Unit-4
Memory

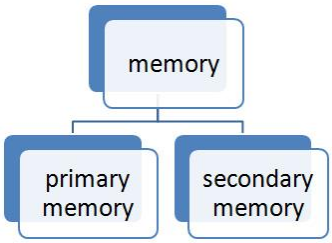
A memory is just like a human brain to store data and instruction. It is the storage space in computer where data is to be processed and instructions required for processing are stored.

Types of Memory:

Computer memory is broadly divided into two groups and they are:

- **Primary memory**
- **Secondary memory**

The diagrammatic representation of the classification of computer memory is shown below:



○ Primary Memory or Main Memory:

Primary memory is the only type of memory which is directly accessed by the CPU. The CPU continuously reads instructions stored in the primary memory and executes them. Any data that has to be operated by the CPU is also stored. The information is transferred to various locations through the BUS**. Primary memories are of two types**. They are:

- **RAM, or Random Access Memory**
- **ROM, or Read-Only Memory**

1) RAM

RAM stands for random access memory, and the name suggest that data stored in random access memory can be accessed in any random order.The most important things to understand about RAM are that RAM memory is very fast, it can be written to as well as read, it is volatile (so all data stored in RAM memory is lost when it loses power) and, finally, it is very expensive compared to all types of secondary memory in terms of cost per gigabyte. It is because of the relative high cost of RAM compared to secondary memory types that most computer systems use both primary and secondary memory.Data that is required for imminent processing is moved to RAM where it can be accessed and modified very quickly, so that the CPU is not kept waiting. When the data is no longer required it is shunted out to slower but cheaper secondary memory, and the RAM space that has been freed up is filled with the next chunk of data that is about to be used.

Types of RAM:

RAM is of two types

- **DRAM**: DRAM stands for Dynamic RAM, and it is the most common type of RAM used in computers. The oldest type is known as single data rate (SDR) DRAM, but newer computers use faster dual data rate (DDR) DRAM. DDR comes in several versions including DDR2 , DDR3, and DDR4, which offer better performance and are more energy efficient than DDR. However different versions are incompatible, so it is not possible to mix DDR2 with DDR3 DRAM in a computer system. DRAM consists of a transistor and a capacitor in each cell.
- **SRAM**: SRAM stands for Static RAM, and it is a particular type of RAM which is faster than DRAM, but more expensive and bulker, having six transistors in each cell. For those reasons SRAM is generally only used as a data cache within a CPU itself or as RAM in very high-end server systems. A small SRAM cache of the most imminently-needed data can result in significant speed improvements in a system

The key differences between DRAM and SRAM is that SRAM is faster than DRAM - perhaps two to three times faster - but more expensive and bulkier. SRAM is usually available in megabytes, while DRAM is purchased in gigabytes.

DRAM uses more energy than SRAM because it constantly needs to be refreshed to maintain data integrity, while SRAM - though volatile – does not need constant refreshing when it is powered up.

2) ROM

ROM stands for read-only memory, and the name suggest that while data can be read from this type of computer memory, data cannot normally be written to it. It is a very fast type of computer memory which is usually installed close to the CPU on the motherboard.

ROM is a type of non-volatile memory, which means that the data stored in ROM persists in the memory even when it receives no power – for example when the computer is turned off. In that sense it is similar to secondary memory, which is used for long term storage.

When a computer is turned on, the CPU can begin reading information stored in ROM without the need for drivers or other complex software to help it communicate. The ROM usually contains "bootstrap code" which is the basic set of instructions a computer needs to carry out to become aware of the operating system stored in secondary memory, and to load parts of the operating system into primary memory so that it can start up and become ready to use.

ROM is also used in simpler electronic devices to store firmware which runs as soon as the device is switched on.

Types of ROM:

Following are the various types of ROM –

- **PROM (Programmable Read Only Memory)**

  PROM is read-only memory that can be modified only once by a user. The user buys a blank PROM and enters the desired contents using a PROM programmer. Inside the PROM chip there are small fuses which are burnt open during programming. It can be programmed only once and is not erasable.

- **EPROM (Erasable and Programmable Read Only Memory)**

  The EPROM can be erased by exposing it to ultra-violet light for a duration of upto 40 minutes. Usually, an EPROM eraser achieves this function. During programming an electrical charge is trapped in an insulated gate region. The charge is retained for more than ten years because the charge has no leakage path.

