Computer Memory (Introduction)

Computer memory is an internal storage area used to store data and programs.

It can be divided into two groups:

1) primary memory and
2) secondary memory.

- While the main memory holds instructions and data when a program is executing, the auxiliary or the secondary memory holds data and programs not currently in use and provides long-term storage.

- The primary memory is volatile, so the data can be retained in it, only when the power is on. Moreover, it is very expensive and therefore limited in capacity.

- The secondary memory stores data or instructions permanently, even when the power is turned off. It is cheap and can store large volumes of data. Moreover, data stored in auxiliary memory is highly portable, as the users can easily move it from one computer to the other.

- The only drawback of secondary memory is that data can be accessed from it at a very slow speed as and when compared with the data access speed of primary memory.
- A memory is just like a human brain. The memory is divided into large number of small parts called cells. Each location or cell has a unique address.
  For example, *if the computer has 64k words, then this memory unit has 64 * 1024 = 65536 memory locations. The address of these locations varies from* ***0 to 65535***.

Primary Memory (Main Memory)

- Primary memory holds only those data and instructions on which the computer is currently working.
- It has a limited capacity and data is lost when power is switched off.
- It is generally made up of semiconductor device.
- These memories are not as fast as registers.
- It is divided into two subcategories RAM and ROM.

Characteristics of Main Memory

- These are semiconductor memories.
- It is known as the main memory.
- A computer cannot run without the primary memory.
- Usually volatile memory.
- Data is lost in case power is switched off.
- It is the working memory of the computer.
- Faster than secondary memories.

RAM is of two types −

- Static RAM (SRAM)
- Dynamic RAM (DRAM)

Static RAM (SRAM)

- The word **static** indicates that the memory retains its contents as long as power is being supplied.
- SRAM chips use a matrix of 6-transistors and no capacitors.
- SRAM is used as cache memory and has very fast access.

Characteristic of Static RAM

- Long life
- No need to refresh
- Faster and Used as cache memory
- Expensive

Dynamic RAM (DRAM)

- DRAM, unlike SRAM, must be continually **refreshed** in order to maintain the data.
- This is done by placing the memory on a refresh circuit that rewrites the data several hundred times per second.
- DRAM is used for most system memory as it is cheap and small. All DRAMs are made up of memory cells, which are composed of one capacitor and one transistor.

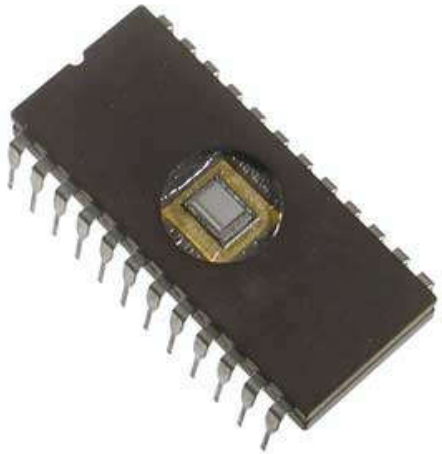
Characteristics of Dynamic RAM

- Short data lifetime
- Needs to be refreshed continuously
- Slower as compared to SRAM
- Used as RAM
- Smaller in size, Less expensive
- Less power consumption

Read Only Memory

- ROM stands for **Read Only Memory**.
- The memory from which we can only read but cannot write on it.
- This type of memory is non-volatile. The information is stored permanently in such memories during manufacture.
- A ROM stores such instructions that are required to start a computer.
- This operation is referred to as **bootstrap**.

- ROM chips are not only used in the computer but also in other electronic items like washing machine and microwave oven.



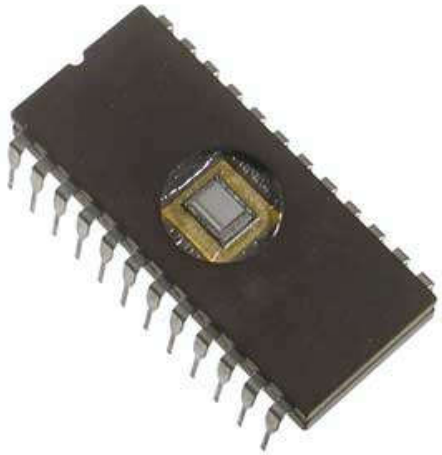
Types of ROMs and their characteristics-

MROM (Masked ROM)

- The very first ROMs were hard-wired devices that contained a pre-programmed set of data or instructions.
- These kind of ROMs are known as masked ROMs, which are inexpensive.

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 program.
- 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)

- EPROM can be erased by exposing it to ultra-violet light. Usually, an EPROM eraser achieves this function.
- During programming, an electrical charge is trapped in an insulated gate region.
- For erasing this charge,ultra-violet light is passed through a lid.

EEPROM (Electrically Erasable and Programmable Read Only Memory)

- EEPROM is programmed and erased electrically.
- It can be erased and reprogrammed about ten thousand times.
- 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.

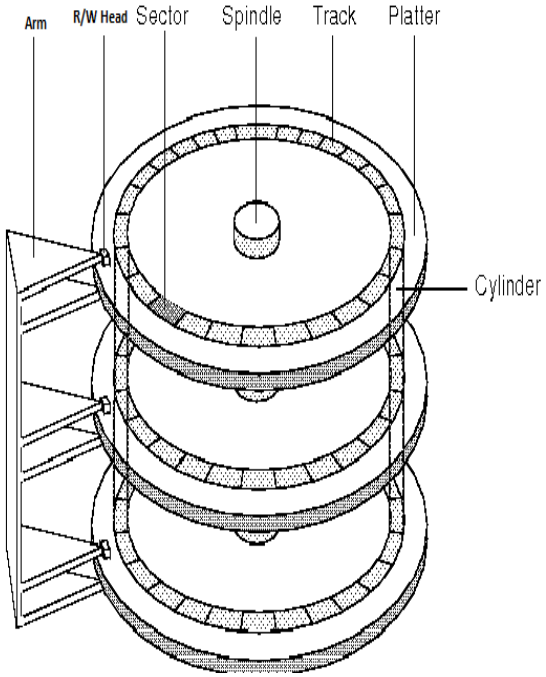
Secondary Memory (Auxiliary Memory)

- This type of memory is also known as external memory or non-volatile.
- It is slower than the main memory.
- These are used for storing data/information permanently.
- CPU directly does not access these memories, instead they are accessed via input-output routines.
- The contents of secondary memories are first transferred to the main memory, and then the CPU can access it.

