Biju Patnaik Institute of IT & Management Studies

Lecture Notes

on

IT FOR MANAGEMENT

Module-I

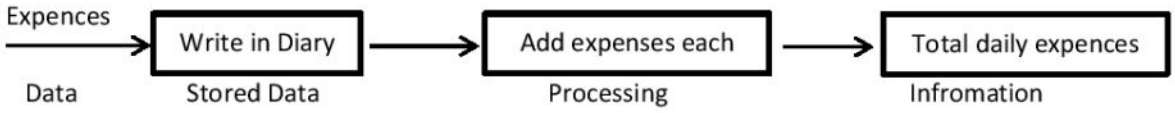
Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Introduction to Information Technology

Information Technology (IT) may be defined as the technology that is used to acquire, store, organize, process and disseminate processed data which can be used in specified applications.

Information is processed data that enable us to take decisions and initiate actions.

Example: A home maker byes groceries and write in a diary about her expences every day.



The hardware and software of computing and communication form the basic too of technology.

COMPUTER

A computer is an electronic device that operates under the control of a set of instructions that are stored in its memory units. A computer accepts data from an Input device and process in to useful information, which it displays on its output device.

APPLICATIONS OF IT IN BUSINESS AND MANAGEMENT

IT has become a vital and integral part of every business plan. From multi-national corporations who maintain mainframe systems and databases to small businesses that own a single computer, IT plays a role. The important applications of information technology in the field of business and management are given as follows.

1. **Business Communication** : Now a day's companies are using email, live chat system, online meeting tools, video conferencing system and smart phones as the principal means of communication between employees, suppliers and customers. Communication by email is faster and costs less than sending a paper letter in the mail.
2. **Product Development :** Information technology can speed up the time it takes new products to reach the market. Companies can now understand the requirements of consumers by collecting marketing intelligence from proprietary databases, customers and sales representatives. IT helps businesses respond quickly to changing customer requirements.
3. **Process Improvement** : Process improvement is another important IT application in business. Enterprise resource planning (ERP) system allow managers to review sales, costs and other operating figures on one integrated software platform, usually in real time.
4. **Marketing :** One of the main applications of IT is in the area of marketing. Both large and small businesses can now play on a same level and status on the internet. They can have a web site, take orders, buy goods, sell excess or even operate some businesses entirely online.
5. **Inventory Management :** IT helps business to manage inventory effectively. Organizations are now able to maintain enough stock to meet demand without investing in more than they require. Inventory management systems track the quantity of each item a company maintains, placing an order of additional stock when the quantities fall below a pre-determined reorder level.
6. **Customer Relationship Management :** Companies are using IT to design and manage customer relationships. Customer Relationship Management (CRM) systems capture every interaction a

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

company has with a customer. The entire interaction is stored in the CRM system, ready to be recalled if the customer calls again.

7. **Data Management :** Through IT, companies are able to store and maintain a tremendous amount of historical data economically and employees benefit from immediate access to the documents they need.
8. **Management Information Systems (MIS) :** Storing data is only beneficial if that data can be used effectively. MIS enable companies to track sales data, expenses and productivity levels. The information can be used to track profitability over time, maximize return on investment and identify the areas of improvement.
9. **Globalization :** IT is at the core of operating models essential for globalization, such as telecommuting and outsourcing. A company can outsource most of its noncore functions, such as human resources and finances, to offshore companies and use network technologies to stay in contact with its overseas employees, customers and suppliers.
10. **Competitive Advantage :** Cost savings, rapid product development and process improvements help companies gain and maintain a competitive advantage in the market place. Companies can use rapid prototyping, software simulations and other IT based systems to bring a product to market cost effectively and quickly.
11. **Cost Efficiencies :** Although the initial IT implementation costs can be substantial, that resulting long term cost savings are usually worth the investment. IT allows companies to reduce transaction and implementation costs.

THE APPLICATION OF COMPUTERS IN VARIOUS FIELDS

Business:

- Payroll calculations
- Budgeting
- Sales analysis
- Financial forecasting
- Managing employee database
- Maintenance of stocks, etc.

Banking : Today, banking is almost totally dependent on computers.
Banks provide the following facilities –

- Online accounting facility, which includes checking current balance, making deposits and overdrafts, checking interest charges, shares, and trustee records.
- ATM machines which are completely automated are making it even easier for customers to deal with banks.

Insurance: Insurance companies are keeping all records up-to-date with the help of computers. Insurance companies, finance houses, and stock broking firms are widely using computers for their concerns.
Insurance companies are maintaining a database of all clients with information showing –

- Procedure to continue with policies
- Starting date of the policies
- Next due installment of a policy
- Maturity date
- Interests due
- Survival benefits
- Bonus

Education: The computer helps in providing a lot of facilities in the education system.

- The computer provides a tool in the education system known as CBE (Computer Based Education).

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- CBE involves control, delivery, and evaluation of learning.
- Computer education is rapidly increasing the graph of number of computer students.
- There are a number of methods in which educational institutions can use a computer to educate the students.
- It is used to prepare a database about performance of a student and analysis is carried out on this basis.

Marketing: In marketing, uses of the computer are following –

- **Advertising** – With computers, advertising professionals create art and graphics, write and revise copy, and print and disseminate ads with the goal of selling more products.
- **Home Shopping** – Home shopping has been made possible through the use of computerized catalogues that provide access to product information and permit direct entry of orders to be filled by the customers.

Healthcare : Computers have become an important part in hospitals, labs, and dispensaries. They are being used in hospitals to keep the record of patients and medicines. It is also used in scanning and diagnosing different diseases. ECG, EEG, ultrasounds and CT scans, etc. are also done by computerized machines.
Following are some major fields of health care in which computers are used.

- **Diagnostic System** – Computers are used to collect data and identify the cause of illness.
- **Lab-diagnostic System** – All tests can be done and the reports are prepared by computer.
- **Patient Monitoring System** – These are used to check the patient's signs for abnormality such as in Cardiac Arrest, ECG, etc.
- **Pharma Information System**–Computer used to check drug labels,expiry dates,harmful side effects etc.
- **Surgery** – Nowadays, computers are also used in performing surgery.

Engineering Design: Computers are widely used for Engineering purpose.
One of the major areas is CAD (Computer Aided Design) that provides creation and modification of images. Some of the fields are –

- **Structural Engineering –** Requires stress and strain analysis for design of ships, buildings, budgets, airplanes, etc.
- **Industrial Engineering –** Computers deal with design, implementation, and improvement of integrated systems of people, materials, and equipment.
- **Architectural Engineering –** Computers help in planning towns, designing buildings, determining a range of buildings on a site using both 2D and 3D drawings.

Military : Computers are largely used in defense. Modern tanks, missiles, weapons, etc. Military also employs computerized control systems. Some military areas where a computer has been used are

- Missile Control, Military Communication, Military Operation and Planning, Smart Weapons

Communication : Communication is a way to convey a message, an idea, a picture, or speech that is received and understood clearly and correctly by the person for whom it is meant. Some main areas in this category are –

- E-mail
- Chatting
- Usenet
- FTP
- Telnet
- Video-conferencing

Government : Computers play an important role in government services. Some major fields in this category are –

- Budgets
- Sales tax department
- Income tax department
- Computation of male/female ratio
- Computerization of voters lists
- Computerization of PAN card
- Weather forecasting

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

BRIEF HISTORY OF COMPUTERS

The computer as we know it today had its beginning with a 19th century English mathematics professor name Charles Babbage. He designed the Analytical Engine and it was this design that the basic framework of the computers of today are based on.

History of computers has been provided in a chronological order from the beginning of 13th Century till now.

Abacus: Abacus is known to be the first mechanical calculating device, which was used to be performed addition and subtraction easily and speedily. This device was invented in the 600 B.C. Abacus is made up of wooden frame in which rod where fitted across with rounds beads sliding on the rod. It is divided into two parts called 'Heaven' and 'Earth'. Heaven was the upper part and Earth was the lower one. Thus any number can be represented by placing the beads at proper place.

NAPIER'S BONES: As the necessity demanded, scientist started inventing better calculating device. In thus process John Napier's of Scotland invented a calculating device, in the year 1617 called the Napier Bones. In the device, Napier's used the bone rods of the counting purpose where some number is printed on these rods. These rods that one can do addition, subtraction, multiplication and division easily.

PASCAL'S CALCULATOR: In the year 1642, Blaise Pascal a French scientist invented an adding machine called Pascal's calculator which represents the position of digit with the help of gears in it.

LEIBNZ CALCULATOR: In the year 1671, a German mathematics, Gottfried Leibniz modified the Pascal calculator and he developed a machine which could perform various calculation based on multiplication and division as well.

ANALYTICAL ENGINE: In the year 1833, a scientist form England knows to be Charles Babbage invented such a machine, which could keep our data safely. This device was called Analytical engine and it deemed the first mechanical computer. It included such feature which is used in today's computer language. For this great invention of the computer, Sir Charles Babbage is also known as the father of the computer.

His design contained the five key characteristics of modern computers i.e. An input device, Storage, A processor or number calculator, A unit to control the task and the sequence of its calculations, An output device.

Some of the early well-known computers are:

The Harvard Mark I: The Harvard Mark I, also known as the IBM Automatic Sequence Controlled Calculator is said to be a monument in the development of the art of computing. Mark I was designed in 1937 by a Harvard graduate student, Howard H. Aiken to solve advanced mathematical physics problems encountered in his research. The machine was capable of multiplying two 23 digit numbers in 6 Seconds.

The Atanasoff–Berry Computer (ABC): It was the first automatic electronic digital computer developed at Bell Telephone Laboratories and demonstrated at American Mathematical Society in 1940. The complex computer or Model-2 was capable of adding, subtracting, multiplying & dividing two complex numbers with very high reliability of operation.

The ENIAC: In 1946, J. Presper Eckert and John W. Mauchly invented giant ENIAC machine at the University of Pennsylvania. ENIAC stands for Electronic Numerical Integrator and Calculator was the first machine to use large number of vacuum tubes. The machine required a big space and lot of energy to keep it cool. Further, it had punched-card input and output. The instructions had to be fed in to the machine by way of switches because there was no internal memory within the machine.

The EDVAC (1946-52) : The EDVAC was a serial binary machine and economical than ENIAC in terms of electronic tubes. The other major difference was that EDVAC had a much larger memory than ENIAC. Therefore, with EDVAC, the invention of the modern digital computer was complete. This became the source of inspiration for the vast number of computer projects during the late 1940.

Harvard Mark II 1948 : In this series, in 1948 Harvard Mark II was designed by Aiken based on relay technology. As relay lent themselves to complex circuit arrangements, therefore, floating point arithmetic operation could be performed which was not possible earlier. Later J. Presper Eckert and John W. Mauchly started the concept of large scale electronic computer at the Moor School of Electrical Engineering in Philadelphia.

Since 1950, computers have advanced at a pace unparalleled in the history of technology. Processing speed and memory capacity have increased; size and cost have decreased by several orders of magnitude.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

THE GENERATIONS OF COMPUTERS

In 19th Century, after invention of Vacuum Tube, Transistor, Register, Capacitors, Diodes the Electronic computers are invented. The Electronic computers in generation wise are as follows.

Generation in computer terminology is a change in technology a computer is/was being used. It also reflects the growth of computer Industry. The advancement in technology existed not only in hardware but also software. There are five computer generations known till date.

The First Generation (1950-1960):

The First Generation of computer began around 1950, with the introduction of commercial computers manufactured and sold in quantity. Computers of the first generation stored their programs internally and used Vacuum Tubes as their switching technology. The main computers introduced during this era were UNIVAC I and IBM 704, which could be programmed in assembly language.

UNIVAC (Universal Automatic Computer). It was developed by two scientists **J. Presper Exkert** and **John Mauchly,** at the Census Department of United States and was first delivered in 1951. Most customers used the UNIVAC for accounting, statistical, and other applications that would later be known as data processing.

The UNIVAC used binary-coded decimal (BCD), arithmetic performed in four general-purpose accumulators. Word length was 45 bits, each word could represent 11 decimal digits plus a sign, or 6 alphabetic characters (6 bits per character plus 1 parity bit). Basic clock speed was 2.25 MHz, and the multiplication time was about 2ms.

The UNIVAC was ruggedly designed and built. Its central processor contained over 5,000 tubes, installed in cabinets that were arranged in a 10-foot (3m) by 14-foot (4.5m) rectangle.

IBM 704 IBM also initiated development of a smaller machine, whose origin lay in proposals for extensions of punched card equipment. In late 1955, IBM began deliveries of the 704, its successor to the 701. The 704's most notable features were its core memory. The 704's processor had hardware floating-point arithmetic and three addressable index registers. In January 1957, IBM announced the **709** as a compatible upgrade to the 704, but it did not enjoy the same success.

Machines like the **ERA 1103** also belonged to first generation and they were too big, too expensive, and required too much specialized programming skill to be found anywhere. But, it was found at the largest aerospace firms or Government research laboratories.