- **EEPROM (Electrically Erasable and Programmable Read Only Memory)**

  The EEPROM is programmed and erased electrically. It can be erased and reprogrammed about ten thousand times. Both erasing and programming take about 4 to 10 ms (millisecond). In EEPROM, any location can be selectively erased and programmed. EEPROMs can be erased one byte at a time, rather than erasing the entire chip. Hence, the process of re-programming is flexible but slow.

Differences between ROM and RAM

ROM:

- Non-volatile
- Fast to read
- Usually used in small quantities
- Cannot be written to quickly
- Used to store boot instructions or firmware
- Relatively expensive per megabyte stored compared to RAM

RAM:

- Volatile
- Fast to read and write
- Used as system memory to store data (including program code) that the CPU needs to process imminently
- Relatively cheap per megabyte stored compared to ROM, but relatively expensive compared to secondary memory

Characteristic of Primary Memory:

- The computer can't run without primary memory
- It is known as the main memory.
- You can lose data in case power is switched off
- It is also known as volatile memory
- It is a working memory of the computer.
- Primary memory is faster compares to secondary memory.

○ Secondary Memory:

Secondary memory is not directly accessible by the CPU. It is external and permanent in nature. Primary memory has limited storage capacity and is volatile. Secondary memory overcome this limitation by providing permanent storage of data and in bulk quantity. It is also termed as external memory and refers to the various storage media on which a computer can store data and programs.

The Secondary storage media can be fixed or removable. Fixed Storage media is an internal storage medium like hard disk that is fixed inside the computer. Storage medium that are portable and can be taken outside the computer are termed as removable storage media.

Types of Secondary Memory:

There are the following main types of secondary memory-

○ Floppy Disk :

A floppy disk is a magnetic storage medium for computer systems. The floppy disk is composed of a thin, flexible magnetic disk sealed in a square plastic carrier. In order to read and write data from a floppy disk, a computer system must have a floppy disk drive (FDD). A floppy disk is also referred to simply as a floppy. Since the early days of personal computing, floppy disks were widely used to distribute software, transfer files, and create back-up copies of data. When hard drives were still very expensive, floppy disks were also used to store the operating system of a computer.

A number of different types of floppy disks have been developed. The size of the floppy got smaller, and the storage capacity increased. However, in the 1990s, other media, including hard disk drives, optical drives started to replace floppy disks as the primary storage medium.

The first floppy disks that came on the market were 8 inches (200 mm) in diameter. The disk was protected by a flexible plastic jacket. An 8-inch disk back in the late 1970s could store about 1 MB of data. This was quickly followed by a smaller version of the same design, the 5.25-inch (133 mm) floppy, which could store about the same amount of information using higher-density media and recording techniques.

In the early 1980s, the 3.5-inch (90 mm) floppy, or micro floppy, came on the market, and this type became the dominant storage medium for personal computers for many

years. Each of these floppy disks required a different type of floppy disk drive. These were typically built into the computer case itself.

Floppy disks were quite vulnerable. The disk medium was very sensitive to dust, moisture, and heat. The flexible plastic carrier was also not very sturdy. The hard plastic case of the 3.5-inch floppy presented a substantial improvement in this respect. The most common format of this floppy became the double-sided, high-density 1.44 MB disk drive.

o Hard Disk:

A hard disk consists of one or more circular disks called platters which are mounted on a common spindle. Each surface of a platter is coated with a magnetic material. Both surfaces of each disk are capable of storing data except the top and bottom disk where only the inner surface is used. The information is recorded on the surface of the rotating disk by magnetic read/write heads. These heads are joined to a common arm known as access arm.

o CD :

Compact Discs (CDs) are one of the oldest forms of external memory.It was made initially to replace cassette tapes until it grew and led to several types of CDs being made. The differences between each type lie in the read and write restrictions for users.

- **CD Read-Only Memory (CD-ROM) —** This type of CD doesn't allow users to erase nor write over whatever is in the disc.
- **CD-Recordable (CD-R) —** A CD like this gives users the ability to write on the CD only once but has an unlimited number of reads available.
- **CD Re-Writable (CD-RW) —** With a CD like this, you can erase and write over it for up to 1000 times.

o DVD

Digital Versatile Discs (DVDs) are much like CDs in that it also uses laser light to store and retrieve data. This particular way of storing or retrieving data is characteristic of optical storage devices.

