INTRODUCTION TO

INFORMATION TECHNOLOGY

Introduction to Information Technology

SYLLABUS

Course Title: Introduction to Information Technology **Full Marks:** 60 + 20 + 20
Course No: CSC109   **Pass Marks:** 24 + 8 + 8
Nature of the Course: Theory + Lab   **Credit Hrs:** 3
Semester: I

Course Description: This course covers the basic concepts of computers and information technology including introduction, hardware, software, memory, input/output, data representation, database, networks and data communication, Internet, multimedia, and computer security.

Course Objectives: The main objective of this course is to provide students knowledge of fundamental concepts of computers and information technology.

Course Contents:
Unit 1: Introduction to Computer (3 Hrs.)
Introduction; Digital and Analog Computers; Characteristics of Computer; History of Computer; Generations of Computer; Classification of Computer; The Computer System; Application of Computers

Unit 2: The Computer System Hardware (3 Hrs.)
Introduction; Central Processing Unit; Memory Unit; Instruction Format; Instruction Set; Instruction Cycle; Microprocessor; Interconnecting the Units of a Computer; Inside a Computer Cabinet

Unit 3: Computer Memory (4 Hrs.)
Introduction; Memory Representation; Memory Hierarchy; CPU Registers; Cache Memory; Primary Memory; Secondary Memory; Access Types of Storage Devices; Magnetic Tape; Magnetic Disk; Optical Disk; Magneto-Optical Disk; How the Computer uses its memory

Unit 4: Input and Output Devices (4 Hrs.)
Introduction; Input-Output Unit; Input Devices; Human Data Entry Devices; Source Data Entry Devices; Output Devices; I/O Port; Working of I/O System

Unit 5: Data Representation (6 Hrs.)
Introduction; Number System; Conversion from Decimal to Binary, Octal, Hexadecimal; Conversion of Binary, Octal, Hexadecimal to Decimal; Conversion of Binary to Octal, Hexadecimal; Conversion of Octal, Hexadecimal to Binary; Binary Arithmetic; Signed and Unsigned Numbers; Binary Data Representation; Binary Coding Schemes; Logic Gates

Unit 6: Computer Software (6 Hrs.)
Introduction; Types of Software; System Software; Application Software; Software Acquisition; Operating System (Introduction, Objectives of Operating System, Types of OS, Functions of OS, Process Management, Memory Management, File

Management, Device Management, Protection and Security, User Interface, Examples of Operating Systems)

Unit 7: Data Communication and Computer Network (5 Hrs.)
Introduction; Importance of Networking; Data Transmission Media; Data Transmission across Media; Data Transmission and Data Networking; Computer Network; Network Types; Network Topology; Communication Protocol; Network Devices; Wireless Networking

Unit 8: The Internet and Internet Services (4 Hrs.)
Introduction; History of Internet; Internetworking Protocol; The Internet Architecture; Managing the Internet; Connecting to Internet; Internet Connections; Internet Address; Internet Services; Uses of Internet; Introduction to Internet of Things (IoT), Wearable Computing, and Cloud Computing, Introduction to E-commerce, E-governance, and Smart City, and GIS

Unit 9: Fundamentals of Database (4 Hrs.)
Introduction; Database; Database System; Database Management System; Database System Architectures; Database Applications; Introduction to Data Warehousing, Data mining, and BigData

Unit 10: Multimedia (3 Hrs.)
Introduction; Multimedia - Definition; Characteristics of Multimedia; Elements of Multimedia; Multimedia Applications

Unit 11: Computer Security (3 Hrs.)
Introduction; Security Threat and Security Attack; Malicious Software; Security Services; Security Mechanisms (Cryptography, Digital Signature, Firewall, Users Identification and Authentication, Intrusion Detection Systems); Security Awareness; Security Policy

Laboratory Works: After completing this course, students should have practical knowledge of different hardware components of computer, operating systems, Word Processors, Spreadsheets, Presentation Graphics, Database Management Systems, and Internet and its services.

Text Books:
1. Computer Fundamentals, Anita Goel, Pearson Education India

Reference Books:
1. Introduction to Computers, Peter Norton, 7th Edition, McGraw Hill Education
2. Computer Fundamental, Pradeep K. Sinha and Priti Sinha
3. Data Mining Concepts and Techniques, Third Edition, Jiawei Han, Micheline Kamber and Jian Pei
4. Cloud Computing Bible, Barrie Sosinsky, Wiley

Contents

Introduction to Information Technology

Unit 1: Introduction to Computer

1.1 Introduction;

Nowadays, computers are an integral part of our lives. They are used for the reservation of tickets for airplanes and railways, payment of telephone and electricity bills, deposit and withdrawal of money from banks, processing of business data, forecasting of weather conditions, diagnosis of diseases, searching for information on the Internet, etc. Computers are also used extensively in schools, universities, organizations, music industry, movie industry, scientific research, law firms, fashion industry, etc.

The term computer is derived from the word compute. The word compute means to calculate. A computer is an electronic machine that accepts data from the user, processes the data by performing calculations and operations on it, and generates the desired output results. Computer performs both simple and complex operations, with speed and accuracy.

1.2 Digital and Analog Computers;

A digital computer uses distinct values to represent the data internally. All information are represented using the digits 0s and 1s. The computers that we use at our homes and offices are digital computers.

Analog computer is another kind of a computer that represents data as variable across a continuous range of values. The earliest computers were analog computers. Analog computers are used for measuring of parameters that vary continuously in real time, such as temperature, pressure and voltage. Analog computers may be more flexible but generally less precise than digital computers. Slide rule is an example of an analog computer.

1.3 Characteristics of Computer;

1. **Speed:** The speed of electronic computer is very fast because the signal can pass at the speed of electricity which is near to that of speed of light i.e. 2.997×10^8 m/sec. Thus millions of calculations can be done with in a fragment of second. The processing speed of microprocessor is generally measured in Millions of Instructions per second (MIPS).

$$\text{Frequency (Hz)} = \frac{1}{Time(in\ seconds)}$$

Measuring units for processing speed in terms of frequency and time

Measuring units in frequency	Measuring units in Time
1 Hz (Hertz or 1×10^0 Hz)	1 second (1×10^0 sec)
1 KHz (Kilo Hertz or 1×10^3 Hz)	1 milli second (1×10^{-3} Sec)
1 MHz (Mega Hertz or 1×10^6 Hz)	1 micro second (1×10^{-6} Sec)
1 GHz (Giga Hertz or 1×10^9 Hz)	1 nano second (1×10^{-9} Sec)
1 THz (Tera Hertz or 1×10^{12} Hz)	1 pico second (1×10^{-12} Sec)
1 PHz (Peta Hertz or 1×10^{15} Hz)	1 femto second (1×10^{-15} Sec)
1 EHz (Exa Hertz or 1×10^{18} Hz)	1 atto second (1×10^{-18} Sec)

2. **Accuracy:** The accuracy of a computer is consistently high, and every calculation is performed with the same accuracy i.e. 100% accurate results are always same as per design. The degree of accuracy of a particular computer depends upon its design. Sometimes errors can occur in a computer system due to inaccurate data or instructions fed by human or end users. If the error occurs due to the wrong input by the user then this type of error is called GIGO (Garbage In Garbage Out) and if the error occurs due to the malfunction of the computer hardware and this type of error is called Bug.

3. **Automatic:** Being a programmable machine, modern computer can be programmed in such a way that it can perform a series of tasks automatically after it has been started. Modern digital computer does not need a human operator to execute different types of command.

4. **Diligence:** Unlike human beings, the computer never feels tired and bored to do a same task repeatedly. The computer can perform the same task with same speed and accuracy for a long time if it is provided the suitable environment and correct instruction. It is free from tiredness, weakness, lack of concentration and monotony.

5. **Reliability:** computer system is particularly adept a lot of work without any mistakes and tiredness. That's why, they are widely used everywhere because of their reliability.

6. **Versatility:** Computer can perform different tasks depending upon the different program fed to it, is known as versatility. It has a wide range of application areas i.e. it can do many types of jobs at the same time. Some of the application areas of computers are in education, business, office-automation, bank, medical diagnosis, science and technology. So, it is versatile in nature.

7. **Storage Capacity:** A computer can store a great amount of data. It has data storage area, called memory. Any data stored in the memory can be retrieved at any time and at very fast speed. The storing capacity of a computer is measured in terms of byte, kilobyte, megabyte, gigabyte, terabyte etc.

Storage unit conversion table

Unit	Abbreviation	In Unit
1 Byte	B	8 bits
1 Kilobyte	KB	1024 bytes
1 Megabyte	MB	1024 KB
1 Gigabyte	GB	1024 MB
1 Terabyte	TB	1024 GB
1 Petabyte	PB	1024 TB
1 Exabyte	EB	1024 PB
1 Zeta byte	ZB	1024 EB

Computers have several limitations too. Computer can only perform tasks that it has been programmed to do. Computer cannot do any work without instructions from the user. It executes instructions as specified by the user and does not take its own decisions.

1.4 History of Computer;
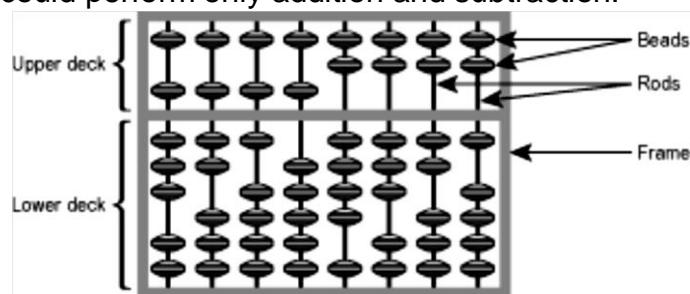
"Necessity is the mother of invention". The saying is true for computer also because computer were inbuilt as a result of man search for fast and accurate calculating device. It is due to the need and curiosity of man for faster way to calculate has led to develop a computer. Following are the major development in the history of computer from early calculating machine to modern high speed digital computer.

The Mechanical Era (Zeroth Generation)
A mechanical computer is built from mechanical components such as levers and gears, rather than electronic components. The most common examples are adding machines and mechanical counters, which use the turning of gears to increment output displays. More complex examples can carry out multiplication and division, and even differential analysis.

1. **ABACUS:**
   The abacus, also called a counting frame, is a calculating tool used primarily in parts of Asia for performing arithmetic processes during 5000 BC to 3000 BC. Today, abaci are often constructed as a bamboo frame with beads sliding on wires, but originally they were beans or stones moved in grooves in sand or on tablets of wood, stone, or metal. It has two parts, heaven and earth. The heaven consists of two beads, whereas the earth part consists of five beads to each rod. It could perform only addition and subtraction.
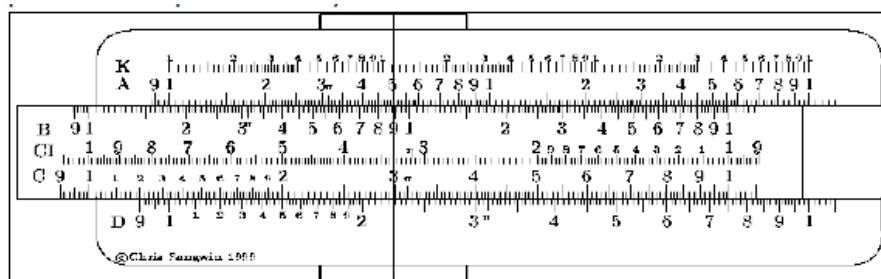
2. **John Napier and his rods (1550-1617 AD):**
   The Scottish mathematician John Napier first published the table of logarithms in 1614 AD and it was very useful in simplifying the multiplication of large numbers. He invented some bone rods, uses bones to demonstrate division by subtraction and multiplication by addition, according to principle of logarithm. These rods were made up of strips of bones on which numbers were carved and painted, that device was called Napier bone.
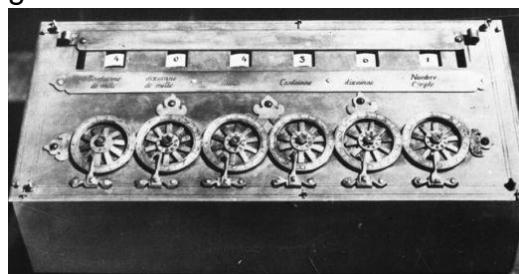
3. **William Oughtred and his Slide Rule (1575-1660 AD):**

   In 1620 AD, the English man William Oughtred invented a rectangular calculating device called slide-rule based on the principles of logarithm. It consists of two graduated scales, one of which slips upon the other. The scales are devised in such a manner that suitable alignment of one scale against the other makes it possible to find products and quotient of any numbers.
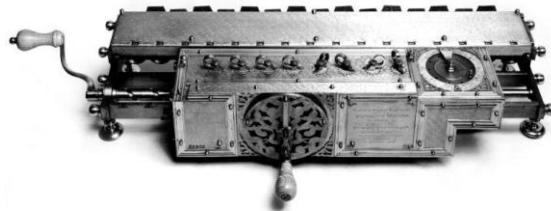
4. **Blaise Pascal and his calculator (1623-1662 AD):**

   A French mathematician Blaise Pascal developed mechanical calculating machine to assist his father in his accounting work. In this machine numbers were entered by dialing a series of numbered wheels each wheel having numbers from 0 to 9. For addition the wheel was moved into forward direction and for subtraction into the reverse direction. It could perform addition and subtraction up to 8 digits. This mechanical calculator is called Pascaline.
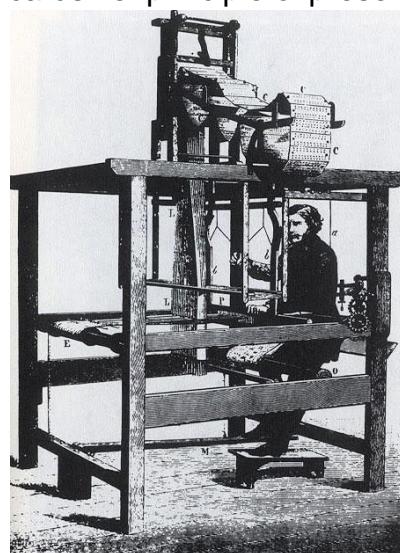
5. **Baron Gottfried Wilhelm Von Leibniz and his Stepped Reckoner (1646-1716):**

   In 1671 AD, Leibniz modified the Pascaline machine and invented a first calculator Stepped Reckoner, which was able to perform automatic addition, subtraction, multiplication, division and could find out square root. Leibniz's machine used stepped cylinder each with nine teeth of varying lengths instead of wheels. It was called Leibniz-calculator or Stepped Reckoner.
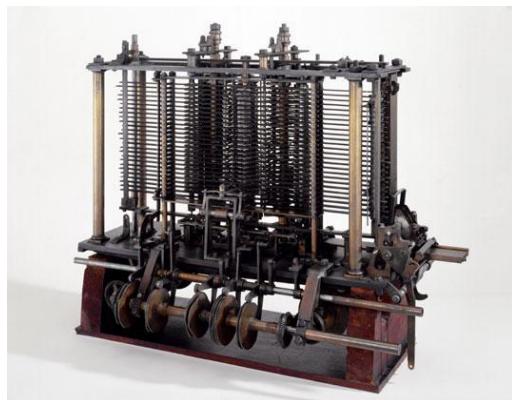
6. **Joseph Marie Jacquard and his Jacquard loom and punch cards (1752-1834 AD):**

   The French textile manufacturer Joseph Marie Jacquard invented a mechanism for automated weaving cloths for the textile industry at Lyon in 1802 AD. This machine was used to automatically control weaving looms to facilitate the production of weaving cloth with complex patterns. This machine was controlled by punch-cards i.e. principle of present and absence of holes.

7. **Charles Babbage and his analytical engine (1791-1871 AD):**

   The English professor and mathematician Charles Babbage invented the Difference Engine at Cambridge University in 1822 AD. Theoretically, it can solve differential equations and calculate various mathematical function, logarithmic tables, polynomial and trigonometric functions. Later Babbage continued working on his difference Engine and developed new idea for constructing general purpose, fully programmable automated machine called Analytical Engine in 1833 AD. He invented the working principle of computers on which the modern computer was based.

8. **Lady Augusta Ada Byron Lovelace (1815-1852 AD):**
   The English woman Lady Augusta created the instruction routines to be fed into the computer and she suggested to Charles Babbage writing a plan for how the engine might calculate using Bernoulli numbers and that plan is now regarded as the first computer program. She wrote the program for Babbage analytical engine so she was considered the first computer programmer and, a software language developed by the US Defense Department was named Ada in her honor in 1979 AD.
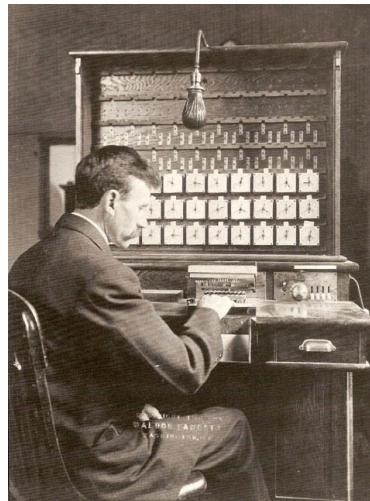
9. **George Boole and his Boolean algebra (1815-1864 AD):**
   The Irish mathematician George Boole developed a method of reducing statements of logic to algebraic statements, using a simple set of symbols. He interprets that algebraic statements in to Boolean algebra in terms of truth values, called the propositional calculus, form the basis of the digital processes in modern computers. His logic provides the basic idea of presenting the electronic circuits with binary input (0 or 1) and its output.

10. **Herman Hollerith and his tabulating machine (1860-1929 AD):**
    An American inventor Herman Hollerith also applied the Jacquard loom concept in computing and applies for patents for automatic punch-card tabulating machine in 1884 AD. He invented a machine known as tabulating Machine (TM) in 1886 AD. He used cards to store data and each punch on a card represented one number and combinations of punches represented a letter. 80 variables could be stored on a single card. His machine was first used to compute the U.S. census. Later his machine became International Business Machine (IBM). He was founder of IBM and founded in 1924 AD. in USA.

The Electro-Mechanical Era
1. **Mark-I:**
   A professor of Physics Howard H. Aiken designed a general purpose mechanical computer at Harvard University and IBM, while working on his doctorate in physics, in the year 1937 and the machine was called IBM Automatic Sequence Controlled Calculator (IBM ASCC) and later as Harvard Mark-I. It was relay based computer on the based on the concept of Charles Babbage's analytical engine principle. Later he modified Mark-I and made Mark-II by using 19000 valves.
   Features of Mark-I
   - It used about 18 thousand vacuum tubes as main memory device with seven lakhs 50 thousands parts.
   - It is about 51 feet long, 8 feet height and 3 feet wide i.e. bulky in size.
   - Technically, it was very complicated machine, consumed huge amount of power and generated lot of heat during the operation.
   - The time taken for multiplication was about four seconds and for division was about eleven seconds.
   - The results were printed at the rate of one result per five seconds.

2. **ABC (Atanasoff Berry Computer):**
   In 1938 John J. Atanasoff and Clifford Berry designed ABC for solving systems of simultaneous equation. It used 18000 calves and other 45 valves for internal logic and capacitors for storage of electrical charges. It used punched cards as secondary storage.

3. **John Von Neumann (1903-1957 AD):**
   The Hungarian mathematician Jon von Neumann gave the idea of stored program computer in the sense that program is stored internally in the main memory of the computer along with its associated data in 1945. Therefore, he is called father of stored program. Before that, modification of a program was not possible.

The electronic computers Era

1. **ENIAC (Electronic Numerical Integrator and Calculator):**
   In 1946, John W. Mauchly and J. Prosper Eckert developed ENIAC at Pennsylvania, USA. ENIAC was the first-popular general-purpose electronic digital computer. It was very large machine weighing about 30 tons and containing about 17468 vacuum tubes, 70,000 resisters, 5 million soldered joins and it consumed 160 kilowatts. It took 200 microseconds for addition and 3 milliseconds to perform a 10-digit multiplication. It was first and last computer, which used decimal number system instead of binary system.

2. **EDSAC (Electronic Delay Storage Automatic Computer):**
   Maurice Wilkes invented EDSAC in 1949. It also used vacuum tubes. Although started after EDVAC, it was completed before EDVAC. Therefore, it became the first stored program computer.

3. **EDVAC (Electronic Discrete Variable Automatic Computer):**
   It was developed by J.P. Eckert and J. Mauchly in 1952 and used for more school personnel and the Ballistics Research Laboratory of the USA Army, built a computer named EDVAC. John Von Neumann supervised the construction of EDVAC. It also used vacuum tubes and some internal storage.

4. **UNIVAC (Universal Automatic Computer):**
   J.P. Eckert and J. Mauchly developed it in 1961. It was the first computer manufactured for commercial use and general-purpose digital computer. Before this, all the computers were used for either defense or census.

1.5 Generations of Computer;
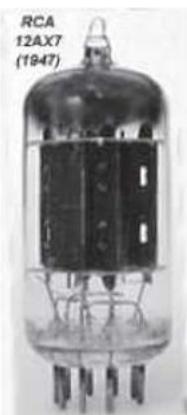
The computer has evolved from a large—sized simple calculating machine to a smaller but much more powerful machine. The evolution of computer to the current state is defined in terms of the generations of computer. Each generation of computer is designed based on a new technological development, resulting in better, cheaper and smaller computers that are more powerful, faster and efficient than their predecessors. Currently, there are five generations of computer. In the following subsections, we will discuss the generations of computer in terms of—

1. the technology used by them (hardware and software),
2. computing characteristics (speed, i.e., number of instructions executed per second),
3. physical appearance, and
4. their applications.

1.5.1 First Generation (1940 to 1956): Using Vacuum Tubes

- *Hardware Technology* The first generation of computers used vacuum tubes (Figure) for circuitry and magnetic drums for memory. The input to the computer was through punched cards and paper tapes. The output was displayed as printouts.

- *Software Technology* The instructions were written in machine language. Machine language uses 0s and 1s for coding of the instructions. The first generation computers could solve one problem at a time.
- *Computing Characteristics* The computation time was in milliseconds.
- *Physical Appearance* These computers were enormous in size and required a large room for installation.
- *Application* They were used for scientific applications as they were the fastest computing device of their time.
- *Examples* UNIVersal Automatic Computer (UNIVAC), Electronic Numerical Integrator And Calculator (ENIAC), and Electronic Discrete Variable Automatic Computer (EDVAC).

The first generation computers used a large number of vacuum tubes and thus generated a lot of heat. They consumed a great deal of electricity and were expensive to operate. The machines were prone to frequent malfunctioning and required constant maintenance. Since first generation computers used machine language, they were difficult to program.

1.5.2 Second Generation (1956 to 1963): Using Transistors

- *Hardware Technology* Transistors (Figure) replaced the vacuum tubes of the first generation of computers. Transistors allowed computers to become smaller, faster, cheaper, energy efficient and reliable. The second generation computers used magnetic core technology for primary memory. They used magnetic tapes and magnetic disks for secondary storage. The input was still through punched cards and the output using printouts. They used the concept of a stored program, where instructions were stored in the memory of computer.

- *Software Technology* The instructions were written using the assembly language. Assembly language uses mnemonics like ADD for addition and SUB for subtraction for coding of the instructions. It is easier to write instructions in assembly language, as compared to writing instructions in machine language. High-level programming languages, such as early versions of COBOL and FORTRAN were also developed during this period.
- *Computing Characteristics* The computation time was in microseconds.
- *Physical Appearance* Transistors are smaller in size compared to vacuum tubes, thus, the size of the computer was also reduced.
- *Application* The cost of commercial production of these computers was very high, though less than the first generation computers. The transistors had to be assembled manually in second generation computers.
- *Examples* PDP-8, IBM 1401 and CDC 1604.

Second generation computers generated a lot of heat but much less than the first generation computers. They required less maintenance than the first generation computers.

1.5.3 Third Generation (1964 to 1971): Using Integrated Circuits

- *Hardware Technology* The third generation computers used the Integrated Circuit (IC) chips. Figure shows IC chips. In an IC chip, multiple transistors are placed on a silicon chip. Silicon is a type of semiconductor. The use of IC chip increased the speed and the efficiency of computer, manifold. The keyboard and monitor were used to interact with the third generation computer, instead of the punched card and printouts.
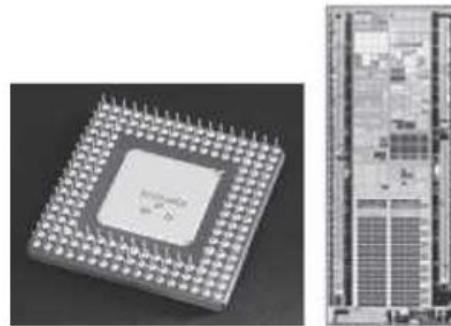
- **Software Technology** The keyboard and the monitor were interfaced through the operating system. Operating system allowed different applications to run at the same time. High-level languages were used extensively for programming, instead of machine language and assembly language.
- **Computing Characteristics** The computation time was in nanoseconds.
- **Physical Appearance** The size of these computers was quite small compared to the second generation computers.
- **Application Computers** became accessible to mass audience. Computers were produced commercially, and were smaller and cheaper than their predecessors.
- **Examples** IBM 370, PDP 11.

The third generation computers used less power and generated less heat than the second generation computers. The cost of the computer reduced significantly, as individual components of the computer were not required to be assembled manually. The maintenance cost of the computers was also less compared to their predecessors.

1.5.4 Fourth Generation (1971 to present): Using Microprocessors

- **Hardware Technology** They use the Large Scale Integration (LSI) and the Very Large Scale Integration (VLSI) technology. Thousands of transistors are integrated on a small silicon chip using LSI technology. VLSI allows hundreds of thousands of components to be integrated in a small chip. This era is marked by the development of microprocessor. Microprocessor is a chip containing millions of transistors and components, and, designed using LSI and VLSI technology. A microprocessor chip is shown in Figure. This generation of computers gave rise to Personal Computer (PC). Semiconductor memory replaced the earlier magnetic core memory, resulting in fast random access to memory. Secondary storage device like magnetic disks became smaller in physical size and larger in capacity. The linking of computers is another key development of this era. The computers were linked to form networks that led to the emergence of the Internet. This generation also saw the development of pointing devices like mouse, and handheld devices.

- *Software Technology* Several new operating systems like the MS-DOS and MSWindows developed during this time. This generation of computers supported Graphical User Interface (GUI). GUI is a user-friendly interface that allows user to interact with the computer via menus and icons. High-level programming languages are used for the writing of programs.
- *Computing Characteristics* The computation time is in picoseconds.
- *Physical Appearance* They are smaller than the computers of the previous generation. Some can even fit into the palm of the hand.
- *Application* They became widely available for commercial purposes. Personal computers became available to the home user.
- *Examples* The Intel 4004 chip was the first microprocessor. The components of the computer like Central Processing Unit (CPU) and memory were located on a single chip. In 1981, IBM introduced the first computer for home use. In 1984, Apple introduced the Macintosh.

The microprocessor has resulted in the fourth generation computers being smaller and cheaper than their predecessors. The fourth generation computers are also portable and more reliable. They generate much lesser heat and require less maintenance compared to their predecessors. GUI and pointing devices facilitate easy use and learning on the computer. Networking has resulted in resource sharing and communication among different computers.

1.5.5 Fifth Generation (Present and Next): Using Artificial Intelligence

The goal of fifth generation computing is to develop computers that are capable of learning and self-organization. The fifth generation computers use Super Large Scale Integrated (SLSI) chips that are able to store millions of components on a single chip. These computers have large memory requirements.
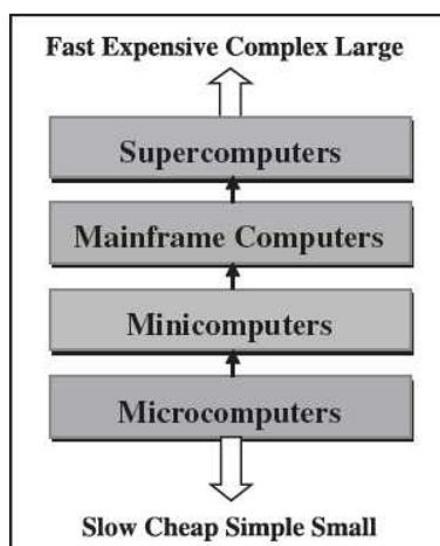
This generation of computers uses parallel processing that allows several instructions to be executed in parallel, instead of serial execution. Parallel processing results in faster processing speed. The Intel dualcore microprocessor uses parallel processing.

The fifth generation computers are based on Artificial Intelligence (AI). They try to simulate the human way of thinking and reasoning. Artificial Intelligence includes areas like Expert System (ES), Natural Language Processing (NLP), speech recognition, voice recognition, robotics, etc.

1.6 Classification of Computer;

The digital computers that are available nowadays vary in their sizes and types. The computers are broadly classified into four categories (Figure) based on their size and type—
(1) Microcomputers,
(2) Minicomputers,
(3) Mainframe computers, and
(4) Supercomputer.

1.6.1 Microcomputers
Microcomputers are small, low-cost and single-user digital computer. They consist of CPU, input unit, output unit, storage unit and the software. Although microcomputers are stand-alone machines, they can be connected together to create a network of computers that can serve more than one user. IBM PC based on Pentium microprocessor and Apple Macintosh are some examples of microcomputers. Microcomputers include desktop computers, notebook computers or laptop, tablet computer, handheld computer, smart phones and netbook, as shown in Figure.

PC   Laptop   Netbook

Tablet   PDA   Smart Phone

- ***Desktop Computer or Personal Computer (PC)*** is the most common type of microcomputer. It is a stand-alone machine that can be placed on the desk. Externally, it consists of three units—keyboard, monitor, and a system unit containing the CPU, memory, hard disk drive, etc. It is not very expensive and is suited to the needs of a single user at home, small business units, and organizations. Apple, Microsoft, HP, Dell and Lenovo are some of the PC manufacturers.

- ***Notebook Computers or Laptop*** resemble a notebook. They are portable and have all the features of a desktop computer. The advantage of the laptop is that it is small in size (can be put inside a briefcase), can be carried anywhere, has a battery backup and has all the functionality of the desktop. Laptops can be placed on the lap while working (hence the name). Laptops are costlier than the desktop machines.

- ***Netbook*** These are smaller notebooks optimized for low weight and low cost, and are designed for accessing web-based applications. Starting with the earliest netbook in late 2007, they have gained significant popularity now. Netbooks deliver the performance needed to enjoy popular activities like streaming videos or music, emailing, Web surfing or instant messaging. The word netbook was created as a blend of Internet and notebook.

- ***Tablet Computer*** has features of the notebook computer but it can accept input from a stylus or a pen instead of the keyboard or mouse. It is a portable computer. Tablet computer are the new kind of PCs.

- ***Handheld Computer or Personal Digital Assistant (PDA)*** is a small computer that can be held on the top of the palm. It is small in size. Instead of the keyboard, PDA uses a pen or a stylus for input. PDAs do not have a disk drive. They have a limited memory and are less powerful. PDAs can be connected to the Internet via a wireless connection. Casio and Apple are some of the manufacturers of PDA. Over the last few years, PDAs have merged into mobile phones to create smart phones.

- *Smart Phones* are cellular phones that function both as a phone and as a small PC. They may use a stylus or a pen, or may have a small keyboard. They can be connected to the Internet wirelessly. They are used to access the electronic-mail, download music, play games, etc. Blackberry, Apple, HTC, Nokia and LG are some of the manufacturers of smart phones.

1.6.2 Minicomputers

Minicomputers (Figure) are digital computers, generally used in multi-user systems. They have high processing speed and high storage capacity than the microcomputers. Minicomputers can support 4–200 users simultaneously. The users can access the minicomputer through their PCs or terminal. They are used for real-time applications in industries, research centers, etc. PDP 11, IBM (8000 series) are some of the widely used minicomputers.

1.6.3 Mainframe Computers

Mainframe computers (Figure) are multi-user, multi-programming and high performance computers. They operate at a very high speed, have very large storage capacity and can handle the workload of many users. Mainframe computers are large and powerful systems generally used in centralized databases. The user accesses the mainframe computer via a terminal that may be a dumb terminal, an intelligent terminal or a PC. A dumb terminal cannot store data or do processing of its own. It has the input and output device only. An intelligent terminal has the input and output device, can do processing, but, cannot store data of its own. The dumb and the intelligent terminal use the processing power and the storage facility of the mainframe computer. Mainframe computers are used in organizations like banks or companies, where many people require frequent access to the same data. Some examples of mainframes are CDC 6600 and IBM ES000 series.

1.6.4 Supercomputers

Supercomputers (Figure) are the fastest and the most expensive machines. They have high processing speed compared to other computers. The speed of a supercomputer is generally measured in FLOPS (FLoating point Operations Per Second). Some of the faster supercomputers can perform trillions of calculations per second. Supercomputers are built by interconnecting thousands of processors that can work in parallel.

Supercomputers are used for highly calculation-intensive tasks, such as, weather forecasting, climate research (global warming), molecular research, biological research, nuclear research and aircraft design. They are also used in major universities, military agencies and scientific research laboratories. Some examples of supercomputers are IBM Roadrunner, IBM Blue gene and Intel ASCI red. PARAM is a series of supercomputer assembled in India by C-DAC (Center for Development of Advanced Computing), in Pune. PARAM Padma is the latest machine in this series. The peak computing power of PARAM Padma is 1 Tera FLOP (TFLOP).

1.7 The Computer System;
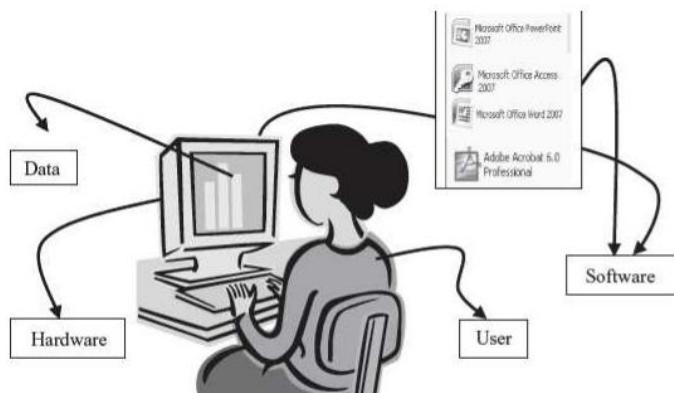
Computer is an electronic device that accepts data as input, processes the input data by performing mathematical and logical operations on it, and gives the desired output. The computer system consists of four parts•(1) Hardware, (2) Software, (3) Data, and (4) Users. The parts of computer system are shown in Figure.

Hardware consists of the mechanical parts that make up the computer as a machine. The hardware consists of physical devices of the computer. The devices are required for input, output, storage and processing of the data. Keyboard, monitor, hard disk drive, floppy disk drive, printer, processor and motherboard are some of the hardware devices.

Software is a set of instructions that tells the computer about the tasks to be performed and how these tasks are to be performed. Program is a set of instructions, written in a language understood by the computer, to perform a specific task. A set of programs and documents are collectively called software. The hardware of the computer system cannot perform any task on its own. The hardware needs to be instructed about the task to be performed. Software instructs the computer about the task to be performed. The hardware carries out these tasks. Different software can be loaded on the same hardware to perform different kinds of tasks.

Data are isolated values or raw facts, which by themselves have no much significance. For example, the data like 29, January, and 1994 just represent values. The data is provided as input to the computer, which is processed to generate some meaningful information. For example, 29, January and 1994 are processed by the computer to give the date of birth of a person.

Users are people who write computer programs or interact with the computer. They are also known as skinware, liveware, humanware or peopleware. Programmers, data entry operators, system analyst and computer hardware engineers fall into this category.

1.7.1 The Input-Process-Output Concept
A computer is an electronic device that
(1) accepts data,
(2) processes data,
(3) generates output, and
(4) stores data.

The concept of generating output information from the input 4 data is also referred to as input-process-output concept.

The input-process-output concept of the computer is explained as follows—
- *Input* The computer accepts input data from the user via an input device like keyboard. The input data can be characters, word, text, sound, images, document, etc.
- *Process* The computer processes the input data. For this, it performs some actions on the data by using the instructions or program given by the user of the data. The action could be an arithmetic or logic calculation, editing, modifying a document, etc. During processing, the data, instructions and the output are stored temporarily in the computer's main memory.
- *Output* The output is the result generated after the processing of data. The output may be in the form of text, sound, image, document, etc. The computer may display the output on a monitor, send output to the printer for printing, play the output, etc.
- *Storage* The input data, instructions and output are stored permanently in the secondary storage devices like disk or tape. The stored data can be retrieved later, whenever needed.
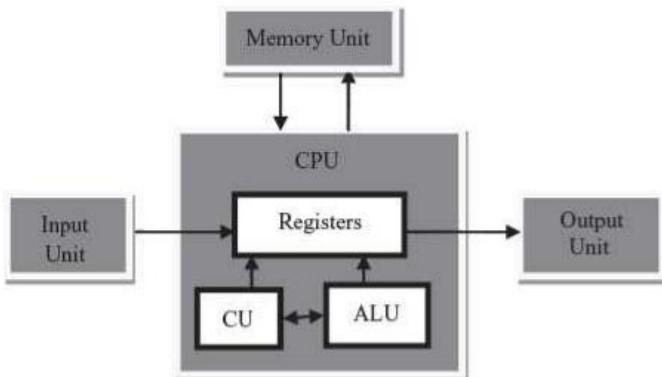
1.7.2 Components of Computer Hardware
The computer system hardware comprises of three main components —
1. Input/Output (I/O) Unit,
2. Central Processing Unit (CPU), and
3. Memory Unit.

The I/O unit consists of the input unit and the output unit. CPU performs calculations and processing on the input data, to generate the output. The memory unit is used to store the data, the instructions and the output information. Figure illustrates the typical interaction among the different components of the computer.

- ***Input/Output Unit*** The user interacts with the computer via the I/O unit. The Input unit accepts data from the user and the Output unit provides the processed data i.e. the information to the user. The Input unit converts the data that it accepts from the user, into a form that is understandable by the computer. Similarly, the Output unit provides the output in a form that is understandable by the user. The input is provided to the computer using input devices like keyboard, trackball and mouse. Some of the commonly used output devices are monitor and printer.
- ***Central Processing Unit*** CPU controls, coordinates and supervises the operations of the computer. It is responsible for processing of the input data. CPU consists of Arithmetic Logic Unit (ALU) and Control Unit (CU).
  - ALU performs all the arithmetic and logic operations on the input data.
  - CU controls the overall operations of the computer i.e. it checks the sequence of execution of instructions, and, controls and coordinates the overall functioning of the units of computer.

Additionally, CPU also has a set of registers for temporary storage of data, instructions, addresses and intermediate results of calculation.

- ***Memory Unit*** Memory unit stores the data, instructions, intermediate results and output, temporarily, during the processing of data. This memory is also called the main memory or primary memory of the computer. The input data that is to be processed is brought into the main memory before processing. The instructions required for processing of data and any intermediate results are also stored in the main memory. The output is stored in memory before being transferred to the output device. CPU can work with the information stored in the main memory. Another kind of storage unit is also referred to as the secondary memory of the computer. The data, the programs and the output are stored permanently in the storage unit of the computer. Magnetic disks, optical disks and magnetic tapes are examples of secondary memory.

1.8 Application of Computers
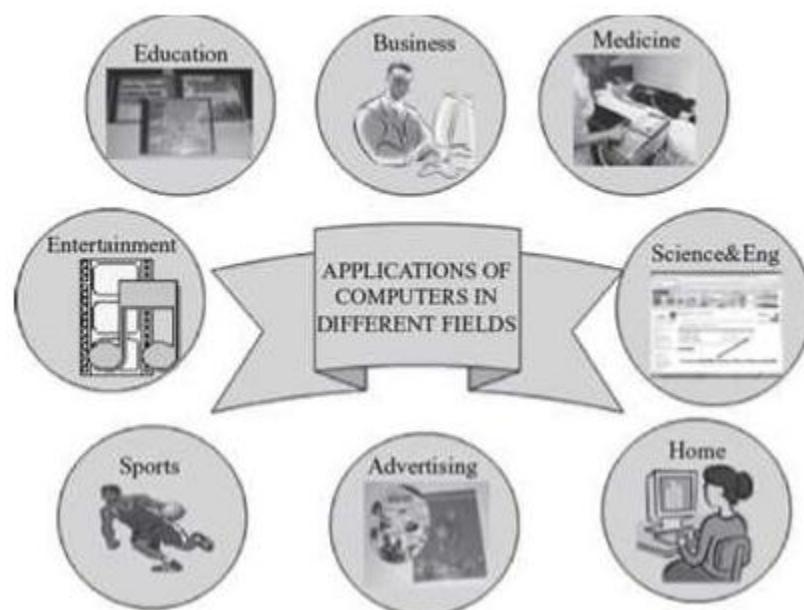
Computers have proliferated into various areas of our lives. For a user, computer is a tool that provides the desired information, whenever needed. You may use computer to get information about the reservation of tickets (railways, airplanes and cinema halls), books in a library, medical history of a person, a place in a map, or the

dictionary meaning of a word. The information may be presented to you in the form of text, images, video clips, etc.

Figure shows some of the applications of computer. Some of the application areas of the computer are listed below—

- *Education* Computers are extensively used, as a tool and as an aid, for imparting education. Educators use computers to prepare notes and presentations of their lectures. Computers are used to develop computer-based training packages, to provide distance education using the e-learning software, and to conduct online examinations. Researchers use computers to get easy access to conference and journal details and to get global access to the research material.
- *Entertainment* Computers have had a major impact on the entertainment industry. The user can download and view movies, play games, chat, book tickets for cinema halls, use multimedia for making movies, incorporate visual and sound effects using computers, etc. The users can also listen to music, download and share music, create music using computers, etc.
- *Sports* A computer can be used to watch a game, view the scores, improve the game, play games (like chess, etc.) and create games. They are also used for the purposes of training players.
- *Advertising* Computer is a powerful advertising media. Advertisement can be displayed on different websites, electronic-mails can be sent and reviews of a product by different customers can be posted. Computers are also used to create an advertisement using the visual and the sound effects. For the advertisers, computer is a medium via which the advertisements can be viewed globally. Web advertising has become a significant factor in the marketing plans of almost all companies. In fact, the business model of Google is mainly dependent on web advertising for generating revenues.

- *Medicine* Medical researchers and practitioners use computers to access information about the advances in medical research or to take opinion of doctors globally. The medical history of patients is stored in the computers. Computers are also an integral part of various kinds of sophisticated medical equipments like ultrasound machine, CAT scan machine, MRI scan machine, etc. Computers also provide assistance to the medical surgeons during critical surgery operations like laparoscopic operations, etc.
- *Science and Engineering* Scientists and engineers use computers for performing complex scientific calculations, for designing and making drawings (CAD/CAM applications) and also for simulating and testing the designs. Computers are used for storing the complex data, performing complex calculations and for visualizing 3–dimensional objects. Complex scientific applications like the launch of the rockets, space exploration, etc., are not possible without the computers.
- *Government* The government uses computers to manage its own operations and also for e-governance. The websites of the different government departments provide information to the users. Computers are used for the filing of income tax return, paying taxes, online submission of water and electricity bills, for the access of land record details, etc. The police department uses computers to search for criminals using fingerprint matching, etc.
- *Home* Computers have now become an integral part of home equipment. At home, people use computers to play games, to maintain the home accounts, for communicating with friends and relatives via Internet, for paying bills, for education and learning, etc. Microprocessors are embedded in house hold utilities like, washing machines, TVs, food processors, home theatres, security devices, etc.

The list of applications of computers is so long that it is not possible to discuss all of them here. In addition to the applications of the computers discussed above, computers have also proliferated into areas like banks, investments, stock trading, accounting, ticket reservation, military operations, meteorological predictions, social networking, business organizations, police department, video conferencing, telepresence, book publishing, web newspapers, and information sharing.

