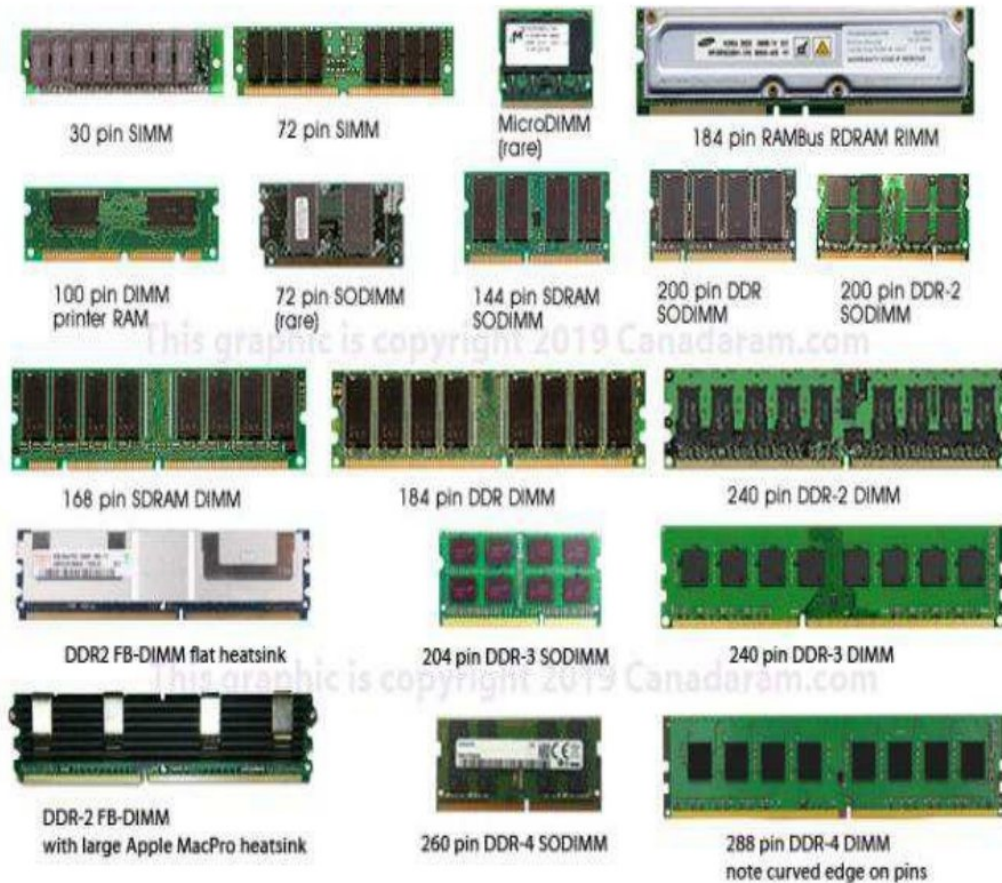


Fig showing different levels of cache.

Memory modules

A broad term used to refer to a series of dynamic random access memory integrated circuits modules mounted on a printed circuit board and designed for use in personal computer, workstations and servers. A memory module is the computer part that holds the Random Access Memory (RAM), measured in megabytes or gigabytes. A megabyte being 1,000,000 bytes of storage and Gigabyte being 10,000 000 000 bytes of storage. All of these bytes are placed in small circuits on a memory module stick making it easy for anyone to replace or add more RAM as the computer needs it.

Types of Memory Module



- DIP (Dual In-line Package)
- SIPP (Single In-Line Pin Package)
- SIMM (single in-line memory module)
- DIMM (Dual In-Line Memory Module)
- SO-DIMM (Small Outline Dual In-Line Memory Module)
- SDR SDRAM (Single Data Rate Synchronous dynamic random-access memory)
- DDR(Double Data Rate), GORZ, DDR3 SDRAM
- Trans Flash Memory Module