Although a CD and a DVD are similar in many ways, they're also different in just as many aspects too. One difference is their storage capacity.

In comparison to CDs, DVDs have much *higher* storage capacity. .

Usually, people use DVDs to store movies and videos, while CDs are more commonly used to store music.

- **USB flash Drives :**

  A USB flash drive -- also known as a USB stick, USB thumb drive or pen drive -- is a plug-and-play portable storage.
  USB flash drives are often used for storage, data back-up and transfer of computer files. Compared with floppy disks or CDs, they are smaller, faster, have significantly more capacity, and are more durable due to a lack of moving parts. It consists of a small printed circuit board carrying the circuit elements and a USB connector, insulated electrically and protected inside a plastic, metal, or rubberized case, which can be carried in a pocket or on a key chain. The USB connector may be protected by a removable cap or by retracting into the body of the drive.
  Most flash drives use a standard type-A USB connection allowing connection with a port on a personal computer, but drives for other interfaces also exist. USB flash drives draw power from the computer via the USB connection. Some devices combine the functionality of a portable media player with USB flash storage; they require a battery only when used to play music on the go.

Characteristic of Secondary Memory:

- These are magnetic and optical memories
- Secondary memory is known as a backup memory
- It is a non-volatile type of memory
- Data is stored permanently even when the power of the computer is switched off
- It helps store data in a computer
- The machine can run without secondary memory
- Slower than primary memory

Difference between Primary Memory and Secondary Memory:

Parameter	Primary Memory	Secondary Memory
Nature	The primary memory is categorized as volatile & nonvolatile memories.	The secondary memory is always a non-volatile memory.
Alias	These memories are also called internal memory.	Secondary memory is known as a Backup memory or Additional memory or Auxiliary memory.

Access	Data is directly accessed by the processing unit.	Data cannot be accessed directly by the processor. It is first copied from secondary memory to primary memory. Only then CPU can access it.
Formation	It's a volatile memory meaning data cannot be retained in case of power failure.	It's a non-volatile memory so that that data can be retained even after power failure.
Storage	It holds data or information that is currently being used by the processing unit. Capacity is usually in 16 to 32 GB	It stores a substantial amount of data and information. Capacity is generally from 200GB to terabytes.
Accesses	Primary memory can be accessed by the data bus.	Secondary memory is accessed by I/O channels.
Expense	Primary memory is costlier than secondary memory.	Secondary memory is cheaper than primary memory.

Cache Memory

Cache memory is a special very high-speed memory that acts as a buffer between RAM and the CPU. It holds frequently requested data and instructions so that they are immediately available to the CPU when needed.It is costlier than main memory or disk memory but economical than CPU registers.

Cache memory is used to reduce the average time to access data from the Main memory. The cache is a smaller and faster memory which stores copies of the data from frequently used main memory locations. There are various different independent caches in a CPU, which store instructions and data.

Functions of Cache Memory:

The basic purpose of cache memory is to store program instructions that are frequently re-referenced by software during operation. Fast access to these instructions increases the overall speed of the software program.

The main function of cache memory is to speed up the working mechanism of computer.

Advantages:

- Cache memory is faster than main memory.
- It consumes less access time as compared to main memory.
- It stores the program that can be executed within a short period of time.
- It stores data for temporary use.

Disadvantages

- Cache memory has limited capacity.
- It is very expensive.

BUS

Inside computers, there are many internal components. In order for these components to communicate with each other, they make use of wires that are known as a '**bus**'.
A **bus** is a **common pathway** through which information flows from one computer component to another. This pathway is used for communication purpose and it is established between two or more computer components.

Functions of Buses in Computers:

1. Data sharing –

All types of buses found in a computer transfer data between the computer peripherals connected to it.The buses transfer or send data either in the serial or parallel method of data transfer. This allows for the exchange of 1, 2, 4 or even 8 bytes of data at a time. (A byte is a group of 8 bits). Buses are classified depending on how many bits they can move at the same time, which means that we have 8-bit, 16-bit, 32-bit or even 64-bit buses.