Unit 2: The Computer System Hardware

2.1 Introduction;

When we talk of computer hardware, the three related terms that require introduction are— computer architecture, computer organization and computer design. Computer architecture refers to the structure and behavior of the computer. It includes the specifications of the components, for example, instruction format, instruction set and techniques for addressing memory, and how they connect to the other components. Given the components, computer organization focuses on the organizational structure. It deals with how the hardware components operate and the way they are connected to form the computer. Given the system specifications, computer design focuses on the hardware to be used and the interconnection of parts. Different kinds of computer, such as a PC or a mainframe computer may have different organization; however, basic organization of the computer remains the same.

A computer consists of three main components—
(1) Input/Output (I/O) Unit,
(2) Central Processing Unit (CPU), and
(3) Memory Unit.

The computer user interacts with the computer via the I/O unit. The purpose of I/O unit is to provide data and instructions as input to the computer and to present relevant information as output from the computer. CPU controls the operations of the computer and processes the received input to generate the relevant output. The memory unit stores the instructions and the data during the input activity, to make instructions readily available to CPU during processing. It also stores the processed output. This chapter discusses the hardware components of the computer and the interaction between them.

2.2 Central Processing Unit;
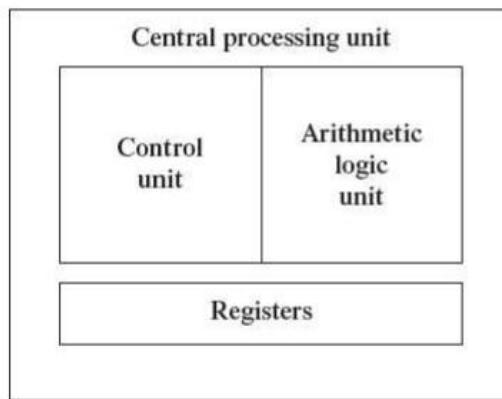
Central Processing Unit (CPU) or the processor is also often called the brain of computer. CPU (Figure) consists of Arithmetic Logic Unit (ALU) and Control Unit (CU). In addition, CPU also has a set of registers which are temporary storage areas for holding data, and instructions. ALU performs the arithmetic and logic operations on the data that is made available to it. CU is responsible for organizing the processing of data and instructions. CU controls and coordinates the activity of the other units of computer. CPU uses the registers to store the data, instructions during processing.

CPU executes the stored program instructions, i.e. instructions and data are stored in memory before execution. For processing, CPU gets data and instructions from the memory. It interprets the program instructions and performs the arithmetic and logic operations required for the processing of data. Then, it sends the processed

data or result to the memory. CPU also acts as an administrator and is responsible for supervising operations of other parts of the computer.

The CPU is fabricated as a single Integrated Circuit (IC) chip, and is also known as the microprocessor. The microprocessor is plugged into the motherboard of the computer (Motherboard is a circuit board that has electronic circuit etched on it and connects the microprocessor with the other hardware components).

```
┌─────────────────────────────────────────┐
│          Central processing unit         │
│  ┌──────────────┬──────────────────┐     │
│  │              │                  │     │
│  │   Control    │    Arithmetic    │     │
│  │    unit      │      logic       │     │
│  │              │       unit       │     │
│  │              │                  │     │
│  ├──────────────┴──────────────────┤     │
│  │            Registers             │     │
│  └──────────────────────────────────┘     │
└─────────────────────────────────────────┘
```

2.2.1 Arithmetic Logic Unit

- ALU consists of two units—arithmetic unit and logic unit.
- The arithmetic unit performs arithmetic operations on the data that is made available to it. Some of the arithmetic operations supported by the arithmetic unit are—addition, subtraction, multiplication and division.
- The logic unit of ALU is responsible for performing logic operations. Logic unit performs comparisons of numbers, letters and special characters. Logic operations include testing for greater than, less than or equal to condition.
- ALU performs arithmetic and logic operations, and uses registers to hold the data that is being processed.

2.2.2 Registers

- Registers are high-speed storage areas within the CPU, but have the least storage capacity. Registers are not referenced by their address, but are directly accessed and manipulated by the CPU during instruction execution.
- Registers store data, instructions, addresses and intermediate results of processing. Registers are often referred to as the CPU's working memory.
- The data and instructions that require processing must be brought in the registers of CPU before they can be processed. For example, if two numbers are to be added, both numbers are brought in the registers, added and the result is also placed in a register.
- Registers are used for different purposes, with each register serving a specific purpose. Some of the important registers in CPU (Figure) are as follows—
  - Accumulator (ACC) stores the result of arithmetic and logic operations.
  - Instruction Register (IR) contains the current instruction most recently fetched.
  - Program Counter (PC) contains the address of next instruction to be processed.

- o Memory Address Register (MAR) contains the address of next location in the memory to be accessed.
  - o Memory Buffer Register (MBR) temporarily stores data from memory or the data to be sent to memory.
  - o Data Register (DR) stores the operands and any other data.

PC	IR
MAR	ACC
MBR	DR

- The number of registers and the size of each (number of bits) register in a CPU helps to determine the power and the speed of a CPU.
- The overall number of registers can vary from about ten to many hundreds, depending on the type and complexity of the processor.
- The size of register, also called word size, indicates the amount of data with which the computer can work at any given time. The bigger the size, the more quickly it can process data. The size of a register may be 8, 16, 32 or 64 bits. For example, a 32–bit CPU is one in which each register is 32 bits wide and its CPU can manipulate 32 bits of data at a time. Nowadays, PCs have 32–bit or 64–bit registers.
- 32-bit processor and 64-bit processor are the terms used to refer to the size of the registers. Other factors remaining the same, a 64-bit processor can process the data twice as fast as one with 32-bit processor.

2.2.3 Control Unit

- The control unit of a computer does not do any actual processing of data. It organizes the processing of data and instructions. It acts as a supervisor and, controls and coordinates the activity of the other units of computer.
- CU coordinates the input and output devices of a computer. It directs the computer to carry out stored program instructions by communicating with the ALU and the registers. CU uses the instructions in the Instruction Register (IR) to decide which circuit needs to be activated. It also instructs the ALU to perform the arithmetic or logic operations. When a program is run, the Program Counter (PC) register keeps track of the program instruction to be executed next.
- CU tells when to fetch the data and instructions, what to do, where to store the results, the sequencing of events during processing etc.
- CU also holds the CPU's Instruction Set, which is a list of all operations that the CPU can perform.

The function of a (CU) can be considered synonymous with that of a conductor of an orchestra. The conductor in an orchestra does not perform any work by itself but
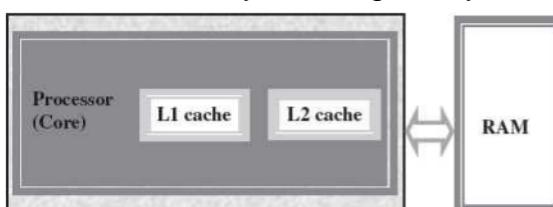
manages the orchestra and ensures that the members of orchestra work in proper coordination.

2.3 Memory Unit;

The memory unit consists of cache memory and primary memory. Primary memory or main memory of the computer is used to store the data and instructions during execution of the instructions. Random Access Memory (RAM) and Read Only Memory (ROM) are the primary memory. In addition to the main memory, there is another kind of storage device known as the secondary memory. Secondary memory is non-volatile and is used for permanent storage of data and programs. A program or data that has to be executed is brought into the RAM from the secondary memory.

2.3.1 Cache Memory

- The data and instructions that are required during the processing of data are brought from the secondary storage devices and stored in the RAM. For processing, it is required that the data and instructions are accessed from the RAM and stored in the registers. The time taken to move the data between RAM and CPU registers is large. This affects the speed of processing of computer, and results in decreasing the performance of CPU.
- Cache memory is a very high speed memory placed in between RAM and CPU. Cache memory increases the speed of processing.
- Cache memory is a storage buffer that stores the data that is used more often, temporarily, and makes them available to CPU at a fast rate. During processing, CPU first checks cache for the required data. If data is not found in cache, then it looks in the RAM for data.
- To access the cache memory, CPU does not have to use the motherboard's system bus for data transfer. (The data transfer speed slows to the motherboard's capability, when data is passed through system bus. CPU can process data at a much faster rate by avoiding the system bus.)

- Cache memory is built into the processor, and may also be located next to it on a separate chip between the CPU and RAM. Cache built into the CPU is faster than separate cache, running at the speed of the microprocessor itself. However, separate cache is roughly twice as fast as RAM.
- The CPU has a built-in Level 1 (L1) cache and Level2 (L2) cache, as shown in Figure. In addition to the built-in L1 and L2 cache, some CPUs have a separate cache chip on the motherboard. This cache on the motherboard is called Level 3 (L3) cache. Nowadays, high-end processor comes with built-in L3 cache, like in Intel core i7. The L1, L2 and L3 cache store the most
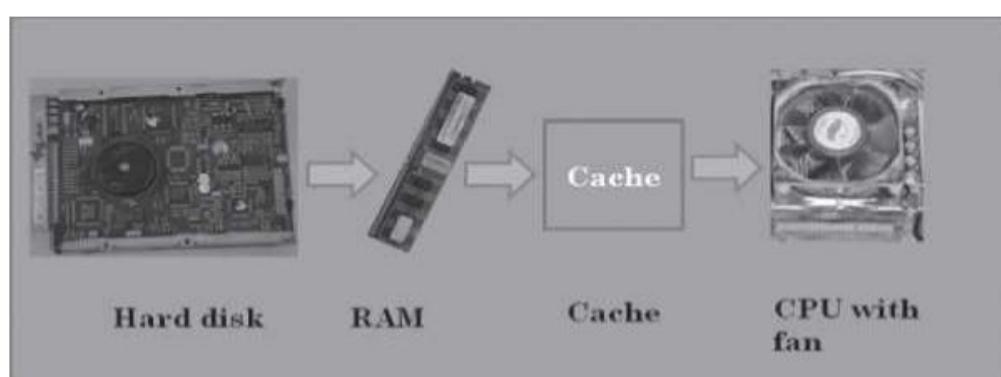
recently run instructions, the next ones and the possible ones, respectively. Typically, CPUs have cache size varying from 256KB (L1), 6 MB (L2), to 12MB (L3) cache.

- Cache memory is very expensive, so it is smaller in size. Generally, computers have cache memory of sizes 256 KB to 2 MB.

2.3.2 Primary Memory

- Primary memory is the main memory of computer. It is used to store data and instructions during the processing of data. Primary memory is semiconductor memory.
- Primary memory is of two kinds—Random Access Memory (RAM) and Read Only Memory (ROM).
- RAM is volatile. It stores data when the computer is on. The information stored in RAM gets erased when the computer is turned off. RAM provides temporary storage for data and instructions.
- ROM is non-volatile memory, but is a read only memory. The storage in ROM is permanent in nature, and is used for storing standard processing programs that permanently reside in the computer. ROM comes programmed by the manufacturer.
- RAM stores data and instructions during the execution of instructions. The data and instructions that require processing are brought into the RAM from the storage devices like hard disk. CPU accesses the data and the instructions from RAM, as it can access it at a fast speed than the storage devices connected to the input and output unit (Figure).
- The input data that is entered using the input unit is stored in RAM, to be made available during the processing of data. Similarly, the output data generated after processing is stored in RAM before being sent to the output device. Any intermediate results generated during the processing of program are stored in RAM.
- RAM provides a limited storage capacity, due to its high cost.

Hard disk   RAM   Cache   CPU with fan

2.3.3 Secondary Memory

- The secondary memory stores data and instructions permanently. The information can be stored in secondary memory for a long time (years), and is generally permanent in nature unless erased by the user. It is a non-volatile memory.

- It provides back-up storage for data and instructions. Hard disk drive, floppy drive and optical disk drives are some examples of storage devices.
- The data and instructions that are currently not being used by CPU, but may be required later for processing, are stored in secondary memory.
- Secondary memory has a high storage capacity than the primary memory.
- Secondary memory is also cheaper than the primary memory.
- It takes longer time to access the data and instructions stored in secondary memory than in primary memory.

Magnetic tape drives, disk drives and optical disk drives are the different types of storage devices.
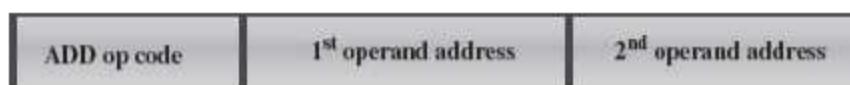
2.4 Instruction Format;

A computer program is a set of instructions that describe the steps to be performed for carrying out a computational task. The program and the data, on which the program operates, are stored in main memory, waiting to be processed by the processor. This is also called the stored program concept.

An instruction is designed to perform a task and is an elementary operation that the processor can accomplish. An instruction is divided into groups called fields. The common fields of an instruction are— Operation (op) code and Operand code (Figure). The remainder of the instruction fields differs from one computer type to other. The operation code represents action that the processor must execute. It tells the processor what basic operations to perform. The operand code defines the parameters of the action and depends on the operation. It specifies the locations of the data or the operand on which the operation is to be performed. It can be data or a memory address.

Operation code	Operand code

The number of bits in an instruction varies according to the type of data (could be between 8 and 32 bits). Figure shows the instruction format for ADD command.

ADD op code	1st operand address	2nd operand address

2.5 Instruction Set;

A processor has a set of instructions that it understands, called as instruction set. An instruction set or an instruction set architecture is a part of the computer architecture. It relates to programming, instructions, registers, addressing modes, memory architecture, etc. An Instruction Set is the set of all the basic operations that a processor can accomplish. Examples of some instructions are shown in Figure. The instructions in the instruction set are the language that a processor understands. All

programs have to communicate with the processor using these instructions. An instruction in the instruction set involves a series of logical operations (may be thousands) that are performed to complete each task. The instruction set is embedded in the processor (hardwired), which determines the machine language for the processor. All programs written in a high-level language are compiled and translated into machine code before execution, which is understood by the processor for which the program has been coded.

LOAD R1, A

ADD R1, B
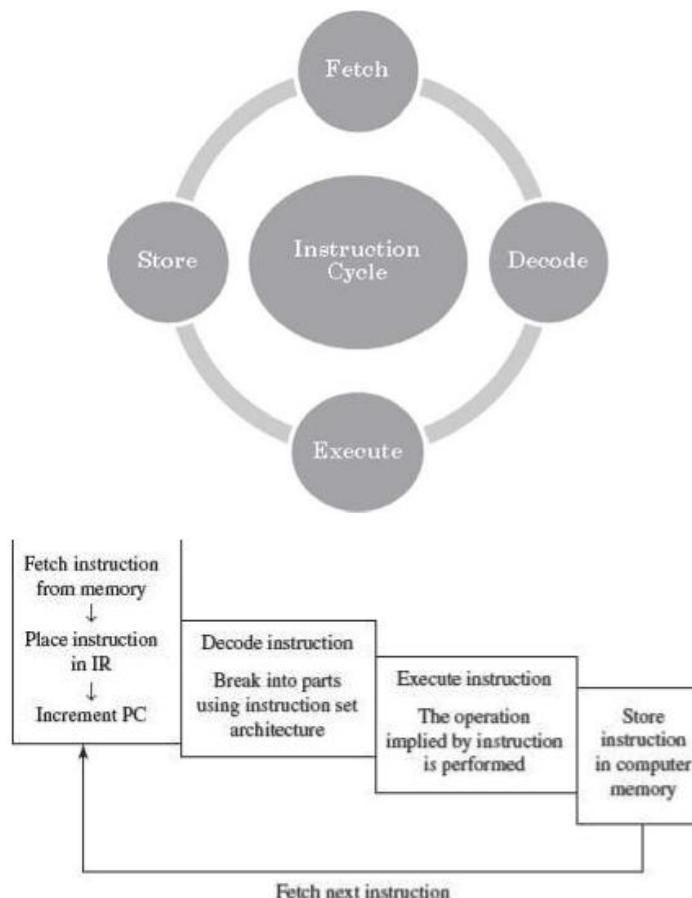
STORE R1, X

Two processors are different if they have different instruction sets. A program run on one computer may not run on another computer having a different processor. Two processors are compatible if the same machine level program can run on both the processors. Therefore, the system software is developed within the processor's instruction set.

2.6 Instruction Cycle;

The primary responsibility of a computer processor is to execute a sequential set of instructions that constitute a program. CPU executes each instruction in a series of steps, called instruction cycle (Figure).

- A instruction cycle involves four steps (Figure)—
  - *Fetching* The processor fetches the instruction from the memory. The fetched instruction is placed in the Instruction Register. Program Counter holds the address of next instruction to be fetched and is incremented after each fetch.
  - *Decoding* The instruction that is fetched is broken down into parts or decoded. The instruction is translated into commands so that they correspond to those in the CPU's instruction set. The instruction set architecture of the CPU defines the way in which an instruction is decoded.
  - *Executing* The decoded instruction or the command is executed. CPU performs the operation implied by the program instruction. For example, if it is an ADD instruction, addition is performed.
  - *Storing* CPU writes back the results of execution, to the computer's memory.

- Instructions are of different categories. Some categories of instructions are—
    - Memory access or transfer of data between registers.
    - Arithmetic operations like addition and subtraction.
    - Logic operations such as AND, OR and NOT.
    - Control the sequence, conditional connections, etc.

A CPU performance is measured by the number of instructions it executes in a second, i.e., MIPS (million instructions per second), or BIPS (billion instructions per second).

2.7 Microprocessor;

A processor's instruction set is a determining factor in its architecture. On the basis of the instruction set, microprocessors are classified as—Reduced Instruction Set Computer (RISC), and Complex Instruction Set Computer (CISC). The x86 instruction set of the original Intel 8086 processor is of the CISC type. The PCs are based on the x86 instruction set.

- **CISC** architecture hardwires the processor with complex instructions, which are difficult to create otherwise using basic instructions. CISC combines the different instructions into one single CPU.
    - CISC has a large instruction set that includes simple and fast instructions for performing basic tasks, as well as complex instructions that correspond to statements in the high level language.

- o An increased number of instructions (200 to 300) results in a much more complex processor, requiring millions of transistors.
- o Instructions are of variable lengths, using 8, 16 or 32 bits for storage. This results in the processor's time being spent in calculating where each instruction begins and ends.
- o With large number of application software programs being written for the processor, a new processor has to be backwards compatible to the older version of processors.
- o AMD and Cyrix are based on CISC.

- **RISC** has simple, single-cycle instructions, which performs only basic instructions. RISC architecture does not have hardwired advanced functions. All high-level language support is done in the software.
    - o RISC has fewer instructions and requires fewer transistors, which results in the reduced manufacturing cost of processor.
    - o The instruction size is fixed (32 bits). The processor need not spend time in finding out where each instruction begins and ends.
    - o RISC architecture has a reduced production cost compared to CISC processors.
    - o The instructions, simple in nature, are executed in just one clock cycle, which speeds up the program execution when compared to CISC processors.
    - o RISC processors can handle multiple instructions simultaneously by processing them in parallel.
    - o Apple Mac G3 and PowerPC are based on RISC.

Processors like Athlon XP and Pentium IV use a hybrid of both technologies.

Pipelining improves instruction execution speed by putting the execution steps into parallel. A CPU can receive a single instruction, begin executing it, and receive another instruction before it has completed the first. This allows for more instructions to be performed, about, one instruction per clock cycle.
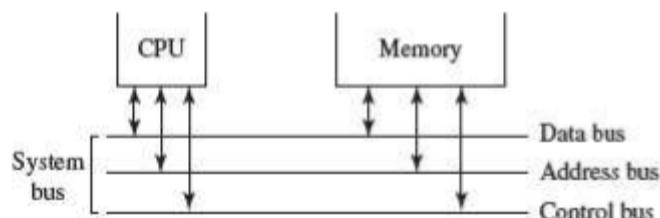
Parallel Processing is the simultaneous execution of instructions from the same program on different processors. A program is divided into multiple processes that are handled in parallel in order to reduce execution time.

2.8 Interconnecting the Units of a Computer;

CPU sends data, instructions and information to the components inside the computer as well as to the peripherals and devices attached to it. Bus is a set of electronic signal pathways that allows information and signals to travel between components inside or outside of a computer. The different components of computer, i.e., CPU, I/O unit, and memory unit are connected with each other by a bus. The data, instructions and the signals are carried between the different components via a bus. The features and functionality of a bus are as follows—

- A bus is a set of wires used for interconnection, where each wire can carry one bit of data.
- A bus width is defined by the number of wires in the bus.
- A computer bus can be divided into two types—Internal Bus and External Bus.
- The Internal Bus connects components inside the motherboard like, CPU and system memory. It is also called the System Bus. Figure shows interaction between processor and memory.
- The External Bus connects the different external devices, peripherals, expansion slots, I/O ports and drive connections to the rest of computer. The external bus allows various devices to be attached to the computer. It allows for the expansion of computer's capabilities. It is generally slower than the system bus. It is also referred to as the Expansion Bus.
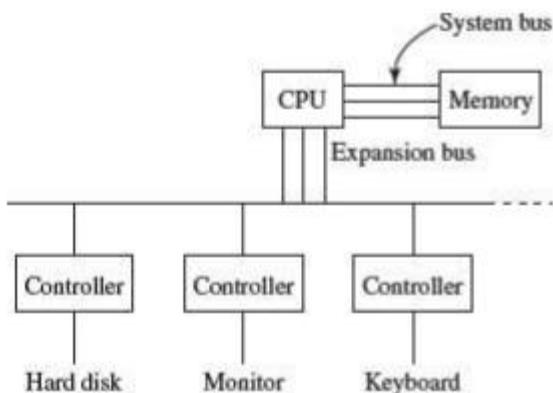
- A system bus or expansion bus comprise of three kinds of buses — data bus, address bus and control bus.
- The interaction of CPU with memory and I/O devices involves all the three buses.
  - The command to access the memory or the I/O device is carried by the control bus.
  - The address of I/O device or memory is carried by the address bus.
  - The data to be transferred is carried by the data bus.

2.8.1 System Bus

The functions of data bus, address bus and control bus, in the system bus, are as follows—

- *Data Bus* transfers data between the CPU and memory. The bus width of a data bus affects the speed of computer. The size of data bus defines the size of the processor. A processor can be 8, 16, 32 or 64-bit processor. An 8–bit processor has 8 wire data bus to carry 1 byte of data. In a 16–bit processor, 16–wire bus can carry 16 bits of data, i.e., transfer 2 bytes, etc.

- ***Address Bus*** connects CPU and RAM with set of wires similar to data bus. The width of address bus determines the maximum number of memory locations the computer can address. Currently, Pentium Pro, II, III, IV have 36–bit address bus that can address 236 bytes or 64 GB of memory.
- ***Control Bus*** specifies whether data is to be read or written to the memory, etc.

2.8.2 Expansion Bus

The functions of data bus, address bus and control bus, in the expansion bus, are as follows—

- The expansion bus connects external devices to the rest of computer. The external devices like monitor, keyboard and printer connect to ports on the back of computer. These ports are actually a part of the small circuit board or expansion card that fits into an expansion slot on the motherboard. Expansion slots are easy to recognize on the motherboard.
- Expansion slots make up a row of long plastic connectors at the back of the computer with tiny copper 'finger slots' in a narrow channel that grab the connectors on the expansion cards. The slots are attached to tiny copper pathways on the motherboard (the expansion bus), which allows the device to communicate with the rest of computer.
- ***Data Bus*** is used to transfer data between I/O devices and CPU. The exchange of data between CPU and I/O devices is according to the industry standard data buses. The most commonly used standard is Extended Industry Standard Architecture (EISA) which is a 32-bit bus architecture. Some of the common bus technologies are—
  - o Peripheral Component Interconnect (PCI) bus for hard disks, sound cards, network cards and graphics cards,
  - o Accelerated Graphics Port (AGP) bus for 3–D and full motion video,
  - o Universal Serial Bus (USB) to connect and disconnect different devices.
- ***Address Bus*** carries the addresses of different I/O devices to be accessed like the hard disk, CD ROM, etc.
- ***Control Bus*** is used to carry read/write commands, status of I/O devices, etc.
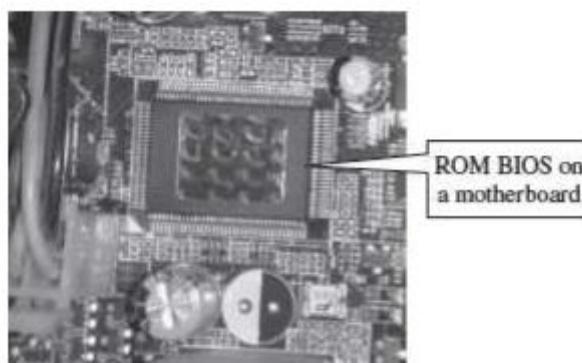
2.9 Inside a Computer Cabinet

The computer cabinet encloses the components that are required for the running of the computer. The components inside a computer cabinet include the power supply, motherboard, memory chips, expansion slots, ports and interface, processor, cables and storage devices.

2.9.1 Motherboard

The computer is built up around a motherboard. The motherboard is the most important component in the PC. It is a large Printed Circuit Board (PCB), having many chips, connectors and other electronics mounted on it. The motherboard is the hub, which is used to connect all the essential components of a computer. The RAM, hard drive, disk drives and optical drives are all plugged into interfaces on the motherboard. The motherboard contains the processor, memory chips, interfaces and sockets, etc.

The motherboard may be characterized by the form factor, chipset and type of processor socket used. Form factor refers to the motherboard's geometry, dimensions, arrangement and electrical requirements. Different standards have been developed to build motherboards, which can be used in different brands of cases. Advanced Technology Extended (ATX) is the most common design of motherboard for desktop computers. Chipset is a circuit, which controls the majority of resources (including the bus interface with the processor, cache memory and RAM, expansion cards, etc.) Chipset's job is to coordinate data transfers between the various components of the computer (including the processor and memory). As the chipset is integrated into the motherboard, it is important to choose a motherboard, which includes a recent chipset, in order to maximize the computer's upgradeability. The processor socket may be a rectangular connector into which the processor is mounted vertically (slot), or a square-shaped connector with many small connectors into which the processor is directly inserted (socket). The Basic Input Output System (BIOS) and Complementary Metal-Oxide Semiconductor (CMOS) are present on the motherboard

ROM BIOS on a motherboard
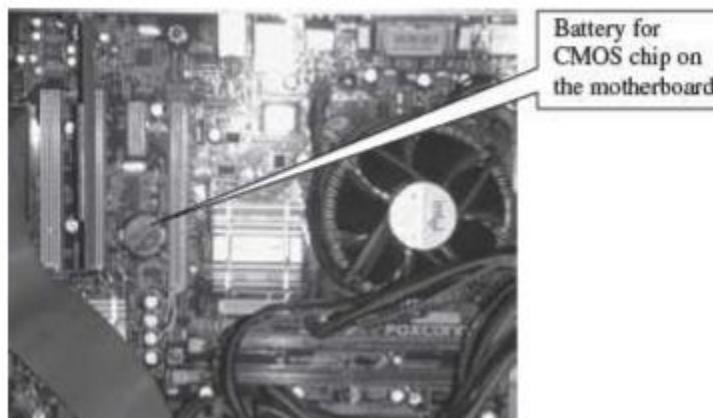
- *BIOS* It is the basic program used as an interface between the operating system and the motherboard. The BIOS (Figure) is stored in the ROM and cannot be rewritten. When the computer is switched on, it needs instructions to start. BIOS contain the instructions for the

starting up of the computer. The BIOS runs when the computer is switched on. It performs a Power On Self Test (POST) that checks that the hardware is functioning properly and the hardware devices are present. It checks whether the operating system is present on the hard drive. BIOS invokes the bootstrap loader to load the operating system into memory. BIOS can be configured using an interface named BIOS setup, which can be accessed when the computer is booting up (by pressing the DEL key).
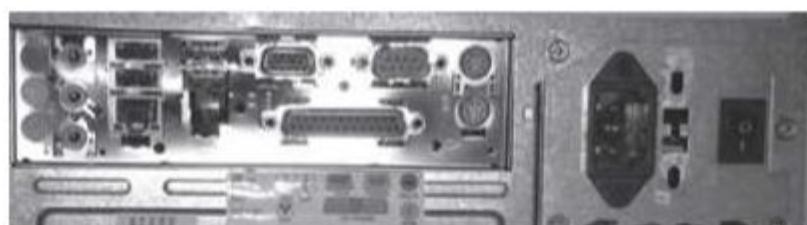
- **CMOS Chip** BIOS ROMs are accompanied by a smaller CMOS (CMOS is a type of memory technology) memory chip. When the computer is turned off, the power supply stops providing electricity to the motherboard. When the computer is turned on again, the system still displays the correct clock time. This is because the CMOS chip saves some system information, such as time, system date and essential system settings. CMOS is kept powered by a button battery located on the motherboard (Figure). The CMOS chip is working even when the computer power is switched off. Information of the hardware installed in the computer (such as the number of tracks or sectors on each hard drive) is stored in the CMOS chip.

2.9.2 Ports and Interfaces

Motherboard has a certain number of I/O sockets that are connected to the ports and interfaces found on the rear side of a computer (Figure 2.16). You can connect external devices to the ports and interfaces, which get connected to the computer's motherboard.

- Serial Port— to connect old peripherals.
- Parallel Port— to connect old printers.
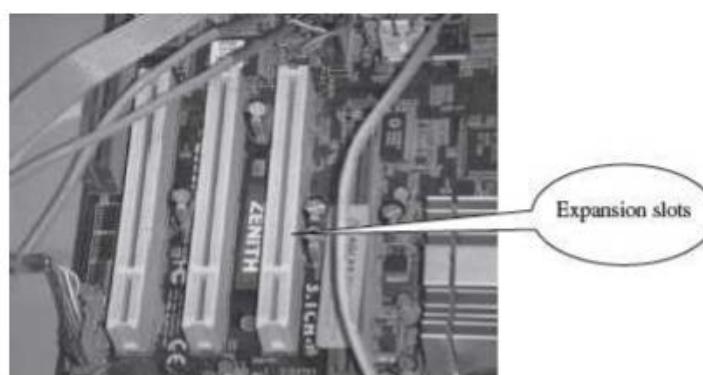
- USB Ports—to connect newer peripherals like cameras, scanners and printers to the computer. It uses a thin wire to connect to the devices, and many devices can share that wire simultaneously.
- Firewire is another bus, used today mostly for video cameras and external hard drives.
- RJ45 connector (called LAN or Ethernet port) is used to connect the computer to a network. It corresponds to a network card integrated into the motherboard.
- VGA connector for connecting a monitor. This connector interfaces with the built-in graphics card.
- Audio plugs (line-in, line-out and microphone), for connecting sound speakers and the microphone. This connector interfaces with the built-in sound card.
- PS/2 port to connect mouse and keyboard into PC.
- SCSI port for connecting the hard disk drives and network connectors.

2.9.3 Expansion Slots

The expansion slots (Figure) are located on the motherboard. The expansion cards are inserted in the expansion slots. These cards give the computer new features or increased performance. There are several types of slots:

- ISA (Industry Standard Architecture) slot—To connect modem and input devices.
- PCI (Peripheral Component InterConnect) slot—To connect audio, video and graphics. They are much faster than ISA cards.
- AGP (Accelerated Graphic Port) slot—A fast port for a graphics card.
- PCI (Peripheral Component InterConnect) Express slot—Faster bus architecture than AGP and PCI buses.
- PC Card—It is used in laptop computers. It includes Wi-Fi card, network card and external modem.

2.9.4 Ribbon Cables

Ribbon cables (Figure) are flat, insulated and consist of several tiny wires moulded together that carry data to different components on the motherboard. There is a wire for each bit of the word or byte and additional wires to coordinate the activity of moving information. They also connect the floppy drives, disk drives and CD-ROM drives to the connectors in the motherboard. Nowadays, Serial Advanced
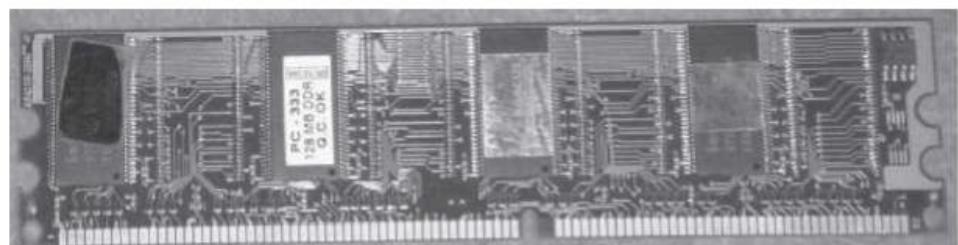
Technology Attachment (SATA) cables have replaced the ribbon cables to connect the drives to the motherboard.
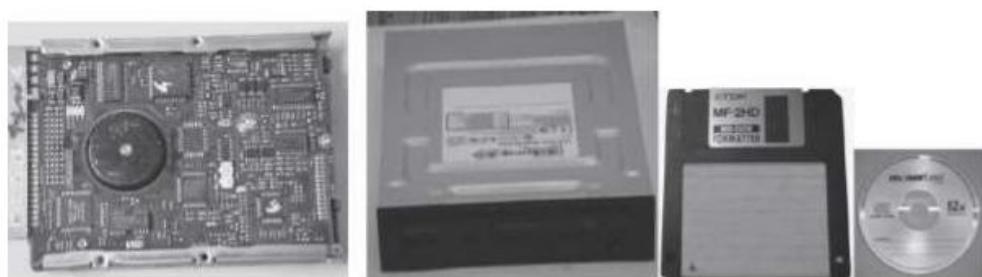
2.9.5 Memory Chips

The RAM consists of chips on a small circuit board (Figure). Two types of memory chips— Single In-line Memory Module (SIMM) and Dual In-line Memory Module (DIMM) are used in desktop computers. The CPU can retrieve information from DIMM chip at 64 bits compared to 32 bits or 16 bits transfer with SIMM chips. DIMM chips are used in Pentium 4 onwards to increase the access speed.

2.9.6 Storage Devices

The disk drives are present inside the machine. The common disk drives in a machine are hard disk drive, floppy drive and CD drive or DVD drive. High-storage devices like hard disk, floppy disk and CDs are inserted into the hard disk drive, floppy drive and CD drive, respectively. These storage devices can store large amounts of data, permanently.

2.9.7 Processor

The processor or the CPU is the main component of the computer. Select a processor based on factors like its speed, performance, reliability and motherboard support. Pentium Pro, Pentium 2 and Pentium 4 are some of the processors.

Unit 3: Computer Memory

3.1 Introduction;

The computer's memory stores data, instructions required during the processing of data, and output results. Storage may be required for a limited period of time, instantly, or, for an extended period of time. Different types of memories, each having its own unique features, are available for use in a computer. The cache memory, registers, and RAM are fast memories and store the data and instructions temporarily during the processing of data and instructions. The secondary memory like magnetic disks and optical disks have large storage capacities and store the data and instructions permanently, but are slow memory devices. The memories are organized in the computer in a manner to achieve high levels of performance at the minimum cost.

3.2 Memory Representation;

The computer memory stores different kinds of data like input data, output data, intermediate results, etc., and the instructions. Binary digit or bit is the basic unit of memory. A bit is a single binary digit, i.e., 0 or 1. A bit is the smallest unit of representation of data in a computer. However, the data is handled by the computer as a combination of bits. A group of 8 bits form a byte. One byte is the smallest unit of data that is handled by the computer. One byte can store 28, i.e., 256 different combinations of bits, and thus can be used to represent 256 different symbols. In a byte, the different combinations of bits fall in the range 00000000 to 11111111. A group of bytes can be further combined to form a word. A word can be a group of 2, 4 or 8 bytes.

1 bit = 0 or 1
1 Byte (B) = 8 bits
1 Kilobyte (KB) = 2^{10} = 1024 bytes
1 Megabyte (MB) = 2^{20} = 1024KB
1 Gigabyte (GB) = 2^{30} = 1024 MB = 1024 *1024 KB
1 Terabyte (TB) = 2^{40}= 1024 GB = 1024 * 1024 *1024 KB
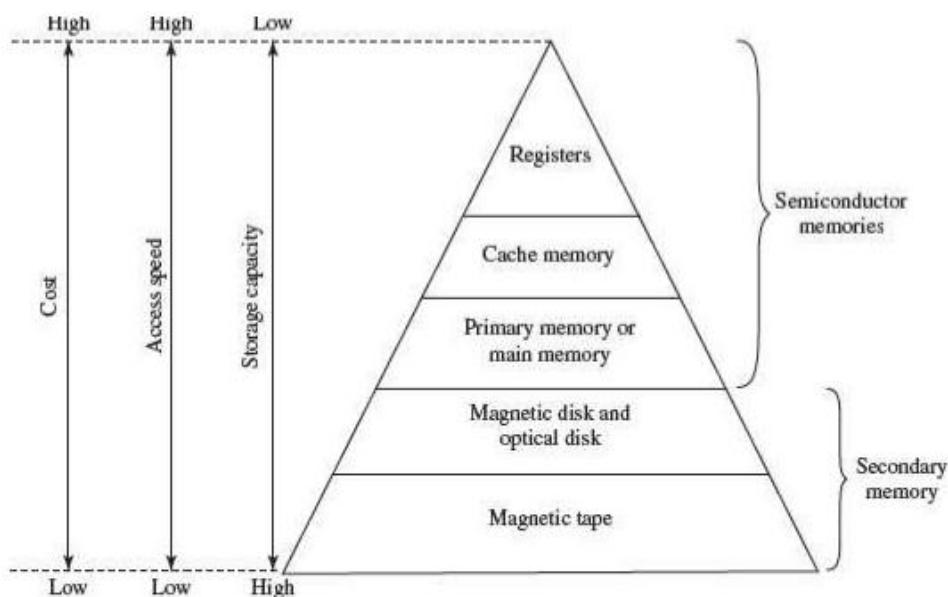
3.3 Memory Hierarchy;

The memory is characterized on the basis of two key factors—capacity and access time. Capacity is the amount of information (in bits) that a memory can store. Access time is the time interval between the read/ write request and the availability of data. The lesser the access time, the faster is the speed of memory. Ideally, we want the memory with fastest speed and largest capacity. However, the cost of fast memory is very high. The computer uses a hierarchy of memory that is organized in a manner to enable the fastest speed and largest capacity of memory. The hierarchy of the different memory types is shown in Figure.
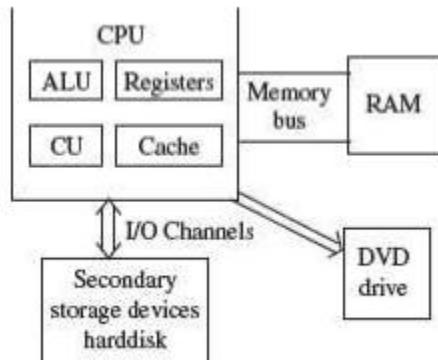
The internal memory and external memory are the two broad categories of memory used in the computer. The internal memory consists of the CPU registers, cache memory and primary memory. The internal memory is used by the CPU to perform the computing tasks. The external memory is also called the secondary memory. The secondary memory is used to store the large amount of data and the software .

In general, referring to the computer memory usually means the internal memory.

- *Internal Memory*—The key features of internal memory are—(1) limited storage capacity, (2) temporary storage, (3) fast access, and (4) high cost. Registers, cache memory, and primary memory constitute the internal memory. The primary memory is further of two kinds—RAM and ROM. Registers are the fastest and the most expensive among all the memory types. The registers are located inside the CPU, and are directly accessible by the CPU. The speed of registers is between 1—2 ns (nanosecond). The sum of the size of registers is about 200B. Cache memory is next in the hierarchy and is placed between the CPU and the main memory. The speed of cache is between 2—10 ns. The cache size varies between 32 KB to 4MB. Any program or data that has to be executed must be brought into RAM from the secondary memory. Primary memory is relatively slower than the cache memory. The speed of RAM is around 60ns. The RAM size varies from 512KB to 3GB.
- *Secondary Memory*—The key features of secondary memory storage devices are—(1) very high storage capacity, (2) permanent storage (non-volatile), unless erased by user, (3) relatively slower access, (4) stores data and instructions that are not currently being used by CPU but may be required later for processing, and (5) cheapest among all memory. The storage devices consist of two parts—drive and device. For example, magnetic tape drive and magnetic tape, magnetic disk drive and disk, and, optical disk drive and disk. The speed of magnetic disk is around 60ms. The capacity of a hard disk ranges from 160 GB

to 1,600 GB (1.6 Tera Bytes). Figure shows the interaction between CPU and memory.

- To get the fastest speed of memory with largest capacity and least cost, the fast memory is located close to the processor. The secondary memory, which is not as fast, is used to store information permanently, and is placed farthest from the processor. With respect to CPU, the memory is organized as follows—
  - Registers are placed inside the CPU (small capacity, high cost, very high speed)
  - Cache memory is placed next in the hierarchy (inside and outside the CPU)
  - Primary memory is placed next in the hierarchy
  - Secondary memory is the farthest from CPU (large capacity, low cost, low speed)

The speed of memories is dependent on the kind of technology used for the memory. The registers, cache memory and primary memory are semiconductor memories. They do not have any moving parts and are fast memories. The secondary memory is magnetic or optical memory, has moving parts and has slow speed.

3.4 CPU Registers;

Registers are very high-speed storage areas located inside the CPU. After CPU gets the data and instructions from the cache or RAM, the data and instructions are moved to the registers for processing. Registers are manipulated directly by the control unit of CPU during instruction execution. That is why registers are often referred to as the CPU's working memory. Since CPU uses registers for the processing of data, the number of registers in a CPU and the size of each register affect the power and speed of a CPU. The more the number of registers (ten to hundreds) and bigger the size of each register (8 bits to 64 bits), the better it is.

3.5 Cache Memory;

Cache memory is placed in between the CPU and the RAM. Cache memory is a fast memory, faster than the RAM. When the CPU needs an instruction or data during processing, it first looks in the cache. If the information is present in the cache, it is called a cache hit, and the data or instruction is retrieved from the cache. If the information is not present in cache, then it is called a cache miss and the information is then retrieved from RAM. The content of cache is decided by the cache controller (a circuit on the motherboard). The most recently accessed information or instructions help the controller to guess the RAM locations that may be accessed next. To get good system performance, the number of hits must far outnumber the misses. The two main factors that affect the performance of cache are its size and level (L1, L2 and L3).

3.6 Primary Memory;

Primary memory is the main memory of computer. It is a chip mounted on the motherboard of computer. Primary memory is categorized into two main types
- Random Access Memory (RAM), and
- Read Only Memory (ROM)

RAM is used for the temporary storage of input data, output data and intermediate results. The input data entered into the computer using the input device, is stored in RAM for processing. After processing, the output data is stored in RAM before being sent to the output device. Any intermediate results generated during the processing of program are also stored in RAM. Unlike RAM, the data once stored in ROM either cannot be changed or can only be changed using some special operations. Therefore, ROM is used to store the data that does not require a change. Flash memory is another form of rewritable read-only memory that is compact, portable, and requires little energy.

3.6.1 Random Access Memory
- RAM is used to store data and instructions during the operation of computer.
  - The data and instructions that need to be operated upon by CPU are first brought to RAM from the secondary storage devices like the hard disk.
  - CPU interacts with RAM to get the data and instructions for processing.
- RAM loses information when the computer is powered off. It is a volatile memory. When the power is turned on, again, all files that are required by the CPU are loaded from the hard disk to RAM. Since RAM
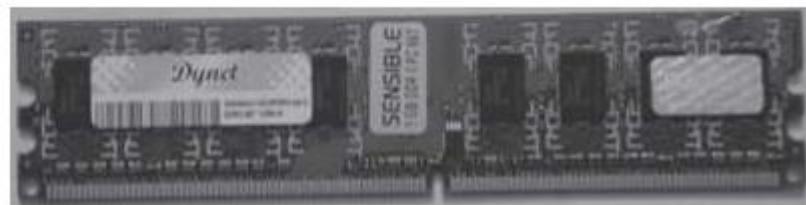
is a volatile memory, any information that needs to be saved for a longer duration of time must not be stored in RAM.