Characteristics of First Generation Computers

a) Vacuum tubes were the only electronic components available during those days. Vacuum tube technology made possible the advent of electronic digital computers.
b) As more than ten thousands of vacuum tubes were there in first generation computers, power consumption was very high since each of vacuum tube required nearly half a watt of power.
c) Vacuum tubes emitted large amount of heat, due to this there were many frequent burnt outs of vacuum tubes. This led to frequent hardware failures. Therefore, constant maintenance and proper air conditioning was required.
d) These computers were the fastest calculating devices for the time. They could perform computations in milliseconds.
e) Computer during this era was non-portable. As they were too bulky in size, required large area for installation.
f) These computers required manual assembly of thousands of components for individual functioning unit. Therefore, commercial production was difficult and costly.
g) These computers had limited commercial use, because they were difficult to program.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

The Second Generation (1960-1965):

The second generation of computing lasted from about 1960 to 1965, and was characterized by discrete transistors for switching elements and ferrite magnetic core planes for internal memory. The scientists at Bell laboratories developed **transistor** in 1947. These scientists include John Barden, William Brattain and William Shockley. Transistors, replacing vacuum tubes on a one-to-one basis, solved the problems of a tube's unreliability, heat and power consumption. As they solved those problems, they exposed another problem, which proved to be more fundamental. That was the complexity of inter-connecting many thousands of simple circuits to obtain a complete system. Some manufacturers laboured under the burden of hiring and training workers to hand-wire and solder the components to one another.

In software, this era saw the acceptance of high-level programming languages like Fortran and Cobol, although assembly language programming remained common.

The most important transistorized computers were **IBM 7094 series**, **IBM 1400 series** and **CDC 164** etc. These machines employed a character-oriented processing.

Characteristics of Second Generation Computer

(a) The computation speed of these computers was faster than first generation computer. Thus, computation time was reduced from millisecond to microsecond.
(b) The size of second generation computers were smaller than first generation computers, thus, required smaller space for installation.
(c) Although heat emission was less, computers were manufactured using transistors. However, air-conditioning was required.
(d) Frequent maintenance required, but less prone to hardware failures than first generation computers.
(e) In comparison to first generation computers, they were easy to program and use, thus they were used commercially.
(f) These computers also required manual assembly of individual components into a functioning unit. Therefore, commercial production was difficult and costly.

The Third Generation (1965-1970):

The Third generation computers used the **Integrated circuits (IC)** technology. Jack Kilby developed the concept of integrated circuit in 1958. A single IC chip may contain thousands of transistors. A second innovation was the use of microprogramming to achieve compatibility obtained instruction set from a read-only memory containing a micro program. The **IBM system 360** used channels for I/O - independent processor that handled the transfer of data between primary memory and peripheral devices. When IBM upgraded the 360 series to the **System 370** in 1970, its architecture was extended to provide this feature and virtual memory as well. The semi-conductor memory loses its information when power is switched off, the S/370 needed a way to store its micro programmed instructions in a non-volatile fashion. IBM engineers invented the **floppy disk** for this purpose. The floppy became the pivotal technology for establishing the personal computer class later in that decade.

The 360 adopted the 8-bit byte as the standard for representing characters, and it used multiple-spindle disk systems with removable disk packs. Micro-programming soon became the most common way to implement architectures.

Characteristics of Third Generation Computers

(a) These computers were smaller in size as compared to first generation. Therefore, easily portable and required lesser space for installation.
(b) These computers were able to reduce computational times from milliseconds to nanoseconds.
(c) These computers had less power requirement and generated less heat than previous generation computers.
(d) Maintenance cost was low because hardware failures were rare.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(e) Commercial production was cheaper and easier. This is because of the fact that manual assembly of individual components into a function unit was not required. Thus, human labour and cost involved at assembly stage reduced considerably.
(f) Even more reliable than second generation computers.
(g) These computers were totally general purpose. They were used not only for commercial applications but also for scientific applications all over the world.
(h) Highly sophisticated technology required for the manufacture of IC chips.
(i) The advancement in technology was not only in hardware but also in software, such that operating system and high level programming languages were also evolved during this time.
(j) Operating systems and other application software allowed the users to work in time-sharing and real time on-line environment.
(k) The standardization of operating system and programming languages allowed the user to develop and execute their programs even at remote locations.
(l) Minicomputers of the third generation computer were also economical and easily affordable even by smaller companies.

The Fourth Generation (1970-1989):

It was Large Scale Integration (LSI) era (Refer to key terms at end of this Chapter), when electronic components on a silicon chip was doubled, as it was based on semi-conductor technology. During this time it was possible to integrate more than 30,000 electronic components on a single chip. This progress led to Very Large Scale Integration (VLSI), where integration of more than one million electronic components on a single chip was possible. The fourth generation computers started with the invention of Microprocessor. The Microprocessor contains thousands of ICs. Ted Hoff produced the first microprocessor in 1971 for **Intel.** It was known as Intel 4004. The microprocessor is a single chip which contains the needed circuit, it could perform arithmetic and logical functions. This further led to personal computer revolution. As personal computers (PCs) were amazingly small, compact and inexpensive, it was possible for everyone to have their own computer. During this era, it was not only the work place where computers were deployed but also at homes. Initially, Apple Computer Inc. launched the PCs, Apple IT in 1978. Then in 1980, PCs market was dominated by IBM, popularly known as IBM Pc. During fourth generation period, several other manufacturers designed their own PCs known as IBM compatible PCs, as this became standard for the PC industry.

In the beginning of this era, the memory sizes were of the order 16 Megabyte with random access memory having fast access time of 200 nanosecond whereas; secondary storage devices were smaller in size and larger in capacity. During the fourth generation period, improved processing and storage capabilities of mainframe led to the advent of supercomputers. The first supercomputer manufacturers were Cray Research and ETA system. Later on IBM, Silicon Graphics and Digital Equipment Corporation built the supercomputers with multiprocessing capability.

Achievements in Hardware and Software in Fourth Generation Computers

(a) The PCs were the most important achievement among all other computers (viz mainframe, minicomputers and supercomputers) during fourth generation period. However, highly sophisticated technology was required for the manufacture of LSI chips.
(b) PCs were smallest in size because of high component density, and powerful in computing capability.
(c) No air-conditioning required for PCs, as heat generated is negligible. However, air-conditioning was essential for supercomputers, mainframe and minicomputers.
(d) Less power required for this generation computers.
(e) These computers were more reliable, less prone to hardware failure. Therefore, maintenance cost is negligible.

(g) Much faster in accessing speed and computation than previous generations.
(h) High storage capability.
(i) Easily portable because of small size.
(j) These computers were general purpose machines used both for commercial, scientific, art design, etc. It also supported multi-media applications.
(k) Minimal labour and cost at assembly stage. Therefore, cheapest among all generations and easily affordable by masses.
(l) This generation computers hardware supported Graphic User Interface (GUI). It made easy learning computers and its applications.
(m) The development of programming languages such as C and C++ (object oriented language) during this period allowed the use of the system effectively, in further software and hardware development.
(n) The networking of computers wireless technology and internet was possible during this period. It enabled haring of resources like disk, printer, information, etc.
(o) As the information from any part of the world is readily available to the computer user through World Wide Web (WWW), several new applications such as electronic commerce, virtual libraries, virtual classroom, etc., emerged during this period.

The Fifth Generation (1980-till date):

In the fifth generation, VLSI technology became ULSI (Ultra Large Scale Integration) technology, resulting in the production of microprocessor chips having ten million electronic components. According to the Moor's law, the speed and size of the main memory and of hard disk, doubled in almost every eighteenth month. This progress led to the powerful features, which were generally used to be found in mainframes and will be available in microprocessor architecture in the fifth generation computers.

This generation is based on parallel processing hardware and AI (Artificial Intelligence) software. AI is an emerging branch in computer science, which interprets the means and method of making computers think like human beings. All the high-level languages like C and C++, Java, .Net etc., are used in this generation. Some computer types of this generation are Desktop, Laptop, NoteBook, UltraBook, ChromeBook

Characteristics of Fifth Generation Computers (now and future)

(a) The common opinion is that future fifth generation computer will be based on artificial intelligence allowing them to learn think and reason.
(b) Voice recognition will likely to be a primary means of input. Computers may be constructed differently than they are today.
(c) The size of the personal computers (called notebook computers) will continue to be small and will have the feature of parallel processing.
(d) They will have significantly faster and larger primary memory and auxiliary memory compared to previous generations.
(e) These machines will be used anywhere and for every purpose.
(f) More complex problems can be solved which were not possible to solve earlier.
(g) Due to such fast advancement in computer technology, their manufacturing does not require manual assembly of individual components. Therefore, commercial production of these systems will be easier and cheaper. However, very few organizations are manufacturer of ULSI chips because to manufacture ULSI Chip highly sophisticated technology and expensive setup is required.
(h) The communication technology during fifth generation period supported tremendous growth of computer networks. This resulted in the emergence of the Internet and associated technologies and applications.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

CLASSIFICATION OF COMPUTERS

Information can be stored and processed in number of ways. A computer may be digital or analog or Hybrid. They do have some principles in common, but they employ different types of data representation and are in general, suited to different kind of work. The way computers process the information can be broadly classified as

a) Analog Computers
b) Digital Computers
c) Hybrid Computers.

a) Analog Computer: An analog computer is a computer which is used to process analog data. Analog computers store data in a continuous form of physical quantities and perform calculations with the help of measures. It is quite different from the digital computer, which makes use of symbolic numbers to represent results. Analog computers are excellent for situations which require data to be measured directly without converting into numerals or codes. Analog computers, although available and used in industrial and scientific applications like control systems and aircraft, have been largely replaced by digital computers due to the wide range of complexities involved.

b) Digital Computers: Digital computer operates on digital data such as numbers. It uses binary number system in which there are only two digits 0 and 1. Each one is called a bit. The digital computer is designed using digital circuits in which there are two levels for an input or output signal. These two levels are known as logic 0 and logic 1. Digital Computers can give more accurate and faster results. Digital computer is well suited for solving complex problems in engineering and technology. Hence digital computers have an increasing use in the field of design, research and data processing. The main characteristic of Digital computer are that it is automatic, general purpose and electronic.

Differentiate between Analog and Digital Computers

Analog Computers	Digital computers
Analog Computers Work on continuous values.	Digital computers Work on discrete values.
Analog Computers have low memory.	Digital computers have a very large memory
Analog computers have Slow speed.	Digital computers have fast speed.
Analog computers are less reliable.	Digital computers are more reliable.
Analog computers used in engineering and science and medical fields.	Digital computers are used in all fields of life.
Analog computers are used to calculate / measure analog quantities like speed and temperature.	Digital computers are used to calculate mathematical and logical operations. It can solve addition, subtraction, division, multiplication and other mathematical and statistical operations.
Analog computers provide less accurate results.	Digital computers provide 100% accurate results.
Normally Analog Computers are specific purpose	Digital Computers are general purpose
Analog computers are difficult to use	Digital computers are easy to use
Examples of Analog computers are: thermometer, analog clock, speedometer etc.	Examples of digital computers are: Personal Computer, laptops, smart phones etc.

c) Hybrid computer: A hybrid computer combines the desirable features of analog and digital computers. It is mostly used for automatic operations of complicated physical processes and machines. Now-a day's analog-to-digital and digital-to-analog converters are used for transforming the data into suitable form for either type of computation. For example, in hospital's ICU, analog devices might measure the patient's temperature, blood pressure and other vital signs. These measurements which are in analog might then be converted into Numbers and supplied to digital components in the system. These components are used to monitor the patient's vital sign and send signals if any abnormal readings are detected. Hybrid computers are mainly used for specialized tasks.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Clasification of Digital computers

Digital computers can be classified in different ways- by size, by function, propose and/or by processing capacity. All computers, no matter how small or large they are, generally have the same functionalities, i.e., processing, storage, input and output.

Computers are generally divided into the following classes:

(1) Supercomputer
(2) Mainframe
(3) Minicomputer
(4) Workstations
(5) Microcomputer.
(6) Portable computer

1. SUPER COMPUTERS

Super Computer is high speed, high performance, multi-processing system which contains more than one processor; hence, more than one job can be carried out simultaneously. Moreover, it also uses parallel processing technologies to solve complex problems. Parallel processing enables a complex problem to be divided into smaller problems that can be assigned to different processors of the system and processed in parallel.The most powerful computer of any time having high speed, high performance and high end computers are called Super Computers. The first supercomputer '**Cray**' was invented by **Seymour Cray**.

Terminals are used by the user to communicate with the computer. This is possible by a predefined set of rules for how to communicate, known as a protocol. Different terminals have different characteristics, so the computer must have some way of knowing what type of terminal it is communicating with. A PC (Personal Computer) can be used as a terminal with communication software such as kermit.

The speed of modern supercomputers is measured in megaflops, gigaflops and teraflops. A gigaflop is 108 floating point arithmetic operations per second, and a teraflop is 1012 floating point arithmetic operations per second. Super Computers are massively parallel processor (MPP) computers.

Example of one of the fastest supercomputer in the world is IBM Blue Gene/L, contains 65,536 nodes containing 131,072 CPUs and 65,536 nodes (2 CPUs for each node). It cost $ 100 million, perform at speed 360 teraflops, runs Linux operating system. It is used to conduct nuclear weapon simulation.

Uses : Users of these computers are governmental agencies, such as National Weather Service, and the National Defense Agency. Also, they are used in the making of movies, space exploration, and the design of many other machines.

Advantages : They are used solely for number crunching. They are the most powerful, the most expensive, and the fastest. They are capable of processing **trillions of instructions per second**. Supercomputers are used for tasks that require numerous data manipulations.

2. MAINFRAME

Mainframe computer is a high-level large and expensive computer, designed to handle intensive computational tasks. Mainframe computers have the ability to simultaneously support many users connected to the computer by terminal. The term mainframe comes from the use of "frame" as a device to hold electronics (rack is also frequently used), and the frame holding the electronics. The UNIVAC I, optimized for business data processing, was an early mainframe.