2. Addressing –

A bus has address lines, which match those of the processor. This allows data to be sent to or from specific memory locations.

3. Power –

A bus supplies power to various peripherals connected to it.

4. Timing –

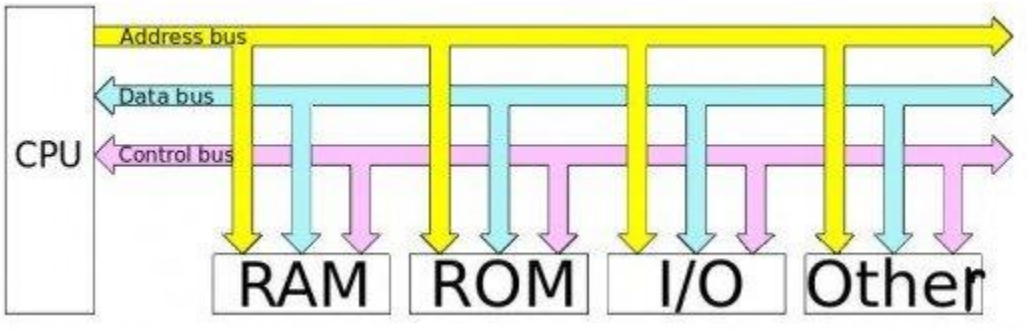
The bus provides a system clock signal to synchronize the peripherals attached to it with the rest of the system.

How Does Computer Bus Work?

A bus transfers electrical signals from one place to another. An actual bus appears as an endless amount of etched copper circuits on the motherboard's surface. The bus is connected to the CPU through the Bus Interface Unit.
Data travels between the CPU and memory along the data bus. The location (address) of that data is carried along the address bus. A clock signal which keeps everything in synch travels along the control bus.
The clock acts like a traffic light for all the PC's components; the "green light" goes on with each clock tick. A PC's clock can "tick" anywhere from 20 to 65 million times per second, which makes it seem like a computer is really fast. But since each task (such as saving a file) is made

up of several programmed instructions, and each of those instructions takes several clock cycles to carry out, a person sometimes has to sit and wait for the computer to catch up.

Different Types of Computer Buses

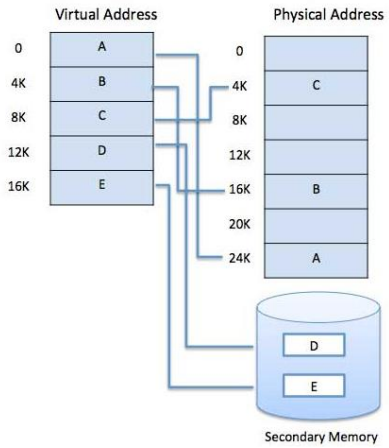


Data Bus: The data bus allows data to travel back and forth between the microprocessor (CPU) and memory (RAM).
Address Bus: The address bus carries information about the location of data in memory.
Control Bus : The control bus carries the control signals that make sure everything is flowing smoothly from place to place

.

Virtual Memory Concept

- In the most of the computer system, the physical main memory is not as large as address space of the processor.
- Suppose user tries to run a program.
- If the program run by the user does not completely fit into the main memory then the parts of its currently being executed are stored in main memory and remaining portion is stored in secondary storage device such as HDD.
- When a new part of program is to be brought into main memory for execution and if the memory is full, it must replace another part which is already is in main memory.
- As this secondary memory is not actually part of system memory, so for CPU, **secondary memory is considered as Virtual Memory.**
- Virtual memory is a memory management technique that is implemented using both hardware and software.
- It maps memory addresses used by a program, called virtual addresses, into physical addresses in computer memory.



Benefits of Virtual Memory:

- Large programs can be written, as virtual space available is huge compared to physical memory.
- Less I/O required, leads to faster and easy swapping of processes.
- More physical memory available, as programs are stored on virtual memory, so they occupy very less space on actual physical memory.

