B.COM(HONS.)
Introduction to Information Technology-107

Unit I

Basics of Computer and its Evolution: Evolution of Computer, Data, Instruction and Information, Characteristics of Computers, Input-output Devices (Hardware, Software, Human ware and Firmware),Function of Different Units of Computer, Classification of Computers. Data Representation : Different Number System (Decimal, Binary, Octal and hexadecimal) and their Inter
Conversion (Fixed Point Only), Binary Arithmetic (Addition, Subtraction, Multiplication and Division), Computer Memory: Primary Memory, Secondary memory, Magnetic Disks, Optical Disks, Flash Memory, Mass storage devices, NAJ & SAN`s.

Unit II

Computer Software: Types of Software, Introduction to Operating System; Function of OS, Types of Operating Systems, Booting Procedure, Start-up Sequence, Details of Basic System Configuration, Important Terms like Directory, File, Volume, Label, Drive Name, etc; Introduction
to GUI using Windows Operating System, Compiler, Interpreter and assembler, Types of languages;
Word processor and software.

Unit III

Operating System Concept:: Introduction to DBMS, Structure of a DBMS, Advantages of DBMS, Data Models, E-R Modelling, SQL queries: Select, union; Intersect, except operators, Aggregate operators, creating and altering tables and views, Introduction to MS-Word, Spread Sheets and Graphical Solutions. All Directory Manipulations. Operating system commands.

Unit IV

Computer Networks and IT applications:
Data communication concepts, types of communication media, Concepts of Computer Networks, Network topologies, Networking devices, OSI model, Internet, Intranet and Extranets; Applications
of internet. Information Technology and Society: Indian IT Act, Application of information Technology in Railways, Airlines, Banking, Insurance, Inventory Control, Financial systems, Hotel
management, Education, entertainment and health, security issues in information technology.

UNIT-1

Computers:-

Computer is an electronic device that is designed to work with Information. The term 'computer is derived from the Latin term 'computare'; this means to calculate. Computer cannot do anything without a Program. It represents the decimal numbers through a string of binary digits. The Word 'Computer' usually refers to the Centre Processor Unit plus Internal memory.

Computer is an advanced electronic device that takes raw data as input from the user and processes these data under the control of set of instructions (called program) and gives the result (output) and saves output for the future use. It can process both numerical and non-numerical (arithmetic and logical) calculations. The basic components of a modern digital computer are: Input Device, Output Device, and Central Processor.

Charles Babbage is called the "Grand Father" of the computer. The First mechanical computer designed by Charles Babbage was called Analytical Engine. It uses read-only memory in the form of punch cards.

Four Functions about computer are:

accepts data	Input
processes data	Processing
produces output	Output
stores results	Storage

Input (Data):

Input is the raw information entered into a computer from the input devices. It is the collection of letters, numbers, images etc.

Process:

Process is the operation of data as per given instruction. It is totally internal process of the computer system.

Output:

Output is the processed data given by computer after data processing. Output is also called as Result. We can save these results in the storage devices for the future use.

Evolution of Digital Computers

The successful general purpose mechanical computers were developed. In 1930, mechanical calculations were built for automatic addition, subtraction, multiplication & division. A calculator is not a programmable device. The different eras of the evolution of the computer are listed below:

(1) **Mechanical Era** : There were many attempts to create a m/c that could help to perform various calculations. In 1823, Charles Babbage tried to build a mechanical as computing m/c capable of performing automatic mathematical calculations. This was designed to compute tables of functions such as logs functions etc. In 1830's Babbage made a more powerful mechanical computer. This m/c was designed to perform any mathematical calculation automatically. It could perform addition etc. It had a memory unit. Its capacity was 1000 numbers, each no. consisting of 50 digits. The m/c was a programmable m/c. It had mechanism for enabling a program to change the sequence of its operations automatically. In the late 19th century punched cards were commercially used. Soon IBM was formed in 1924. Konand Zuse developed a mechanical computer, the Z1, in 1938 in Germany.

(2) **The Electronic Era** : The first electronic computer using. Valves were developed by John V. Atanas off in the late 1930's. It contained add subtract unit. It was relatively a small computer and used about 300 valves. Its memory unit consisted of capacitors mounted on a rotating drum. It used a no. of I/O devices including a card punch and a card reader. The first popular general electronic digital computer was the ENIAC (Electronic Numerical Interpreter and calculator). John von Neumann was the consultant of the ENIAC project. The ENIAC used a high speed memory to store both programs as well as data during program execution. Neumann and his colleagues designed and build the IAS Computers. It used RAM consisting of a cathode ray tube. The transistors were invented in 1948 at AT&T bell laboratories. Slowly they replaced Vacuum tubes. IC's were first introduced, ie, designed and fabricated in 1958-59. The examples of computers using IC's are-: IBM – 370 & PDP-8. In 1970 LSI chips were introduced is form of memory units. Computers built in 1970's & onwards used micro process and other LSI, VLSI and ULSI components.

what is data: Data are plain facts or Raw Figures. When data are processed, organized, structured or presented in a given context so as to make them useful, they are called Information.

Instructions: set of rules and regulation

Information is data that has been processed in such a way as to be meaningful to the person who receives it.

Types of Information Systems:-

- Transaction Processing System (TPS)
- Management Information System (MIS)
- Support Systems
- Intelligent Systems

Characteristic of a computer

Basic characteristics about computer are:

1. Speed: - As we know computer can work very fast. It takes only few seconds for calculations that we take hours to complete. computer can perform millions (1,000,000) of instructions and even more per second.

Therefore, we determine the speed of computer in terms of microsecond (10-6 part of a second) or nanosecond (10 to the power -9 part of a second).

2. Accuracy: - The degree of accuracy of computer is very high and every calculation is performed with the same accuracy. The accuracy level is 7

3. Diligence: - A computer is free from tiredness, lack of concentration, fatigue, etc. It can work for hours without creating any error. If millions of calculations are to be performed, a computer will perform every calculation with the same accuracy. Due to this capability it overpowers human being in routine type of work.

4. Versatility: - It means the capacity to perform completely different type of work.

5. Power of Remembering: - Computer has the power of storing any amount of information or data..

6. No IQ: - Computer is a dumb machine and it cannot do any work without instruction from the user.

7. No Feeling: - It does not have feelings or emotion, taste, knowledge and experience. Thus it does not get tired even after long hours of work. It does not distinguish between users.

8. Storage: - The Computer has an in-built memory where it can store a large amount of data. We can also store data in secondary storage devices such as floppies, which can be kept outside your computer and can be carried to other computers.

Various fields of application of Computers

- Education

- Health and Medicine
- Science
- Business
- Recreation and Entertainment
- Government
- Defence

The various generations of computers an listed below :

(i) First Generation

The period of first generation was 1946-1959.

First generation of computer started with using vacuum tubes as the basic components for memory and circuitry for CPU(Central Processing Unit). These tubes like electric bulbs produced a lot of heat and were prone to frequent fusing of the installations, therefore, were very expensive and could be afforded only by very large organisations.

In this generation mainly batch processing operating system were used. In this generation Punched cards, Paper tape, Magnetic tape Input & Output device were used.

There were Machine code and electric wired board languages used.

The main features of First Generation are:

- Vacuum tube technology
- Unreliable
- Supported Machine language only
- Very costly
- Generate lot of heat
- Slow Input/Output device
- Huge size
- Need of A.C.
- Non portable
- Consumed lot of electricity

Limitations of First Generation Computer

1. They used valves or vacuum tubes as their main electronic component.
2. They were large in size, slow in processing and had less storage capacity.
3. They consumed lots of electricity and produced lots of heat.

4. Their computing capabilities were limited.
5. They were not so accurate and reliable.
6. They used machine level language for programming.
7. They were very expensive.

Example: ENIAC, UNIVAC, IBM 650 etc

(ii) **Second Generation (1955-1964)** : The second-generation computer used transistors for CPU components & ferrite cores for main memory & magnetic disks for secondary memory. They used high-level languages such as FORTRAN (1956), ALGOL (1960) & COBOL (1960 - 1961). I/O processor was included to control I/O operations.

Around 1955 a device called Transistor replaced the bulky Vacuum tubes in the first generation computer. Transistors are smaller than Vacuum tubes and have higher operating speed. They have no filament and require no heating. Manufacturing cost was also very low. Thus the size of the computer got reduced considerably.

It is in the second generation that the concept of Central Processing Unit (CPU), memory, programming language and input and output units were developed. The programming languages such as COBOL, FORTRAN were developed during this period.

Some of the computers of the Second Generation were:-

1. IBM 1620: Its size was smaller as compared to First Generation computers and mostly used for scientific purpose.

2. IBM 1401: Its size was small to medium and used for business applications.

3. CDC 3600: Its size was large and is used for scientific purposes.

Features:

1. Transistors were used instead of Vacuum Tube.

2. Processing speed is faster than First Generation Computers (Micro Second)

3. Smaller in Size (51 square feet)

4. The input and output devices were faster.

Example: IBM 1400 and 7000 Series, Control Data 3600 etc.

(iii) **Third Generation (1964-1977)** : By the development of a small chip consisting of the capacity of the 300 transistors. These ICs are popularly known as Chips. A single IC has many transistors, registers and capacitors built on a single thin slice of silicon. So it is quite obvious that the size of the computer got further reduced. Some of the computers developed during this period were IBM-360, ICL-1900, IBM-370, and VAX-750. Higher level language such as

BASIC (Beginners All purpose Symbolic Instruction Code) was developed during this period. Computers of this generation were small in size, low cost, large memory and processing speed is very high. Very soon ICs Were replaced by LSI (Large Scale Integration), which consisted about 100 components. An IC containing about 100 components is called LSI.

Features:

1. They used Integrated Circuit (IC) chips in place of the transistors.

2. Semi conductor memory devices were used.

3. The size was greatly reduced, the speed of processing was high, they were more accurate and reliable.

4. Large Scale Integration (LSI) and Very Large Scale Integration (VLSI) were also developed.

5. The mini computers were introduced in this generation.

6. They used high level language for programming.

Example: IBM 360, IBM 370 etc.

(iv) **Fourth Generation**: An IC containing about 100 components is called LSI (Large Scale Integration) and the one, which has more than 1000 such components, is called as VLSI (Very Large Scale Integration). It uses large scale Integrated Circuits (LSIC) built on a single silicon chip called microprocessors. Due to the development of microprocessor it is possible to place computer's central processing unit (CPU) on single chip. These computers are called microcomputers. Later very large scale Integrated Circuits (VLSIC) replaced LSICs. Thus the computer which was occupying a very large room in earlier days can now be placed on a table. The personal computer (PC) that we see in your school is a Fourth Generation Computer Main memory used fast semiconductors chips up to 4 M bits size. Hard disks were used as secondary memory. Keyboards, dot matrix printers etc. were developed. OS-such as MS-DOS, UNIX, Apple's Macintosh were available. Object oriented language, C++ etc were developed.

Features:

1. They used Microprocessor (VLSI) as their main switching element.

2. They are also called as micro computers or personal computers.

3. Their size varies from desktop to laptop or palmtop.

4. They have very high speed of processing; they are 100% accurate, reliable, diligent and versatile.

5. They have very large storage capacity.

Example: IBM PC, Apple-Macintosh etc.

(v) **Fifth Generation (1991- continued)** : 5th generation computers use ULSI (Ultra-Large Scale Integration) chips. Millions of transistors are placed in a single IC in ULSI chips. 64 bit microprocessors have been developed during this period. Data flow & EPIC architecture of these processors have been developed. RISC & CISC, both types of designs are used in modern processors. Memory chips and flash memory up to 1 GB, hard disks up to 600 GB & optical disks up to 50 GB have been developed. fifth generation digital computer will be Artificial intelligence.

- VLSI technology used
- Very cheap
- Portable and reliable
- Use of PC's
- Very small size
- Pipeline processing
- No A.C. needed
- Concept of internet was introduced
- Great developments in the fields of networks
- Computers became easily available

Various Input and Output Devices

The devices which are used to input the data and the programs in the computer are known as "Input Devices'. or Input device can read data and convert them to a form that a computer can use. Output Device can produce the final product of machine processing into a form usable by humans. It provides man to machine communication. Some of the I/O devices are explained below:

(1) **Keyboard** : Keyboard is used in the input phase of a computer-based information system. Keyboard is most common input device is used today. The data and instructions are input by typing on the keyboard. The message typed on the keyboard reaches the memory unit of a computer. It's connected to a computer via a cable. Apart from alphabet and numeral keys, it has other function keys for performing different functions.

(2) **Mouse :** It's a pointing device. The mouse is rolled over the mouse pad, which in turn controls the movement of the cursor in the screen. We can click, double click or drag the mouse. Most of the mouse's have a ball beneath them, which rotates when the mouse in moved. The ball has 2 wheels of the sides, which in turn mousse with the movement of the ball. The sensor notifies the speed of its movements to the computer, which in turn moves the cursor/pointer on the screen.

(3) **Scanner** : Scanners are used to enter information directly in to the computers memory. This device works like a Xerox machine. The scanner converts any type of printed or written information including photographs into digital pulses, which can be manipulated by the computer.

(4) **Track Ball** : Track ball is similar to the upside- down design of the mouse. The user moves the ball directly, while the device itself remains stationary. The user spins the ball in various directions to effect the screen movements.

(5) **Light Pen** : This is an input device which is used to draw lines or figures on a computer screen. It's touched to the CRT screen where it can detect raster on the screen as it passes.

(6) **Optical Character Reader**: It's a device which detects alpha numeric characters printed or written on a paper. The text which is to be scanned is illuminated by a low frequency light source. The light is absorbed by the dark areas but reflected from the bright areas. The reflected light is received by the photocells.

(7) **Bar Code Reader** : This device reads bar codes and coverts them into electric pulses to be processed by a computer. A bar code is nothing but data coded in form of light and dark bars.

(8) **Voice Input Systems** : This devices converts spoken words to M/C language form. A micro phone is used to convert human speech into electric signals. The signal pattern is then transmitted to a computer when it's compared to a dictionary of patterns that have been

previously placed in a storage unit of computer. When a close match is found, the word is recognized.

(9) **Plotter** : Plotter is an O/P device that is used to produce graphical O/P on papers. It uses single color or multi color pens to draw pictures as blue print etc.

(10) **Digital Camera** : It converts graphics directly into digital form. It looks like an ordinary camera, but no film is used therein, instead a CCD (changed coupled Divide) Electronic chip in used. When light falls, on the chip though the lens, it converts light waves into electrical waves.

(11) Barcode Reader

When we buy products from shops we usually see small labels containing white and black stripes called barcodes. These labels contain information on the product. In order to read this information, one must pass the barcode in front of a special reader. This reader can be in a form of a pen or installed on the shop's counter. The barcode reading consists of a light emitting diode that will emit light on the code. The reflected light from the barcode is translated to the computer by a light sensitive diode.

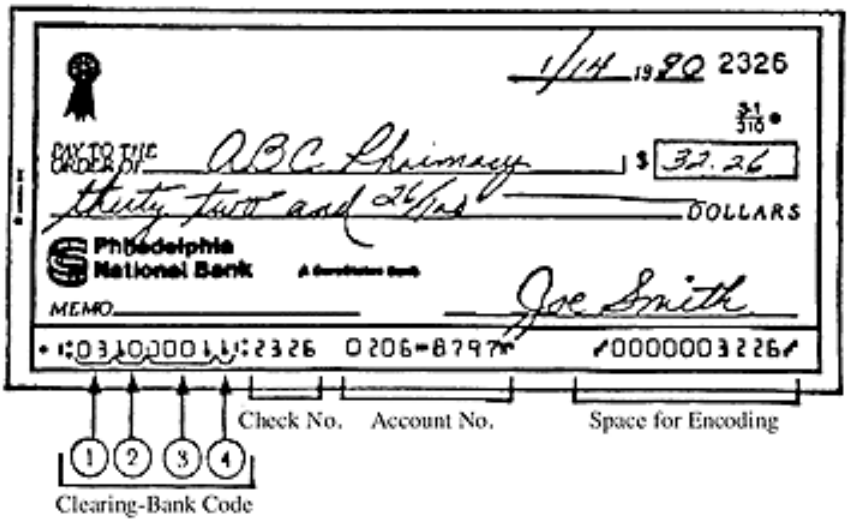
(12) **Optical Mark Reader**

This device reads marks in predetermined places on specially prepared cards. These are usually used in multiple-choice questions. The student will have to fill a space with a pencil, usually HB. The filled examination paper is then inserted into the OMR so as to
enable it to read the answers marked.

(13) **Magnetic Ink**

Magnetic Ink Character Reader (MICR)

This system uses special ink to write numbers with coded magnetic ink. This special ink contains the written numbers magnetically. This system is widely used on cheques issued by banks. The MICR is much faster than OCR or other methods when reading codes and thus identifying cheques.



(14)Touchpad / Touch screen

This is a small sensitive pad used on portable computers (laptops). The touch pad is used as a pointing device. The pointer is moved on screen by touching the touch pad with the finger. Tapping on the touch pad will give the same results as when using the mouse buttons. The touchpad is particularly useful for persons with disability whom can have limitations in moving their hand.

Some other Output Devices:-

- Softcopy – Visual (monitor) or Sound (speakers)
- Hardcopy – Output on a tangible (something you can touch) such as a printer printout.

- **Monitor or Visual Display Unit (VDU)**

The monitor is the most common type of output device and is also called Visual Display Unit (VDU).

Visual Display Unit (VDU)

Monochrome – This type of display uses only one type of colour or a greyscale

2. Colour – This type of display uses more than one type of colour and is generally called RGB monitors. The RGB stand for Red, Green and Blue and a combination of these colours are used to display the image

on the monitor. Colour monitors can display range from 16 to 16.7 million colours.

Resolution

This term refers to the clarity or sharpness of a printout or display screen. The more pixels there are per square inch, the better and greater the resolution. One can find different standards for

monitors. They support different colour depths (number of colours it supports). The most common standards for monitors are:-

- **VGA (Video Graphics Array)** mode appeared in 1987. It offered a resolution of 720x400 in text mode and a resolution of 640 by 480 (640x480) in 16-colour graphics mode. It also offered a resolution of 1024 x 768. The VGA quickly became the baseline display mode for PCs.
- **SVGA (Super Video Graphics Array)** is a graphics mode which can display 256 colours at resolutions of 640x200, 640x350 and 640x480. SVGA can also display higher definitions such as 800x600 or 1024x768 by using fewer colours.
- **XGA (extended Graphics Array).** is an IBM display standard introduced in 1990. Today, it is the most common appellation of the 1024 × 768 pixels display resolution, but the official definition is broader than that. It was not a new and improved replacement for Super VGA, but rather became one particular subset of the broad range of capabilities covered under the "Super VGA" umbrella. The initial version of XGA expanded upon IBM's VGA, adding support for two resolutions:
  - 800 × 600 pixels with high color (16 bits per pixel, i.e. 65,536 colors).
  - 1024 × 768 pixels with a palette of 256 colors (8 bits per pixel).

The Different Types of Monitors

- **Cathode Ray Tube (CRT)** – uses the same technology as used in television sets. The CRT uses a vacuum tube in which an electron gun is installed.When electrons fired by the electron gun touch the phosphor layer situated at the internal side of the screen, this will glow. This dot of light is called a Pixel.

- **Liquid Display Unit / Liquid Crystal Display (LCD**) – we usually find this installed in laptops and handheld computers. The technology used is different from that of the CRT. LCD uses tiny capsules filled with liquid crystals. When the molecules of the liquid are subjected to an electric field, the liquid crystals align and light reflects off them. Without the field, their alignment reverts to its original, non-reflecting arrangement, so the elements appear dark. Combinations of LCD's are arranged to form patterns of reflected light that spell out numbers, symbols or letters. TFT (Thin Film Transistor) is a variant of LCD which uses thin film transistor technology to improve image quality.

Printers:-

Printers are output devices that transfer the output of the computer on paper (hardcopy).

The printers can be divided into two categories:-

Impact and Non-Impact

• **Impact** – refers to all those printers whose printing heads touch the paper. An example is the Dot matrix printer.

• **Non-Impact** – are all those printers whose printing heads do not touch the paper. An example is a Laser or an Ink-Jet printer.

The following is a list of the different types of printers that exist:

- **Daisy Wheel Printers**

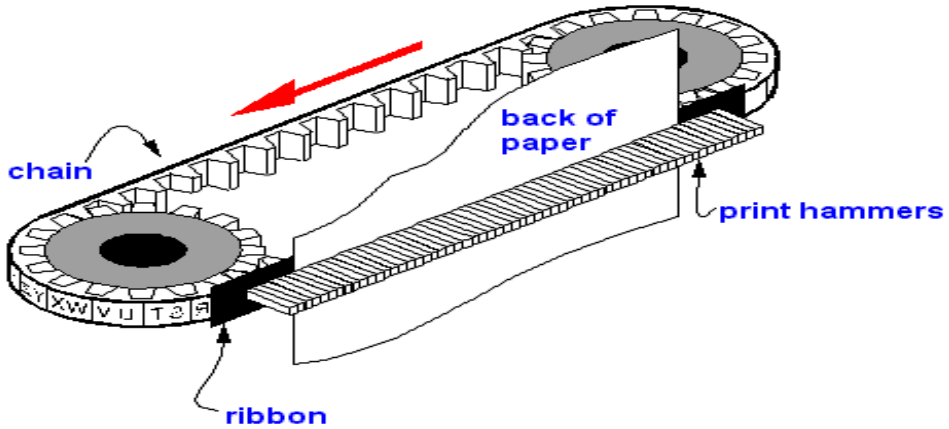
This type of printer is rarely used any more since new printer technology, which is more efficient and economic, has been discovered. The principle of this type of printer is that number and letter are each situated on metal leafs as shown.

These leafs make part of the daisy wheel. In order to print, the daisy wheel rotates while a hammer will strike each character, which will make contact with the printer ribbon and the paper.

- **Line Printers**

As the name suggests these type of printers will print one whole document line at one go. There are two basic types of line printers:

1. **Barrel Printers** – This printer has a number of disks each containing all the letters and numbers. These disks are situated one near each other on a shaft and can rotate independently from each other. Lines are written by revolving each disk in the correct position in order to form a complete line. A row of hammers will strike the ribbon and papers.
2. **Chain Printers** – All the alphabet and numbers are imprinted on a chain. In order to write on paper, a hammer will strike the right characters to print.



3. **Dot Matrix Printers**

Dot matrix printers are categorised as character printers because they write one character at a time. The printing head is made up of several pins, which form a column. The printing head is attached to a belt which moves it from side to side along the width of the paper and at the same time the paper is rolled up by means of a roller. In order to write, the head pins are fired with the aid of an electro magnet, which hit the ribbon and paper.

4.Inkjet Printers

These types of printers are the most popular nowadays. It is basically what most home users have today. These printers have good resolution, are silent and for small amount of printing are not particularly expensive. These printers are called page printers.

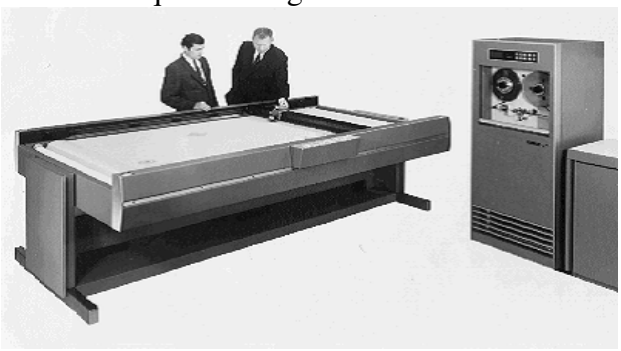
5) **Laser Printers**

Laser printers offer the best quality in terms of printout quality and resolution. On the other hand they are the most expensive to run. Laser printers are called page printers since they print one page at a go. The laser printer uses a laser beam, a drum and a toner.

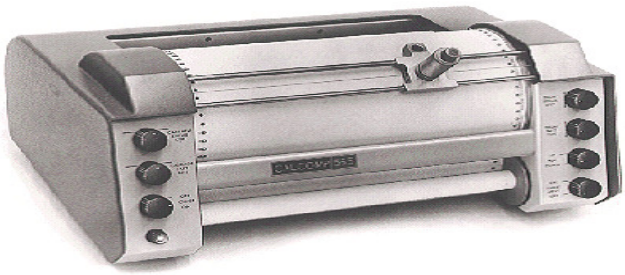
Plotters

The plotter uses pens to write on paper and is usually used for engineering drawings. There are two types of plotters:

1. **Flatbed Plotter** – This uses a flat paper which is fixed, and pens (or styluses) that move on it draw the required design.



2. Drum Plotter – Here the paper moves along a drum vertically while pens move side to side horizontally.



Soundcard and Speakers

These devices are used to produce sound from the computer system. The soundcard is a card installed in the tower case of the computer. Its role is to convert the signal coming out of the computer to a format that can be reproduced by the external speakers.

Components of Computer

The components can be classified into two broad categories as :

Hardware:- Hardware refers to any physical component of a computer. For example, CPU, Monitor, Keyboard, Hard Disk, Floppy Disk, Printer, etc. Are physical components and thus, is hardware.

Software :- Software refers to the programs, which are required to operate the computer. In other words, we can describe software as the logical entity, which enables the physical entities (hardware) to function properly. " A set of programs is also called as Software."

Firmware:- Micro programs written in machine languages are called firmware. They are permanently stored on ROM.

Human ware:- Human ware is hardware and software that emphasizes user capability and empowerment and the design of the user interface

Advantages Of Computers:-

- Computer greatly improves the speed of data input, output, processing, and transmission.
- Computer ensures consistent and error free processing of data.
- Digitization of all kinds of information including sounds and images, combined with massive information processing capabilities of the computer has resulted in development of application to produce physical products of very high quality at great speed and very economically.

Major disadvantages of computers include:

- Computer is highly dependent on the quality of input data fed to it. Though computers are very fast in tasks that are pre-programmed, it lake the ability of human brain to detect and correct errors that it is not specifically programmed to do.
- The task of programming a computer for a computer application is very costly and time consuming. This reduces the utility of computers for applications that are non-repetitive.
- Computer systems are rather rigid. Once a computers system is designed and programmed, making even minor corrections or improvements can be quite costly and time consuming. For this reason a great care is required in design and development of computer systems.
- Computers require use of sophisticated equipment and support facilities. For example, a person solving a problem of maths using just a pen and paper can carry these with him or her anywhere with ease. However if the same problem is to be solved using a computer, the person will not only need access to a suitable computer, it will also be necessary to have the required software and suitable electric power to run the computer.

Block diagram of computer and its various components :-

A computer can process data, pictures, sound and graphics. They can solve highly complicated problems quickly and accurately. A computer as shown in Fig. performs basically five major computer operations or functions irrespective of their size and make. These are

1) It accepts data or instructions by way of input,

2) It stores data,

3) It can process data as required by the user,

4) It gives results in the form of output, and

5) It controls all operations inside a computer.

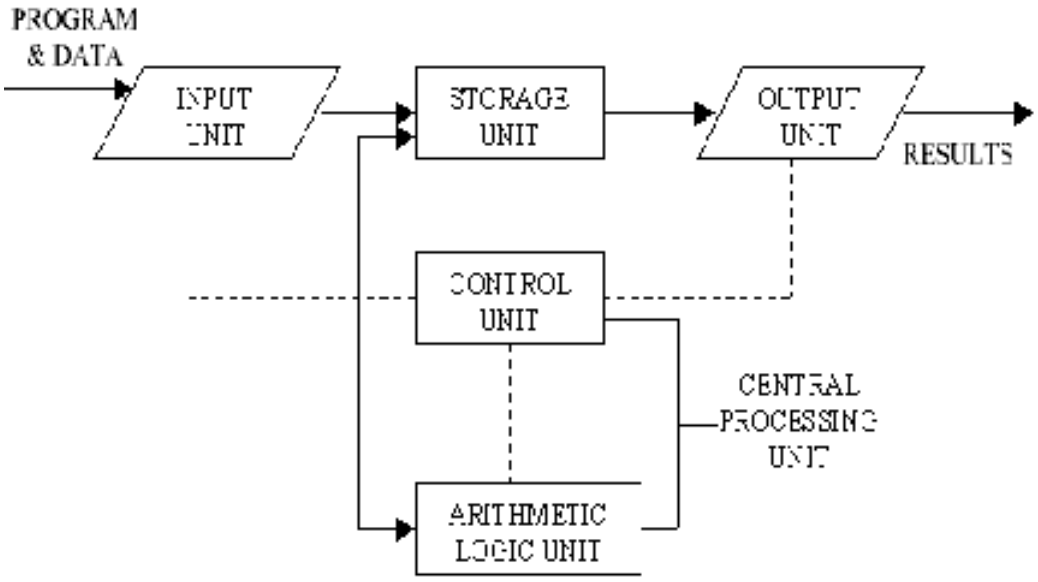


Fig : ***Basic computer Operations***

1. **Input**: This is the process of entering data and programs in to the computer system. We know that computer is an electronic machine like any other machine which takes as inputs raw data and performs some processing giving out processed data. Therefore, the input unit takes data from us to the computer in an organized manner for processing.

2. **Storage**: The process of saving data and instructions permanently is known as storage. Data has to be fed into the system before the actual processing starts. It is because the processing speed of Central Processing Unit (CPU) is so fast that the data has to be provided to CPU with the same speed. Therefore the data is first stored in the storage unit for faster access and processing. This storage unit or the primary storage of the computer system is designed to do the above functionality. It provides space for storing data and instructions.

The storage unit performs the following major functions:

• All data and instructions are stored here before and after processing.

• Intermediate results of processing are also stored here.

3. **Processing**: The task of performing operations like arithmetic and logical operations is called processing. The Central Processing Unit (CPU) takes data and instructions from the storage unit and makes all sorts of calculations based on the instructions given and the type of data provided. It is then sent back to the storage unit.

4. **Output**: This is the process of producing results from the data for getting useful information.