For example, disk, CD-ROM, DVD, etc.

Characteristics of Secondary Memory

- **Non-volatile memory** − Data is not lost when power is cut off.
- **Reusable** − The data stays in the secondary storage on permanent basis until it is not overwritten or deleted by the user.
- **Reliable** − Data in secondary storage is safe because of high physical stability of secondary storage device.
- **Convenience** − With the help of a computer software, authorised people can locate and access the data quickly.
- **Capacity** − Secondary storage can store large volumes of data in sets of multiple disks.
- **Cost** − It is much lesser expensive to store data on a tape or disk than primary memory.



Example of Secondary Memory:

HARD DISK

A hard disk is a set of disks stacked together like phonograph records, that has data recorded electromagnetically in concentric circles also known as tracks.

A single hard disk includes several *platters (or disks)* that are covered with a magnetic recording medium.

- Each platter requires two read/write heads, one for each side.
- Data is actually stored on the surface of a platter in sectors and tracks.

Cache Memory

- Cache memory is a very high speed semiconductor memory which can speed up the CPU.
- It acts as a buffer between the CPU and the main memory.
- It is used to hold those parts of data and program which are most frequently used by the CPU.
- The parts of data and programs are transferred from the disk to cache memory by the operating system, from where the CPU can access them.

- Cache memory is basically a portion of memory made of high-speed static RAM (SRAM) instead of the slower and cheaper dynamic RAM (DRAM) which is used for main memory.

Advantages of cache memory −

- 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.

SEQUENTIAL AND RANDOM ACCESS:

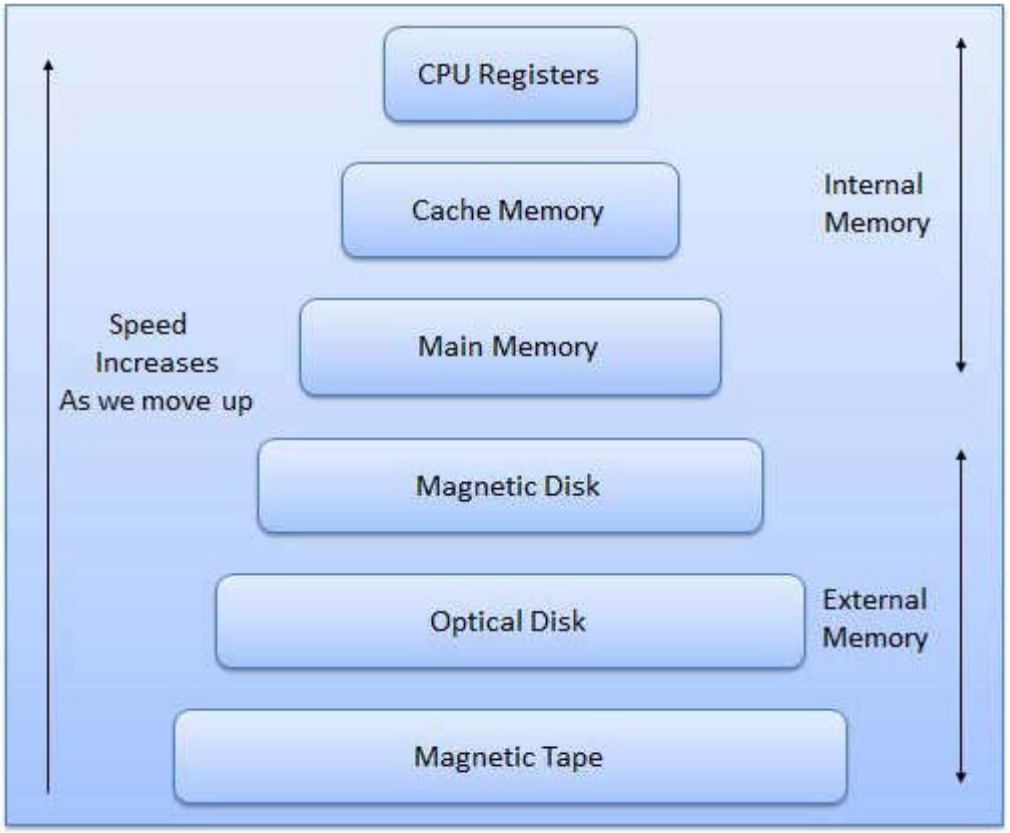
Memory devices can be accessed either randomly or sequentially.

Sequential access	Random access
Data is read sequentially in a specified order. 	Data is read in an arbitrary manner. 
If the 99th record is desired after the 1st one, then all the records have to be traversed to reach the desired one. Therefore, the time required to return data can vary depending on the position of the record.	Random access always returns data in constant time.
Sequential access devices can store a large number of records at a very low cost.	Random access devices are expensive than sequential access devices.
Magnetic tapes support sequential access.	RAM (Random Access Memory) supports random access.
Sequential-access is faster if records are to be accessed in the same order.	Random-access is faster if records are to be accessed in a random order.

MEMORY HIERARCHY

- Characteristics of Memory Hierarchy are following when we go **from top to bottom**.

- Capacity in terms of storage increases.
- Cost per bit of storage decreases.
- Frequency of access of the memory by the CPU decreases.
- Access time by the CPU increases.



Computer - Memory Units

- Memory unit is the amount of data that can be stored in the storage unit. This storage capacity is expressed in terms of Bytes.

S.No	Unit & Description
1	**Bit (Binary Digit)** A binary digit is logical 0 and 1 representing a passive or an active state of a component in an electric circuit.
2	**Nibble** A group of 4 bits is called nibble.
3	**Byte** A group of 8 bits is called byte. A byte is the smallest unit, which can represent a data item or a character.
4	**Word** A computer word, like a byte, is a group of fixed number of bits processed as a unit. The length of a computer word is called word-size or word length. It may be as small as 8 bits or may be as long as 96 bits.