Mainframes are often 'servers' or host computers, computers that control the networks of computers. They are multi-processor having multiple CPUs. Host and other computers are located in the systems room to which entry is restricted to administrators and maintenance staff. One or more console terminals are

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

also located in the system room. The mainframe generally has many indicator lights (sometimes as part of operator's console) to show fault conditions, memory contents, etc. The central processor and main memory are housed together as an aid in increasing processing speeds (cable lengths will be short) and improving reliability (e.g. both will be at similar temperatures and humidity). For hard copy of outputs, mainframes generally have several types of printers and plotters which are connected to back-end.

The Processing speed of mainframe computers is measured in MIPS. It stands for millions of instructions per second i.e. 200-1200 MIPS.

Uses : Mainframes are designed for multiple users and process vast amount of data quickly. Banks, insurance companies, manufacturers, mail- order companies and airlines are typical users or mainframe computer technology.

Advantages

- Mainframe computers process data at very high rate of speed, measured in the millions of instructions per second. They are very expensive, costing millions of dollars in some cases.
- Using mainframe technology a large amount of data can be processed and hundreds of users may be able to carry out lots of different tasks at the same time. Often not only is it necessary to have a large processing unit, but also to be able store large volumes of data and produce many different reports.
- It is often more secured and convenient to have all this processing storage and input-output capability located at one secure site.

3. MINICOMPUTERS

Minicomputer refers to a mid-level computer built to perform complex computations. It also deals efficiently with a high level of Input and Output from users connected via terminals. Minicomputers are used merely in transaction processing applications and as interfaces between mainframe computer system and in Wide Area Networks.

Until about 1968, most computers were mainframes, which were expensive and required considerable floor space, electric power and air conditioning. A typical minicomputer is micro programmed and transfers data internally over a high-speed channel called a bus. To gain access to more memory, they could be directly addressed by a short word and their central processor contains sets of registers for indirect, indexed, and other types of addressing. Minicomputers could be broadly classified as having 12-, 16-,18-,24-, or 32-bit word lengths with memory sizes of 128 MB or more to one gigabyte provided in modules of 256 KB or 1 MB or more. Nearly all minicomputers employ a parallel internal processor structure with a high- speed bus and a clock rate of 4-40 MHz. The basic configuration ranges in price from $4,000-$500,000, with the cost of peripheral devices usually far outstripping the cost of the machine. The use of low-cost LSI and VLSI logic removes many of the initial design constraints.

Uses

- The greatest use of minicomputers is in areas other than general-purpose computing. These areas include: Industrial applications, such as control of power generation, petrochemical systems, data acquisition, and testing of equipment and devices.
- Biomedical control for experiment monitoring.
- As a part of larger computer systems which require communication and peripheral control, such as data concentrators, satellite peripherals, and intelligent terminals.

Advantages

- A mini is a compact, solid-state computer with random access memory whose internal structure is characterized by a short word length and a variety of memory addressing modes.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- A minicomputer is small and rugged enough to fit in a standard equipment rack and thus serve as an embedded controller for other systems.
- A minicomputer is low in cost, small in size, and intended for use by a single individual, small department, or for a dedicated application.
- Intelligent graphics terminals and interactive graphics systems are part of general-purpose, graphics-oriented minicomputers. IBM also introduced minicomputers, but the smaller firms propelled the industry.

4. WORKSTATION

Workstation is a hardware which allows the user to interact with a computer system; it may be main, mini or micro. Basically, a workstation is a powerful desktop computer designed to meet the computing needs of engineers, architects and other professionals. It has a powerful high resolution colour graphic displays with hardware support for computationally intensive operation such as computer animation and CAD/CAM (computer aided designing/ computer aided manufacturing) application. It has greater processing power and larger storage capacity than any average PC, and is intended for a single user. Processors of workstations are designed using special technology known as Reduced Instruction Set Computers (RISC). RISC based architecture increases the speed of the processor to a great extent.

Uses

- Workstation operating system is usually designed to support multiuser environment.
- Every PC had a powerful processor. The distinction between workstation and PC is a matter of function than hardware. The factor that distinguishes between the two is that most workstations have built-in hardware to connect to Local Area Network (LAN).

Advantages

- Besides the five operating systems used in PCs viz: MS-DOS, MS- Windows, Windows-NT, Linux and Unix, all workstations run on variations of UNIX operating system such as Solaris used in SUN workstation, AIX in IBM workstations.
- Workstations equipped with 32-bit or 64-bit micro-processor, and operating system with virtual memory.
- It contains large main memory of at least 512MB to 1 GB and auxiliary memory of several hundred giga bytes (GB).

5. MICROCOMPUTERS

Microcomputer refers to a computer that is built around a single chip microprocessor. It is less powerful than workstations. These categories of computers were not only reliable but cost effective also. Microcomputers can be divided into two groups - personal computers and workstations. Workstations are specialized computers that approach the speed of mainframes, as discussed above. Often personal computers are connected to networks of other computers.

Uses

- The home user, schools and small businesses all use microcomputers. Microprocessors are found in many home appliances.
- Schools, universities and other educational organizations also use microcomputers.

Advantages

- Microcomputers are an affordable way to improve efficiency by storing and processing data.
- Microcomputer is good at word-processing and using as a communication tool.
- Microcomputers provide access to the Internet for millions of users.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMPARISON BETWEEN WORKSTATION AND MICROCOMPUTERS

Workstation

(1) They are more powerful than average personal computers

(2) It has powerful higher resolution color graphic displays with hardware support for computationally intensive operations

(3) They are constructed differently than microcomputers, as they are based on different architecture of CPU called reduced instruction set computing (RISC)

(4) Workstations run on Unix operating system or a variation of it.

Microcomputers

(1) Relatively less powerful than workstations

(2) Capabilities are less compared to workstations

(3) These computers use a microprocessor as its CPU which contains the entire circuitry of the computer

(4) Microcomputers can run any of the four major operating systems such as DOS, Unix, OS/2 & Microsoft Windows NT

6. PORTABLE COMPUTERS

A portable computer is a small, light enough to be carried, and can operate on internal battery.

In terms of decreasing weight capabilities and the order in which they were developed, portable computers can be classified as per their form factor: hand held, laptop, notebook, PDA, etc.

Hand held (Palmtop) computer is a miniature portable computer small enough to be resembled with pocket calculators but possesses many of the capabilities of a personal or even a mainframe computer. These computers serve as dictionaries, language translator, notepads, and calculators. They are also used by police and investigating agencies.

Laptops are another classification of computers which offer complete computer functionality in a size convenient for holding on a person's lap. A larger, heavier version is called a laptop computer, however, a N**otebook** computer can fit into a briefcase and weighs lesser than two pounds, and approximately 9X12X2 inches in size, yet it can compete with the microcomputer. The enabling technologies for the laptop computer, the most prevalent type of portable computer in 1999, include: flat-panel displays; electronic components and disk drives that use very little power; long-life rechargeable batteries; and a compact keyboard that provides some special way to move the display cursor other than through an external mouse (which can be added as an option).

Personal digital assistant or a PDA is the smallest hand held computer PDAs are used to track appointments and shipments as well as names and addresses. PDAs are called pen-based computers because they utilize a pen-like stylus that accepts hand-written input directly on a touch-sensitive screen. It may be noted that delivery employees use this type of computer.

Features of Portable Systems

The main features of portable system design are:-

(i) Portable systems are designed to be smaller and lighter than desktops.

(ii) Portable systems are specifically designed to use less power than desktops.

(iii) The battery technologies like lithium ion make power supplies more consistent and reliable and meeting the power requirements.

(iv) Portable systems have such power management that when operating systems and utilities that turnoff specific system when not in use, components such as disk to drives this greatly increase battery life.

(v) The portable systems manufactured these days are upgradable and repairable.

(vi) The difference between portable systems and desktop systems is the display screen.

(vii) Portable system has a flat screen, which is less than half an inch thick. This is called LCD or liquid crystal display and usually it is coloured.

(viii) LCD display in portable system is designed to operate at a specific resolution, because the size of the pixel on an LCD panel cannot be changed. (Generally 800 X 600 arrangement of pixels display on 640 X 480 pixels screen).

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

CHARACTERISTICS OF COMPUTER:-

(1) Speed: - A computer is very fast device. It can perform large amount of work in a few seconds. Where human being worked a particular work for whole day, computer does the same in very short time. Today, computers can perform 100 million computations in one second. The speed of computers are measured in terms of microseconds, Nano seconds and even in Pico seconds.

Where 1 second=10^{-6} micro second =10^{-9} Nano second =10^{-12} Pico second

(2) Accuracy: – The computer is 100% accurate and capable to perform arithmetical calculation and logic operations with the same accuracy. It can never make mistakes. All mistakes are done by users. The cause of errors may be due to inaccurate feeding data or due to wrong setting of the programmer.

(3) Diligence: – A computer can operate twenty four hours continuously without taking any rest. It has no feelings or no emotions, if you work continuously for 3 hours, you feel lack of concentrate but a computer is free from these and you will get the result you want with the same speed and same accuracy.

(4) Versatility: – Versatility is one of the most wonderful features about the computer. One moment, it is preparing the results of a particular examination, the next moment it is busy in preparing electricity bills, and in between it may be helping an office secretary to trace an important letter is seconds. It can do multiple works at a same time.

(5) High Memory: – computer has made more memory or storage capacity than human beings. It can store millions of data and instructions, which can be retrieved and recalled even after a number of years. This is not possible in case of human brain.

(6) Automation: – Computers are automatic in operation. It means once the data and instructions are fed to a computer, human interventions are not required.

(7) Superiority of Manufacture: – Computer hardware is manufactured with the best materials by the most sophisticated processor. It has superiority of manufacture as compared to other machines.

(8) Reduce the Paper Work:–

(9) Solve Complete Problems: – Computer solved those problems which are not solved by human being computer solved the computer solved the problems in seconds.

LIMITATIONS OF COMPUTER:-

Garbage-In Garbage-Out (GIGO) : Computer lacks common sense and 10, thus incorrect-input would lead to incorrect output result.

Software Limitation: To do specialized tasks, special purpose soft- ware are required, which becomes very costly.

Hardware Limitation: Some of the applications require greater speed and storage space, which lead to more investment.

Data Tampering: Data handling through computers is prone to the loss of information and being tampered. Thus, it requires special protection routines.

Dependency: – Fully Dependent upon human beings.

Lack of Intelligence: – Computers do not have their own Intelligence and hence cannot think for example, a computer can create music but cannot tell its quality. In other words, computer has no brain.

Lack of Detecting Errors: – Computer can detect only syntax error; it cannot detect flaw or error in logic.

Lack of Capability: – It has no Capability to take any decision by itself, like a human being.

No Feelings:–Computers have no feelings and no emotion because they are machine.

Lack of Creativity and imagination: – It does not have intuition and cannot draw conclusion without going through all the intermediate steps.

Experience:–Computers don't learn from experience.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

BASIC COMPUTING ARCHITECTURE

Computer Architecture is concerned with the structure and behavior of the computer as seen by the user. It includes the information, formats, the instruction set, and techniques for addressing memory. The architectural design of a computer system is concerned with the specifications of the various functional modules, such as processors and memories, and structuring them together into a computer system.

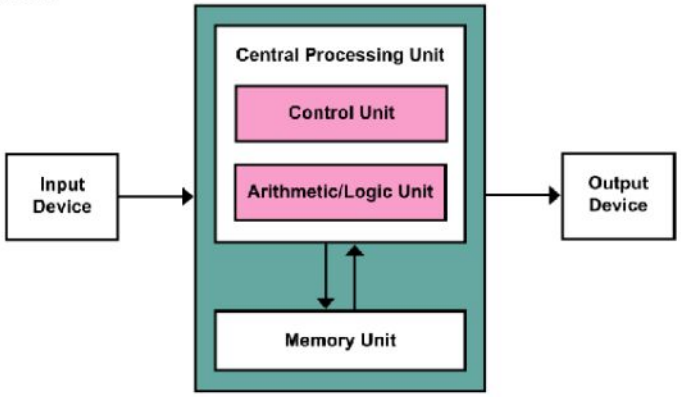
Two basic types of computer architecture are:

- Von Neumann Architecture
- Harvard architecture

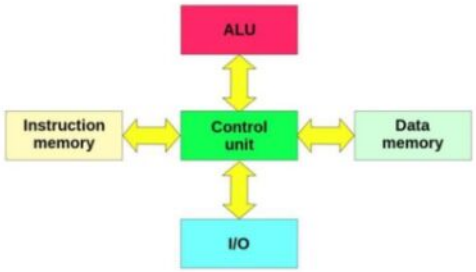
Von Neumann Architecture: This architecture describes a general framework or structure, that a computer's hardware, programming, and data should follow. Although other structures for computing have been invented and implemented, the vast majority of computers in use today operate according to the von Neumann architecture.

Von Neumann architecture is composed of three distinct components (or sub-systems): a central processing unit (CPU), memory, and input/output (I/O) interfaces.