5. **Control**: The manner how instructions are executed and the above operations are performed. Controlling of all operations like input, processing and output are performed by control unit. It takes care of step by step processing of all operations inside the computer.

FUNCTIONAL UNITS

In order to carry out the operations mentioned in the previous section the computer allocates the task between its various functional units. The computer system is divided into three separate units for its operation. They are

1) Arithmetic logical unit

2) Control unit.

3) Central processing unit.

Arithmetic Logical Unit (ALU) Logical Unit

Logical Unit : After we enter data through the input device it is stored in the primary storage unit. The actual processing of the data and instruction are performed by Arithmetic Logical Unit. The major operations performed by the ALU are addition, subtraction, multiplication, division, logic and comparison. Data is transferred to ALU from storage unit when required. After processing the output is returned back to storage unit for further processing or getting stored.

Control Unit (CU)

The next component of computer is the Control Unit, which acts like the supervisor seeing that things are done in proper manner. Control Unit is responsible for co ordinating various operations. The control unit determines the sequence in which computer programs and instructions are executed. Things like processing of programs stored in the main memory, interpretation of the instructions and issuing of signals for other units of the computer to execute them.

Central Processing Unit (CPU)

The ALU and the CU of a computer system are jointly known as the central processing unit. We can call CPU as the brain of any computer system. It is just like brain that takes all major decisions, makes all sorts of calculations and directs different parts of the computer functions by activating and controlling the operations

Classification of Computers

According to functionality, computers are classified as

Analog Computer

An analog computer (spelt analogue in British English) is a form of computer that uses continuous physical phenomena such as electrical, mechanical, or hydraulic quantities to model the problem being solved

Digital Computer

A computer that performs calculations and logical operations with quantities represented as digits, usually in the binary number system

Hybrid Computer (Analog + Digital)

A combination of computers those are capable of inputting and outputting in both digital and analog signals. A hybrid computer system setup offers a cost effective method of performing complex simulations.

On the basis of Size

Super Computer

The fastest and most powerful type of computer Supercomputers are very expensive and are employed for specialized applications that require immense amounts of mathematical calculations. For example, weather forecasting requires a supercomputer. Other uses of supercomputers include animated graphics, fluid dynamic calculations, nuclear energy research, and petroleum exploration.

The chief difference between a supercomputer and a mainframe is that a supercomputer channels all its power into executing a few programs as fast as possible, whereas a mainframe uses its power to execute many programs concurrently.

Mainframe Computer

A very large and expensive computer capable of supporting hundreds, or even thousands, of users simultaneously. In the hierarchy that starts with a simple microprocessor (in watches, for example) at the bottom and moves to supercomputers at the top, mainframes are just below supercomputers. In some ways, mainframes are more powerful than supercomputers because they support more simultaneous programs. But supercomputers can execute a single program faster than a mainframe.

Desktop Computer: a personal or micro-mini computer sufficient to fit on a desk.

Laptop Computer: a portable computer complete with an integrated screen and keyboard. It is generally smaller in size than a desktop computer and larger than a notebook computer.

Palmtop Computer/Digital Diary /Notebook /PDAs: a hand-sized computer. Palmtops have no keyboard but the screen serves both as an input and output device.

Workstations

A terminal or desktop computer in a network.

General Purpose Computers: On the other hand, these can perform more than one task provided that the correct instructions are loaded in the computer, therefore depending on what PROGRAM we load, the computer may act differently.

Examples:
- Computer + 'Word' used for word processing – typing
- Computer + 'Excel' used for accounting purposes – spreadsheet
- Computer + 'Paint' used for drawing
- Computer + 'AutoCAD' used for designing layouts – architects, engineers, etc.\

Data Representation (NUMBER SYSTEM)

Decimal Number System

The number system that we use in our day-to-day life is the decimal number system. Decimal number system has base 10 as it uses 10 digits from 0 to 9. In decimal number system, the successive positions to the left of the decimal point represent units, tens, hundreds, thousands and so on.

Each position represents a specific power of the base (10). For example, the decimal number 1234 consists of the digit 4 in the units position, 3 in the tens position, 2 in the hundreds position, and 1 in the thousands position, and its value can be written as

(1x1000)+ (2x100)+ (3x10)+ (4xl)

(1x103)+ (2x102)+ (3x101)+ (4xl00)

1000 + 200 + 30 + 1

1234

Binary Number System

Characteristics

- Uses two digits, 0 and 1.
- Also called base 2 number system
- Each position in a binary number represents a 0 power of the base (2). Example 2^0

- Last position in a binary number represents a x power of the base (2). Example 2^x where x represents the last position - 1.

Example

Binary Number: 10101_2

Calculating Decimal Equivalent:

Step	Binary Number	Decimal Number
Step 1	10101_2	$((1 \times 2^4) + (0 \times 2^3) + (1 \times 2^2) + (0 \times 2^1) + (1 \times 2^0))_{10}$
Step 2	10101_2	$(16 + 0 + 4 + 0 + 1)_{10}$
Step 3	10101_2	21_{10}

Note: 10101_2 is normally written as 10101.

Octal Number System

Characteristics

- Uses eight digits, 0,1,2,3,4,5,6,7.
- Also called base 8 number system
- Each position in a octal number represents a 0 power of the base (8). Example 80
- Last position in a octal number represents a x power of the base (8). Example 8x where x represents the last position - 1.

Example

Octal Number: 12570_8

Calculating Decimal Equivalent:

Step	Octal Number	Decimal Number
Step 1	12570_8	$((1 \times 8^4) + (2 \times 8^3) + (5 \times 8^2) + (7 \times 8^1) + (0 \times 8^0))_{10}$
Step 2	12570_8	$(4096 + 1024 + 320 + 56 + 0)_{10}$
Step 3	12570_8	5496_{10}

Note: 12570_8 is normally written as 12570.

Hexadecimal Number System

Characteristics

- Uses 10 digits and 6 letters, 0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F.
- Letters represents numbers starting from 10. A = 10. B = 11, C = 12, D = 13, E = 14, F = 15.
- Also called base 16 number system
- Each position in a hexadecimal number represents a 0 power of the base (16). Example 160
- Last position in a hexadecimal number represents a x power of the base (16). Example 16x where x represents the last position - 1.

Example

Hexadecimal Number: 19FDE16

Calculating Decimal Equivalent:

Step	**Binary Number**	**Decimal Number**
Step 1	$19FDE_{16}$	$((1 \times 16^4) + (9 \times 16^3) + (F \times 16^2) + (D \times 16^1) + (E \times 16^0))_{10}$
Step 2	$19FDE_{16}$	$((1 \times 16^4) + (9 \times 16^3) + (15 \times 16^2) + (13 \times 16^1) + (14 \times 16^0))_{10}$
Step 3	$19FDE_{16}$	$(65536+ 36864 + 3840 + 208 + 14)_{10}$
Step 4	$19FDE_{16}$	106462_{10}

CONVERSIONS

1) DECIMAL NUMBER TO BINARY NUMBER

To convert a decimal number into binary number it requires successive division by 2writing down each quotient and its remainder. The remainders are taken in the reverse order, which is the

binary equivalent of the decimal number. For example, let it is required to convert the decimal number 25 to its binary equivalent.

2 25

2 12 -1

2 6 - 0

2 3 -0

1 -1

The binary equivalent for $(25)_{10}$ $=(11001)_2$. The binary equivalent for $(25)_{10}$ $=(11001)_2$.To convert decimal fractions into equivalent binary fractions repeatedly double t h e decimal fraction. The number (0 or 1) that appears on the left is written separately. The bits that are written in this manner are read from top to bottom with a decimal point on the left. For example if the given number is 0.0625, conversion is done in the following manner.

0.625 *2 = 0 .1250

0.125 *2 = 0 .25

0.25 *2 = 0 .5

0.5 *2 = 1 .0

Bits on the left: 0

0

0

1

The Multiplication cannot be continued further, as the fractional part in the previous step has already become zero. Therefore, 0.062510 = .(0001)2

2) **DECIMAL NUMBER TO OCTAL NUMBER**

Conversion from decimal to octal can be performed by repeatedly dividing the decimal number by 8 and using each remainder as a digit in the octal number being formed. For instance, to convert decimal number 200 to an octal representation, we divide as follows.

8 200

8 25 - 0

3 - 1

Therefore $(200)_{10} = (310)_8$

3) **DECIMAL NUMBER TO HEXADECIMAL NUMBER**

One way to convert decimal to Hexadecimal is the hex dabbles. The idea is as divide successively by 16, writing down the remainders. Here is a sample of how it is done. To convert decimal 2429 to hexadecimal,

16 2429

16 154 - 15 (F)

9 - 10 (A)

Therefore $(2429)_{10} = (9AF)_{16}$

4) **BINARY NUMBER TO DECIMAL NUMBER**

For converting the value of Binary numbers to decimal equivalent we have to find itsquantity, which is found by multiplying a digit by its place value. The following example illustrates the conversion of binary numbers to decimal system.

101 =1*23-1 +0*23-2+1*23-3

= 1*22 +0*21+1*20

= 4 + 0 +1

= 5

1001 =1*24-1 +0*24-2+0*24-3 + 1*24-4

= 1*23 + 0*22 + 0*21 + 1*20

= 8+1

= 9

11.011 =1*22-1 +1*22-2+0*22-3 +1*22-4+1*22-5

=1*21 +1*20+0*2-1 +1*2-2+1*2-3

= 2+1+0+1/4+1/8

= 2+1+1/4+1/8

=3 3/8

5). **BINARY NUMBER TO OCTAL NUMBER**

There is a simple trick for converting a binary number to an octal number. Simply group the binary digits into groups of 3, starting at the octal point, and read each set of three binary digits according to the following table .

BINARY	OCTAL
000	0
001	1
010	2
011	3
100	4
101	5
110	6
111	7

Let us convert the binary number 011101 into octal. First, we break binary number into 3 digits (011 101). Then converting each group of three binary digits, we get 35 in octal. Therefore 011101 binary = 35 in octal.

011101 = 011 101

= 3 5

= (3 5) 8

6) **BINARY NUMBER TO HEXADECIMAL NUMBER**

To convert binary to hexadecimal, we simply break a binary number into groups of four digits and convert each group of four digits according to the preceding code. Here are some examples:

Example:-1 (10111011)2 = 1011 1011

= B B

= (B B)16

Example:2 (10010101)2 = 1001 0101

= 9 5

= (9 5)16

7) **OCTAL NUMBER TO DECIMAL NUMBER**

To convert an octal number to a decimal number, we use the same sort of polynomial as was used in the binary case, except that we now have a radix of 8 instead of 2. Therefore 1213 in octal is :-

=1*83+2*82+1*81+3*80

=512 + 128 + 8 + 3 = 651 in decimal. Also, 1.123 in octal is 1*80+1*8-1+2*8-2+3*8-3

=1 83/512 in decimal

8) **OCTAL NUMBER TO BINARY NUMBER**

The conversion from octal number to binary number is easily accomplished. Each octal bit is converted to its three digit binary equivalent.

Example: 1 (2 6 1 5 3 . 7 4 0 6)8

(010 110 001 101 011. 111 100 000 110)2

9.) **OCTAL NUMBER TO HEXADECIMALNUMBER**

The method of converting octal to hexadecimal number is to convert the given octal number to binary number and then arrange the binary digits into groups of 4 starting at the binary point.

Example :1

Convert octal number 714.06 to hexadecimal.

(714.06)8 =(111 001 100.000 110)2

=0001 1100 1100 . 0001 1000

=(1 C C . 1 8)

The hexadecimal equivalent of (714.06)8 is (1CC.18)16

10). **HEXADECIMAL NUMBER TO BINARY NUMBER**

To convert a hexadecimal number to a binary number, convert each hexadecimal digit to its 4-bit equivalent using the code. For instance, here's how 9AF converts to binary.

(3 0 6 . D)16

(0011 0000 0110 . 1101)2

11). **HEXADECIMAL NUMBER TO OCTAL NUMBER**

The conversion of Hexadecimal number to octal number involves two steps. First the method suggests to go from hexadecimal to binary numbers and second from binary to octal numbers Convert the hexadecimal into binary by writing 4 bits binary value for each bit in hexadecimal number and then arrange the binary digits into groups of three starting at the binary point.

Example: convert Hexadecimal (1E.C) to octal conversion

(1E.C)16 = (0001 1110.1100)2

= (011 110.110)

= (3 6. 6)

The octal equivalent of (1E.C)16 is (36.6)8

12). **HEXADECIMAL NUMBER TO DECIMAL NUMBER**

The conversion of Hexadecimal to decimal is straightforward but time consuming. In Hexadecimal number system each digit position corresponds to a power of 16. The weights of the digit positions in a hexadecimal number are as follows: For instance, BB represents BB=B*161 +B*160

=11*16+11*1

=176+11

=(187)10

13) BINARY ADDITION AND SUBTRACTION

Binary addition is performed in the same manner as decimal addition. The complete table for binary addition is as follows:

0+0=0

0+1=1

1+0=1

1+1=0 plus a carryover of 1

'Carry over' are performed in the same manner as in decimal arithmetic. Since 1 is the largest digit in the binary system, any sum greater than 1 requires that a digit be carried.

Examples:

Decimal	Binary
5	101
6	110
11	1011

Subtraction is the inverse operation of addition. To subtract, it is necessary to establish a procedure for subtracting a larger from a smaller digit. The only case in which this occurs with binary numbers is when 1 is subtracted from 0. It is necessary to borrow 1 from the next column to the left. This is the binary subtraction table.

0-0=0

1-0=1

0-1=1 with a borrow of 1

1-1=0

Examples:

Decimal	Binary
9	1001
-5	-101

4 100

14) **BINARY MULTIPLICATION AND DIVISION**

The table for binary multiplication is very short, with only four entries instead of the many for normal decimal multiplication

0*0=0

0*1=0

1*0=0

1*1=1

The following examples of binary multiplication illustrate the simplicity of each operation. It is only necessary to copy the multiplicand if the digit in the multiplier is 1 and to copy all 0's if the digit in the multiplier is a 0.

111		10.1	
101	*	10.1	*
111		101	
000*		000*	
111**		101**	
100011		110.01	

The complete table for binary division is as follows:

0/1=0

1/1=1

Examples: 111

110 101010

110

1001

110

110

110

0

111.1

110 101101

110

1010

110

1001

110

0110

110

0

BINARY CODED DECIMAL NUMBER REPRESENTATION

In BCD number system a group of binary bit is used to represent each of 10 decimal digits. For instances, an obvious and natural code is a simple weighted binary code as shown in table

BINARY CODE	DECIMAL DIGIT
0000	0
0001	1
0010	2
0011	3
0100	4

0101	5
0110	6
0111	7
1000	8
1001	9

Binary Coded Decimal

This is known as a binary coded decimal 8421 code or simply BCD. Notice that 4 binary bits are required for each decimal digit and each bit is assigned a weight; for instance the rightmost bit has a weight of 1, and the leftmost bit in each code group has a weight of 8. By adding the weights of the positions in which 1's appear, the decimal digit represented by a code group may be derived. This is somewhat uneconomical since 24=16, and thus the 4 bits could actually represent 15 different values. For the decimal number 1246 to be represented, 16 bits are required: 0001 0010 0100 0110

Examples:

Convert decimal 4019 to BCD

4 0 1 9

0100 0000 0001 1001

The BCD equivalent of (4019)10 is 0100 0000 0001 1001

Convert BCD number 0001 1001 0000 0111 to decimal

0001 1001 0000 0111

1 9 0 7

The decimal equivalent of BCD Number 0001 1001 0000 0111 is (1907)10. BCD numbers are useful wherever decimal information is transferred into a computer. The pocket calculator is one of the best examples for the application of BCD numbers. Other examples of BCD system are electronic counters, digital voltmeter and digital clocks.

DECIMAL	BCD
0	0000

1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
10	0001 000
11	0001 0001

Computer Memory

Primary Memory (ROM and it's type – PROM, EPROM,EEPROM, RAM)

Primary storage, also known as main storage or memory, is the main area in a computer in which data is stored for quick access by the computer's processor. On today's smaller computers, especially personal computers and workstations, the term random access memory (RAM) - or just memory - is used instead of primary or main storage.

RAM typically contains portions of the operating system as well as applications and data in active or frequent use. The part of the operating system stored in the primary memory is called the kernel. It is the first part that loads when the computer is turned on, and it stays there until the computer is turned off. Parts of the programs that are used when the computer is on, like text or a spreadsheet, also stay in the RAM.

Primary Memory Capacity

- In deciding how much memory you need, follow a simple rule: the more, the better. Larger primary memory allows the computer to perform operations faster. RAM capacity is measured in gigabytes. Most computers, as of 2010, come with a 2 gig RAM, which can be easily expanded to 4 gigs.

Read Only Memory (**ROM**) is an example of nonvolatile memory. ROM is a class of storage medium used in computers and other electronic devices. Read Only Memory (ROM), also known as firmware, is an integrated circuit programmed with specific data when it is manufactured. The instructions for starting the computer are housed on Read only memory chip. A ROM chip is also non volatile so data stored in it is not lost when power is turned off. ROM is a semiconductor memory that is capable of operating at electronics speed.

ROM Types:

PROM : Short for programmable read-only memory, a memory chip on which data can be written only once. Once a program has been written onto a PROM, it remains there forever. Unlike RAM, PROMs retain their contents when the computer is turned off. The difference between a PROM and a ROM (read-only memory) is that a PROM is manufactured as blank memory, whereas a ROM is programmed during the manufacturing process. To write data onto a PROM chip, you need a special device called a PROM programmer or PROM burner. The process of programming a PROM is sometimes called burning the PROM.

EPROM : Acronym for erasable programmable read-only memory, and pronounced ee-prom, EPROM is a special type of memory that retains its contents until it is exposed to ultraviolet light. The ultraviolet light clears its contents, making it possible to reprogram the memory. To write to and erase an EPROM, you need a special device called a PROM programmer or PROM burner.

EEPROM : Short form of electrically erasable programmable read-only memory. EEPROM is a special type of PROM that can be erased by exposing it to an electrical charge. Like other types of PROM, EEPROM retains its contents even when the power is turned off. Also like other types of ROM, EEPROM is not as fast as RAM

Secondary memory-SASD, DASD

The **secondary memory** is employed for bulk storage of programs, data and other information. It has much larger capacity than the main memory. It stores system software, assemblers, compilers, useful package, large data files, etc. The secondary memory is non-volatile in nature.

The magnetic memory retains the information stored in it. The magnetic memories such as hard disks and floppy disk are the most common secondary memories used in computer. Back up memory use to store the copy of the important programs such as operating system, compilers etc. Floppy disks and magnetic tape can be employed as back up storage. These programs are generally available in the secondary memory but they are copies are also kept in the back up memory so that they can be loaded into the secondary memory in case the original program stored in the secondary memory are lost accidentally or due to any other reason.

SASD: SEQUENTIAL ACCESS STORAGE DEVICESAND MEDIA

sequential access memory is a class of data storage devices that read their data in sequence.

Example**:-** magnetic tape

DIRECT ACCESS STORAGE DEVICES ANDMEDIA (DASD)

Direct access storage device, or **DASD** is any secondary storage device which has relatively low access time relative to its capacity.Example:

§ magnetic floppy disks

§ magnetic hard disks

§ optical discs (CD-ROM/DVD)

§ Flash Memory Cards

Secondary storage devices

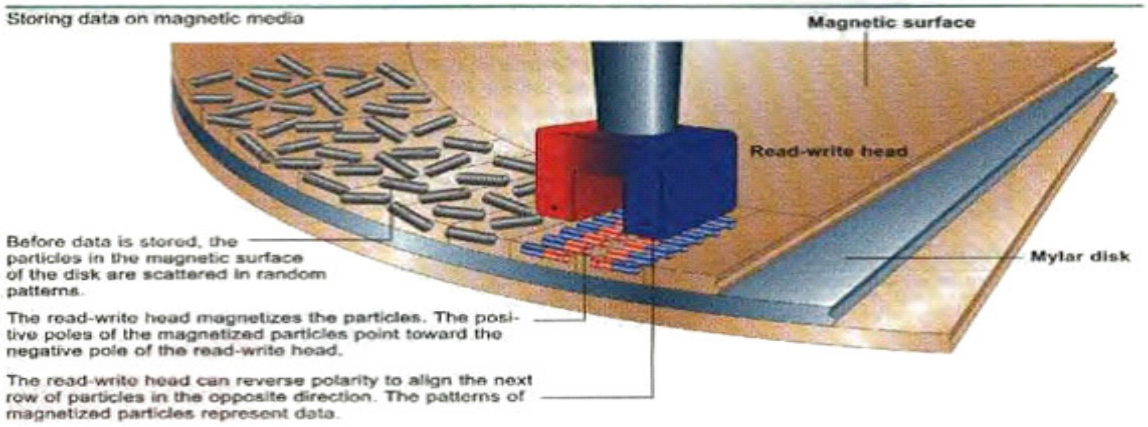
Secondary storage devices cannot be classified under input or output devices. These devices are used to store data for later use.
Secondary storage devices can be categorised in 3 main groups:
1. Magnetic
2. Optical
3. Electronic

1. **Magnetic Storage**

Magnetic storage devices store the data on a magnetic layer. Examples of such devices are:

a. Floppy Disk
b. Hard Disk
c. Magnetic Tape

The technology used is very simple to understand. The magnetic layer is magnetised when there is a binary 1 and left un-magnetised when there is a binary 0.

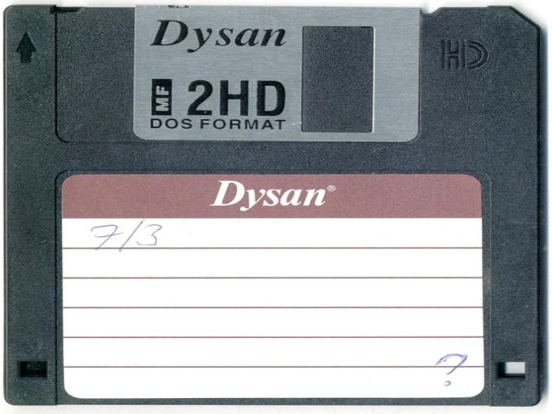


DEVI C
ESNDARY STORAGE

Floppy Disk

Floppy disks store small amount of data typically not larger than 1.44MB. These small diskettes are used mainly to transport small files from one computer to another. The construction of the floppy is basically a plastic disk coated with magnetic material, enclosed in a PVC jacket to protect the magnetic material.

In order to write data to a floppy this must be first formatted. This procedure involves the creation of sectors and tracks. On the diskette, data is recorded on closed concentric rings (tracks). Each track is divided into sectors. These are invisible wedged-shaped sections used by the computer for storage reference purposes.

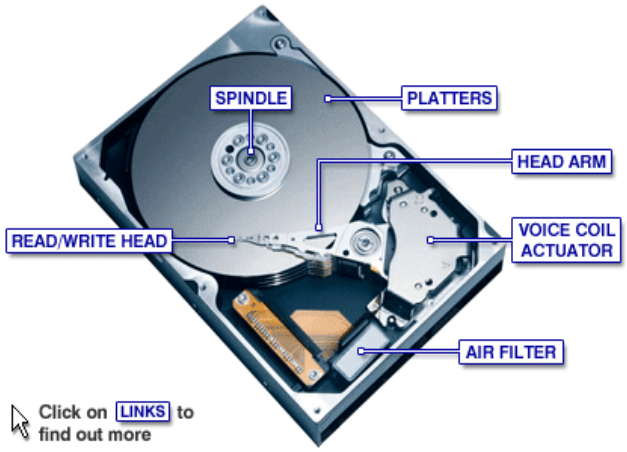


NDARY STOR

Hard Disks

As the name implies, hard disks are built with metal or glass platters covered with a substance that allows data to be held in the form of magnetised spots. Today hard disks are composed of

more than one disk (also named as platter), which are stacked on each other. This technology gives the possibility of having large storage capacity. Typically computer systems today have an 80GB (or more) hard disk capacity.



Magnetic Tape

This type of storage media is used in large computers where large amounts of data are stored. The tape is made up of a thin plastic tape having horizontal running tracks and vertical frames. The data to be stored is divided into:

a. The Header – Including file name and number of blocks
b. Block – of data separated by inter block gaps
c. Trailer – containing number of blocks, to be used as cross reference with the header

Examples – Reel to Reel tape drive and Cassette Tape drive

Magnetic tape is ideally suited to store a list of records whose sequential order is sorted for later processing. Payroll systems and record updating is ideal for tape. A sorted tape can be used to issue pay cheques, update a list of customers etc.

Optical Storage

- **The CD-ROM Disk**

CD-ROM stands for Compact Disk – Read Only Memory. This is an optical disk where information is stored at the time of manufacturing.

Examples of CD-ROM are Music CD's and software that one buys from computer shops such as games. Optical disks write data with a high power laser beam, which records data by burning tiny pits onto the surface of the disk. In order to read data, a low power laser beam reads data by reflecting smooth areas, which are interpreted as 1 bits, and not reflecting pitted areas which are interpreted as 0 bits.

The typical storage capacity of a CD-ROM is 650Mb, which makes them a very versatile data storage media when it comes to store data for short and long term periods of time.

- **CD-R**

CR-R which stands for Compact Disk- Recordable is a CD format that allows users to use a CD-R drive to write data, only once, onto a specially manufactured disk that can then be read by a standard CD-Rom drive.

- **CD-RW**

CD-RW stands for Compact Disk – Rewritable. This type of media allows users to copy and erase data. Thus, this disk can be used over and over again. DVD-Rom
DVD-Rom stands for Digital Video Disk or Digital Versatile Disk. This optical disk looks like a normal CD but in fact, it can store much more data than a CD

Electronic Storage
Electronic memory is the latest technology in secondary storage media. This is also referred to as **Flash memory.**
Flash memory is non-volatile solid-state computer memory storage that can be electrically erased and reprogrammed. It is a technology that is primarily used in memory cards and USB flash drives for general storage and transfer of data between computers and other digital products. It is a specific type of EEPROM (Electrically Erasable Programmable Read-Only Memory) This type of memory media has no mechanical moving parts and thus it offers some advantages when compared to the other type of media.
Advantages:
1. Very small
2. Does not consume a lot of energy in order to function
3. Noiseless
4. Limited heat emission
5. Portable

Disadvantages:
Limited amount of memory capacity when compared to the other type of media (Currently big advancements are being made and the maximum capacity of this media is 60GB). Because of their small size and big capacity, electronic memory is suitable to be used in portable devices. Today electronic memory is being used in Digital cameras, Mp3 Players, PDAs, Mobile phones and Laptops.

Data Access Methods
We have two types of data access methods:
a. **Serial Access**: - Cassette tape
- Tape Stream
b. **Random (Direct) Access**: - Floppy Disk
- Hard Disk
- Zip Disk
- CD-Rom
- DVD
- Flash Memory

Serial Access VS Random Access

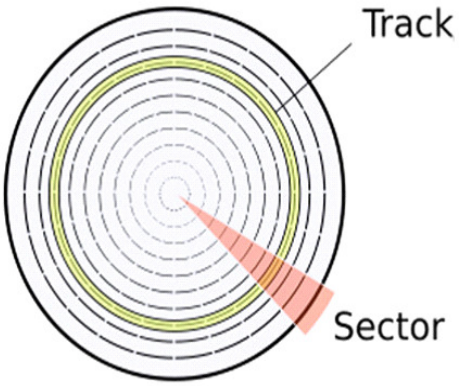
Imagine one has a thousand records stored on a secondary storage device and one needs to access the 543rd record. In serial access, one has to access the preceding 542 records before homing in on to the 543rd record.
On the other hand, in random (direct) access, one is given the possibility to go directly to the particular record required.

Serial Access Method

Serial access is where the items are read, one at a time, from the physical start of the file, in the order in which they are stored. Data is stored on tape in Binary Coded Decimal Format. A tape is made up of horizontally running tracks and vertical frames. When a file is stored on tape, the program is divided into:-

a. **a file header** (including the file name and number of blocks)
b**. number of blocks** (of a fixed number of records), each separated by inter-block gaps
c. a **trailer (**with the number of blocks for cross checking with the program header)

Direct Access Method – Hard disk, Zip disk and floppy disk



Formatting a floppy disk or hard disk means preparing the medium to receive data. Mostly when a new disk is bought, the surface are is free and not organized. In order to start writing data onto a disk, one must format it or prepare the area into chunks where the data is stored (like boxes in a room to organise our storage). Formatting prepares the disk into round concentric circles called TRACKS (like an athletics track). Tracks are then subdivided into smaller pieces to form SECTORS. Even though the outer sector is larger in area than the inner sectors, they will still hold the same amount of data.

Floppy disks rotate at about 360rpm. Data transfer is around 10k to 100k per second. They consist of 80 tracks and of 256 sectors per track. Floppy disks are lightweight, portable, relatively cheap, rewritable, and can store 1.44Mb of data (they can store more if special software is used for compression of data). Magnetic medium is brown in colour and soft. Disk allows random access. This is achieved by storing each new file at the start of a new sector. If for example, program one occupies sectors 0, 1, 2 and half of sector 3, file two will be stored in sector 4, leaving half sector 3 empty. All programs stored on disk are catalogued in a directory for easy searching. This is done by the operating system..

Zip disks are very popular for making backups both in the industry, offices and at home. They are magnetic disks but the media is made of high density magnetic material – meaning that the surface is more densely packed with magnetic particles. This allows for high density of data too. Zip disks use special compression software which packs large amounts of data into smaller sized data. Zip disks require special zip drives and special software compression. Typical zip disks are of 100Mb and 250Mb capacities and are small cartridge like in shape.

CD-Rom (medium)

CD-Rom stands for Compact Disk Read Only Memory – note that these disks are for reading and only store the information that the manufacturer records on it. The user cannot erase, change, or add to it, only read it. This type of optical disk is used primarily for storing huge amounts of data – such as government statistics, encyclopaedias, medical reference books, and dictionaries. CD-Rom's have an access time of about 160ms and a transfer rate of up to 800KB/s The CD-Rom is an optical storage disk – meaning it can be read by light sensing equipment. The
difference between this and magnetic type storage is the large amount of data it can hold. In fact, a CDRom can store up to 600Mb of data. Unlike the hard disk it is portable. A CD-Rom allows manufacturers to store video clips, still pictures and sound tracks apart from text to create what we now call multimedia.