- RAM provides random access to the stored bytes, words, or larger data units. This means that it requires same amount of time to access information from RAM, irrespective of where it is located in it.
- RAM can be read from and written to with the same speed.
- The size ofRAM is limited due to its high cost. The size of RAM is measured in MB or GB.
- The performance of RAM is affected by—
    - Access speed (how quickly information can be retrieved). The speed of RAM is expressed in nanoseconds.
    - Data transfer unit size (how much information can be retrieved in one request).
- RAM affects the speed and power of a computer. More the RAM, the better it is. Nowadays, computers generally have 512 MB to 4 GB of RAM.
- RAM is a microchip implemented using semiconductors.
- There are two categories of RAM, depending on the technology used to construct a RAM— (1) Dynamic RAM (DRAM), and (2) Static RAM (SRAM).
- *DRAM* is the most common type of memory chip. DRAM is mostly used as main memory since it is small and cheap.
    - It uses transistors and capacitors. The transistors are arranged in a matrix of rows and columns. The capacitor holds the bit of information 0 and 1. The transistor and capacitor are paired to make a memory cell. The transistor acts as a switch that lets the control circuitry on the memory chip read the capacitor or change its state.
    - DRAM must be refreshed continually to store information. For this, a memory controller is used. The memory controller recharges all the capacitors holding a 1 before they discharge. To do this, the memory controller reads the memory and then writes it right back.
    - DRAM gets its name from the refresh operation that it requires to store the information; otherwise it will lose what it is holding. The refresh operation occurs automatically thousands of times per second. DRAM is slow because the refreshing takes time.
    - Access speed of DRAM ranges from 50 to 150 ns.
- *SRAM* chip is usually used in cache memory due to its high speed.
    - SRAM uses multiple transistors (four to six), for each memory cell. It does not have a capacitor in each cell.
    - A SRAM memory cell has more parts so it takes more space on a chip than DRAM cell.
    - It does not need constant refreshing and therefore is faster than DRAM.

- SRAM is more expensive than DRAM, and it takes up more space.
- It stores information as long as it is supplied with power.
- SRAM are easier to use and very fast. The access speed of SRAM ranges from 2– 10 nanosecond.

- The memory chips (Figure4) are available on a separate Printed Circuit Board (PCB) that is plugged into a special connector on the motherboard. Memory chips are generally available as part of a card called a memory module. There are generally two types of RAM modules—Single Inline Memory Module (SIMM) and Dual Inline Memory Module (DIMM).

- SIMM modules have memory chip on one side of the PCB. SIMM modules can store 8 bits to 32 bits of data simultaneously.
- DIMM modules have memory chips on both sides of the PCB. DIMM format are 64–bit memories. Smaller modules known as Small Outline DIMM (SO DIMM) are designed for portable computers. SO DIMM modules have 32–bit memory.

3.6.2 Read Only Memory

ROM is a non-volatile primary memory. It does not lose its content when the power is switched off. The features of ROM are described as follows—

- ROM, as the name implies, has only read capability and no write capability. After the information is stored in ROM, it is permanent and cannot be corrected.
- ROM comes programmed by the manufacturer. It stores standard processing programs that permanently reside in the computer. ROM stores the data needed for the start up of the computer. The instructions that are required for initializing the devices attached to a computer are stored in ROM.
- The ROM memory chip (Figure) stores the Basic Input Output System (BIOS). BIOS provides the processor with the information required to boot the system. It provides the system with the settings and resources that are available on the system. BIOS is a permanent part of the computer. It does not load from disk but instead is stored in a ROM memory chip. The program code in the BIOS differs from ordinary software since it acts as an integral part of the computer. When the computer is turned on, the BIOS does the following things—
  - ***Power On Self Test (POST)*** is a program that runs automatically when the system is booted. BIOS performs the

power-on self-test. It checks that the major hardware components are working properly.
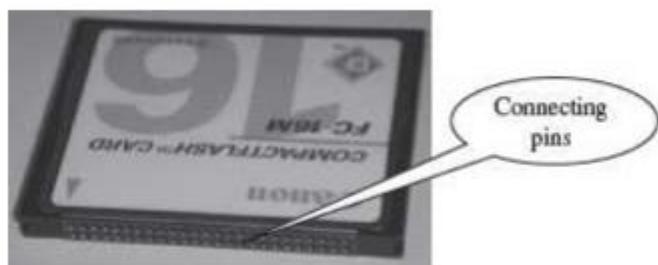
- **BIOS setup program**, which is a built-in utility in BIOS, lets the user set the many functions that control how the computer works. BIOS displays the system settings and finds the bootable devices. It loads the interrupt handlers and device drivers. It also initializes the registers.

- **Bootstrap Loader** is a program whose purpose is to start the computer software for operation when the power is turned on. It loads the operating system into RAM and launches it. It generally seeks the operating system on the hard disk. The bootstrap loader resides in the ROM. The BIOS initiates the bootstrap sequence.

- ROMs are of different kinds. They have evolved from the fixed read only memory to the ones that can be programmed and re-programmed. They vary in the number of re-writes and the method used for the re-writing. Programmable ROM (PROM), Erasable Programmable ROM (EPROM) and Electrically Erasable Programmable ROM (EEPROM) are some of the ROMs. All the different kinds of ROM retain their content when the power is turned off. o
    - **PROM** can be programmed with a special tool, but after it has been programmed the contents cannot be changed. PROM memories have thousands of fuses (or diodes). High voltage (12 V) is applied to the fuses to be burnt. The burnt fuses correspond to 0 and the others to 1.
    - **EPROM** can be programmed in a similar way as PROM, but it can be erased by exposing it to ultra violet light and re-programmed. EPROM chips have to be removed from the computer for re-writing. o EEPROM memories can be erased by electric charge and re-programmed. EEPROM chips do not have to be removed from the computer for re-writing.

- ***Flash Memory*** is a kind of semiconductor-based non-volatile, rewritable computer memory that can be electrically erased and reprogrammed (Figure). It is a specific type of EEPROM.

- It combines the features of RAM and ROM. It is a random access memory and its content can be stored in it at any time. However, like ROM, the data is not lost when the machine is turned off or the electric power is cut. Flash memory stores bits of data in memory cells.
- Flash memories are high-speed memories, durable, and have low-energy consumption. Since flash memory has no moving part, it is very shock-resistant. Due to these features, flash memory is used in devices such as digital camera, mobile phone, printer, laptop computer, and record and play back sound devices, such as MP3 players.

3.7 Secondary Memory;

In the previous section, we saw that RAM is expensive and has a limited storage capacity. Since it is a volatile memory, it cannot retain information after the computer is powered off. Thus, in addition to primary memory, an auxiliary or secondary memory is required by a computer. The secondary memory is also called the storage device of computer. In this chapter, the terms secondary memory and storage device are used interchangeably. In comparison to the primary memory, the secondary memory stores much larger amounts of data and information (for example, an entire software program) for extended periods of time. The data and instructions stored in secondary memory must be fetched into RAM before processing is done by CPU.

Magnetic tape drives, magnetic disk drives, optical disk drives and magneto-optical disk drives are the different types of storage devices.

3.8 Access Types of Storage Devices;

The information stored in storage devices can be accessed in two ways—
1. Sequential access
2. Direct access

3.8.1 Sequential Access Devices
Sequential access means that computer must run through the data in sequence, starting from the beginning, in order to locate a particular piece of data. Magnetic tape is an example of sequential access device. Let us suppose that magnetic tape

consists of 80 records. To access the 25th record, the computer starts from first record, then reaches second, third etc. until it reaches the 25th record. Sequential access devices are generally slow devices.

3.8.2 Direct Access Devices

Direct access devices are the ones in which any piece of data can be retrieved in a non-sequential manner by locating it using the data's address. It accesses the data directly, from a desired location. Magnetic disks and optical disks are examples of direct access devices. There is no predefined order in which one can read and write data from a direct access device. In a magnetic disk consisting of 80 records, to access the 25th record, the computer can directly access the 25th record, without going past the first 24 records. Based on access, magnetic tapes are sequential access devices, and, magnetic disks, optical disk and magneto-optical disks are direct access devices.

3.9 Magnetic Tape;

Magnetic tape is a plastic tape with magnetic coating (Figure). It is storage medium on a large open reel or in a smaller cartridge or cassette (like a music cassette). Magnetic tapes are cheaper storage media. They are durable, can be written, erased, and re-written. Magnetic tapes are sequential access devices, which mean that the tape needs to rewind or move forward to the location where the requested data is positioned in the magnetic tape. Due to their sequential nature, magnetic tapes are not suitable for data files that need to be revised or updated often. They are generally used to store back-up data that is not frequently used or to transfer data from one system to other.

The features of magnetic tape are—
- Inexpensive storage device
- Can store a large amount of data
- Easy to carry or transport
- Not suitable for random access data
- Slow access device
- Needs dust prevention, as dust can harm the tape
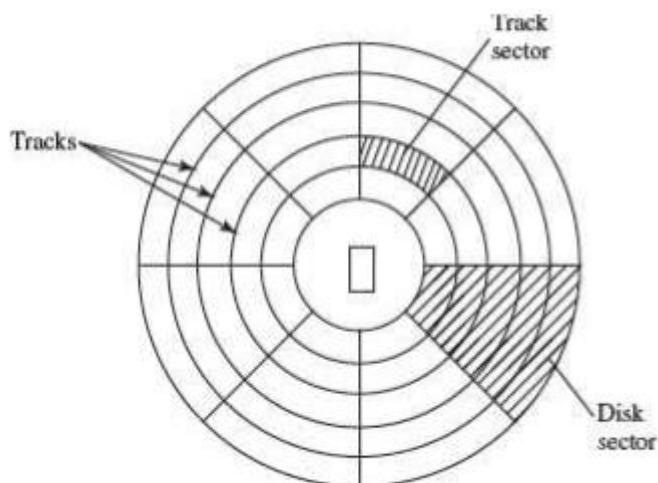
- Suitable for back-up storage or archiving

3.10 Magnetic Disk;

Magnetic disk is a direct access secondary storage device. It is a thin plastic or metallic circular plate coated with magnetic oxide and encased in a protective cover. Data is stored on magnetic disks as magnetized spots. The presence of a magnetic spot represents the bit 1 and its absence represents the bit 0.

The working of magnetic disk is explained as follows—

- The surface of disk is divided into concentric circles known as tracks. The outermost track is numbered 0 and the innermost track is the last track. Tracks are further divided into sectors. A sector is a pie slice that cuts across all tracks. The data on disk is stored in sector. Sector is the smallest unit that can be read or written on a disk. A disk has eight or more sectors per track (Figure).

- Magnetic disk is inserted into a magnetic disk drive for access. The drive consists of a read/write head that is attached to a disk arm, which moves the head. The disk arm can move inward and outward on the disk.
- During reading or writing to disk, the motor of disk drive moves the disk at high speed (60–150 times/sec.)
- Accessing data on the disk requires the following—
  - The read/write head is positioned to the desired track where the data is to be read from or written to. The time taken to move the read/write head to the desired track is called the seek time.
  - Once the read/write head is at the right track, then the head waits for right sector to come under it (disk is moving at high speed). The time taken for desired sector of the track to come under read/write head is called the latency time.
  - Once the read/write head is positioned at the right track and sector, the data has to be written to disk or read from disk. The rate at which data is written to disk or read from disk is called data transfer rate.
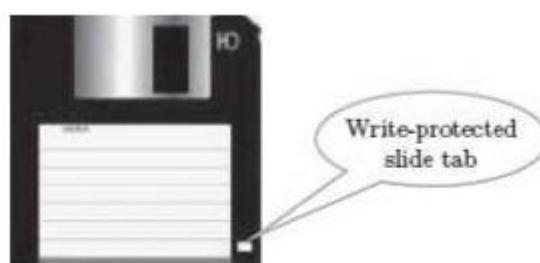
- The sum of seek time, latency time and time for data transfer is the access time of the disk.
- The storage capacity of disk drive is measured in gigabytes (GB).
- Large disk storage is created by stacking together multiple disks. A set of same tracks on all disks forms a cylinder. Each disk has its own read/write head which work in coordination.
- A disk can also have tracks and sectors on both sides. Such a disk is called double-sided disk.

The features of magnetic disk are—
- Cheap storage device
- Can store a large amount of data
- Easy to carry or transport
- Suitable for frequently read/write data
- Fast access device
- More reliable storage device
- To be prevented from dust, as the read/write head flies over the disk. Any dust particle in between can corrupt the disk.

3.10.1 Floppy Disk
- Floppy disk (FD) is a flat, round, single disk made of Mylar plastic and enclosed in square plastic jacket (Figure).
- Floppy Disk Drive (FDD) is the disk drive for floppy disk.
- The floppy disk is inserted into the floppy disk drive to read or write data to it.
- Floppy disk has a write-protect slide tab that prevents a user from writing to it.
- A floppy disk may be single-sided or double-sided disk, i.e., data can be read and written on one and both sides of floppy disk, respectively.
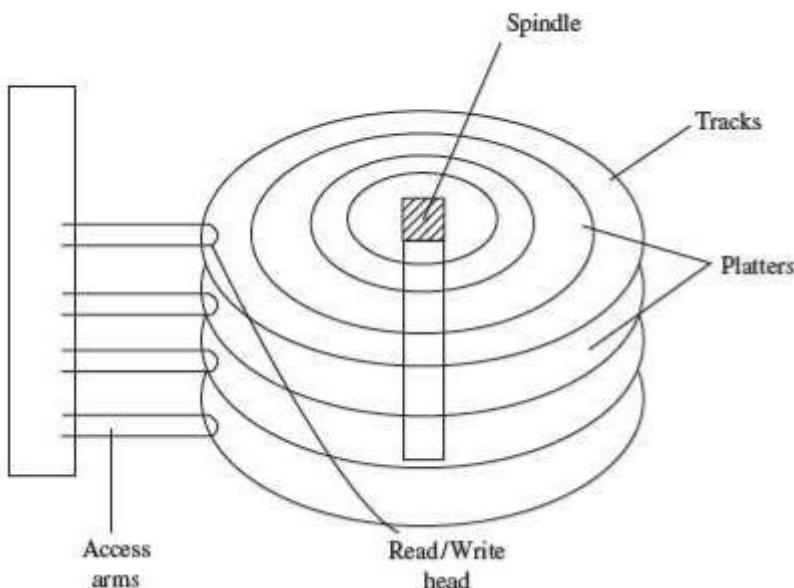
- They are portable. They can be removed from the disk drive, carried or stored separately.
- They are small and inexpensive.
- Floppy disks are slower to access than hard disk. They have less storage capacity and are less expensive than hard disk.
- They come in two basic sizes—5-¼ inch and 3-½ inch.
- The 5-¼ inch disk came around 1987. It can store 360 KB to 1.2 MB of data.

- The 3-½ inch disk has capacity of 400 KB to 1.44 MB. It usually contains 40 tracks and 18 sectors per track and can store 512 bytes per sector.

3.10.2 Hard Disk

- A hard disk (HD) consists of one or more platters divided into concentric tracks and sectors. It is mounted on a central spindle, like a stack. It can be read by a read/write head that pivots across the rotating disks. The data is stored on the platters covered with magnetic coating.

- Hard disk is a fixed disk. The disk is not removable from the drive, unlike floppy disk.
- The hard disk and Hard Disk Drive (HDD) is a single unit.
- Hard disk can store much more data than floppy disk. The data in hard disk are packed more closely (because fast spinning uses smaller magnetic charges) and they have multiple platters, with data being stored on both sides of each platter. Large capacity hard disks may have 12 or more platters.
- Unlike floppy disk, the read/write head of hard disk does not touch the disk during accessing.
- Hard disk can spin at the speed of up to 10,000 revolutions per minute and have an access time of 9—14 ms. It stores 512 bytes per sector but the number of sectors are more per track (54 or more) than floppy disk.
- Nowadays, hard disks are available that can store up to 500 GB of data. Generally, PCs come with 160 GB hard disk.
- Hard disk is the key secondary storage device of computer. The operating system is stored on the hard disk. The performance of computer like speed of computer boot up, loading of programs to

primary memory, loading of large files like images, video, audio etc., is also dependent on the hard disk.

- Nowadays, portable external hard disk drive is available which can be attached to the USB drive of the computer. They come in the storage capacities of 80 GB to 500 GB.
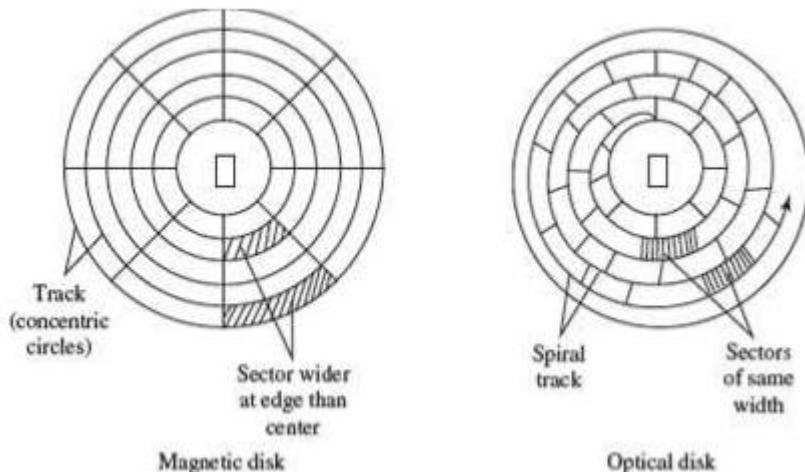
3.10.3 Zip Disk

- They are high-capacity removable disk and drive.
- They have the speed and capacity of hard disk and portability of floppy disk.
- Zip disk are of the same size as floppy disk, i.e., 3–½ inch but have a much higher capacity than the floppy disk.
- Zip disk and drive were made by Iomega Corp. It comes as a complete unit—disk, drive, connection cable, power cord and operating system. It can be connected to the computer system externally using a parallel chord or SCSI cable.
- Their capacity ranges from 100 MB to 750 MB. They can be used to store large files, audio and video data.

3.11 Optical Disk;

Optical disk (Figure) is a flat and circular disk which is coated with reflective plastic material that can be altered by laser light. Optical disk does not use magnetism. The bits 1 and 0 are stored as spots that are relatively bright and light, respectively.

- An optical disk consists of a single spiral track that starts from the edge to the centre of disk. Due to its spiral shape, it can access large amount of data sequentially, for example music and video. The random access on optical disk is slower than that of magnetic disk, due to its spiral shape.
- The tracks on optical disk are further divided into sectors which are of same length. Thus, the sectors near the centre of disk wrap around the disk longer than the sectors on the edges of disk. Reading the disk thus requires spinning the disk faster when reading near the centre and slower when reading near the edge of disk. Optical disks are generally slower than hard disks. Figure shows the tracks and sectors in a magnetic disk and optical disk.
- Optical disks can store large amount of data, up to 6 GB, in a small space. Commonly used optical disks store 600–700 MB of data.
- The access time for an optical disk ranges from 100 to 200 ms.
- There are two most common categories of optical disks—read-only optical disks and recordable optical disks.

Magnetic disk   Optical disk

3.11.1 CD-ROM

- Originally, Compact Disk (CD) was a popular medium for storing music. Now, it is used in computers to store data and is called Compact Disk-Read Only Memory (CD-ROM).
- As the name suggests, CD-ROM (Figure) is an optical disk that can only be read and not written on. CD-ROM is written on by the manufacturer of the CD-ROM using the laser light.
- A CD-ROM drive reads data from the compact disk. Data is stored as pits (depressions) and lands (flat area) on CD-ROM disk. When the laser light is focused on the disk, the pits scatter the light (interpreted as 0) and the lands reflect the light to a sensor (interpreted as 1).
- As CD-ROM is read only, no changes can be made into the data contained in it.
- Since there is no head touching the disk, but a laser light, CD-ROM does not get worn out easily.
- The storage density of CD-ROM is very high and cost is low as compared to floppy disk and hard disk.
- Access time of CD-ROM is less. CD-ROM drives can read data at 150Kbps. They come in multiples of this speed like—2x, 4x, 52x, 75x, etc.
- It is a commonly used medium for distributing software and large data.

3.11.2 DVD-ROM

- Digital Video Disk-Read Only Memory (DVD-ROM) is an optical storage device used to store digital video or computer data (Figure).
- DVDs look like CDs, in shape and physical size.
- It improves on CD technology.
- It is a high-density medium with increased track and bit density.
- DVD-ROM uses both sides of the disk and special data compression technologies. The tracks for storing data are extremely small.
- A full-length movie can be stored on a single disk.

- Each side of DVD-ROM can store 4.7 GB of data, so a single DVD can store 9.4 GB of data.
- New DVD-ROMs use layers of data track, to double its capacity. Such dual layer disks can store 17 GB of data.

3.11.3 Recordable Optical Disk

In addition to the read only CDs and DVDs, recordable optical disks (Figure) are also available. Users can record music, video, audio and data on it. The recordable optical disks are—

- *Compact Disk-Recordable (CD-R)* is a Write Once-Read Many (WORM) disk. A CD-R disk allows the user to write data permanently on to the disk. Once the data is written, it cannot be erased. CD-R disk uses a laser that burns pits into the disk surface. It looks like a CD disk externally. To write to a CD-R disk, a device named CD-Writer or CD–burner is required. A CD-R disk can store 700 MB of data that can run for 80 minutes. CD-R is used to create music CDs in home computers, back up data from other storage devices, archives of large data, etc.
- *Compact Disk-ReWritable (CD-RW)* allows data to be written, erased and re-written on. The capacity of CD-RW is same as a CD. They generally do not play on all CD-ROM drives.
- *Digital Video Disk-Recordable (DVD-R)* allows recording of data on a DVD. A DVD writer device is required to write the data to DVD. The data once written on a DVD cannot be erased or changed.

3.12 Magneto-Optical Disk;

Magneto-optical disks use laser beam to read data and magnetic field to write data to disk.

- These are optical disks where data can be written, erased and re-written.
- They are expensive and outdated. They were used during the mid 1990s. They have now been replaced by CD-RW and DVD-R.

3.13 How the Computer uses its memory

The computer starts using the memory from the moment the computer is switched on, till the time it is switched off. The list of steps that the computer performs from the time it is switched on are—

- Turn the computer on.
- The computer loads data from ROM. It makes sure that all the major components of the computer are functioning properly.
- The computer loads the BIOS from ROM. The BIOS provides the most basic information about storage devices, boot sequence, security, plug and play capability and other items.
- The computer loads the OS from the hard drive into the system's RAM. CPU has immediate access to the OS as the critical parts of the OS are maintained in RAM as long as the computer is on. This enhances the performance and functionality of the overall system.
- Now the system is ready for use.
- When you load or open an application it is loaded in the RAM. Since the CPU looks for information in the RAM, any data and instructions that are required for processing (read, write or update) is brought into RAM. To conserve RAM usage, many applications load only the essential parts of the program initially and then load other pieces as needed. Any files that are opened for use in that application are also loaded into RAM.
- The CPU requests the data it needs from RAM, processes it and writes new data back to RAM in a continuous cycle. The shuffling of data between the CPU and RAM happens millions of times every second.
- When you save a file and close the application, the file is written to the secondary memory as specified by you. The application and any accompanying files usually get deleted from RAM to make space for new data.
- If the files are not saved to a storage device before being closed, they are lost.

Unit 4: Input and Output Devices

4.1 Introduction;

A computer interacts with the external environment via the input-output (I/O) devices attached to it. Input device is used for providing data and instructions to the computer. After processing the input data, computer provides output to the user via the output device. The I/O devices that are attached, externally, to the computer machine are also called peripheral devices. Different kinds of input and output devices are used for different kinds of input and output requirements. In this chapter, we shall discuss different kinds of input devices and output devices.

4.2 Input-Output Unit;

An I/O unit is a component of computer. The I/O unit is composed of two parts—input unit and output unit. The input unit is responsible for providing input to the computer and the output unit is for receiving output from the computer.

4.2.1 Input Unit

- The input unit gets the data and programs from various input devices and makes them available for processing to other units of the computer.
- The input data is provided through input devices, such as—keyboard, mouse, trackball and joystick. Input data can also be provided by scanning images, voice recording, video recording, etc.
- Irrespective of the kind of input data provided to a computer, all input devices must translate the input data into a form that is understandable by the computer, i.e., in machine readable form. The transformation of the input data to machine readable form is done by the input interface of input device.

In brief, the input unit accepts input data from the user via input device, transforms the input data in computer acceptable form using input interface for the input device and provides the transformed input data for processing.

4.2.2 Output Unit

- The output unit gets the processed data from the computer and sends it to output devices to make them available to the user of computer.
- The output data is provided through output devices like display screen, printer, plotter and speaker.
- The processed data sent to the output device is in a machine understandable form. This processed data is converted to human readable form by the output interface of output device.

Introduction to Information Technology

In brief, the output unit accepts output data from computer via output device, transforms the output information to human readable form using the output interface of output device and provides the transformed output to user.

4.3 Input Devices;

Input devices allow users and other applications to input data into the computer, for processing. The data input to a computer can be in the form of text, audio, video, etc. The data is entered manually by the user or with minimal user intervention. Input devices are classified as follows:

- Human data entry devices
  - Keyboard
  - Pointing devices—mouse, trackball, joystick, digitizing tablet
  - Pick devices—light pen, touch screen
- Source data entry devices
  - Audio input—speech recognition
  - Video input—digital camera
  - Scanner—hand-held scanner, flat-bed scanner
  - Optical Scanner—OCR, OMR, MICR, barcode reader

The input is provided to the computer using an input device, and must be translated to a form that the computer can understand. The translation is done by the input interface of the input device.

In addition to the above devices, the input to a computer can also be provided from a storage device on the computer, another computer, or another piece of equipment, such as a musical instrument, thermometer or sensors.

4.4 Human Data Entry Devices;

Input devices that require data to be entered manually to the computer are identified as human data entry devices. The data may be entered by typing or keying in, or by pointing a device to a particular location.

a. **Keyboard:** Keyboard is the most common means of hardware by which user will input the data. It looks like an electronic typewriter keyboard but contains keys for specific purposes. It is quite similar to the keyboard of typewriter. It contains alphanumeric keys, punctuation keys and some special keys. They are of extended technology, advanced technology and enhanced technology.

b. **Mouse:** Mouse is a small hand held pointing and clicking device used to control the position of the cursor on the screen. It is a small plastic box with two or three buttons on the top. Two rotating wheels inside the mouse detect

the movement and the direction of the movement. By pressing the button of the mouse, the user communicates their choice to the computer. Mouse are of three types: mechanical mouse, Opto-mechanical mouse and optical mouse.

c. **Trackball:** It is another form of pointing device. It is a variant on the mouse. It has a moveable ball on top of a stationary device that is rotated with the fingers or palm of the hand. The trackball looks like a mouse turned upside down. Instead of moving the mouse around on the desktop, the trackball is moved with the tips of the finger.

d. **Touch pad:** It is one of the latest pointing device. It looks like a small gray window, about two inches wide. The user slides the finger tips across the window to move the cursor on the screen. When window is tapped, it works as the clicking of mouse button.

e. **Joystick:** It is a pointing device that consists of a vertical handle like a gearshift lever mounted on a base with one or two buttons. It is used principally in video games and in some computer-aided design systems. As the handle is moved around, the movement is translated into binary signals with the help of electronic contacts in its base.

f. Light pen: The light pen is a light-sensitive stylus, or pen-like device, connected by a wire to the computer terminal. The user brings the pen to a desired point on the display screen and presses the pen button which detects the light emitted from the screen and generates electric pulses, which are transmitted to a processor that identifies the particular pixel.

g. Touch screen: It is a video display screen that has been sensitized to receive input from the touch of a finger or some pointed object. It is a screen with touch sensitive points or program icons. When a specific point is touched by finger tips, related program is activated.

h. Digitizing tablet: It consist a tablet connected by a wire to a stylus or puck. A stylus is a pen-like device with which the user sketches an image. A puck is a copying device with which the user copies, or traces an image. When used with drawing and painting software, a digitizing tablet and stylus allows shading and many other effects similar to those artists achieve with pencil, pen or charcoal.

i. Digital camera: Digital still and video camera are used to capture picture and video which is in the digitized form and can be stored in computers. It can take photo and load them directly from the camera digitally to a main storage or secondary storage device.

4.5 Source Data Entry Devices;
---

a. **Scanner:** It is an input device which is used to enter text, images, and graphs printed on paper directly into the computer. It provides faster and more accurate data entry. It digitizes an image by dividing it into small grid of boxes and representing each box with either zero or one, whether the box is filled or not.

b. **Bar code reader (BCR):** It is a photo-electric scanner that reads bare codes. Bar code is an identification code printed in the form of light and dark bars with coded spacing and thickness which contains identification of an item such as price, quality, manufacturer etc. It is mostly used in supermarket, bookshop, library etc.

c. **Optical character reader (OCR):** It reads printed characters directly, converts them into appropriate codes and stores them in a computer. When light source is passed over a document, written or typed character reflects less light than the background area. Thus the shape of the character is determined.

d. **Optical Mark reader (OMR):** It is a device that reads pencil marks or shades and converts them into computer usable form. It is mainly used to read answer sheets of multiple choice questions and survey form, Graduate record Examination (GRE) etc.

e. **Magnetic Ink Character Reader (MICR):** It is an input device, which reads the characters printed using magnetic ink. Magnetic ink is a special type of ink containing iron-oxide particles, which is used to write characters. It is mostly used in bank for processing cheque.

f. **Smart card:** A smart card looks like a credit card but contains a microprocessor and memory chip. It can be used for public telephone, bank card, as well as for carrying their medical histories on them.

g. **Microphone:** It is an input device used for recording voice or sound. It is widely used for the sound input to the computer for voice mail, videoconferencing, video chat etc.

4.6 Output Devices;

Output unit consists of devices that translate information processed by the computer into a form that humans can understand. The computer-processed information consists of 0s and 1s, which needs to be translated into words, numbers, or pictures that people can understand. There are mainly two kinds of output devices: Softcopy output and hardcopy output. Softcopy refers to the data that is shown on a display screen or sound produced by speaker and hard copy refers to printed output printed

on the paper or some material that can be touched and carried for being shown to others.

Softcopy output devices:
Monitor: Monitor is the display screen that is used to display the text and graphics, allowing users to view the result of the processing. It is the most common and popularly used output device for producing soft copy output. It displays the generated output on a television like screen.

Classification of monitor on the basis of the colors displayed:
   a. **Monochrome monitor:** It displays only one color such as green or white against a contrasting background which is usually black. This monitor is used for text only displays.
   b. **Grayscale monitor:** It is a special monochrome monitor such that it can show the varying intensity of gray against a white or off-white background.
   c. **Color monitor:** It can display 16 colors to several million colors. Color CRT in graphics system is designed as RGB monitor. Color graphics system can be designed by several types of CRT display device.

Classification of monitor on the basis of their display technique:
   a. **Cathode Ray tube (CRT):** It operates much like a television set. A single electron gun (negatively charged heated metal) in a monochrome CRT sends a beam of electrons towards (the positively charged) phosphor coated screen surface. The screen's phosphor coating in organized into a grid of dots. The smallest phosphor dot that an electron gun (gun which is composed of three primary colors red, green and blue) can focus on is called a pixel. When the electrons strike the phosphor-coated screen, it emits visible light. The electron gun systematically aims at every pixel on the screen, starting at the top-left corner and scanning to the right edge, then dropping down a tiny distance and scanning another line.
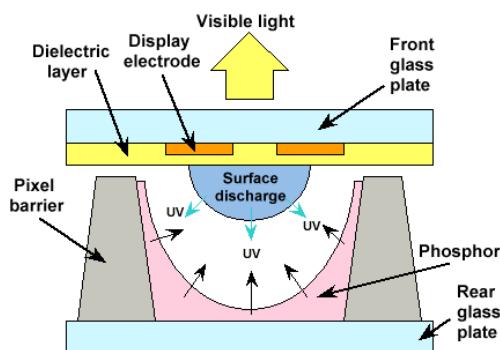
   b. **Flat panel monitor:** There are different types of flat panel monitors like:
      a. **Liquid crystal display (LCD) monitor:** It is the most common flat-panel monitor. It creates images with a special kind of liquid crystal that is normally transparent but becomes opaque when charged with electricity. The liquid crystalline material is sandwiched between two glass or plastic plates. The front plate is transparent and the back is reflective. LCD is lighter in weight, not bulky but expensive compared to CRT monitors. It consumes low voltage and power consumption. The user should sit in front to monitor, if he sits with angle, contents are not visible.

b. **Electroluminescent (EL) display:** It contains a substance that glows when it is charged by an electric current. A pixel is formed on the screen when current is sent to the intersection of the appropriate row and column. The combined voltages from the row and column cause the screen to glow at that point.

c. **Gas Plasma Display:** It is like a neon-bulb, in which the display uses a gas that emits light in the presence of an electric current. That is, the technology uses predominantly neon gas and electrodes above and below the gas. When, electric current passes between the electrodes, the gas glows.

The Advantages of LCD Monitors Over Traditional CRT Monitors

Older cathode ray tube (CRT) monitors are being replaced by liquid crystal display (LCD) monitors in many homes and offices. LCD television panels are increasingly replacing CRT televisions in homes and hotels. LCD panels offer a number of advantages over CRT monitors, especially once the cost of purchasing LCD monitors matches the cost of the CRT models they are replacing.

SPACE
LCD monitors take up much less space than CRT monitors. LCD monitors are also known as "flat panel" monitors because they do not need the space taken up by the cathode ray tube in a conventional CRT monitor. This means that there is more space on the physical desktop for documents and other peripheral computer equipment.

POWER CONSUMPTION
LCD monitors consume less power than CRT monitors. Energy consumption of LCD monitors does increase as screen size increases, but still remains significantly lower

than that of CRT monitors. Typical power consumption for an LCD monitor is between 25 and 50 watts, depending on the size, while even a 15-inch CRT monitor can use between 60 and 80 watts, while 19-inch CRT monitors use anything between 70 and 150 watts.

BRIGHTNESS
LCD monitors output a much brighter image than CRT monitors due to the high peak intensity that they generate. High peak intensity results from the fact that LCD panels have a constant back light that illuminates the screen. This makes LCD monitors a much better choice for use in brightly lit areas.

SCREEN FLICKER
CRT monitors scan the screen viewing area as horizontal lines, and the rate at which the whole screen is covered by these horizontal line scans is known as the refresh rate. CRT monitors typically have low refresh rate, and this causes a flickering effect. This can be detrimental to health when viewed over long periods, causing eye strain and headaches. LCD monitors are capable of much higher refresh rates, with rates of 75 and 85 hertz being common. This means that there is less visible flicker with an LCD monitor, making LCD monitors more comfortable to use.

Voice output device:
It produce digitized sounds ranging from beep, chirp to music. All these sounds are nonverbal. To output sound both the necessary software and the sound card or digital audio circuit board are required.

Hard copy output devices:
    a. **Printer:** Printer is a peripheral device used to print text, picture, illustration, figures and images. It is used to produce hardcopy output i.e. in the permanent readable form. The printers are of two types:
        a. **Impact printer:** The printer which forms character or image by striking a mechanism such as a print hammer or wheel against an inked ribbon, leaving an image on paper. For example, daisy wheel, dot matrix, line printers etc.
        b. **Non-impact printer:** The printer which forms characters and images without making direct physical contact between printing mechanism and paper is called non-impact printer. It is faster than the impact printer. It can also print in different fonts.

Printers can also be classified by how they print and how fast they operate:
    a. **Character printer:** It prints one character at a time. It is low-speed printer with the printing speed in the range of 30-600 characters per second depending upon the type of the printer.
    b. **Line printer:** It prints one line at a time. It uses impact mechanism. It composes one line at a time i.e. the character to be printed in each position is known to the printer before it prints the lien.
    c. **Page printer:** It prints one page at a time. An entire page is processed and printed at a time.

Dot-Matrix Printer:

It is a more common type of microcomputer impact printer which contains a print head of small pins that strikes an inked ribbon forming characters or images. Print heads are available with 9, 18, or 24 pins, with 24-pin head offering the best print quality. It operate at two or three speeds, lower the speed better the printing quality. It is noisy and not good for continuous printing but it can print multi pages i.e. it is possible to use carbon papers to get extra copies of the same document.

Daisy wheel printer: It is an older type of impact printer used with microcomputers, which has a removable print wheel-the flower like daisy wheel consisting of spokes. Each spokes ends with a raised character, which is turned to align the desired letter and then struck with a hammer. It is noisy and slow. It has a fixed font and cannot print graphics. It has a speed of about 90 characters per second.

Inkjet Printer: It is a non-impact character printer. Like dot-matrix printer, ink-jet printer also form images with little dots, but the dots are much smaller and more numerous. It sprays small, electrically charged droplets of ink from nozzles through holes in a matrix at high speed onto paper. Most color printing is done on ink-jet because the nozzles can hold four different colors, CYMK (Cyan, Yellow, Meganta and Black). Other colors are derived by the combination of these four colors.

Laser printer: It is a non-impact page printer. An entire page is processed at a time. It produces high quality print. Print resolution ranges from 300 to more than 2400 dpi. It is expensive than other printers. It uses the principle of dot-matrix printers in creating dot like images. However, these images are created on a drum, treated with a magnetically charged ink powder known as toner, and then transferred from drum to paper.

Thermal Printer: It uses colored waxes and heat to produce images by burning dots onto special heat sensitive paper. It can provide the highest-quality desktop color printing available. They are generally used for high volume printing.

Plotter:

A plotter is a specialized output device designed to produce high-quality graphics in a variety of colors. Plotter is especially useful for creating maps and architectural drawings, although they may also produce less complicated charts and graphs. It uses ink pen or inkjet mechanism to draw graphics or drawings. Either single or multicolor pens can be employed. The pens are driven by motor. Drawings can be prepared on paper or plastic sheets. Pen plotter is slower than inkjet plotters. The graphics and drawings produced by pen plotter are uniform, precise and are of good quality.

4.7 I/O Port;
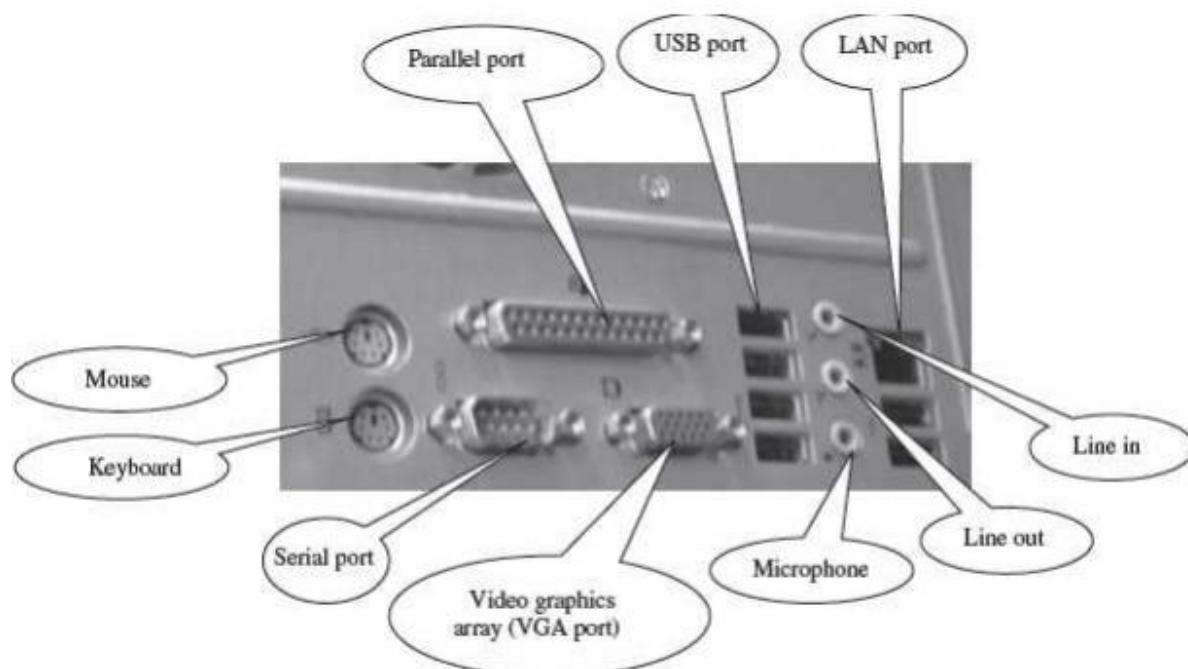
The peripheral devices can be connected to computer in several ways. Devices such as network adapters and sound cards are connected to expansion slots inside the computer. Printers and scanners are connected to ports on the backside of the computer. Also in a portable computer, the PC Card connects to the PC Card slot on it.

The I/O ports are the external interfaces that are used to connect input and output devices like printer, modem and joystick to the computer. The I/O devices are connected to the computer via the serial and parallel ports, Universal Serial Bus (USB) port, Firewire port, etc.

- *Parallel Port* A parallel port is an interface for connecting eight or more data wires. The data flows through the eight wires simultaneously. They can transmit eight bits of data in parallel. As a result, parallel ports provide high speed data transmission. Parallel port is used to connect printer to the computer.
- *Serial Port* A serial port transmits one bit of data through a single wire. Since data is transmitted serially as single bits, serial ports provide slow speed data transmission. Serial port is used to connect external modems, plotters, barcode reader , etc.
- *USB Port* Nowadays, USB is a common and popular external port available with computers. Normally, two to four USB ports are provided on a PC. USB allows different devices to be connected to the computer without requiring re-boot of the computer. USB also has the plug and play feature which allows devices ready to be run simply by plugging them to the USB port. A single USB port can support connection of up to 127 devices.
- *Firewire (IEEE 1394)* It is used to connect audio and video multimedia devices like video camera. It is an expensive technology and is used for large data movement. Hard disk drive and new DVD drives connect through firewire. It has data transfer rate of up to 400 MB/sec.

In addition to the above ports, other ports also exist like Musical Instrument Digital Interface (MIDI) port to connect musical instruments like synthesizers and drum machines, PC expansion boards, and PC card and many more.

4.8 Working of I/O System

The working of I/O system combines I/O hardware and I/O software. The I/O hardware includes ports, buses and device controllers for different devices, and I/O devices. The I/O software is the device driver software that may be embedded with operating system or comes with each device. The working of I/O system is described as follows—

- *I/O Devices* are attached to computer via the ports of computer. There are many standard ports available on the backside of the computer case like serial port and parallel port. If one or more devices use a common set of wires, it is called a bus. For example, PCI bus, PCI Express bus, etc.
- *Device Controller* operates on a bus, a port or a device. It controls the signals on the wires of port or bus. The controllers have one or more registers for data and control signals. Controller may be simple like a serial port controller for a serial port, or, complex like a SCSI controller. Some devices have their own built-in controllers.
- *Device Driver* is software via which the operating system communicates with the device controllers. Each device has its own device driver, and a device controller which is specific to the device. The device drivers hide the differences among the different device controller and present a uniform interface to the operating system.
- *Application programs* use an I/O device by issuing commands and exchanging data with the device driver. The device driver provides correct commands to the controller, interprets the controller register, and transfers data to and from device controller registers as required for the correct device operation.

Unit 5: Data Representation

5.1 Introduction;

The data stored in the computer may be of different kinds, as follows—
- Numeric data (0, 1, 2, …, 9)
- Alphabetic data (A, B, C, …, Z)
- Alphanumeric data—Combination of any of the symbols—(A, B, C… Z), (0, 1… 9), or special characters (+,−, Blank), etc.