1. The CPU, which can be considered the heart of the computing system, includes three main components: the control unit (CU), one or more arithmetic logic units (ALUs), and various registers. The control unit determines the order in which instructions should be executed and controls the retrieval of the proper operands. ALUs perform all mathematical and Boolean operations. The registers are temporary storage locations to quickly store and transfer the data and instructions being used. A CPU that is implemented on a single chip is called a microprocessor.
2. The computer's memory is used to store program instructions and data. Two of the commonly used types of memories are RAM (random-access memory) and ROM (read-only memory). RAM stores the data and general-purpose programs that the machine executes. RAM is temporary; its contents can be changed at any time and it is erased when the computer is power off. ROM is permanent and is used to store the initial boot up instructions of the machine.
3. The I/O interfaces allow the computer's memory to receive information and send data to output devices. Also, they allow the computer to communicate to the user and to secondary storage devices like disk and tape drives.



Harvard architecture



Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

FUNCTIONS OF COMPUTER SYSTEM

There are basically six building blocks of digital computer system where each block performs certain functions. Thus, all types of computer systems perform the following operations:

(1) Interfacing: We live in an age when people and machine can communicate. This activity is made Possible by the interfacer component of a computer system which takes in data from the user and translates it into machine language and vice versa.

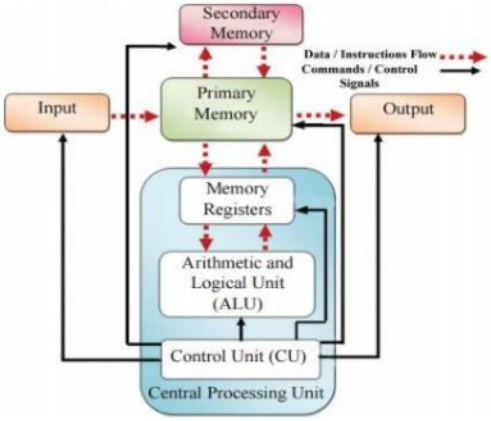
(2) Inputting: It accepts input from the user.

(3) Storing: Stores data and instructions for processing.

(4) Processing: Processes data, by performing arithmetic operation or logical operation based on the instructions coded in the program.

(5) Outputting: It produces output as meaningful information or result for the user, may be as printed report or visual display.

(6) Controlling: It controls and directs the sequence of operations to be performed.



Basic Organisation of a Computer System

INPUT UNIT An input unit is responsible for accepting data and instruction in the form that depends upon the device used. It is generally an electromechanical device that accesses the data that computer uses, which is communicated from outside. This is the actual means by which data is converted into electronic pulses and further transformed and stored as binary codes. This task is performed by input interface. Input unit functions can be summarized as follows :

1. It communicate with its external environment by accepting (or reading) data and the list of instructions. The common input devices are: Card reader, Paper tape reader, Magnetic tape drive, Disk drive, Keyboard, etc.
2. It converts these instructions and data in computer acceptable form.
3. It supplies the converted instructions and data to the computer system for further processing.

OUTPUT UNIT The output devices of the computer enable it to communicate results to the outside world or user. It accepts the results produced by the computer which are in coded form and hence cannot be easily understood. It converts these coded results to human acceptable (readable) form. This activity is performed by output interfaces, which are designed in compatible with output devices. It supplies the converted results to the outside world. Output devices fall into two categories:

1. Devices which produce output that is understood by human beings, e.g. Printers and Monitors (VDU)
2. Auxiliary storage devices that hold data may be intended for further processing (e.g. Magnetic disks)

CENTRAL PROCESSING UNIT

The CPU is known to be heart or brain of the computer system as it follows the instructions of the software/program to manipulate data into information.

All types of units of a digital computer are interconnected and functioning under central control. The "central processor" or Central Processing Unit (CPU) carries out a variety of essential data manipulation and controlling tasks in the computer which is also known as **microprocessor** i.e., *microprocessor = ALU + CU*. In this unit actual execution of the instruction takes place. Every microprocessor contains a system clock. The speed at which processor executes instruction is known as clock speed and it is measured in Megahertz (MHz).

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Where mega means millions and hertz means cycles. Therefore 550 MHz processor performs 550 million cycles per second. Generally, clock speed determines the system performance. The latest generation of processors operate in gigahertz (GHz) i.e. a billion cycles per second.

Clock speed is the speed at which system clock oscillates. For microcomputers it is measured in Megahertz (MHz) and Gigahertz (GHz)

Megahertz (MHz): a measure of frequency equivalent to 1 million cycles per second

Gigahertz (GHz): a billion cycles per second

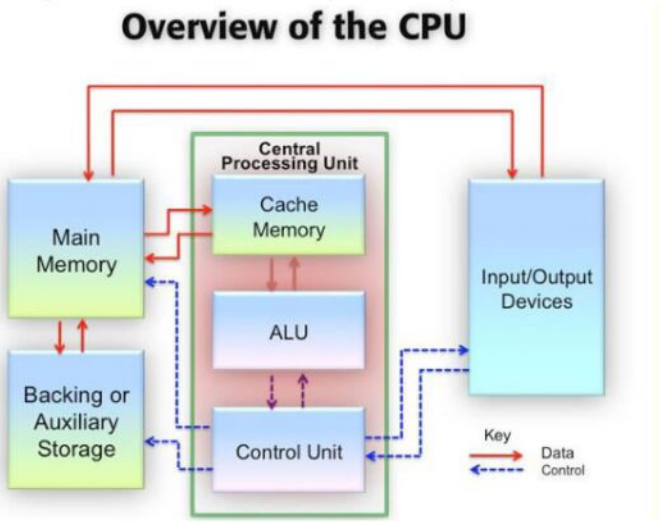
A computer performs its operation in fraction of a second. The increasing speed given below are:

Millisecond = 10^{-3} second
Microsecond = 10^{-6} second
Nanosecond = 10^{-9} second
Picosecond = 10^{-12} second.

CPU consists of two parts :

i. Control Unit (CU) and
ii. An Arithmetic-Logic Unit (ALU).

Both of them contain high storage areas and all are linked by electronic path known as bus.



i. Arithmetic and Logical Unit

ALU is the principal element of the CPU. It performs the basic four arithmetic operations. The arithmetic logic unit also allows the control unit to make logical decisions that affect the flow of the program. The arithmetic and control units heavily depend on third element of a computer, the main storage or memory unit. The arithmetic unit needs numbers to operate on and to store the immediate results until the end of calculation.

Arithmetic/Logic unit, performs arithmetic and logical operation and controls the speed of these operations

ii. Control Unit

Control unit acts as automatic operator; it provides the means of communication within the machine, by moving, advancing or transferring information. CU does not perform any actual processing on data. Rather, it manages and co-ordinates the entire computer system. Basically, *control unit is incharge of decoding instructions and sequencing the actions of various functional units*. It fetches program instructions one by one in sequence, stored in main memory. It interprets (or decodes) them, and sends appropriate control signals to the various units. It resolves conflicts between on chip resources and responds to external and internal interrupts.

Control unit refers to a functioning unit which carries out four basic operations, there are :

1. *Fetches an Instrunctions*
2. *Decodes the instrunctions*
3. *Executes the instrunctions*
4. *It stores the result*

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

MEMORY (Storage) UNIT

Different types of memory (or storage) are usually used in a computer system. Any storage unit of a computer may have following characteristics:

(i) Access time is the amount of time needed for locating and delivering data from memory (storage device) to the processor after the address for the data has been selected. Access time is usually measured in millisecond and is used as a performance measure for disks and CDs.

(ii) Storage capacity is the amount of information/ data a storage unit can hold. Higher storage capacity devices are desirable.

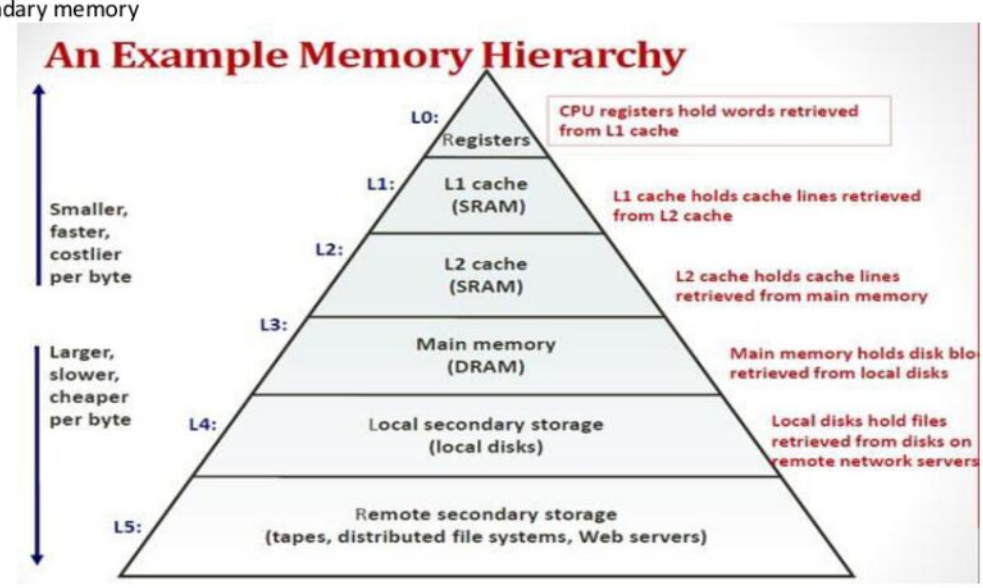
(iii) Cost per bit of storage is generally higher with fast devices, though the objective is to minimize the cost.

These may be fast or slow. The availability of fast, inexpensive memory has had enormous impact on computer technology. Fast is relatively small and expensive storage units, used as fast access by the central processing unit. Slow are significantly large and inexpensive storage units.

Since the 1970s, memory performance and cost have reduced at the rate of 27% per year. Over this period, the price per bit of main memory has decreased at the astonishing rate of approximately 27% per year. These advances have taken place primarily, due to improvement in the integrated microelectronic technologies.

Storage unit of computer system consist of four types of storage, namely:

1. Registers
2. Cache memory
3. Primary memory
4. Secondary memory



1. Register

Registers are high-speed storage areas that temporarily store data during processing to enhance computer's performance. The size of the internal registers indicates how much information the processor can operate on at one time and how data moves around internally within the chip. A Register is capable of holding one number and its location is with an execution unit.

It is usually a small chip data array that is addressable by program instructions. Registers are included in all major microprocessor and provide storage with fastest possible access time. Registers may be of three types namely:

General purpose Registers (Accumulator) : These are the registers which hold temporary data values during computations. This is because reading a value from register is much faster than reading from memory locations, thus registers greatly improve the execution speed. There may be any number (a few or hundreds) of general-purpose registers, depending on the architecture of the computer.

Dedicated Registers : These registers are meant for execution of some instructions. Generally, modern architecture program counter and condition codes are placed in these registers. However, some older CPUs used to have single accumulator register for each arithmetic operation and to receive the result.

Floating-point Registers: These are general-purpose registers, accessed by floating point and data transfer instructions. They provide storage space for lengthy floating-point numbers; this speeds up the execution significantly.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

2. Cache Memory

Cache memory is a small high-speed buffer memory used to hold instructions temporarily that the processor is likely to use frequently, to speed up processing.

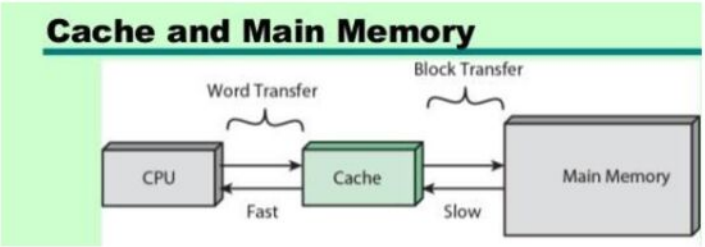
The CPU of a computer system commonly use Cache memory where it holds or buffers the contents of the main memory because CPU runs much faster than main memory RAM. Thus to reduce the waiting time of CPU cache is used. Cache memory also holds the contents of the secondary storage (viz. disk or tape). Caching is used in various purposes such as for look up files in directory system and web pages on a web server. Cache memory reduces traditional system bottleneck because system RAM is much slower than CPU. This prevents the processor from having to wait for program and data from slower main memory.

A cache memory is generally smaller and faster memory (than main memory). The CPU generates a main memory address, when it wishes to read or write data. When a cache is added to this design, it is intermediate memory between CPU and Main Memory.

There are two kinds of cache Level 1 and Level 2

i. **Level1 (L1)** cache is a memory that built into i486 and higher level processors to improve processing speed. The L1 cache, typically containing 8KB, can be read in a single clock cycle, so it is tried first then followed by L2 cache. Level1 cache is the fastest type of cache but holds less data than other levels of cache. Without the Level1 (L1) Cache, a processor would be forced to wait until system memory is caught up. Its capacity ranges from 8 to 256 kilobyte

ii. **Level 2 (L2)** cache is not part of microprocessor chip. There is another secondary external L2 Cache memory of extremely fast static RAM (SRAM) chip, which is also used in pentium based systems which is faster than DRAM but slower than L1 cache. Its capacities range from 128KB to 2MB.



3. Main Memories or Integrated Circuits (Semi-conductor memories)

In general, the main memory will contain instructions and data that are accessed by a program while it is executing. Main memory stores data and instructions for the CPU. It has four basic purposes:

- To hold data and instructions for a given job, from an input unit (keyboard, terminal, etc.) before processing.
- To hold intermediate results, while processing.
- To hold data after processing, until they are transferred to output device
- It also holds program and instructions not only received from input devices but also from secondary storage devices.
- It also stores complex programs such as operating system.

Thus, primary storage or main memory consists of mainly four areas:

(i) Input Storage Area : It holds data and instruction required for processing.

(ii) Working Storage Area : It holds the data being processed and intermediate results of such processing.