CD-R – WORM (medium)

WORM stand for Write Once Read Many
This type of video disk allows for recording onto the next free tracks until the disk is full.
This type of medium is recorded only once.

CD-RW – Erasable Optical Disks (medium)

Erasable optical disks are an alternative to large capacity hard disks. They store 1,400 times as much data as diskettes. In contrast to CD-Rom's, erasable disks can be modified and erased. The removable disks provide convenience and security along with huge storage capacity. Optical disks can also be used to back up fixed hard disks. Many experts think that optical storage will become the most efficient inexpensive storage method. They offer 15,000 tpi (tracks per inch), compared to 96 tpi of floppy disks.

.

DVD – Digital Video Disk (medium)

The DVD medium is similar to the CD-Rom disk. There are no visible differences except that the tracks are more closely packed. Therefore there are more tracks on the DVD than on the CD-Rom disk. These new disks allow far more storage than the previous CD-Rom and today are available with a capacity of around 17GB. This large amount of storage enables the inclusion of more video clips; in fact, the most significant development sector for the DVD is in the film industry. One can already rent full feature films on DVD's. The advantage of DVD's over the original video cassette tapes is the high quality digital sound and video. These films are sometimes also provided with different language tracks pre-recorded so one can choose the language in which to listen to the sound track

Optical Disk Drives

A CD-Rom is similar to an audio CD but with the recording tracks much thinner and closer to each other.Therefore, the disk would be holding much more tracks for recording. At first there were two or three major standards, which meant that not all CD-Rom drives were compatible to read all CD Rom's. However agreement was reached and a standard was set up. The CD-Rom drive uses a laser light to read data on the tracks while the disk is spinning at a standard speed. These speeds have been increasing rapidly from 1x to 40x. A high-speed CD-Rom drive, allows for better access time and betterrefresh rates.

NAS and SAN

A storage area network (SAN) is a dedicated network that provides access to consolidated, block level data storage. SANs are primarily used to make storage devices, such as disk arrays, tape libraries, and optical jukeboxes, accessible to servers so that the devices appear like locally attached devices to the operating system. A SAN typically has its own network of storage devices that are generally not accessible through the local area network by other devices. The cost and complexity of SANs dropped in the early 2000s to levels allowing wider adoption across both enterprise and small to medium sized business environments.

NAS, Network Attached Storage was a solution to the problems of Direct Attached Storage (DAS). The solution was that servers would share a connection to the storage devices via a network connection through the LAN. This set up allows the server to be used loaded with software and applications instead of being split between two duties. With the old DAS setup the server would be split between application use and storage use. With NAS there is no longer a need for support of the traditional storage interface (SCSI) and now the server or client may access NAS storage with a network connection. The drawback to this is that there is no longer a high-speed connection between the CPU and storage units – they still must use the LAN to communicate and this creates bandwidth bottlenecks. In addition requests are processed using file access protocols and CPU cycles must be used to convert into block requests that a server may use to retrieve files and information. This has regulated NAS to be used as data backup more than anything else.

UNIT-II

Computer Software

Computer Software can be defined in the following ways:Computer software**, or** just software, is an any set of machine-readable instructions that directs a computer's processor to perform specific operations. It is a program that tells a computer what to do.**Software,** is a collection of programs that includes some input and outputs interface.

Types of software computer Software:

In most computer platforms, software can be grouped into two broad categories:
1.System Software
2.Application Software
1.System software :
System software is a program that manages and supports the computer resources and operations of **a** computer system while it executes various tasks such as processing data and information, controlling hardware components, and allowing users to use application software.
That is, systems software functions as a bridge between computer system hardware and the application software**.** System software is made up of many control programs, including the operating system, communications software and database manager.
Systems software consists of three kinds of programs.
1.The system management programs
2.system support programs and
3.system development programs

2. Application software :
Application software consists of programs that direct computers to perform specific information processing activities for end users. These programs are called application packages because they direct the processing required for a particular use, or application, which users want to accomplish. Thousands of application packages are available because there are thousands of different jobs end users want computers to do. Examples of application software are:
Word Processing Software, Database Software etc.

Difference between Compiler and Interpreter

No	Compiler	Interpreter

1	Compiler Takes **Entire** program as input	Interpreter Takes **Single** instruction as input .
2	Intermediate Object Code is Generated	No Intermediate Object Code is Generated
3	Conditional Control Statements are Executes **faster**	Conditional Control Statements are Executes **slower**
4	Memory Requirement **:** More **(Since Object Code is Generated)**	Memory Requirement **is** Less
5	Program need not be **compiled** every time	Every time higher level program is converted into lower level program
6	Errors **are displayed after** entire program **is checked**	Errors **are displayed for** every instruction **interpreted (if any)**
7	**Example** : C Compiler	**Example** : BASIC

Generation of language (Machine Level, Assembly, High Level, 4GL)

The **first generation languages**, or **1GL** are low-level languages that are **machine language**.The native language of the computer. In order for a program to run, it must be presented to the computer as binary-coded machine instructions that are specific to that CPU family. Although programmers are sometimes able to modify machine language in order to fix a running program , they do not create it. Machine language is created by software called "assemblers," "compilers" and "interpreters." These conversion programs turn the programmer's source code into machine language (machine code).

The **second generation languages**, or **2GL** are also low-level languages that consist of assembly languages.A programming language that is one step away from machine language. Each assembly language statement is translated into one machine instruction by the assembler. Programmers must be well versed in the computer's architecture, and, undocumented assembly language programs are difficult to maintain. It is hardware dependent; there is a different assembly language for each CPU series.

The **third generation languages**, or **3GL** are high-level languages such as C.

The **fourth generation languages**, or **4GL** are languages that consist of statements similar to statements in a human language. Fourth generation languages are commonly used in database programming and scripts. .A non-procedural programming language that requires less coding than lower-level languages. Command-line languages that come with operating systems and database management systems (DBMSs) are fourth-generation languages (4GLs), as are query

languages and report writers. Any language with English-like commands that does not require traditional input-process-output logic falls into this category.

Operating System Concept

What is Operating System

The Operating System (OS) is an integral part of your computer. It acts as an interpreter between the different application software programs and your computer. The OS can understand the data, information and commands used by different programs because those programs are written or coded specifically for that Operating System's interface, using it's standard rules, commands and protocols. It then translates the many commands and information into procedures the computer can understand. Computers are not of much use without an operating system.

So we can say that the Operating System have the Following Characteristics:-

1) Operating System is a Collection of Programs those are Responsible for the Execution of other Programs.

2) Operating System is that which Responsible is for Controlling all the Input and Output Devices those are connected to the System.

3) Operating System is that which Responsible is for Running all the Application Software's.

4) Operating System is that which Provides Scheduling to the Various Processes Means Allocates the Memory to various Process those Wants to Execute.

5) Operating System is that which provides the Communication between the user and the System.

6) Operating System is Stored into the BIOS Means in the Basic Input and Output System means when a user Starts his System then this will Read all the instructions those are Necessary for Executing the System Means for Running the Operating System, Operating System Must be Loaded into the Computer For this, this will use the Floppy or Hard Disks Which Stores the Operating System.

Functions of Operating System

- **Operating System as a Resource Manager**
- **Storage Management**
- **Process Management**
- **Memory Management**
- **Extended Machine**

Types of Operating System

1)**Serial Processing**: The Serial Processing Operating Systems are those which Performs all the instructions into a Sequence Manner or the Instructions those are given by the user will be executed by using the FIFO Manner means First in First Out. All the Instructions those are Entered First in the System will be Executed First and the Instructions those are Entered Later Will be Executed Later.

2)**Batch Processing**: The Batch Processing is same as the Serial Processing Technique. But in the Batch Processing Similar Types of jobs are Firstly Prepared and they are Stored on the Card. and that card will be Submit to the System for the Processing. The System then Perform all the Operations on the Instructions one by one. And a user can't be Able to specify any input. And Operating System wills increments his Program Counter for Executing the Next Instruction. The Main Problem is that the Jobs those are prepared for Execution must be the Same Type and if a job requires for any type of Input then this will not be Possible for the user. And Many Time will be wasted for Preparing the Batch. The Batch Contains the Jobs and all those jobs will be executed without the user Intervention. And Operating System will use the LOAD and RUN Operation. This will first LOAD the Job from the Card and after that he will execute the instructions. By using the RUN Command. The Speed of the Processing the Job will be Depend on the Jobs and the Results those are produced by the System in difference of Time which is used for giving or submit the Job and the Time which is used for Displaying the Results on the Screen.

3)**Multi-Programming**: As we know that in the Batch Processing System there are multiple jobs Execute by the System. The System first prepare a batch and after that he will Execute all the jobs those are Stored into the Batch. But the Main Problem is that if a process or job requires an Input and Output Operation, then it is not possible and second there will be the wastage of the Time when we are preparing the batch and the CPU will remain idle at that Time.

But With the help of Multi programming we can Execute Multiple Programs on the System at a Time and in the Multi-programming the CPU will never get idle, because with the help of Multi-Programming we can Execute Many Programs on the System and When we are Working with the Program then we can also Submit the Second or Another Program for Running and the CPU will then Execute the Second Program after the completion of the First Program. And in this we can also specify our Input means a user can also interact with the System. The Multi-programming Operating Systems never use any cards because the Process is entered on the Spot by the user. But the Operating System also uses the Process of Allocation and De-allocation of the Memory Means he will provide the Memory Space to all the Running and all the Waiting Processes. There must be the Proper Management of all the Running Jobs.

- **Real Time System**: There is also an Operating System which is known as Real Time Processing System. In this Response Time is already fixed. Means time to Display the Results after Possessing has fixed by the Processor or CPU. Real Time System is used at those Places in which we Requires higher and Timely Response. These Types of Systems are used in Reservation. So when we specify the Request, the CPU will perform at that Time.

There are two Types of Real Time System

- **Hard Real Time System**: In the Hard Real Time System, Time is fixed and we can't Change any Moments of the Time of Processing. Means CPU will Process the data as we Enters the Data.

- **Soft Real Time System**: In the Soft Real Time System, some Moments can be Change. Means after giving the Command to the CPU, CPU Performs the Operation after a Microsecond.

- **Distributed Operating System**. - Distributed Means Data is Stored and Processed on Multiple Locations. When a Data is stored on to the Multiple Computers, those are placed in Different Locations. Distributed means In the Network, Network Collections of Computers are connected with Each other. Then if we want to Take Some Data From other Computer, Then we uses the Distributed Processing System. And we can also Insert and Remove the Data from out Location to another Location. In this Data is shared between many users. And we can also Access all the Input and Output Devices are also accessed by Multiple Users.

- **Multiprocessing**: Generally a Computer has a Single Processor means a Computer have a just one CPU for Processing the instructions. But if we are Running multiple jobs, then this will decrease the Speed of CPU. For Increasing the Speed of Processing then we uses the Multiprocessing, in the Multi Processing there are two or More CPU in a Single Operating System if one CPU will fail, then other CPU is used for providing backup to the first CPU. With the help of Multi-processing, we can Execute Many Jobs at a Time. All the Operations are divided into the Number of CPU's. if first CPU Completed his Work before the Second CPU, then the Work of Second CPU will be divided into the First and Second.

- **Parallel operating systems** are used to interface multiple networked computers to complete tasks in parallel.

What is Booting? Type of Booting

Booting: When we start our Computer then there is an operation which is performed automatically by the Computer which is also called as Booting. In the Booting, System will check all the hardware's and Software's those are installed or Attached with the System and this will also load all the Files those are needed for running a system.

In the Booting Process all the Files those are Stored into the ROM Chip will also be Loaded for Running the System. In the Booting Process the System will read all the information from the Files those are Stored into the ROM Chip and the ROM chip will read all the instructions those are Stored into these Files. After the Booting of the System this will automatically display all the information on the System. The Instructions those are necessary to Start the System will be read at the Time of Booting.

There are two Types of Booting

1) Warm Booting: When the System Starts from the Starting or from initial State means when we Starts our System this is called as warm Booting. In the Warm Booting the System will be Started from its beginning State.

2) Cold Booting : The Cold Booting is that in which System Automatically Starts when we are Running the System, For Example due to Light Fluctuation the system will Automatically Restarts.

Start Up Procedure and Details of Basic System Configuration

Booting is a startup sequence that starts the operating system of a computer when it is turned on. A boot sequence is the initial set of operations that the computer performs when it is switched on. Every computer has a boot sequence. The average computer doesn't understand the boot sequence but is important to know for customizing and troubleshooting your computer.
When you hit the power button on your computer a whole lot of stuff happens. We call this the boot process. In the days when I first started using computers there was literally a "boot disk", a floppy (5.25″ not a 3.5″) disk that told the system where to go and what to do so that the operating system would start up. Since then the boot sequence has become somewhat more complicated. So let me take you thru the steps the computer takes to get started. An example for a Windows XP system.

- First is the POST, this stands for Power On Self Test, for the computer. This process tests memory as well as a number of other subsystems. You can usually monitor this as it runs each test. After that is complete the system will run POST for any device that has a BIOS (Basic Input-Output System). An AGP has its own BIOS, as do some network cards and various other devices.
- Once the POST is complete and the BIOS is sure that everything is working properly, the BIOS will then attempt to read the MBR (Master Boot Record). This is the first sector of the first hard drive (called the Master or HD0). When the MBR takes over it means that Windows is now in control.
- The MBR looks at the BOOT SECTOR (the first sector of the active partition). That is where NTLDR is located, NTLDR is the BOOT LOADER for Windows XP. NTLDR will allow memory addressing, initiate the file system, read the boot.ini and load the boot menu. NTLDR has to be in the root of the active partition as do NTDETECT.COM, BOOT.INI, BOOTSECT.DOS (for multi-OS booting) and NTBOOTDD.SYS (if you have SCSI adapters)
- Once XP is selected from the Boot Menu, NTLDR will run NTDETECT.COM, BOOT.INI and BOOTSECT.DOS to get the proper OS selected and loaded. The system starts in 16-bit real mode and then moves into 32-bit protected mode.
- NTLDR will then load NTOSKRNL.EXE and HAL.DLL. Effectively, these two files are windows XP. They must be located in %SystemRoot%System32.

- NTLDR reads the registry, chooses a hardware profile and authorizes device drivers, in that exact order.
- At this point NTOSKRNL.EXE takes over. It starts WINLOGON.EXE that in turn starts LSASS.EXE, this is the program that display the Logon screen so that you can logon.

An example for a Windows XP system:-

1.First is the POST, this stands for Power On Self Test, for the computer. This process tests memory as well as a number of other subsystems. You can usually monitor this as it runs each test. After that is complete the system will run POST for any device that has a BIOS (Basic Input-Output System). An AGP has its own BIOS, as do some network cards and various other devices.

- Once the POST is complete and the BIOS is sure that everything is working properly, the BIOS will then attempt to read the MBR (Master Boot Record). This is the first sector of the first hard drive (called the Master or HD0). When the MBR takes over it means that Windows is now in control.
- The MBR looks at the BOOT SECTOR (the first sector of the active partition). That is where NTLDR is located, NTLDR is the BOOT LOADER for Windows XP. NTLDR will allow memory addressing, initiate the file system, read the boot.ini and load the boot menu. NTLDR has to be in the root of the active partition as do NTDETECT.COM, BOOT.INI, BOOTSECT.DOS (for multi-OS booting) and NTBOOTDD.SYS (if you have SCSI adapters)
- Once XP is selected from the Boot Menu, NTLDR will run NTDETECT.COM, BOOT.INI and BOOTSECT.DOS to get the proper OS selected and loaded. The system starts in 16-bit real mode and then moves into 32-bit protected mode.
- NTLDR will then load NTOSKRNL.EXE and HAL.DLL. Effectively, these two files are windows XP. They must be located in %SystemRoot%System32.
- NTLDR reads the registry, chooses a hardware profile and authorizes device drivers, in that exact order.
- At this point NTOSKRNL.EXE takes over. It starts WINLOGON.EXE that in turn starts LSASS.EXE, this is the program that display the Logon screen so that you can logon.

Directory

Organization folder keeping all the files in your computer. **Directories** are found in hierarchical file system such as DOS, OS/2, Unix, etc. When referring to a directory, a user commonly indicates the name of the directory. Other common names for directories are the root directory, home directory, and current directory. In a GUI such as Microsoft Windows directories are referred to as folders.Below is an example of what a directory path would like in MS-DOS.

C:\Windows\System32>

In the above example, C: is the drive letter and the current directory is System32, which is a subdirectory of the Windows directory.

File

In a computers data is stored in files. When you run a program, MS-DOS processes the data stored in the file and passes it to the system. Depending on its size a hard drive may contain several thousands of files. However, the more files there are, the more difficult it is to manage them, this is when we need to store them in directories.

Data File Volumes

Definition

When extracting data, the system may create one or more sequential files to store the extracted data. If you specify a maximum file size smaller than the size that can accommodate the extracted data, then the system creates multiple physical files.

The physical files created for an extraction can be stored in the same directory, or they can be stored in different directories. A file volume defines the location of a single physical file.

Structure

If the data extracted fits into a single physical file, then the system creates a single sequential file and assigns it to a single volume. This volume is considered Volume 0 (zero). For example, if the user specifies DATA1998.TXT as the file name for the extract, the system creates DATA1998.TXT and assigns it to volume 0.

If the data extracted is too large to fit into a single sequential file, then the system creates several physical sequential files. The system assigns these files to volumes in the following manner: the first file created for the extract is assigned to Volume 0, the second file is assigned to Volume 1, and so on.

The following table reiterates the assignment of files to volumes:

Volume ID	Identifies the…
0 (zero)	First file created by the extract
1	Second file created by the extract
2	Next file created by the extract
…	

Volume Labels

Definition: A volume label is a unique name assigned to a drive, disc, or other media. In Windows, a volume label is *not* required but it's often useful to give a name to a drive to help

identify its use in the future. A drive's volume label can be changed at any time but is usually set during the formatting of the drive.

Certain restrictions apply when assigning volume labels depending on which file system is on the drive - NTFS or FAT:

Volume Label on NTFS Drives:

- Maximum of 32 characters
- No tabs

Volume Label on FAT Drives:

- Maximum of 11 characters
- No * ? . , ; : / \ | + = < > [ ]
- No tabs

Spaces are allowed in the volume label no matter which of the two file systems is used.

Drive name

To partition a hard drive means to divide the hard drive into parts and make those parts available to the operating system. It's also possible for an entire hard drive to have a single partition, which is actually the most common way to partition a hard drive .But we partitioned hard disk into more than one, and give the name to them like C drive ,D drive etc.

Introduction to GUI using Windows Operating System:

A graphical user interface is a type of user interface that allows users to interact with electronic devices with images rather than text commands. GUIs can be used in computers, hand-held devices such as MP3 players, portable media players or gaming devices, household appliances and office equipment. A GUI is a type of computer human interface on a computer. GUI usually have common characteristic such as windows, icons, menus, and push-buttons (WIMP). Collectively, WIMP are pictures that bring forth a certain action or an action space. The user issues commands via the GUI to computer applications. GUI usually have three major components**.** These three components are: ***a windowing system, an imaging model, and an application program interface (API)***.

The windowing system builds the windows, menus, and dialog boxes that appear on the screen.

The imaging model defines the fonts and graphics that appear on the screen. WIMP are products of both the windowing system and imaging model.

Finally, **the API** is the means in which the user specifies how and what windows and graphics appear on the screen.

Unlike a command line operating system or CUI like Unix or MS-DOS, GUI Operating Systems are much easier for end-users to learn and use because commands do not need to be known or memorized. Because of their ease of use, GUI Operating Systems have become the dominant operating system used by end-users today.

A few examples of a GUI Operating Systems

- Microsoft Windows
- Apple System 7 and Mac OS X

Assembler: A computer will not understand any program written in a language, other than its machine language. The programs written in other languages must be translated into the machine language. Such translation is performed with the help of software. A program which translates an assembly language program into a machine language program is called an assembler. If an assembler which runs on a computer and produces the machine codes for the same computer then it is called self assembler or resident assembler. If an assembler that runs on a computer and produces the machine codes for other computer then it is called Cross Assembler.

Assemblers are further divided into two types: One Pass Assembler and Two Pass Assembler. One pass assembler is the assembler which assigns the memory addresses to the variables and translates the source code into machine code in the first pass simultaneously. A Two Pass Assembler is the assembler which reads the source code twice. In the first pass, it reads all the variables and assigns them memory addresses. In the second pass, it reads the source code and translates the code into object code.

Compiler: It is a program which translates a high level language program into a machine language program. A compiler is more intelligent than an assembler. It checks all kinds of limits, ranges, errors etc. But its program run time is more and occupies a larger part of the memory. It has slow speed. Because a compiler goes through the entire program and then translates the entire program into machine codes. If a compiler runs on a computer and produces the machine codes for the same computer then it is known as a self compiler or resident compiler. On the other hand, if a compiler runs on a computer and produces the machine codes for other computer then it is known as a cross compiler.

Interpreter: An interpreter is a program which translates statements of a program into machine code. It translates only one statement of the program at a time. It reads only one statement of program, translates it and executes it. Then it reads the next statement of the program again translates it and executes it. In this way it proceeds further till all the statements are translated and executed. On the other hand, a compiler goes through the entire program and then translates the entire program into machine codes. A compiler is 5 to 25 times faster than an interpreter.

By the compiler, the machine codes are saved permanently for future reference. On the other hand, the machine codes produced by interpreter are not saved. An interpreter is a

small program as compared to compiler. It occupies less memory space, so it can be used in a smaller system which has limited memory space.

Linker: In high level languages, some built in header files or libraries are stored. These libraries are predefined and these contain basic functions which are essential for executing the program. These functions are linked to the libraries by a program called Linker. If linker does not find a library of a function then it informs to compiler and then compiler generates an error. The compiler automatically invokes the linker as the last step in compiling a program.

Not built in libraries, it also links the user defined functions to the user defined libraries. Usually a longer program is divided into smaller subprograms called modules. And these modules must be combined to execute the program. The process of combining the modules is done by the linker.

Loader: Loader is a program that loads machine codes of a program into the system memory. In Computing, a **loader** is the part of an Operating System that is responsible for loading programs. It is one of the essential stages in the process of starting a program. Because it places programs into memory and prepares them for execution. Loading a program involves reading the contents of executable file into memory. Once loading is complete, the operating system starts the program by passing control to the loaded program code. All operating systems that support program loading have loaders. In many operating systems the loader is permanently resident in memory.

Unit-III

A ***database*** is a set of data that has a regular structure and that is organized in such a way that a computer can easily find the desired information.

A ***database management system*** **(DBMS)** is software that has been created to allow the efficient use and management of databases, including ensuring that data is consistent and correct and facilitating it's updating.

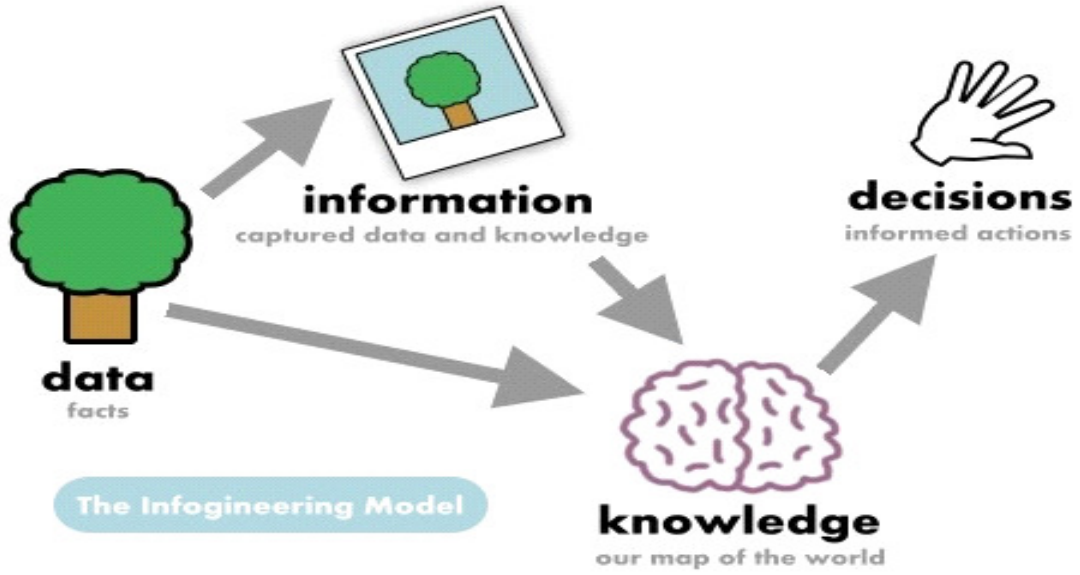
For small, single user databases, all functions are often managed by a single program; for larger and multi-user databases, multiple programs are usually involved and a client-server architecture is generally employed. The first DBMSs were developed in the 1960s in an attempt to make more effective use of the new direct access storage devices (i.e., hard disk drives the new direct access storage devices (i.e., hard disk drives) that were becoming available as supplements and eventual replacements for punched cards and magnetic tape. The word *database* is commonly used in a broad sense to refer not only just too structured data but also to the DBMS that is used with it.

Data and information

- **Data** is raw material for data processing. Data relates to fact, event and transactions.
- **Information** is data that has been processed in such a way as to be meaningful to the person who receives it. It is anything that is communicated.

  Data refers to the lowest abstract or a raw input which when processed or arranged makes meaningful output. It is the group or chunks which represent quantitative and qualitative attributes pertaining to variables. Information is usually the processed outcome of data. More specifically speaking, it is derived from data. Information is a concept and can be used in many domains.

  There are two sources that the brain uses to build this knowledge - information and data.
- **Knowledge:** Data and information organized and processed to convey understanding, experience, accumulated learning and expertise as they apply to a current problem or activity.



Summary:
1. Information is processed data whereas knowledge is information that is modeled to be useful.
2. You need information to be able to get knowledge.
3. Information deals with the way data is related while knowledge examines patterns within a given set of information.
4. To get knowledge you need some cognitive and analytical ability while for information you do not need cognitive ability.

Data	Information

Data is raw facts and figures	Information is a process form of data
Data refers to the lowest abstract or meaningless input.	Information arranged makes meaningful output
Data does not help in decision making.	Information helps in decision making

Why Use a DBMS?

- Data independence and efficient access.
- Reduced application development time.
- Data integrity and security.
- Uniform data administration.
- Concurrent access, recovery from crashes.

Why Study Databases??

- Shift from *computation* to *information*

At the "low end": scramble to web space (a mess!)

At the "high end": scientific applications

- Datasets increasing in diversity and volume.
- Digital libraries, interactive video, Human gnome project

Advantages and disadvantages of DBMS

Advantages

- Reduced data redundancy
- Reduced updating errors and increased consistency
- Greater data integrity and independence from applications programs
- Data can be shared
- Improved data security
- Reduced data entry, storage, and retrieval costs
- Facilitated development of new applications program.

Disadvantages

- Database systems are complex, difficult, and time-consuming to design
- Substantial hardware and software start-up costs
- Damage to database affects virtually all applications programs
- Extensive conversion costs in moving form a file-based system to a database system
- Initial training required for all programmers and users

Explanation Advantages

1. Redundancies and inconsistencies can be reduced

In conventional data systems, an organization often builds a collection of application programs often created by different programmers and requiring different components of the operational data of the organization. The data in conventional data systems is often not centralized. Some applications may require data to be combined from several systems. These several systems could well have data that is redundant as well as inconsistent (that is, different copies of the same data may have different values). Data inconsistencies are often encountered in everyday life. For example, we have all come across situations when a new address is communicated to an organization that we deal with (e.g. a bank, or Telecom, or a gas company), we find that some of the communications from that organization are received at the new address while others continue to be mailed to the old address. Combining all the data in a database would involve reduction in redundancy as well as inconsistency. It also is likely to reduce the costs for collection, storage and updating of data.

2. Better service to the Users

a DBMS is often used to provide better service to the users. In conventional systems, availability of information is often poor since it normally is difficult to obtain information that the existing systems were not designed for. Once several conventional systems are combined to form one centralized data base, the availability of information and its up-to-datedness is likely to improve since the data can now be shared and the DBMS makes it easy to respond to unforeseen information requests.

Centralizing the data in a database also often means that users can obtain new and combined information that would have been impossible to obtain otherwise. Also, use of a DBMS should allow users that do not know programming to interact with the data more easily.

The ability to quickly obtain new and combined information is becoming increasingly important in an environment where various levels of governments are requiring organizations to provide

more and more information about their activities. An organization running a conventional data processing system would require new programs to be written (or the information compiled manually) to meet every new demand.

3. Flexibility of the system is improved

Changes are often necessary to the contents of data stored in any system. These changes are more easily made in a database than in a conventional system in that these changes do not need to have any impact on application programs.

4. Cost of developing and maintaining systems is lower

As noted earlier, it is much easier to respond to unforeseen requests when the data is centralized in a database than when it is stored in conventional file systems. Although the initial cost of setting up of a database can be large, one normally expects the overall cost of setting up a database and developing and maintaining application programs to be lower than for similar service using conventional systems since the productivity of programmers can be substantially higher in using non-procedural languages that have been developed with modern DBMS than using procedural languages.

5. Standards can be enforced

since all access to the database must be through the DBMS, standards are easier to enforce. Standards may relate to the naming of the data, the format of the data, the structure of the data etc.

6. Security can be improved

in conventional systems; applications are developed in an ad hoc manner. Often different system of an organization would access different components of the operational data. In such an environment, enforcing security can be quite difficult.

Setting up of a database makes it easier to enforce security restrictions since the data is now centralized. It is easier to control that has access to what parts of the database. However, setting up a database can also make it easier for a determined person to breach security. We will discuss this in the next section.