All kinds of data, be it alphabets, numbers, symbols, sound data or video data, is represented in terms of 0s and 1s, in the computer. Each symbol is represented as a unique combination of 0s and 1s.

This chapter discusses the number systems that are commonly used in the computer. The number systems discussed in this chapter are—
(1) Decimal number system,
(2) Binary number system,
(3) Octal number system, and
(4) Hexadecimal number system. The number conversions described in this chapter are:
- Decimal (Integer, Fraction, Integer.Fraction) to Binary, Octal, Hexadecimal
- Binary, Octal, Hexadecimal (Integer, Fraction, Integer.Fraction) to Decimal
- Binary to Octal, Hexadecimal
- Octal, Hexadecimal to Binary

5.2 Number System;

A number system in base r or radix r uses unique symbols for r digits. One or more digits are combined to get a number. The base of the number decides the valid digits that are used to make a number. In a number, the position of digit starts from the right-hand side of the number. The rightmost digit has position 0, the next digit on its left has position 1, and so on. The digits of a number have two kinds of values:
- Face value, and
- Position value.

The **face value** of a digit is the digit located at that position. For example, in decimal number 52, face value at position 0 is 2 and face value at position 1 is 5.

The **position value** of a digit is (base^{position}). For example, in decimal number 52, the position value of digit 2 is 10^0 and the position value of digit^5 is 10^1. Decimal numbers have a base of 10.

The number is calculated as the sum of, face value * base^{position}, of each of the digits. For decimal number 52, the number is $5*10^1 + 2*10^0 = 50 + 2 = 52$

In computers, we are concerned with four kinds of number systems, as follows—
- Decimal Number System —Base 10
- Binary Number System —Base 2
- Octal Number System —Base 8
- Hexadecimal Number System—Base 16

The numbers given as input to computer and the numbers given as output from the computer, are generally in decimal number system, and are most easily understood by humans. However, computer understands the binary number system, i.e., numbers in terms of 0s and 1s. The binary data is also represented, internally, as octal numbers and hexadecimal numbers due to their ease of use.

A number in a particular base is written as $(number)_{base\ of\ number}$ For example, $(23)_{10}$ means that the number 23 is a decimal number, and $(345)_8$ shows that 345 is an octal number.

5.2.1 Decimal Number System
- It consists of 10 digits—0, 1, 2, 3, 4, 5, 6, 7, 8 and 9.
- All numbers in this number system are represented as combination of digits 0—9. For example, 34, 5965 and 867321.
- The position value and quantity of a digit at different positions in a number are as follows:

Position:	3	2	1	0	.	−1	−2	−3
Position Value:	10^3	10^2	10^1	10^0		10^{-1}	10^{-2}	10^{-3}
Quantity:	1000	100	10	1		1/10	1/100	1/1000

5.2.2 Binary Number System
- The binary number system consists of two digits—0 and 1.
- All binary numbers are formed using combination of 0 and 1. For example, 1001, 11000011 and 10110101.
- The position value and quantity of a digit at different positions in a number are as follows:

Position:	3	2	1	0	.	−1	−2	−3
Position Value:	2^3	2^2	2^1	2^0		2^{-1}	2^{-2}	2^{-3}
Quantity:	8	4	2	1		1/2	1/4	1/8

5.2.3 Octal Number System
- The octal number system consists of eight digits—0 to 7.
- All octal numbers are represented using these eight digits. For example, 273, 103, 2375, etc.
- The position value and quantity of a digit at different positions in a number are as follows:

Position:	3	2	1	0		−1	−2	−3
Position Value:	8^3	8^2	8^1	8^0		8^{-1}	8^{-2}	8^{-3}
Quantity:	512	64	8	1		1/8	1/64	1/512

5.2.4 Hexadecimal Number System

- The hexadecimal number system consists of sixteen digits—0 to 9, A, B, C, D, E, F, where (A is for 10, B is for 11, C-12, D-13, E-14, F-15).
- All hexadecimal numbers are represented using these 16 digits. For example, 3FA, 87B, 113, etc.
- The position value and quantity of a digit at different positions in a number are as follows:

Position:	3	2	1	0		−1	−2	−3
Position Value:	16^3	16^2	16^1	16^0		16^{-1}	16^{-2}	16^{-3}
Quantity:	4096	256	16	1		1/16	1/256	1/4096

	Base	Digits	Largest Digit
Decimal	10	0–9	9
Binary	2	0,1	1
Octal	8	0–7	7
Hexadecimal	16	0–9, A, B, C, D, E, F	F (15)

Summary of Number System

Decimal	Binary	Octal	Hexadecimal
0	0000	000	0
1	0001	001	1
2	0010	002	2
3	0011	003	3
4	0100	004	4
5	0101	005	5
6	0110	006	6
7	0111	007	7
8	1000	010	8
9	1001	011	9
10	1010	012	A
11	1011	013	B
12	1100	014	C
13	1101	015	D
14	1110	016	E
15	1111	017	F
16	10000	020	10

Decimal, Binary, Octal and Hexadecimal equivalents

5.3 Conversion from Decimal to Binary,

a. Divide the given number by 2
b. Write the quotient under the number. This now becomes the new number.
c. Write the remainder in right side.
d. Repeat step (a) to (c) until 0 is produced as the new number.
e. The 1's and 0's written as remainder in reverse order (i.e. bottom to top) is the required binary number.

For example: $(24)_{10} = (?)_2$

```
Remainder
2|24
2|12    0      ↑
2|6     0      |
2|3     0      |
2|1     1      |
  |0    1      |
```

Therefore, $(24)_{10}=(11000)_2$

Fractional decimal to binary

a. Multiply the fractional part by 2. The result contains an integer part and fractional part.
b. Write the integer number and the fractional number in their respective column.
c. Now, the fractional part becomes new fraction.
d. Repeat step (a) to (c) until the fractional part becomes 0 or the desired place after decimal is obtained.
e. The 1's or 0's written in integer part from top to bottom is the required fractional binary number.

For example: $(0.625)_{10}=(?)_2$

Fractional Decimal	Operation	Product	Fractional part	Integer part
0.625	Multiply by 2	1.250	0.250	1
0.250	Multiply by 2	0.500	0.500	0
0.500	Multiply by 2	1.000	0.000	1

Hence $(0.625)_{10}=(0.101)_2$

5.4 Conversion from Decimal to Octal,

a. Divide the given number by 8
b. Write the quotient under the number. Now it becomes new number.
c. Write the remainder in right side.
d. Repeat steps (a) to (c) until 0 is produced as the new number.
e. The number written as remainder in reverse order (i.e. bottom to top) is the required octal number. For example: $(405)_{10}=(?)_8$

```
8 | 405        Remainder
8 | 50 _____ 5
8 | 6 _____ 2        ↑
     0 _____ 6
```
Therefore $(405)_{10}=(625)_8$

Fractional decimal to octal

    a. Multiply the fractional part by 8. The result contains an integer part and fractional part.

    b. Write the integer number and the fractional number in their respective column.

    c. Now, the fractional part becomes new fraction.

    d. Repeat step (a) to (c) until the fractional part becomes 0 or the desired place after decimal is obtained.

    e. The numbers in the integer part from top to bottom is the required fractional octal number.

For example: $(0.0625)_{10}=(?)_8$

Fractional Decimal	Operation	Product	Fractional part	Integer part
0.0625	Multiply by 8	0.5	0.5	0
0.5	Multiply by 8	4.0	0.0	4

Hence $(0.0625)_{10}=(0.04)_8$

5.5 Conversion from Decimal to Hexadecimal;

    a. Divide the given number by 16

    b. Write the quotient under the number. Now it becomes new number.

    c. Write the remainder in right side.

    d. Repeat steps (a) to (c) until 0 is produced as the new number.

    e. The number written as remainder in reverse order (i.e. bottom to top) is the required octal number. For example: $(495)_{10}=(?)_{16}$

```
16 | 495 _____ 15=F     Remainder
16 | 30 _____ 14=E        ↑
16 | 1 _____ 1
      0
```
Therefore, $(495)_{10}=(1EF)_{16}$

Fractional Decimal to Hexadecimal

    a. Multiply the fractional part by 16. The result contains an integer part and fractional part.

    b. Write the integer number and the fractional number in their respective column.

    c. Now, the fractional part becomes new fraction.

    d. Repeat step (a) to (c) until the fractional part becomes 0 or the desired place after decimal is obtained.

    e. The numbers in the integer part from top to bottom is the required fractional hexadecimal number.

For example: $(0.62)_{10} = (?)_{16}$

Fractional Decimal	Operation	Product	Fractional part	Integer part
0.62	Multiply by 16	9.92	0.92	9
0.92	Multiply by 16	14.72	0.72	14=E
0.72	Multiply by 16	11.52	0.52	11=B

Hence $(0.62)_{10} = (0.9EB...)_{16}$

5.6 Conversion of Binary to decimal,

a) Write binary digits as power of 2 increasing from right to left starting from 0.
b) Convert each power of two into its decimal equivalent term.
c) Add these terms to give the decimal number.

For example: $(1110)_2 = (?)_{10}$

$1110_2 = 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0$

$= 8 + 4 + 2 + 0$

$= 14$

Hence $(1110)_2 = (14)_{10}$

Fractional Binary to Decimal

a) Write the binary numbers as the negative powers of 2 from left to right starting from point as -1, -2 and so on.
b) Convert each power of 2 into its decimal equivalent.
c) Add all decimal equivalent numbers.

For example: $(0.001)_2 = (?)_{10}$

$0.001_2 = 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3}$

$= 0 + 0 + 1/8$

$= 0.125$

Hence, $(0.001)_2 = (0.125)_{10}$

5.7 Conversion of Octal to decimal,

a) Write octal digits as power of 8 increasing from right to left starting from 0.
b) Convert each power of 8 into its decimal equivalent term.
c) Add these terms to give the decimal number.

For example: $(521)_8 = (?)_{10}$

$521_8 = 5 \times 8^2 + 2 \times 8^1 + 1 \times 8^0$

$= 320 + 16 + 1$

$= 337$

Hence $(521)_8 = (337)_{10}$

Fractional Octal to Decimal

a) Write the octal numbers as the negative powers of 8 from left to right starting from point as -1, -2 and so on.
b) Convert each power of 8 into its decimal equivalent.
c) Add all decimal equivalent numbers.

For example: $(127.54)_8 = (?)_{10}$

$127.54_8 = 1 \times 8^2 + 2 \times 8^1 + 7 \times 8^0 + 5 \times 8^{-1} + 4 \times 8^{-2}$

$= 64 + 16 + 7 + 0.625 + 0.0625$

$= 87.6875$

Hence, $(127.54)_8 = (87.6875)_{10}$

5.8 Conversion of Hexadecimal to Decimal;

a. Write hexadecimal digits as power of 16 increasing from right to left starting from 0.
b. Convert each power of 16 into its decimal equivalent term.
c. Add these terms to give the decimal number.

For example: $(345)_{16} = (?)_{10}$
$345_{16} = 3 \times 16^2 + 4 \times 16^1 + 5 \times 16^0$
$= 768 + 64 + 5$
$= 837$
Hence $(345)_{16} = (837)_{10}$

Fractional Hexadecimal to Decimal

a. Write the hexadecimal numbers as the negative powers of 16 from left to right starting from point as -1, -2 and so on.
b. Convert each power of 16 into its decimal equivalent.
c. Add all decimal equivalent numbers.
For example: $(2B.C4)_{16} = (?)_{10}$
$2B.C4_{16} = 2 \times 16^1 + 11 \times 16^0 + 12 \times 16^{-1} + 4 \times 16^{-2}$
$= 32 + 11 + 0.75 + 0.015625$
$= 43.765625$
Hence, $(2B.C4)_{16} = (43.765625)_{10}$

5.9 Conversion of Binary to Octal,

Method-I

a. Separate the given binary number into group of three bits (from right to left, add 0 in the left most side if required).
b. Replace each group by its decimal equivalent.

For example: $(10111)_2 = (?)_8$
Now grouping into group of 3 bits i.e. 010 and 111
Now converting these sets into its octal equivalent.

$010_2 = 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0$ $111_2 = 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$
$= 0 + 2 + 0$ $= 4 + 2 + 1$
$= 2$ $= 7$

Replacing each group by its octal equivalent.
Hence, $(10111)_2 = (27)_8$

Method-II
   a. Convert given binary number in to decimal.
   b. Again convert this decimal number to octal number.

5.10 Conversion of Binary to Hexadecimal;

Method-I

   a. Separate the given number into group of 4 bits (from right to left, add 0 in the left most side if required).
   b. Replace each group by its hexadecimal equivalent.

For example: $(11011110111)_2 = (?)_{16}$

Grouping binary digits in groups of 4 bits

   0110   1111   0111

Now converting each group to its decimal equivalent and then to its hexadecimal equivalent.

$(0110)_2 = (6)_{10} = (6)_{16}$

$(1111)_2 = (15)_{10} = (F)_{16}$

$(0111)_2 = (7)_{10} = (7)_{16}$

Replacing each group by its hexadecimal equivalent.

Hence, $(11011110111)_2 = (6F7)_{16}$

Method-II
   a. Convert binary to decimal.
   b. Then, convert decimal to hexadecimal.

5.11 Conversion of Octal to binary,

Method-I
   a. Take each digit individually, assuming it to be a decimal digit.
   b. Convert it into binary and place it in the combination of 3 bit each.

For example: $(27)_8 = (?)_2$
Converting each digit individually to its' binary equivalent.
   2= 010   7 = 111
Placing it in its order $(27)_8 = (010111)_2$

Method-II
   a. Convert octal number to decimal equivalent.
   b. Again convert the obtained decimal into binary.

P a g e | 80

5.12 Conversion of Hexadecimal to Binary;

Method-I
        a. Take each digit individually
        b. Convert it into binary and place it in the combination of 4 bit each.

For example: $(6F7)_{16} = (?)_2$
Converting each digit individually to its' binary equivalent
        6 = 0110   F=1111   7 = 0111

        Placing it in its order $(6F7)_{16} = (011011110111)_2$

Method-II
a. Convert hexadecimal number to decimal number.
b. Again convert the obtained decimal into binary.

5.13 Binary Arithmetic;

Arithmetic operation with binary number or any number with base r follow the same rules as for decimal numbers. Because binary number system is also the positional number system, the rules for the positional number apply to the binary numbers also.

1. Binary Addition:
As with decimal addition, to add two binary numbers we need to add two corresponding bits from the two numbers at a time. Following are the rules for binary addition:

0+0=0	0+1=1	1+0=1	1+1=10 (with sum 0 and carry 1)

For example:

```
carry →0    1    1    1    1
        1    0    1    1    0    1
      + 0    0    1    1    1    1
      _____
sum  →1    1    1    1    0    0
```

2. Binary Subtraction:
Similar to decimal subtraction, to subtract a binary numbers from other we need to subtract a corresponding bit from other. Following are the rules for binary subtraction:

1-1=0	1-0=1	0-1=1 (with borrowing 1)	0-0=0

For example:

```
borrow⟶ 0    0    1    1    0
         1    1    1    0    0    1 ←Minuend
       - 0    1    0    0    1    0 ←subtrahend
difference⟶ 1  0  0    1    1    1
```

3. Binary Multiplication:

Following are the rules for binary multiplication:

1x1=1	1x0=0	0x1=0	0x0=0

For example:

```
                1    0    1    1 ←multiplicand
           x    1    0    1    1 ←multiplier
                1    0    1    1
           1    0    1    1    X
      0    0    0    0    X    X
 1    0    1    1    X    X    X
 1    1    1    1    0    0    1 ←Product
```

5. Binary Division:

Following are the rules for binary division

1÷1=1	1÷0=not defined	0÷1=0	0÷0=not defined

For example:

Divide 101011 by 110

```
      110)101011(111 ⟵ quotient
          - 110
            1001
          -  110
             111
            - 110
              1  ⟵ remainder
```

5.14 Signed and Unsigned Numbers;

Complement:

In computer system, subtraction is not performed directly as arithmetic subtraction. It is performed by the technique called complement. It is the process of repeated addition.

There are two types of complements:

i. r's complement
ii. (r-1)'s complement

Where, r is the base of a number system. In binary number system, there are two types of complements 2's complement and 1's complement. Similarly, decimal number system has 10's and 9's complement.

1. 9's complement:
The 9's complement of decimal number can be obtained by subtracting each digit of the number from 9.
For example: 9's complement of 3 is: (9-3)=6
And, 9's complement of 234 is: (999-234)=765

2. 10's complement:
The 10's complement of decimal number can be obtained by adding 1 to the least significant digit of 9's complement of that number.
For example: 10's complement of 3 is 7 (9-3=6+1=7)
And, 10's complement of 123 is 877 (999-123=876+1=877)

3. 1's complement:
1's complement of binary number is obtained by subtracting each bit by 1. We can get complement by simply replacing 1 by 0 and 0 by 1. For example: 1's complement of 1011 = 0100

4. 2's complement:
2's complement of a binary number is obtained by adding binary 1 to the 1's complement of the number.
For example: 2's complement of 1101101 is:
    1's complement of 1101101 = 0010010
    2's complement of 1101101 = 0010010+1=0010011

Subtraction of decimal numbers by using 9's complement:
Steps:
  a. Make the both number having same number of digits.
  b. Determine the 9's complement of the number to be subtracted (subtrahend).
  c. Add the 9's complement to the given number from which we subtract (minuend).
  d. If there exists' any additional digit (carry) in the result after addition, remove it and add it to the remaining digits to find the final result. If there exists' no any carry then determine the 9's complement of the result and prefix by negative sign to get the final result.

i)   Subtracting smaller number from greater number.

For example: Subtract $(123)_{10}$ from $(345)_{10}$

Solution:

9's complement of 123 = (999-123) = 876

Adding it with minuend (345)   = + 345

                                         1221

In the result there exist one carry i.e. 1 so, adding this carry to the remaining digit:

                                         221

                                         + 1

                                         222

Hence, $(222)_{10}$ is the required result after subtracting $(123)_{10}$ from $(345)_{10}$

ii)   Subtracting greater number from smaller number.

For example: Subtract $(345)_{10}$ from $(123)_{10}$

Solution:

9's complement of 345 = (999-345) = 654

Adding it with minuend (123)   = + 123

                                         777

In the result there doesn't exist any carry, so calculating 9's complement of the result and prefix by negative sign.

9's complement of 777 = (999-777)=222

Hence, $(-222)_{10}$ is the required result after subtracting $(345)_{10}$ by $(123)_{10}$

Subtraction of decimal numbers by using 10's complement:

Steps:

    a. Make the same numbers having same number of digits.

    b. Determine the 10's complement of the number to be subtracted (subtrahend).

    c. Add the 10's complement to the given number from which we subtract (minuend).

    d. If there exists' any additional digit (carry) in the result after addition, remove it from the result and the remaining digits form the final result. If there exists' no any carry then determine the 10's complement of the result and prefix by negative sign to get the final result.

i)   Subtracting smaller number from greater number.

For example: Subtract $(123)_{10}$ from $(345)_{10}$

Solution:

10's complement of 123 = (999-123) = 876+1 = 877

Adding it with minuend (345)   = + 345

                                                    1222

In the result there exist one carry i.e. 1 so, removing it to find the result.

Hence, $(222)_{10}$ is the required result after subtracting $(123)_{10}$ from $(345)_{10}$

ii) Subtracting greater number from smaller number.

For example: Subtract $(345)_{10}$ from $(123)_{10}$

Solution:

> 10's complement of 345 = (999-345) = 654+1 = 655
> Adding it with minuend (123)   = + 123
>   778

In the result there doesn't exist any carry, so calculating 10's complement of the result and prefix by negative sign.

10's complement of 778 = (999-778) =221+1 = 222

Hence, $(-222)_{10}$ is the required result after subtracting $(345)_{10}$ by $(123)_{10}$

Subtraction of binary number using 1's complement:

Steps:

a. Make the both numbers having same number of bits.
b. Determine the 1's complement of the number to be subtracted (subtrahend).
c. Add the 1's complement to the given number from which we subtract (minuend).
d. If there exists' any additional bit (carry) in the result after addition, remove and add it to the result. If there exists' no carry determine the 1's complement of the result and prefix by negative sign to get the final result.

i)   Subtracting smaller number from greater number.

For example: Subtract $(110000)_2$ from $(111000)_2$

Solution:

> 1's complement of 110000   = 001111
> Adding it with minuend (111000) = + 111000
>   1000111

In the result there exist one carry i.e. 1 so, adding this carry to the remaining digit:

> 000111
> +   1
> 001000

Hence, $(001000)_2$ is the required result after subtracting $(110000)_2$ from $(111000)_2$

ii)   Subtracting greater number from smaller number.

For example: Subtract $(111000)_2$ from $(110000)_2$

Solution:

> 9's complement of 111000   = 000111
> Adding it with minuend (110000) = + 110000
>   110111

In the result there doesn't exist any carry, so calculating 1's complement of the result and prefix by negative sign.

1's complement of 110111 = 001000

Hence, $(-001000)_2$ is the required result after subtracting $(111000)_2$ by $(110000)_2$

Subtraction of binary number using 2's complement:
Steps:
  a. Make the both numbers having same number of bits.
  b. Determine the 2's complement of the number to be subtracted (subtrahend). c. Add the 2's complement to the given number from which we subtract (minuend). d. If there exists' any additional bit (carry) in the result after addition, remove and add it to the result. If there exists' no carry determine the 2's complement of the result and prefix by negative sign to get the final result.

i) *Subtracting smaller number from greater number.*
For example: Subtract $(110000)_2$ from $(111000)_2$
Solution:

  1's complement of 110000 = 001111
  2's complement of 110000 = 001111
              +1
             010000
  Adding it with minuend (111000) = + 111000
             1001000

In the result there exist one carry i.e. 1 so, removing the carry:
Hence, $(001000)_2$ is the required result after subtracting $(110000)_2$ from $(111000)_2$

ii) *Subtracting greater number from smaller number.*
For example: Subtract $(111000)_2$ from $(110000)_2$
Solution:

  1's complement of 111000 = 000111
  2's complement of 111000 = 000111
             + 1
             001000
  Adding it with minuend (110000) = + 110000
             111000

In the result there doesn't exist any carry, so calculating 2's complement of the result and prefix by negative sign.
1's complement of 111000 = 000111
2's complement of 111000 = 000111
        + 1
        001000
Hence, $(-001000)_2$ is the required result after subtracting $(111000)_2$ by $(110000)_2$

5.15 Binary Data Representation;

A binary number may also have a binary point, in addition to the sign. The binary point is used for representing fractions, integers and integer-fraction numbers. Registers are high-speed storage areas within the Central Processing Unit (CPU) of the computer. All data are brought into a register before it can be processed. For

example, if two numbers are to be added, both the numbers are brought in registers, added, and the result is also placed in a register. There are two ways of representing the position of the binary point in the register—fixed point number representation and floating point number representation.

The fixed point number representation assumes that the binary point is fixed at one position either at the extreme left to make the number a fraction, or at the extreme right to make the number an integer. In both cases, the binary point is not stored in the register, but the number is treated as a fraction or integer. For example, if the binary point is assumed to be at extreme left, the number 1100 is actually treated as 0.1100.

The floatingpoint number representation uses two registers. The first register stores the number without the binary point. The second register stores a number that indicates the position of the binary point in the first register.

5.16 Binary Coding Schemes;

The alphabetic data, numeric data, alphanumeric data, symbols, sound data and video data, are represented as combination of bits in the computer. The bits are grouped in a fixed size, such as 8 bits, 6 bits or 4 bits. A code is made by combining bits of definite size. Binary Coding schemes represent the data such as alphabets, digits 0−9, and symbols in a standard code. A combination of bits represents a unique symbol in the data. The standard code enables any programmer to use the same combination of bits to represent a symbol in the data.

The binary coding schemes that are most commonly used are—
- Extended Binary Coded Decimal Interchange Code (EBCDIC),
- American Standard Code for Information Interchange (ASCII), and
- Unicode

In the following subsections, we discuss the EBCDIC, ASCII and Unicode coding schemes.

5.16.1 EBCDIC
- The Extended Binary Coded Decimal Interchange Code (EBCDIC) uses 8 bits (4 bits for zone, 4 bits for digit) to represent a symbol in the data.
- EBCDIC allows 28 = 256 combinations of bits.
- 256 unique symbols are represented using EBCDIC code. It represents decimal numbers (0−9), lower case letters (a−z), uppercase letters (A−Z), Special characters, and Control characters (printable and non−printable, e.g., for cursor movement, printer vertical spacing, etc.).
- EBCDIC codes are mainly used in the mainframe computers.

5.16.2 ASCII

- The American Standard Code for Information Interchange (ASCII) is widely used in computers of all types.
- ASCII codes are of two types—ASCII−7 and ASCII−8.
- ASCII-7 is a 7-bit standard ASCII code. In ASCII-7, the first 3 bits are the zone bits and the next 4 bits are for the digits. ASCII-7 allows 27 = 128 combinations. 128 unique symbols are represented using ASCII-7. ASCII-7 has been modified by IBM to ASCII-8.
- ASCII-8 is an extended version of ASCII-7. ASCII-8 is an 8-bit code having 4 bits for zone and 4 bits for the digit. ASCII-8 allows 28 = 256 combinations. ASCII-8 represents 256 unique symbols. ASCII is used widely to represent data in computers.
- The ASCII-8 code represents 256 symbols.
    - Codes 0 to 31 represent control characters (non−printable), because they are used for actions like, Carriage return (CR), Bell (BEL), etc.
    - Codes 48 to 57 stand for numeric 0−9.
    - Codes 65 to 90 stand for uppercase letters A−Z.
    - Codes 97 to 122 stand for lowercase letters a−z.
    - Codes 128 to 255 are the extended ASCII codes.

5.10.3 Unicode

- Unicode is a universal character encoding standard for the representation of text which includes letters, numbers and symbols in multi−lingual environments. The Unicode Consortium based in California developed the Unicode standard.
- Unicode uses 32 bits to represent a symbol in the data.
- Unicode allows 232 = 4164895296 (~ 4 billion) combinations.
- Unicode can uniquely represent any character or symbol present in any language like Chinese, Japanese, etc. In addition to the letters; mathematical and scientific symbols are also represented in Unicode codes.
- An advantage of Unicode is that it is compatible with the ASCII−8 codes. The first 256 codes in Unicode are identical to the ASCII-8 codes.
- Unicode is implemented by different character encodings. UTF-8 is the most commonly used encoding scheme. UTF stands for Unicode Transformation Format. UTF-8 uses 8 bits to 32 bits per code.
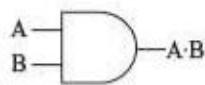
5.17 Logic Gates

The information is represented in the computer in binary form. Binary information is represented using signals in two states off or on which correspond to 0 or 1, respectively. The manipulation of the binary information is done using logic gates. Logic gates are the hardware electronic circuits which operate on the input signals to produce the output signals. Each logic gate has a unique symbol and its operation is described using algebraic expression. For each gate, the truth table shows the output that will be outputted for the different possible combinations of the input

signal. The AND, OR and NOT are the basic logic gates. Some of the basic combination of gates that are widely used are—NAND, NOR, XOR and XNOR.

AND			
X = A.B or X = AB	• Two or more binary inputs • The output is 1 if all the inputs are 1, otherwise the output is 0. • Represented using a multiplication symbol "."		

A	B	A.B
0	0	0
0	1	0
1	0	0
1	1	1

OR			
X = A + B	• Two or more binary inputs • The output is 1 if at least one input is 1, otherwise the output is 0. • Represented using a "+"		

A	B	A+B
0	0	0
0	1	1
1	0	1
1	1	1
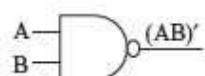
NOT			
A = A′	• One binary input • The output is complement (opposite) of input. If input is 1 output is 0 and if input is 0 output is 1. • Represented using a " ′ "		

A	A′
0	1
1	0

NAND

A ———|‾‾\
B ———|__/o— (AB)'

X = (AB)'
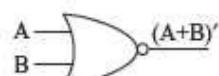
- Two or more binary inputs
- NAND is complement of AND

A	B	(A.B)'
0	0	1
0	1	1
1	0	1
1	1	0

NOR

A ———\‾‾\
B ———/__/o— (A+B)'

X = (A + B)'

- Two or more binary inputs
- NOR is complement of OR.

A	B	(A+B)'
0	0	1
0	1	0
1	0	0
1	1	0

XOR

A ———\\‾‾\
B ———//__/— A⊕B

X = (A ⊕ B)
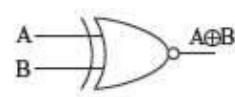
- Two or more binary inputs
- The output is 1 if the odd number of inputs is 1.
- Represented using a " ⊕ "

A	B	(A⊕B)
0	0	0
0	1	1
1	0	1
1	1	0

XNOR

A ———\\‾‾\
B ———//__/o— A⊕B

X = (A ⊕ B)'

- Two or more binary inputs
- XNOR is complement of XOR.

A	B	(A⊕B)'
0	0	1
0	1	0
1	0	0
1	1	1

Unit 6: Computer Software

6.1 Introduction;

A computer system consists of hardware and software. The computer hardware cannot perform any task on its own. It needs to be instructed about the tasks to be performed. Software is a set of programs that instructs the computer about the tasks to be performed. Software tells the computer how the tasks are to be performed; hardware carries out these tasks. Different sets of software can be loaded on the same hardware to perform different kinds of tasks. For example, a user can use the same computer hardware for writing a report or for running a payroll program. The components like monitor, keyboard, processor, and mouse, constitute the hardware.

6.2 Types of Software;

Software can be broadly classified in two categories:
1. System Software, and
2. Application Software.

System software provides the basic functions that are performed by the computer. It is necessary for the functioning of a computer. Application software is used by the users to perform specific tasks. The user may choose the appropriate application software, for performing a specific task, which provides the desired functionality. The system software interacts with hardware at one end and with application software at the other end. The application software interacts with the system software and the users of the computer.
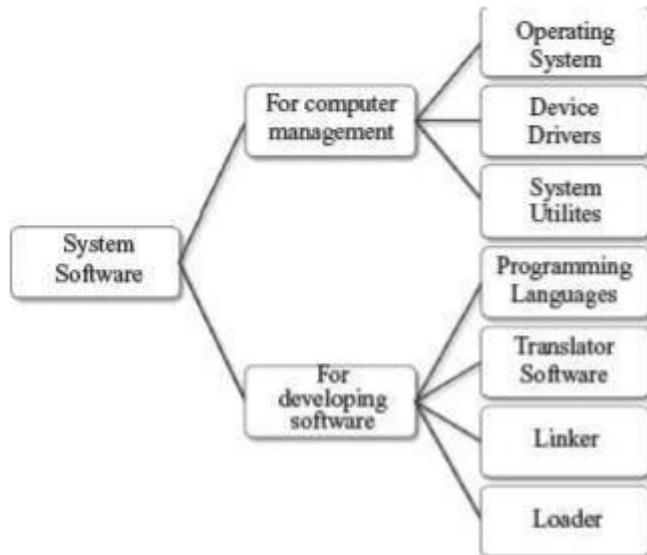
6.3 System Software;

System software provides basic functionality to the computer. System software is required for the working of computer itself. The user of computer does not need to be aware about the functioning of system software, while using the computer. For example, when you buy a computer, the system software would also include different device drivers. When you request for using any of the devices, the corresponding device driver software interacts with the hardware device to perform the specified request. If the appropriate device driver for any device, say a particular model of a printer, is installed on the computer, the user does not need to know about the device driver, while printing on this printer.

The purposes ofthe system software are:
- To provide basic functionality to computer,
- To control computer hardware, and
- To act as an interface between user, application software and computer hardware.

On the basis of their functionality, system software may be broadly divided into two categories (Figure) as follows:

- *System software for the management and functionality of computer relates* to the functioning of different components of the computer, like, processor, input and output devices etc. System software is required for managing the operations performed by the components of computer and the devices attached to the computer. It provides support for various services, as requested by the application software. Operating system, device drivers, and system utilities constitute the system software for management of computer and its resources.

- System software for the *development of application software* provides services required for the development and execution of application software. System software provides the software tools required for the development of application software. The programming language software, translator software, loader, and linker are also categorized as system software, and are required for the application software development.

6.3.1 Operating System

Operating System (OS) is an important part of a computer. OS intermediates between the user of a computer and the computer hardware. Different kinds of application software use specific hardware resources of a computer like CPU, I/O devices and memory, as needed by the application software. OS controls and coordinates the use of hardware among the different application software and the users. It provides an interface that is convenient for the user to use, and facilitates efficient operations of the computer system resources. The key functions of OS are:

- It provides an environment in which users and application software can do work.
- It manages different resources of the computer like the CPU time, memory space, file storage, I/O devices etc. During the use of computer by other programs or users, operating system manages various resources and allocates them whenever required, efficiently.

- It controls the execution of different programs to prevent occurrence of error.
- It provides a convenient interface to the user in the form of commands and graphical interface, which facilitates the use of computer.

6.3.2 Device Driver

A device driver acts as a translator between the hardware and the software that uses the devices. In other words, it intermediates between the device and the software, in order to use the device.

Some devices that are commonly connected to the computer are—keyboard, mouse, hard disk, printer, speakers, microphone, joystick, webcam, scanner, digital camera, and monitor. For proper working of a device, its corresponding device driver must be installed on the computer. For example, when we give a command to read data from the hard disk, the command is sent to the hard disk driver and is translated to a form that the hard disk can understand. The device driver software is typically supplied by the respective device manufacturers.

Programmers can write the higher-level application code independently of whatever specific hardware devices it will ultimately use, because code and device can interface in a standard way, regardless of the software superstructure, or of the underlying hardware. Each version of a device, such as a printer, requires its own hardware-specific specialized commands. In contrast, most applications instruct devices (such as a file to a printer) by means of high level generic commands for the device, such as PRINTLN (print a line). The device-driver accepts these generic high-level commands and breaks them into a series of low-level, device-specific commands, as required by the device being driven.

Nowadays, the operating system comes preloaded with some commonly used device drivers, like the device driver for mouse, webcam, and keyboard. The device drivers of these devices are preinstalled on the computer, such that the operating system can automatically detect the device when it is connected to the computer. Such devices are called plug and play devices. In case the computer does not find the device driver, it prompts the user to insert the media (like a CD which contains the corresponding device driver) provided along with the device. Most device manufacturers, host the device drivers for their devices on their companies' websites; users can download the relevant driver and install it on their computer.

6.3.3 System Utilities

System utility software is required for the maintenance of computer. System utilities are used for supporting and enhancing the programs and the data in computer. Some system utilities may come embedded with OS and others may be added later on. Some examples of system utilities are:
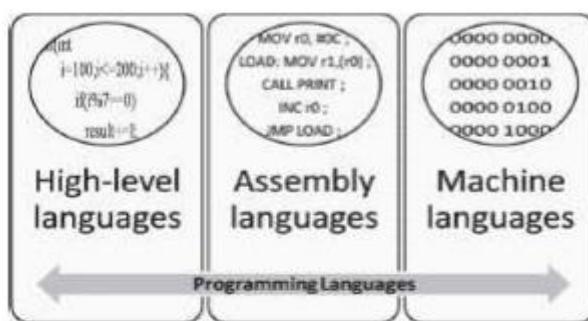
- *Anti-virus* utility to scan computer for viruses.
- *Data Compression* utility to compress the files.
- *Cryptographic* utility to encrypt and decrypt files.

- *Disk Compression* utility to compress contents of a disk for increasing the capacity of a disk.
- *Disk Partitioning* to divide a single drive into multiple logical drives. Each drive is then treated as an individual drive and has its own file system.
- *Disk Cleaners* to find files that have not been used for a long time. It helps the user to decide what to delete when the hard disk is full.
- *Backup Utility* to make a copy of all information stored on the disk. It also restores the backed up contents in case of disk failure.
- *System Profiling Utility* provides detailed information about the software installed on the computer and the hardware attached to it.
- *Network Managers* to check the computer network and to log events.

6.3.4 Programming Languages

A Programming Language consists of a set of vocabulary and grammatical rules, to express the computations and tasks that the computer has to perform. Programming languages are used to write a program, which controls the behavior of computer, codify the algorithms precisely, or enables the human-computer interface. Each language has a unique set of keywords (words that it understands) and a special syntax for organizing program instructions. The programming language should be understood, both by the programmer (who is writing the program) and the computer. A computer understands the language of 0·s and 1·s, while the programmer is more comfortable with English-like language. Programming Language usually refers to high-level languages like COBOL, BASIC, FORTRAN, C, C++, Java etc. Programming languages fall into three categories.

- *Machine Language* is what the computer can understand but it is difficult for the programmer to understand. Machine languages consist of numbers only. Each kind of CPU has its own unique machine language.
- *Assembly Language* falls in between machine language and high-level language. They are similar to machine language, but easier to program in, because they allow the programmer to substitute names for numbers.
- *High-level Language* is easier to understand and use for the programmer but difficult for the computer.

Regardless of the programming language used, the program needs to be converted into machine language so that the computer can understand it. In order to do this a program is either compiled or interpreted.

6.3.4.1 Machine Language

A program written in machine language is a collection of binary digits or bits that the computer reads and interprets. It is a system of instructions and data executed directly by a computer·s CPU. It is also referred to as machine code or object code. It is written as strings of 0's and 1·s, as shown in Figure. Some of the features of a program written in machine language are as follows:

```
000000001010000100000000000011000
000000001000111100001100000100001
10001100011000100000000000000000
10001100111100100000000000000100
10101100111100100000000000000000
10101100011000100000000000000100
00000011111100000000000000001000
```

- The computer can understand the programs written in machine language directly. No translation of the program is needed.
- Program written in machine language can be executed very fast (Since no translation is required).
- Machine language is defined by the hardware of a computer. It depends on the type of the processor or processor family that the computer uses, and is thus machine-dependent. A machine- level program written on one computer may not work on another computer with a different processor.
- Computers may also differ in other details, such as memory arrangement, operating systems, and peripheral devices; because a program normally relies on such factors, different computer may not run the same machine language program, even when the same type of processor is used.
- Most machine-level instructions have one or more opcode fields which specify the basic instruction type (such as arithmetic, logical, jump, etc), the actual operation (such as add or compare), and some other fields.
- It is difficult to write a program in machine language as it has to be written in binary code. For e.g., 00010001 11001001. Such programs are also difficult to modify.
- Since writing programs in machine language is very difficult, programs are hardly written in machine language.

6.3.4.2 Assembly Language

A program written in assembly language uses symbolic representation of machine codes needed to program a particular processor (CPU) or processor family. This representation is usually defined by the CPU manufacturer, and is based on abbreviations (called mnemonics) that help the programmer remember individual instructions, registers, etc. Small, English-like representation is used to write the program in assembly language, as shown in Figure. Some of the features of a program written in assembly language are as follows:

```
MOV   B, A
MVI   C, 06H
LXI   H, XX50H
ADD   M
JNC   NXTITM
INR   B
INX   H
DCR   C
JNZ   NXTBIT
```

- Assembly language programs are easier to write than the machine language programs, since assembly language programs use short, English-like representation of machine code. For e.g.:

  ADD 2, 3
  LOAD A
  SUB A, B

- The program written in assembly language is the source code, which has to be converted into machine code, also called object code, using translator software, namely, assembler.
- Each line of the assembly language program is converted into one or more lines of machine code. Hence assembly language programs are also machine-dependent.
- Although assembly language programs use symbolic representation, they are still difficult to write.
- Assembly language programs are generally written where the efficiency and the speed of program are the critical issues, i.e. programs requiring high speed and efficiency.

6.3.4.3 High-level Language

A program in a high-level language is written in English-like language. Such languages hide the details of CPU operations and are easily portable across computers. A high-level language isolates the execution semantics of computer architecture from the specification of the program, making the process of developing a program simpler and more understandable with respect to assembly and machine level languages. Some of the features of a program written in high-level language are as follows:

- Programs are easier to write, read or understand in high-level languages than in machine language or assembly language. For example, a program written in C++ is easier to understand than a machine language program.
- Programs written in high-level languages is the source code which is converted into the object code (machine code) using translator software like interpreter or compiler.
- A line of code in high-level program may correspond to more than one line of machine code.
- Programs written in high-level languages are easily portable from one computer to another.
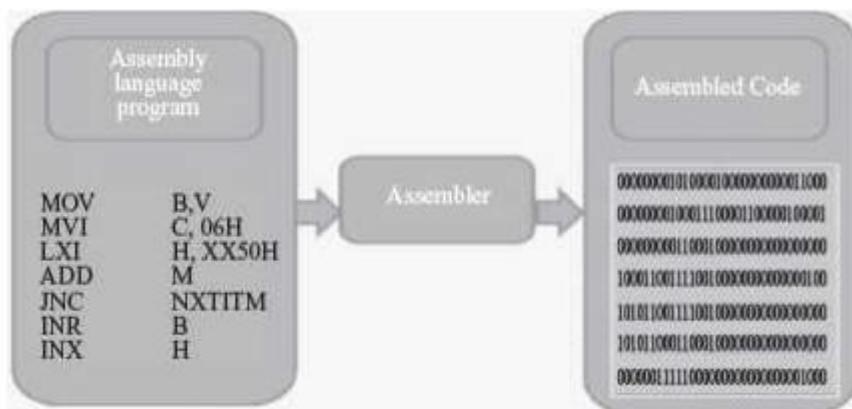
6.3.5 Translator Software

Translator software is used to convert a program written in high-level language and assembly language to a form that the computer can understand. Translator software converts a program written in assembly language, and high-level language to a machine-level language program. The translated program is called the object code. There are three different kind of translator software:

6.3.5.1 Assembler

Assembly language is also referred to as a symbolic representation of the machine code. Assembler is a software that converts a program written in assembly language into machine code (Figure). There is usually a one-to-one correspondence between simple assembly statements and machine language instructions. The machine language is dependent on the processor architecture, though computers are generally able to carry out the same functionality in different ways. Thus the corresponding assembly language programs also differ for different computer architectures.

6.3.5.2 Compiler

A program written in a high-level language has to be converted to a language that the computer can understand, i.e. binary form. Compiler is the software that translates the program written in a high-level language to machine language. The program written in high-level language is referred to as the source code and compiled program is referred as the object code. The object code is the executable code, which can run as a stand-alone code. It does not require the compiler to be present during execution. Each programming language has its own compiler. Some languages that use a compiler are C++, COBOL, Pascal, and FORTRAN. In some languages, compilation using the compiler and linking using the linker are required for creating the executable object code.

The compilation process generally involves two parts—breaking down the source code into small pieces and creating an intermediate representation, and, constructing the object code for the intermediate representation. The compiler also reports syntax errors, if any, in the source code.

6.3.5.3 Interpreter

The purpose of interpreter is similar to that of a compiler. The interpreter is used to convert the high-level language program into computer-understandable form. However, the interpreter functions in a different way than a compiler. Interpreter performs line-by-line execution of the source code during program execution. Interpreter reads the source code line-by-line, converts it into machine understandable form, executes the line, and then proceeds to the next line. Some languages that use an interpreter are BASIC and Python.

Difference Between a Compiler and An Interpreter:

Compiler and Interpreter are used to convert a program written in high-level language to machine language; however, they work differently. The key differences between a compiler and an interpreter are as follows:

- Interpreter looks at a source code line-by-line. Compiler looks at the entire source code.
- Interpreter converts a line into machine executable form, executes the line, and proceeds with the next line. Compiler converts the entire source code into object-code and creates the object code. The object code is then executed by the user.
- For a given source code, once it is compiled, the object code is created. This object code can be executed multiple number of times by the user. However, interpreter executes line-by-line, so executing the program using an interpreter means that during each execution, the source code is first interpreted and then executed.
- During execution of an object code, the compiler is not required. However, for interpretation, both interpreter and the source code is required during execution (because source code is interpreted during execution).
- Since interpreter interprets line-by-line, the interpreted code runs slower than the compiled code.