(iii) Output Storage Area: This holds the final result of processing.

(iv) Program Storage Area: This area holds the sequence of instructions for the computer.

Since memory addresses are binary numbers, the memory size is typically related to a power of two and described in units of Kilobytes (KB), Megabytes (MB), Gigabytes (GB), or (for mass storage) Terabytes (TB).

A Bit = each 0 or 1, short for binary digit
A Byte = group of 8 bit, represent one character

Where 8 bits = 1 byte = 1 character

A Kilobyte = 2^{10} or 1,024 bytes
A Megabyte = 1024 Kilobytes
A Gigabyte = 1024 Megabytes
A Terabyte = 1024 Gigabytes and
A Petabyte = 1024 Terabytes

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

There are two types of built-in memory, permanent and temporary, known as ROM and RAM.

(a) Read Only Memory

Read Only Memory (ROM) is simple type of memory, whose contents cannot be changed even after loss of electrical power. There are two types of read only memory, namely, (a) manufacturer programmed, and (b) user-programmed.

(i) Manufacturer programmed ROM: ROM is one in which the contents are programmed by the manufacturer during the construction of the chip, it is unalterable. Usually, ROMs are used for high volume controlled applications such as the set of instructions known as "system boot program" that is needed to make the computer system ready for use when its power is switched on. These instructions are startup instructions, commonly found in ROM. ROM-BIOS means Read only Memory **Basic Input Output Services**. These services or instructions perform the basic control over hardware including I/O devices. Generally, ROM chip of a computer stores Operating System. Disk Drivers, Printers drivers, etc. Thus, ROM stores program and data which do not change and are frequently used.

(ii) User-programmed ROM: The user-programmed ROM is one in which user can load and store read only programs and data. The user- programmed ROM include the following:

- **Programmable Read only Memory (PROM) :** It can be written once via an irreversible process. PROM is a non-volatile storage, i.e. the stored information cannot be changed even if the power is switched off.
- **Erasable Programmable Read only Memory (EPROM) :** Data on EPROM can be erased to store new information. However, the information on this chip can only be read. When EPROM is in use, information in the chip remains until it is erased. EPROM chips are generally used by R&D personnel who test the efficiency of the system with new programs. EPROMs are of two types namely, **Electrically Erasable Programmable Read only Memory (EEPROM)** and **Ultra Violet Erasable Programmable Read only Memory (UVEPROM)**. Data on EEPROM can be erased by using high electric voltage and due to this ease with which stored program can be altered. Whereas data on UVEPROM can be reset by exposure to intense ultra violet rays.
- **Flash memory :** Flash memory is small variant of Electrically Erasable programmable Read Only Memory. These memory chips are the latest known storage devices, available in capacities ranging from 512 MB to 2TB. They are non-volatile memories as they retain the data stored in them without a power source. They may be used to replace magnetic storage i.e., floppy diskette. *Flash memory chips can be erased and reprogrammed more than once.*

(b) Random Access Memory

*RAM (random access memory) chip, are meant for primary storage. They hold temporarily **(a) software/ program instructions and (b) data before and after processing**.*

"Random access" means that any location can be referenced in the same time and in same manner, as it is independent of the address or location in the memory. It is a volatile memory. It holds data and instructions, during their execution. The additional RAM chip can be plugged into the special socket on the motherboard known as **Single In-Line Memory Module (SIMM)**. RAM chips are of two types, namely, **dynamic RAM (DRAM)** and static RAM (SRAM).

DRAM is denser and needs less power. The densities achieved by DRAMs have made multi-megabyte primary memory affordable. DRAM regenerates the storage charge to retain the stored data by using external circuitry periodically. **SRAM** is 10 times faster but more costlier and consume more power than DRAM. SRAM is used mainly for specialized applications.

COMPARISON BETWEEN DYNAMIC RAM AND STATIC RAM

Dynamic RAM	Static RAM
DRAM operates at a high speed	It is 10 times faster than DRAM
It keeps refreshing continuously and thus regenerating the storage charge as it is dynamic in nature, otherwise information may be erased	It does not require refreshing the memory contents. Being static in nature, it remains stable as long as power supply is provided
It has higher density	Less densely packed than DRAM
Storage capacity is more	Lesser storage capacity
Requires less power consumption	Requires more and constant power supply
It is cheaper	Expensive
Base element used in DRAM is capacitor, therefore, bit stored and regenerated is in charged form	Base element used is flip flop, thus, bit stored is in voltage form

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMPARISON BETWEEN RAM AND ROM

ROM	RAM
Read only memory i.e., information can only be read and cannot be written on to it.	Random Access Memory as it randomly selects and uses location of the memory to access the information.
It stores information permanently.	It holds information temporarily.
Information is not lost even if power supply is switched off.	Information is lost when power supply is switched off.
Known as non-volatile memory.	Known as volatile memory.
Holds system software's such as operating systems.	Holds application programs which are currently in use.
Types of ROMs are PROM, EPROM EEPROM.	Types of RAMs are Dynamic RAM & Static RAM.

(c) Video RAM (VRAM)

Video RAM is a special type of dynamic RAM (DRAM) used in high speed video applications. VRAM uses special circuitry that transfers pixels from video RAM to screen serially (bit by bit).

Monitor has its own memory known as VRAM that is used to accelerate the display of graphics on the screen. It does this by the use of two ports, one port is connected to the screen and other to CPU. Data from CPU flow to the screen VRAM very smoothly. That is why we see clean graphic display on the screen. A variation of this is Window RAM (WRAM) that supports memory.

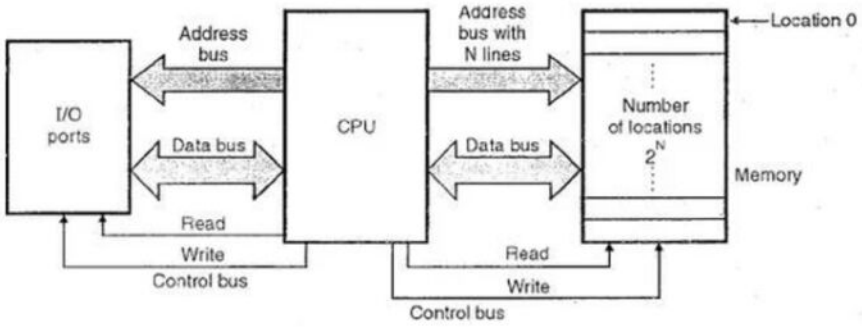
4. Secondary Storage

Secondary storage is required to supplement the limited capacity and volatile characteristics of primary storage. The secondary storage is inexpensive, having large capacity and is non-volatile in nature. It stores program and data, which may be required at later date. The most commonly used secondary storage devices are magnetic disks, tapes, or optical disks.

CHARACTERISTICS OF PRIMARY AND SECONDARY STORAGE

Primary Storage	Secondary Storage
It is the main memory as part of the CPU.	It is auxiliary memory which works under the control of CPU.
It is most expensive.	Relatively less expensive than primary memory.
Storage capacity is generally in Megabytes (MB).	Storage Capacity is in Giga bytes (GB).
Information can be retrieved and processed very very fast.	Retrieval and processing of information is relatively higher than primary memory.
Program and data are processed directly on RAM component of primary memory.	Program and data cannot be processed directly but first it is placed on RAM of main memor
Based on semi conductor Technology.	Made up of plastic coated with iron oxide material.
Access is random	Access can be sequential or random
It holds system software such as operating system, Application programs and user programs that are currently in use	It holds program and data other than those which are not currently in use i. e., used for backup purposes.

MEMORY BUSES



Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

When any program execution takes place in the CPU, it requires frequent interaction between the CPU and main memory= The connectivity between CPU and memory is established through a channel known as BUS. Physically bus is a set of wires which carry group of bits.
Buses are electrical data pathways within CPU and between CPU and other components of the motherboard through which bits are transmitted.
The capacity of a bus is called bus width which means the number of bits they carry at one time.
Hence, like the buses between CPU and main memory, in similar manner system uses buses for connecting CPU and I/O devices, which are known as data bus and address bus.
Data Bus: The data bus is called bi-directional, as it can serve to write data to and read data from the memory system. New data can be read from or written to the memory system every cycle time. Physically bus is a group of parallel wires, that form an electrical path that connects the CPU, Memory, and I/O devices. Thus, a chip may be 8-bit, 16-bit or 32-bit chip. The number of wires in the bus affects the speed of the data transfer, because each wire can transfer one bit at a time, thus, 8 wire bus can move eight bits or one byte at one time. The 16-bit bus that transmits data along 8 or 16 data lines depending on the adapter card in the expansion slot, is known as ISA bus (Indian Standard Architecture). Further, advancement in computer improved the performance by the use of 32-bit bus that transmits data along 32 data lines, known as MCA (Micro Channel Architecture). It is faster than ISA bus, because MCA expansion slot requires only 32-bit adapter card. Since MCA bus could not work with ISA bus, therefore, EISA bus (Extended Indian Standard Architecture) was designed to solve this problem. Although, EISA is 32-bit bus, it could accept old ISA expansion slot. EISA is faster than ISA.
Address Bus : It transmits signals for locating a given address in the primary memory (RAM). The numbers of wires in address bus determines the maximum number of memory address, i.e. If address bus could carry 8-bit at one time, then CPU could address 256 (= 28) bytes of RAM.
Similarly, **Control bus** transmits signals of "read" or "write" data to or from a primary storage address, input device, or output device.

COMPARISON BETWEEN DATA BUS AND ADDRESS BUS

Data Bus	Address Bus
It is a set of parallel wires that connect CPU, Memory & I/O devices	It is a set of wires that connect CPU and RAM (Primary Memory)
It is bi-directional as it serves to write data to and read data from the memory system	It transmit signals for locating a given address in the primary memory
The number of wires determine the number of bits it carry at one time	The number of wires determine the maximum number of memory addresses
Data Bus can be 8-bit, 16-bit or 32-bits that means it can carry1 byte, 2 bytes or 4 bytes respectively	If address bus carry 8-bits at a time, then CPU could address 256 bytes of RAM. Most PCs have 20-bit address bus and 32 bit address bus that can address 1 MB or 4 GB of RAM respectively
Data buses are designed to match the capabilities of the devices attached to them	Whereas for address buses operating system is taken into consideration.

MEMORY AND ITS IMPACT ON PROCESSING

As we know, RAM holds all the information the CPU needs to process. Thus, if a computer has more RAM, it would be able to load bigger programs as well as data files, it can make the computer runs faster to execute the loaded program. This is because the exchange of data between microprocessor and RAM are so intricately timed, the distance between them becomes a critical performance factor. Although RAM is smaller compared to disk storage space. A 2MB RAM is capable of executing a program that actually occupies 10MB of disk storage space by loading only the most essential files into its memory. However, when other part of the program is required, it swaps or unloads the part which is not required and loads the program/ data it requires.
Thus, whenever we double-click on application icon to run a program, the microprocessor loads the program files from the hard disk into main memory (RAM) and it remains there as long as we are working with that program. When we save the same program, it is written on to the disk and RAM is freed. The memory modules are ready to receive fresh instruction from the processor.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

SECONDARY STORAGE DEVICES

Secondary storage devices are supplement to the main memory that permanently holds program instruction and data as well as output (information).

CHARACTERISTICS OF SECONDARY STORAGE DEVICES

The storage devices come in many sizes, speeds and costs, because of economical and technological reasons. The four most important factors that determine the efficiency of the auxiliary memory are:

- **Total capacity:** The amount of information it can store. That may be in megabytes(MB), gigabytes (GB), or terabytes(TB)
- **Access time :** The time needed to acquire information, that is, how long it takes to find the first item of data needed. The requested data may be supplied by memories in very small intervals of time measured in nanoseconds (billionths of a second), typically 60ns or less.
- **Data Transfer rate:** The rate at which information can be streamed into or out of the computer memory in units of (millions or billions of) bits per second.
- **Cost per unit** of data expressed in, say, $ per megabyte.

TYPES OF SECONDARY STORAGE DEVICES

Auxiliary memory comprises all other memories whose contents must be fetched into main memory before processing by the CPU. It is used to store both current programs and data files and to provide backup storage as a safeguard against loss and for programs and data not in current use and which may be needed in future. Most auxiliary memory is rewritable, i.e. it can be read, written, rewritten etc. many times without deterioration. Certain types of optical auxiliary memory are read only.

Auxiliary memory generally uses electromagnetic or optical digital technology for storing data. There are at least six different types of auxiliary memory, and their variety and number continue to grow:

- Magnetic tape
- Magnetic drum
- Diskette
- Fixed-head hard disks
- Moving-head hard disks
- Optical disks

There are two ways to access the information: sequential or direct access. Thus, above devices may be classified as:

- **Sequential Access Devices :** Sequential access refers to the method of storing or retrieving information that requires the program to start reading at the beginning and continue until it finds the desired data. Magnetic tape is an example of sequential access storage device.