The following table lists some higher storage units −

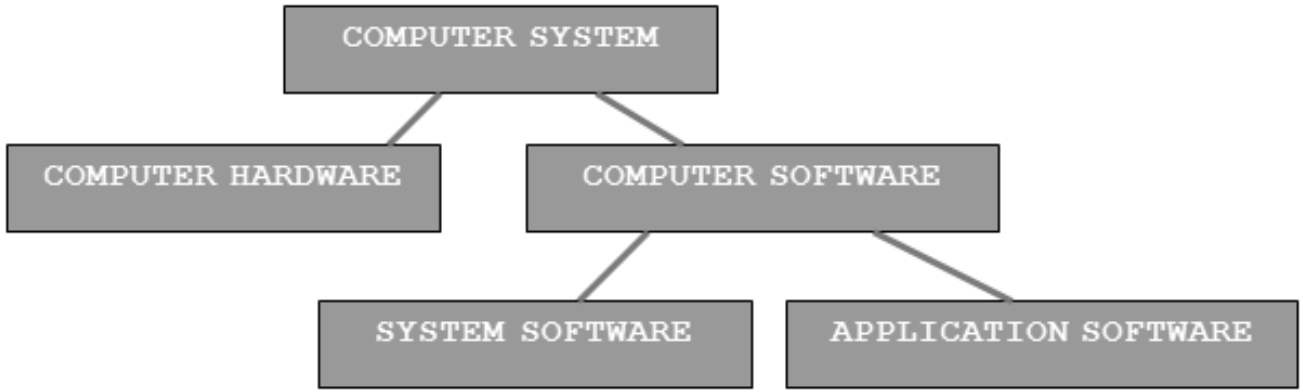
S.No.	Unit & Description
1	**Kilobyte (KB)** 1 KB = 1024 Bytes
2	**Megabyte (MB)** 1 MB = 1024 KB
3	**GigaByte (GB)** 1 GB = 1024 MB
4	**TeraByte (TB)** 1 TB = 1024 GB
5	**PetaByte (PB)** 1 PB = 1024 TB

Tabular Representation of various Memory Sizes

NAME	EQUAL TO	SIZE(IN BYTES)
Bit	1 bit	1/8
Nibble	4 bits	1/2 (rare)
Byte	8 bits	1
Kilobyte	1024 bytes	1024
Megabyte	1, 024kilobytes	1, 048, 576
Gigabyte	1, 024 megabytes	1, 073, 741, 824
Terrabyte	1, 024 gigabytes	1, 099, 511, 627, 776
Petabyte	1, 024 terrabytes	1, 125, 899, 906, 842, 624
Exabyte	1, 024 petabytes	1, 152, 921, 504, 606, 846, 976
Zettabyte	1, 024 exabytes	1, 180, 591, 620, 717, 411, 303, 424
Yottabyte	1, 024 zettabytes	1, 208, 925, 819, 614, 629, 174, 706, 176

Prepared by: Amit Kr Mishra, Department of CSE, GEHU, DDUN

Software & Hardware



- The computer hardware cannot think and make decisions on its own.
- The hardware needs a software (a set of programs) to instruct what has to be done.
- A program is a set of instructions that is arranged in a sequence to guide a computer to find a solution for the given problem.
- The process of writing a program is called programming.

Computer software is written by computer programmers using a programming language.

Examples of computer software include:

- Driver Software
- Educational software
- Media Players and Media Development Software
- Productivity Software such as word processors, database management utilities, and presentation software, Operating Systems software , etc.
- Computer Games

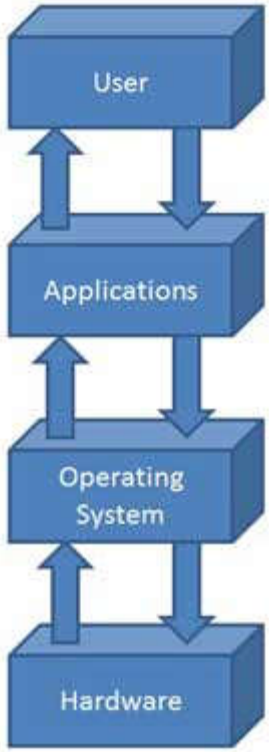

System Software	Application Software
It is a collection of programs that enable the users to interact with hardware components efficiently	It is a collection of programs written for a specific application. Like, we have library system, inventory control system, etc
It controls and manages the hardware	It uses the services provided by the system software to interact with hardware components
System software is machine dependant	It is machine independent
The programmer must understand the architecture of the machine and hardware details to write a system software	The programmer ignores the architecture of the machine and hardware details to write an application software
Interacts with the hardware directly	Interacts with the hardware indirectly through system calls provided by system software
Writing a system software is a complicated task	Writing application programs is relatively very easy
Example: compiler, operating system	Example: MS-WORD, PAINT

Operating System:

- The primary goal of an operating system is to make the computer system c*onvenient and efficient to use.*
- An operating system ensures that the system resources (such as CPU, memory, I/O devices, etc) are utilized efficiently.

  For example, there may be many service requests on a web server and each user request need to be serviced. Similarly, there may be many programs residing in the main memory.

  Therefore, the system needs to determine which programs are active and which need to wait for some I/O operation.



Objectives of Operating System:

- Manages the computer hardware
- Provides a user interface
- Process Management
- Memory Management
- Security Management

Examples of Operating System

- MS DOS System
- Windows Operating System
- UNIX
- LINUX
- Mobile Operating Systems

  such as- iOS, Android, Windows Phone , BlackBerry OS

Few other software:

TRANSLATORS

Compiler: A compiler is a special type of program that transforms source code written in a programming language into machine language comprising of just two digits- 1s and 0s. The resultant code in 1s and 0s is known as the object *code*.

Interpreter: Like the compiler, the interpreter also converts machine level language instructions into machine level language. However, it translates the instructions into an intermediate form, which it then executes. Usually, a compiled program executes faster than an interpreted program.

- **Assembler:** An assembler takes an assembly language program as an input and gives a code in machine language (also called an object code) as output.

There is a one-to-one correspondence between the assembly language code and the machine language code. However, if there is an error, the assembler gives a list of errors. The object file is created only when the assembly language code is free from errors.