7. Integrity can be improved

since the data of the organization using a database approach is centralized and would be used by a number of users at a time, it is essential to enforce integrity controls.

Integrity may be compromised in many ways. For example, someone may make a mistake in data input and the salary of a full-time employee may be input as Rs4,000 rather than Rs40,000. A student may be shown to have borrowed books but has no enrolment. Salary of a staff member in one department may be coming out of the budget of another department.

If a number of users are allowed to update the same data item at the same time, there is a possibility that the result of the updates is not quite what was intended. For example, in an airline DBMS we could have a situation where the number of bookings made is larger than the capacity of the aircraft that is to be used for the flight. Controls therefore must be introduced to prevent such errors to occur because of concurrent updating activities. However, since all data is stored only once, it is often easier to maintain integrity than in conventional systems.

8. Data model must be developed

Perhaps the most important advantage of setting up a database system is the requirement that an overall data model for the enterprise be built. In conventional systems, it is more likely that files will be designed as needs of particular applications demand. The overall view is often not considered. Building an overall view of the enterprise data, although often an expensive exercise is usually very cost-effective in the long term.

Disadvantages

1. Cost of Hardware & Software

A processor with high speed of data processing and memory of large size is required to run the DBMS software. It means that you have to upgrade the hardware used for file-based system. Similarly, DBMS software is also very costly.

2. Cost of Data Conversion

When a computer file-based system is replaced with a database system, the data stored into data file must be converted to database file. It is very difficult and costly method to convert data of data files into database. You have to hire database and system designers along with application programmers. Alternatively, you have to take the services of some software house. So a lot of money has to be paid for developing software. •

3. Cost of Staff Trailing

Most DBMSs are often complex systems so the training for users to use the DBMS is required. Training is required at all levels, including programming, application development, and database administration. The organization has to be paid a lot of amount for the training of staff to run the

DBMS.

4. Appointing Technical Staff

The trained technical persons such as database administrator, application programmers, data entry operators etc. Are required to handle the DBMS. You have to pay handsome salaries to these persons. Therefore, the" system cost increases.

5. Database Damage

In most of the organizations, all data is integrated into a single database. If database is damaged due to electric failure or database is corrupted on the storage media, then your valuable data may be lost forever.

FILE PROCESSING SYSTEMS

- A file system is a method of storing and organizing computer files and the data they contain to make it easy to find and access them. File systems may use a storage device such as hard disk or CD-ROM and involve maintaining the physical location of the files. File processing system store data in separate computer files. More formally, a file system is a special-purpose database for the storage, organization, manipulation, and retrieval of data

Flat file systems - In a flat file system, there are no subdirectories—everything is stored at the same (root) level on the media, be it a hard disk, floppy disk, etc. While simple, this system rapidly becomes inefficient as the number of files grows, and makes it difficult for users to organize data into related groups.

Characteristics of File processing systems

- It is a group of files storing data of an organization.
- Each file is independent from one another
- Each file is called flat file
- Each file contained and processed information for one specific function, such as accounting or inventory.
- files are designed by using programs written in programming languages such as C, C++

Disadvantages of File Processing Systems include:

1. **Program-Data Dependence**. File descriptions are stored within each application program that accesses a given file.

2**. Duplication of Data**. Applications are developed independently in file processing systems leading to unplanned duplicate files. Duplication is wasteful as it requires additional storage space and changes in one file must be made manually in all files. This also results in loss of data

integrity. It is also possible that the same data item may have different names in different files, or the same name may be used for different data items in different files.

3**. Limited data sharing.** Each application has its own private files with little opportunity to share data outside their own applications. A requested report may require data from several incompatible files in separate systems.

4. **Lengthy Development Times.** There is little opportunity to leverage previous development efforts. Each new application requires the developer to start from scratch by designing new file formats and descriptions.

5. **Excessive Program Maintenance**. The preceding factors create a heavy program maintenance load.

6. **Integrity Problem**. The problem of integrity is the problem of ensuring that the data in the database is accentuate.

7. **Inconsistence data**

8. **Security**

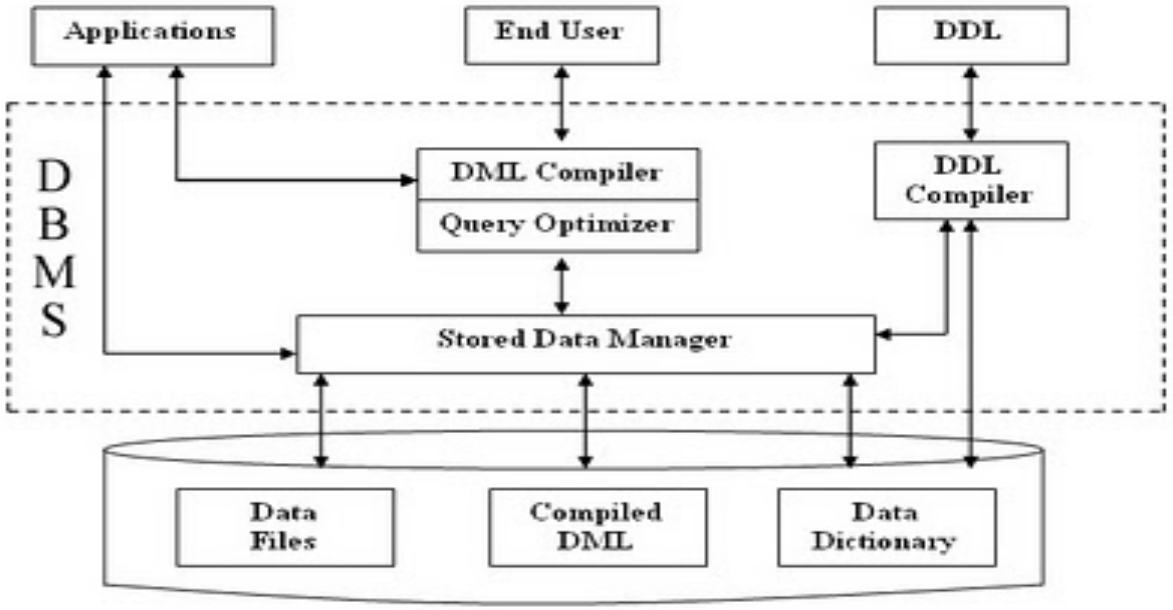
Comparison of file management system with database management system

File management system with **In C,++ or COBOL**	**Database management system** **with MS Access, SQL & Oracle**
Small system	Large system
Relatively cheap	Relatively expensive

Few files	Many files
Files are in the form of files	Files are in the form of table
Simplex structure	Complex structure
Redundant data	Reduce redundancy
Changes of inconsistency	Consistent data
Isolated data	Data can be shared
Little preliminary design	Vast preliminary design
Integrity left to application programmer	Rigorous inbuilt integrity checking
No security	Rigorous security
Simple, primitive backup	Complex & sophisticated backup
No recovery mode	Fully furnished Recovery mode
Often single user	Multiple user

STRUCTURE OF DBMS

DBMS (Database Management System) acts as an interface between the user and the database. The user requests the DBMS to perform various operations (insert, delete, update and retrieval) on the database. The components of DBMS perform these requested operations on the database and provide necessary data to the users. The various components of DBMS are shown below: -



1. DDL Compiler - Data Description Language compiler processes schema definitions specified in the DDL. It includes metadata information such as the name of the files, data items, storage details of each file, mapping information and constraints etc.

2. DML Compiler and Query optimizer - The DML commands such as insert, update, delete, retrieve from the application program are sent to the DML compiler for compilation into object code for database access. The object code is then optimized in the best way to execute a query by the query optimizer and then send to the data manager.

3. Data Manager - The Data Manager is the central software component of the DBMS also knows as Database Control System.

The Main Functions Of Data Manager Are: –

- Convert operations in user's Queries coming from the application programs or combination of DML Compiler and Query optimizer which is known as Query Processor from user's logical view to physical file system.
- Controls DBMS information access that is stored on disk.
- It also controls handling buffers in main memory.
- It also enforces constraints to maintain consistency and integrity of the data.
- It also synchronizes the simultaneous operations performed by the concurrent users.
- It also controls the backup and recovery operations.

4. Data Dictionary - Data Dictionary is a repository of description of data in the database. It contains information about

- Data - names of the tables, names of attributes of each table, length of attributes, and number of rows in each table.
- Relationships between database transactions and data items referenced by them which is useful in determining which transactions are affected when certain data definitions are changed.
- Constraints on data i.e. range of values permitted.
- Detailed information on physical database design such as storage structure, access paths, files and record sizes.
- Access Authorization - is the Description of database users their responsibilities and their access rights.
- Usage statistics such as frequency of query and transactions..Data dictionary is used to actually control the data integrity, database operation and accuracy. It may be used as an important part of the DBMS.

Importance of Data Dictionary -

Data Dictionary is necessary in the databases due to following reasons:

• It improves the control of DBA over the information system and user's understanding of use of the system.

• It helps in documentation the database design process by storing documentation of the result of every design phase and design decisions.

• It helps in searching the views on the database definitions of those views.

• It provides great assistance in producing a report of which data elements (i.e. data values) are used in all the programs.

• It promotes data independence i.e. by addition or modifications of structures in the database application program are not effected.

5. Data Files - It contains the data portion of the database.

6. Compiled DML - The DML complier converts the high level Queries into low level file access commands known as compiled DML.

7. End Users - end-user as the person who uses an Application.

Data models:A data model is a collection of concepts for describing data, its relationships, and its constraints Provides a clearer and more accurate description and representation of data Standard platform that enables database designers and end-users to communicate Come in three varieties

- Object-based models (conceptual schema)

- Record-based models (external schema)
- Physical data models (internal schema)

Importance of Data models

- Data models representations, usually graphical, of complex real-world data structures
- Facilitate interaction among the designer, the applications programmer and the end user
- End-users have different views and needs for data
- Data model organizes data for various users

Object based data model

Object based data models use concepts such as entities, attributes and relationship. An entity is a distinct object (a person, place ,concept, event) in the organization that is to represented in the databse.

Physical Data Model

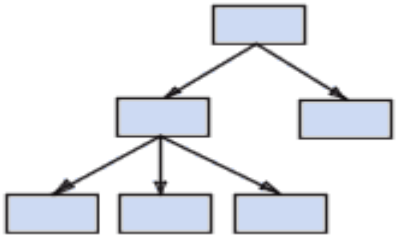
Physical data model describe how data is stored in the computer representing information such as record ordering and access paths.

What is a record-based data model?

A record based data model is used to specify the overall logical structure of the database.
In this model the database consists of a no. of fixed formats of different types.
Each record type defines a fixed no. of fields having a fixed length.
There are 3 principle types of record based data model. They are:

- **Hierarchical data model.**
- **Network data model.**
- **Relational data model.**

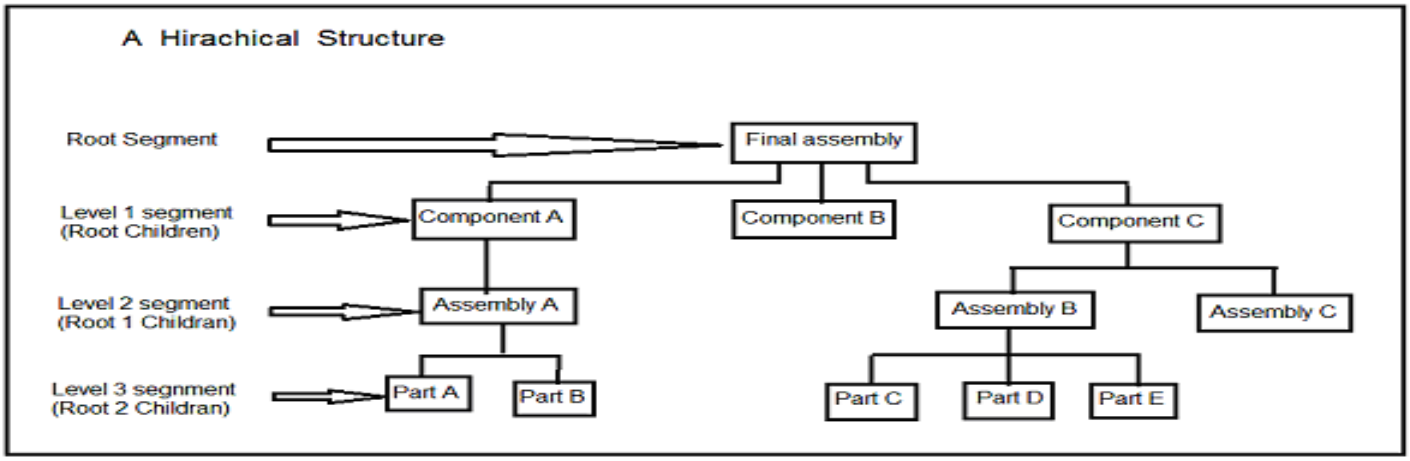
Hierarchical Model: Data is sorted hierarchically, using a downward tree. This model uses pointers to navigate between stored data. It was the first DBMS model.



- The hierarchical data model organizes data in a tree structure. There is a hierarchy of parent and child data segments.
- This structure implies that a record can have repeating information, generally in the child data segments.
- Data in a series of records, which have a set of field values attached to it.

- It collects all the instances of a specific record together as a record type.
  These record types are the equivalent of tables in the relational model, and with the individual records being the equivalent of rows.
- To create links between these record types, the hierarchical model uses Parent Child Relationships. These are a 1:N mapping between record types.
- This is done by using trees, like set theory used in the relational model, "borrowed" from maths.
- For example, an organization might store information about an employee, such as name, employee number, department, salary.
  The organization might also store information about an employee's children, such as name and date of birth.

The employee and children data forms a hierarchy, where the employee data represents the parent segment and the children data represents the child segment.If an employee has three children, then there would be three child segments associated with one employee segment.In a hierarchical database the parent-child relationship is one to many.This restricts a child segment to having only one parent segment.



Logical structure represented as an upside-down "tree"

- Hierarchical structure contains levels or segments
- Depicts a set of one-to-many (1: M) relationships Between a parent and it's children segments
- Each parent can have many children
- Each child has only one parent

Advantages & Disadvantages of Hierarchical model

Advantages

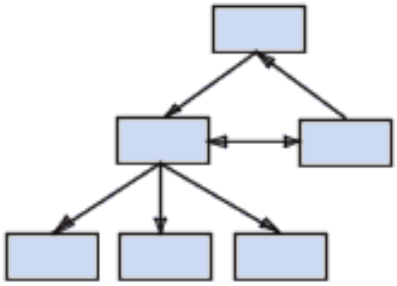
- Many features form the foundation for current data models
- Generated a large installed base of programmers
- Who developed solid business applications

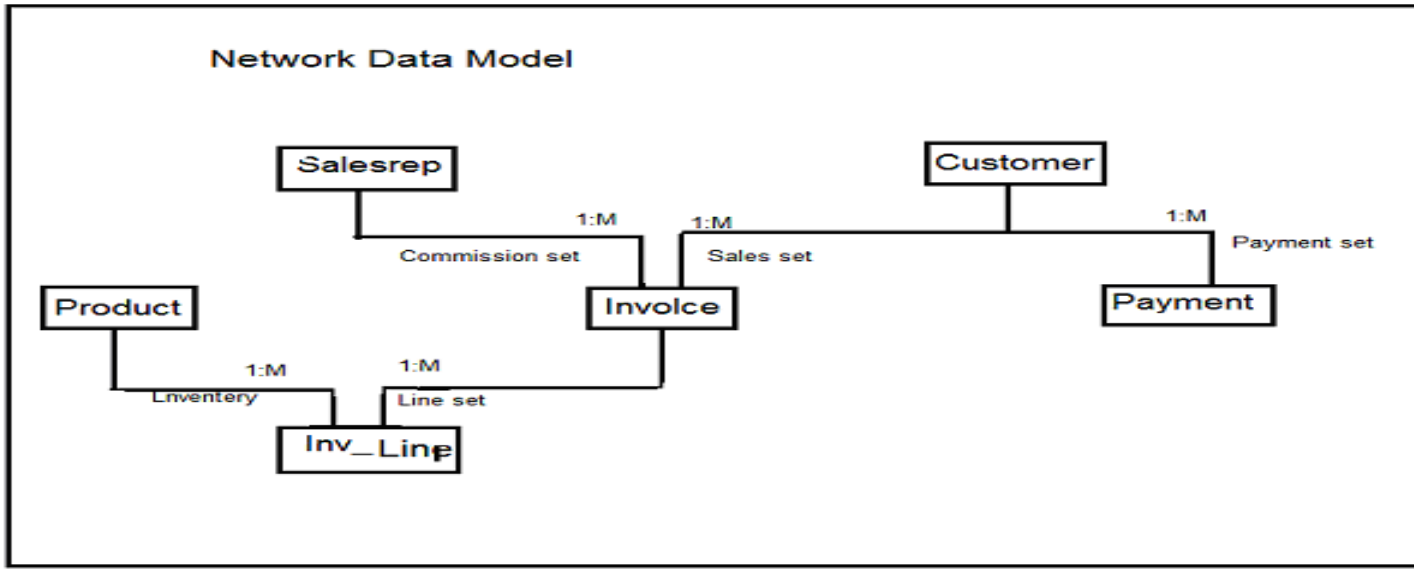
Disadvantages

- Complex to implement
- Difficult to manage
- Lacks structural independence
- Implementation limitations
- Lack of standards (Company vs. Industry or Open)

2. Network Data Model

Like the hierarchical model, this model uses pointers toward stored data. However, it does not necessarily use a downward tree structure.





- Resembles hierarchical model
- Difference child can have multiple parents

• Collection of records in 1: M relationships
• Set – Relationship of at least two record types
• Owner – Equivalent to the hierarchical model's parent
• Member – Equivalent to the hierarchical model's child

ADVANTAGES

- Provide very- Simplicity The network model is conceptually,efficient "High-speed" retrieval
- Ability to handle more relationship types Thesimple and easy to design.
- Easenetwork model can handle the one-to-many and many-to- many relationships. of data access In the network database terminology, a relationship is a set.
- Each set comprises of two types of records:- an owner record and a member record, In a network model an application can access an owner record and all the **Data Integrity** In a network model, no member canmember records within a set. exist without an owner. A user must therefore first define the owner record and then the member record. This ensures the integrity.
- Data Independence The network model draws a clear line of demarcation between programs and the complex physical storage details. The application programs work independently of the data. Any changes made in the data characteristics do not affect the application program.

DISADVANTAGES

System complexity

In a network model, data are accessed one record at a time. This males it essential for the database designers, administrators, and programmers to be familiar with the internal data structures to gain access to the data. Therefore, a user friendly database management system cannot be created using the network model

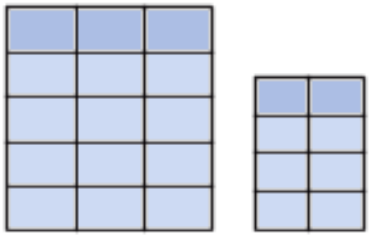
Lack of Structural independence.

Making structural modifications to the database is very difficult in the network database model as the data access method is navigational. Any changes made to the database structure require the application programs to be modified before they can access data. Though the network model achieves data independence, it still fails to achieve structural independence.

3. RELATIONAL DATA MODEL

The relational model (**RDBMS**, *Relational database management system*): The data is stored in two-dimensional tables (rows and columns). The data is manipulated based on the relational

theory of mathematics.

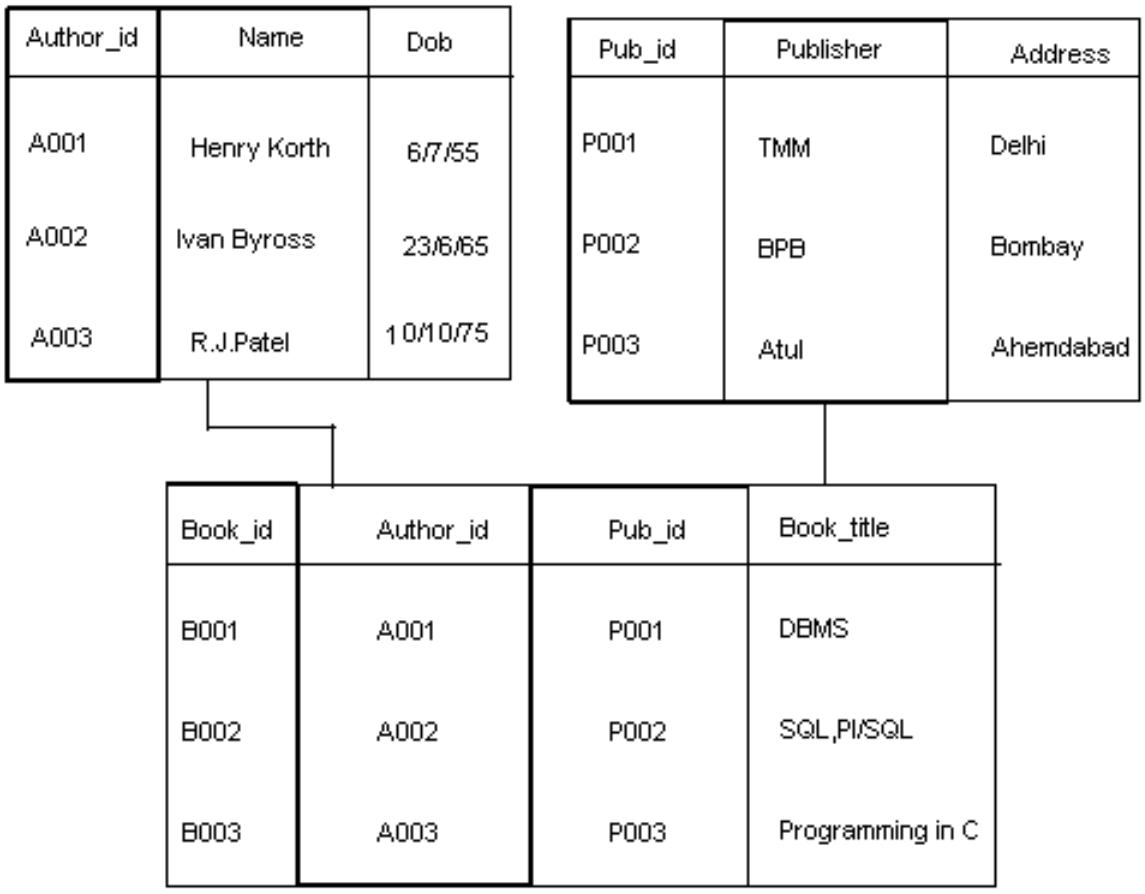


A database based on the relational model developed by E.F. Codd. A relational database allows the definition of data structures, storage and retrieval operations and integrity constraints. In such a database the data and relations between them are organised in tables. A table is a collection of records and each record in a table contains the same fields.

Properties of Relational Tables:

- Values Are Atomic
- Each Row is Unique
- Column Values Are of the Same Kind
- The Sequence of Columns is Insignificant
- The Sequence of Rows is Insignificant
- Each Column Has a Unique Name

Certain fields may be designated as keys, which means that searches for specific values of that field will use indexing to speed them up. Where fields in two different tables take values from the same set, a join operation can be performed to select related records in the two tables by matching values in those fields. Often, but not always, the fields will have the same name in both tables. For example, an "orders" table might contain (customer-ID, product-code) pairs and a "products" table might contain (product-code, price) pairs so to calculate a given customer's bill you would sum the prices of all products ordered by that customer by joining on the product-code fields of the two tables. This can be extended to joining multiple tables on multiple fields. Because these relationships are only specified at retreival time, relational databases are classed as dynamic database management system. The RELATIONAL database model is based on the Relational Algebra.



Author_id	Name	Dob
A001	Henry Korth	6/7/55
A002	Ivan Byross	23/6/65
A003	R.J.Patel	10/10/75

Pub_id	Publisher	Address
P001	TMM	Delhi
P002	BPB	Bombay
P003	Atul	Ahemdabad

Book_id	Author_id	Pub_id	Book_title
B001	A001	P001	DBMS
B002	A002	P002	SQL,PI/SQL
B003	A003	P003	Programming in C

Comparison of Data Model

The following table gives a comparative study of the three traditional data models.

S.NO	HIERARCHICAL DATA MODEL	NETWORK DATA MODEL	RELATIONAL DATA MODEL
1.	Relationship between records is of the parent child type	Relationship between records is expressed in the form of pointers or links	Relationship between records is represented by a relation that contains a key for each record involved in the relationship
2.	Many to many relationship cannot be expressed in this model	Many to many relationship can also be implemented	Many to many relationship can be easily implemented
3.	It is a simple, straightforward and natural method of implementing record relationships	Record relationship implementation is very complex due to the use of pointers	Relationship implementation is very easy through the use of a key or composite key field(s)
4.	This type of model is useful only when there is some hierarchical character in the database.	Network model is useful for representing such records which have many to many relationships	Relational model is useful for representing most of the real world objects and relationships among them
5.	In order to represent links among records, pointers are used. Thus relations among records are physical.	In Network model also the record relations are physical.	Relational model does not maintain physical connection among records. Data is organized logically in the form of rows and columns, and stored in table
6.	Searching for a record is very difficult since one can retrieve a child only after going through its parent record.	Searching a record is easy since there are multiple access paths to data elements.	A unique indexed key field is used to search for a data element.

Database Languages

Data Definition Language or **Data Description Language** (**DDL**) is a computer language for defining data structures.

A DDL is a language used to define data structures within a database. It is typically considered to be a subset of SQL, the Structured Query Language, but can also refer to languages that define other types of data.

A Data Definition Language has a pre-defined syntax for describing data. For example, to build a new table using SQL syntax, the CREATE command is used, followed by parameters for the table name and column definitions. The DDL can also define the name of each column and the associated data type. Once a table is created, it can be modified using the ALTER command. If the table is no longer needed, the DROP command will delete the table

- **Data Manipulation** is:
    - **retrieval** of information from the database
    - **insertion** of new information into the database
    - **deletion** of information in the database
    - **modification** of information in the database
- A DML is a language which enables users to access and manipulate data. The goal is to provide efficient human interaction with the system.
- There are two types of DML:
    - **procedural**: the user specifies *what* data is needed and *how* to get it
    - **nonprocedural**: the user only specifies *what* data is needed
        - Easier for user
        - May not generate code as efficient as that produced by procedural languages

Data control language

- DATA CONTROL LANGUAGE is known as DCL.
- DCL Statement is used for securing the database.
- DCL Statement control access to database.

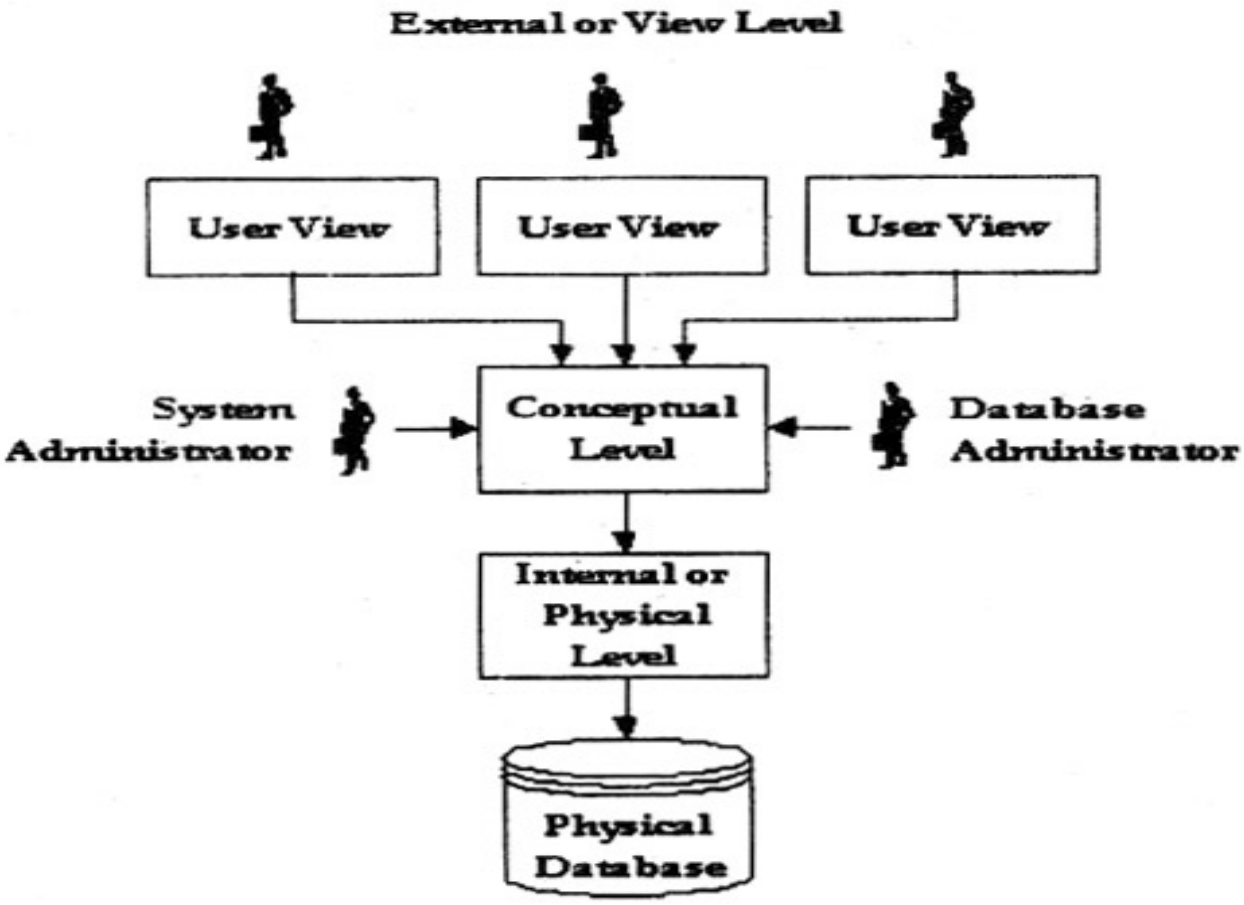
- As data is important part of whole database system we must take proper steps to check that no invalid user access the data and invalidate the information created by us. To kept such a kind of watch we must have to execute certain DCL statement.
- Two main DCL statement are Grant and Revoke
- GRANT to allow specified users to perform specified tasks.
- REVOKE to cancel previously granted or denied permissions.