6.3.6 Linker

Linker is a program that links several object modules and libraries to a single executable program. A source code of a program is often very large consisting of several hundred or more lines. The source code may also include reference to libraries. All these independent modules may not be stored in a single object file. The code is broken down into many independent modules for easy debugging and maintenance. Before execution of the program, these modules and the required libraries are linked together using the linker software. The compiled and the linked program are called the executable code.

6.3.7 Loader

The loader software is used to load and re-locate the executable program in the main memory. Software has to be loaded into the main memory during execution. Loader assigns storage space to the program in the main memory for execution. The java source file Sample.java is compiled and converted into a set of bytecodes and are stored in Sample.class file. At runtime, the class loader loads the bytecodes from

the hard disk, checks it, and runs it in an interpreter. The interpreter executes the bytecode and makes calls to the underlying hardware.

6.4 Application Software;

The software that a user uses for accomplishing a specific task is the application software. Application software may be a single program or a set of programs. A set of programs that are written for a specific purpose and provide the required functionality is called software package. Application software is written for different kinds of applications—graphics, word processors, media players, database applications, telecommunication, accounting purposes etc.

Some examples of application software packages are as follows:

- *Word Processing Software:* For writing letter, reports, documents etc. (e.g. MS-WORD).
- *Image Processing Software:* For assisting in drawing and manipulating graphics (e.g. Adobe Photoshop).
- *Accounting Software:* For assisting in accounting information, salary, tax returns (Tally software).
- *Spreadsheet Software:* Used for creating budget, tables etc. (e.g. MS-Excel).
- *Presentation Software:* To make presentations, slide shows (e.g. MS-PowerPoint)
- *Suite of Software having Word Processor, Spreadsheet and Presentation Software:* Some examples are MS-Office, Google Docs, Sun Openoffice, Apple iWork.
- *CAD/CAM Software:* To assist in architectural design. (e.g. AutoCAD, Autodesk)
- *Geographic Information Systems:* It captures, stores, analyzes, manages, and presents data, images and maps that are linked to different locations. (e.g. ArcGIS)
- *Web Browser Software:* To access the World Wide Web to search documents, sounds, images etc. (e.g. Internet Explorer, Netscape Communicator, Chrome).

6.5 Software Acquisition;

Different kinds of software are made available for use to users in different ways. The user may have to purchase the software, can download for free from the Internet, or can get it bundled along with the hardware. Nowadays with the advent of Cloud computing, many application software are also available on the cloud for use through the Internet, e.g. Google Docs. The different ways in which the software are made available to users are:

- **Retail Software** is off-the-shelf software sold in retail stores. It comes with printed manuals and installation instructions. For example, Microsoft Windows operating system.
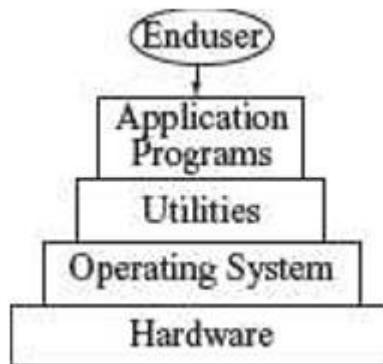
- **OEM Software stands** for "Original Equipment Manufacturer" software. It refers to software which is sold, and bundled with hardware. Microsoft sells its operating system as OEM software to hardware dealers. OEM software is sold at reduced price, without the manuals, packaging and installation instructions. For example, Dell computers are sold with the "Windows 7" OS pre-loaded on them.
- **Demo Software** is designed to demonstrate what a purchased version of the software is capable of doing and provides a restricted set of features. To use the software, the user must buy a fully- functional version.
- **Shareware** is a program that the user is allowed to try for free, for a specified period of time, as defined in the license. It is downloadable from the Internet. When the trial period ends, the software must be purchased or uninstalled.
- **Freeware** is software that is free for personal use. It is downloadable from the Internet. The commercial use of this software may require a paid license. The author of the freeware software is the owner of the software, though others may use it for free. The users abide by the license terms, where the user cannot make changes to it, or sell it to someone else.
- **Public Domain Software** is free software. Unlike freeware, public domain software does not have a copyright owner or license restrictions. The source code is publicly available for anyone to use. Public domain software can be modified by the user.
- **Open-Source Software** is software whose source code is available and can be customized and altered within the specified guidelines laid down by the creator. Unlike public domain software, open-source software has restrictions on their use and modification, redistribution limitations, and copyrights. Linux, Apache, Firefox, OpenOffice are some examples of open-source software.

6.6 Operating System

6.6.1 Introduction,
The computer system comprises of a functional set of hardware, software, user and data. Hardware consists of the components of computer like memory, processor, storage devices, and Input/Output devices. The software may be of different kinds—application software and system softwa re. A computer system may be a single stand-alone system or may consist of several interconnected systems. The user uses the application software to perform various tasks, for example, the user uses word processing software for document preparation. While using the application software, the user uses the storage of a computer—to store a document on the hard disk, to execute a command on the CPU, to retrieve a document from a peripheral device or to print document on printer. For using the hardware, there is a need for software that interacts with both the hardware and the application software. Operating system (OS) is the software that provides an interface between the computer hardware, and the application programs or users.

6.6.2 Objectives of Operating System,

Operating system is system software that controls and coordinates the use of hardware among the different application software and users. OS intermediates between the user of computer and the computer hardware. The user gives a command and the OS translates the command into a form that the machine can understand and execute. OS has two main objectives—

(1) to make the computer system convenient and easy to use, for the user, and—

(2) to use the computer hardware in an efficient way, by handling the details of the operations of the hardware.

- OS hides the working of the hardware from the user and makes it convenient for the user to use the machine. The application program used by the user requires the use of the hardware during processing. Some examples are— display of application·s user interface, loading a program into memory, using I/O devices, allocating CPU to different processes during execution, and store or load data from hard disk. When using the machine, the user gives the command to perform the required actions to the OS and the OS handles all the operational steps. The user is not bothered about how these actions will be performed. This is the job of OS. OS provides an interface to the application programs to interact with the hardware. The user need not get into the details of this interaction.

- At the other end, the different resources of computer hardware have to be managed and controlled. This includes managing the communication between different devices, controlling the sequence and execution of processes, allocating space on hard disk, providing error handling procedures etc. OS supervises and manages the hardware of the computer.

Some of the commonly used operating systems are Microsoft Disk Operating System (MSDOS), Windows 7, Windows XP, Linux, UNIX, and Mac OS X Snow Leopard.

6.6.3 Types of OS,

OS are classified into different types depending on their capability of processing—

(1) Single user,

(2) Multiuser,
(3) Multitasking,
(4) Multiprocessing,
(5) Real time, and
(6) Embedded.

- **Single User and Single Task OS** is for use by a single user for a standalone single computer for performing a single task. Operating system for Personal Computers (PC) are singleuser OS. For example, if the user is editing a document, then a document cannot be printed on the printer simultaneously. Single user OS are simple operating system designed to manage one task at a time. MS-DOS is an example of single user OS.
- **Single User and Multitasking OS** allows execution of more than one task or process concurrently. For this, the processor time is divided amongst different tasks. This division of time is also called time sharing. The processor switches rapidly between processes. For example, the user can listen to music on the computer while writing an article using a word processor software. The user can switch between the applications and also transfer data between them. Windows 95 and all later versions of Windows are examples of multitasking OS.
- **Multiuser OS** is used in computer networks that allow same data and applications to be accessed by multiple users at the same time (Figure 7.4). The users can also communicate with each other. Linux, UNIX, and Windows 7 are examples of multiuser OS.
- **Multiprocessing OS** have two or more processors for a single running process. Processing takes place in parallel and is also called parallel processing. Each processor works on different parts of the same task, or, on two or more different tasks. Since execution takes place in parallel, they are used for high speed execution, and to increase the power of computer. Linux, UNIX and Windows 7 are examples of multiprocessing OS.
- **Real Time OS** are designed to respond to an event within a predetermined time. These operating systems are used to control processes. Processing is done within a time constraint. OS monitors the events that affect the execution of process and respond accordingly. They are used to respond to queries in areas like medical imaging system, industrial control systems etc. LynxOS is an example of real time OS.
- **Embedded OS** is embedded in a device in the ROM. They are specific to a device and are less resource intensive. They are used in appliances like microwaves, washing machines, traffic control systems etc.

6.6.4 Functions of OS

Operating system is a large and complex software consisting of several components. Each component of the operating system has its own set of defined inputs and outputs. Different components of OS perform specific tasks to provide the overall functionality of the operating system (Figure). Main functions of the operating system are as follows:

- **Process Management**—The process management activities handled by the OS are—
  (1) control access to shared resources like file, memory, I/O and CPU,
  (2) control execution of applications,
  (3) create, execute and delete a process (system process or user process),
  (4) cancel or resume a process
  (5) schedule a process, and
  (6) synchronization, communication and deadlock handling for processes.
- **Memory Management**—The activities of memory management handled by OS are—
  (1) allocate memory,
  (2) free memory,
  (3) re-allocate memory to a program when a used block is freed, and
  (4) keep track of memory usage.
- **File Management**—The file management tasks include—
  (1) create and delete both files and directories,
  (2) provide access to files,
  (3) allocate space for files,
  (4) keep back-up of files, and
  (5) secure files.
- **Device Management**—The device management tasks handled by OS are—
  (1) open, close and write device drivers, and
  (2) communicate, control and monitor the device driver.
- **Protection and Security**—OS protects the resources of system. User authentication, file attributes like read, write, encryption, and back-up of data are used by OS to provide basic protection.

- **User Interface or Command Interpreter**—Operating system provides an interface between the computer user and the computer hardware. The user interface is a set of commands or a graphical user interface via which the user interacts with the applications and the hardware.

6.6.5 Examples of Operating Systems
1 MS-DOS

- MS-DOS was the first widely-installed operating system for PCs in 1980s.
- MS-DOS is easy to load and install. It neither requires much memory for the operating system, nor a very powerful computer to run on.
- MS-DOS is a command line user interface operating system. This means that the user has to type single line commands through the command interface. So, user has to remember the different commands and their syntax.
- It is a single-user and single-tasking operating system for the PC. Only one user can use it and only one task can be executed, at a given point of time. Also, it does not have a builtin support for networking.
- MS-DOS is a 16-bit OS, meaning thereby that it can send or receive 16 bits of data at a time and can process 16 bits of data. It is not able to take the advantage of 32-bit processors.
- To use MS-DOS, user must know where the programs and data are stored and how to interact with it. In the MS-DOS command mode, command.com routine interprets the typed in command from the keyboard.

2 Windows Family of OS

- Windows is a personal computer operating system from Microsoft.
- The Windows family of OS which is currently in use includes the Windows 9x family (Windows 95, Windows 98 and Windows 2000), Windows XP, Windows Vista, and Windows 7 operating systems.
- Windows family of OS is GUI-based operating system. Since GUI interfaces are easy to use and are user-friendly, these have become very popular.
- Windows support multi-tasking. It means Windows OS allows simultaneous execution of multiple tasks
- Windows contains built-in networking, which allows users to share files and applications with each other, if their PCs are connected to a network.
- Windows 7 comes in six different editions, Starter, Home Basic, Home Premium, Professional, Enterprise and Ultimate.
- With each new version of the Windows OS, the user interface undergoes some changes and the user has to learn to use the new interface. This becomes troublesome for the user.

3 Linux OS

- Linux is a Unix-like OS. Unix OS has a user interface called shell. The kernel provides interface for the programs to interact with the hardware, and provides services like process management and memory management. The shell interacts with the kernel through the system calls.

- Linux was developed by Linus Torvalds in 1992. Linux is copyright under the GNU Public License. Linux is a "free" operating system that is easily available. Since Linux follows the open development model, it is being constantly upgraded by programmers across the globe.
- Some organizations offer Linux with add-on features and capabilities. Red Hat, Mandrake, Debian and Novell are the popular vendors of Linux OS.
- Tux, the Linux penguin is the official mascot of Linux.
- Linux is a command line user interface OS. Linux has GUI interfaces called desktop environments like GNOME and K Desktop Environment (KDE). The GUI interface is convenient for the user to use.
- Linux is a 32-bit, multi-tasking OS. It supports multiple users and multiple processors.
- Linux is a reliable and secure OS, and is available almost for free. So, Linux is fast becoming very popular and powerful OS.
- Linux OS is easily available, such as Redhat Linux ver. 9, and, Debian·s— Ubuntu, Kubuntu, and Edubuntu.

Unit 7: Data Communication and Computer Network

7.1 Introduction;

The communication process involves—sender of information, receiver of information, language used for communication, and medium used to establish the communication. Communication between computers also follows a similar process.

7.2 Importance of Networking;

Networking of computers provides a communication link between the users, and provides access to information. Networking of computers has several uses, described as follows:

- **Resource Sharing**—In an organization, resources such as printers, fax machines and scanners are generally not requiredby each person at all times. Moreover, for small organizations it may not be feasible to provide such resources to each individual. Such resources can be made available to different users of the organization on the network. It results in availability of the resource to different users regardless of the physical location of the resource or the user, enhances optimal use of the resource, leads to easy maintenance, and saves cost too.
- **Sharing of Information**—In addition to the sharing of physical resources, networking facilitates sharing of information.Information stored on networked computers located at same or different physical locations, becomes accessible to the computers connected to the network.
- **As a Communication Medium**—Networking helps in sending and receiving of electronic-mail (email) messages from anywhere in the world. Data in the form of text, audio, video and pictures can be sent via e-mail. This allows the users to communicate online in a faster and cost effective manner. Video conferencing is another form of communication made possible via networking. People in distant locations can hold a meeting, and they can hear and see each other simultaneously.
- **For Back-up and Support**—Networked computers can be used to take back-up of critical data. In situations where there is a requirement of always-on computer, another computer on the network can take over in case of failure of one computer.

7.3 Data Transmission Media;

The data is sent from one computer to another over a transmission medium. The transmission media can be grouped into guided media, and unguided media.

In the guided media, the data signals are sent along a specific path, through a wire or a cable. Copper wire and optical fibers are the most commonly used guided media. Copper wire transmits data as electric signals. Copper wires offer low resistance to current signal, facilitating signals to travel longer distances. To

minimize the effect of external disturbance on the copper wire, two types of wiring is used:
(1) Twisted Pair, and
(2) Coaxial Pair.
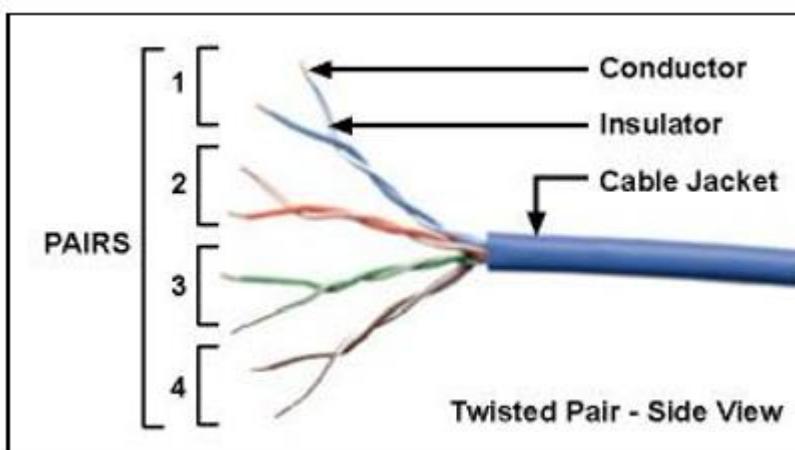
Optical fibers transmit data as light signals.

In the unguided media, the data signals are not bounded by a fixed channel to follow. The data signals are transmitted by air. Radio, microwave, and satellite transmissions fall into this category.

1 Twisted Pair
- A twisted pair cable consists of four pairs of copper wires coated with an insulating material like plastic or Teflon, twisted together. The twisting of wires reduces electromagnetic interference from externalsources.
- Twisted pair cabling is often used in data networks for short and medium length connections because of its relatively lower costs compared to optical fiber and coaxial cable.
- Twisted pair is of two kinds—Shielded Twisted Pair (STP), and Unshielded Twisted Pair (UTP).
- STP cable has an extra layer of metal foil between the twisted pair of copper wires and the outer covering. The metal foil covering provides additional protection from external disturbances. However, the covering increases the resistance to the signal and thus decreases the length of the cable. STP is costly and is generally used in networks where cables pass closer to devices that cause external disturbances.
- UTPis the most commonly used medium for transmission over short distances up to 100m. Out of the four pairs of wires in a UTP cable, only two pairs are used for communication. Figure shows the cross-section of STP and UTP cables.
- UTP cables are defined in different categories. The commonly used UTP cable is the Cat-5 cable which is used with fast Ethernet.
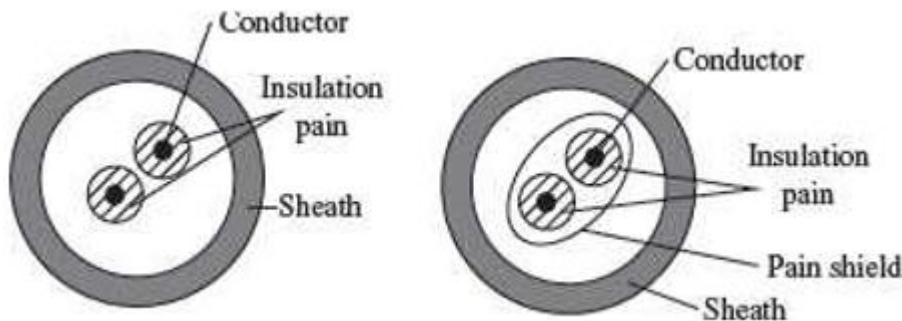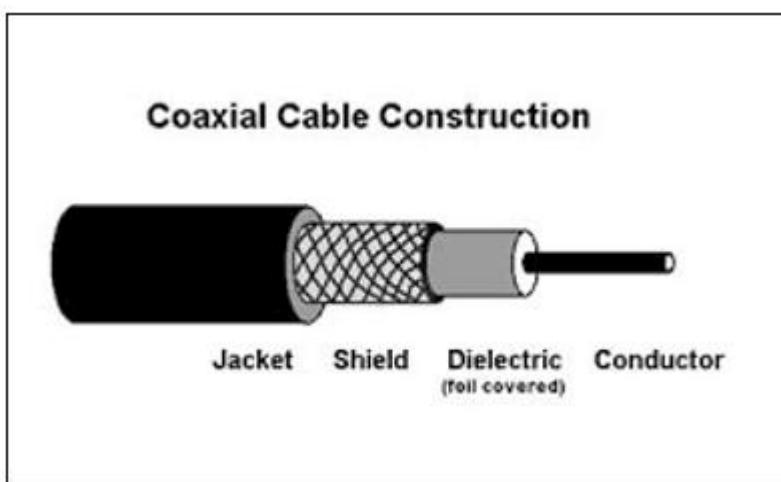
Fig: Cross section of UTP and STP

2. Coaxial Cable

- A coaxial cable has a single inner conductor that transmits electric signals; the outer conductor acts as a ground. The two conductors are separated by insulation. The inner conductor, insulator, and the outer conductor are wrapped in a sheath of Teflon or PVC.
- The copper wire is used for both inner and outer conductor. The signal is transmitted over the surface of the inner conductor.
- In an ideal coaxial cable the electromagnetic field carrying the signal exists only in the space between the inner and outer conductors. This allows coaxial cable runs to be installed next to metal objects such as gutters without the power losses that occur in other transmission lines, and provides protection of the signal from external electromagnetic interference.
- A thickercoaxial cable can transmit more data than a thinner one.
- The commonly used coaxial cable is10 base 2that transmits over a distance of 185 m, and10 base 5that transmits over a distance of 500 m.
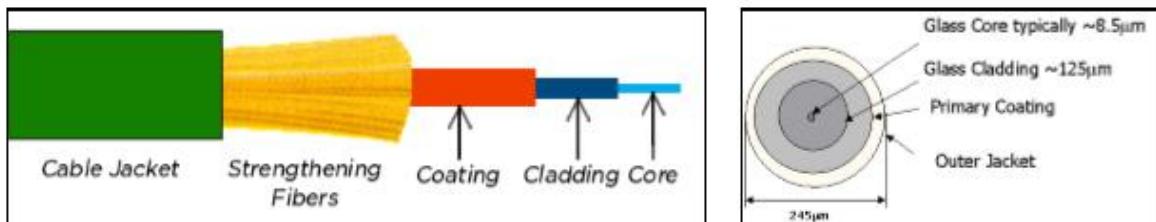

3. Optical Fiber

- Optical fibers are being used for transmission of information over large distances more costeffectively than the copper wire connection. Communication systems are now unthinkable without fiber optics.
- Optical fiber transmits data as light signals instead of electric signals.
- An optical fibercable consists of

(1) core–optical fiber conductor (glass) that transmits light,
(2) cladding–an optical material that surrounds the core to prevent any light from escaping the core, and
(3) jacket—outer covering made of plastic to protect the fiber from damage.

- Modern optical fiber cables can contain up to a thousand fibers in a single cable, so the performance of optical networks easily accommodate large demands for bandwidth on a pointto-point basis.

The Advantages of Optical Fibers over wires are:

- Optical fibers do not cause electrical interference in other cables, since they use light signals.
- Due to much lowerattenuation and interference, optical fiber has large advantages over existing copper wire in long-distance and high-demand applications.
- A fiber can carry a pulse of light much farther than a copper wire carrying a signal.
- Optical fiber can carry more information than a wire (light can encode more information than electrical signal).
- A single optical fiber is required for light to travel from one computer to another (two wires are required for electric connection).
- Because signals in optical fibers degradeless, lower-power transmitters can be used instead of the high-voltage electrical transmitters needed for copper wires. Again, this saves your provider and you, money.
- No amplification of the optical signal is needed over distances of hundreds of kilometers. This has greatly reduced the cost of optical networking, particularly over undersea spans where the cost reliability of amplifiers is one of the key factors determining the performance of the whole cable system.
- Optical fibers are ideally suited for carrying digital information, which is especially useful in computer networks.
- They are highly secure as they cannot be tapped and for lack of signal radiation.
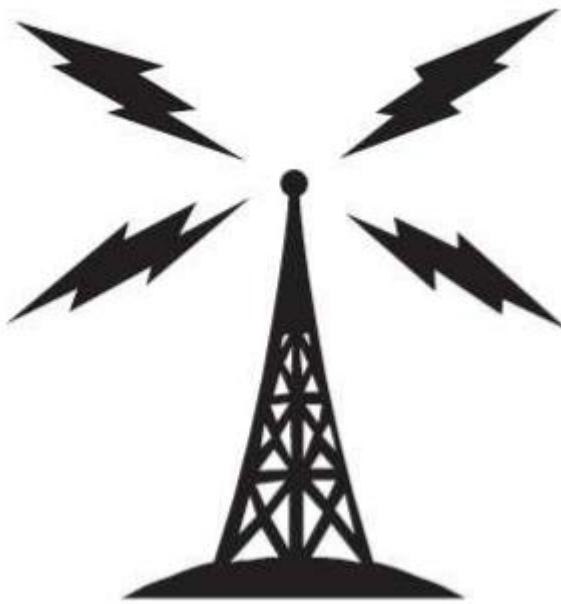
The Disadvantages of Optical Fiber are:

- Installing an optical fiber requires special equipment.
- If a fiber breaks, finding the broken location is difficult.
- Repairing a broken optical fiber is difficult and requires special equipment.
- Due to its high installation costs, they are economical when the bandwidth utilization is high.
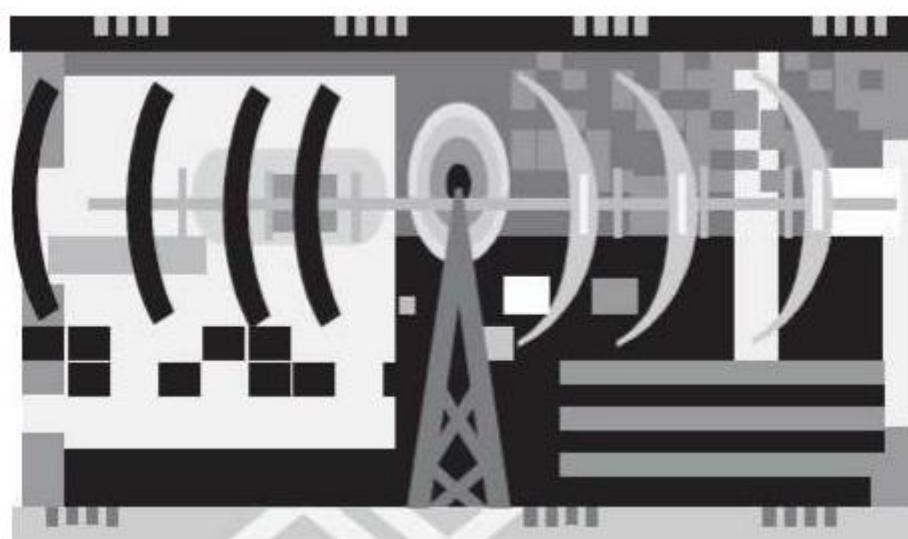
4. Radio Transmission

The electromagnetic radio waves that operate at the adio frequency are also used to transmit computer data. rhis transmission is also known as Radio Frequency RF) transmission. The computers using RF ransmission do not require a direct physical connection ike wires or cable. Each computer attaches to an antenna hat can both send and receive radio transmission.

5 Microwave Transmission

Microwave transmission refers to the technique of transmitting information over a microwave link. Microwaves have a higher frequency than radio waves. Microwave transmission can be aimed at a single direction, instead of broadcasting in all directions (like in radio waves). Microwaves can carry more information than radio waves but cannot penetrate metals. Microwaves are used where there is a clear path between the transmitter and the receiver.
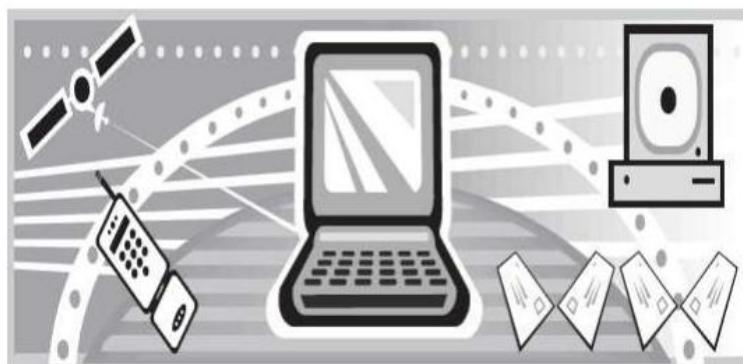
Microwave transmission has the advantage of not requiring access to all contiguous land along the path of the system, since it does not need cables. They suffer from the disadvantages: a) needing expensive towers and repeaters, and b) are subject to

interference from passing airplanes and rain. Because microwave systems are line-of-sight media, radio towers must be spaced approximately every 42 km along the route.

6. Satellite Transmission

The communication across longer distances can be provided by combining radio frequency transmission with satellites. Geosynchronous satellites are placed in an orbit synchronized with the rotation of the earth at a distance of 36,000 km above the surface of the earth. Geosynchronous satellites appear to be stationary when viewed from the earth. The satellite consists of transponder that can receive RF signals and transmit them back to the ground at a different angle. A ground station on one side of the ocean transmits signal to the satellite which in turn sends the signal to the ground station on the other side of the ocean.

7.4 Data Transmission across Media;

Transmitting data across media implies sending bits through the transmission medium. Physically, the data is sent as electric signals, radio waves or as light signals. Let's now discuss the use of electric current to transfer digital information. For this, the bits are encoded and sent as characters.

7.4.1 Transmission Modes

The direction in which data can be transmitted between any two linked devices is of three types—
(1) Simplex,
(2) Half-duplex, and
(3) Full-duplex, or duplex.

Simplex transmission is unidirectional data transmission. Of the two linked devices, only one of them can send data and the other one can only receive data.

Half–duplex transmission is bi-directional data transmission, but the linked devices cannot send and receive at the same time. When one device is sending data the other can only receive.

Full-duplex transmission is bi-directional and the linked devices can send and receive data simultaneously. The linked devices can send data and at the same time receive data.

Figure shows the different kinds of transmission modes used for interaction.

7.4.2 Transmission Speed

- When the signals are transmitted between twocomputers, two factors need to be considered— (1) Bandwidth, and (2) Distance.
- Bandwidthis the amount of data that can be transferred through the underlying hardware i.e. the communication medium, in a fixed amount of time. Bandwidth is measured incyclesper second (cps) or Hertz (Hz). The bandwidth of the transmission medium determines the data transfer rate.
- Throughputis the amount of data that is actually transmitted between the two computers. Throughput is specified inbits per second (bps). The throughput capability of the communication medium is also calledbandwidth. The bandwidth of the communication medium is the upper bound on the throughput, because data cannot be sent at a rate more than the throughput of the communication medium.
- Higher throughput is achieved by using a large part of the electromagnetic spectrum (large bandwidth). Technology that uses large part of the electromagnetic spectrum to achieve higher throughput is known asbroadband technology. The technology that uses small part ofthe electromagnetic spectrum is known asbaseband technology.
- Throughput is affected by thedistancebetween the connected computers or devices. Even if a transmission medium is designed for a specific bandwidth, the throughput is affected by the distanceof communication.

- The bandwidth of transmission medium is limited by the distance over which the medium needs to transmit the signal. The bandwidth decreases with the increase in the distance between the connected devices. When a signal has to travel longdistance, the signal strength decreases; the signal strength is utilized to overcome the resistance offered by the connecting medium (cable or wire). The gradual deterioration of signal strength across long distances is called attenuation.
- Moreover, with increasing distance the external disturbance increases, which causes the signal to deteriorate and results in less amount of data to be transferred. The degradation of signal due to internal or external disturbances is called distortion.
- The bandwidth and distance of the transmission medium is selected so that it offers minimum attenuation and minimum distortion.
- Thecat-5 UTP cable has a throughput of 100 Mbps over a distance of 100m. The 10 base2 coaxial cable has a throughput up to 10Mbps over a distanceof 185 m. The10 base5 coaxial cable has a throughput up to 10Mbps over a distance of 500 m.

7.4.3 Fundamentals of Transmission

Telecommunication systems use electromagnetic waves to transfer information. Electromagnetic waves can travel through transmission media like copper wires, fiber optics or as radio waves. They can also travel in vacuum. Wireless communication uses electromagnetic waves for transmission of information. The transmission media through which the waves propagate are not perfect. As a result, the waves propagated via the transmission media get attenuated and distorted.

The information to be transmitted does not always exist in a form that is compatible with the transmission medium. Waves that are compatible with the transmission medium must be generated to carry information. A signal is a wave that is suitable for carrying information over a transmission medium. Signals can be electric signals, light signals, electromagnetic signals or radio signals. Electric signals are used to carry information through copper wires, light signals for fiber optic cables, and radio signals for carrying information in free space. Electrical signals have limited bandwidth and cannot be used in long distance communication. They need to be amplified or regenerated. Light signals have a high bandwidth and are suited for long distance communication.

7.4.3.1 Analog and Digital Signals

- Information carrying signals are of two types—(a) analog signal, and (b) digital signal.
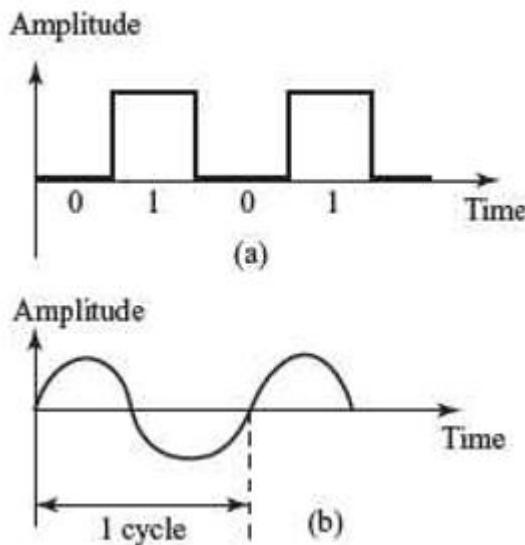
Fig: (a) Digital Signal   (b) Analog Signal

- *Analog Signal:* An analog signal is a wave that continuously changes its information carrying properties over time. The wave may vary in amplitude or frequency in response to changes in sound, light, heat, position, or pressure etc. For example a telephone voice signal is analog. The intensity of the voice causeselectric current variations. At the receiving end, the signal is reproduced in the same proportion.
- *Digital Signal:* A digital signal is a wave that takes limited number of values at discrete intervals of time. Digital signals are non-continuous, they change in individual steps. They consist of pulses or digits with discrete levels or values. The value of each pulse is constant, but there is an abrupt change from one digit to the next. Digital signals have two amplitude levels called nodes. The value of which are specified as one of two possibilities such as 1 or 0, HIGH or LOW, TRUE or FALSE, and so on.
- Analog and digital signals are compared on the basis of—(1) impact of noise, (2) loss of information, and (3) introduction of error.
- Analog signal has the potential for an infinite amount of signal resolution. Another advantage with analog signals is that they can be processed more easily than their digital equivalent. The primary disadvantage of the analog signals is the noise. The effects of noise create signal loss and distortion, which is impossible to recover, since amplifying the signal to recover attenuated parts of the signal, also amplifies the noise. Even if the resolution of an analog signal is higher than a comparable digital signal, the differencecan be overshadowed by the noise in the signal. In digital systems, degradation can not only be detected, but corrected as well.
- *Amplifier* is any device or a circuit that changes, usually increases, the amplitude of an analog signal.
- *Repeater* is an electronic device that receives a signal and retransmits it at a higher level and/or higher power, so that the signal can cover longer distances. With physical media like Ethernet or Wi-Fi, data transmissions can

only span a limited distance before the quality of the signal degrades. Repeaters attempt to preserve signal integrity and extend the distance over which data can safely travel. Actual network devices that serve as repeaters usually have some other name. Active hubs, for example, are repeaters. Active hubs are sometimes also called "multiport repeaters," but more commonly they are just "hubs."

7.4.3.2 Modulation and Demodulation

- *Modulation:* Signals consist of two components—the information signal and the carrier signal. The transmission of any signal oversome communication medium usually involvesmodulation of a carrier. Prior to their transmission the information signal and the carrier signal are combined and the process of combining these two signals is calledmodulation. Characteristics of the carriersignal are varied in proportion to the amplitude of the information-carrying signal. Modulation results in the transfer of the signal information to higher frequency carrier signal. In simple English terms, the information signal sits on top of the carrier signal and rides on it from the receiver to the transmitter.

- *Need for Modulation:* Let's understand the need for modulation by using a simple example. Stereophonic radio signal consist of frequency ranges from 30 Hz (Hertz) to 15 KHz (Kilo Hertz). Hence they need a bandwidth of 15 KHz. If ten different radio stations start transmitting their voice signals between 30 Hz and 15 KHz frequencies, then a combination of these signals would only create noise and the receiver would not be able to discriminate between the signals of each radio station. To overcome this, usually the FM broadcast band, used for broadcasting FM radio stations, goes from 87.5 to 108.0 MHz (Mega Hertz). For example a radio channel-1 may be broadcast using a carrier signal of 102 MHz and would typically use band of frequencies between 101.9 to 102.1 MHz. Radio channel2 using a carrier signal of 102.2 MHz would use band of frequencies between 102.1 and 102.3 MHz. Similarly for other channels, the same method of allocation would be followed.This eliminates the problem of discrimination and decoding signals of each of the radio stations at the receiving end.

There are three primary reasons which necessitate modulation:

 i. To make efficient use of the lines or media used for communication
 ii. To makeradio communications feasible: The lower the frequency of signal, the larger is the size of the antenna needed for transmission and reception. A signal of 10 KHz would require an antenna whose dimensions are in the range of a few kilometers.
 iii. To simplify signal processing: It is simpler to design electronic systems for narrow frequency bands.

Modulation is of three kinds, which are defined as follows:

- **Amplitude Modulation**—The amplitude of the carrier wave is modified in proportion to the data signal. The frequency and phase of the carrier signal remains unchanged.
- **FrequencyModulation**—The frequency of the carrier signal is modified in proportion to the data signal. The amplitude and phase of the carrier signal remains unchanged.
- **Phase Shift Modulation**—The phase of the carrier signal is modified in proportion to the data signal. The amplitude and frequency of the carrier signal remains unchanged.
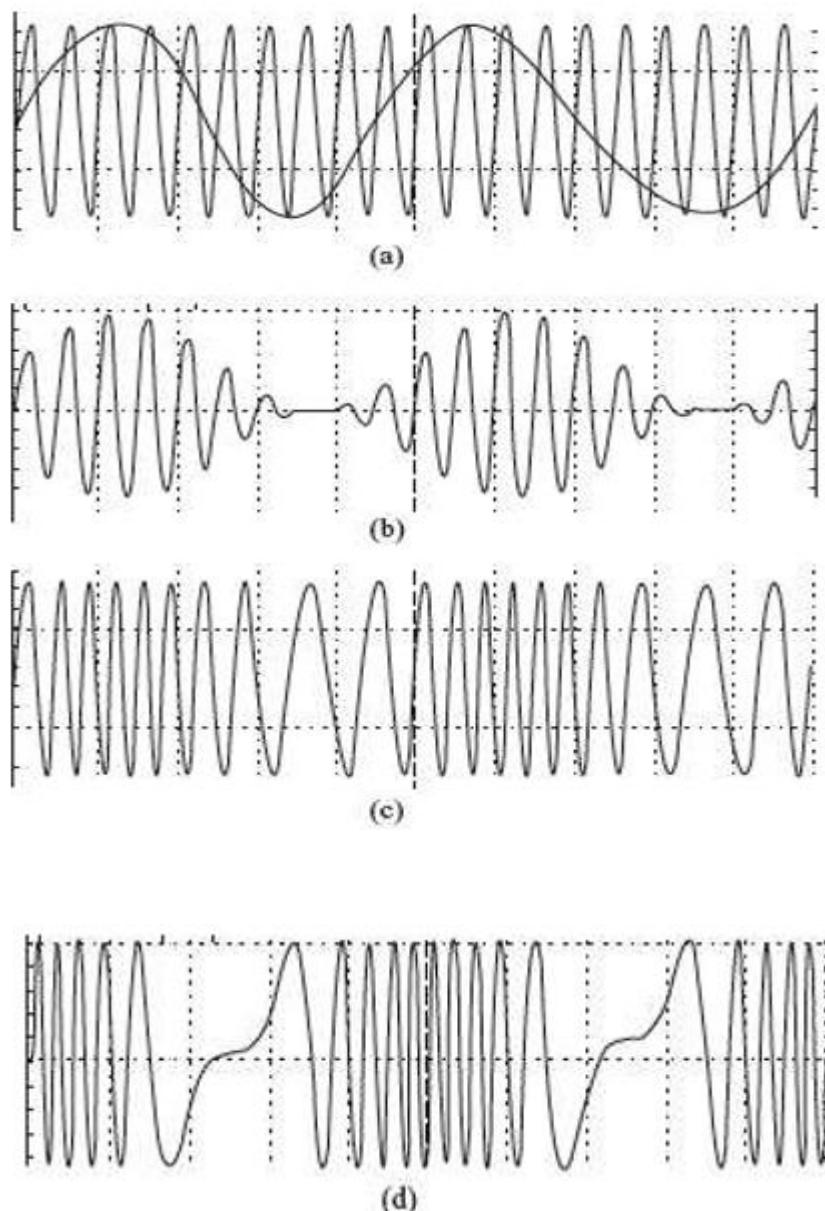

Fig: (a) Carrier wave with signal (b) Amplitude Modulation (c) Frequency Modulation (d) Phase Modulation

7.5 Data Transmission and Data Networking;

Data transmission at physical level involves the hardware required for handling individual bits and encoding bits in signals. The details of the underlying hardware are generally handled by the engineers who design the hardware.

Any two devices directly linked via a communication medium (point to point communication) can send and receive data, to and from each other respectively. If a large number of computers need to interact with each other, point to point communication will require direct link between all the computers. This is not a practical solution. The communication circuits and the associated hardware required for communication (like modem) are expensive. Moreover, there may not be a need to transmit data all the time, which will result in the communication medium lying idle for most of the time. For long distance communication, instead of point to point connection, a network of nodes is used as a communication medium. The different computers attached to the network share the communication facility.
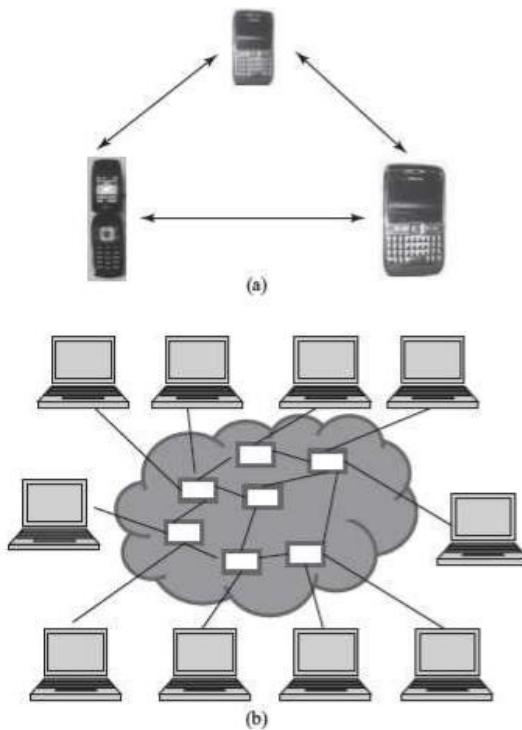
Fig: (a) Point to point communication (b) Switching

The computer network provides a convenient interface that handles sending of multiple bytes of data across the network instead of handling data transmission at physical level.