- **Linker:** Also called *link editor* and *binder,* a linker is a program that combines object modules to form an executable program. Generally, in case of a large program, the programmers prefer to break a code into smaller modules. Eventually, when the source code of all the modules has been converted into object code, all the modules must be put together. This is the job of the linker. The compiler automatically invokes the linker as the last step in compiling a program.

- **Loader**: A loader is a special type of program that copies programs from a storage device to main memory, where they can be executed. Most loaders are transparent to the users.

- **Debugge**r: Debugging tools, commonly known as *debugger* s , are used to identify coding errors at different stages of software (or program) development. A debugger is a program that runs other programs allowing users to exercise some degree of control over their programs so that they can examine them when things go wrong.

Acquiring Computer software

Software can be acquired in following three ways:

1) Buying pre-written software
2) Having customized software
3) Downloading public domain software

Computer Hardware represents the physical and tangible components of a computer, i.e. the components that can be seen and touched.

Examples of Hardware are the following −

- **Input devices** − keyboard, mouse, etc.
- **Output devices** − printer, monitor, etc.
- **Secondary storage devices** − Hard disk, CD, DVD, etc.
- **Internal components** − CPU, motherboard, RAM, etc.

COMPUTER NETWORKS

A computer network is a collection of computers and devices interconnected to facilitate sharing of resources, information and electronic documents.

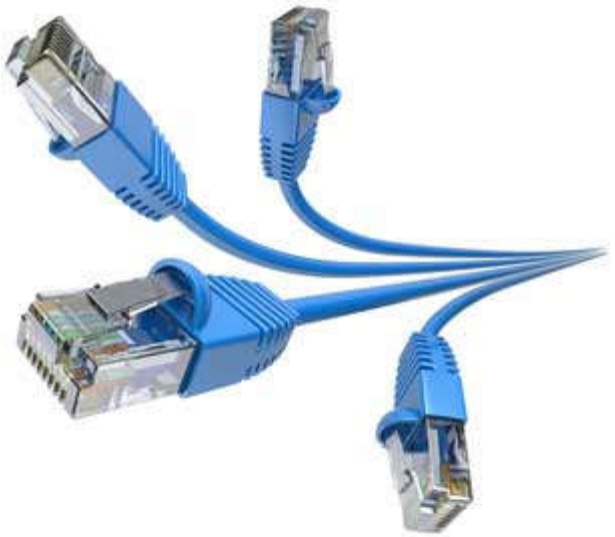
Advantages of Comp. Networks

- File Sharing
- Resource Sharing
- Increased Storage Capacity
- Load Sharing
- Facilitate communications

- Following is the list of hardware's required to set up a computer network.

1) Network Cables

Network cables are used to connect computers. The most commonly used cable is Category 5 cable RJ-45.



2) Distributors

- A computer can be connected to another one via a serial port but if we need to connect many computers to produce a network, this serial connection will not work.
- The solution is to use a central body to which other computers, printers, scanners, etc. can be connected and then this body will manage or distribute network traffic.

3) Router

- A router is a type of device which acts as the central point among computers and other devices that are a part of the network. It is equipped with holes called ports. Computers and other devices are connected to a router using network cables. Now-a-days router comes in wireless modes using which computers can be connected without any physical cable.

4) Network Card

- Network card is a necessary component of a computer without which a computer cannot be connected over a network. It is also known as the network adapter or Network Interface Card (NIC). Most branded computers have network card pre-installed. Network cards are of two types: Internal and External Network Cards.

Internal Network Cards

External Network Cards

5) Universal Serial Bus (USB)

- USB card is easy to use and connects via USB port. Computers automatically detect USB card and can install the drivers required to support the USB network card automatically.

TYPES OF NETWORKS

LOCAL AREA NETWORK (LAN) was first invented for communication between two computers but later was used to connect computers and devices in a limited geographical area such as home, school, computer laboratory, office building, or closely positioned group of buildings. Owing to limited scope and cost of operation, LANS are typically owned, controlled, and managed by a single person or organization.

A LAN can be one of two types: *wired* or wireless. These days, most of the wired LANS are based on Ethernet technology. However, they can also be created using wires like coaxial cables, phone lines and power lines.

METROPOLITAN AREA NETWORK (MAN) The size of the network varies in between LAN and WAN. A MAN covers an area of 5 – 50 kms.

A MAN (like a WAN) is usually not owned by a single individual or an organization. It is owned by either a consortium of users or by a network service provider who sells the service to the users.

A MAN is a high speed network that allows sharing of regional resources. Though MAN is not a very widely used area network but it has its own importance for some government bodies and organizations on larger scale.

WIDE AREA NETWORKS (WAN) spans a large geographic area such as a city, country, or even intercontinental distances, using a communications channel that combines many types of media such as telephone lines, cables, and air waves. WAN can be created by linking LANs together.

When individual networks connect together to form a larger network (or a bigger WAN), the resulting network is called an internetwork, which is generically abbreviated to ‘an internet’. Moreover, when all WANs all over the world connect to form a global internet, it is called The Internet.

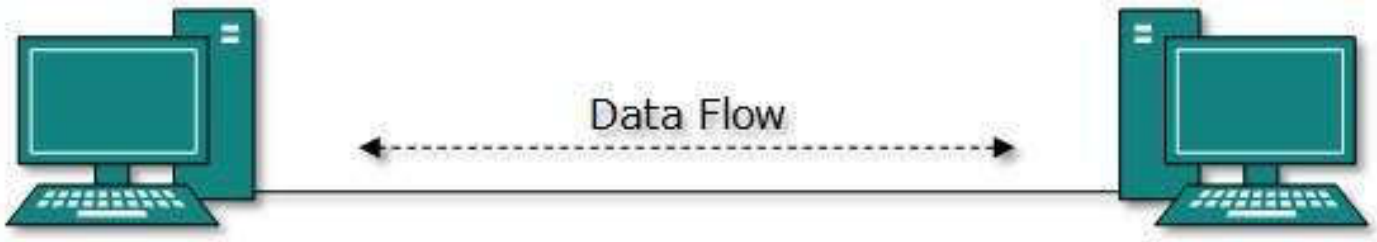
- **CAMPUS AREA NETWORKS** (CANs) is a computer network created by interconnecting LANs within a limited geographical area. The network is almost entirely owned by the campus of an enterprise, university, government, military bases, etc. The size of the area that CANs covers is larger than LAN and smaller than MAN or WAN.
- These days, CANs are mostly formed using the wireless communication mediums rather than cabling and wirings. CANs are economical, beneficial and easy to implement in the specific kilometers of locality.
- **PERSONAL AREA NETWORK** (PAN) is a computer network designed for communication between computer devices like mobile computers, cell phones and PDAs that are close to one person. The scope or the reach of a PAN is a few meters (less than 10 meters).
- PANs are used to communicate with the personal devices themselves or for connecting to a higher level network and the Internet. PANs can either be wired with computer buses such as USB and FireWire or be wireless with network technologies such as Infra red and Bluetooth. Bluetooth PANs are also called piconets.