THREE-LEVEL ARCHITECTURE

Three level architecture is Describing how storing data in data base management.

The three level database architecture allows a clear separation of the information meaning (conceptual view) from the external data representation and from the physical data structure layout. A database system that is able to separate the three different views of data is likely to be flexible and adaptable. This flexibility and adaptability is data independence that we have discussed earlier.

We now briefly discuss the three different views.



The external level is the view that the individual user of the database has. This view is often a restricted view of the database and the same database may provide a number of different views for different classes of users. In general, the end users and even the applications programmers are only interested in a subset of the database. For example, a department head may only be interested in the departmental finances and student enrolments but not the library information. The librarian would not be expected to have any interest in the information about academic staff. The payroll office would have no interest in student enrolments.

The conceptual view is the information model of the enterprise and contains the view of the whole enterprise without any concern for the physical implemenation. This view is normally more stable than the other two views. In a database, it may be desirable to change the internal view to improve performance while there has been no change in the conceptual view of the database. The conceptual view is the overall community view of the database and it includes all the information that is going to be represented in the database. The conceptual view is defined by the conceptual schema which includes definitions of each of the various types of data.

The internal view is the view about the actual physical storage of data. It tells us what data is stored in the database and how.

EXTERNAL LEVEL (highest level)

• The user's view of the database.

• Consists of a number of different external views of the DB.

• Describes part of the DB for particular group of users.

• Provides a powerful and flexible security mechanism by

hiding parts of the DB from certain users. The user is not

aware of the existence of any attributes that are missing

from the view.

• It permits users to access data in a way that is customized to their needs, so that the same data can be seen by different

users in different ways, at the same time.

CONCEPTUAL LEVEL

• The logical structure of the entire database as seen by DBA.

• What data is stored in the database.

• The relationships among the data.

• Complete view of the data requirements of the organization, independent of any storage consideration.

• Represents:

- entities, attributes, relations

- constraints on data

- semantic information on data

- security, integrity information

Supports each external view: any data available to a user must be contained in, or derivable from the conceptual level.

INTERNAL LEVEL (last level)

• Physical representation of the DB on the computer.

• How the data is stored in the database.

• Physical implementation of the DB to achieve optimal run–

time performance and storage space utilization.

- Storage space allocation for data and indexes

- Record description for storage

- Record placement

- Data compression, encryption

PHYSICAL LEVEL

Managed by the OS under the direction of the DBMS.

SCHEMAS, MAPPINGS, INSTANCES

DB schema: overall description of the DB.

Three different schemas according to the level of abstraction.

DBMS: mapping between schemas consistency of schemas

conceptual/internal mapping: to find the actual record (combinations) in physical storage that constitute a logical record in the conceptual schema.

external/conceptual mapping: map names in the user's view onto the relevant part of the conceptual schema.

DATA INDEPENDENCE

Data independence is a form of database management that keeps data separated from all programs that make use of the data. As a cornerstone for the idea of a DBMS or database management system, data independence ensures that the data cannot be redefined or reorganized by any of the programs that make use of the data. In this manner, the data remains accessible, but is also stable and cannot be corrupted by the applications using it.

There are two kinds:

Logical data independence

• The ability to modify the conceptual scheme without causing application programs to be rewritten.

• Immunity of external schemas to changes in the conceptual schema.

• Usually done when logical structure of database is altered

Physical data independence

• The ability to modify the internal scheme without having to change the conceptual or external schemas.

• Modifications at this level are usually to improve performance.

Define primary key foreign key and unique key.

Primary and foreign key

- Primary keys enforce entity integrity by uniquely
- identifying entity instances.
- Foreign keys enforce referential integrity by completing an association between two entities
- Primary key is unique key but foriegn key always refers to primary key

- Primary key is unique
- Primary key is not NULL and foreign key is NULL
- Foreign key reference as Primary key in another table
- Primary key will not allow "Null values" and "Duplicate values"
- Foreign key will allow "Null values" and "Duplicte values"and it refers to a primary key in anoter table.

Unique Key.

- Null values are accepted.
- More than one unique key will be there in a table.
- Non-Clustered index is created in unique key.
- Unique key constraint is used to prevent the duplication of key values within the rows of a table and allow null values.

Entity Relationship model

What is it?

The entity-relationship model (or ER model) is a way of graphically representing the logical relationships of entities (or objects) in order to create a database. The ER model was first proposed by Peter Pin-Shan Chen of Massachusetts Institute of Technology (MIT) in the 1970s.

The whole purpose of ER modelling is to create an accurate reflection of the real world in a database. The ER model doesn't actually give us a database description. It gives us an intermediate step from which it is easy to define a database. Let's look at an example. Suppose you are presented with the following situation and are told to create a database for it:

Every department within our company is in only one division. Each division has more than one department in it. We don't have an upper limit on the number of departments that a division can have. For example, the **New Business Development**---the one managed by Mackenzie---and **Higher Education** departments are both in the **Marketing** division.

This is a fairly clear description of a situation. Many things are left unsaid that we understand about the situation. For example: each division has a name, and that name is unique within the company. For now, though, let's focus on the description as it is given.

The first step is to figure out the items of interest in this situation. (In this document you will come across **Problems**. You should attempt to perform these before continuing the reading. Simply reading the problem and then reading the answer is not sufficient---you should attempt the problem yourself before you continue reading. Understanding these problems are integral to understanding the text. The answer to the problem appears in the text immediately after the problem.)

In ER modeling, the structure for a database is portrayed as a diagram, called an entity-relationship diagram (or ER diagram), that resembles the graphical breakdown of a sentence into its grammatical parts. **Entities are rendered** as points, polygons, circles, or ovals. Relationships are portrayed as lines connecting the points, polygons, circles, or ovals. Any ER diagram has an equivalent relational table, and any relational table has an equivalent ER diagram. ER diagramming is an invaluable aid to engineers in the design, optimization, and debugging of database programs.

Basic Constructs of E-R Modeling

Entity

One of the basic components of ER model is entity. An entity is any distinguishable object about which information is stored. These objects can be person, place, thing, event or a concept. Entities contain descriptive information. **Each entity is distinct.**

An entity may be physical or abstract. A person, a book, car, house, employee etc. are all physical entities whereas a company, job, or a university course, are abstract entities.

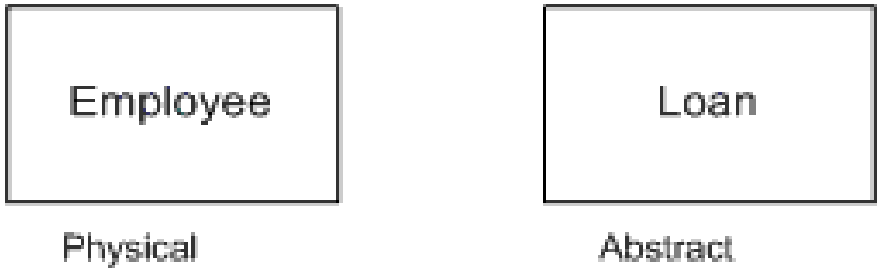


Fig - Physical and Abstract Entity

Another classification of entities can be independent or dependent (strong or weak) entity.

Entities are classified as independent or dependent (in some methodologies, the terms used are strong and weak, respectively). An independent entity is one, which does not rely on another entity for identification. A dependent entity is one that relies on another entity for identification. An independent entity exists on its own whereas dependent entity exists on the existence of some other entity. For example take an organization scenario. Here department is independent entity. Department manager is a dependent entity. It exists for existing depts. There won't be any department manager for which there is no dept.

Some entity types may not have any key attributes of their own. These are called weak entity types. Entities belonging to a weak entity type are identified by being related to specific entities from another entity type in combination with some of their attribute values. For example, take the license entity. It can't exist unless it is related to a person entity.

Attributes

After you identify an entity, then you describe it in real terms, or through its attributes. Attributes are basically properties of entity. We can use attributes for identifying and expressing entities.

For example, Dept entity can have DeptName, DeptId, and DeptManager as its attributes. A car entity can have modelno, brandname, and color as its attributes.

A particular instance of an attribute is a value. For example, "Bhaskar" is one value of the attribute Name. Employee number 8005 uniquely identifies an employee in a company.

The value of one or more attributes can uniquely identify an entity.

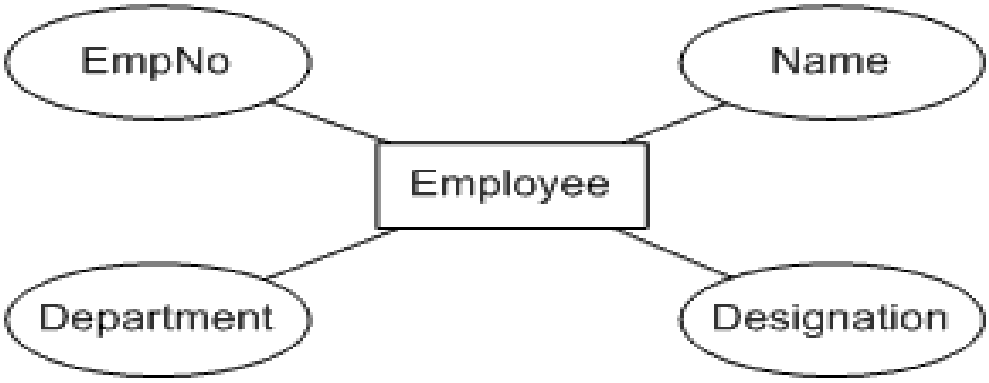


Fig - Entity and its attributes

In the above figure, employee is the entity. EmpNo, Name, Designation and Department are its attributes.

An entity set may have several attributes. Formally each entity can be described by set of <attribute, data value> pairs.

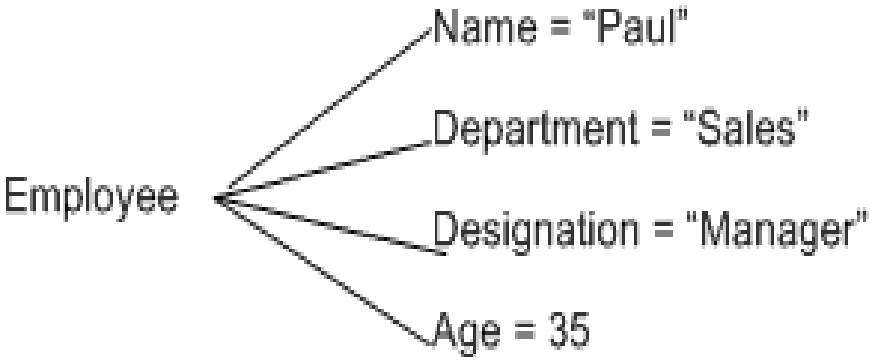


Fig - Employee entity and its attribute values

Derived attributes are attributes whose values are generated from other attributes using calculations, algorithms or procedures. For example, Account Balance is derived by subtracting Total Debit from Total Credit.

Relationship

After identification of entities and their attributes, the next stage in ER data modeling is to identify the relationships between these entities.

We can say a relationship is any association, linkage, or connection between the entities of interest to the business. Typically, a relationship is indicated by a verb connecting two or more entities. Suppose there are two entities of our library system, member and book, then the relationship between them can be "borrows".

Member borrows book

Each relationship has a name, degree and cardinality.

Relationships exhibit certain characteristics like degree, connectivity, and cardinality. Once the relationships are identified their degree and cardinality are also specified.

Degree: The degree of a relationship is the number of entities associated with the relationship. The n-ary relationship is the general form for degree n. Special cases are the binary, and ternary, where the degree is 2, and 3, respectively.

Binary relationships, the association between two entities are the most common type in the real world.

Fig - shows a binary relationship between member and book entities of library system

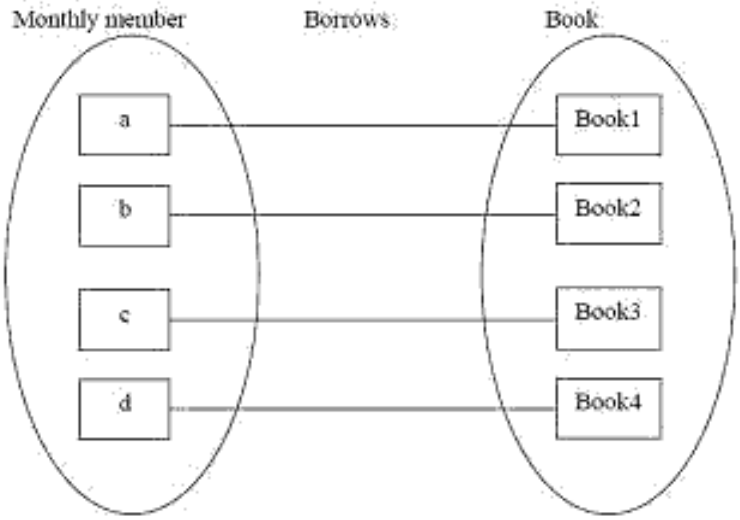


Fig.-Binary Relationship

A ternary relationship involves three entities and is used when a binary relationship is inadequate. Many modeling approaches recognize only binary relationships. Ternary or n-ary relationships are decomposed into two or more binary relationships.

Connectivity and Cardinality

By connectivity we mean how many instances of one entity are associated with how many instances of other entity in a relationship. Cardinality is used to specify such connectivity. The connectivity of a relationship describes the mapping of associated entity instances in the relationship. The values of connectivity are "one" or "many". **The cardinality of a relationship**

is the actual number of related occurrences for each of the two entities. The basic types of connectivity for relations are: one-to-one, one-to-many, and many-tomany.

A **one-to-one (1:1)** relationship is when at most one instance of an entity A is associated with one instance of entity B. For example, take the relationship between board members and offices, where each office is held by one member and no member may hold more than one office.

A **one-to-many (1:N)** relationship is when for one instance of entity A, there are zero, one, or many instances of entity B but for one instance of entity B, there is only one instance of entity A. An example of a 1:N relationships is

a department has many employees;
each employee is assigned to one department.

A **many-to-many (M:N)** relationship, sometimes called non-specific, is when for one instance of entity A, there are zero, one, or many instances of entity B and for one instance of entity B there are zero, one, or many instances of entity A. An example is employees may be assigned to no more than three projects at a time; every project has at least two employees assigned to it.

Here the cardinality of the relationship from employees to projects is three; from projects to employees, the cardinality is two. Therefore, this relationship can be classified as a many-to-many relationship.

If a relationship can have a cardinality of zero, it is an optional relationship. If it must have a cardinality of at least one, the relationship is mandatory. Optional relationships are typically indicated by the conditional tense

SQL is a standard language for accessing and manipulating databases.**What is SQL?**

- SQL stands for Structured Query Language
- SQL lets you access and manipulate databases
- SQL is an ANSI (American National Standards Institute) stand

What Can SQL do?

- SQL can execute queries against a database
- SQL can retrieve data from a database
- SQL can insert records in a database
- SQL can update records in a database
- SQL can delete records from a database
- SQL can create new databases
- SQL can create new tables in a database

- SQL can create stored procedures in a database
- SQL can create views in a database
- SQL can set permissions on tables, procedures, and views

SQL is a Standard - BUT....

Although SQL is an ANSI (American National Standards Institute) standard, there are many different versions of the SQL language.

However, to be compliant with the ANSI standard, they all support at least the major commands (such as SELECT, UPDATE, DELETE, INSERT, WHERE) in a similar manner.

Note: Most of the SQL database programs also have their own proprietary extensions in addition to the SQL standard!Using SQL in Your Web Site

To build a web site that shows some data from a database, you will need the following:

- An RDBMS database program (i.e. MS Access, SQL Server, MySQL)
- A server-side scripting language, like PHP or ASP
- SQL
- HTML / CSS

RDBMS- RDBMS stands for Relational Database Management System.

RDBMS is the basis for SQL, and for all modern database systems like MS SQL Server, IBM DB2, Oracle, MySQL, and Microsoft Access.

The data in RDBMS is stored in database objects called tables. A table is a collections of related data entries and it consists of columns and rows.

Database Tables

A database most often contains one or more tables. Each table is identified by a name (e.g. "Customers" or "Orders"). Tables contain records (rows) with data.

Below is an example of a table called "Persons":

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes

3	Pettersen	Kari	Storgt 20	Stavanger

The table above contains three records (one for each person) and five columns (P_Id, LastName, FirstName, Address, and City).

SQL Statements

Most of the actions you need to perform on a database are done with SQL statements. The following SQL statement will select all the records in the "Persons" table:

```
SELECT * FROM Persons
```

Keep in Mind That...

- SQL is not case sensitive

Semicolon after SQL Statements?

Some database systems require a semicolon at the end of each SQL statement.

Semicolon is the standard way to separate each SQL statement in database systems that allow more than one SQL statement to be executed in the same call to the server.

We are using MS Access and SQL Server 2000 and we do not have to put a semicolon after each SQL statement, but some database programs force you to use it.

SQL DML and DDL

SQL can be divided into two parts: The Data Manipulation Language (DML) and the Data Definition Language (DDL).

The query and update commands form the DML part of SQL:

- **SELECT** - extracts data from a database
- **UPDATE** - updates data in a database
- **DELETE** - deletes data from a database
- **INSERT INTO** - inserts new data into a database

The DDL part of SQL permits database tables to be created or deleted. It also define indexes (keys), specify links between tables, and impose constraints between tables. The most important DDL statements in SQL are:

- **CREATE DATABASE** - creates a new database

- **ALTER DATABASE** - modifies a database
- **CREATE TABLE** - creates a new table
- **ALTER TABLE** - modifies a table
- **DROP TABLE** - deletes a table
- **CREATE INDEX** - creates an index (search key)
- **DROP INDEX** - deletes an index

The SQL SELECT Statement

The SELECT statement is used to select data from a database. The result is stored in a result table, called the result-set.

SQL SELECT Syntax

```
SELECT column_name(s)
FROM table_name
```

And

```
SELECT * FROM table_name
```

Note: SQL is not case sensitive. SELECT is the same as select.

An SQL SELECT Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select the content of the columns named "LastName" and "FirstName" from the table above. We use the following SELECT statement:

```
SELECT LastName,FirstName FROM Persons
```

The result-set will look like this:

LastName	FirstName

Hansen	Ola
Svendson	Tove
Pettersen	Kari

SELECT * Example

Now we want to select all the columns from the "Persons" table.

We use the following SELECT statement:

```
SELECT * FROM Persons
```

Tip: The asterisk (*) is a quick way of selecting all columns!

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

The SQL SELECT DISTINCT Statement

In a table, some of the columns may contain duplicate values. This is not a problem, however, sometimes you will want to list only the different (distinct) values in a table.

The DISTINCT keyword can be used to return only distinct (different) values.

SQL SELECT DISTINCT Syntax

```
SELECT DISTINCT column_name(s)
FROM table_name
```

SELECT DISTINCT Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select only the distinct values from the column named "City" from the table above. We use the following SELECT statement:

```
SELECT DISTINCT City FROM Persons
```

The result-set will look like this:

City
Sandnes
Stavanger

SQL WHERE Clause

The WHERE clause is used to filter records.

The WHERE Clause

The WHERE clause is used to extract only those records that fulfill a specified criterion.

SQL WHERE Syntax

```
SELECT column_name(s)
FROM table_name
WHERE column_name operator value
```

WHERE Clause Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City

1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select only the persons living in the city "Sandnes" from the table above. We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE City='Sandnes'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes

Quotes Around Text Fields

SQL uses single quotes around text values (most database systems will also accept double quotes). Although, numeric values should not be enclosed in quotes.

For text values:

```
This is correct:

SELECT * FROM Persons WHERE FirstName='Tove'

This is wrong:

SELECT * FROM Persons WHERE FirstName=Tove
```

For numeric values:

```
This is correct:

SELECT * FROM Persons WHERE Year=1965

This is wrong:
```

```
SELECT * FROM Persons WHERE Year='1965'
```

Operators Allowed in the WHERE Clause

With the WHERE clause, the following operators can be used:

Operator	Description
=	Equal
<>	Not equal
>	Greater than
<	Less than
>=	Greater than or equal
<=	Less than or equal
BETWEEN	Between an inclusive range
LIKE	Search for a pattern
IN	If you know the exact value you want to return for at least one of the columns

Note: In some versions of SQL the <> operator may be written as !=

SQL AND & OR Operators

The AND & OR operators are used to filter records based on more than one condition.

The AND & OR Operators -The AND operator displays a record if both the first condition and the second condition is true. The OR operator displays a record if either the first condition or the second condition is true.

AND Operator Example -The "Persons" table:

P_Id	LastName	FirstName	Address	City

1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select only the persons with the first name equal to "Tove" AND the last name equal to "Svendson":

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE FirstName='Tove'
AND LastName='Svendson'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
2	Svendson	Tove	Borgvn 23	Sandnes

OR Operator Example

Now we want to select only the persons with the first name equal to "Tove" OR the first name equal to "Ola":

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE FirstName='Tove'
OR FirstName='Ola'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes

Combining AND & OR

You can also combine AND and OR (use parenthesis to form complex expressions).

Now we want to select only the persons with the last name equal to "Svendson" AND the first name equal to "Tove" OR to "Ola":

We use the following SELECT statement:

```
SELECT * FROM Persons WHERE
LastName='Svendson'
AND (FirstName='Tove' OR FirstName='Ola')
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
2	Svendson	Tove	Borgvn 23	Sandnes

The ORDER BY keyword is used to sort the result-set.

The ORDER BY Keyword

The ORDER BY keyword is used to sort the result-set by a specified column.

The ORDER BY keyword sort the records in ascending order by default.

If you want to sort the records in a descending order, you can use the DESC keyword.

SQL ORDER BY Syntax

```
SELECT column_name(s)
FROM table_name
ORDER BY column_name(s) ASC|DESC
```

ORDER BY Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes

3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Tom	Vingvn 23	Stavanger

Now we want to select all the persons from the table above, however, we want to sort the persons by their last name.

We use the following SELECT statement:

```
SELECT * FROM Persons
ORDER BY LastName
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
4	Nilsen	Tom	Vingvn 23	Stavanger
3	Pettersen	Kari	Storgt 20	Stavanger
2	Svendson	Tove	Borgvn 23	Sandnes

ORDER BY DESC Example

Now we want to select all the persons from the table above, however, we want to sort the persons descending by their last name.

We use the following SELECT statement:

```
SELECT * FROM Persons
ORDER BY LastName DESC
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

4	Nilsen	Tom	Vingvn 23	Stavanger
1	Hansen	Ola	Timoteivn 10	Sandnes

SQL INSERT INTO Statement

The INSERT INTO statement is used to insert new records in a table.

The INSERT INTO Statement

The INSERT INTO statement is used to insert a new row in a table.

SQL INSERT INTO Syntax

It is possible to write the INSERT INTO statement in two forms.

The first form doesn't specify the column names where the data will be inserted, only their values:

```
INSERT INTO table_name
VALUES (value1, value2, value3,...)
```

The second form specifies both the column names and the values to be inserted:

```
INSERT INTO table_name (column1, column2, column3,...)
VALUES (value1, value2, value3,...)
```

SQL INSERT INTO Example

We have the following "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to insert a new row in the "Persons" table.We use the following SQL statement:

```
INSERT INTO Persons
VALUES (4,'Nilsen', 'Johan', 'Bakken 2', 'Stavanger')
```

The "Persons" table will now look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Johan	Bakken 2	Stavanger

Insert Data Only in Specified Columns

It is also possible to only add data in specific columns.

The following SQL statement will add a new row, but only add data in the "P_Id", "LastName" and the "FirstName" columns:

```
INSERT INTO Persons (P_Id, LastName, FirstName)
VALUES (5, 'Tjessem', 'Jakob')
```

The "Persons" table will now look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Johan	Bakken 2	Stavanger
5	Tjessem	Jakob		

SQL UPDATE Statement

The UPDATE statement is used to update records in a table.

The UPDATE Statement

The UPDATE statement is used to update existing records in a table.

SQL UPDATE Syntax

```
UPDATE table_name
```

```
SET column1=value, column2=value2,...
WHERE some_column=some_value
```

Note: Notice the WHERE clause in the UPDATE syntax. The WHERE clause specifies which record or records that should be updated. If you omit the WHERE clause, all records will be updated!

SQL UPDATE Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Johan	Bakken 2	Stavanger
5	Tjessem	Jakob		

Now we want to update the person "Tjessem, Jakob" in the "Persons" table. We use the following SQL statement:

```
UPDATE Persons
SET Address='Nissestien 67', City='Sandnes'
WHERE LastName='Tjessem' AND FirstName='Jakob'
```

The "Persons" table will now look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Johan	Bakken 2	Stavanger
5	Tjessem	Jakob	Nissestien 67	Sandnes

SQL UPDATE Warning

Be careful when updating records. If we had omitted the WHERE clause in the example above, like this:

```
UPDATE Persons
SET Address='Nissestien 67', City='Sandnes'
```

The "Persons" table would have looked like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Nissestien 67	Sandnes
2	Svendson	Tove	Nissestien 67	Sandnes
3	Pettersen	Kari	Nissestien 67	Sandnes
4	Nilsen	Johan	Nissestien 67	Sandnes
5	Tjessem	Jakob	Nissestien 67	Sandnes

SQL DELETE Statement

The DELETE statement is used to delete records in a table.

The DELETE statement is used to delete rows in a table.

SQL DELETE Syntax

```
DELETE FROM table_name
WHERE some_column=some_value
```

Note: Notice the WHERE clause in the DELETE syntax. The WHERE clause specifies which record or records that should be deleted. If you omit the WHERE clause, all records will be deleted!

SQL DELETE Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes

2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Johan	Bakken 2	Stavanger
5	Tjessem	Jakob	Nissestien 67	Sandnes

Now we want to delete the person "Tjessem, Jakob" in the "Persons" table.

We use the following SQL statement:

```
DELETE FROM Persons
WHERE LastName='Tjessem' AND FirstName='Jakob'
```

The "Persons" table will now look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger
4	Nilsen	Johan	Bakken 2	Stavanger

Delete All Rows

It is possible to delete all rows in a table without deleting the table. This means that the table structure, attributes, and indexes will be intact:

```
DELETE FROM table_name

or

DELETE * FROM table_name
```

Note: Be very careful when deleting records. You cannot undo this statement!

SQL LIKE Operator

The LIKE operator is used in a WHERE clause to search for a specified pattern in a column. The LIKE operator is used to search for a specified pattern in a column.

SQL LIKE Syntax

```
SELECT column_name(s)
FROM table_name
WHERE column_name LIKE pattern
```

LIKE Operator Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select the persons living in a city that starts with "s" from the table above.

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE City LIKE 's%'
```

The "%" sign can be used to define wildcards (missing letters in the pattern) both before and after the pattern.

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Next, we want to select the persons living in a city that ends with an "s" from the "Persons" table.

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE City LIKE '%s'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes

Next, we want to select the persons living in a city that contains the pattern "tav" from the "Persons" table.

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE City LIKE '%tav%'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
3	Pettersen	Kari	Storgt 20	Stavanger

It is also possible to select the persons living in a city that NOT contains the pattern "tav" from the "Persons" table, by using the NOT keyword.

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE City NOT LIKE '%tav%'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes

SQL IN Operator

The IN operator allows you to specify multiple values in a WHERE clause.

SQL IN Syntax

```
SELECT column_name(s)
FROM table_name
WHERE column_name IN (value1,value2,...)
```

IN Operator Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select the persons with a last name equal to "Hansen" or "Pettersen" from the table above.

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE LastName IN ('Hansen','Pettersen')
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

SQL BETWEEN Operator

The BETWEEN operator is used in a WHERE clause to select a range of data between two values.The BETWEEN Operator

The BETWEEN operator selects a range of data between two values. The values can be numbers, text, or dates.

SQL BETWEEN Syntax

```
SELECT column_name(s)
```

```
FROM table_name
WHERE column_name
BETWEEN value1 AND value2
```

BETWEEN Operator Example

The "Persons" table:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

Now we want to select the persons with a last name alphabetically between "Hansen" and "Pettersen" from the table above.