7.5.1 Switching
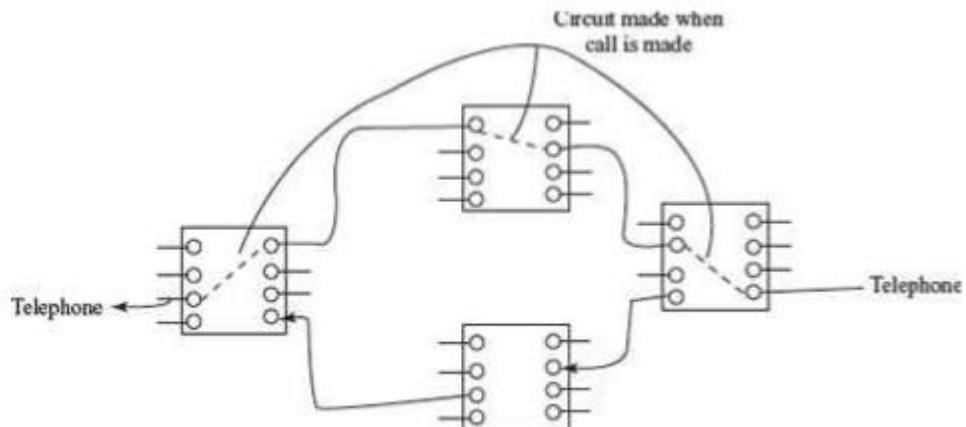
A network cannot allow or deny access to a shared communication facility. All computers attached to the network can use it to send and receive data. Networks allow sharing of communication medium using switching. Switching routes the traffic (data traffic) on the network. It sets up temporary connections between the network nodes to facilitate sending of data. Switching allows different users, fair access to the shared communication medium. There are three kinds of switching techniques—

(1) Packet switching,
(2) Circuit switching, and
(3) Message switching. Computer networks generally use packet switching, occasionally use circuit switching but do not use message switching.
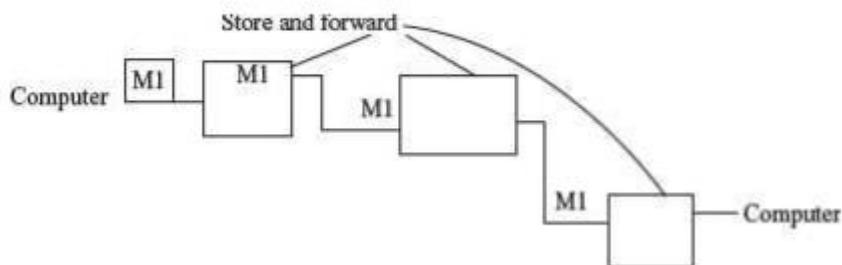
7.5.1.1 Circuit Switching

Circuit switching sets up end-to-end communication path between the source and the destination, before the data can be sent. The path gets reserved during the duration of the connection. Circuit switching is commonly used in the telephone communication network.
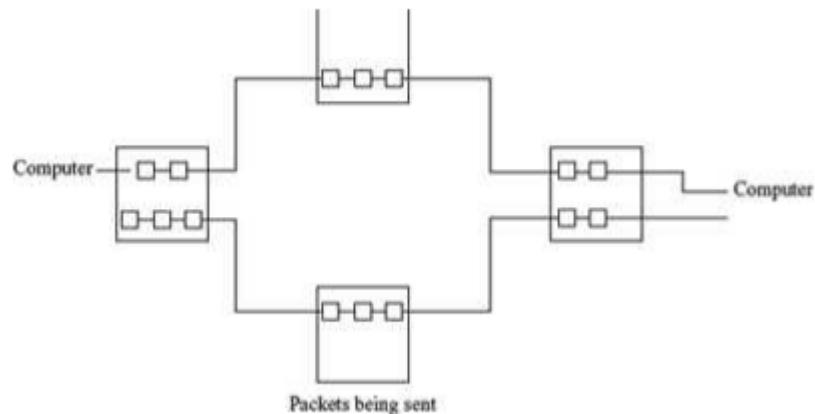
7.5.1.2 Message Switching

Message switching does not establish a physical path in advance, between the sender and the receiver. It uses the 'store and forward' mechanism. In this mechanism, the network nodes have large memory storage. The message is received from the sender and stored in the network node, and when it finds a free route, it forwards the message to the next node till it reaches the destination. Message switching requires large data storage capacity and incurs delay in storing and forwarding of message. Message switching may block the network nodes for a long time. They are thus not suitable for interactive communication. Message switching is no more used in computer networks.

7.5.1.3 Packet Switching

Like message switching, packet switching does not establish a physical path between the sender and the receiver, in advance. Packet switching also uses the 'store and forward' mechanism. However, instead of a complete message, packets are sent over the network. Packet switching splits a message into small "packets" of defined size to be sent over the network. Each packet is numbered.

Packets being sent

7.6 Computer Network;

A computer network is an interconnection of two or more computers that are able to exchange information. The computers may be connected via any data communication link, like copper wires, optical fibers, communication satellites, or radio links. The computers connected to the network may be personal computers or large main frames. The computers in a network may be located in a room, building, city, country, or anywhere in the world.

7.7 Network Types;

The mostly classification is based on the distance covered, geographical area and bandwidth of the data transmission. This type of network has specific features on its own types. Some of them used within room and some of them used to share data from one continent to another. There are different types of computer network on the basis of geographical area covered.

A. **Personal Area Network (PAN):** PAN is the computer network used to communicate among electronic devices, including phones, television, personal digital assistance (PDA) etc. The range of these devices bounds within a room, ranging upto a few meters. The PAN can be use to communicate among personal devices themselves. Personal area networks may be wired or wireless in nature. Infrared communication (Remote controller), Bluetooth, Z-Wave and ZigBee are the common example of wireless personal area network.
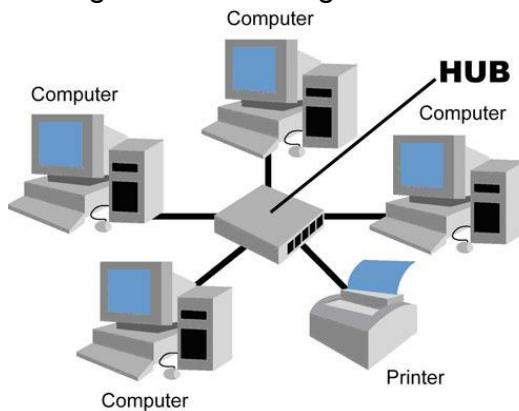
Advantages of PAN:
i. It has high bandwidth.
ii. It is easy to troubleshoot because minimum devices are used.
iii. It has low cost in comparison to other network types.
iv. It has minimum chance of hacking or data lost.

Disadvantages of PAN:
i. It cannot use for general purpose.
ii. The transmission generally has one to one communication.

B. **Local Area Network (LAN):** A group of computer and other devices dispersed over a relatively limited area and connected by a communications link that enables any device to interact with any other on the network. LANs commonly include PCs and shared resources such as laser printer and hard disks. The devices on a LAN are known as nodes and the nodes are connected by cables through which messages are transmitted.

Advantages of LAN:
i. It provides high bandwidth for resource sharing.
ii. It provides cheaper and reliable communication medium.
iii. It provides a great deal of design flexibility, easy maintenance and cost efficiency.
iv. It can support centralized data storage.
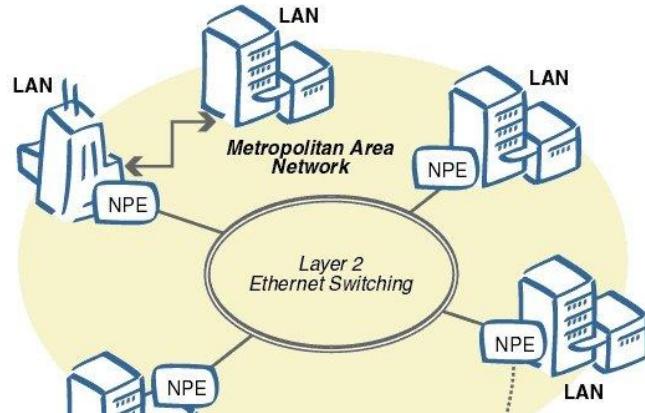
Disadvantages of LAN:
i. Large number of nodes cannot handle efficiently.
ii. Fault on network can lost data and information.
iii. One time installation cost is high.
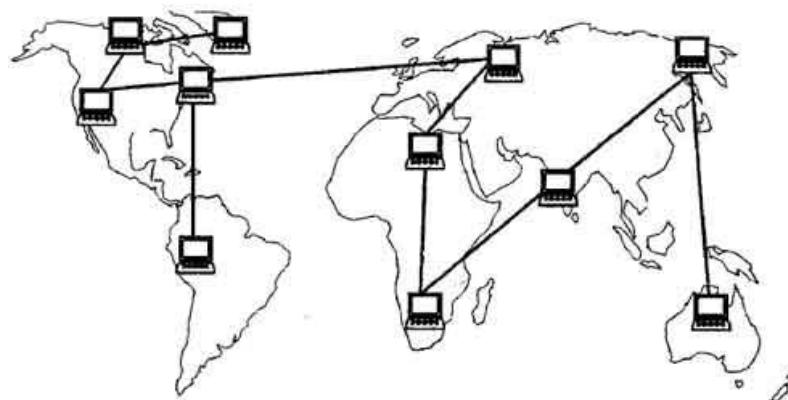
C. **Metropolitan Area Network (MAN):** A high speed network that can carry voice, data, video and images up to 512 Mbps or faster over distances up to 75 Km. Based on network architecture, the transmission speed can be higher for shorter distances and lower for long distances. A MAN can include one or more LANs as well as telecommunication equipment, microwave relay

stations and satellite base stations. A MAN is larger than LANs and smaller than Wide Area Network (WAN) and operates at high speed. Cable television, cable internet, wireless internet, telephone exchange station, bank networking within city is the most common examples of MAN.

D. Wide Area Network (WAN): A geographically widespread network capable for communicating and sharing all types of data and information all around the world is called WAN. WAN is a large network also called network of network and consists large number of LANs and MANs. The span of WAN is in thousands kilometers around the world connecting countries and continents. The bandwidth is lesser in comparison to the LANs and MANs. The operator of the WAN is by internet service provider (ISP) or by a telephone company.

Wide area network

Difference between LAN and WAN

S.No.	Local Area Network	S.No.	Wide Area Network
1	Area covered within a local site.	1	Distance up to thousands of kilometers.
2	Higher data rates higher than 10 Mbps to 1 Gbps even more.	2	Typical data rates less than 1 Mbps.
3	It has low error rates.	3	It has higher error rates.
4	It uses digital signal over private cables (generally CAT5,6,7)	4	It uses telephone to optical fiber back bone.
5	It is managed by the same	5	It is generally managed by
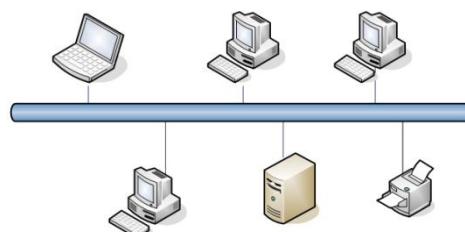
|				
	company who owns it.			organizations apart from users.
6	It uses simple protocols, low cost devices and low cost installation.	6		It uses complex protocols, expensive devices and high cost installation.
7	It can support limited number of hosts (computers) generally less than 1024.	7		It can support large number of hosts theoretically not bound in numbers i.e. infinite.
8	For example: star topology, cellular topology etc.	8		For example: Internet and intranet.
9	Generally, LANs use wireless or digital transmission.	9		WANs use digital or analog signal transmission.

7.8 Network Topology;

Network topology is the layout pattern or configuration of interconnected various components. Most of the all network topology is used to create Local Area Network (LAN) and connect two distinct LANs between computer and computer related components. Hence, topology refers to the shape of a network. How different nodes in a network are connected to each other and how they communicate with each other is determined by the network topology. Some of the topologies are as follows:

A. **Bus Topology:** In bus topology all the devices are connected to a central cable called bus or backbone. The bus topology connects workstations using a single cable. Each workstation is connected to the next workstation in point-to-point fashion. Each node monitors the activities on the bus. The data sent are detected by all the nodes but accepted only those nodes which are addressed. A malfunction of node cannot disturb the network operation.
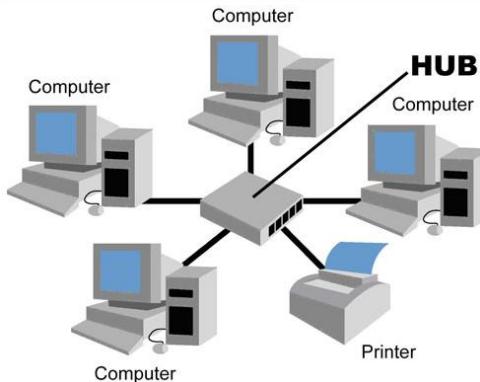
Advantages of Bus Topology:
i.   Installation is cheap and easy.
ii.   Addition of nodes/new connection is easily made.
iii.   Less cable is required.
iv.   If one computer down it does not affect the network.

Disadvantages of Bus Topology:
i.   It is used only for small network.
ii.   Fault identification is difficult.
iii.   A single fault in cable disturbs network transmission.
iv.   All computers share the same path hence performance decreases while network grows.

B. Star Topology: Star topology uses a central hub through which all the components are connected. The central hub is the host and the end of each connection is a terminal. Nodes communicate across the network by passing data through the hub. The central hub makes decision to transmit data around the network and make simple communication channel among devices.
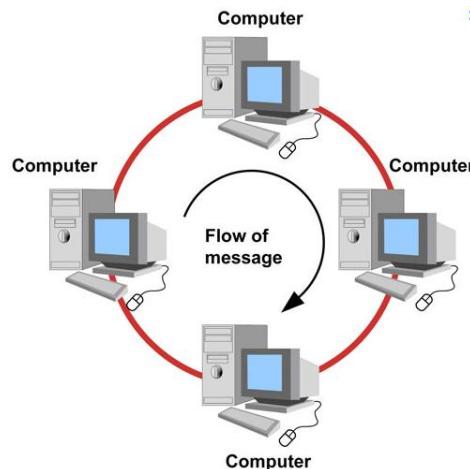
Advantages of star topology:
i.   Installation and configuration of network is easy.
ii.   Less expensive when compare to Mesh Topology.
iii.   Fault in the network can be easily tracked.
iv.   Expansion and modification of star topology is easy.
v.   Single computer failure does not affect the network.
vi.   It supports multiple cable type like STP, UTP, telephone cable etc.

Disadvantages of star topology:
i.   Failure of central hub brings the entire network halt.
ii.   More cabling is required in comparison to tree, bus or ring topology.

C. Ring Topology: All the devices are interconnected to one another in the shape of closed loop so that each device is connected directly to other two devices, forming a ring is called Ring Topology. The data transmitted on the closed ring is in one direction only. Information travels around the ring from one workstation to the next. Each packet of data sent by fixing its destination address, when data arrives in one computer it simply checks destination address is same as its own. If both address matches then it accepts the data otherwise it sends the data to other workstation in the ring.
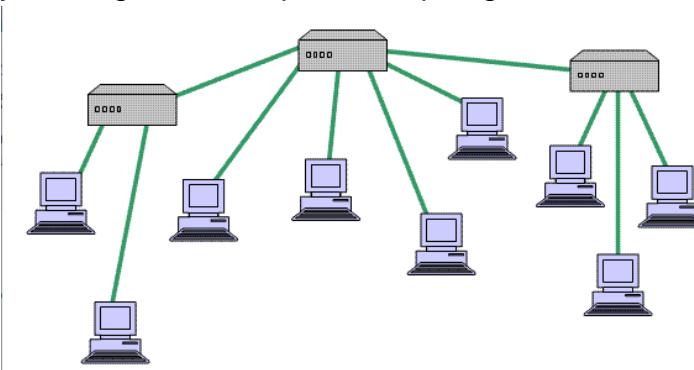
Advantages of ring topology:
i.   It is easy to install and modify the network.
ii.   The fault isolation is simplified.
iii.   No signal/data loss in ring topology.

Disadvantages of ring topology:
i.   Adding and removing computers disturbs the entire network.
ii.   A break in the ring can stop transmission in the network.
iii.   It cannot use for large scale network

D. **Tree Topology:** Tree topology is a LAN topology in which only one path exists between any two nodes on the network. The pattern of the connection branching further and construct tree. The top most devices from where roots are generated is called root of the network. The last ending point of the network which does not have further nodes is called leaf node. It is a type of hybrid topology forming after multiple star topologies.

Advantages of Tree Topology:
i.   Installation and configuration of network is easy.
ii.   The fault can be easily tracked.
iii.   Expansion of the network is easy.
iv.   It supports for multiple cable type.
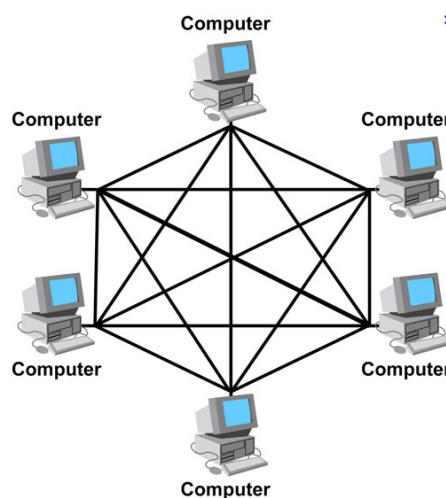
Disadvantages of Tree Topology:
i.   Failure of root brings communication breakup.

ii.   The more cable required for interconnection.
iii.   More devices are needed to interconnect topology.
iv.   Data bandwidth is higher in root nodes.

E. **Mesh Topology:** Devices are connected in one to one system between multiple nodes. In well-connected mesh topology, every node has a connection to every other node in the network. Mesh topology is used in critical connection of host computers. Alternate path allow each computer to balance the load of other computer systems in the network by using more than one connection path available. A fully connected mesh network has n*(n-1)/2 physical links between n devices.
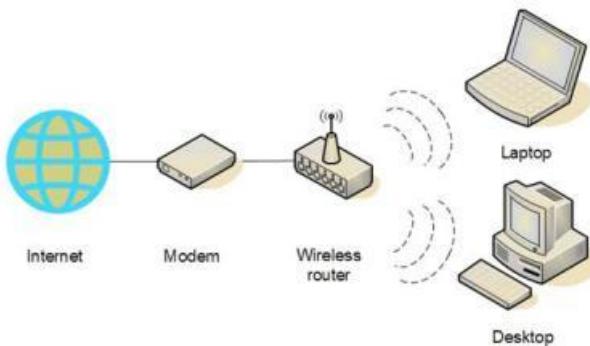
Advantages of Mesh topology:
i.   Use of dedicated links eliminates traffic problems.
ii.   Failure in one of the computer does not affect the entire network.
iii.   Point-to-point link makes fault isolation easy.
iv.   It is robust.
v.   Privacy between computers is maintained as message travel along dedicated path.

Disadvantages of Mesh topology:
i.   The length of cabling required is high.
ii.   A large number of input/output ports are required.
iii.   Cost of the networking is high because of cable and ports.

F. **Cellular topology:** Cellular topology divides the area into cells. In wireless media, the range or geographical area of wireless device within certain boundary is called cell. Each cell represents the network area to communicate between devices. All the devices communicate via the central hub called access point. Multiple access points create a broader range of network to share among devices. This topology is applicable only in case of wired medium is not accessible.

Advantages of cellular topology:
i.   Troubleshooting is easy.
ii.   Cost is low because wire is not used.
iii.   Good looking network topology because no wire is used.

Disadvantages of cellular topology:
i.   If access point (wireless hub) fails all services are affected.
ii.   Generally, the bandwidth is less than star topology.

OSI model
- **Physical Layer**—This layer specifies the basic network hardware. Some of the characteristics defined in the specification are—interface between transmission media and device, encoding of bits, bit rate, error detection parameters, network topology, and the mode of transmission (duplex, half-duplex or simplex).
- **Data Link Layer**—This layer specifies the functions required for node-to-node transmission without errors. It specifies the organization of data into frames, error detection in frames during transmission, and how to transmit frames over a network.
- **Network Layer**—The network layer specifies the assignment of addresses (address structure, length of address etc.) to the packets and forwarding of packets to the destination i.e. routing.
- **Transport Layer**—It specifies the details to handle reliable transfer of data. It handles end-toend error control and flow control, breaking up data into frames and reassembling the frames.
- **Session Layer**—The session layer maintains a session between the communicating devices. It includes specifications for password and authentication, and maintaining synchronization between the sender and the receiver.
- **Presentation Layer**—This layer specifies the presentation and representation of data. Its functions include translation of the representation of the data into an identifiable format at the receiver end, encryption, and decryption of data etc.
- **Application Layer**—This layer specifies how an application uses a network. It deals with the services attached to the data. It contains the protocols used by users like HTTP, protocol for file transfer and electronic mail.
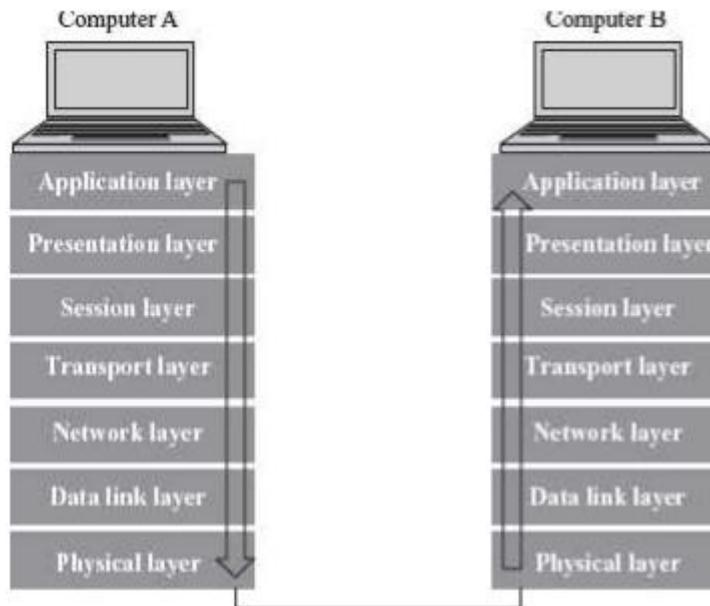
Fig: OSI reference model

7.9 Communication Protocol;

A network/communication protocol is a formal description of digital message formats and the rules for exchanging those messages in between computer system. Protocols define a set of formal rules describing how to transmit data especially across a network. Low level protocol define the electrical and physical standards to be observed, bit- and byte-ordering and the transmission of the bit stream. High level protocols deal with the data formatting, including the syntax of messages, character sets, sequencing of messages etc.

1) **TCP/IP:** The Transmission Control Protocol (TCP) and Internet Protocol (IP) are commonly referred layer 4 and layer 3 protocols respectively. But TCP/IP represents a protocol suit. It is a collection of different protocols and work as an implemented version of protocols. TCP represents transport layer protocol that provides end to end reliable transmission. For the same purpose, TCP includes such functions such as flow control, error control and status exchange information. IP represents unreliable source of transmission between networks. TCP delivers message in reliable manner i.e. message will not discarded or dropped but the delivery time is high. TCP/IP is not a single protocol but a set of more than a dozen protocols. Each protocol within the TCP/IP family is dedicated to a different task.

2) **SMTP:** Simple Mail Transfer Protocol (SMTP) is an application layer protocol used for email transmission across the internet. The mail applications use SMTP for sending messages to a mail server on the internet. SMTP is a delivery protocol only. It cannot pull messages from a remote server on demand. Other protocols such as Post Office Protocol (POP) and the Internet Message Access protocol (IMAP) are specifically designed for retrieving messages and managing mail boxes. However, SMTP have a feature to

initiate mail queue processing on a remote server so that the requesting system may receive any messages destined for it.

3) **POP3:** The Post Office Protocol version 3 (POP3) is an application layer protocol used by local email clients to retrieve email from a remote server over internet connection. POP and IMAP are the most common internet standard protocols for email retrieval. Virtually all modern email clients and servers support both. POP3 is supported by most webmail services such as Hotmail, Gmail and Yahoo! Mail.

4) **FTP:** The File Transfer Protocol (FTP) is an application layer protocol for transferring files between computer systems. FTP is not hardware dependent hence it can work anywhere in the network. FTP is built on client server architecture and utilizes separate control and data connections between the client and server. FTP client and server exist in the network, where client sends commands to the FTP server and the server responds.

5) **HTTP:** It is the underlying protocol used by the World Wide Web. HTTP (Hyper Text Transfer Protocol) defines how messages are formatted and transmitted over the Internet. It decides what actions should be taken on web-servers and web-browsers for the various commands. HTTP has been in use by the World Wide Web since 1990. HTTP is called a stateless protocol because each command is executed independently. The HTTP is a networking protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web.

6) **Telnet:** Telnet is a remote login protocol for executing commands on remote host. The telnet protocol runs in a client server mode and uses the TCP for data transmission. The telnet data will encoded before transmission by telnet client.

7.10 Network Devices;

The cables are used to transmit data in the form of signals from one computer to another. But cables cannot transmit signals beyond a particular distance. Moreover there is a need to connect multiple computers and devices. A concentrator is a device having two or more ports to which the computers and other devices can be connected. A concentrator has two main functions—(1) it amplifies the signal to restore the original strength of the signal, and (2) it provides an interface to connect multiple computers and devices in a network. Repeater, hub, switch, bridge, and gateway are examples of network connecting devices.

Two or more LANs using different protocols may not be able to communicate with the computers attached to their network. For example, a LAN connected using Ethernet may not be able to communicate with a LAN connected using Token Ring. Bridge, Router, and Gateway are devices used to interconnect LANs.
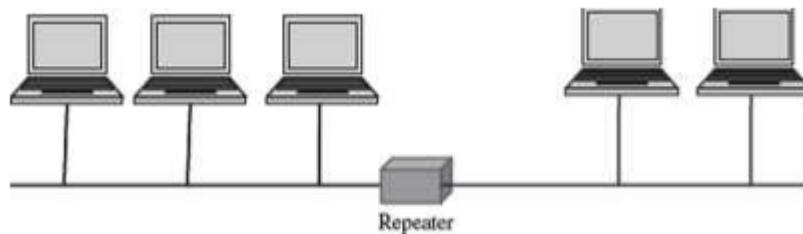
Introduction to Information Technology

1. Network Interface Card

- A Network Interface Card (NIC) is a hardware device through which the computer connects to a network.
- NIC is an expansion card, it can be either ISA or PCI, or can be on-board integrated on a chipset. NIC has an appropriate connector to connect the cable to it. NIC for different LAN are different (NIC for token ring is different from NIC for Ethernet).
- NIC work at both the data link layer and physical layer of the OSI reference model.
- At the data link layer, NIC converts the data packets into data frames, adds the Media ACcess address (MAC address) to data frames. At the physical layer, it converts the data into signals and transmits it across the communication medium. The MAC address is a globally unique hardware number present on the NIC and is specified by the NIC manufacturer.
- NIC depends upon the configuration of the computer, unlike hub or switches that perform independently.

2 Repeater
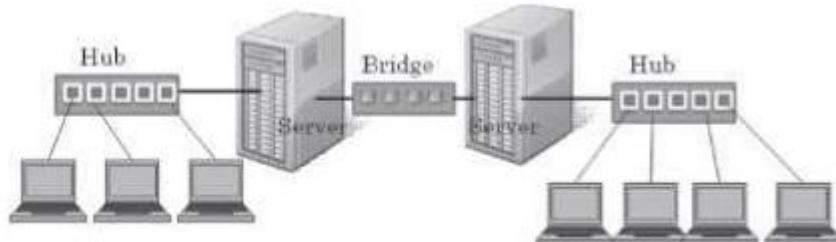
- Repeaters (Figure) are used to extend LAN. It has only two ports and can connect only two segments ofa network. Multiple repeaters can be used to connect more segments. (Segment is a logical section of the same network).

Repeater

- Repeaters operate at the Physical layer of OSI reference model.
- They are useful when computers in a network arelocated far away from each other.
- Repeaters amplify the signal so that the signal is as strong as the original signal. They can thus extend the reach of a network.
- Repeaters cannot be used if multiple computers need to be interconnected or multiple segments need to be interconnected.
- Repeaters cannot identify complete frames. Thus, in addition to the valid transmissions from one segment to another, repeater also propagates any electrical interference occurring on a segment to other segment.

3. Bridge

- Bridge (Figure) is used to connect two LAN segments like a repeater; it forwards complete and correct frames to the other segment. It does not forward any electrical interference signals to the other segment.
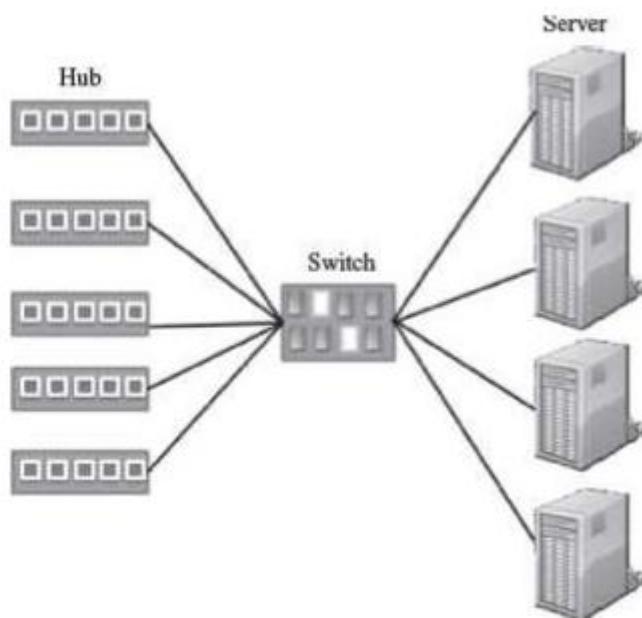
- Bridge forwards a copy of the frame to the other segment, only if necessary. If a frame is meant for a computer on the same segment, then bridge does not forward a copy of the frame to other segment.
- Bridge connects networks that use different protocol at the Data Link Layer. The frame format of data in the two networks is different. The bridge converts the frame format before transmitting data from one network to another, with translation software included in the bridge.
- A bridge is also used to divide a network into separate broadcast domains to reduce network traffic while maintaining connectivity between the computers.

4. Hub
- It is like a repeater with multiple ports. But, hub does not amplify the incoming signal.
- Hub operates at the Physical layer of OSI reference model, hence treats data as a signal.
- Hubs are used to connect multiple segments of the same network.
- Hubs are also used to connect computers to network that use Star topology.
- The port on the hubs can alsobe used to connect another hub, switch, bridge or router.
- Hubs increase the network traffic because they broadcast data to all the device connected all the ports of the hub.
- It is preferable to use a hub in a small LAN having about 8–10 computers connectedto it.
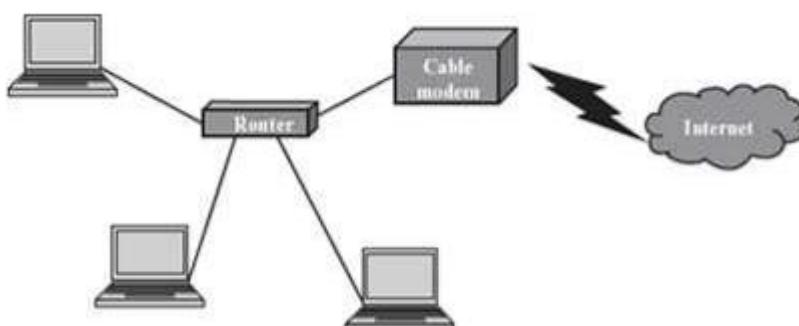
5. Switch
- Like hub, switch also connects multiple computers in a network or different segments of the same network. A hub simulates a single segment that is shared by all computers attached to it (hub transmits the data to all computers attached to it). In a hub, at most two computers can interact with each other at a given point of time. However, in a switch each computer attached to a switch has a simulated LAN segment.
- Switches (Figure) work at the Data Link Layer of the OSI reference model.Hence, switches consider data as frames and not as signals

- A data frame contains the MACaddress of the destination computer. A switch receives a signal as a data frame from a source computer on a port, checks the MAC address of the frame, forwards the frame to the port connected to the destination computer having the same MAC addresses, reconverts the frame back into signal and sends to the destination computer. (Switching is a technique that reads the MAC address of the data frame and forwards the data to the appropriate port). Switches, thus, regenerate the signals.
- Since a switch does not broadcast data, but sends the data from the source computer to the destination computer, a half of the computers attached to the switch can send data at the same time.
- Switch is also referred to as a multi-port bridge. In general, bridges are used to extendthe distance of the network, and switches are primarily used for their filtering capabilities to create a multiple and smaller virtual LAN (a LAN segment can be connected to each port of the switch) from a single large LAN.

6. Router
- Router (Figure) is used to connect heterogeneous networks.
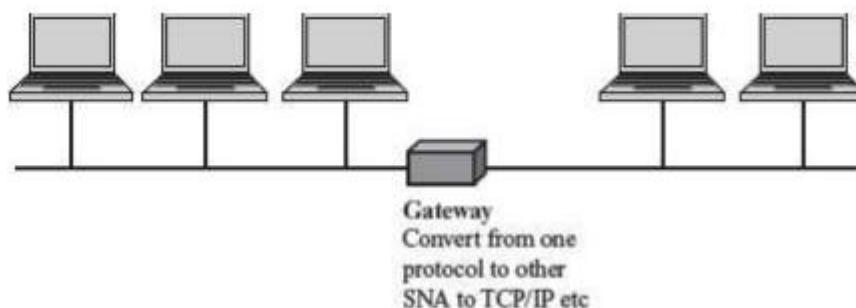
- A router has a processor, memory, and I/O interface for each network to which it connects.

- A router connects networks that use different technologies, different media, and physical addressing schemes or frame formats.
- A router can connect two LANs, a LAN and a WAN, or two WANs.
- A router is used to interconnect the networks in the Internet.
- Router operates at the Network layer of the OSI model (layer 3).
- Physically, a router resembles a bridge, but is different from a bridge. A router determines which way is the shortest or fastest in a network, and routespackets accordingly. Since it works at the Network layer, it moves packets based on the IP addresses etc. In contrast, a bridge connects two LANs almost permanently.

7. Gateway
- Gateway(Figure) is a generic term used to represent devices that connect two dissimilar networks.

Gateway
Convert from one
protocol to other
SNA to TCP/IP etc

- A gateway at the transport layer converts protocols among communications networks. It can accept a packet formatted for one protocol and convert it to a packet formatted for another protocol, before forwarding it. An application gateway can translate messages from one format to the other.
- A gateway can be implemented inhardware, software, or in both hardware and software. Generally, gateway is implemented by software installed within a router.

The network connecting devices—repeater and hub operate at the physical layer, bridge and switch operate at the data link layer, and the router operates at the network layer of the OSI model.

7.11 Wireless Networking

Wireless technology, as the name suggests, is used to establish a wire-free connection or communication between two or more devices. In contrast to the wired technology where data is encoded as electric current and signals travel through wires, in wireless technology data is encoded on electromagnetic waves that travel through air. The wireless technology is used for broadcasting in radio and television communication, for communication using mobile phones and pagers, for connecting components of computers using Bluetooth technology, for Internet connection using Wi-Fi, Wireless LAN, PDA, and in remote controls for television, doors etc.

- Wireless network is a computer network connected wirelessly. The communication is done through a wireless media like radio waves, infrared or Bluetooth.
- The wireless networks have two main components—the wireless access points that include the transmitter along with the area it can cover, and the wireless clients like mobile handsets, laptops with Ethernet cards etc.
- The access point receives data frames from the computers attached to it wirelessly, checks the frames, and transmits them to their destination. The coverage area of a transmitter depends on the output power of the transmitter, its location, and the frequency used to transmit the data. Higher frequencies require a clear line of sight as compared to lower frequencies.
- The speed of wireless connection is determined by the distance of the wireless client device from the access point, the obstruction-free path (walls, trees etc.), interference, and the number of users using the network at a given time.
- Wireless networks can be divided into three categories based on their use:
  - *Bluetooth technology* to connect the different components of the computer in a room, a small office or home.
  - *Wireless LAN* is used to connect computers and devices wirelessly in a LAN, for example, different computers or devices in an office or campus.
  - *Wireless WAN* is used to connect wide area systems, for example access to Internet via mobile devices like cell phone, PDAs and laptops.

Unit 8: The Internet and Internet Services

8.1 Introduction;

Internet is defined as an interconnection of networks. Internet allows computers on different kinds of networks to interact with each other. Any two computers, often having different software and hardware, can exchange information over the Internet, as long as they obey the technical rules of Internet communication. The exchange of information may be among connected computers located anywhere, like military and research institutions, different kinds of organizations, banks, educational institutions (elementary schools, high schools, colleges), public libraries, commercial sectors etc.

8.2 History of Internet;

Internet has evolved from a research prototype project to a full-grown commercial computer communication system. The growth of Internet can be discussed in three steps, as follows:

- Internetworking Protocol—Transmission Control Protocol/Internet Protocol (TCP/IP) in 1970s
- Usenet groups and Electronic mail in 1980s
- World Wide Web in 1990s

The networking of computers has its origin at the US Department of Defense Advanced Research Projects Agency (DARPA).

- During 1970's DARPA developed the ARPANET as a WAN to connect different computers and later to connect computers ondifferent networks (Internetworking). Internetworking became the focus of research at ARPA and led to the emergence of Internet. During their research, DARPA set up design goals for themselves, which included—(1) the ability to interconnect different types of network, (2) to connect through alternate paths if some path gets destroyed, and (3) to support applications of various types like audio, video, text etc.
- Based on the design goals, a protocol namedTransmission Control Protocol/Internet Protocol (TCP/IP)was developed for computer communication (Protocolis a network term used to indicate the set of rules used by a network for communication). TCP/IP has become the protocol for Internet.
- In late 1970s, the US National Science Foundation (NSF) designeda successor to ARPANET, called NSFNET, which was open for use to all university research groups, libraries and museums. This allowed scientists across the country to share data and interact with each other for their research projects. Internet grew exponentially when ARPANET was interconnected with NSFNET.
- In 1980s, many Internet applications like electronic mail, newsgroups, file transfer facility and remote login were developed. TheElectronic mailfacility

allowed users to compose, send, and receive messages. Users having common interests could exchange messages using forums like Newsgroups. The Telnet command allowed users to login to a remote computer. The File Transfer Protocolprogram was used to copy files from one computer to another on the Internet.

- In the early 1990s, a new application World Wide Web (WWW) changed the way in which Internet was used.WWWis a system of creating, organizing, and linking documents, and was created by British scientist Tim Berners Lee. A protocol based on hypertext was developed that allowed the documents and content on WWW to be connected via hyperlink.
- In 1993, Marc Andreessen at the University of Illinois developed theMosaic browser. The WWW along with the browser made it possible to set up number of web pages thatmay consist of text, pictures or sound, and with link to other pages.

8.3 Internetworking Protocol;

- TCP/IP is the communication protocol for the Internet.
- The TCP/IP protocol has two parts: TCP and IP.
- Transmission Control Protocol (TCP) provides reliable transport service, i.e. it ensures that messages sent from sender to receiver are properly routed andarrive intact at the destination.
- TCP converts messages into a set of packets at the source, which are then reassembled back into messages at the destination. For this, TCP operates with the packet switchingtechnique, which is described as follows:
  - o The message is divided into small packets.
  - o Each packet contains address, sequencing information, and error control information.
  - o The address is used to route the packet to its destination.
  - o Since multiple users can send or receive information over the same communication line, the packets can arrive out of order at the destination. The sequencing information in the packet is used to reassemble the packets in order, at their destination.
  - o The error control information is used to check that the packet arrived at the destination is the same as that sent from the source (i.e. has not got corrupted)
- Internet Protocol (IP) allows different computers to communicate by creating a network of networks.
- IP handles the dispatch of packets over the network.
- It handles the addressing of packets, and ensures that a packet reaches its destination traveling through multiple networks with multiple standards.

The computers connected to Internet may be personal computers or mainframes; the computers could have a slow or fast CPU, small or large memory, connected to different networks having slow or fast speed. TCP/IP protocol makes it possible for

any pair of computers connected to Internet to communicate, despite their hardware differences.

8.4 The Internet Architecture;

Internet is a network of interconnected networks and is designed to operate without a central control. If a portion of the network fails, connection is made through alternative paths available. The architecture of Internet is hierarchical in nature. A brief description of the architecture of Internet is as follows:

- **Client** (user of computer) at home or in a LAN network is at the lowest level in hierarchy.
- Local **Internet Service Provider (ISP)** is at the next higher level.
- An ISP is an organization that has its own computersconnected to the Internet and provides facility to individual users to connect to Internet through their computers.
- Local ISP is the local telephone company located in the telephone switching office, where the telephone of client terminates. Examples of local ISP in India are Bharat Sanchar Nigam Ltd. (BSNL), Mahanagar Telephone Nigam Ltd. (MTNL), and Airtel.
- The client calls local ISP using a modem or Network Interface Card.
- **Regional ISP** is next in the hierarchy. The local ISP is connected to regional ISP
  - A router is a special hardware system consisting of a processor, memory, and an I/O interface, used for the purpose of interconnecting networks. A router can interconnect networks having different technologies, different media, and physical addressing schemes or frame formats.
  - The regional ISP connects the local ISP's located in various cities via routers. o If the packet received by regional ISP is for a client connected to this regional ISP, then the packet is delivered; otherwise, packet is sent to the regional ISP's backbone.
- **Backbone** is at top of the hierarchy.
  - Backbone operators are large corporations like AT&T which have their own server farms connected to the backbone. There are many backbones existing in the world.
  - The backbone networks are connected to Regional ISP's with a large number of routers through high speed fiber-optics.
  - Network Access Point (NAP) connects different backbones, so that packets travel across different backbones.
  - If a packet at the backbone is for a regional ISP connected to this backbone, the packet is sent to the closest router to be routed to local ISP and then to its destination; otherwise, packet is sent to other backbone via NAP. The packet traverses different backbones until it reaches the backbone of regional ISP for which it is destined.
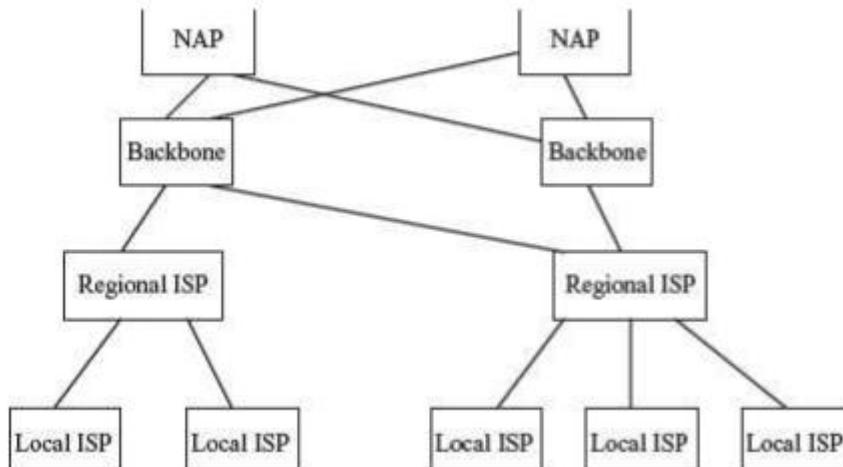
Fig: Internet Hierarchy

8.5 Managing the Internet;

Internet is not controlled by any one person or an organization. A number of organizations manage the Internet. Some of the governing bodies of the Internet and their functions are shown:
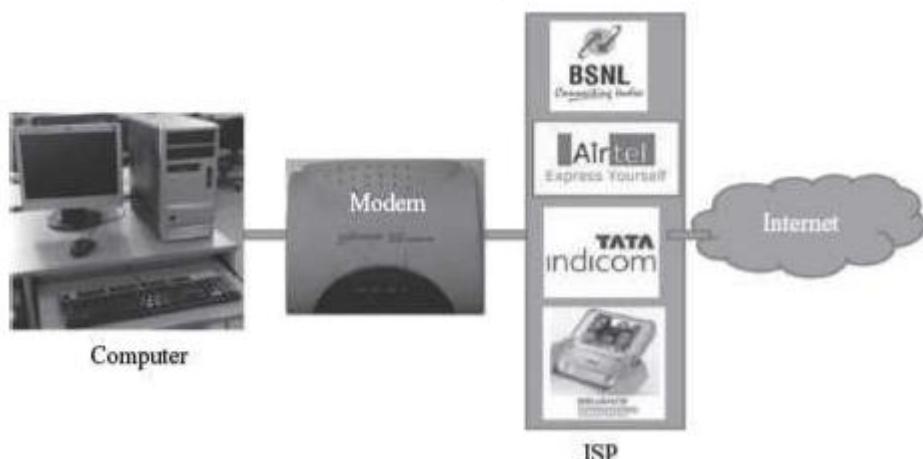
Governing Bodies of Internet and its Functions:

- **Internet Society (ISOC)**
  - Provides information about Internet
  - Responsible for development of standards and protocols related to Internet
- **Internet Architecture Board (IAB)**
  - Advisory group of ISOC
  - Responsiblefor development of Internet architecture
- **Internet Engineering Task Force (IETF)**
  - Community of network designers, operators, vendors, and researchers
  - Responsible for evolution of Internet
  - Open to all individuals
- **Internet Engineering Steering Group (IESG)**
  - Reviews standards developed by IETF
- **Internet Research Task Force (IRTF)**
  - Focuses on research towards the future of Internet (Internet protocol, architecture etc.)
- **Internet Assigned Number Authority (IANA)**
  - Allots IP address to organizations andindividuals
- **Internet Network Information Center (InterNIC)**
  - Responsible for domain name registration
- **World Wide Web Consortium (W3C)**
  - Responsible for development of technologies for World Wide Web

8.6 Connecting to Internet;
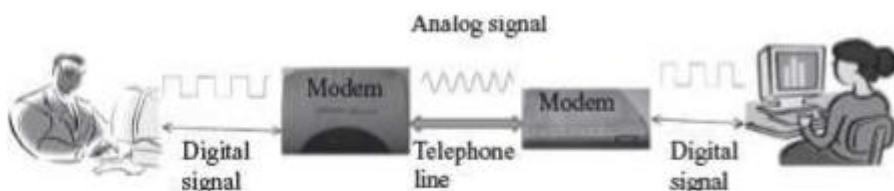
To be able to connect your computer to the Internet, you require—(1) a TCP/IP enabled computer, (2) web browser software, (3) an account with an ISP, (4) a telephone line, and (5) a modem or Network Interface Card (NIC) to connect the telephone line to the computer.