NETWORK TOPOLOGIES

- Topology refers to the schematic description of the arrangement of a network.
- Network topology refers to the actual geometric layout of computers and other devices connected to the network.
- A Network Topology is the arrangement with which computer systems or network devices are connected to each other. Topologies may define both physical and logical aspect of the network. Both logical and physical topologies could be same or different in a same network.

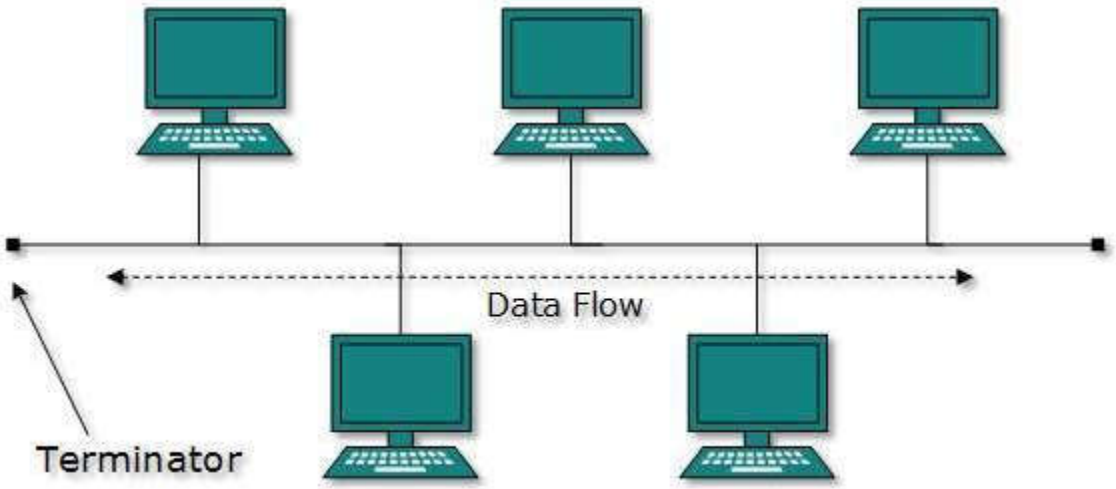
Point-to-Point

- Point-to-point networks contains exactly two hosts such as computer, switches or routers, servers connected back to back using a single piece of cable. Often, the receiving end of one host is connected to sending end of the other and vice-versa.
- If the hosts are connected point-to-point logically, then may have multiple intermediate devices. But the end hosts are unaware of underlying network and see each other as if they are connected directly.



Bus Topology

- In case of Bus topology, all devices share single communication line or cable.Bus topology may have problem while multiple hosts sending data at the same time. Therefore, Bus topology either uses CSMA/CD technology or recognizes one host as Bus Master to solve the issue. It is one of the simple forms of networking where a failure of a device does not affect the other devices. But failure of the shared communication line can make all other devices stop functioning.
- Both ends of the shared channel have line terminator. The data is sent in only one direction and as soon as it reaches the extreme end, the terminator removes the data from the line.


Data Flow
Terminator

BUS TOPOLOGY in which each computer is connected to the single cable. When a node wants to send data, it creates a message and adds the address of the recipient to it. Then, it checks whether the line is free or not. If the line is free it puts the message on the line else wait till it becomes available.

When the message is placed, every node connected to the bus checks the destination address mentioned on it. If the node's address does not match the intended address, it ignores the message otherwise it picks it and sends an acknowledgement to the sending device before it frees the line.

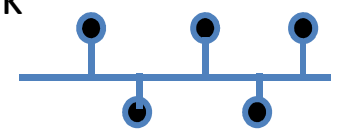
Advantages: Easy to install and connect a new device to the network

Requires less cable length than a star topology., thus inexpensive

Failure of a single node does not affect the network

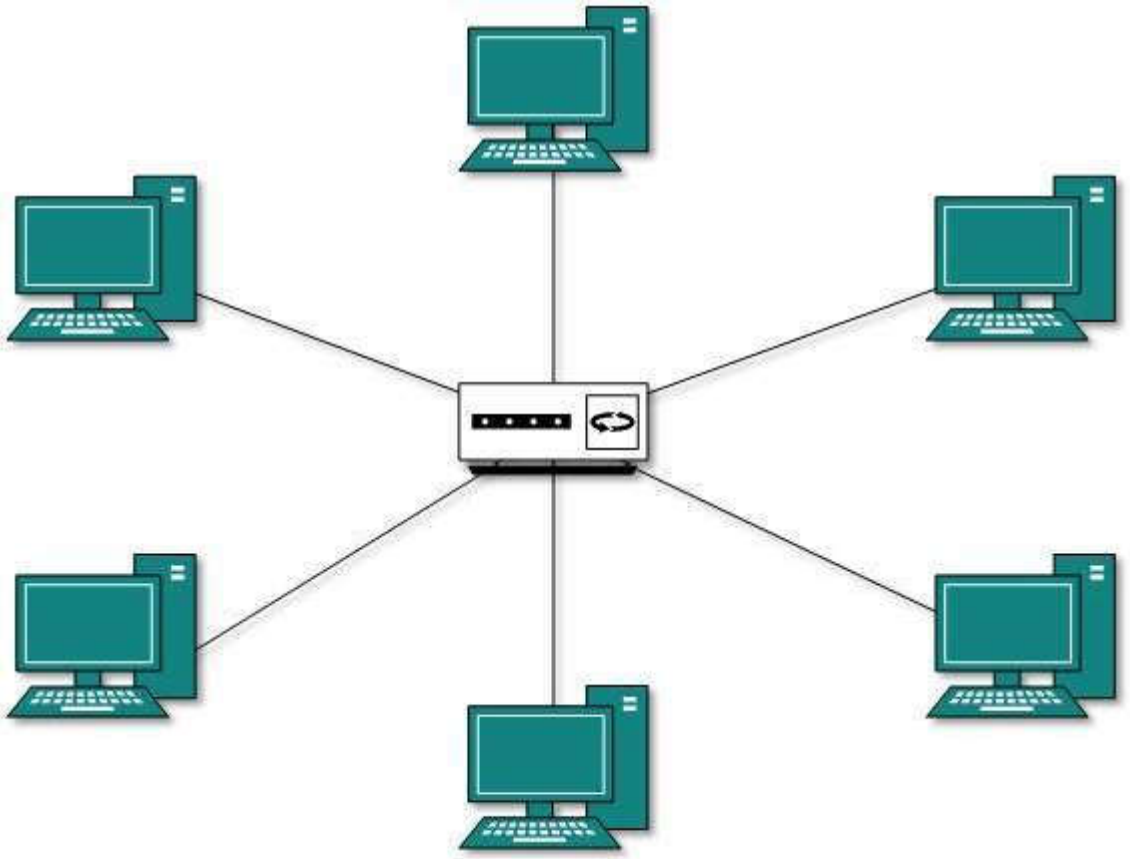
Disadvantages: Failure in the cable results in shut down of the entire network