We use the following SELECT statement:

```
SELECT * FROM Persons
WHERE LastName
BETWEEN 'Hansen' AND 'Pettersen'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
1	Hansen	Ola	Timoteivn 10	Sandnes

Note: The BETWEEN operator is treated differently in different databases.

Example 2

To display the persons outside the range in the previous example, use NOT BETWEEN:

```
SELECT * FROM Persons
WHERE LastName
NOT BETWEEN 'Hansen' AND 'Pettersen'
```

The result-set will look like this:

P_Id	LastName	FirstName	Address	City
2	Svendson	Tove	Borgvn 23	Sandnes
3	Pettersen	Kari	Storgt 20	Stavanger

SQL Alias

With SQL, an alias name can be given to a table or to a column.You can give a table or a column another name by using an alias. This can be a good thing to do if you have very long or complex table names or column names. An alias name could be anything, but usually it is short.

SQL Alias Syntax for Tables

```
SELECT column_name(s)
FROM table_name
AS alias_name
```

SQL Alias Syntax for Columns

```
SELECT column_name AS alias_name
FROM table_name
```

Alias Example

Assume we have a table called "Persons" and another table called "Product_Orders". We will give the table aliases of "p" an "po" respectively.

Now we want to list all the orders that "Ola Hansen" is responsible for.

We use the following SELECT statement:

```
SELECT po.OrderID, p.LastName, p.FirstName
FROM Persons AS p,
Product_Orders AS po
WHERE p.LastName='Hansen' AND p.FirstName='Ola'
```

The same SELECT statement without aliases:

```
SELECT Product_Orders.OrderID, Persons.LastName, Persons.FirstName
FROM Persons,
Product_Orders
WHERE Persons.LastName='Hansen' AND Persons.FirstName='Ola'
```

As you'll see from the two SELECT statements above; aliases can make queries easier to both write and to read.

GROUP FUNCTIONS(Aggregate operators)

1. SQL SUM() Function - The SUM() function returns the total sum of a numeric column.

SQL SUM() Syntax

```
SELECT SUM(column_name) FROM table_name
```

SQL SUM() Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to find the sum of all "OrderPrice" fields". We use the following SQL statement:

```
SELECT SUM(OrderPrice) AS OrderTotal FROM Orders
```

The result-set will look like this:

OrderTotal
5700

2. SQL AVG() Function

The AVG() function returns the average value of a numeric column.

SQL AVG() Syntax

```
SELECT AVG(column_name) FROM table_name
```

SQL AVG() Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to find the average value of the "OrderPrice" fields.

We use the following SQL statement:

```
SELECT AVG(OrderPrice) AS OrderAverage FROM Orders
```

The result-set will look like this:

OrderAverage
950

Now we want to find the customers that have an OrderPrice value higher than the average OrderPrice value. We use the following SQL statement:

```
SELECT Customer FROM Orders
WHERE OrderPrice>(SELECT AVG(OrderPrice) FROM Orders)
```

The result-set will look like this:

Customer

Hansen
Nilsen
Jensen

3. SQL MIN() Function

The MIN() function returns the smallest value of the selected column.

SQL MIN() Syntax

```
SELECT MIN(column_name) FROM table_name
```

SQL MIN() Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to find the smallest value of the "OrderPrice" column.

We use the following SQL statement:

```
SELECT MIN(OrderPrice) AS SmallestOrderPrice FROM Orders
```

The result-set will look like this:

SmallestOrderPrice
100

4. SQL MAX() Function

The MAX() function returns the largest value of the selected column.

SQL MAX() Syntax

```
SELECT MAX(column_name) FROM table_name
```

SQL MAX() Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to find the largest value of the "OrderPrice" column. We use the following SQL statement:

```
SELECT MAX(OrderPrice) AS LargestOrderPrice FROM Orders
```

The result-set will look like this:

LargestOrderPrice
2000

5. SQL COUNT() Function

The COUNT() function returns the number of rows that matches a specified criteria.

SQL COUNT(column_name) Syntax

The COUNT(column_name) function returns the number of values (NULL values will not be counted) of the specified column:

```
SELECT COUNT(column_name) FROM table_name
```

SQL COUNT(*) Syntax

The COUNT(*) function returns the number of records in a table:

```
SELECT COUNT(*) FROM table_name
```

SQL COUNT(DISTINCT column_name) Syntax

The COUNT(DISTINCT column_name) function returns the number of distinct values of the specified column:

```
SELECT COUNT(DISTINCT column_name) FROM table_name
```

Note: COUNT(DISTINCT) works with ORACLE and Microsoft SQL Server, but not with Microsoft Access.

SQL COUNT(column_name) Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to count the number of orders from "Customer Nilsen".

We use the following SQL statement:

```
SELECT COUNT(Customer) AS CustomerNilsen FROM Orders
WHERE Customer='Nilsen'
```

The result of the SQL statement above will be 2, because the customer Nilsen has made 2 orders in total:

CustomerNilsen
2

SQL COUNT(*) Example

If we omit the WHERE clause, like this:

```
SELECT COUNT(*) AS NumberOfOrders FROM Orders
```

The result-set will look like this:

NumberOfOrders
6

which is the total number of rows in the table.

SQL COUNT(DISTINCT column_name) Example

Now we want to count the number of unique customers in the "Orders" table.

We use the following SQL statement:

```
SELECT COUNT(DISTINCT Customer) AS NumberOfCustomers FROM Orders
```

The result-set will look like this:

NumberOfCustomers
3

which is the number of unique customers (Hansen, Nilsen, and Jensen) in the "Orders" table.

SQL GROUP BY Statement

Aggregate functions often need an added GROUP BY statement.

The GROUP BY Statement

The GROUP BY statement is used in conjunction with the aggregate functions to group the result-set by one or more columns.

SQL GROUP BY Syntax

```
SELECT column_name, aggregate_function(column_name)
FROM table_name
WHERE column_name operator value
GROUP BY column_name
```

SQL GROUP BY Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to find the total sum (total order) of each customer.

We will have to use the GROUP BY statement to group the customers. We use the following SQL statement:

```
SELECT Customer,SUM(OrderPrice) FROM Orders
GROUP BY Customer
```

The result-set will look like this:

Customer	SUM(OrderPrice)
Hansen	2000

Nilsen	1700
Jensen	2000

Let's see what happens if we omit the GROUP BY statement:

```
SELECT Customer,SUM(OrderPrice) FROM Orders
```

The result-set will look like this:

Customer	SUM(OrderPrice)
Hansen	5700
Nilsen	5700
Hansen	5700
Hansen	5700
Jensen	5700
Nilsen	5700

The result-set above is not what we wanted.

Explanation of why the above SELECT statement cannot be used: The SELECT statement above has two columns specified (Customer and SUM(OrderPrice). The "SUM(OrderPrice)" returns a single value (that is the total sum of the "OrderPrice" column), while "Customer" returns 6 values (one value for each row in the "Orders" table). This will therefore not give us the correct result. However, you have seen that the GROUP BY statement solves this problem.

GROUP BY More Than One Column

We can also use the GROUP BY statement on more than one column, like this:

```
SELECT Customer,OrderDate,SUM(OrderPrice) FROM Orders
GROUP BY Customer,OrderDate
```

SQL HAVING Clause

The HAVING Clause

The HAVING clause was added to SQL because the WHERE keyword could not be used with aggregate functions.

SQL HAVING Syntax

```
SELECT column_name, aggregate_function(column_name)
FROM table_name
WHERE column_name operator value
GROUP BY column_name
HAVING aggregate_function(column_name) operator value
```

SQL HAVING Example

We have the following "Orders" table:

O_Id	OrderDate	OrderPrice	Customer
1	2008/11/12	1000	Hansen
2	2008/10/23	1600	Nilsen
3	2008/09/02	700	Hansen
4	2008/09/03	300	Hansen
5	2008/08/30	2000	Jensen
6	2008/10/04	100	Nilsen

Now we want to find if any of the customers have a total order of less than 2000.

We use the following SQL statement:

```
SELECT Customer,SUM(OrderPrice) FROM Orders
GROUP BY Customer
HAVING SUM(OrderPrice)<2000
```

The result-set will look like this:

Customer	SUM(OrderPrice)
Nilsen	1700

Now we want to find if the customers "Hansen" or "Jensen" have a total order of more than 1500.

We add an ordinary WHERE clause to the SQL statement:

```
SELECT Customer,SUM(OrderPrice) FROM Orders
WHERE Customer='Hansen' OR Customer='Jensen'
GROUP BY Customer
HAVING SUM(OrderPrice)>1500
```

The result-set will look like this:

Customer	SUM(OrderPrice)
Hansen	2000
Jensen	2000

The SQL UNION Operator

The UNION operator is used to combine the result-set of two or more SELECT statements.

Notice that each SELECT statement within the UNION must have the same number of columns. The columns must also have similar data types. Also, the columns in each SELECT statement must be in the same order.

SQL UNION Syntax

```
SELECT column_name(s) FROM table1
UNION
SELECT column_name(s) FROM table2;
```

Note: The UNION operator selects only distinct values by default. To allow duplicate values, use the ALL keyword with UNION.

SQL CREATE VIEW Statement

In SQL, a view is a virtual table based on the result-set of an SQL statement.

A view contains rows and columns, just like a real table. The fields in a view are fields from one or more real tables in the database.

You can add SQL functions, WHERE, and JOIN statements to a view and present the data as if the data were coming from one single table.

SQL CREATE VIEW Syntax

```
CREATE VIEW view_name AS
SELECT column_name(s)
FROM table_name
WHERE condition
```

Note: A view always shows up-to-date data! The database engine recreates the data, using the view's SQL statement, every time a user queries a view.

SQL CREATE VIEW Examples

If you have the Northwind database you can see that it has several views installed by default.

The view "Current Product List" lists all active products (products that are not discontinued) from the "Products" table. The view is created with the following SQL:

```
CREATE VIEW [Current Product List] AS
SELECT ProductID,ProductName
FROM Products
WHERE Discontinued=No
```

We can query the view above as follows:

```
SELECT * FROM [Current Product List]
```

Another view in the Northwind sample database selects every product in the "Products" table with a unit price higher than the average unit price:

```
CREATE VIEW [Products Above Average Price] AS
SELECT ProductName,UnitPrice
FROM Products
WHERE UnitPrice>(SELECT AVG(UnitPrice) FROM Products)
```

We can query the view above as follows:

```
SELECT * FROM [Products Above Average Price]
```

Another view in the Northwind database calculates the total sale for each category in 1997. Note that this view selects its data from another view called "Product Sales for 1997":

```
CREATE VIEW [Category Sales For 1997] AS
SELECT DISTINCT CategoryName,Sum(ProductSales) AS CategorySales
FROM [Product Sales for 1997]
GROUP BY CategoryName
```

SQL: INTERSECT Query

The SQL INTERSECT query allows you to return the results of 2 or more "select" queries. However, it only returns the rows selected by all queries. If a record exists in one query and not in the other, it will be omitted from the INTERSECT results.

Each SQL statement within the SQL INTERSECT query must have the same number of fields in the result sets with similar data types.

The syntax for the SQL INTERSECT query is:

```
select field1, field2, . field_n

from tables

INTERSECT

select field1, field2, . field_n

from tables;
```

SQL INTERSECT Query - Single field example

The following is an example of an SQL INTERSECT query that has one field with the same data type:

```
select supplier_id
from suppliers
INTERSECT
select supplier_id
from orders;
```

In this SQL INTERSECT query example, if a supplier_id appeared in both the suppliers and orders table, it would appear in your result set.

SQL INTERSECT Query - Using ORDER BY Clause example

The following is an SQL INTERSECT query that uses an SQL ORDER BY clause:

```
select supplier_id, supplier_name
from suppliers
where supplier_id > 2000
INTERSECT
select company_id, company_name
from companies
where company_id > 1000
ORDER BY 2;
```

Since the column names are different between the two "select" statements, it is more advantageous to reference the columns in the SQL ORDER BY clause by their position in the result set. In this example, we've sorted the results by supplier_name / company_name in ascending order, as denoted by the "ORDER BY 2".

SQL – EXCEPT COMMANDS

The SQL **EXCEPT** clause/operator is used to combine two SELECT statements and returns rows from the first SELECT statement that are not returned by the second SELECT statement. This means EXCEPT returns only rows, which are not available in second SELECT statement.

Just as with the UNION operator, the same rules apply when using the EXCEPT operator. MySQL does not support EXCEPT operator.

Syntax:

The basic syntax of **EXCEPT** is as follows:

```
SELECT column1 [, column2 ]
FROM table1 [, table2 ]
[WHERE condition]

EXCEPT

SELECT column1 [, column2 ]
FROM table1 [, table2 ]
[WHERE condition]
```

Here given condition could be any given expression based on your requirement.

Example:

Consider the following two tables, (a) CUSTOMERS table is as follows:

```
+----+----------+-----+-----------+----------+
| ID | NAME     | AGE | ADDRESS   | SALARY   |
+----+----------+-----+-----------+----------+
|  1 | Ramesh   |  32 | Ahmedabad |  2000.00 |
|  2 | Khilan   |  25 | Delhi     |  1500.00 |
```

```
| 3 | kaushik  |  23 | Kota      |  2000.00 |
| 4 | Chaitali |  25 | Mumbai    |  6500.00 |
| 5 | Hardik   |  27 | Bhopal    |  8500.00 |
| 6 | Komal    |  22 | MP        |  4500.00 |
| 7 | Muffy    |  24 | Indore    | 10000.00 |
+----+----------+-----+-----------+----------+
```

(b) Another table is ORDERS as follows:

```
+-----+---------------------+-------------+--------+
|OID  | DATE                | CUSTOMER_ID | AMOUNT |
+-----+---------------------+-------------+--------+
| 102 | 2009-10-08 00:00:00 |           3 |   3000 |
| 100 | 2009-10-08 00:00:00 |           3 |   1500 |
| 101 | 2009-11-20 00:00:00 |           2 |   1560 |
| 103 | 2008-05-20 00:00:00 |           4 |   2060 |
+-----+---------------------+-------------+--------+
```

Now, let us join these two tables in our SELECT statement as follows:

```
SQL> SELECT  ID, NAME, AMOUNT, DATE
    FROM CUSTOMERS
    LEFT JOIN ORDERS
    ON CUSTOMERS.ID = ORDERS.CUSTOMER_ID
EXCEPT
    SELECT  ID, NAME, AMOUNT, DATE
    FROM CUSTOMERS
    RIGHT JOIN ORDERS
    ON CUSTOMERS.ID = ORDERS.CUSTOMER_ID;
```

This would produce the following result:

```
+----+---------+--------+---------------------+
| ID | NAME    | AMOUNT | DATE                |
+----+---------+--------+---------------------+
|  1 | Ramesh  |   NULL | NULL                |
|  5 | Hardik  |   NULL | NULL                |
|  6 | Komal   |   NULL | NULL                |
|  7 | Muffy   |   NULL | NULL                |
+----+---------+--------+---------------------+
```

SQL - ALTER TABLE COMMAND

The SQL **ALTER TABLE** command is used to add, delete or modify columns in an existing table.

You would also use ALTER TABLE command to add and drop various constraints on a an existing table.

Syntax:

The basic syntax of **ALTER TABLE** to add a new column in an existing table is as follows:

```
ALTER TABLE table_name ADD column_name datatype;
```

The basic syntax of ALTER TABLE to **DROP COLUMN** in an existing table is as follows:

```
ALTER TABLE table_name DROP COLUMN column_name;
```

The basic syntax of ALTER TABLE to change the **DATA TYPE** of a column in a table is as follows:

```
ALTER TABLE table_name MODIFY COLUMN column_name datatype;
```

The basic syntax of ALTER TABLE to add a **NOT NULL** constraint to a column in a table is as follows:

```
ALTER TABLE table_name MODIFY column_name datatype NOT NULL;
```

The basic syntax of ALTER TABLE to **ADD UNIQUE CONSTRAINT** to a table is as follows:

```
ALTER TABLE table_name
ADD CONSTRAINT MyUniqueConstraint UNIQUE(column1, column2...);
```

The basic syntax of ALTER TABLE to **ADD CHECK CONSTRAINT** to a table is as follows:

```
ALTER TABLE table_name
ADD CONSTRAINT MyUniqueConstraint CHECK (CONDITION);
```

The basic syntax of ALTER TABLE to **ADD PRIMARY KEY** constraint to a table is as follows:

```
ALTER TABLE table_name
ADD CONSTRAINT MyPrimaryKey PRIMARY KEY (column1, column2...);
```

All Directory Manipulation and Operating System commands :

Creating directory, Sub directory, Renaming, Coping and Deleting the directory:-

commands in DOS. commands with their syntax

DOS commands are small programs, which are made to perform a particular job. Every DOS command performs different task. It is not possible to work on the computer without these commands. There are two types of DOS commands

(1) Internal Commands

(2) External Commands

Internal Commands : These commands enter into the computer memory during computer booting. These commands are not in the form of any file; so neither they can be viewed nor can be edited or detected. For example : MD, CD, TIME, DATE, COPY, COPY CON, TYPE ETC.

External Commands : These commands are stored in the computer list in the form of files. These Commands can be viewed, copied, changed or deleted. For example : FORMAT, COPY, PRINT, SYS, EDIT, TREE, SORT, PROMPT etc.

Important Internal DOS Commands :

(i) **MD (Make Directory) :** We use this commands to make a new **directory or sub directory**.

Syntax : C\:>MD DIRECTORY NAME

Example : MD STUDENT

(ii) Rename a directory

Syntax: ren [old name] [new name]-> used for renaming a director

for example if you got a file named one.txt, then to rename this file to two.txt write the below command

Example: ren one.txt two.txt

(iii) **XCOPY(Coping directory)**

Syntax: XCOPY source [destination]

Example: xcopy c:\test c:\test2

For example there is directory C:/test(source directory) and C:/test2(destination directory)

(iv) **CD (Change Directory) :** This commands is used to move from one directory to another.

Syntax : C:\> CD Directory name

Example : > CD Student

Exit to Directory :

CD… The command move the subdirectory to parent directory.

CD The command is used to move directly to the root directory.

(v) RD (Remove/delete Directory) : If a Directory which was earlier is ;not required than such directory can be removed by using

Syntax : C:\> RD Directory name

Example : >RD student

Operating system commands

File Manipulation: Creating a file, deleting, coping, Renaming a file

(i) Copy Con : Its command is used to create a file. The name of the file, which is to be created, is written after the copy Con leaving one space in between

Syntax : Copy Con file name

To created a file, the following steps are :

(i) Type Copy Con Monu and press Enter.

(ii) Type whatever is to be typed in the file.

(iii) Press F6 function key or CTRL + Z keys '?Z' will be displayed on the screen, which indicates that the file is complete.

(iv) Press Enter and after that the DOS will save the file and will display the message '1 File(s) copied.

(ii) Del : This command is used to erase the files which are no longer required.

Syntax : C:\> Del < File name >

Example : >Del Monu

(iii) **Type :** This command is used to view the contents of text file.

Syntax : Type <file name>

(iv) **Copy :** This command is used to copy of file from one place to another place. A copy of file is another file with the same contents.

Syntax : C:1> copy <source> < destination path>

(v) **Ren :** This command is used to rename the file. In REN command two parameters are used. The first is the file we want to rename and the second is the new name for the file.

Syntax : > Ren <old file name> <New file name>

(vi) **DIR :** This command is used to display of directory and files.

Syntax : C:\> DIR

(viii) **CLS :** This command is used to clear the screen.

Syntax : C:\> CLS

DOS Boot-up Sequence

IO.SYS: A hidden file in the root directory of the primary drive. This file provides the basic I/O capabilities for the system, allowing it the ability to communicate with the different peripherals. IO.SYS directs the overall process of loading the Operating System.

MSDOS.SYS: Also a hidden file in the root directory sometimes called the kernal for DOS. When an application needs to access a device or peripheral, this file translates the request into actions that IO.SYS can perform.

CONFIG.SYS: A user-configurable text file that usually contains device drivers and system setup values.

COMMAND.COM: This is the Command Interpreter. It can accept commands from the user, launch programs and pass this Information to MSDOS.SYS.

AUTOEXEC.BAT: Another user-configurable text file that is used to set system variables and load TSRs.

Once DOS is loaded into memory, the system looks for two files; Config.sys and Autoexec.bat. These files are "user configurable". This means that you can make changes to these files that will affect the way your computer starts up. DOS then configures your system and carries out any start up commands that are in these files. If the final command in autoexec.bat is to open windows3.x, then DOS opens windows and leaves you in the program manager. If not, DOS leaves you with the Command Prompt (C:\>) and waits for your command.

The "C:" means you are currently using or "on" your C drive, which is usually your hard drive on a stand alone computer, and the "\" means you are at the "root" of that drive (the very beginning of the drive's directory).

Word processors

A WORD PROCESSOR is a program, which is used for preparing documents, typing letters, and producing reports.

Advantages of word processors over traditional typewriters:

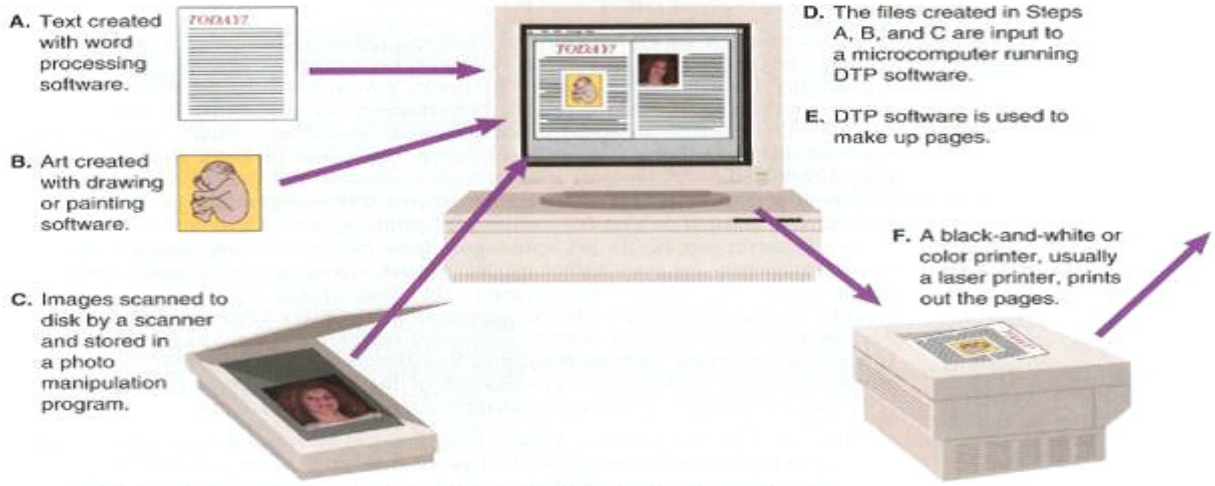
- Ease of correction
- Automatic Formatting
- Multiple copies and Form Letters
- Automatic Features
- Search and replace facilities
- Spelling checkers and thesaurus
- Font changes
- Mail Merge
- Other features: Multi column text, grammar checker, table formatting, inclusion of drawings

Desk top publishing

Desk Top Publishing (DTP), involves using a computer, mouse, scanner, printer and Desk Top Publishing application software for mixing text and graphics to produce high quality printed output for commercial printing. The emphasis is not on the entry of text (as it is with wordprocessing), but the ability to compose text and format the document in a manner that comes close to the quality achieved with professional typesetting equipment.

Desktop Publishing Software is used for:

- Catalogues
- Brochures
- Presentations
- Reports
- Desktop Publishing Software is NOT used for:
- Memos
- News releases
- Addressing
- Mailing list



DTP software must offer the following capabilities:

Allow high – quality composition of text:

The results must be of typeset quality, meaning that proportional spacing, kerning, tracking and leading control, and justification of both right and left margins is possible. Automatic hyphenation is also usually supported.

Support page composition:

This allows the user to arrange text and images on the display screen as desired – in columns, sidebars, and so on. The arrangement on the screen is the same as that which appears when the page is printed.

Incorporate graphics:

This makes it possible for the user to move text and images as desired. The DTP software does not usually allow the user to create the text and the image data – these must be created using other software (a word processor for the text, and a graphics package for the graphics) and then imported into the DTP program for exact placement.

Provide editing capabilities:

All DTP software allows the user to edit text. Such text editing facilities are usually minimal, and only meant for performing small corrections or adding captions and headings.

Printed results should be of high quality:

Although this of course depends more on the printer than the DTP software, it does mean that the software must support high-resolution printers (typically at least 300 dpi resolution).

Nowadays, Word Processors such as MS Word will also provide desktop publishing features such as automatic table of contents and index creation, multi-column documents, tables, frames, embedded graphics objects etc.

Even providing such DTP features however, software such as MS Word still is considered a Word Processor and not a Desktop Publisher.

Spread sheets and Graphical solutions

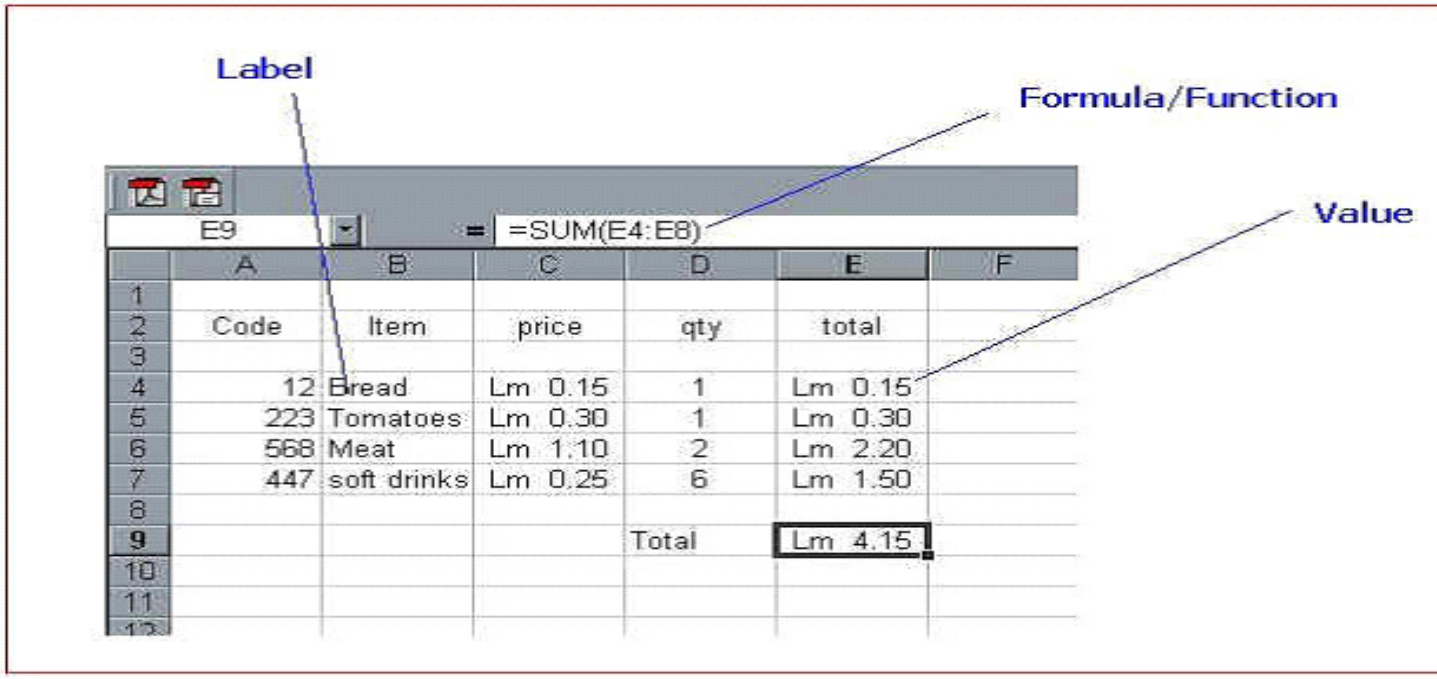
An electronic spreadsheet is a tool for performing calculations on data. Used generally for accounting purposes, the spreadsheet is made up of columns and rows. By using a spreadsheet program, one is able to automatically recalculate a whole sheet of figures every time a single value is altered.

1. **Components of a spreadsheet**.

The spreadsheet is made up of columns labelled with letters (A to Z, AA to AZ, BA to BZ etc.) and rows labelled with numbers. The intersection of a column and a row is called a CELL. Each cell has an address or coordinates for (e.g. A1, B6, F17, etc.)

2. **Labels, Values and Formulas.**

In each cell, the user may insert one of three types of items – LABELS, VALUES and FORMULAS (or FORMULAE). The following example spreadsheet will be used for illustration.



	A	B	C	D	E	F
1						
2	Code	Item	price	qty	total	
3						
4	12	Bread	Lm 0.15	1	Lm 0.15	
5	223	Tomatoes	Lm 0.30	1	Lm 0.30	
6	568	Meat	Lm 1.10	2	Lm 2.20	
7	447	soft drinks	Lm 0.25	6	Lm 1.50	
8						
9				Total	Lm 4.15	
10						
11						

Lables are words to describe particular items such as Bread, Tomatoes, Meat etc.
Values are numeric quantities like LM 0.15, LM 0.30 etc.

Value cells are sometimes called Input cells, since the user inputs data to the spreadsheet via these cells, and represent the variables or parameters of the system being modelled by the spreadsheet.

Formulas describe computations to be made on the contents of the cells (formulas are not displayed on the spreadsheet – they are computed, and their results displayed instead). For this reason, formula cells are sometimes called Output cells or Result cells, since they serve to display the results of computations performed by the spreadsheet. The real power of a spreadsheet comes from the fact that formulas are recalculated every time the value in a value cell changes – the output cells change dynamically as the input parameters are varied.

Spreadsheet formulas can perform:
I. Financial calculations
II. Mathematics
III. Statistics
IV. Date calculations

3. Operations on Cell.
Spreadsheets can carry operations on cells or group of cells such as: clear, move, delete and copy whole rows and columns.

4. Data Graphing.
A nice feature of spreadsheets is to display selected values in graphical form:

- Line graphs
- Bar graphs
- Scatter plots
- Area graphs

5. Spreadsheet publishing.
Modern spreadsheets support desktop publishing features such as fonts, colours, graphics and borders. These features are useful for preparing presentation and reports. Databases are typically used for storing and accessing data.
A Database can be simply defined as a collection of information that is organized so that it can easily be accessed, managed, and updated.
Basic terminology used in Database environments:

- **Record:**

A structured collection of data regarding a single person, and is composed of fields. By structured we mean that, for example, an employee's record is not just a haphazard collection of facts and figures, but is a form which entries for the employee's name, number, address, rank and any other item of data thought desirable for record purposes.

- **Fields:**

A field holds a single item of information about that individual (for example: name, surname, address etc.) Fields come in various types, depending on the type of data to be stored.

Some common field types :-
a. Numeric
b. Date fields
c. Logical fields (or Boolean fields)

d. Text fields
e. Memo fields
f. Picture fields

UNIT-IV

Concept of Data Communication and Networking:

Communication Networks & Services Data Communication
When we communicate, we are sharing information. This sharing can be local or remote. In every form of communication there are 5 components:-

- Transmitter (sender, source)
- Receiver (destination)
- Message to be communicated
- Medium (how message is carried)

Message. The message is the information (data) to be communicated. Information may include text, numbers, pictures, audio, and video.
Sender. The sender is the device that sends the data message. It can be a computer, workstation, telephone handset, video camera, and so on.
Receiver. The receiver is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.
Transmission medium. The transmission medium is the physical path by which a message travels from sender to receiver. Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable, and radio waves.
Protocol. A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices. Without a protocol, two devices may be connected but not communicating, just as a person speaking French cannot be understood by a person who speaks only Japanese.

Computer Network :-
The term "computer network" to mean a collection of autonomous computers interconnected by a single technology. Two computers are said to be interconnected if they are able to exchange information. The connection need not be via a copper wire; fiber optics, microwaves, infrared, and communication satellites can also be used. Networks come in many sizes, shapes and forms.
Uses of computer networks are in Business Applications, Home Applications, Mobile Users, and Social Issues.
The **goals of a computer network** include:
• Resource sharing: programs (O.S., applications), data, equipment (printers, disks) are available to all users of the network regardless of location.

• High reliability: By replicating files on different machines and having spare cpus, users are more immune from hardware/software failure.
• Less cost: Small machines have about 1/10 the power of a mainframe but 1/1000 the cost. By using such machines with file server machine(s), a local area network LAN can be cheaply installed. It is easy to increase the capacity by adding new machines.
• **Communications medium**: Users have access to email and the Internet.

What is a communication network?
The equipment (hardware & software) and facilities that provide the basic communication service is called communication network. Equipment are Routers, servers, switches, multiplexers, hubs, modems,.. Facilities are Copper wires, coaxial cables, optical fiber,..

Types of Computer Networks or Categories of Network:
Today when we speak of networks, we are generally referring to three primary categories:

- Local Area Network (LAN)
- Metropolitan Area Network (MAN)
- Wide Area Network (WAN)

- **Local Area Network (LAN):**

When the computers in a network are located close together (usually less than 1or 2 Kms), the network is called a LAN. A Local Area Network is a group of computers and associated devices that share a common communications lines or wireless link and share the resources of a single processor or server within a small geographic area usually within an office building. Usually, a LAN is installed in industrial plants, office buildings, college or university campuses or similar locations. in these locations, it is feasible for the owning organization to install high quality, high speed communication links interconnecting nodes. Typical data transmission speeds are one to 100 megabits per second.

- LAN size is limited to few kilometers.
- LAN will use only one type of transmission media
- LANs have the data rates in the 4 to 16 Mbps range.
  Today, however speeds are increasing and can reach 100 Mbps with gigabit systems in development.

- **Metropolitan Area Network (MAN):**

A MAN is a network that interconnects users with computer resource in a geographical area larger than that covered by even a large LAN but smaller than the area covered by a WAN. A Metropolitan area network is designed to extend over an entire city. It may be a single network such as a cable television network, or it may be a means of connecting a number of LANs into a larger network so that resources may be shared LAN-to-LAN as well as device-to-device. MAN may be wholly owned and operated by a private company or it may be a service provided by a public company such as local telephone company.

- **Wide Area Network (WAN):**

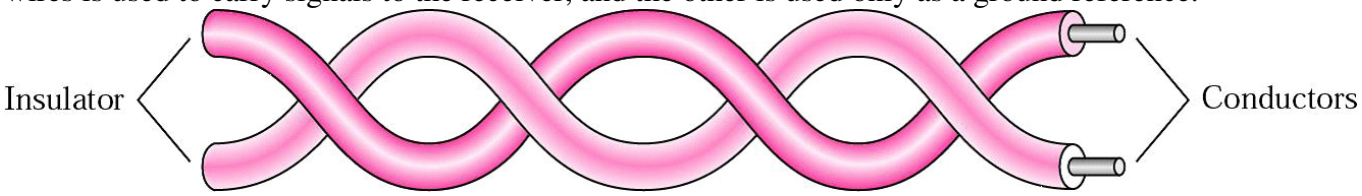
When the computers in a networks are separated by long distances (from a few Kms to global distances) the network is called a WAN. WAN is a telecommunication network which covers a large geographical area, and uses communications circuits to connect the intermediate nodes. WAN spans a wide geographical area such as a state or country. Numerous WAN have been constructed, including public packet networks, large corporate networks, military networks, banking networks, stock brokerage networks and airline reservation networks. The transmission rates are typically 2 Mbps, 34 Mbps, 45 Mbps, 155 Mbps, 625 Mbps or sometimes even more. Many WANs are used to transfer and consolidate corporate data, such as daily transaction summaries from branches.

Communication Media (Transmission Medium)

The transmission medium is the physical path by which a message travels from sender to receiver. Computers and telecommunication devices use signals to represent data. These signals are transmitted from a device to another in the form of electromagnetic energy.

- **Twisted Pair Cable**

Twisted pair consists of two conductors (normally copper), each with its own plastic insulation, twisted together. The twisting helps to reduce the interference (noise) and crosstalk. One of the wires is used to carry signals to the receiver, and the other is used only as a ground reference.



Twisted pairs can be used for transmitting either analog or digital signals. The bandwidth depends on the thickness of the wire and the distance traveled, but several megabits/sec can be achieved for a few kilometers in many cases. Due to their adequate performance and low cost, twisted pairs are widely used.

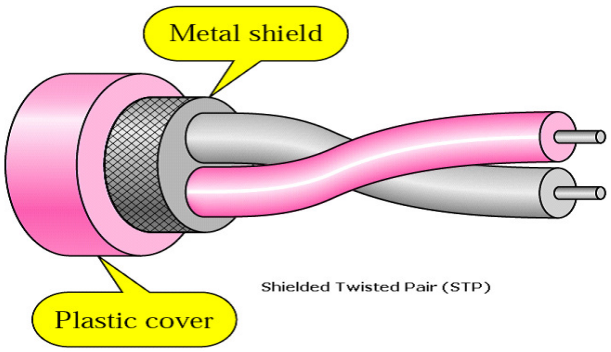
Twisted-pair cable comes in two forms:

· **Shielded Twisted Pair (STP)**

· **Unshielded Twisted pair (UTP)**

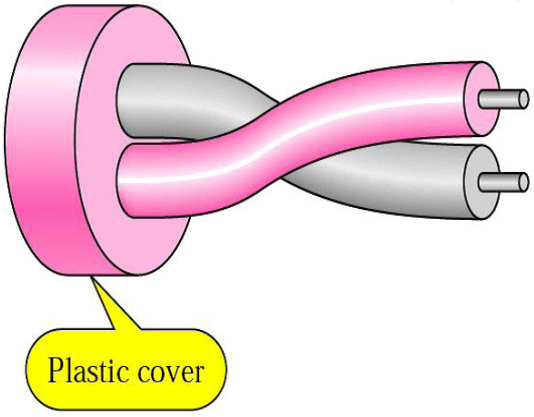
Shielded Twisted Pair (STP)

IBM has also produced a version of twisted-pair cable for its use called shielded twisted-pair (STP). STP cable has a metal foil or braided-mesh covering
that enhances each pair of insulated conductors. The metal casing prevents the penetration of electromagnetic noise. Materials and STP is costly than UTP but less susceptible to noise.



Shielded Twisted Pair (STP)

Unshielded Twisted Pair (UTP) UTP cable is the most common type of



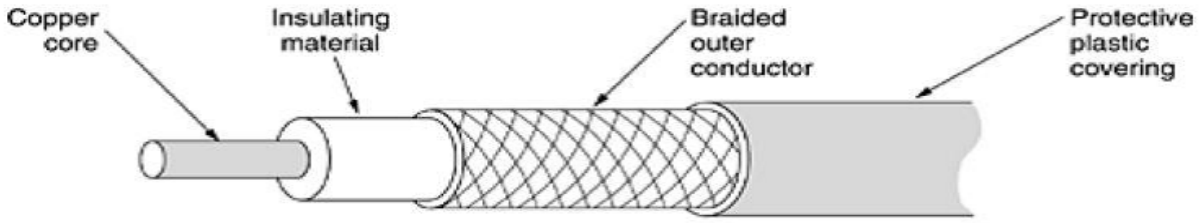
Unshielded Twisted pair (UTP)

Telecommunication medium in use today. The range is suitable for transmitting both data and video. UTP is cheap, flexible, and easy to install.

- **Coaxial Cable**

Coaxial cable carries signals of higher frequency ranges than twisted-pair cable. Two kinds of coaxial cable are widely used. One kind, 50-ohm cable, is commonly used when it is intended for digital transmission from the start. The other kind, 75-ohm cable, is commonly used for analog transmission and cable television but is becoming more important with the advent of Internet over cable.

A coaxial cable consists of a stiff copper wire as the core, surrounded by an insulating material. The insulator is encased by a cylindrical conductor, often as a closely-woven braided mesh. The outer conductor is covered in a protective plastic sheath.



The construction and shielding of the coaxial cable give it a good combination of high bandwidth and excellent noise immunity. The bandwidth possible depends on the cable quality, length, and signal-to-noise ratio of the data signal. Modern cables have a bandwidth of close to 1 GHz.

CHARACTERISTICS OF COAXIAL CABLE

- Low cost
- Easy to install
- Up to 10Mbps capacity
- Medium immunity form EMI
- Medium of attenuation

ADVANTAGES COAXIAL CABLE

- Inexpensive
- Easy to wire
- Easy to expand
- Moderate level of EMI immunity

DISADVANTAGE COAXIAL CABLE

Single cable failure can take down an entire network

- **Optical Fiber Cable**

Metal cables transmit signals in the form of electric current. Optical fiber is made of glass or plastic and transmits signals in the form of light. Light, a form of electromagnetic energy, travels at 300,000 Kilometers/second in a vacuum. The speed of the light depends on the density of the medium through which it is traveling (the higher density, the slower the speed).

An optical transmission system has three key components: the light source, the transmission medium, and the detector. Conventionally, a pulse of light indicates a 1 bit and the absence of light indicates a 0 bit. The transmission medium is an ultra-thin fiber of glass. The detector generates an electrical pulse when light falls on it. By attaching a light source to one end of an optical fiber and a detector to the other, we have a unidirectional data transmission system that accepts an electrical signal, converts and transmits it by light pulses, and then reconverts the output to an electrical signal at the receiving end.

Optical fibers use reflection to guide light through a channel.

Types of Optical Fiber

There are two basic types of fiber: **multimode fiber** and **single-mode fiber**.

Multimode fiber is best designed for short transmission distances, and is suited for use in LAN systems and video surveillance.
Single-mode fiber is best designed for longer transmission distances, making it suitable for long-distance telephony and multichannel television broadcast systems.
Advantages of Optical Fiber
The major advantages offered by fiber-optic cable over twisted-pair and coaxial cable are **noise resistance, less signal attenuation, and higher bandwidth**.

• **Noise Resistance**: Because fiber-optic transmission uses light rather than electricity, noise is not a factor. External light, the only possible interference, is blocked from the channel by the outer jacket.

• **Less signal attenuation:** Fiber-optic transmission distance is significantly greater than that of other guided media. A signal can run for miles without requiring regeneration.

• **Higher bandwidth:** Currently, data rates and bandwidth utilization over fiber-optic cable are limited not by the medium but by the signal generation and reception technology available.

Disadvantages of Optical Fiber
The main disadvantages of fiber optics are c**ost, installation/maintenance, and fragility**.

• Cost. Fiber-optic cable is expensive. Also, a laser light source can cost thousands of dollars, compared to hundreds of dollars for electrical signal generators.

• Installation/maintenance

• Fragility. Glass fiber is more easily broken than wire, making it less useful for applications where hardware portability is required.

- **Radio Transmission**

Radio waves are easy to generate, can travel long distances, and can penetrate buildings easily, so they are widely used for communication, both indoors and outdoors. Radio waves also are omni directional, meaning that they travel in all directions from the source, so the transmitter and receiver do not have to be carefully aligned physically.

The properties of radio waves are frequency dependent. At low frequencies, radio waves pass through obstacles well, but the power falls off sharply with distance from the source, roughly as 1/r2 in air. At high frequencies, radio waves tend to travel in straight lines and bounce off obstacles. They are also absorbed by rain. At all frequencies, radio waves are subject to interference from motors and other electrical equipment.

- **Infra red Light**

Unguided infrared waves are widely used for short-range communication. The remote controls used on televisions, VCRs, and stereos all use infrared communication. They are relatively directional, cheap, and easy to build but have a major drawback: they do not pass through solid objects (try standing between your remote control and your television and see if it still works). In general, as we go from long-wave radio toward visible light, the waves behave more and more like light and less and less like radio.

TRANSMISSION MODES

The direction of signal flow between two linked devices is called transmission modes. There are three types of transmission modes:

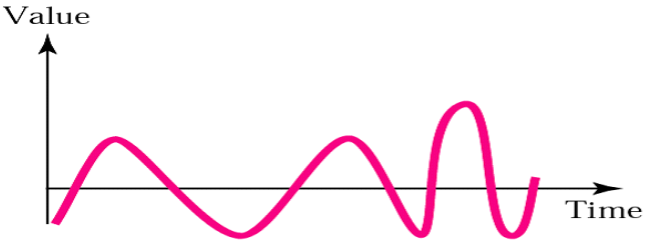
1. **Simplex:** Information is transmitted in one direction only and the roles of transmitter and receiver are fixed. This form is not used for conventional data transmission. Example: Pager, instructions / command send from computer to printer.
2. **Half Duplex** (HDX): transmission is allowed in both directions but in only one direction at a time. Data communication systems that use the telephone network usually transmit in HDX. Example: walkie-talkie.
3. **Full Duplex** (FDX): sender/receiver can transmit and receive from each other at the same time. In order to transmit in FDX, the user usually has private direct lines. Example: Telephone, Mobile.

Analog and Digital Transmission

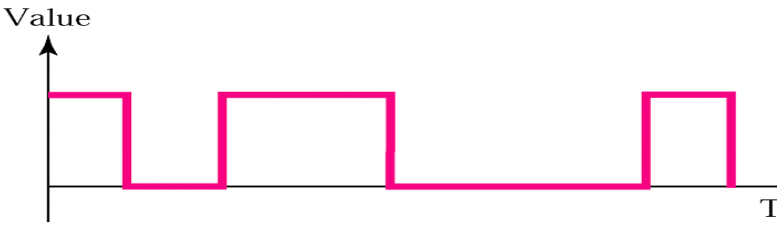
Both data and the signals that represent them can be either analog or digital in form.

Analog and Digital Signal

- Analog signals can have an infinite number of values in a range.
- Digital signals can have only a limited number of values.



a. Analog signal



b. Digital signal

Comparison of Analog and Digital Signals **Analog**	**Digital**
• Analog signal is a continuous signal which transmits information as a response to changes in physical phenomenon.	• Digital signals are discrete time signals generated by digital modulation.
• Data transmission is poor quality	• Data transmission is high (good)

	quality
• Response to Noise is More affected.	• Response to Noise is Less affected.
• Uses continuous range of values to represent information	• Uses discrete or discontinuous values to represent information
• Denoted by sine waves	• Denoted by square waves
• Analog technology records waveforms as they are.	• Converts analog waveforms into set of numbers and records them. The numbers are converted into voltage stream for representation.
• Example Human voice in air	• Example Electronic devices

Analog and digital signals are used to transmit information, usually through electric signals. In both these technologies, the information, such as any audio or video, is transformed into electric signals. The difference between analog and digital technologies is that in analog technology, information is translated into electric pulses of varying amplitude. In digital technology, translation of information is into binary format (zero or one) where each bit is representative of two distinct amplitudes.

Difference between Synchronous and Asynchronous Transmission

- Serial communication occurs in either synchronous or asynchronous format.
- In synchronous transmission a receiver and a transmitter are synchronized and a block of character is transmitted along with the synchronization transmission while asynchronous transmission is character oriented means each character carries start and stop bits .
- A Synchronous transmission is used for high speed transmission (more than 20 kbps) while asynchronous transmission is generally used in low speed transmission. (less than 20 kbps).
- In asynchronous transmission when no data transmitted a receiver stays high at logic 1 called mark. Logic 0 is called space. Transmission begins with one start bit followed by a character and one or two stop bits.

Advantages and disadvantages

	Advantages	Disadvantages
Asynchronous transmission	• Simple, doesn't require synchronization of both communication sides • Cheap, timing is not as critical as for synchronous transmission, therefore hardware can be made cheaper • Set-up is faster than other transmissions, so well suited for applications where	• Large relative overhead, a high proportion of the transmitted bits are uniquely for control purposes and thus carry no useful information

	messages are generated at irregular intervals, for example data entry from the keyboard and the speed depends on different applications.	
Synchronous transmission	Lower overhead and thus, greater throughput	Slightly more complex Hardware is more expensive

Network Topology:-

A Topology is a schematic description of the arrangement of a network, including its nodes and connecting lines. The physical topology of a network is the actual geometric layout of workstations. There are several common physical topologies, as described below and as shown in the figure. There are five basic topologies BUS, STAR, TREE, RING, MESH

Bus

A linear bus topology consists of a main run of cable with a terminator at each end. All nodes (file server, workstations, and peripherals) are connected to the linear cable.

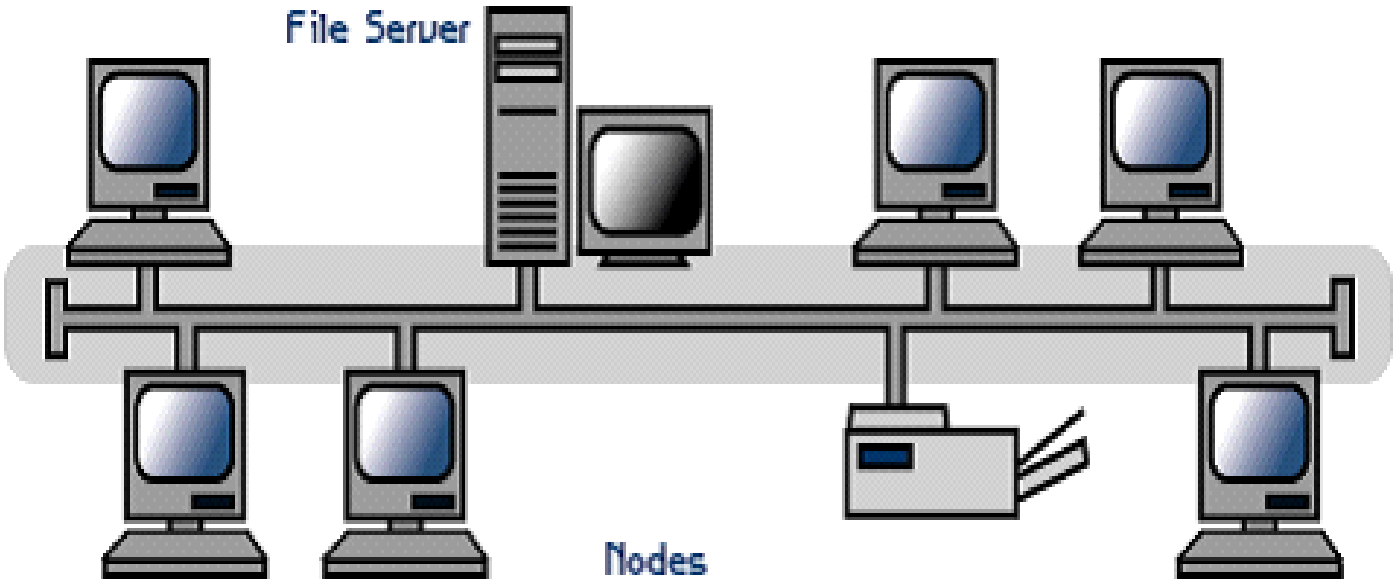


Fig. Linear Bus topology

In the bus network topology, every workstation is connected to a main cable called the bus. Therefore, in effect, each workstation is directly connected to every other workstation in the network.

Advantages of a Linear Bus Topology

Easy to connect a computer or peripheral to a linear bus.
Requires less cable length than a star topology.

Disadvantages of a Linear Bus Topology

- Entire network shuts down if there is a break in the main cable.
- Terminators are required at both ends of the backbone cable.
- Difficult to identify the problem if the entire network shuts down.
- Not meant to be used as a stand-alone solution in a large building.

Star

A star topology is designed with each node (file server, workstations, and peripherals) connected directly to a central network hub, switch, or concentrator.

In the star network topology, there is a central computer or server to which all the workstations are directly connected. Every workstation is indirectly connected to every other through the central computer.

Data on a star network passes through the hub, switch, or concentrator before continuing to its destination. The hub, switch, or concentrator manages and controls all functions of the network. It also acts as a repeater for the data flow. This configuration is common with twisted pair cable; however, it can also be used with coaxial cable or fiber optic cable.

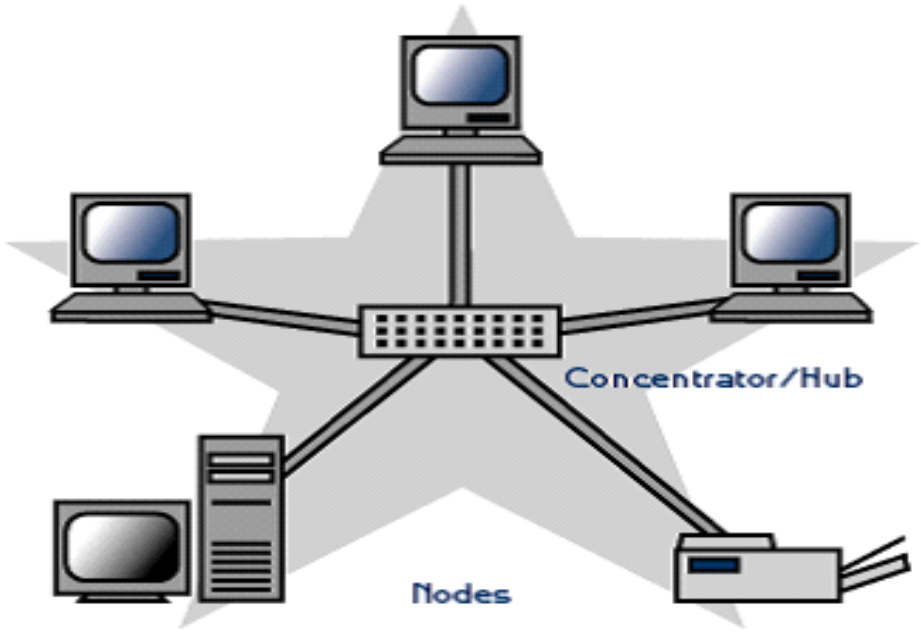


Fig. Star topology

Advantages of a Star Topology

- Easy to install and wire.
- No disruptions to the network when connecting or removing devices.
- Easy to detect faults and to remove parts.

Disadvantages of a Star Topology

- Requires more cable length than a linear topology.
- If the hub, switch, or concentrator fails, nodes attached are disabled.
- More expensive than linear bus topologies because of the cost of the hubs, etc.

Tree

A tree topology combines characteristics of linear bus and star topologies. It consists of groups of star-configured workstations connected to a linear bus backbone cable.

The tree network topology uses two or more star networks connected together. The central computers of the star networks are connected to a main bus. Thus, a tree network is a bus network of star networks. Tree topologies allow for the expansion of an existing network, and enable schools to configure a network to meet their needs.

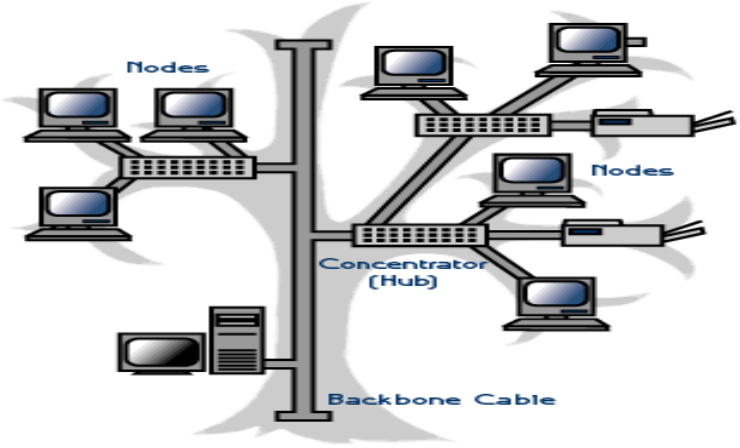


Fig. Tree topology

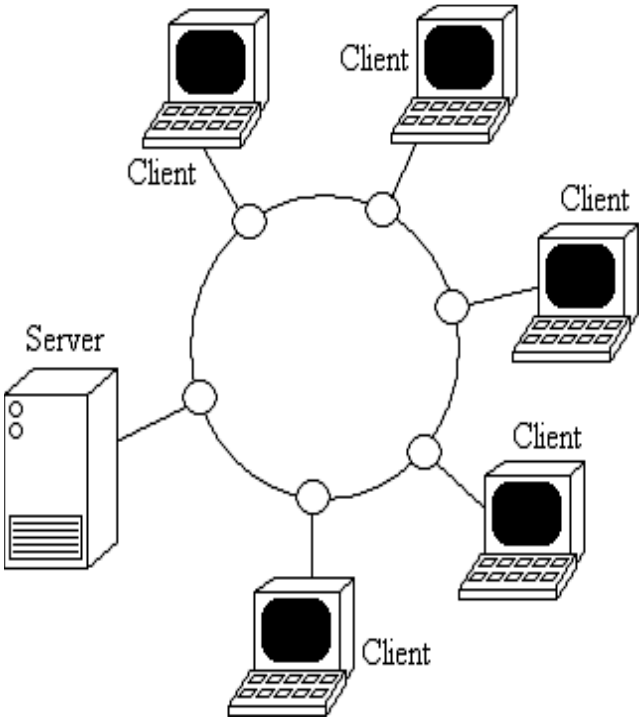
Advantages of a Tree Topology

- Point-to-point wiring for individual segments.
- Supported by several hardware and software venders.

Disadvantages of a Tree Topology

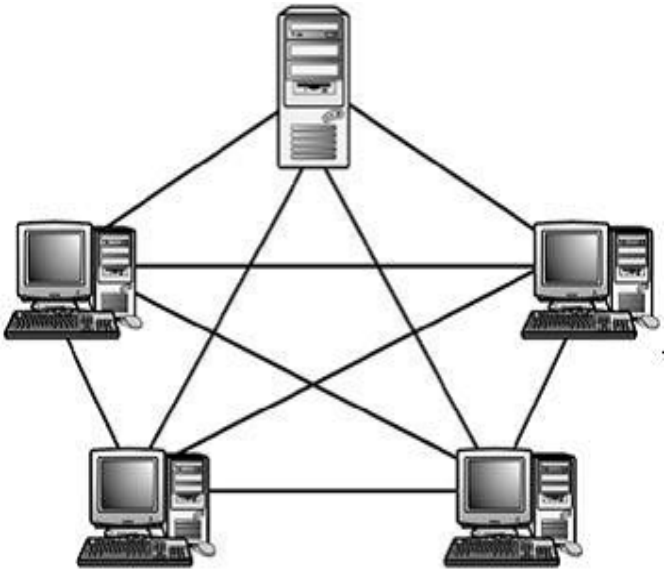
- Overall length of each segment is limited by the type of cabling used.
- If the backbone line breaks, the entire segment goes down.
- More difficult to configure and wire than other topologies.

Ring: In the ring network topology, the workstations are connected in a closed loop configuration. Adjacent pairs of workstations are directly connected. Other pairs of workstations are indirectly connected, the data passing through one or more intermediate nodes.



- A ring is relatively easy to install and reconfigure
- Unidirectional traffic can be disadvantage
- A break in the ring can disable the entire network

- **Mesh**: Mesh networking (topology) is a type of networking where each node must not only capture and distribute its own data, but also serve as a relay for other nodes, that is, it must collaborate to propagate the data in the network.



- The use of dedicated link guarantees that each connection can carry its data load, thus eliminating the traffic problems
- Mesh topology is robust. If one link becomes unusable, it does not incapacitate the entire system
- Privacy and security are ensured
- Fault identification would be easy

NETWORKING DEVICES

Networking hardware may also be known as **network equipment**, **computer networking devices**. Units which are the last receiver or generate data are called hosts or data terminal equipment.

HUB

Networks using a Star topology require a central point for the devices to connect. Originally this device was called a concentrator since it consolidated the cable runs from all network devices. The basic form of concentrator is the hub.

As shown in Figure; the hub is a hardware device that contains multiple, independent ports that match the cable type of the network. Most common hubs interconnect Category 3 or 5 twisted-pair cable with RJ-45 ends, although Coax BNC and Fiber Optic BNC hubs also exist. The hub is considered the least common denominator in device concentrators. Hubs offer an inexpensive option for transporting data between devices, but hubs don't offer any form of intelligence. Hubs can be active or passive.

An **active hub** strengthens and regenerates the incoming signals before sending the data on to its destination.

Passive hubs do nothing with the signal.

Ethernet Hubs

An Ethernet hub is also called a multiport repeater. A repeater is a device that amplifies a signal as it passes through it, to counteract the effects of attenuation. If, for example, you have a thin

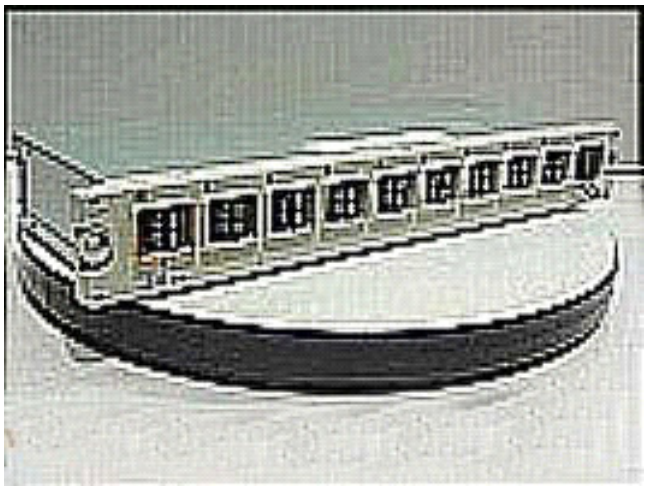
Ethernet network with a cable segment longer than the prescribed maximum of 185 meters, you can install a repeater at some point in the segment to strengthen the signals and increase the maximum segment length. This type of repeater only has two BNC connectors, and is rarely seen these days.

8 Port mini Ethernet Hub

The hubs used on UTP Ethernet networks are repeaters as well, but they can have many RJ45 ports instead of just two BNC connectors. When data enters the hub through any of its ports, the hub amplifies the signal and transmits it out through all of the other ports. This enables a star network to have a shared medium, even though each computer has its own separate cable. The hub relays every packet transmitted by any computer on the network to all of the other computers, and also amplifies the signals.

The maximum segment length for a UTP cable on an Ethernet network is 100 meters. A segment is defined as the distance between two communicating computers. However, because the hub also functions as a repeater, each of the cables connecting a computer to a hub port can be up to 100 meters long, allowing a segment length of up to 200 meters when one hub is inserted in the network.

Multistation Access Unit



A **Multistation Access Unit (MAU)** is a special type of hub used for token ring networks. The word **"hub"** is used most often in relation to Ethernet networks, and MAU only refers to token ring networks. On the outside, the MAU looks like a hub. It connects to multiple network devices, each with a separate cable.

Unlike a hub that uses a logical bus topology over a physical star, the MAU uses a logical ring topology over a physical star.

OSI MODEL

The **Open Systems Interconnection (OSI) model** (ISO/IEC 7498-1) is a conceptual model that characterizes and standardizes the internal functions of a communication system by partitioning it into abstraction layers. The model is a product of the Open Systems Interconnection project at the International Organization for Standardization (ISO).