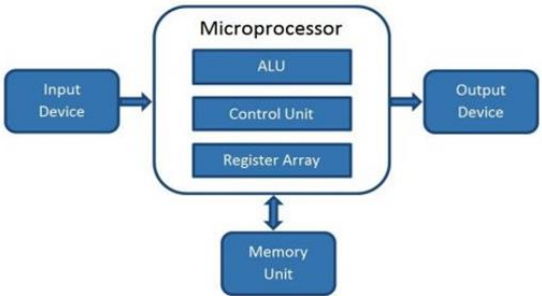
Disadvantages of Virtual Memory:

- Applications run slower if the system is using virtual memory.
- It takes more time to switch between applications.
- Less hard drive space for your use.
- It reduces system stability.

Microprocessor

A Microprocessor is an important part of a computer architecture without which you will not be able to perform anything on your computer. It is a programmable device that takes in input perform some arithmetic and logical operations over it and produce desired output. In simple words, a Microprocessor is a digital device on a chip which can fetch instruction from memory, decode and execute them and give results.
A Microprocessor takes a bunch of instructions in machine language and executes them, telling the processor what it has to do. Microprocessor performs three basic things while executing the instruction:

- It performs some basic operations like addition, subtraction, multiplication, division and some logical operations using its Arithmetic and Logical Unit (ALU).
- Data in Microprocessor can move from one location to another.
- It has a Program Counter (PC) register that stores the address of next instruction based on the value of PC, Microprocessor jumps from one location to another and takes decision.



A microprocessor consists of an ALU, control unit and register array. Where **ALU** performs arithmetic and logical operations on the data received from an input device or memory. Control unit controls the instructions and flow of data within the computer. And, **register array** consists of registers identified by letters like B, C, D, E, H, L, and accumulator , which acts as temporary fast access memory locations for processing data

Working of Microprocessor:

The microprocessor follows a sequence to execute the instruction: Fetch, Decode, and then Execute.
Initially, the instructions are stored in the storage memory of the computer in sequential order. The microprocessor fetches those instructions from the stored area (memory), then decodes it and executes those instructions till STOP instruction is met. Then, it sends the result in binary form to the output port. Between these processes, the register stores the temporary data and ALU (Arithmetic and Logic Unit) performs the computing functions.

Features of Microprocessor:

- **Low Cost** - Due to integrated circuit technology microprocessors are available at very low cost. It will reduce the cost of a computer system.
- **High Speed** - Due to the technology involved in it, the microprocessor can work at very high speed. It can execute millions of instructions per second.
- **Small Size** - A microprocessor is fabricated in a very less footprint due to very large scale and ultra large scale integration technology. Because of this, the size of the computer system is reduced.
- **Versatile** - The same chip can be used for several applications, therefore, microprocessors are versatile.
- **Low Power Consumption** - Microprocessors are using metal oxide semiconductor technology, which consumes less power.
- **Less Heat Generation** - Microprocessors uses semiconductor technology which will not emit much heat as compared to vacuum tube devices.
- **Reliable** - Since microprocessors use semiconductor technology, therefore, the failure rate is very less. Hence it is very reliable.
- **Portable** - Due to the small size and low power consumption microprocessors are portable.

Types of Microprocessors:

- **Complex Instruction Set Computer (CISC) -**

  CISC or Complex Instruction Set Computer is a computer architecture where instructions are such that a single instruction can execute multiple low level operations like loading from memory, storing into memory or an arithmetic

operation etc. It has multiple addressing nodes within single instruction.CISC makes use of very few registers.

- **Reduced Instruction Set Computer (RISC)-**

  RISC or Reduced Instruction Set Computer is a computer architecture where instruction are simple and designed to get executed quickly. Instructions get completed in one clock cycle this is because of the optimization of instructions and pipelining (a technique that allows for simultaneous execution of parts, or stages, of instructions to more efficiently process instructions). RISC makes use of multiple registers to avoid large interactions with memory. It has few addressing nodes.

Difference between RISC and CISC

S.No.	RISC	CISC
1.	Simple instruction set	Complex instruction set
2.	Consists of Large number of registers.	Less number of registers
3.	Larger Program	Smaller program
4.	Simple processor circuitry (small number of transistors)	Complex processor circuitry (more number of transistors)
5.	More RAM usage	Little Ram usage
6.	Simple addressing modes	Variety of addressing modes
7.	Fixed length instructions	Variable length instructions
8.	Fixed number of clock cycles for executing one instruction	Variable number of clock cycles for each instructions