As the number of nodes increases, the speed of the network slows down



Star Topology

- All hosts in Star topology are connected to a central device, known as hub device, using a point-to-point connection. That is, there exists a point to point connection between hosts and hub. The hub device can be any of the following:
- Layer-1 device such as hub or repeater
- Layer-2 device such as switch or bridge
- Layer-3 device such as router or gateway
- As in Bus topology, hub acts as single point of failure. If hub fails, connectivity of all hosts to all other hosts fails. Every communication between hosts, takes place through only the hub.Star topology is not expensive as to connect one more host, only one cable is required and configuration is simple.



STAR TOPOLOGY

It is considered the easiest topology to design and implement. In this topology, each node is connected to a central hub with a point-to-point connection. All traffic that traverses the network passes through the central hub. So, the host node that controls communication between other nodes.

The hub also regenerates the message (because signals become weak due to noise) and then send it to the destination node.

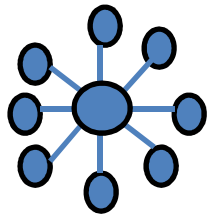
Advantages: Easy to install

New nodes can be connected easily

the network does not get disturbed when any device is added or removed from it

Easy to detect faults

Failure of any other node (except the hub) does not affect the network



Disadvantages:

Requires more cable length than bus topology.

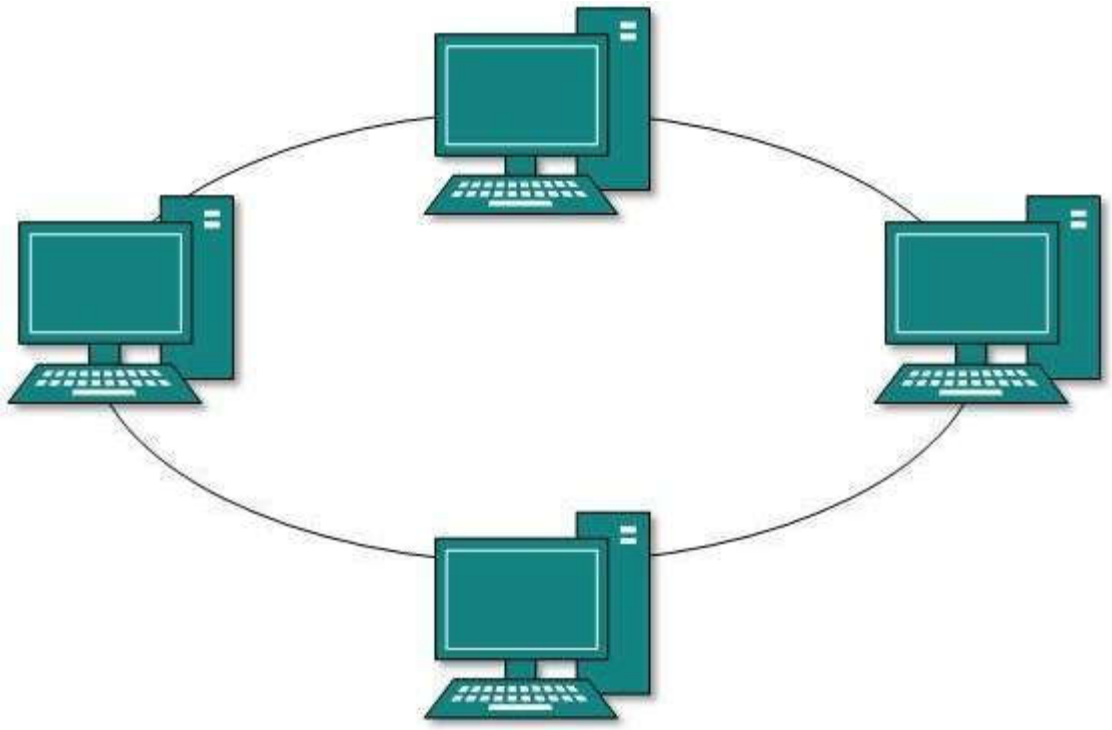
If the central hub fails, the entire network is shut down

More expensive bus topologies because of the cost of the hubs, cables etc.

Ring Topology

- In ring topology, each host machine connects to exactly two other machines, creating a circular network structure. When one host tries to communicate or send message to a host which is not adjacent to it, the data travels through all intermediate hosts. To connect one more host in the existing structure, the administrator may need only one more extra cable.
- Failure of any host results in failure of the whole ring.Thus, every connection in the ring is a point of failure. There are methods which employ one more backup ring.

•



RING TOPOLOGY

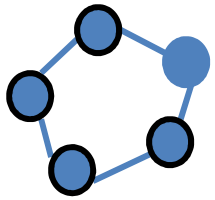
All the nodes are connected to each other in the shape of a closed loop, so that every node is connected directly to two other nodes, one on either side. The messages travel through the ring in a circular fashion. A failure in any cable or device breaks the loop and can cause the network to shut down.