The model groups similar communication functions into one of seven logical layers. A layer serves the layer above it and is served by the layer below it. For example, a layer that provides error-free communications across a network provides the path needed by applications above it, while it calls the next lower layer to send and receive packets that make up the contents of that path. Two instances at one layer are connected by a horizontal connection on that layer.

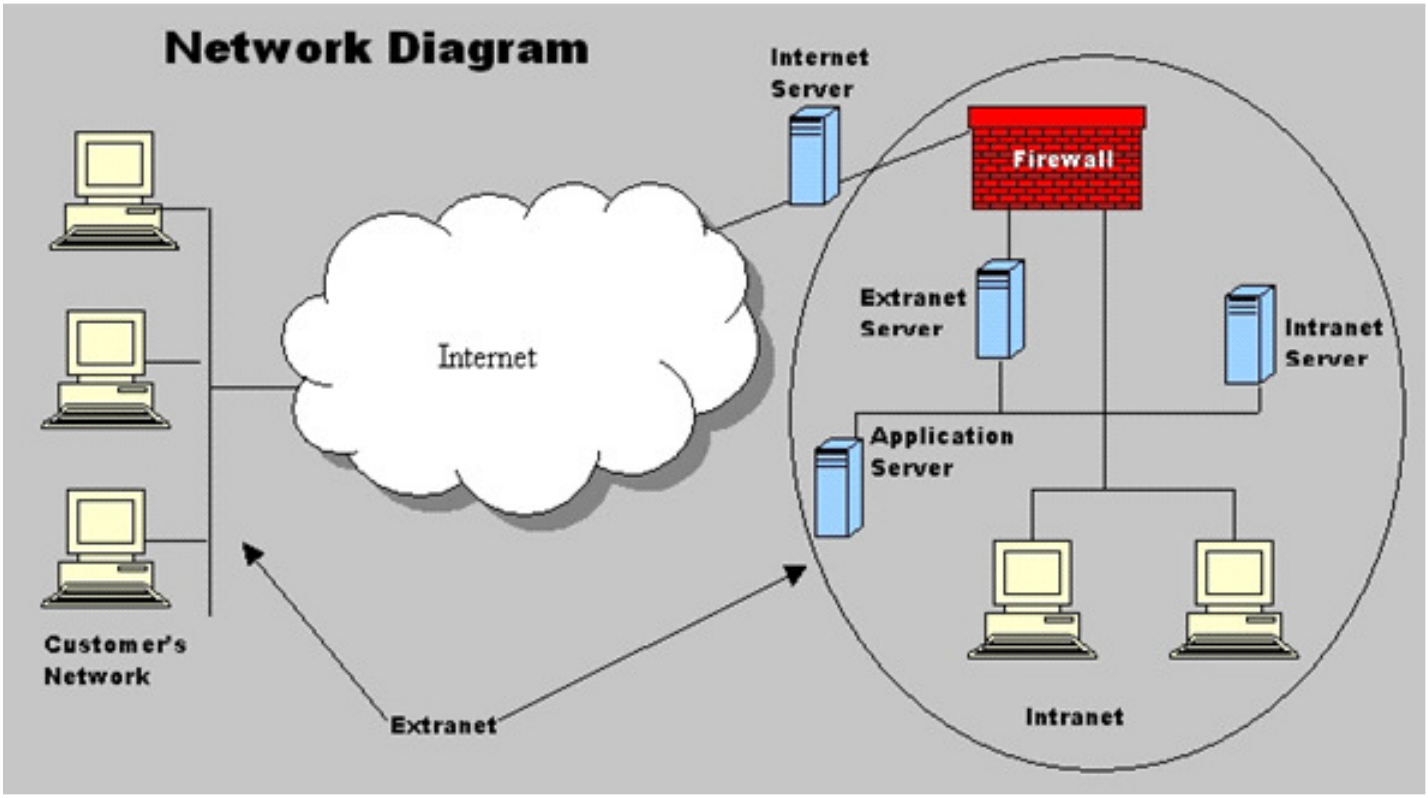
OSI Model				
	Data unit	**Layer**	**Function**	

Host layers	Data	7. Application	Network process to application	
		6. Presentation	Data representation, encryption and decryption, convert machine dependent data to machine independent data	
		5. Session	Interhost communication, managing sessions between applications	
	Segments	4. Transport	Reliable delivery of packets between points on a network.	
Media layers	Packet/Datagram	3. Network	Addressing, routing and (not necessarily reliable) delivery of datagrams between points on a network.	
	Bit/Frame	2. Data link	A reliable direct point-to-point data connection.	
	Bit	1. Physical	A (not necessarily reliable) direct point-to-point data connection	

INTERNET,INTERANET AND EXTRANET

- **Intranet** is shared content accessed by members within a single organization.
- **Extranet** is shared content accessed by groups through cross-enterprise boundaries.
- **Internet** is global communication accessed through the Web.

For better comprehension, take a look at this drawing:



Summary:

The Internet, extranets, and intranets all rely on the same TCP/IP technologies. However, they are different in terms of the levels of access they allow to various users inside and outside the organization and the size of the network. An intranet allows for restricted access to only members of an organization; an extranet expands that access by allowing non-members such as suppliers and customers to use company resources. The difference between the Internet and extranets is that while the extranet allows limited access to non-members of an organization, the Internet generally allows everyone to access all network resources.

APPLICATION OF IT

The term "*information*"has a diversity of meanings, from everyday usage to technical interpretations. Generally speaking, the concept of information is associated with knowledge derived from study, experience, or instruction. *Technology*, on the other hand, refers to the application of knowledge to the practical aims of human life, or to changing and manipulating the human environment. Technology includes the use of materials, tools, techniques and sources

of power to make life easier or more pleasant and work more productive. Technology began to influence human endeavour as soon as people began using tools. Technology also started being used for managing information when the amount and variety of information grew to such vast proportions that the human brain could neither store nor process it efficiently.

Definition

The term Information Technology (IT) was coined by Jim Domsic of Michigan in November 1981. Domsic created the term to modernize the outdated phrase "*data processing*". Information Technology is a general term that describes any technology that helps to produce, manipulate, store, communicate and/or disseminate information. Presumably, when speaking of Information Technology as a whole, it is noted that the use of computers and information are associated. "Information Technology" as defined by the Information Technology Association of America (ITAA), is "the study, design, development, implementation, support or management of computer-based information systems, particularly software applications and computer hardware." IT deals with the use of computers and computer software to convert, store, protect, process, transmit and securely retrieve information.

Encompassing the computer and information systems industries, Information Technology is the capability to electronically input, process, store, output, transmit, and receive data and information, including text, graphics, sound, and video, as well as the ability to control machines of all kinds electronically. Information Technology is comprised of computers, networks, satellite communications, robotics, video, text, cable television, electronic mail ("e-mail"), electronic games, and automated office equipment. The information industry consists of all computer, communications, and electronics-related organizations, including hardware, software, and services. Completing tasks using Information Technology results in rapid processing and information mobility, as well as improved reliability and integrity of processed information.

Another related term, **Information and Communications Technology** (ICT) is sometimes used in preference to Information Technology, particularly in the fields of education and governance. In common usage, it is often assumed that ICT is synonymous with IT; ICT in fact encompasses any medium to record information, technology for broadcasting information;

and technology for communicating through voice and/or images. It includes the wide variety of computing hardware (PCs, servers, mainframes, networked storage), the rapidly developing personal hardware market (mobile phones, personal devices, MP3 players), application software (from the smallest home-developed spreadsheet to online software services); and the hardware and software needed to operate networks for transmission of information. Thus, ICT makes more explicit that technologies such as broadcasting and wireless mobile telecommunications are included.

History of Information Technology

The basic concept of Information Technology can be traced to the World War II alliance of the military and industry in the development of electronics, computers, and information theory. After the 1940s, the military remained the major source of research and development funding for the expansion of automation to replace manpower with machine power.

Since the 1950s, four generations of computers have evolved. Each generation reflected a change to hardware of decreased size but increased capabilities to control computer operations. The first generation used vacuum tubes, the second used transistors, the third used integrated circuits, and the fourth used integrated circuits on a single computer chip. Advances in artificial intelligence that will minimize the need for complex programming characterize the fifth generation of computers, still in the experimental stage.

The first commercial computer was the UNIVAC I, developed by John Eckert and John W. Mauchly in 1951. It was used by the Census Bureau to predict the outcome of the 1952 presidential election. For the next twenty-five years, mainframe computers were used in large corporations to do calculations and manipulate large amounts of information stored in databases. Supercomputers were used in science and engineering, for designing aircraft and nuclear reactors, and for predicting worldwide weather patterns. Minicomputers came on to the scene in the early 1980s in small businesses, manufacturing plants, and factories.

In 1975, the Massachusetts Institute of Technology developed microcomputers. The market for microcomputers increased dramatically when IBM introduced the first personal computer in the fall of 1981. Because of dramatic improvements in computer components and

manufacturing, personal computers today do more than the largest computers of the mid-1960s at about a thousandth of the cost.

Indian IT Industry

The Indian Information Technology industry has played a key role in putting India on the global map. Thanks to the success of the IT industry, India is now a power to reckon with. According to the National Association of Software and Service Companies (NASSCOM), the apex body for software services in India, the revenue of the IT sector has grown from 1.2 per cent of the gross domestic product (GDP) in FY 1997-98 to an estimated 5.5 per cent in FY 2007-08. The net value added by this sector, to the economy, is estimated to be 3.3 to 3.9 per cent for FY 2007-08. Direct employment in Indian IT-BPO crossed the 2 million mark, an increase of about 389,000 professionals over FY2007; indirect job creation is estimated at about 8-9 million. IT-BPO exports (including hardware exports) reached USD 40.9 billion in FY2008 as against USD 31.8 billion in FY2007, a growth of 28 per cent. Many of the global IT companies have development centres in India where a lot of new products are being designed. India's most prized resource in today's knowledge economy is its readily available technical work force. India has the second largest English-speaking scientific professionals in the world, second only to the U.S.

The phenomenal growth of the Indian IT Software & Services, IT Enabled Services (ITES) and Business Process Outsourcing (BPO) sector has had a perceptible multiplier effect on the Indian economy as a whole. In addition to the direct positive impact on national income, the sector has grown to become the biggest employment generator, and has spawned the mushrooming of several ancillary industries such as transportation, real estate and catering, and has created a rising class of youthful consumers with high disposable incomes. This, in turn, has triggered a rise in direct-tax collections and propelled an increase in consumer spending.

Applications of Information Technology

Every day, people use computers in new ways. Computers and other electronic devices are becoming increasingly affordable. They continue to be more powerful as information-processing tools as well as easier to use. Humans are continually becoming dependant on IT-

enabled devices for carrying out simple tasks like remembering a phone number to complex ones like flying a fighter plane. Information Technology has applications in almost all aspects of our life. Some of the important ones are:

Science and Engineering: Scientific progress in fields like biotechnology is almost entirely dependent on the use of computers and other microprocessor-controlled devices. Using supercomputers, meteorologists predict future weather by using a combination of observations of weather conditions from many sources, a mathematical representation of the behavior of the atmosphere, and geographic data. Computer-aided design (CAD) and computer-aided manufacturing (CAM) programs have led to improved products in many fields, especially where designs tend to be very detailed. Computer programs make it possible for engineers to analyze designs of complex structures such as power plants and space stations.

Financial system(Business &Commerce): One of the first and largest applications of computers is keeping and managing business and financial records. Most large companies keep the employment records of all their workers in large databases that are managed by computer programs. Similar programs and databases are used in business functions like billing customers; tracking payments received and payments to be made; and tracking supplies needed and items produced, stored, shipped, and sold. In fact, practically all the information companies need to do business involves the use of computers and Information Technology. Almost all the financial transactions in the world are done electronically. Newer technologies like m-commerce have enabled almost everybody to carry out routine financial transactions on the move.

On a smaller scale, many businesses have replaced cash registers with point-of-sale (POS) terminals. These POS terminals not only print a sales receipt for the customer but also send information to a computer database when each item is sold to maintain an inventory of items on hand and items to be ordered. Computers have also become very important in modern factories. Computer-controlled robots now do tasks that are hot, heavy, or hazardous. Robots are also used to do routine, repetitive tasks in which boredom or fatigue can lead to poor quality work.

With today's sophisticated hardware, software, and communications technologies, it is often difficult to classify a system as belonging uniquely to one specific application program. Organizations increasingly are consolidating their information needs into a single, integrated information system. Management Information System (MIS), with the Chief Information Officer (CIO) at its head, is a whole, new branch of enterprise management.

Education: The advent of Information Technology has changed the meaning of the term "literate", with computer literacy being almost as important as basic literacy in many cases. Computer education is an essential course at the primary level in most schools across the world. With more information getting digitized every day, and the internet making it accessible to anyone across the world, students are increasingly relying on electronic sources of information rather than physical libraries for their needs. Instructional methodology has also undergone a sea change with use of images, animations, videos, presentations and e-learning to complement traditional techniques.

Governance: The concept of e-governance is one of the most novel applications of Information Technology whereby it is changing the lives of millions across the globe. Computerization of Government activities makes it easier to supervise and audit, and makes the administration more responsive to the needs of society. It also bridges the divide between the Government and the people. Technologies like touch-screen kiosks help disseminate information on land records, photo identity cards, pending bills etc. and enable even illiterate people to take more informed decisions. India is leading the world in the effective use of IT for elections.

Health: Information Technology plays an important role in medicine. For example, a scanner takes a series of pictures of the body by means of computerized axial tomography (CAT) or magnetic resonance imaging (MRI). A computer then combines the pictures to produce detailed three-dimensional images of the body's organs. In addition, the MRI produces images that show changes in body chemistry and blood flow. Most critical life support equipment are programmed to respond to changes in the patient's status in split-seconds, thereby reducing the response time and risk of human error. Newer concepts like robotic surgery enable specialists to perform surgeries from remote locations. Genomic studies greatly depend on supercomputing power to develop technologies for the future.

Entertainment: IT has changed the lifestyle of most people. The convergence of various technologies has created various options for entertainment like games, streaming music and video, digital television broadcasts, satellite radio, animated movies etc. which can be accessed with the help of mobile phones, PDAs, notebook computers or on television either with a cable connection or wirelessly using newer-generation WiFi, CDMA or GPRS technologies.

Information Technology plays a vital role in most of our daily activities. There is hardly anyone who has not been affected or influenced by IT. With each passing day, newer applications of IT are being developed which increase our interaction with and dependence on IT-enabled devices. Therefore, understanding this technology and using it creatively is imperative to human progress

SECURITY ISSUES IN IT

Information security, sometimes shortened to **InfoSec**, is the practice of defending information from unauthorized access, use, disclosure, disruption, modification, perusal, inspection, recording or destruction. It is a general term that can be used regardless of the form the data may take (electronic, physical, etc...

Two major aspects of information security are:

- **IT security**: Sometimes referred to as computer security, Information Technology Security is information security applied to technology (most often some form of computer system). It is worthwhile to note that a computer does not necessarily mean a home desktop. A computer is any device with a processor and some memory (even a calculator). IT security specialists are almost always found in any major enterprise/establishment due to the nature and value of the data within larger businesses. They are responsible for keeping all of the technology within the company secure from malicious cyber attacks that often attempt to breach into critical private information or gain control of the internal systems.
- **Information assurance**: The act of ensuring that data is not lost when critical issues arise. These issues include but are not limited to: natural disasters, computer/server malfunction, physical theft, or any other instance where data has the potential of being lost. Since most information is stored on computers in our modern era, information assurance is typically dealt with by IT security specialists. One of the most common methods of providing information assurance is to have an off-site backup of the data in case one of the mentioned issues arise.

INDIAN IT ACT

The **Information Technology Act 2000** (also known as **ITA-2000**, or the **IT Act**) is an Act of the Indian Parliament (No 21 of 2000) notified on October 17, 2000. This act is being opposed by Save Your Voice campaign and other civil society organizations in India. User-review and consumer social networking site MouthShut.com has filed a writ petition in the Supreme Court of India to repeal and nullify parts of IT Act 2000.

Information technology Act 2000 consisted of 94 sections segregated into 13 chapters. Four schedules form part of the Act. In the 2008 version of the Act, there are 124 sections (excluding 5 sections that have been omitted from the earlier version) and 14 chapters. Schedule I and II have been replaced. Schedules III and IV are deleted.

Information Technology Act 2000 addressed the following issues:

1. Legal Recognition of Electronic Documents
2. Legal Recognition of Digital Signatures
3. Offenses and Contraventions
4. Justice Dispensation Systems for Cybercrimes

APLICATION OF INFORMATION TECHNOLOGY IN RAILWAYS

Indian Railways has been spearheading the use of Information Technology for providing better services. Information Technology was adopted in early 60′s when computerized passenger and Freight Revenue accounting, Operating Statistics, Payroll and inventory management were introduced using IBM 1401 computers placed in Zonal data centers. In the following years, Indian Railways has continued using technologies for making significant improvements in their services. Introduction of Online Passenger Reservation System is one such significant milestone. The online reservation system allows the users make reservations sitting at the comfort of their homes. The online reservation system also integrates other related services like irctc pnr status check.

Being the 4th largest Railway network in the world, there is lot of pressure on Indian Railways to use more and more state of the art technologies for making significant improvements in the current system. Recently, A Technology Mission on Indian Railways has been launched. The

effort is sponsored jointly by the Ministry of Human Resource Development, Ministry of Railways and a consortium of private industries. The goals and objectives of the mission include

1. Development and adoption of state-of-art safety and control technologies through projects aimed towards achieving higher throughput, lower cost of transmission and safer train movement

2. Encouraging and initiating R & D activities pertinent to Railways in academic institutions and laboratories and establish convergence and synergy among them

3. Dissemination of technologies through participatory approach to other application areas

This is a really good step by Indian Railways towards the right direction. The latest technologies related to Civil Engineering, Mechanical and Electrical Engineering, Telecommunications and material science must be harnessed to run trains effectively, especially in current competitive scenario.

Similarly information is a critical resource for train operations. Harnessing information technology for effective collection and dissemination of information is not a luxury but rather a necessity for any railway operator. The effective usage of IT saves costs rather than increasing expenditure in today's competitive environment.

APLICATION OF INFORMATION TECHNOLOGY IN AIRLINES

Passenger Services

- When you book a flight, regardless of the method, your reservation information is processed and stored by the airline's computer system. If you book this online, your registration information is directly stored with the company. If you book your flight over the phone, a customer service representative will enter this information for you. This computer-based reservation system allows you to easily modify travel arrangements at any airport, and even to use multiple airline companies over the course of a single trip.

Airport Services

- Many aspects of an airport rely heavily on computers. Security screening machines such as X-rays may not make use of personal computers, but they do rely on computer technology for a great deal of their operations. Furthermore, computers are necessary for the use of pre-screening measures such as the current U.S. counter-terrorism efforts.

Air Traffic

- Computers are crucial to an airport's air traffic control services. While the communication between air traffic controllers and pilots occurs through radio, the system by which this communication is synthesized with radar and weather data is based on computers. Computers allow air traffic controllers to visualize and track the location of planes in the air, and then instruct pilots as to the correct course of action.

Plane Mechanics

- An airplane depends on computers for a number of essential flight functions. The autopilot, which many modern airplanes are equipped with, frequently employs a computer. This relieves the pilots of many in-flight tasks, allowing the plane to process navigational data and flight control systems. Computerized "fly-by-wire" technology, which many airplanes have, transmits pilot or autopilot signals to a computer. This allows planes to respond to subtle atmospheric variables and make other adjustments easily. Additionally, some airplanes have wing flaps that can be controlled by a computer. Others have power systems that can be regulated by computer. Finally, computers are central to the efficient communication and recording of an airplane's flight data. Such data is vital to air traffic controllers, and greatly assists in the event of an emergency situation.

APLICATION OF INFORMATION TECHNOLOGY IN BANKING

The advantages accruing from computerization are three-directional - to the customer, to the bank and to the employee.

For the customer. Banks are aware of customer's need for new services and plan to make them available. IT has increased the level of competition and forced them to integrate the new technologies in order to satisfy their customers. They have already developed and implemented a certain number of solutions among them:

- Self-inquiry facility: Facility for logging into specified self-inquiry terminals at the branch to inquire and view the transactions in the account.
- Remote banking: Remote terminals at the customer site connected to the respective branch through a modem, enabling the customer to make inquiries regarding his accounts, on-line, without having to move from his office.
- Anytime banking- Anywhere banking: Installation of ATMs which offer non-stop cash withdrawal, remittances and inquiry facilities. Networking of computerized branches inter-city and intra-city, will permit customers of these branches, when interconnected, to transact from any of these branches.
- Telebanking: A 24-hour service through which inquiries regarding balances and transactions in the account can be made over the phone.
- Electronic Banking: This enables the bank to provide corporate or high value customers with a Graphical User Interface (GUI) software on a PC, to inquire about their financial

transactions and accounts, cash transfers, cheque book issue and inquiry on rates without visiting the bank. Moreover, LC text and details on bills can be sent by the customer, and the bank can download the same. The technology used to provide this service is called electronic data interchange (EDI). It is used to transmit business transactions in computer-readble form between organizations and individuals in a standard format.

- As information is centralized and updates are available simultaneously at all places, single-window service becomes possible, leading to effective reduction in waiting time.

For the bank. During the last decade, banks applied IT to a wide range of back and front office tasks in addition to a great number of new products. The major advantages for the bank to implement IT are:

- Availability of a wide range of inquiry facilities, assisting the bank in business development and follow-up.
- Immediate replies to customer queries without reference to ledger-keeper as terminals are provided to Managers and Chief Managers.
- Automatic and prompt carrying out of standing instructions on due date and generation of reports.
- Generation of various MIS reports and periodical returns on due dates.
- Fast and up-to-date information transfer enabling speedier decisions, by interconnecting computerized branches and controlling offices.

For the employees. IT has increased their productivity through the followings:

- Accurate computing of cumbersome and time-consuming jobs such as balancing and interest calculations on due dates.
- Automatic printing of covering schedules, deposit receipts, pass book / pass sheet, freeing the staff from performing these time-consuming jobs, and enabling them to give more attention to the needs of the customer.
- Signature retrieval facility, assisting in verification of transactions, sitting at their own terminal.
- Avoidance of duplication of entries due to existence of single-point data entry.

APLICATION OF INFORMATION TECHNOLOGY IN INSURANCE

Introduction of technology in the insurance sector has improved every aspect of the industry. Technologies play a major role in data management process of an insurance agency by providing flawless services from underwriting policies, producing documents to collating various ratings and data. The state-of the-art implementations offers instantaneous accurate information about different insurances to the clients. Insurance firms regularly spend a part of their yearly premiums on modern technology that aids in enhancing the overall performance of the organization. Insurance technologies help the insurance agents to immediately respond to the requirements of the customers and technology has managed to cut back the annual expenditure of the organizations.

The basic purpose of insurance technologies is to reduce the paperwork of proposals and policies and address the customer services effectively in a shorter time than any other traditional methods. Information technology in insurance has made it easier for the customers too. Online availability of the insurance agencies allow the clients in dealing with application procedures, signing proposals and policies as well as in receiving quotes without even visiting the insurance office in person.

The best part of technology in insurance is that it helps the firms in reducing the costs by eliminating the mail rooms, paper files as well as the data entry clerks. The elaborate underwriting, data processing and the rating take place online and the customers or brokers receive the emailed policy documents within no time. This online advantage, however, comes with a price. The system requires a substantial initial investment in its primary stages but the owner certainly gets his returns on investment over the years that come equipped with superior services and response timings.

Various innovative technological applications allow the insurers to recognize the risks and opportunities easily. The modeling device examines the loss histories and compares them with the risk characteristics while it searches for correlations. Such insurance technologies help the insurance companies to charge higher prices for the higher-risk client base and lower prices for the safer opportunities.

There is a variety of insurance technologies available in the market. The hardware and the insurance software should be chosen depending on the business necessities of the insurance agencies. Different insurance management systems and comparative rating systems enable the firms to generate more revenues by decreasing the span of output and input procedures.

Insurance technologies have made insurance services mobile with the availability of smart phones and such devices. Insurance companies use these devices to provide faster services like view policies, obtain quotes, and report claims through live chat application. Such improvements would have been impossible if there were no insurance technologies available within the industry.

APLICATION OF INFORMATION TECHNOLOGY IN INVENTORY CONTROL

Speed and Efficiency

A computerized inventory management system makes everything from inputting information to taking inventory easier. Doing a hand count of inventory can take days, but with a computerized inventory management system, the same process can be done in a matter of hours.

Document Generation

Once the computerized inventory management system is in place, managers and workers can use it to automatically generate all kinds of documents, from purchase orders and checks to invoices and account statements. Managers can also use the system to automatically order products when they run low.

Timely Data

With a manual system, the data is only as accurate and up to date as the last hand count. With a computerized inventory management system, the management team can pull a report and instantly see how many units are on the floor, how many have sold and which products are selling the fastest.

Reliance on Technology

With a computerized inventory management system, the company is at the mercy of its technology. Outside factors like a power failure or the loss of Internet or network connectivity can render the system temporarily useless.

Accuracy Issues

A computerized system alone does not ensure accuracy, and the inventory data is only as good as the data entry that created it. Companies that plan to use a computerized inventory management system need to have a system in place to validate their data and check the numbers reported by the system. A select hand count or targeted audit may be necessary to ensure the integrity of the system.

Risk of Fraud

Any computerized system carries the risk of intrusion, and with a computerized inventory management system comes the risk of fraud as well. A dishonest vendor could hack the system to receive payment for products never delivered, or a dishonest employee could redirect checks to themselves.

APLICATION OF INFORMATION TECHNOLOGY IN HOTEL MANAGEMENT

Traditionally, hotels were largely dependent on cards and paperwork at the front desk to keep in touch with old and current customers. They were largely at the mercy of the desires of

vacationers to arrive, and on their own efforts and staff to be ready for potential surges or long droughts of occupancy. Luckily, such inconvenience and old-fashioned methods are long since past, thanks to advances in information technology.

The first area in which information technology became important was in regards to billing. Old-fashioned paper-based book-keeping was time consuming and inefficient, and was not able to quickly tell a hotel owner what the situation of their hotel was. Luckily, advances in modern record keeping allow for a hotel owner to keep track of what they have on hand, how much of it they have, and how much it costs. Accounting is complicated, but advanced accounting software, especially that tailored to the unique needs of the hospitality industry, helps to enable hotel owners to make smart decisions. Services and products that are no longer used can be quickly cut off to save money, while those who show demand can be increased in quantity or modified so as to reduce the heavy usage.

Most hotels are familiar with booking rooms and reservations over the phone, but information technology has expanded well beyond that. Hotels can now work with various online travel companies and booking services to have their rooms booked online, with no need to employ expensive staff. This also allows a hotel to advertise their open rooms and special deals directly to persons who would be most likely to purchase them, instead of wasting lots of money advertising in an unfocused manner. High quality information technology thus allows for better arrangement and management of bookings in order to allow a hotel to better maximize occupancy, and to know in advance when large groups or lean times are approaching. This allows a hotel manager to make plans regarding temporary staff, good times to renovate or expand, or other concerns, because he/she can determine the state of their hotel currently and for the next few months with only a few clicks on the computer.

The advances in information technology extend well beyond booking, however. The internet is essential for vacationers who wish to contact those back home, and for those traveling on business to get in touch with the office. Therefore, wireless internet has become a very common and very useful service for hotels to provide. Many business minded persons even require that a hotel offer internet services so that they can keep working while on the road. Luckily, such services are easy to provide, as all that is required is a wireless router and various devices to ensure the entire hotel is filled with the network. Modern advances in wireless internet also allow for the wireless internet provided for hotel visitors to be used to network the hotel itself. Security cameras, door locks, and other devices essential to hotel security and safety can be wired into the network, so that staff are alerted whenever a door is propped open, a fire alarm goes off or suspicious activity occurs. Though the hotel guests are wholly unaware of it, this sort of added safety and security keeps them safe, and in the event of a problem they will most certainly appreciate the benefits of such a system.

As advanced as it is, information technology in the hospitality industry is still going forward. Intelligent booking systems enable rapid and efficient guest feedback, along with the ability to predict who is likely to use the hotel again and inform them via e-mail or text messages when good deals arrive. Hotels with room service or other guest services can offer their menus online, allowing for quick updates, high-quality photos, and other ways to allow guests to see and order services before they even arrive. There are also advances in terms of payroll and inventory which make information technology a valuable asset for saving money and maximizing profits. The unique nature of the hospitality industry makes it a great place for new and emerging information technology, and forward-thinking hotel owners and managers are always looking for smart equipment and software to invest in.

Reference:

- Sinha, Kr. Pradeep & Sinha, Preeti, "*Foundations of Computing*"
- Leon and Leon, "Introduction to Information Technology"
- V. rajaraman," fundamental Of computers"
- http://www.di.unisa.it
- www.os-book.com
- www.cs.uah.edu