  Sequential access devices are the devices whose access time varies according to the location of the information stored. It has to pass over the previous information to access the information at a particular location. The most common sequential access device is magnetic tape. It stores and retrieves the information in same sequence as songs are stored and retrieved in music cassette, that is, to play 5th song in the cassette, you need to pass through or forward first four songs. They are generally slow devices.
- **Direct Access Devices :** Direct access devices are the devices whose access time remains the same irrespective of the location of the information stored on the device. It directly accesses the information at desired location. The most common direct access devices are magnetic disks and optical disks. It stores and retrieves the information in same manner as songs are stored and retrieved on a CD, that is, to play 5th song on the CD, unlike cassette you don't need to pass through or forward first four songs but directly play the desired song.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

MAGNETIC TAPES

Magnetic Tape is a form of auxiliary memory. Initially magnetic tape was considered to be the main epitome of data processing. However, magnetic tape is not much in use because of the fact that the inexpensive and large capacity randomaccess storage devices (disks) are available. *Magnetic Tape is thin plastic reel coated with iron oxide substance that can be magnetized. Data is represented as magnetized or non-magnetized spots representing Is and Os respectively. Because of continuous length, read/write head cannot jump to a desired point, thus tape must be read or written sequentially.*

Tape cartridges resemble with audio cassettes that contain tape in rectangular plastic housings.

Thus, the storage capacity of the tape = Data recording density X Length

For example, a fully written reel of tape, recorded at 38,000 frames per inch. For one gigabyte it can be computed as follows: 2,400 feet X 12 inches/foot X 38,000 bytes linch = 1,094,400,000 bytes.

Data Transfer : *Datatransfer rate refers to the transferring the number of characters per second from the tape to the primary storage.*

The nominal rate of data transfer in terms of characters/bytes per second is derived from the tape speed, the recording density and the number of tracks.

Data Transfer rates (bytes per second) = Tape speed X density (in bpi)

Data Transfer rates are quoted in 40,000 characters or bytes per second (e.g. 40 kc/s or 40 kb/s) and normally these are the maximum speeds.

Uses

Magnetic tapes are used for following applications:

- For taking backups of data that are stored on an on-line storage device (i.e. disks) so that data can be retrieved from the backup if data on the disk is lost accidently.
- Archiving of data which may be required rarely.
- Magnetic tapes are generally used for distribution of software by vendors.
- Magnetic tapes are used for the applications which require sequential data processing.

Advantages

Low cost: Magnetic tape reels and cartridges have high data recording densities. Thus, cost per bit of storage is less.

Reusability: Magnetic tapes can be erased and reused many times.

Easy-to-handle: Magnetic tapes and cartridges are compact and light in weight.

Portability: Magnetic tapes are often used for transferring data and program from one computer to another because of its compact size, light in weight and high storage capacity.

Limitations

- Magnetic tape is not suitable for applications which require random access of data, because tape can only be accessed sequentially.
- Tape must be stored in dust-free and moisture free environment, because dust or humidity may result in loss of stored data.
- It need to be labelled in order to avoid the ambiguity.

DIRECT ACCESS DEVICES

Magnetic Disks

Magnetic disks are most commonly used direct access secondary storage device. All magnetic disks are usually circular platter coated with iron oxide material. They are available in different sizes and capacities. They can be portable and permanently mounted in the storage devices called disk drives. They also differ in data accessing capability and thus, accordingly, in data transfer rate. They are usually made up of rigid metal or flexible plastic. The most commonly used devices are floppy disks and hard disks.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

a) Floppy Disks

Floppy disks are also known as floppies or diskette. Floppy disks were introduced in 1972 by IBM as a portable storage media for personal computers. A floppy disk is a single disk, which is flat and round plastic coated with magnetic oxide. The disk surface of the floppy is encased in a square plastic jacket for protection. When data can be recorded on one side of the disk, then it is known as single sided (SS) diskette and when both the surfaces are used, then it is known as double sided (DS) diskette. Thus, a floppy disk drive may be able to read the record from the both sides of a disk (a double-sided drive) or only on one side (a single-sided drive). The capacity of a floppy disk and the rate of data transfer is substantially smaller and slower than the hard disk. Floppy disks are very cheap storage medium for microcomputer systems, intelligent terminals and for some mini-computer systems. There are two standard sizes 5.25 inch and 3.50 inch frequently referred to as mini-floppy and micro-floppy, respectively.

Precautions required in using and storing a diskette:

- The surface of the diskette should not be touched with hand or any other object.
- Write-protect notch should be observed by peeling off or sticking on the aluminium square (depending on application).
- Disk must be correctly entered in the disk drive.
- The diskette should be inserted and removed only when power to the computer system is on.
- The diskette must be protected from direct sunlight, heat, dust, liquid,moisture and magnets.
- Disk drive must be cleaned in order to remove the dust regularly.
- Diskette should not be compressed.

b) Hard disk: Winchester Disks

A hard disk is a high-capacity, high-speed rotational thin but rigid metal platters covered with a substance that allows data to be held in the form of magnetized spots.

A hard disk is large 14 inch based on Winchester technology, placed permanently in a contamination free container. Hard disks are sometimes called fixed disks because they usually cannot be removed from the computer. Hard disk drives are also called Winchester drives, a name derived from a 1969 IBM drive that stored 30 MB of information on each of two spindles, a 30-30 arrangement reminiscent of the framed Winches- ter 30-30 rifle.

Data Storage Organization

A disk is stack of diskette called platter, where each platter has two sides. A disk drive may hold as many as 12 sides or surfaces. Each diskette in this stack is divided into concentric circles called tracks. The disk stores data in tracks of magnetically-oriented particles. There may be over 10,000 tracks on a high-capacity disk. Tracks are divided into sectors. Each sector of the disk is assigned a unique number. The formatting program identifies and marks sectors that are bad and lays out the sector locations and stores them. The collection of tracks with the same number arranged vertically on all platters is called a cylinder. As tracks get closer to the centre of the disk, their circumference necessarily gets smaller. But since the number of sectors remains the same and so does the number of bytes per sector (for most drives), the data is more densely packed. If a piece of data need to be accessed, then disk drive locates as per record's disk address.

The disk address consists of cylinder number, sector number, track number and surface number.

Similarly, the storage capacity of the disk system may be determined by the following:-

(a) Number of recording surfaces

(b) Number of tracks per surface

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(c) Number of sectors per track
(d) Number of bytes per sector
(e) Recording density of bits on a track
The storage capacity of the disk system = Number of recording surfaces
X Number of tracks per surface
X Number of sectors per track
X Number of bytes per sector

Disk Formatting: Finding Data on Disk by Operating Systems

Each type of disk requires different disk drives so that dimension of the disk should match with physical component of the corresponding disk drive. To solve this problem, disk formatting was introduced. Thus, to prepare the disk for use so that operating system finds data on disk. Computers operating system locate data on a disk because location of all data kept in special log file on the disk. Since each track and sector is labelled by performing logical format or soft format. This logical formats or Disk Formats are performed by DOS or Windows using Format command. Create four disk areas :-

(a) Boot record is a small program that runs when the computer is switched on. In order to run the operating system successfully, boot record determines whether the disk has the basic components of DOS or Windows. It determines that required files are present and disk has valid format then it boot the system by transferring controls to the operating system program files. Thus the computer can be said to pull itself up by its own bootstrap. The boot record also describes the number of bytes per sector and sectors per track.

Boot record refers to the section of disk that contains the operating system.

Boot sector is the portion of the disk reserved for the self starting portion (boot strap loader) of an operating system.

Boot strap loader is a program that is automatically run when computer is switched on. It first performs few basic hardware tests then passes control to operating system's loader files to load the OS.

(b) The File Allocation Table (FAT) refers to the table or list maintained by some operating systems to manage disk space used for the file storage. Files on a disk are stored, as space allows, thus a single file can thus be scattered in pieces over many separate storage areas. A file allocation table maps available storage space so that it can mark flowed segments that should not be used and can find and link the pieces of files. In MS- DOS, the file allocation table is known as FAT. File Allocation Table (FAT) is a hidden table on a disk that records exactly where I the parts of a given table are stored.

(c) Root folder or directory refers to the point of entry into the directory tree in a disk-based hierarchical directory structure. Branching from this root are various directories and sub directories each of which can contain one and/ or more files and sub directories of its own.

When we view the contents of the folder using OS, it lists files in the folder, their size, time and date of creation and last modified etc.

Root folder is a folder on a drive from which all other folders branch. The root folder's name consists of a single blackslash character. For example, on drive C, this folder would be represented in the file system as C:\ .

(d) Data area is the part of the disk that remains free after the boot sector, FAT and root folder have been created. It is called the data area, because this is the area where program and data files are actually stored.

Data Access

Because the disks are rigid, they can be spun much faster than a floppy-up to 10,800 Revolution Per Minute(rpm). Number of platters mounted on a rotating spindle are contained in the drive itself. Each platter surface has its own read-write head. This actually floats above the surface of the disk on a cushion of air. The heads on hard disks are designed like small airfoils so that they can be efficiently lifted and landed. The heads float very close to the disk surface, the gap being about 1 / 1000,000 of an inch. The disk case isassembled in a controlled and ultra-clean environment because contaminants in the form of airborne

particles can be very destructive. The inside of a disk drive cannot be a vacuum because the heads need air to lift off and land.

A head crash may happen if the surface of the read/write head or particles on its surface come into contact with the surface of the hard disk platter, that causes the loss of data on the disk.

Since the heads are connected to only one access arm, all heads move in unison. For example, when the head is over, say, track 157 on the top platter, it is over track 157 on all platters. For that reason, operating systems, through programmed directions to the hard disk controller, try to store all sectors belonging to the same file on the same cylinder to minimize head movement.

There are two track numbers that serve a special purpose. The landing track is a track where data is never written. This is where the heads go to land when the system is powered down or the heads are "parked" by a utility programme prior to shipment of the disk drive. Some disks have self- parking heads while others require user intervention through software. Access time for data stored on disk is basically determined by seek time.

Seek time: Time required to position a head over proper track. Seek time is zero for fixed read/write head drive.

Search time (latency time): The time required to spin the needed data under the head. This is also known as Rotational time.

Data Transfer time: This is the time taken to activate the read/ write head and read the requested data and transmit them to primary memory.

Disk Drives

There are two major varieties of disk drives

- Fixed head multiple platter
- Moving head (a) Moving head single platter (b) Moving head multiple platter

Fixed head multiple platter: Fixed Head (FH) disks have comparable access times and transfer rates and greater storage capacity may be up to gigabyte. Each fixed head drive contains a head for each track. The delay in accessing data block is due to rotational latency, that is, a to 15ms required for the block to revolve beneath the corresponding R/W head. However, track lengths linearly vary in distance from spindle, R/W heads are calibrated in such a way that track capacities are identicaL Thus, rate for data transfer is same for all the tracks (whether read from inner or outer track). Winchester drives for hard disks are usually fixed head. Winchester drives for hard disks do not themselves contain R/W heads. They are given with the platter by the manufacturer of that disk.

Fixed head multiple platter also known as non removable hard disk (or in short fixed disk) is housed in a micro computer system and it is used to store programs, data and output/ information.

Moving Head: In this only two heads are associated per platter. They are cheaper than Fixed Head drives. They are given with the platter by the manufacturer of that disk. Moving head drives may be either single platter or multiple platter.

Moving Head Single Platter: It contains two R/W heads, that enable reading from or writing to either surface of the disk.

Moving Head Multiple Platter (MHMP): This type of drive usually have 2N-2 heads, where N being the number of platters, this is because the top surface of the top platter and bottom surface of the bottom platter are not used on MHMP packs. As generally, non-winchester moving head drives are exposed to scratches and dust contamination than interior surfaces.

Disk Controller

A disk is controlled by a disk controller, which interprets the commands for operating disk drive. However, they support only Read and Write commands, that is, each such command is specified with a disk address (i.e. surface number, cylinder number, sector number and track number). The most popular disk controller standards are EIDE-Enhanced Integrated Drive Electronics and SCSI-Small Computer System Interface. EIDE drives usually spin at 5,400 rpm compared with 7,200 rpm for SCSI drives.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Characteristics of Magnetic Tape and Magnetic Disk

Characteristics	Magnetic Tape	Magnetic Disk
Accessability	Provides only serial access	Provides sequential as well as direct access
Updation	Not very fast because of serial access, it delays the retrieval	Easier and faster because of fast retrieval time
Access time	Slow	Fast
Data Transfer rate	Slow	Fast
Used as memory	Back up/archival purpose, long term storage.	Virtual memory
Cost	Economical - cheaper	More expensive
Compactness	Very compact	Disk packs are bulky require more physical space.
Environmental Problems	Prone to be affected with moisture, dirt, smoke etc. thus less durable	Durability is high
Human Intervention	Human Intervention needed to load/ unload tapes and tape rewinding is necessary when dismounting. Thus, extra care needed	No human intervention required
Portability	Easily Portable	Difficulty in Portability
Security	Security depends on the operator as after using, dismounting is done.	More chances of errors, less secure.
Reusability	Reusable	Reusable
Storage densities	Same for every track	Vary for inner and outer tracks, as same number of characters are stored on all the track. Inner tracks have more storage densities.

ZIP DRIVES AND JAZ DRIVES

Zip drive are a high capacity disk, that are slightly larger and thicker than normal floppy disks. They have their own special drive which can be attached to a parallel port (socket at the back of the computer that the printer usually connects to), so it can be plugged in when ever we need to use it. For frequent use it can also be permanent fixure in a computer like a floppy disk device. This drive has to have its own software installed so that the drive can be used. A Jaz drive is a removable disk drive capable to taking large capacity disks. They have a fast data transfer rate.

Uses Zip and Jaz drives can be used for backing up hard disks and transferring large volumes of data from one machine to another.

Advantages

- Using these devices it is fairly fast to back-up data. Data is also portable.
- These devices have a large capacity.

Disadvantages

- Disks must be kept safe and are subject to the same hazards as floppy disks such as heat, magnetic and fluid damage.
- Compared to hard disk access to stored data is slow.