A node receives message from any of its two adjacent nodes. Then it checks the destination address. If the message is addressed to it, it accepts it otherwise just regenerates the signals and passes it to the next node in sequence.

Advantages: Best suited for networks that does not have a hub

More reliable , easy to install , can span over larger distances

Every node has equal chance to transmit data



Disadvantages: Delay in communication.

In case of network failure, it is difficult to diagnose the fault

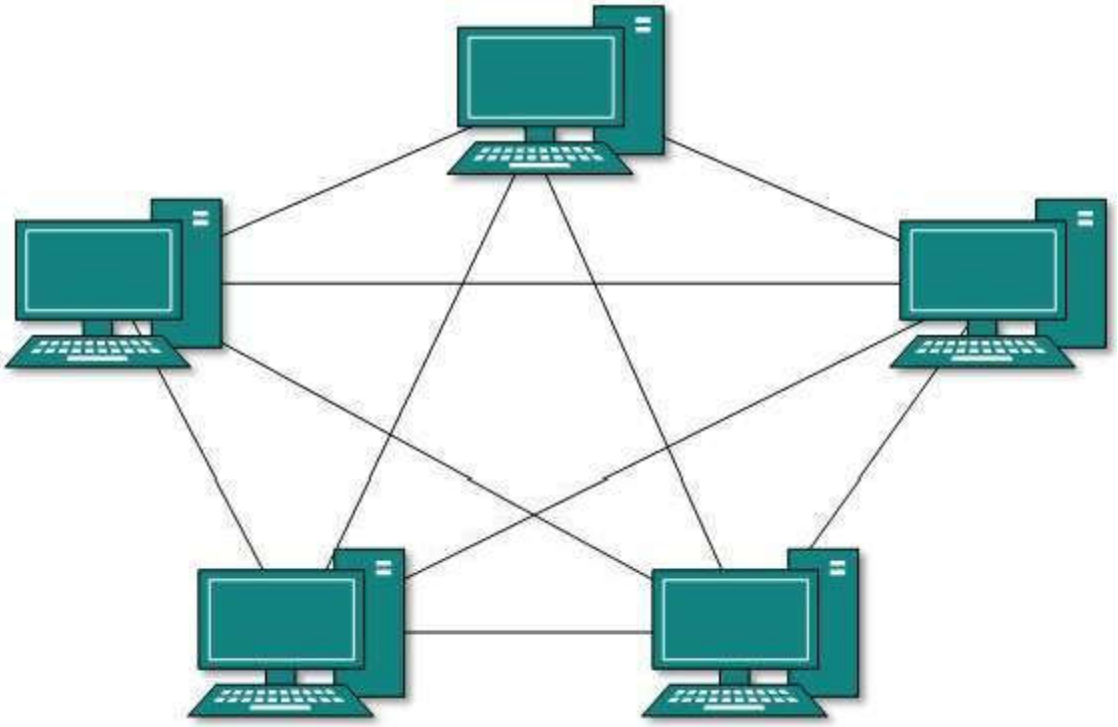
If one node fails, the entire network is shut down because the ring is not complete

Difficult to add or remove nodes from the network

Mesh Topology

- In this type of topology, a host is connected to one or multiple hosts.This topology has hosts in point-to-point connection with every other host or may also have hosts which are in point-to-point connection to few hosts only.
- Hosts in Mesh topology also work as relay for other hosts which do not have direct point-to-point links. Mesh technology comes into two types:
- **Full Mesh**: All hosts have a point-to-point connection to every other host in the network. Thus for every new host n(n-1)/2 connections are required. It provides the most reliable network structure among all network topologies.
- **Partially Mesh**: Not all hosts have point-to-point connection to every other host. Hosts connect to each other in some arbitrarily fashion. This topology exists where we need to provide reliability to some hosts out of all.

- .



MESH TOPOLOGY

Also known as a completely connected network, every node is connected to every other node using a separate physical link. Mesh topology involve the concept of routes. Unlike other topologies, in a mesh network, a message can take any of the several possible paths from the source to the destination. For example, if a message has to be sent from A to B via C, then if node C fails, then the message can be sent to B via any other node in the network.

Advantages: Failure of a node does not affect the entire network.

Communication is fast as there is a direct link between the nodes

Each connection can have its own data load, so the traffic problem is eliminated.

Mesh network ensures security of data because every message travels along a dedicated link.