A modem is a device that connects a computer to Internet. A Network Interface Card or NIC is a device that is required to connect a computer to Internet via a LAN or high-speed Internet connection like cable modem or Digital Subscriber Line (DSL). A web browser is a software that allows the user to view information on WWW. WWW is a large-scale, on-line repository of information that the users search using the web browser. Internet Explorer and Netscape Navigator are examples of web browser.

8.7 Internet Connections;

The ISPs provide Internet connections of different types. Bandwidth and cost are the two factors that help you (the user) in deciding which Internet connection to use. Bandwidth is the amount of data that can be transferred through a communication medium in a fixed amount of time. The speed of Internet access depends on the bandwidth. The speed of Internet access increases with the increase in bandwidth. ISPs offer low speed Internet connection like Dial-up connection, and high-speed Internet connection called broadband connection. Broadband are the services with more bandwidth than standard telephone service. DSL, Cable modem, and Integrated Services Digital Network are some of the existing broadband connections, each, having a different bandwidth and cost. Let's now discuss about some of the Internet connections that are nowadays available for Internet access.

8.7.1 Dial-up Access

- Dial-up access is a method ofconnecting to the Internet using an existing telephone line (above Figure). When your computer is connected to the Internet, you cannot receive voice telephone calls on this telephone line during that time.
- In Dial-up access, you are assigned an account on the server of ISP along with some storage space on the disk of server. For example hari@ntc.net.np is an account with an ISP named NTC. You are also assigned a user-id and password.
- You connect to Internet by dialing-up one of the computers of ISP. For this, you use a telephone number provided by ISP and connect via a 56 Kbps modem. The computer that dials-up is the client or remote machine, and the computer of ISP is theserver or host.
- The client enters the user-id and password, and gets connected to the Internet via the ISP.

8.7.2 Leased Line

- Leased line is a dedicated phone linethat connects a computer (also known as gateway) to Internet, using special kind of modems. At the other end, the gateway is connected to a large number of computers, which access the Internet via the gateway.
- The gateway forms a domain on Internet, e.g. niit.com, which is used to provide connection to the other computers on the Internet to connect to it.
- Leased lines provide reliable and high-speed Internet access.
- The entire bandwidth of leased line is reserved for the traffic between gateway and Internet.
- Leased lines are generally used by large organizations and universities that have their own internal network, and have large number of users.
- The leased lines are on-line, twenty-four hours a day and seven days a week. Theleased lines are normally provided on a yearly contract basis. The charges for the leased line are fixed based on many criteria like the bandwidth, number of users etc. The fixed charges do not vary with the actual usage of Internet.

8.7.3 Integrated Services Digital Network (ISDN)

- ISDN is a digital telephone service that can transmit voice, data and control information over an existing single telephone line.
- Internet access is faster using ISDN than Dial-up access.
- ISDN is commonly used for business purposes. You are able to connect a computer, a fax machine or a telephone to a single ISDN line, and also use them simultaneously.
- ISDN is costlier than Dial-up connection. It requires a special phone service and modem.
- Nowadays, ISDN services are largely being replaced by high speed broadband connection.

8.7.4 Digital Subscriber Line (DSL)

- DSL is a broadband connection that allows connecting to Internet over the existing telephone lines. It does not affect your telephone voice services. DSL uses the modem provided by ISP.
- The data transmission speed of DSL ranges from 128 Kbps to 8.448 Mbps.
- Originally, telephone lines were designed for carrying human voice and the whole system worked according to this requirement. All frequencies less than 300 Hz and above 3.4kHz were attenuated, since 300Hz to 3.4kHz is the range for human speech to be clearly audible. When using DSL, a different kind of switch is used that does not filter the frequencies, thus making entire frequency available. DSL uses frequency beyond 3.4kHz for Internet access.
- Asymmetric DSL (ADSL), a variant of DSL, provides high-speed delivery of download data (from Internet to user), than that for upload (from user to Internet), since most users download much more than they upload.
- The bandwidth of connecting wire is divided into three bands—(1) 0–25kHz for regular telephone, (2) 25kHz–200kHz for user to Internet (upload), and (3) 250kHz–1MHz for Internet to the user (download). The available bandwidth for each direction for Internet is divided into channels of 4 kHz.
- DSL is almost ten times faster than Dial-up access and is an always-on connection.

8.7.5 Cable Modem

- The user can connect to the Internet via a cable modem through cable television. The cable modem provides two connections—one for television and other for computer.
- The cable modem sends andreceives data through the coaxial cable which connects the cable modem to the cable service provider. Coaxial cables allow transmission of Internet data, audio, and video, and control over its several channels simultaneously. The user can access the Internet and watch television at the same time.
- Like DSL, cable modem provides high-speed Internet connection. However, while using cable modem, the bandwidth is shared by many users. If many users access the Internet simultaneously then the available bandwidthfor each of the user reduces.

8.8 Internet Address;

A computer connected to the Internet must have a unique address in order to communicate across the Internet. Internet Protocol (IP) address is assigned uniquely to every computer connected to the Internet. IP address is provided by the ISP whose services you use to connect your computer to the Internet. IP address is a string of numbers consisting of four parts, where each part is a number between 0 and 255. An IP address looks like 201.54.122.107. Since IP addresses are numeric, it is difficult to remember everyone's IP address. So, instead of numeric IP address, domain name is used.

Domain name is a text name (string of words) corresponding to the numeric IP address of a computer on the Internet. Domain names are used for the convenience of the user. A domain name combines a group of hosts on the Internet (e.g. Yahoo, Google, MSN etc.), and a top level domain. Some examples of top-level domain are as follows:

com—for commercial organizations,
edu—for educational institutions,
net—for gateways and administrative hosts,
org—for non-profit organizations,
co—for companies, and

Some examples of domain name are google.com, ntc.net.np, tuiost.edu.np and vsnl.net.

Additionally, top-level domain is also provided based on the two-letter Internet country code. For example, in for India, uk for United Kingdom, au for Australia etc. Example of such domain names are du.ac.in, ibm.co.in etc.

In order to translate numeric IP address that identifies a computer on the Internet to a domain name that is convenient for the user to remember, a mapping is needed between the IP addresses and domain names. Domain Name System (DNS) server is a computer having a database that stores the IP addresses and their domain names. Whenever a user uses the domain name, DNS translates it into its corresponding IP address, to access the computer on Internet. For example, DNS translates google.com to the IP address of the computer that houses Google.

8.9 Internet Services;

Internet is a huge de-centralized network that connects computers. Every computer connected to the Internet has a unique address, which helps to identify the computer on the Internet uniquely. Over the years, Internet has grown as the biggest network for communication and provides several services to its users. Each service has its own features and uses. Some of the important services provided by Internet are—World Wide Web, electronic mail, news, chat, and discussion groups.

8.9.1 World Wide Web (WWW)

WWW (also called as Web) is a large scale, online store of information. It is a system of creating, organizing, and linking of documents. Information is stored on WWW as a collection of documents that are interconnected with each other via links. The interconnected documents may be located on one or more than one computer, worldwide, thus, the name world wide web. The features of WWW and terms linked to WWW are given below—
- The documents on web are created inhypertext format. Hypertext facilitates linking of documents.

- The language used to create a hypertext format document is HyperText Markup Language (HTML). HTML allows the designer of the document to include text, pictures, video, images, sound, graphics, movies etc., and also to link contents on the same document or different documents using ahyperlink.
- The hypertext format document is transferred on theWeb usingHyperText Transfer Protocol (HTTP).
- A single hypertext document is called a web page.
- A group of related web pages is called a web site. A web site displays related information on a specific topic.
- The first web page or main page of a website is called Homepage.
- The web pages are stored on the Internet on theWeb Server. Web servers are host computers that can store thousands of web pages.
- The process of storing a web page on a web server is called uploading.
- The process of retrieving a web page from a web server onto the user's computer is downloading.
- The web pages stored on web server on the Internet, can be viewed from the user's computer using a tool called Web browser.
- Every web page is identified on Internet by its address, also called Uniform Resource Locator (URL).
- A web portalis a web site that presents information from different sources and makes them available in a unified way. A web portal enables the user to search for any type of information from a single location, i.e. the home page of the web portal. A web portal generally consists of a search engine, e-mail service, news, advertisements, and an extensive list of links to other sites etc. www.msn.com and www.google.co.in (igoogle) are popular web portals.

8.10 Uses of Internet;

Internet is used for different purposes by different people. Some uses of the Internet are listed below:

- E-Commerce (auction, buying, selling products etc.)
- Research (on-line journals, magazines, information etc.)
- Education (e-learning courses, virtual classroom, distance learning)
- E-Governance (online filing of application (Income Tax), on-line application forms etc.)
- On-line ticket booking (airplane tickets, rail tickets, cinema hall tickets etc.)
- On-line payments (credit card payments etc.)
- Video conferencing
- Exchange of views, music, files, mails, folders, data, information etc.
- Outsourcing jobs (work flow software)
- Social networking (sites likefacebook, linkedin, twitter, orkut
- E-Telephony (sites likeskype)

Fig: Uses of internet

8.11 Introduction to Internet of Things (IoT),

Internet of Things or IoT is an architecture that comprises specialized hardware boards, Software systems, web APIs, protocols which togather creates a seamless environment which allows smart embedded devices to be connected to internet such that sensory data can be accessed and control system can be triggered over internet.

Also devices could be connected to internet using various means like WiFi, Ethernet and so on. Furthermore devices may not needed to be connected to internet independently. Rather a cluster of devices could be created (for example a sensor network) and the base station or the clusterhead could be connected to internet. This leads to more abstract architecture for communication protocols which ranges from high level to low level.

Internet of Things represents a general concept for the ability of network devices to sense and collect data from the world around us, and then share that data across the Internet where it can be processed and utilized for various interesting purposes.

Some also use the term industrial Internet interchangeably with IoT. This refers primarily to commercial applications of IoT technology in the world of manufacturing. The Internet of Things is not limited to industrial applications, however.
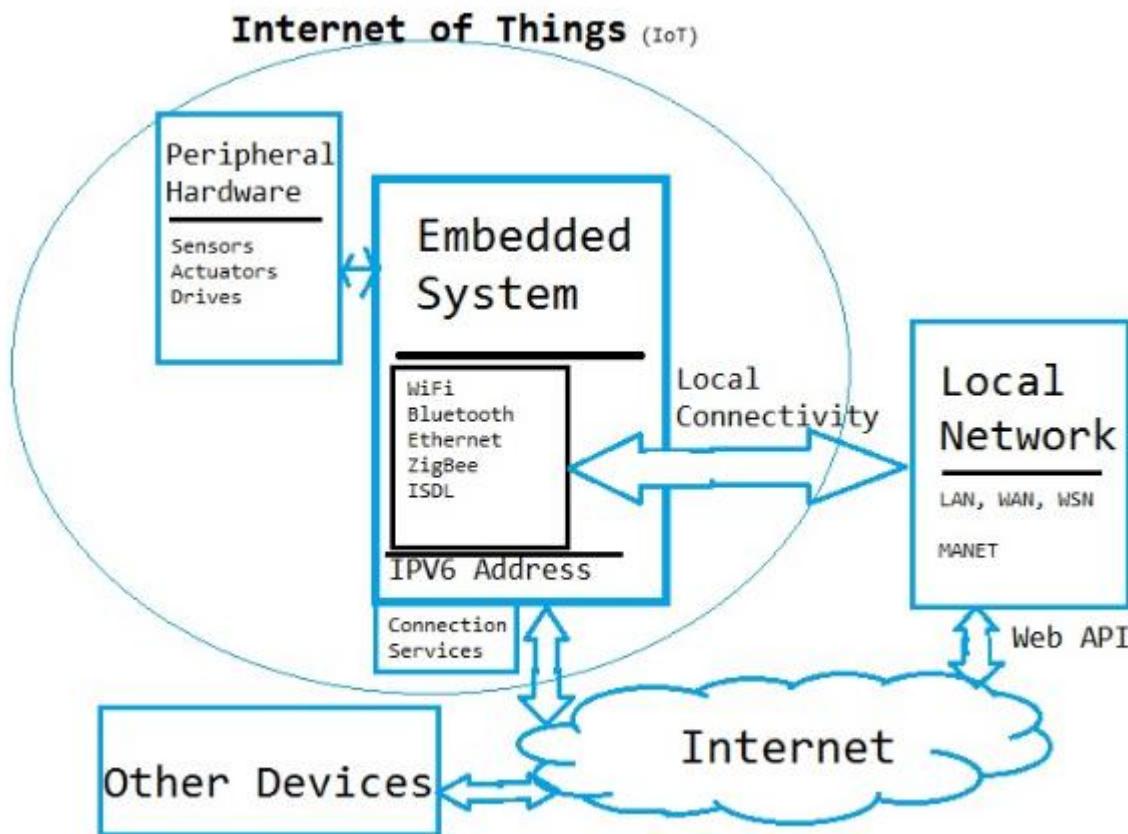
Fig: Internet of Things (IoT) basic architecture

What the Internet of Things Can Do for Us

Some future consumer applications envisioned for IoT sound like science fiction, but some of the more practical and realistic sounding possibilities for the technology include:

- receiving warnings on your phone or wearable device when IoT networks detect some physical danger is detected nearby (think: smart smoke detectors)
- self-parking automobiles (think: Volvo S90)

- automatic ordering of groceries and other home supplies (think: Amazon Dash Wand)
- automatic tracking of exercise habits and other day-to-day personal activity including goal tracking and regular progress reports (think: fitness trackers)

Potential benefits of IoT in the business world include:

- location tracking for individual pieces of manufacturing inventory

- fuel savings from intelligent environmental modeling of gas-powered engines
- new and improved safety controls for people working in hazardous environments

8.12 Wearable Computing,

Wearable computing is the study or practice of inventing, designing, building, or using miniature body-borne computational and sensory devices. Wearable computers may be worn under, over, or in clothing, or may also be themselves clothes (i.e. "Smart Clothing").

Wearable computing is A term that refers to computer-powered devices or equipment that can be worn by a user, including clothing, watches, glasses, shoes and similar items. Wearable computing devices can range from providing very specific, limited features like heart rate monitoring and pedometer capabilities to advanced "smart" functions and features similar to those a smartphone or smartwatch offers.

These more advanced wearable computing devices can typically enable the wearer to take and view pictures or video, read text messages and emails, respond to voice commands, browse the web and more. While wearable computing devices are only just now starting to emerge from the realm of science fiction into reality, rumored devices like Google Glasses and the Apple iWatch may soon bring advanced wearable computing devices into the mainstream.

8.13 Cloud Computing,

Cloud computing is the on-demand delivery of compute power, database storage, applications, and other IT resources through a cloud services platform via the internet with pay-as-you-go pricing.

Cloud Computing Basics

Whether you are running applications that share photos to millions of mobile users or you're supporting the critical operations of your business, a cloud services platform provides rapid access to flexible and low cost IT resources. With cloud computing, you don't need to make large upfront investments in hardware and spend a lot of time on the heavy lifting of managing that hardware. Instead, you can provision exactly the right type and size of computing resources you need to power your newest bright idea or operate your IT department. You can access as many resources as you need, almost instantly, and only pay for what you use.

How Does Cloud Computing Work?

Cloud computing provides a simple way to access servers, storage, databases and a broad set of application services over the Internet. A Cloud services platform such as Amazon Web Services owns and maintains the network-connected hardware required for these application services, while you provision and use what you need via a web application.

Advantages and Benefits of Cloud Computing

Trade capital expense for variable expense
Instead of having to invest heavily in data centers and servers before you know how you're going to use them, you can only pay when you consume computing resources, and only pay for how much you consume.

Benefit from massive economies of scale
By using cloud computing, you can achieve a lower variable cost than you can get on your own. Because usage from hundreds of thousands of customers are aggregated in the cloud, providers such as Amazon Web Services can achieve higher economies of scale which translates into lower pay as you go prices.

Stop guessing capacity
Eliminate guessing on your infrastructure capacity needs. When you make a capacity decision prior to deploying an application, you often either end up sitting on expensive idle resources or dealing with limited capacity. With cloud computing, these problems go away. You can access as much or as little as you need, and scale up and down as required with only a few minutes notice.

Increase speed and agility
In a cloud computing environment, new IT resources are only ever a click away, which means you reduce the time it takes to make those resources available to your developers from weeks to just minutes. This results in a dramatic increase in agility for the organization, since the cost and time it takes to experiment and develop is significantly lower.

Stop spending money on running and maintaining data centers
Focus on projects that differentiate your business, not the infrastructure. Cloud computing lets you focus on your own customers, rather than on the heavy lifting of racking, stacking and powering servers.

Go global in minutes
Easily deploy your application in multiple regions around the world with just a few clicks. This means you can provide a lower latency and better experience for your customers simply and at minimal cost.

Types of Cloud Computing:

Infrastructure as a Service (IaaS):
Infrastructure as a Service, sometimes abbreviated as IaaS, contains the basic building blocks for cloud IT and typically provide access to networking features, computers (virtual or on dedicated hardware), and data storage space. Infrastructure as a Service provides you with the highest level of flexibility and management control over your IT resources and is most similar to existing IT resources that many IT departments and developers are familiar with today.

Platform as a Service (PaaS):
Platforms as a service remove the need for organizations to manage the underlying infrastructure (usually hardware and operating systems) and allow you to focus on the deployment and management of your applications. This helps you be more efficient as you don't need to worry about resource procurement, capacity planning, software maintenance, patching, or any of the other undifferentiated heavy lifting involved in running your application.

Software as a Service (SaaS):
Software as a Service provides you with a completed product that is run and managed by the service provider. In most cases, people referring to Software as a Service are referring to end-user applications. With a SaaS offering you do not have to think about how the service is maintained or how the underlying infrastructure is managed; you only need to think about how you will use that particular piece software. A common example of a SaaS application is web-based email where you can send and receive email without having to manage feature additions to the email product or maintaining the servers and operating systems that the email program is running on.

8.14 Introduction to E-commerce,

- E-commerce involves any business transaction executed electronically between parties. It uses Internet and Web for doing the business. It uses services like e-mail, workflow software tools, Intranet, and, the e-payment services.
- E-commerce involves buying and selling of products and services, electronically.

- The parties involved in e-commerce may be of the following kinds:
  - Companies and Companies (B2B). A data processing company handling data services for a company.
  - Companies and Consumers (B2C).
  - Consumers and Consumers (C2C). A customer selling goods to another customer, like in e-bay.com.
  - Business and the publicsector, and, consumers and the public sector.
- E-commerce web sites are like on-line market places where you can sell and buy items, and facilitate it by advertising your product, establishing newsgroups and blogs, posting job-oriented resumes etc.
- The on-line shopping is a fast growing segment as consumers are becoming more confident to use it, with the widespread use of the Internet.

8.15 E-governance,

Electronic governance or **e-governance** is the application of information and communication technology (ICT) for delivering government services, exchange of information, communication transactions, integration of various stand-alone systems and services between government-to-customer (G2C), government-to-business (G2B).

Benefits of E-governance:

1.Speed – Technology makes communication speedier. Internet, Phones, Cell Phones have reduced the time taken in normal communication.

2. Cost Reduction – Most of the Government expenditure is appropriated towards the cost of stationary. Paper-based communication needs lots of stationary, printers, computers, etc. which calls for continuous heavy expenditure. Internet and Phones makes communication cheaper saving valuable money for the Government.

3. Transparency – Use of ICT makes governing profess transparent. All the information of the Government would be made available on the internet. The citizens can see the information whenever they want to see. But this is only possible when every piece of information of the Government is uploaded on the internet and is available for the public to peruse. Current governing process leaves many ways to conceal the information from all the people. ICT helps make the information available online eliminating all the possibilities of concealing of information.

4. Accountability – Once the governing process is made transparent the Government is automatically made accountable. Accountability is answerability of the Government to the people. It is the answerability for the deeds of the Government. An accountable Government is a responsible Government.

8.16 Smart City,

A smart city is a designation given to a city that incorporates information and communication technologies (ICT) to enhance the quality and performance of urban

services such as energy, transportation and utilities in order to reduce resource consumption, wastage and overall costs. The overarching aim of a smart city is to enhance the quality of living for its citizens through smart technology.

The factors that contribute to a city being classified as smart are:

- The application of a wide variety of digital and electronic technologies to the city and its communities
- The application of ICT to uplift life and the working environments in the region
- The embedding of such ICT within government systems
- The territorialization of practices that bring the people and ICT together in order to foster innovation and enhance the knowledge that they offer

The core infrastructure elements in a smart city would include:
i. adequate water supply,
ii. assured electricity supply,
iii. sanitation, including solid waste management,
iv. efficient urban mobility and public transport,
v. affordable housing, especially for the poor,
vi. robust IT connectivity and digitalization,
vii. good governance, especially e-Governance and citizen participation,
viii. sustainable environment,
ix. safety and security of citizens, particularly women, children and the elderly, and
x. health and education.

ILLUSTRATIVE LIST

Smart Solutions

E-Governance and Citizen Services
1. Public Information, Grievance Redressal
2. Electronic Service Delivery
3. Citizen Engagement
4. Citizens - City's Eyes and Ears
5. Video Crime Monitoring

Waste Management
6. Waste to Energy & fuel
7. Waste to Compost
8. Waste Water to be Treated
9. Recycling and Reduction of C&D Waste

Water Management
10. Smart Meters & Management
11. Leakage Identification, Preventive Maint.
12. Water Quality Monitoring

Energy Management
13. Smart Meters & Management
14. Renewable Sources of Energy
15. Energy Efficient & Green Buildings

Urban Mobility
16. Smart Parking
17. Intelligent Traffic Management
18. Integrated Multi-Modal Transport

Others
19. Tele-Medicine & Tele Education
20. Incubation/Trade Facilitation Centers
21. Skill Development Centers

8.17 GIS

Geographic information systems (GIS) (also known as Geospatial information systems) are Computer Software and hardware systems that enable users to capture, store, analyze and manage spatially referenced data. GISs have transformed the way spatial (geographic) data, relationships and patterns in the world are able to be interactively queried, processed, analyzed, mapped, modeled, visualized, and displayed for an increasingly large range of users, for a multitude of purposes.

In a general sense, the term describes any information system that integrates stores, edits, analyzes, shares, and displays geographic information. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data in maps, and present the results of all these operations. Geographic information science is the science underlying geographic concepts, applications, and systems

It is a special case of information system where the database consists of observations on spatially distributed features, activities or events, which are definable in space as points, lines or area. A geographic information system manipulates data about these points, lines and areas to retrieve data for ad hoc queries and analyses.

These can be categorized into 3 main groups of GIS:
1. **Web based GIS:**
    Web-based GIS or WebGIS are online GIS applications which in most cases are excellent data visualization tools. Their functionality is limited compared to software stored on your computer, but they are user friendly and particularly useful as they not required data download.
2. **Geobrowser:**
    A Geobrowser is better explained with reference to a web-browser. In short, a geobrowser can be understood as an Internet Explorer for geographic information. Like the internet it allows the combination of many types of geographic data from many different sources. The biggest difference between the World Wide Web and the geographic web however is that everything within the latter is *spatially referenced.* Google-Earth is the most popular geobrowser available.
3. **Desktop GIS:**
    A GIS, or GIS software, allows us to interactively work with spatial data. A desktop GIS is a mapping software that needs to be installed onto and runs on a personal computer. ArcGIS is what ESRI refer to as a suite of products which can be tailored to our need. ArcGIS is used for a vast range of activities, covering both commercial and educational users.
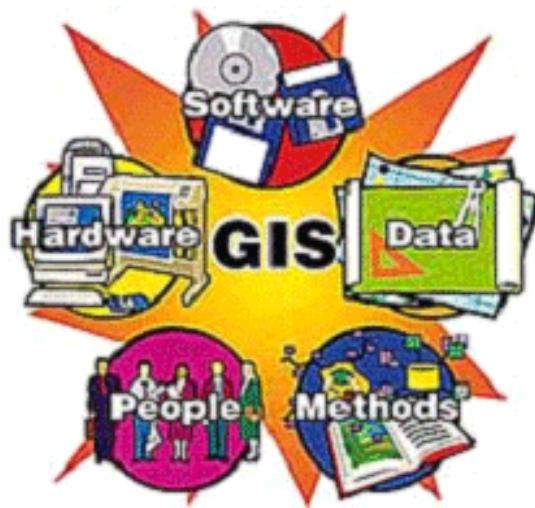
Functional Components of GIS:
An operational GIS also has a series of components that combine to make the system work. These components are critical to a successful GIS.

A working GIS integrates five key components:

1. **Hardware:** Hardware is the computer system on which a GIS operates. Today, GIS software runs on a wide range of hardware types, from centralized computer servers to desktop computers used in stand-alone or networked configurations.

2. **Software:** GIS software provides the functions and tools needed to store, analyze, and display geographic information.

3. **Data:** Perhaps the most important component of a GIS is the data. Geographic data and related tabular data can be collected in house, compiled to custom specifications and requirements, or occasionally purchased from a commercial data provider. A GIS can integrate spatial data with other existing data resources, often stored in a corporate DBMS. The integration of spatial data (often proprietary to the GIS software), and tabular data stored in a DBMS is a key functionality afforded by GIS.

4. **People:** GIS technology is of limited value without the people who manage the system and develop plans for applying it to real world problems. GIS users range from technical specialists who design and maintain the system to those who use it to help them perform their everyday work. The identification of GIS specialists versus end users is often critical to the proper implementation of GIS technology.

5. **Methods:** A successful GIS operates according to a well-designed implementation plan and business rules, which are the models and operating practices unique to each organization.

Unit 9: Fundamentals of Database

9.1 Introduction;

Databases and database systems are essential parts of our life. We have been interacting with non-computerized databases since long time, remember, looking for a word in a dictionary or finding the telephone number of your friend from a telephone directory. Computerized databases use computer to store, manipulate, and manage the database. In our daily lives we interact with the computerized databases when we go for the reservation of railway tickets and movie tickets, for the searching of a book in a library, to get the salary details, to get the balance of our account while using an ATM, to get the rate list while purchasing items from a cash and carry store and, this list can run into several pages.

9.2 Database;

Database is a repository or collection of logically related, and similar data. Database stores similar kind of data that is organized in a manner that the information can be derived from it, modified, data added, or deleted to it, and used when needed. Some examples of databases in real life situations are: dictionary—a database of words organized alphabetically along with their meaning, telephone directory—a database of telephone numbers and addresses organized by the last name of people, railway timetable— a database of trains organized by train names, and, companies listed on Stock Exchange organized by names alphabetically.

A database is defined as—(1) a collection, or repository of data, (2) having an organized structure, and (3) for a specific purpose. A database stores information, which is useful to an organization. It contains data based on the kind of application for which it is required. For example, an airline database may contain data about the airplane, the routes, airline reservation, airline schedules etc.; a college database may contain data about the students, faculty, administrative staff, courses, results etc.; a database for manufacturing application may contain data about the production, inventory, supply chain, orders, sales etc.; and a student database may contain data about students, like student names, student course etc.

9.2.1 File-Oriented Approach and Database Approach

In the early days, data was stored in files. For an application, multiple files are required to be created. Each file stores and maintains its own related data. For example, a student information system would include files like student profile, student course, student result, student fees etc. The application is built on top of the file system. However, there are many drawbacks of using the file system, as discussed below—

- Data redundancymeans storing the same data at multiple locations. In an application, a file may have fields that are common to more than one file. The

data for these common fields is thus replicated in all the files having these fields. This results in data redundancy, as more than one file has the same data values stored in them. For example,student_name and student_course may be stored in two files—"student profile" and "student fees".

- Data inconsistencymeans having different data values for the common fields in different files. During the updating process, the common fields may not get updated in all the files. This may result in different data values for the common fields in different files. For example, a student having different home address in two different files. Data redundancy provides opportunity for data inconsistency.
- The files in which the data is stored can havedifferentfile formats. This results in difficulty in accessing the data from the files since different methods are required for accessing the data from the files having different formats.
- In a file system, theconstraints ofthesystem (for example, student age >17) become part of the program code. Adding new constraints or changing an existing one becomes difficult.
- The files can be accessedconcurrently by multiple users.Uncontrolled concurrent access may lead to inconsistency and security problems. For example, two users may try to update the data in a file at the same time.

9.2.2 Characteristics of Database Approach
The main characteristics of the database approach are defined as follows:
- **Data Redundancy is Minimized:** Database system keeps data at one place in the database. The data is integrated into a single, logical structure. Different applications refer to the data from the centrally controlled location. The storage of the data, centrally, minimizes data redundancy.
- **Data Inconsistency is Reduced**: Minimizing data redundancy using database system reducesdata inconsistency too. Updating of data values becomes simple and there is no disagreement in the stored values. E.g. students' home addresses are stored at a single location and get updated centrally.
- **Data is Shared:** Data sharing means sharing the same data among more than one user. Each user has access to the same data, though they may use it for different purposes. The database is designed to support shared data. Authorized users are permitted to use the datafrom the database. Users are provided with views of the data to facilitate its use. E.g. the students' home addresses stored in the database which is shared by student profile system and library system.
- **Data Independence:** It is the separation of data description (metadata) from the application programs that use the data. In the database approach, data descriptions are stored in a central location called thedata dictionary. This property allows an organization's data to change and evolve (within limits) without changing the application programs that process the data.
- **Data Integrity is Maintained:** Stored data is changed frequently for variety of reasons such asadding new data item types, and changing the data formats. The integrity and consistency of the database are protected using constraints

on values that data items can have. Data constraint definitions are maintained in the data dictionary.

- **Data Security is Improved:** The database is a valuable resource that needs protection. The data base is kept secure by limiting access to the database by authorized personnel. Authorized users are generally restricted to the particular data they can access, and whether they can update it or not. Access is often controlled by passwords.

- **Backup and Recovery Support:** Backup and recovery are supported by the software that logs changes to the database. This support helps in recovering the current state of the database in case of system failure.

- **Standards are Enforced:** Since the data is stored centrally, it is easy to enforce standards on the database. Standards could include the naming conventions, and standard for updating, accessing and protecting data. Tools are available for developing and enforcing standards.

- **Application Development Time is Reduced:** The database approach greatly reduces the cost and time for developing new business applications. Programmer can focus on specific functions required for the new application, without having to worry about design, or low-level implementation details; as related data have alreadybeen designed and implemented. Tools for the generation of forms and reports are also available.

In addition to the advantages highlighted above, there are several other implications of using the database approach like provision of multiple-user interfaces, representation of complex relationships, concurrent data access etc.

9.3 Database System;

A bank, hospital, college, university, manufacturer, government are some examples of organizations or enterprises that are established for specific purposes. All organizations or enterprises have some basic common functions. They need to collect and store data, process data, and disseminate data for their various functions depending on the kind of organization. Some of the common functions include payroll, sales report etc.

A database system integrates the collection, storage, and dissemination of data required for the different operations of an organization, under a single administration. A database system is a computerized record keeping system. The purpose of the database system is to maintain the data and to make the information available on demand.

9.3.1 Components of Database System

A database system has four main components—(1) Users, (2) Hardware, (3) Software, and (4) Data.

- **Users**: Users are the peoplewho interact with the database system. The users of a database system are segregated into three categories based on the way

they interact with the system— (1) Application Programmers, (2) End Users, and (3) Data Administrators.

- o *Application programmers* develop application programs that manipulate the database. The developed application programs operate on the data for its insertion, deletion and retrieval. Programs are also written for interacting with the system through calls, accessing of specialized database applications, creating form requests etc.
- o *End users* are the people who interact with the database system to get information from the database to carry out their business responsibility. They interact with the system throughmenus or forms. These users can also interact with the system using query languages which requires some expertise for its use. End users include people like executives, managers, clerical staff, bank teller, and so on.
- o *Data administrator* is the manager responsible for establishingpolicies for the maintenance and handling of data once it is stored. The creation of database and the implementation of policies of the data administrator is a technical task, which is performed by a technical person called the Data BaseAdministrator (DBA). While the Data administrator's role is managerial in nature, the DataBase Administrator's role is technical in nature.

- **Software** In a database system, software lies between the stored data and the users of data. The database software can be broadly classified into three types—(1) DataBase Management System (DBMS), (2) Application software, and (3) User Interface.
  - o *DBMS* handles requests from the users to access the database for addition, deletion, or retrieval of data. Software components required to design the DBMS are also included here. Some software components are the automated tools like Computer-Aided Software Engineering (CASE) tools for the designing of databases and application programs, software utilities, and report writers.
  - o *Application software* uses DBMS facilities to manipulate the database to achieve a specific business function, such as, providing reports that can be used by users. Application software is generally written in a programming language like C, or it may be written in a language supported by the DBMS.
  - o *User interface* of the database system constitutes the menus and the screens that the user uses to interact with the application programs and the DBMS.
- **Hardware** Hardware is the physical device on which the database system resides. The hardware required for the database system is the computer or a group of connected computers. Hardware also includes devices like magnetic disk drives, I/O device controllers, printers, tape drives, connecting cables and other auxiliary hardware.
- **Data** Data is raw numbers, characters, or facts represented by values. The data can be numeric data, non-numeric data, images or pictures. Some

examples of data are 23, 45, "Divya", "Computer", "India". The data in a database system is integrated and shared.
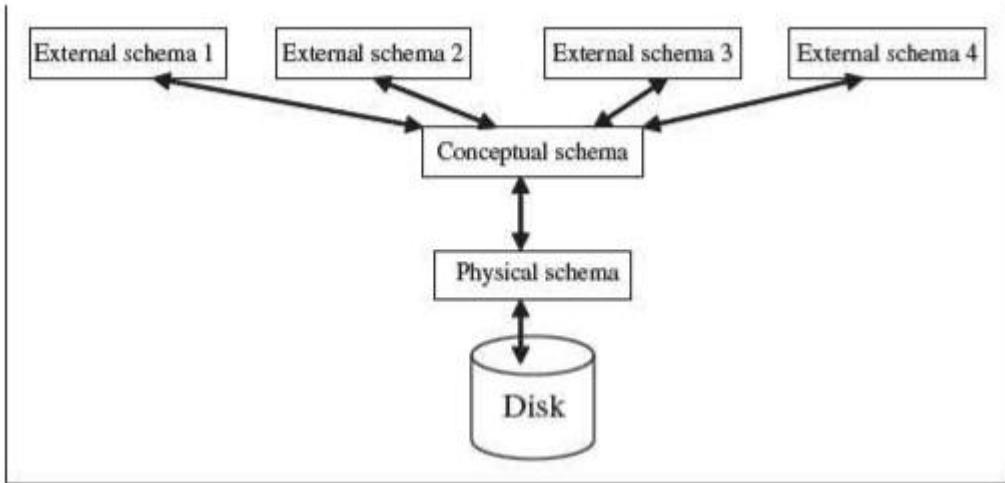
- o *Integrated data* implies that several distinct data files are unified in a manner that redundancy among these data files is wholly or partially eliminated. For example, a database may contain files for STUDENT_PROFILE having student roll no., student name, address,course enrolled etc., and MARKS file having the marks of students in various subjects. To prepare the report card, there is a need to get the complete student profile along with the marks obtained in each subject. For this, the complete profile of the student need not be stored in MARKS file as the information can be retrieved by referring to STUDENT_PROFILE file.
- o *Shared data* implies that the data stored in the database can be shared among multiple users, such that each user has access to the same data, which the user can use for different purposes. For example, data forcourse enrolledin STUDENT_PROFILE file is shared by the NCC department and the Library department—both use the same data for different purposes.

9.3.2 Architecture of Database System
The architecture of a database system provides a general framework for database systems. Different database systems, small or big, may not support all aspects of the architecture; however, different systems can be matched to the framework. The architecture for database system is proposed by ANSI/SPARC study group and is called ANSI/SPARC architecture. The purpose of the architecture is to make databases more independent of the application that is using the database. Schema contains information about the description of the database.

The ANSI/SPARC architecture is divided into three levels (Figure), as follows:

- **Internal (Physical) Level** has an internal schema. The internal schema describes the physical storage structure of the database. It is concerned about how data is stored physically. It does not deal with the physical storage in terms of blocks, neither pages nor the device. Rather, it describes the organization of files, the access path to the database etc. The physical data model is used to describe the physical schema.
- **Conceptual Level** has a conceptual schema. The conceptual schema describes the structure of wholedatabase for the users. It describes the entities in the database, their relationships, and constraints. The representation or implementation data model is used to describe the conceptual schema. Conceptual schema also uses the conceptual data model.
- **External Level or View Level** provides a user's and application's view of the data. It includes one or more external schema. For a particular user, the external schema describes the structure of the database relevant to it, and hides rest of the information. It uses the data model at the conceptual level.

9.4 Database Management System;

The interrelated set of data that forms the database needs to be stored and managed, so that the database can be accessed for the retrieval of data, and for insertion, deletion, or updating of data. DBMS is a software system for creating, organizing and managing the database. DBMS handles all access to the database and manages the database. Managing the database implies that it provides a convenient environment to the user to perform operations on the database for creation, insertion, deletion, updating, and retrieval of data. DBMS defines the scope of the use of database. This keeps data secure from unauthorized access. The functionality of DBMS includes—(1) the database that contains interrelated data, and (2) a set of programs to access the data. The DBMS implements the three schema architecture—internal schema at internal level, conceptual schema at conceptual level, and external schema at external level. The different DBMSs may not provide a clean distinction between the three levels, but generally they follow the three schema architecture.

Mappingis a term used for transforming requests and the results between different levels of the DBMS architecture. In a DBMS, the user interacts with the DBMS through the external view, and the data in the DBMS is stored at the physical level. Any request from the user, would require that the DBMS access the data and then present it to the user as specified in the external schema. This requires transforming the request from one level to the other, or mapping.
Some of the common commercial DBMSs are—Oracle Database, IBMs DB2, Microsoft's SQL Server and Microsoft Access. MySQL is a popular open source DBMS.

9.5 Database System Architectures;

The computers has evolved from big mainframe computers to small desktop personal computers. The advances in computer and its architecture have also

resulted in the advances in databases and its architecture. Here, we discuss the architecture of databases of three kinds—Centralized DBMS Architecture, Client-Server Architecture, and Distributed Databases.

1. Centralized DBMS Architecture

- Centralized databases are the traditional database systems where all database functionality, data, application program and user interface processing are located on one machine. Access to the database from remote locations is via the communication links
- In the early systems, the mainframe computers provided all the functionality required by the users of computer. The users interacted with the mainframe computers via dumb terminals. All processing was performed on the mainframe computer. The DBMS was centralized, and stored on the mainframes. The dumb terminals had only the display facility.
- The centralized database is easy to manage and administer.
- Common examples of centralized databases are personal database, and central computer database.
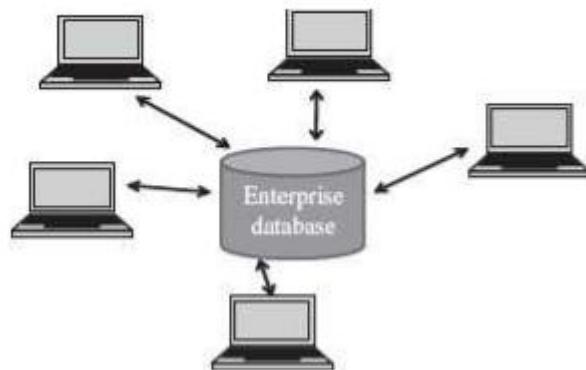
Fig: Centralized database

2. Client-Server Architecture

- The growth of Personal Computer (PC) has resulted in PCs replacing the dumb terminals. PCs are powerful and fast machines that can handle the user interface functionality.
- Client-Server systems involve aclientmachine and aservermachine. Clients are PCs or workstations that have user interface capability and functionality for local processing. Servers are powerful computers that can manage the files (file servers), printer (printer servers), or, emails (e-mail servers). The client is connected to the server via a communication link. The client interacts with the server when it requires access to any additional functionality that does not exist in its own machine. Client provides interfaces to access and utilize the server resources.
- The Client-Server architecture has three components—the user interface programs, the application programs that contain the application logic,and the DBMS that stores the data. The request made by the user interface program

is processed using the application logic which then accesses the database to retrieve the data.