PEN DRIVES

Pen drives are becoming increasingly popular. They are tiny devices which are about the size of small pen and which plug into an USB port. They vary in capacity and can store 64MB, 128MB and go up to an even 1 GB in some cases. The more space on the pen drive, means more price you will pay. Prices start at approximately $20

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

for smaller storage capacities. It is much faster to write to a pen drive than to CD- R Wand due to their size they are a more convenient and compact way of backing-up and transferring data.

Uses

Backing up and transferring data.

Advantages

- Small and compact.
- Very fast to transfer (Read/write) data.

Disadvantages

- They are more expensive than CD-Rs and CD-RWs.

OPTICAL DISKS: CDs AND DVDs

Optical disk is random accessed, removable disk on which data is written and read through the use of laser beam. It is usually 4.75 inches in diameter and less than 1/20 of an inch thick.

The process of recording data into an optical disc is called burning. To burn CD- Rand DVD-R special HW and SW are needed. Operating system Windows XP include CD burning capabilities.

It is also a random access storage media. In contrast to memory, that functions magnetically, an optical storage device uses a laser to etch and detect microscopic pits in the surface of its recording medium. As it contains highly reflective material, laser beam technology is used to read/ write on the disk. The use of this auxiliary computer memory is of three principle kinds, besides storage two are due to the origin as audio and audio-visual entertainment media. All optical disks (also known as laser disks) are round platters. The optical disk contains spiral tracks to store data of multimedia applications. These tracks like magnetic disks are split into sectors. It allows to organize the data to have maximum density over the entire disk.

Storage capacity of the disk = Number of sector X Number of bytes per sector.

Typically, the optical disk contains 3,30,000 sectors, each sector contains 2352 bytes. For reading information from the optical disk, it has to be mounted on a optical disk drive, which contains a motor to rotate the disk. The most commonly used disks are CD-ROM, DVD and many others.

Uses

For distribution of software, electronic version of journals, magazines, books, conference proceedings.

For storing multimedia applications, such as video games, movies, music, etc.

For archiving data, which is not frequently required.

Advantages

Low cost: Cost-per-bit of storage is low and it can store enormous data.

Reliable storage medium : The optical disks are more reliable than magnetic storage, because optical disk drives do not have read/write heads. Thus, no chance of crash into the disk surface.

Read large blocks of data: The data is organized in a single spiral track. This makes optical storage ideal for storing music and video as it can read large blocks of sequential data.

Easy-to-handle: Optical disks are compact, light weight easy-to-store and easily portable.

More data storage life: Optical disks have data storage life more than magnetic tape or disk (more than 30 years app.). Thus used for archiving purposes.

Permanent storage : Data stored on optical disk is permanent as it cannot be erased or overwritten.

Limitations

- It is not reusable storage medium.
- It is read-only storage medium.
- Data recording and reading require a complicated mechanism than magnetic disk.
- Data access speed is slower than magnetic disks.

Optical Disks: CD-ROM

CD·Rom (Compact disk read-only memory) is an optical storage medium used primarily with personal computers. It is used to hold pre-recorded text, graphics and sound.

CD· ROM applications include reference material, games, and software distribution. CD-ROM has revolutionized the encyclopaedia industry, totally replacing some print encyclopaedias.

CD-R (compact disk-recordable) disks can be written only once but can be read many times.

CD· R W (compact disk rewritable) disk, also known as erasable optical disk. It also allows users to record and erase data so that disk can be reused again.

CD· R W with such a drive, re-recording is done by first erasing old inforrnation through intense laser action that heats each amorphous stripe to that temperature just below its melting point where it recrystallizes, then after a few milliseconds of cooling to complete the process, a less intense laser action is used to record new information.

Uses

Large volume applications may be recorded to CD-R when they are finalised, for example a tailor made database system or a catalogue of student work for display. CD- R W s may be used in situations where we need to record the final result and keep it for quite a long time before replacing it with another result. A typical example might be a monthly school magazine which may be required for a number of sessions before being replaced with more recent work.

Advantages

- A portable record of a final result that is fairly secure is created with CR-R because once it has been created the disk becomes a CD-ROM.
- Large volume data is stored safely and securely.

Disadvantages

- Writing to CDs is time consuming. It may taken more than 15 minutes to fill a CD. Consequently they are unlikely to replace hard disks which have a much faster access time.

Optical Disks: Digital Versatile Disk (DVD)

DVD-ROM abbreviated as digital versatile/ video disk is a CD-like disk with read only memory and very high capacity, can store 4.7 to 17 GB.

A DVD (Digital Versatile Disk, formerly Digital Video Disk) has the same 12 em size as a CD but is called "versatile" because it appears destined to replace not only the laser-disk and the audio CD for entertainment but also the CD-ROM for use as auxiliary computer memory. DVD technology uses a higher frequency and shorter wavelength laser to etch pits, to make a CD. So, the surface of DVD contains microscopic pits, which represent as and Is of digital code which can be read by a laser. Furthermore, the DVD's spiral tracks are more dense than those of a CD. Although the diameter and total thickness of CD and DVD are same, yet capacity of DVD is 4.7 GB each side. The differences between CD and DVD are listed in the table given below

Like CD-ROM, an ordinary DVD is DVD-ROM, that is, its pre-recorded pitencoded information can be read but not rewritten. It is used for mass distribution of pre-recorded software programs and multimedia. But, DVD-ROM drives are now available, are WMRM devices. These have now become the optical storage media for use with personal computers, as a matter of choice. It is used for general read-and-write applications in the PC or consumer electronic domain. Similarly, DVD-R is used for low cost, write once recordable media. There also exists rewritable version of DVD- R known as DVD- R W. Primary use of DVDs is for distribution of pre-recorded motion pictures and music, with convenient random access.

CHARACTERISTICS OF CD & DVD

Characteristics	DVD(Digital Versatile/Video Disc)	CD(Compact Disc)
Size	12 cm (4 1/3")	12 cm
Versatility	Very versatile because destined to replace laser disc, audio CD for entertainment and CD-ROM.	Not very versatile as DVD.
Technology used to etch pits.	DVD technology uses a higher frequency (shorter wavelength) laser to etch pits.	CD uses lower frequency laser than DVD to etch pits.
Spiral Tracks	DVD's spiral tracks are more dense than CD circumferentially-0.4 urn, and radially -.74/lm.	Spiral tracks are less dense circumferentially-0.83/lm radially –1.6,
Capacity	8.5 GB (Single Side) 17 GB (When recorded on both sides) It is 13 times more than CD	650MB
Error correcting code	Read - Solomon Product Code(RS-PC). It is ten times robust than the code used on CD	CD uses Cross Interleaved Read Solomon Code
Eight to Fourteen modulation for recording	EFM Plus 8 bit to 16 bit modulation	EFM 8 bit to 14 bit modulation used

MAGNETO-OPTICAL DISKS

Magneto optical disks integrate optical and magnetic disk technology to enable read-write storage. Presently, magneto-optical disks are expensive, slow access time and not very reliable. It is 51A inch disks and can store 1000 MB. As magnetic disks dominate over magnetic tapes similarly, these days optical disk technology is advancing to offer reliability, cost effectiveness, read as well as write operations.

Features

- Very high density over magnetic hard disc.
- Cartridge is removable
- Long life

STORAGE SYSTEMS FOR NETWORKS AND LARGE COMPUTER SYSTEMS

Large computer systems such as mainframe computers and servers require storage hardware as PCs but on a much larger scale. As we commonly see desktop computers containing single hard drive installed within the system unit, instead we are likely to find on a storage server a separate hardware containing multiple high speed hard drives connected to computer system. Large storage servers can contain 2,560 hard drives that each hold 147 GB having total capacity of over 384 TB. These types of storage system referred to as enterprise storage system. Storage servers may also be used in the following:

Network Attached Storage (NAS) : NAS devices are high performance storage servers that provide storage for computers connected to the network.

Storage Area Network (SAN) : Like NAS, SAN also provide storage for a network, but consist of separate network of hard drives (or other storage devices) that is attached to main network.

The difference between NAS and SAN is how storage devices interface with the network. In case of NAS storage devices act as individual network node like pes, printer etc. where as in SAN, it is a separate network of storage devices that is accessible to the main network. Both NAS and SAN are scalable, so new devices can be added or removed without disrupting the network.

RAID (Redundant Arrays of Indepent Disks) : It is a storage method that uses several small hard disks in parallel to do the job of a large disk. RAID can be used to increase performance and to protect data on a storage server.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PC COMPONENTS, FEATURES AND SYSTEM DESIGN

SYSTEM COMPONENTS

These days we assemble PCs, that means many components are used to construct a system. They are integrated together with other components, so that they can perform many functions. The components required to assemble a PC system are as follows:

(a) Motherboard
(b) Processor
(c) Memory (RAM ROM)
(d) Case Chassis
(e) Power supply
(f) Floppy drive
(g) Hard Disk
(h) CD-ROM Drive
(I) Keyboard
(J) Mouse
(k) Video Card
(l) Monitor
(m) Sound Card
(n) Speakers

(a) Motherboard

Motherboard (or system board) is the main circuit board in the system. Motherboard is a flat board containing both soldered, non-removable components and sockets or slots for components that can be removed i.e., microprocessor chip/RAM chips and various expansion cards.

Motherboard is the core of the system, as everything else in the system is connected to it. Motherboard is also known as mainboard, system board and planar. There are several physical dimensions and sizes of the (i.e., form factor) motherboards available. There are some of the standard form factors such as Full sized AT, ATX, etc. and some are Proprietary designs. A modern motherboard has various built-in components, these are:

(i) Processor Socket/Slot
(ii) Chipset
(iii) Super I/O chip
(iv) BIOS
(v) SIMM/DIMM sockets
(vi) Bus Slots
(vii) CPU voltage Regulator
(viii) Battery

(i) Processor socket/slot - The CPU is usually installed in a socket. The socket specifications are generally numbered. This number is based on socket or slot number that is on the motherboard, this lets the designer of the system know that exactly what type of processors can be installed. Earlier before 486 was installed, sockets were not used to be numbered.

(ii) Chipsets - When IBM originally created PC motherboards, there were many discrete chips used to complete the design. There were simple logic chips which were used such as math co-processor. The primary chip components on the motherboard other than processor include clock generator, bus controller, system time, interputs and DMA controllers, CMOS RAM clock and the keyboard controller. All these components are usually Intel licensed except CMOS/clock chip (this is from Motorola). With the advent of Pentium series, systems sold use Intel processor as their motherboard have an Intel chipset on them, the individual chip is identified by a chipset number, that is stamped on it.

(iii) Super I/O Chips - Super I/O chip is the third major chip on most of the PC motherboards. Super I/O chip usually contains the following components such as floppy controller, dual serial port controller, parallel port controller. Super I/ 0 chip on modern PC System motherboard handles two drives. The biggest utility

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

of this super I/ 0 chip is that, as it contains everything into single component, therefore, it saves space and reduces part count on the motherboard.

(iv) BIOS - BIOS stand for Basic Input/ Output System. BIOS is a collection of programs embedded in EPROM (EPROM – Erasable programmable read only memory). When we start computer, this collection of program gets loaded into the computer even before operating system. This BIOS chip contains basically start up program and drivers that are used to get the system running and act as the interface to the basic hardware in the system. In most of the PCs, BIOS performs four main functions:

- **Power-on Self-Test (POST) -** It test computer's processor memory, chipset, video adapter, disk controllers, disk drives, key boards and other components.
- **Bootstrap Loader -** It is a routine that searches for operating system. If operating system is found, then next job of it is to load it into the system, which further enables the system functioning.
- **Basic Input Output System (BIOS) -** This is the collection of actual drivers that are used to provide a basic interface between operating system and hardware.
- **System Configuration and Set up Programs (CMOS)** - CMOS (Complementary metal-oxide Semi-conductor) chips contain flexible start up instructions, powered by a battery and thus don't lose their contents when power is turned off. This is normally a menu driven program that allows the user to configure the motherboard and chipset settings with date, time, passwords, disk drives and other system settings. In earlier systems, there were no set up programs in the ROM, rather, they used to be boot from a special boot disk. (When MSDOS was running OS)

(v) SIMM/DIMM Sockets - Originally, memory is installed via. Individual memory chip, thus, populating boards with such chips was a time consuming and tedious job. These chips were often referred as DIP (Dual Inline Package). These chips had another problem that they used to creep out of this socket as system got heated up and power was on. SIMM-Single Inline Memory Module or DIMM - Dual Inline Memory Module. Both use DRAM chips. A SIMM chips has RAM chips on only one side. A DIMM has RAM chips on both sides. DIMM is an alternative to individual memory chip. DIMM became more popular on Pentium-MMX system.

(vi) Bus Slots/Expansion Slot - The I/O bus or expansion slot enables the CPU of the system to communicate with I/O devices (or peripheral devices). I/O bus enables the user to add I/O devices to the computer to enhance its capabilities. It allows the basic components such as sound card, video card and specialized devices such as network interface cards, adaptors, etc., to be plugged into expansion slots.

Expansion is a way of increasing computer's capabilities by adding hardware to perform tasks that are beyond the scope of the basic system. Upgrading means changing the components (memory chips) to more powerful or sophisticated versions.