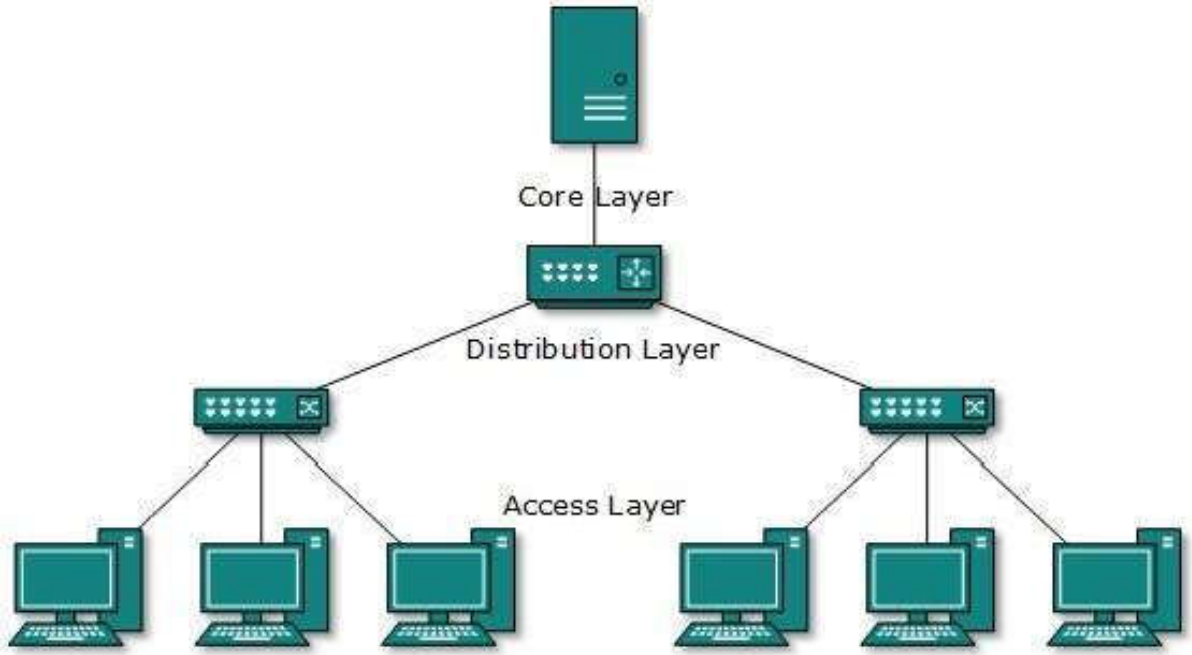
It is easy to detect network errors

Disadvantages: It is the most expensive network as for n nodes, n*(n-1)/2 physical links (cables) are required.

It is difficult to install

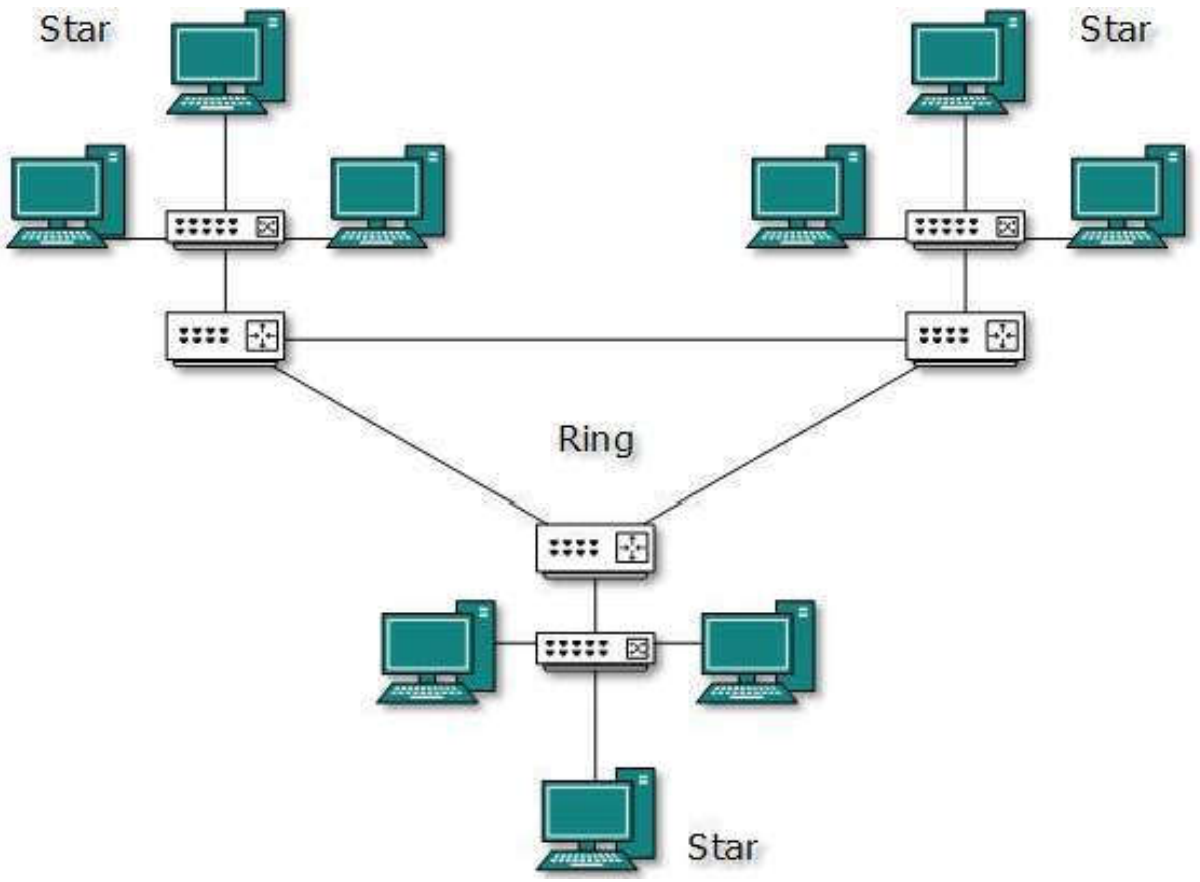
Tree Topology

- Also known as Hierarchical Topology, this is the most common form of network topology in use presently.This topology imitates as extended Star topology and inherits properties of bus topology.
- This topology divides the network in to multiple levels/layers of network. Mainly in LANs, a network is bifurcated into three types of network devices. The lowermost is access-layer where computers are attached. The middle layer is known as distribution layer, which works as mediator between upper layer and lower layer. The highest layer is known as core layer, and is central point of the network, i.e. root of the tree from which all nodes fork.
- All neighboring hosts have point-to-point connection between them.Similar to the Bus topology, if the root goes down, then the entire network suffers even.though it is not the single point of failure. Every connection serves as point of failure, failing of which divides the network into unreachable segment.


Core Layer
Distribution Layer
Access Layer

Hybrid Topology

- A network structure whose design contains more than one topology is said to be hybrid topology. Hybrid topology inherits merits and demerits of all the incorporating topologies.
- The above picture represents an arbitrarily hybrid topology. The combining topologies may contain attributes of Star, Ring, Bus, and Daisy-chain topologies. Most WANs are connected by means of Dual-Ring topology and networks connected to them are mostly Star topology networks. Internet is the best example of largest Hybrid topology


Star
Star
Ring
Star

HYBRID TOPOLOGY

Each of the topologies has their own advantages and disadvantages. So in the real world, a pure start or pure ring or bus is rarely used. Rather a combination of two or more topologies is used.

Hence, hybrid network topology uses a combination of any two or more topologies in such a way that the resulting network does not exhibit one of the standard topologies (e.g., bus, star, ring, etc.). Two very commonly used hybrid network topologies include the star ring network and star bus network.

If an organization has two departments where one department has connected its computers using the bus network and another department is using ring network. Now the networks of the two departments can be connected by a central hub thereby using the star topology.