- The DBMS architecture on the Client-Server systems are of two kinds—two-tier Client-Server architecture and three-tier Client-Server architecture.
- In two-tier Client-Server architecture, the user interface programs and the application programs run on the client side. An Application Program Interface (API) allows client side programs, to call the DBMS which is at the server side. The client programs use Open Data Base Connectivity (ODBC) or Java Data Base Connectivity (JDBC) interfaces to communicate with the database.
- The three-tier Client-Server architecture is commonly used for web applications. In addition to the client and the database servers, it has an intermediate layer or middleware called Application Server or Web Server. The web server stores the application or business logic part of the application. The client stores the user interface. The DBMS is stored at the server side. The web server interacts with client at one end, and with the server for the DBMS at the other end. The web server acts like a pipe for receiving the client request, processing it and accessing the datafrom the DBMS server, and sending it back to the client.
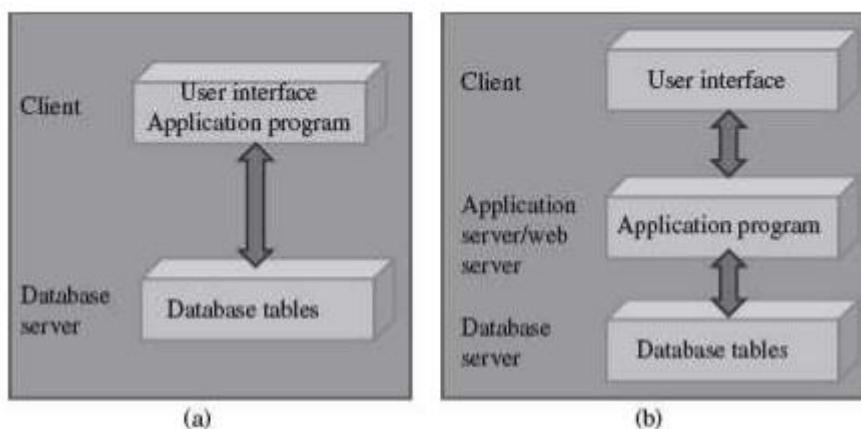
Fig: (a) 2 – Tier (b) 3 – Tier Architecture

3. Distributed Databases

- Adistributed databaseis one with decentralized functionality of the DBMS. It is distributed among a number of computers
- The computers that store the components of database are physically placed at different geographical locations; however, all the components are logically related.
- Access to the distributed databases is from remote locations via the communication links.
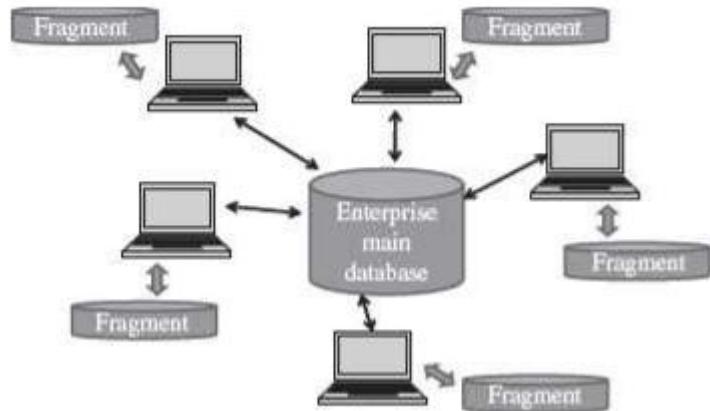
Fig: Distributed Database

9.6 Database Applications;

Databases range from those designed for a single user with a desktop computer to those on mainframe computers with thousands of users. The database applications can be for different purposes like—
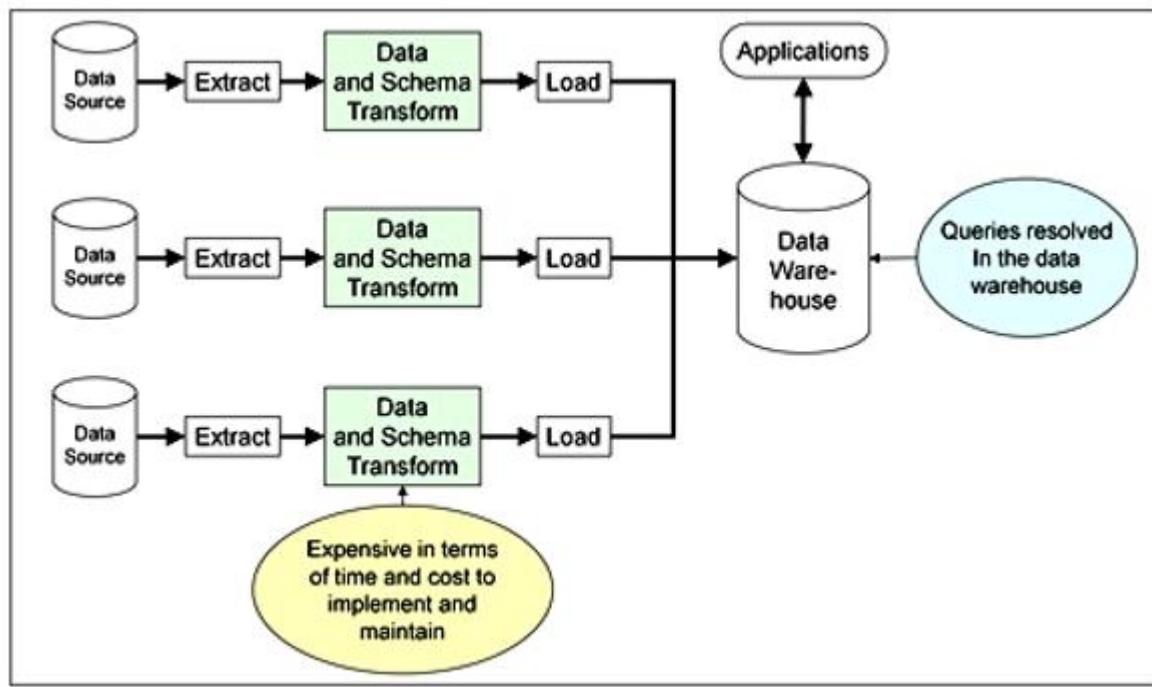
a. personal databases that support one user with a stand-alone personal computer,

b. workgroup databases for a small team of people (less than 25) who work in collaboration on a project,

c. departmental databases designed to support the various functions and activities of a department (a functional unit of an organization), and

d. enterprise databases to support organization-wide operations and decision making. Data warehouse is an enterprise database.

The DBMS commonly provides tools to the application programmer for application development. Some of the tools provided by DBMS include tools for screen, menu, and report generation, application generators, compilers, and data and view definition facilities. Modern database systems provide language components that are much more powerful than those of traditional languages. For example, Developer 2000/PowerBuilder for Oracle, and Visual Basic for Microsoft SQL server.

9.7 Introduction to Data Warehousing,

A data warehouse is a central repository for all or significant parts of the data that an enterprise's various business systems collect. Typically, a data warehouse is housed on an enterprise mainframe server. Data from various online transaction processing (OLTP) applications and other sources is selectively extracted and organized on the data warehouse database for use by analytical applications and user queries. Data warehousing emphasizes the capture of data from diverse sources for useful analysis and access, but does not generally start from the point-of-view of the end user or knowledge worker who may need access to specialized, sometimes local databases.

Advantages

Data warehouses tend to have a high query success, as they have complete control over the four main areas of data management systems:

- Clean data
- Indexes: multiple types
- Query processing: multiple options
- Security: data and access

Other benefits may include:
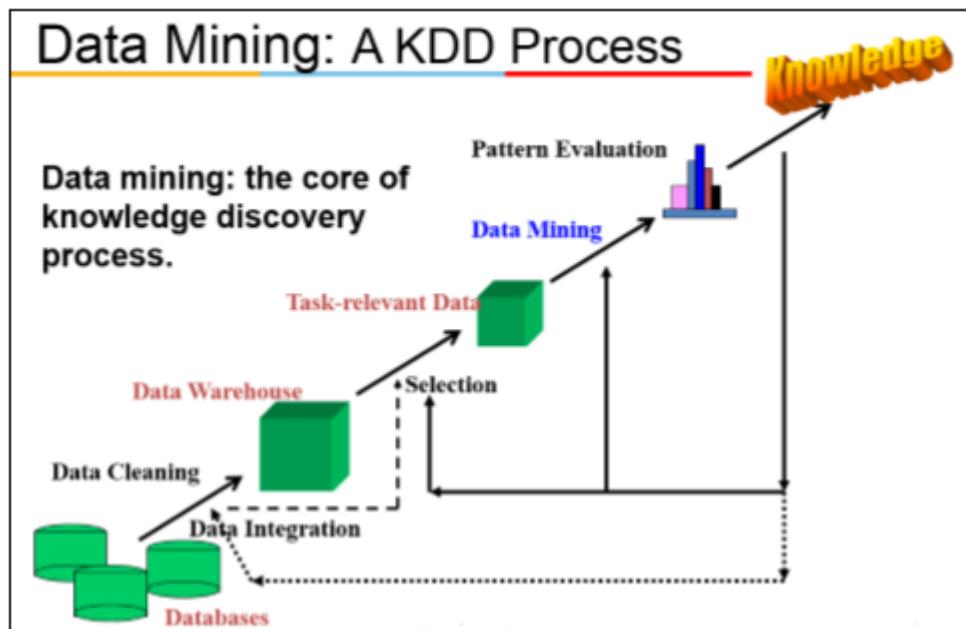
- Integrating data from multiple sources;
- Performing new types of analyses; and
- Reducing cost to access historical data.
- Standardizing data across the organization, a "single version of the truth";
- Improving turnaround time for analysis and reporting;
- Sharing data and allowing others to easily access data;
- Supporting ad hoc reporting and inquiry;
- Reducing the development burden on IS/IT; and
- Removing informational processing load from transaction-oriented databases;

9.8 Data mining,

Data mining (the analysis step of the "Knowledge Discovery in Databases" process, or KDD), an interdisciplinary subfield of computer science, is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data

preprocessing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

9.9 BigData

Big Data is a phrase used to mean a massive volume of both structured and unstructured data that is so large it is difficult to process using traditional database and software techniques. In most enterprise scenarios the volume of data is too big or it moves too fast or it exceeds current processing capacity.

Big Data has the potential to help companies improve operations and make faster, more intelligent decisions. The data is collected from a number of sources including emails, mobile devices, applications, databases, servers and other means. This data, when captured, formatted, manipulated, stored and then analyzed, can help a company to gain useful insight to increase revenues, get or retain customers and improve operations.

An example of big data might be petabytes (1,024 terabytes) or exabytes (1,024 petabytes) of data consisting of billions to trillions of records of millions of people—all from different sources (e.g. Web, sales, customer contact center, social media, mobile data and so on). The data is typically loosely structured data that is often incomplete and inaccessible.

Importance of Big Data:

The importance of big data doesn't revolve around how much data you have, but what you do with it. You can take data from any source and analyze it to find

answers that enable 1) cost reductions, 2) time reductions, 3) new product development and optimized offerings, and 4) smart decision making. When you combine big data with high-powered analytics, you can accomplish business-related tasks such as:

- Determining root causes of failures, issues and defects in near-real time.

- Generating coupons at the point of sale based on the customer's buying habits.

- Recalculating entire risk portfolios in minutes.

- Detecting fraudulent behavior before it affects your organization.

Uses of Big Data:
Banking

With large amounts of information streaming in from countless sources, banks are faced with finding new and innovative ways to manage big data. While it's important to understand customers and boost their satisfaction, it's equally important to minimize risk and fraud while maintaining regulatory compliance. Big data brings big insights, but it also requires financial institutions to stay one step ahead of the game with advanced analytics.

Education

Educators armed with data-driven insight can make a significant impact on school systems, students and curriculums. By analyzing big data, they can identify at-risk students, make sure students are making adequate progress, and can implement a better system for evaluation and support of teachers and principals.

Government

When government agencies are able to harness and apply analytics to their big data, they gain significant ground when it comes to managing utilities, running agencies, dealing with traffic congestion or preventing crime. But while there are many advantages to big data, governments must also address issues of transparency and privacy.

Health Care

Patient records. Treatment plans. Prescription information. When it comes to health care, everything needs to be done quickly, accurately – and, in some cases, with enough transparency to satisfy stringent industry regulations. When big data is managed effectively, health care providers can uncover hidden insights that improve patient care.

Manufacturing

Armed with insight that big data can provide, manufacturers can boost quality and output while minimizing waste – processes that are key in today's highly competitive market. More and more manufacturers are working in an analytics-based culture, which means they can solve problems faster and make more agile business decisions.

Retail

Customer relationship building is critical to the retail industry – and the best way to manage that is to manage big data. Retailers need to know the best way to market to customers, the most effective way to handle transactions, and the most strategic way to bring back lapsed business. Big data remains at the heart of all those things.

Unit 10: Multimedia

10.1 Introduction;

Newspaper and television are the common medium of mass communication. However, they differ in the way they present information to the user. The information in a newspaper is presented as a combination of text, image, and graphics. This has a different impact on the user than the information presented on the television as a combination of image, photo, video, sound and music. Similarly, talking over a telephone (using sound) has a different impact on the user, than, talking using the Internet telephone with a web camera (sound and photo) attached to it. In other words, the same information when presented using different media has a different impact on the user. Or, we can say that the media used for presenting the information affects the way the user perceives and understands the information. Multimedia is a combination of graphics, audio, text, animation, and video using the computer.

10.2 Multimedia - Definition;

Newspaper, radio, television and films are some of the earliest and traditional means of mass communication that use mass media. Newspapers use text, image and graphics, and, television uses text, image, graphics, and sound. In these traditional means of communication, the communication is one-way—from the media to the mass users. The user simply reads the newspaper, listens to the radio, and watches the television and films, and, cannot in any way manipulate the media. Mass media is used by the user in a sequence, linearly. For example, a text book consists of a sequence of combination of text, graphics, and images. The text book is meant to be read linearly from the start to the end. With mass media, the user is in a passive state (receiving whatever is communicated).

The word multimedia consists of two words—multi and media. The word multi means many and the word media (plural of medium) are the means through which information is shared. There are different mediums of sharing information like sound, text, image, graphics, animation or video. Multimedia represents information through a variety of media. Multimedia is a combination of more than one media—text, graphics, images, audio, or video, which is used for presenting, sharing, and disseminating the information. Multimedia is delivered through the computer and microprocessor-based devices, thereby introducing the elements of interactivity, which differentiates it from the traditional forms of media (also called mass media). Multimedia or Interactive multimedia allows the user and the multimedia application to respond to each other. The user is able to control the elements of the multimedia application in terms of what elements will be delivered and when. Since multimedia systems are integrated with computers, they are also referred to as the digital multimedia system.

10.3 Characteristics of Multimedia;

A multimedia system has four basic characteristics:

- *Computer* is an intrinsic part of the multimedia system. As a result, multimedia has become interactive. In multimedia, computer allows the user to interact with the media and thus manipulate it by controlling what is to be communicated and when. Multimedia has resulted in the creation of many new possibilities—
  (1) thecomputational power of computer is utilized for multimedia applications,
  (2) the telecommunication network (Internet, WWW) along with the computer enables transmission and distribution of information, and,
  (3) the use of computer facilitates design and creation of a variety of new applications.

- The different elements of multimedia are combined andintegratedinto a single multimedia system. Special software is required for the integration of different media element files.

- The useof computer in multimedia requires all elements of multimedia to be indigitalformat. In a digital multimedia system, the media streams are digital and are produced, processed, stored, represented and transmitted using computers. The digital nature of multimedia requires special treatment of the multimedia elements. The hardware and software are needed to convert multimedia elementsfrom analog to digital format and vice versa. There is a need to decide about theresolution versus qualityof output required, during storing of data in the computer. Storing multimedia files on computer hard disk takes large amount of disk space, so compression technologiesandfile formatsfor storing the different media elements is required. Moreover, special programs are required toplaythe compressed files. Similarly, special software is required toeditthe different media element files, and to combine and integrate the different elements of the multimedia into a single multimedia system.

- Multimedia system is interactive.The user is active and can manipulate whatever is being communicated. Multimedia allows two-way communication. The user can use devices like keyboard, trackball or joystick to interact with the multimedia system. Interactive multimedia is non-linear. Theuser is able to follow the links and jump from one part of the document to the other.Hypermediaenables a user to gain or provide access to text, audio and video, and computer graphics using links in a non-linear way, using computers. World Wide Web (WWW)is an example of hypermedia application. The user is able to respond and control what to see or hear and when to do it.

10.4 Elements of Multimedia;

A multimedia system consists of several elements like—text, graphics, sound, video, and animation. The data streams of these different elements of the multimedia system are of two kinds—time-dependent, and time-independent. Media like text, graphics, and image are timeindependent. The information is not dependent on the

timing of the data stream when using these media. However, media like audio, video, and animation are time-dependent. In time-dependent media, the data streams must appear at regular time intervals for the purposes of continuity.

The different media elements are described briefly in the following subsections.

1 Text

- Text is an important element of multimedia. The purpose of using text is to write titles, to define menus, to navigate, and to write the content.
- The text in multimedia is different from the traditional text written using paper and pen. Multimedia text is combined with other elements like graphics and pictures to deliver a powerful effect. Multimedia text can be written in a way that the user needs to jump back and forth, written in multiple channels like pictures, sound, animation, and colored text.
- The text must be short and relevant wherever used. Text is often mixed with art resulting in a much greater impact on the user. Text can also be animated. Text has evolved from being displayed in one size and one color under MS-DOS, to being specified using fonts and color on a color monitor having Windows, to the vector-based text by Adobe which uses graphics-based fonts and creates images.
- Text Font—The text can be written in different fonts: A font is composed of three things—typeface, style and size. Arial, Courier, Times are typefaces. Bold and italics are styles. Size is the lengthof the character (from top to bottom). Arial 11-point italic is a font.
- Text effects allow special effects to be added to text fonts by adding depth and visual impact (2D and 3D effects). Text effects can be created using MS-Word's WordArt. In WordArt, text is treated as a drawing object, i.e. text can be manipulated like an object. WordArt provides special effects to the text like, depth, direction, shape, color, and texture.
- Text animationcan be used to make the text move, change or flash. MS-WORD and MS-PowerPoint can be used for text animation.

2. Graphics

Communication via pictures is easier to understand. Graphics is the most important element of multimedia. Multimedia presentations are predominantly graphics-based. Graphics elements in a multimedia system are images that could be still pictures (like photographs) converted to digital format with the help of scanners, or pictures generated on the computer. They may be 2-dimensional such as photographs, or, 3-dimensional such as objects around us. They may be either static graphic elements or animated. In computer graphics, an image is always a digital image.

3. Audio

Sound consists of all possible sounds which may or may not be audible to humans. Audio consists of the sounds humans can hear. For example, the sound emitted by the dog-whistle is heard by dogs but not by humans. The presence of sound enhances the effect of a graphic presentation, video or animation. In a multimedia

project, sound can be used in two ways. It can be used to provide audio content in a multimedia system such as, narration for a clip playing on the screen; audio sound tracks in movies; short instructions; or, music to communicate as in a song. Sound can also be used in the background and for sound effects.

4. Video

Video is another element of multimedia. Video and audio are closely related, and together they are the most effective means of communication that can be a part of the multimedia system. Digital video is used in making of movies, gaming, and IT industry. The Digital Video Disk (DVD) makes it possible to distribute large videos in a compact form.

5. Animation

Animation is creating of an illusion of movement from a series of still drawings. To create a feeling of motion of an image (still drawing), the image is projected on the screen as frames. Generally, 30 frames per second are used to make the object appear to be in smooth motion on the screen.

10.5 Multimedia Applications

Multimedia applications have found their way into different arenas of our life. We are interacting with multimedia applications in the area of education and entertainment that includes sports, laser shows, video games, or animation movies. Multimedia applications have found their way in business—may it be for advertising, marketing, video meetings, result presentations or customer feedbacks. Nowadays, training is also imparted using multimedia applications like simulations and 3D designs. Entertainment parks like Disneyland use virtual reality and multimedia, innovatively design, create, and improve; their games and rides. The following subsections discuss the impact of multimedia applications in different areas of our life.

1. In Education

Using multimedia for education encompasses the use of video clips, speech and sound, music, animations, and graphics. The use of these multimedia elements impacts the whole learning process and pedagogy. It is always better to visually observe and hear about a topic than only reading it from a book.

2. In Entertainment

Video games and animation movies are some of the common uses of multimedia for entertainment. Gaming is one of the oldest uses of multimedia. It uses the different elements of multimedia to make it more realistic and interactive. The users (especially children) are captivated by the beauty of the game and heavily engaged in using them. The playing of games has its own advantages as it results in the sharpening of mental faculties (requires quick response from the user), allows mixing pleasure with work, and results in enjoyment and relaxation. Nowadays, the games can also be played across the Internet (like chess requiring two players), where the users are located at different physical locations and can interact with each other over a game.

3. In Training

Training is an important ingredient in any organization. Imparting training to employees requires significant amount of effort and resources. There are only a few people in an organization who have training skills and knowledge, to train others. Earlier, training the employees in an organization would require the training team to travel to the location of training, incur expenses (lodging, boarding, traveling etc.), and train the people and move on to next location.

4. In Business

Multimedia is crucial for the success of business organizations. In government and nongovernment organizations, no meeting is complete without PowerPoint presentations and reports generated along with graphs and images. Moreover, businesses use multimedia to further their products and to enlighten people about their products. A CD giving details of the product is used. Car manufacturers, white-good manufacturers like television, microwave, refrigerators and washing machines, cell phone manufacturers etc., rely heavily on multimedia to display their product, market them, and to enhance sales of their product. Multimedia is also used by business organizations to train their people, for marketing and advertising, for creating visual and sound effects that are appealing to people, and thus making their product more saleable. Since the information inclusive of all sound and visual effects can be stored on a CD and distributed, it is considered as an inexpensive means of advertising and business promotion.

5. Virtual Reality

Virtual Reality is created using multimedia. Virtual reality is a special environment that is created using multimedia, where the users feel as if they are in a three-dimensional world. It gives the feeling to the users as if they are participating in the scenario. Virtual reality uses a head-mounted display which consists of two display screens, and, data gloves that are used to record the hand movements. The head-mounted display and the gloves are connected to a computer which has the software for virtual reality installed into it. When the user enters into virtual reality, the position of the head of the user, eyes, and hand movement is used to decide the response; the next view is presented accordingly by the virtual reality software.

Unit 11: Computer Security

11.1 Introduction;

We all like to be secure in our home, office, locality, city, country, and in this world. We use different mechanisms to ensure our security. Inside our homes, we keep our valuables safely locked in a cupboard that is accessible by the elders of the house; we keep the gates of our house bolted and even have an intrusion-detection system installed. We have high walls and gates surrounding our locality and also a watchman who guards the open gates. We have police for our security within a city and armed forces for the country. We take all these measures to make ourselves and our valuables, resources, possessions secure.

The widespread use of computers has resulted in the emergence of a new area for security— security of computer. Computer security is needed to protect the computing system and to protect the data that they store and access. Transmission of data using network (Internet) and communication links has necessitated the need to protect the data during transmission over the network. Here, we use the term computer security to refer to both the computer security and the network security.
Computer security focuses on the security attacks, security mechanisms and security services.

- Security attacksare the reasonsfor breach of security. Security attacks comprise of all actions that breaches the computer security.
- Security mechanismsare the tools that include the algorithms, protocols or devices, that are designed to detect, prevent, or recover from a security attack.
- Security servicesare the services that are provided by a system for a specific kind of protection to the system resources.

The purpose of computer security is to provide reliable security services in the environments suffering security attacks, by using security mechanisms. The security services use one or more security mechanism(s).
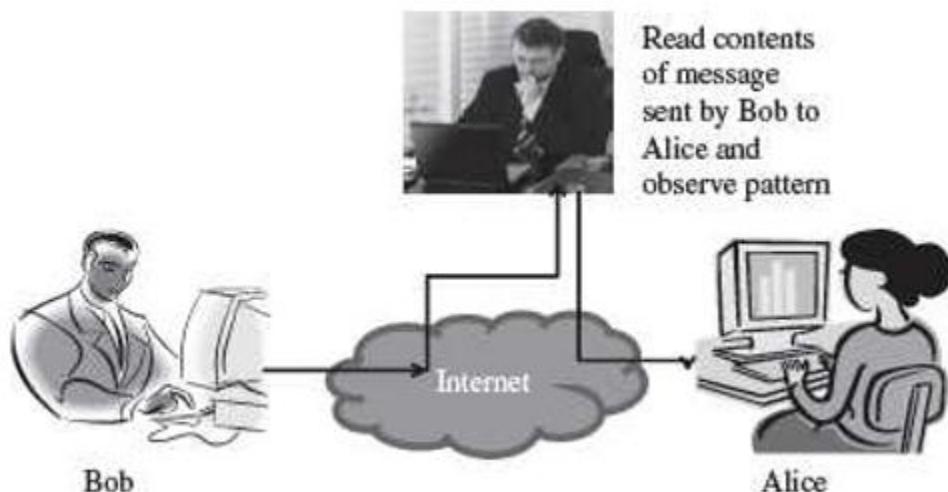
11.2 Security Threat and Security Attack;

A threat is a potential violation of security and causes harm. A threat can be a malicious program, a natural disaster or a thief. Vulnerability is a weakness of system that is left unprotected. Systems that are vulnerable are exposed to threats. Threat is a possible danger that might exploit vulnerability; the actions that cause it to occur are the security attacks. For example, if we leave the house lock open—it is vulnerable to theft; an intruder in our locality (might exploit the open lock)is a security threat; the intruder comes to know of the open lock and gets inside the house—This is a security attack.
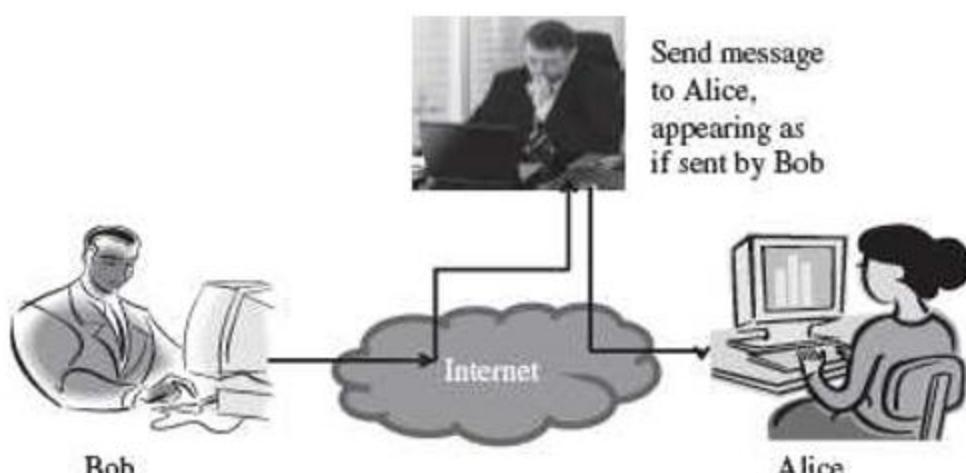
A security attack may be a passive attack or an active attack.

- The aim of apassive attackis to get information from the system but it does not affect the system resources. Passive attacks are similar to eavesdropping (Figure). Passive attacks may analyze the traffic tofind the nature of communication that is taking place, or, release the contents of the message to a person other than the intended receiver of the message. Passive attacks are difficult to detect because they do not involve any alteration of the data. Thus, the emphasis in dealing with passive attacks is on prevention rather than detection.

- Anactive attacktries to alter the system resources or affect its operations. Active attack may modify the data or create a false data (Figure 14.2). An active attack may be a masquerade (an entity pretends to be someone else), replay (capture events and replay them), modification of messages, and denial of service. Active attacks are difficult to prevent. However, an attempt is made to detect an active attack and recover from them.

- Security attacks can be on users, computer hardware and computer software (Figure).

- Attacks on userscould be to the identity of user and to the privacy of user. Identity attacks result in someone else acting on your behalf by using personal information like password, PIN number in an ATM, credit card number, social security number etc. Attacks on the privacy of user involve tracking of users habits and actions—the website user visits, the buying habit of the user etc. Cookies and spam mails are used for attacking the privacy of users.
- Attacks on computer hardwarecould be due to a natural calamity like floods or earthquakes; due to power related problems like power fluctuations etc.; or by destructive actions of a burglar.
- Software attacksharm the data stored in the computer. Software attacks may be due to malicious software, or, due to hacking.Malicious software or malwareis a software code includedinto the system with a purpose to harm the system. Hacking is intruding into another computer or network to perform an illegal act.

11.3 Malicious Software;

Malicious users use different methods to break into the systems. The software that is intentionally included into a system with the intention to harm the system is called malicious software. Viruses, Trojan horse, and Worms are examples of malicious programs. Javascripts and Java applets written with the purpose of attacking, are also malicious programs.

1. Virus
Virus is a software program that is destructive in nature. Virus programs have the following properties:
- It can attach itself to otherhealthy programs.
- It can replicate itself and thus can spread across a network.
- It is difficult to trace a virus after it has spread across a network.
- Viruses harm the computer in many ways—

- o corrupt or delete data or files on the computer,
- o change the functionality of software applications,
- o use e-mail program to spread itself to other computers,
- o erase everything on the hard disk, or,
- o degrade performance of the system by utilizing resources such as memory or disk space.
- Virus infects an executable file or program. The virus executes when a program infected with virus is executed or you start a computer from a disk that has infected system files.
- Once a virus is active, it loads into the computer's memory and may save itself to the hard drive or copies itselfto applications or system files on the disk.
- However, viruses cannot infect write protected disks or infect written documents. Viruses do not infect an already compressed file. Viruses also do not infect computer hardware; they only infect software.
- Viruses are most easily spread by attachments in e-mail messages. Viruses also spread through download on the Internet.

2. Worms
Worm is self-replicating software that uses network and security holes to replicate itself. A copy of the worm scans the network for another machine that has a specific security hole. It copies itself to the new machine using the security hole, and then starts replicating from there, as well. A worm is however different from a virus. A worm does not modify a program like a virus, however, it replicates so much that it consumes the resources of the computer and makes it slow. Some examples of worms are—"Code Red" and "Nimda".
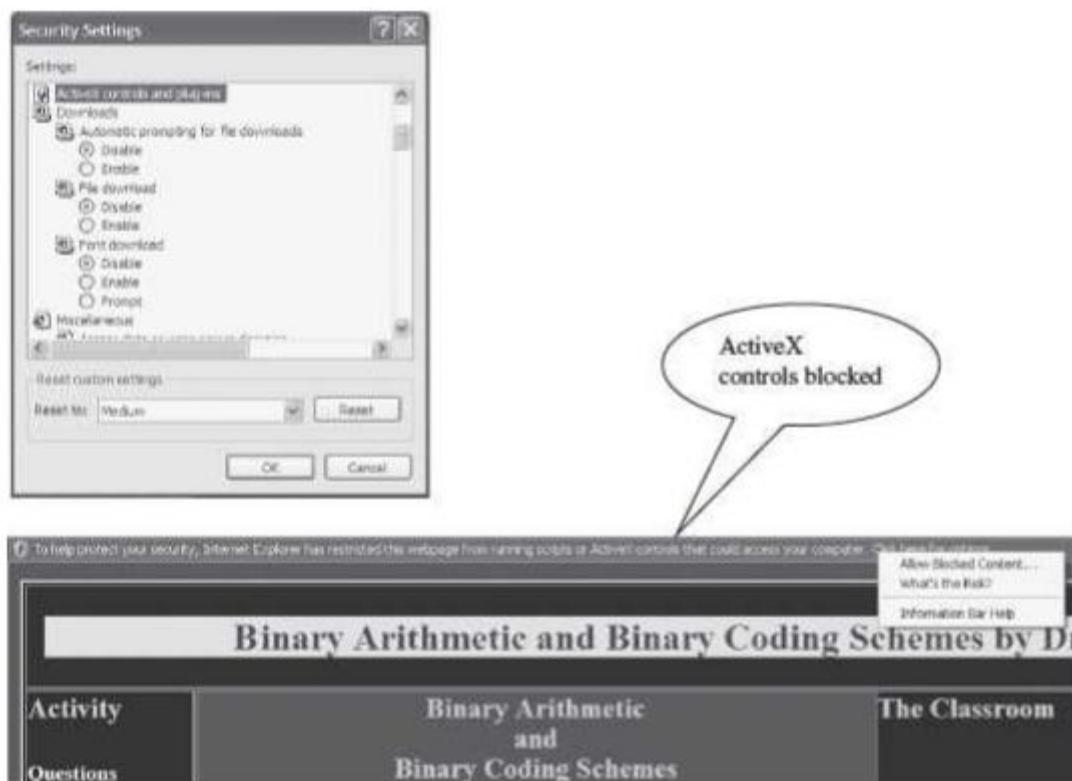
3. Trojan Horse
Trojan horse is destructive programs that masquerade as useful programs. The name "Trojan horse" is given because of the Greek soldiers who reached the city of Troy by hiding themselves inside a large wooden horse (Figure). The people of the city of Troy themselves pulled the horse inside their city, unaware of the fact that the Greek soldiers were hiding inside the horse. Similarly, users install Trojan horses thinking that it will serve a useful purpose such as a game or provide entertainment. However, Trojan horses contain programs that corrupt the data or damage the files. Trojan horses can corrupt software applications. They can also damage files and can contain viruses that destroy and corrupt data and programs. Trojan horse does not replicate themselves like viruses.

4. Javascripts, Java Applets and ActiveX Controls

Applets (Java programs), and ActiveX controls are used with Microsoft technology, which can be inserted in a Web page and are downloaded on the client browser for execution. Applets and ActiveX controls are generally used to provide added functionality such as sound and animation. However, these programs when designed with a malicious intention can be disastrous for the client machine. Java Applets have strong security checks that define what an applet can do and what it cannot. ActiveX controls do not have such security checks. Normally, ActiveX controls must be kept disabled while working on the Internet (Figure)

5. HACKING

Hacking is the act of intruding into someone else's computer or network. A hacker is someone who does hacking. Hacking may result in a Denial of Service (DoS) attack.

The DoS attack prevents authorized users from accessing the resources of the computer. It aims at making the computer resource unusable or unavailable to its intended users. It targets the computer and its network connections, to prevent the user from accessing email, web sites, online accounts (banking, etc.), or other services that rely on the affected computer. In a DoS attack, the services of the entire network, an Internet site or service, may be suppressed or disabled. The affected machine is flooded with spurious requests and messages so as to overload the network. As a result, the affected machine cannot process the valid requests. This is a denial of service to the valid users. Generally, the targets of such attacks are the sites hosted on high-profile web servers such as banks and credit card payment gateways.

Packet sniffing, E-mail hacking and Password cracking are used to get the username and password of the system to gain unauthorized access to the system. These methods gather the information when the data is being transmitted over the network.

6. Packet Sniffing

The data and the address information are sent as packets over the Internet. The packets may contain data like a user name and password, e-mail messages, files etc. Packet sniffing programs are used to intercept the packets while they are being transmitted from source to destination. Once intercepted, the data in the packets is captured and recorded. Generally, packet sniffers are interested in packets carrying the username and password. Packet sniffing attacks normally go undetected. Ethereal and Zx Sniffer are some freeware packet sniffers. Telnet, FTP, SMTP are some services that are commonly sniffed.

7. Password Cracking

Cracking of password is used by hackers to gain access to systems. The password is generally stored in the system in an encrypted form. Utilities like Password cracker is used to crack the encrypted passwords. Password cracker is an application that tries to obtain a password by repeatedly generating and comparing encrypted passwords or by authenticating multiple times to an authentication source.

8. E-mail Hacking

The e-mail transmitted over the network contains the e-mail header and the content. If this header and the content are sent without encryption, the hackers may read or alter the messages in transit. Hackers may also change the header to modify the sender's name or redirect the messages to some other user. Hackers use packet replay to retransmit message packets over a network. Packet replay may cause serious security threats to programs that require authentication sequences. A hacker may replay the packets containing authentication data to gain access to the resources of a computer.

11.4 Security Services;

The security services provide specific kind of protection to system resources. Security services ensure Confidentiality, Integrity, Authentication, and Non-Repudiation of data or message stored on the computer, or when transmitted over the network. Additionally, it provides assurance for access control and availability of resources to its authorized users.

Confidentiality—The confidentiality aspect specifies availability of information to only authorized users. In other words, it is the protection of data from unauthorized disclosure. It requires ensuring the privacy of data stored on a server or transmitted via a network, from being intercepted or stolen by unauthorized users. Data encryption stores or transmits data, in a form that unauthorized users cannot understand. Data encryption is used for ensuring confidentiality.

Integrity—It assures that the received data is exactly as sent by the sender, i.e. the data has not been modified, duplicated, reordered, inserted or deleted before reaching the intended recipient. The data received is the one actually sent and is not modified in transit.

Authentication—Authentication is the process of ensuring and confirming the identity of the user before revealing any information to the user. Authentication providesconfidence in the identity of the user or the entity connected. It also assures that the source of the received data is as claimed. Authentication is facilitated by the use of username and password, smart cards, biometric methods like retina scanning and fingerprints.

Non-Repudiation prevents either sender or receiver from denying a transmitted message. For a message that is transmitted, proofs are available that the message was sent by the alleged sender and the message was received by the intended recipient. For example, if a sender places an order for a certain product to be purchased in a particular quantity, the receiver knows that it came from a specified sender. Non-repudiation deals with signatures.

Access Control—It is the prevention of unauthorizeduse of a resource. This specifies the users who can have access to the resource, and what are the users permitted to do once access is allowed.

Availability—It assures that the data and resources requested by authorized users are available to them when requested.

11.5 Security Mechanisms

Security mechanisms deal with prevention, detection, and recovery from a security attack. Prevention involves mechanisms to prevent the computer from being damaged. Detection requires mechanisms that allow detection of when, how, and by whom an attacked occurred. Recovery involves mechanism to stop the attack, assess the damage done, and then repair the damage.

Security mechanisms are built using personnel and technology.

- Personnel are used to frame security policy andprocedures, and for training and awareness.

- Security mechanisms use technologies like cryptography, digital signature, firewall, user identification and authentication, and other measures like intrusion detection, virus protection, and, data and information backup, as countermeasures for security attack.

11.5.1 Cryptography,
Cryptography is the science of writing information in a "hidden" or "secret" form and is an ancient art. Cryptography is necessary when communicating data over any network, particularly the Internet. It protects the data in transit and also the data stored on the disk. Some terms commonly used in cryptography are:
- Plaintext is the original message that is an input, i.e. unencrypted data.
- Cipher and Code—Cipher is a bit-by-bit or character-by-character transformation without regard to the meaning of the message. Code replaces one word with another word or symbol. Codes are not used any more.
- Cipher text—It is the coded message or the encrypted data.
- Encryption—It is the process of convertingplaintext to cipher text, using an encryption algorithm.
- Decryption—It is the reverse of encryption, i.e. converting cipher text to plaintext, using a decryption algorithm.
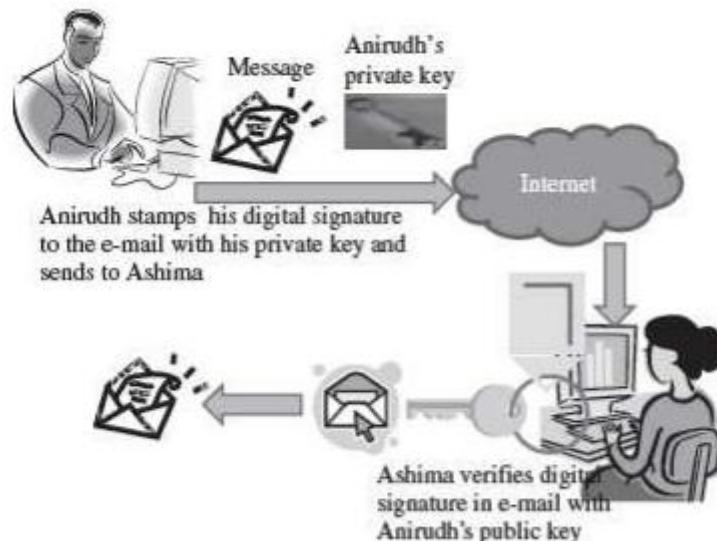
11.5.2 Digital Signature,
A signature on a legal, financial or any other document authenticates the document. A photocopy of that document does not count. For computerized documents, the conditions that a signed document must hold are—(1) The receiver is able to verify the sender (as claimed), (2) The sender cannot later repudiate the contents of the message, (3) The receiver cannot concoct the message himself. A digital signature is used to sign a computerized document. The properties of a digital signature are same as that of ordinary signature on a paper. Digital signatures are easy for a user to produce, but difficult for anyone else to forge. Digital signatures can be permanently tied to the content of the message being signed and then cannot be moved from one document to another, as such an attempt will be detectable.

Digital signature scheme is a type of asymmetric cryptography. Digital signatures use the publickey cryptography, which employs two keys—private key and public key. The digital signature scheme typically consists of three algorithms:

- Key generation algorithm—The algorithm outputs private key and a corresponding public key.
- Signing algorithm—It takes, message + private key, as input, and, outputs a digital signature.
- Signature verifying algorithm—It takes, message + public key + digital signature, as input, and, accepts or rejects digital signature.

11.5.3 Firewall,

A firewall is a security mechanism to protect a local network from the threats it may face while interacting with other networks (Internet). A firewall can be a hardware component, a software component, or a combination of both. It prevents computers in one network domain from communicating directly with other network domains. All communication takes place through the firewall, which examines all incoming data before allowing it to enter the local network.

Functions of Firewall:

The main purpose of firewall is to protect computers of an organization (local network) from unauthorized access. Some of the basic functions of firewall are:

- Firewalls provide security by examining the incoming data packets and allowing them to enter the local network only if the conditions are met.
- Firewalls provide user authentication by verifying the username and password. This ensures that only authorized users have access to the local network.
- Firewalls can be used for hiding the structure and contents of a local network from external users. Network Address Translation (NAT) conceals the internal network addresses and replaces all the IP addresses of the local network with one or more public IP addresses.

Working of Firewall:

The working of firewall is based on a filtering mechanism. The filtering mechanism keeps track of source address of data, destination address of data and contents of data. The filtering mechanism allows information to be passed to the Internet from a

local network without any authentication. It makes sure that the downloading of information from the Internet to a local network happens based only on a request by an authorized user.

11.5.4 Users Identification and Authentication,

Identification is the process whereby a system recognizes a valid user's identity. Authentication is the process of verifying the claimed identity of a user. For example, a system uses userpassword for identification. The user enters his password for identification. Authentication is the system which verifies that the password is correct, and thus the user is a valid user. Before granting access to a system, the user's identity needs to be authenticated. If users are not properly authenticated then the system is potentially vulnerable to access by unauthorized users. If strong identification and authentication mechanisms are used, then the risk that unauthorized users will gain access to a system is significantly decreased. Authentication is done using one or more combinations of—what you have (like smartcards), what you know (Password), and what you are (Biometrics like Fingerprints, retina scans).

We will now briefly discuss the following authentication mechanisms:
- User name and password
- Smart Card
- Biometrics—Fingerprints, Iris/retina scan

Once the user is authenticated, the access controls for the user are also defined. Access controls is what the user can access once he is authenticated.

1. User Name and Password

The combination of username and password is the most common method of user identification and authentication. The systems that use password authentication first require the user to have a username and a password. Next time, when the user uses the system, user enters their username and password. The system checks the username and password by comparing it to the stored password for that username. If it matches, the user is authenticated and is granted access to the system.

2. Smart Card

A smart card is in a pocket-sized card with embedded integrated circuits which can process data. With an embedded microcontroller, smart cards have the unique ability to store large amounts of data, carry out their own on-card functions (e.g. encryption and mutual authentication) and interact intelligently with a smart card reader. A smart card inserted into a smart card reader makes a direct connection to a conductive contact plate on the surface of the card (typically gold plated). Transmission of commands, data, and card status takes place over these physical contact points.

3 Biometric Techniques

Biometrics is the science and technology of measuring and statistically analyzing biological data. In information technology, biometrics refers to technologies that measures and analyzes human traits for authentication. This can include fingerprints,

eye retinas and irises, voice patterns, facial patterns and hand measurements, for authentication purposes.

11.5.5 Intrusion Detection Systems;

They complement firewalls to detect if internal assets are being hacked or exploited. A Network-based Intrusion Detection monitors real-time network traffic for malicious activity and sends alarms for network traffic that meets certain attack patterns or signatures. A Host-based Intrusion Detection monitors computer or server files for anomalies and sends alarms for network traffic that meets a predetermined attack signature.

11.6 Security Awareness;

The aim of the security awareness is to enhance the security of the organization's resources by improving the awareness of the need to secure the system resources. Staff members play a critical role in protecting the integrity, confidentiality, and availability of IT systems and networks. It is necessary for an organization to train their staff for security awareness and accepted computer practices. Security of resources can be ensured when the people using it are aware of the need to secure their resources. Security awareness of staff includes the knowledge of practices that must be adhered to, for ensuring the security and the possible consequences of not using those security practices. For example, not disclosing your password to unauthorized users is a security practice, but if the users are not aware of the possible consequences of disclosing the password, they may disclose their password to other users, unintentionally, thus making their systems prone to security attack. In order to make the users and people in an organization aware of the security practices to be followed, regular training programs are conducted in organizations. Awareness is also promoted by regular security awareness sessions, videotapes, newsletters, posters, and flyers.

11.7 Security Policy

Asecuritypolicyis a formal statement that embodies the organization's overall security expectations, goals, and objectives with regard to the organization's technology, system and information.

- To be practical and implementable, policies must be defined by standards, guidelines, and procedures. Standards, guidelines, and procedures provide specific interpretation of policies and instruct users, customers, technicians, management, and others on how to implement the policies.

- The security policy states what is, and what is not allowed. A security policy must be comprehensive, up-to-date, complete, delivered effectively, and available to all staff. A security policy must also be enforceable. To accomplish this, the security policy can mention that strict action will be taken against employees who violate it, like disclosing a password.

- Generally, security policies are included within asecurity plan. A security plan details how the rules put forward by the security policy will be implemented. The statements within a security plan can ensure that each employee knows the boundaries and the penalties of overstepping those boundaries. For example, some rules could be included in the security policy of an organization, such as, to log off the system before leaving the workstation, or not to share the password with other users.

- The security policy also includes physical security of the computers. Some of the measures taken to ensure the physical security of a computer are—taking regular backups to prevent data loss from natural calamity, virusattack or theft, securing the backup media, keeping valuable hardware resources in locked room (like servers), to avoid theft of systems and storage media.

The End