(vii) CPU Voltage Regulators - The Pentium and its earlier series, most processors and motherboards used to have a single voltage level. Generally it used to vary between 3.3 volts - 5 volts. As CPU's processing speed increased, it used less power, as that produces less heat, I/O voltage is generally higher (3.3V) with other chips on the motherboard. Thus, on Pentium MMX invented in Pentium series, these motherboard designs allow lower voltage CPUs to be used. The voltage regulator on the motherboard is what must be charged to supply the correct voltages to the processor socket. MMX version run on usually 2.8V. However, non-MMX Pentium run both CPU and I/O Pins at the same voltage.

(viii) Batteries/Power supply - All computers require power source in order to run the system. Although a great deal has been done to improve the power management capabilities of today's computer, the motherboard should support enhanced processor with Advanced Power Management (APM) and System Management Mode (SMM) that allows less power to be consumed by various system components. The latest standard for power management is called ACPI - Advanced Configuration and Power Interface that makes the motherboard support minimum power consumption.

(b) Processor/CPU

The Processor is also known as 'engine' of the computer. It is a sophisticated, miniaturized electronic component containing millions of transistors on it. It accomplishes the execution of the program. Although it is the smallest part of the computer system, it is the most expensive part.

(c) Memory (RAM)

The system memory is the primary memory that holds the program and data at a given time. RAM is a volatile memory i.e. as the power goes off or is turned off, whatever is in the memory gets cleared. The special memory known as ROM cannot be easily erased, thus it is also known as non-volatile memory. In a modern system, memory installed is either SIMM or DIMM form.

(d) CASE (Chassis)

The box that houses the computer is referred to as case. This part of the computer system is usually not given any importance, although it serves an important function for the PC, such as protection for the components of the system, maintain cooling, allow installation and access to the system components. The design of form factor of the motherboard and other components greatly depends upon the case.

(e) Power supply

Power supply is a device that converts AC to DC to run the electrical power required by every single part in the PC. The main function of the supply is to convert 110V into 3.3V, 5V that is required by the system for operations. Most new pentium-class processors are known as 'dual power plane' or 'dual voltage' processors, because their core voltage is different from that of the chipset and other devices. However, older pentium class required 3.3V to 3.5V input. These are called 'single power plane' processors and motherboards that accept that they must be sure to provide the proper voltage.

(f) Floppy Disk Drive

Floppy disks are the smallest and lowest form of off-line storage. Floppy disks allow a convenient way to transfer information to main memory. This can hold backup of small files, games, application/system software. With the advent of CD-ROM, the importance of floppy disk has been reduced. Modern PCs use 1.44MB capacity floppy drive.

(g) Hard Disk Drive

The hard disk is the archival storage memory for the system. It is known as hard drive because it consists of spinning platters of aluminum coated with magnetic media. The platters are available in various sizes, densities and in numbers, therefore, hard drives can be created with many different storage capacity. Usually, most of the desktop systems today use 3½ inch platter drives and Laptop computers use 2½ inch platter drives.

(h) CD ROM Drive

CD-ROM is abbreviated as Compact DISK-Read only memory. CD-ROM is identical to music CDs. As the name suggests, they are a read-only media, but now writable CDs are also available. In modern PCs with MMX versions CDs are most popular way to distribute applications, games, multimedia videos, music and large graphic files.

(i) Keyboard

The Keyboard is the widely used input device for the computers. It is used to enter /input text and commands into the computer system.

(j) Mouse

The invention of Graphical User Interface (GUI) led to the development of mouse. It is used in graphical environment, where user can just "Point and Click" to give instruction to the computer. It is simpler than keyboard as operations are easy to perform.

(k) Video Card

The visibility of the information is controlled by Video Card. All video cards consist of four basic parts - a video chip, video RAM, a Digital to Analog Convertor (DAC) and BIOS. The Video chip is responsible for controlling the information on screen, by writing data to the Video RAM.

The DAC serves the purpose of converting digital data into analog signals to drive the monitor by reading the information on Video RAM. The BIOS holds the routine software video driver, that allows the functions to be displayed during boot time. As the windows and its application software are complex, accordingly video drivers get enhanced.

(l) Monitor

The monitor is a specialized, high resolution screen, which varies as per the application. The video card is responsible for sending the contents of its video memory to monitor at the rate of 60 or more per second. The display screen is actually made up of red, green and blue dots.

They are illuminated by electronic beam, that reflects from behind. The video card and DAC chip basically determine the picture on the screen by controlling the electronic beam which further controls which dots would lit up intensity of brightness.

INPUT-OUTPUT DEVICES

I/O OPERATIONS

I/O operations are generally slow in speed as compared to operations carried out in memory and devices. This is because of the fact that functioning of input devices depends upon the speed of the user and speed of the output devices cannot match the processor's speed. This limitation of speed leads to the development of faster I/O devices. This section describes the ports required for serial and parallel I/O devices.

WHAT IS A PORT?

Port is a connecting socket, outside the system into which different types of cables are plugged.

I/O ports are the interfaces through which computers communicate with external devices such as printers, modems, joysticks and terminals. Physically, they are distinguished by their connectors, as they have contact pins of different shapes and numbers.

1) **Parallel Port** A parallel port transmits 8 bits of a byte of data in parallel. It is used for transmitting fast data over short distances.

   A parallel port has eight or more data lines for control of I/O devices. It is commonly used to connect printer to a computer. Since parallel port transmit an entire byte at a time, they operate I/O devices at relatively high speed. However, interference among multiple signals limits the cable to relatively short distances.

2) **Serial Port** Serial ports transmits one bit of a byte, one at a time as a simple stream of bits. It is meant for transmitting slow data over a long distance.

   Communication over a phone system is an example of serial communication. Serial port connects variety of devices such as modems, plotters, other computers, bar code readers and device control circuits.

3) **Universal Serial Bus (USB)** The USB port controller determines the devices by their characteristics and assigns each device a 7-bit address. It reserves one bit-pattern, thus, *USB port can connect up to 127 peripheral devices, theoretically it is used to transmit data to up to 127 devices such as digital cameras, digital speakers, scanners, modems, joysticks, etc*. USB also permits Plug and Play - configuring of expansion cards and peripheral devices as and when they are installed.

4) **SCSI Port** SCSI-Small computer system interface port allows data to be transmitted in a daisy chain to up to 7 devices at a speed (32 bits at a time) higher than those possible with serial and parallel ports.

   SCSI port is meant for transmitting fast data to up to seven devices in a daisy chain. The devices that may be connected are hard-disk drives, CD- ROMs drives, scanners, backup units etc.

5) **Fire Wire (IEEE 1394)** Fire Wire is a relatively new interface bus technology. The need of this arose due to large data movement of audio and video multimedia devices. The devices connected via 1394 are disk drives such as hard disk, Floppy, CD- ROM and new DVD drives. Microsoft has developed drivers to support 1394 in Windows Versions 95,98/NT. It is extremely fast with data transfer rate up to 400 Mb/sec.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

INPUT DEVICES: APPROACHES TO DATA ENTRY

An input device is an electromechanical device through which user feeds data and instruction to the computer. This further converts information into electronic pulses. To increase the throughput, the various approaches to data entry methods exist to maximize the use of high speed computer input devices by minimizing the human interaction. Thus, there is a need to categorize the data entry devices, these are:

(a) On-line or Direct Data Entry Devices: These are the devices which are under the control of CPU and communicate directly with the CPU, that is, these are connected via cable to the CPU. The devices under this category are classified as following:

Interactive input devices

- Keyboard
- Locator devices
- Pick devices

(b) Source Data-Entry Devices: You must have observed in the grocery stores the clerks merely wave the product over a scanner, which automatically enter the price in the computer whereas earlier prices were used to enter via keyboard. As a result they decrease the time, effort and errors associated with data entry. Thus data can be entered directly from the source, without human intervention. The devices used in this category are:

- Video digitizer
- Digital camera
- Scanning devices
- Magnetic stripe reader
- Sound sensor - Microphone
- MIDI Instrument

INTERACTIVE INPUT DEVICES

Interactive input devices can be classified into three categories: Keyboard, Locator Devices and Pick Devices.

a. Keyboard

The most common input device is keyboard. It is used for entering textual data into a computer file which may have been opened under the control of an editor or word processor. The modern keyboards are named as QWERTY keyboard because of the sequence of keys in the top row, and they are similar to those of standard typewriter. This allows the typist to work quicker as keys are not jammed, as frequently used keys are placed at a distance. Because of the electronic nature and speed of the computer, when data is entered via keyboard, computer monitor displays the same contents depending on the position of the cursor. This makes the documenting and formatting easier. The cursor is usually a vertical line, or may be an underscore, or a bright blinking video. The type of cursor is dependent on the application package.

The on-line data entry device keyboard generally contains 101 keys that are arranged in same order as a typewriter.

Uses A Keyboard is used for a wide range of tasks but is well known for its capability to enter data, which is copied from original paper documents (such as name and address details on a form). Copying data from paper documents is called transcription. Keyboards are also used for typing letters, memos, reports and other documents using word processing software.

Advantages

- Most computers have this device attached to it and a skilled typist can enter data very quickly.

Disadvantages

- It is very easy to make mistakes when typing in data.
- It can be very time consuming to enter data this way, especially if you hate not had much practice at typing.
- It is very difficult to enter some data, for example details of diagrams and pictures.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

b. Locator Devices

The position of the objects on the screen is indicated by Locator Devices. It also allows the quick movement of the cursor in text or graphic file. The most current position can be seen visually with an arrow or a cross hand displayed on a screen. Generally, locator devices are good for gross movement but not for precision work. The types of locator devices are the following: Digitizing tablet , Mouse, Track ball and Joystick

I. Digitizing Tablet

Digitizing tablet refers to an electronic plastic board, on which specifying a location corresponds to location on the computers screen.

Uses

Digitizers are used by architects and engineers in Computer Aided Designing (CAD) for designing purposes such as buildings, cars, mechanical parts, robots, etc. Moreover, they are used in the Geographical Information System for digitizing of maps.

II. Mouse

Mouse is used to position the cursor, in application software such as MS WORD and moving an object by dragging it into a new location. Mouse includes two or three buttons, they can be clicked to start or stop an operation or to indicate a chosen position. As mouse is primarily for gross movement, on the screen, thus it is not useful for operations that require high precision. Mouse requires a smooth and dust free surface, that is, a mouse pad. It is generally considered to be a relative locator because if, mouse or mouse pad are lifted, then location of the mouse on the screen is not disturbed. A mouse is classified as:

- Physical mouse
- Optical mouse

- **Physical Mouse** has a ball that protrudes, from its buttons, thus, when mouse is moved, friction causes the ball to move. There are a set of potentiometers inside the mouse that sense the direction and rate of movement of the ball. They are further converted into electrical signals that are interpreted by the computer.
- **Optical mouse** uses a Light Emitting Diode (LED) a light sensor and a special mouse pad. In this, mouse pad facilitates the operation of the mouse. It has a set of light and dark horizontal and vertical lines and a reflective surface. As the mouse is moved, light from the LED is reflected by the same pad to the sensor. This reflected light is darker at the horizontal lines than vertical lines. These pulses of light determine the direction and rate of movement. Optical mouse was introduced by Microsoft in 1999.

Uses

The mouse is used to select and move around items displayed on a screen. It is a point and click device and is used as a way of starting tasks by positioning the pointer and selecting items. It can also be used to create simple graphics.

Advantages

- Some people find it much easier to select items and start tasks by clicking icons or choosing from menus rather than typing in commands.
- It is often a faster method of getting to the tasks that user wants to start.
- It is quicker to use a mouse to help you move around a document than to use the arrow keys on the keyboard.

Disadvantages

- People knew to computers often find it difficult to control the movement of the pointer on the computer screen by moving the mouse around on the desk. They also have difficulty when double clicking the mouse button.
- Mouse balls can become very dirty and stop functioning correctly unless cleaned regularly.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

III. Trackball

Trackball is a movable ball mounted on a stationery device, which can be rotated manually using finger.

Trackball is like a physical mouse which can be turned upside down by the palm of the hand (and not rolling on the surface can be seen on laptops). Three dimensional track balls are available that allow the movement not only traditional left/right and forward/backward but also up and down as well.

IV. Joystick

Joystick is a remote control device for a computer which is merely used for playing video games to indicate position.

These are relative locators (joysticks), where the direction of push determines the change of position and amount of deflection determines a change of speed.

Joystick not only indicates two dimensional positions by moving the ball rationally but also three dimensional, where third dimension is indicated by the rotation of the rod.

Uses

Games software often takes input from a joystick.

Advantages

- There is an immediate feel of direction due to the movement of the stick.

Disadvantages

- Some people find the joystick difficult to control rather than other point and click devices. This is probably because more arm and wrist movement is required to control the pointer than with a mouse or tracker ball.
- Joysticks are not particularly strong and can break easily when used with games software.

V. Pick Devices

Pick devices are used to select an object on the screen. The object may be text or graphic. The examples of pick device are light pens and touch screens.

- **Light Pens** Light pen is a pen like light-sensitive device. It is connected by a wire to the computer terminal to detect the CRT beam when pointed towards the screen and generate a narrow electrical pulse that can be fed to the computer as input signal.

  Light pens may be used in place of mouse or key board particularly for menu based applications where an option may be selected by pointing on it. Light pens are also used for Computer Aided Designing (CAD) and drawing purposes.
- **Touch Screens (Title case)**

  Touch screens are monitors which detect where it is being touched. The user makes selections by touching the screen, rather than moving a cursor to the point on the screen with a mouse, joystick, touched or tracker ball.

USES

Touch screens are often used in situations where users are likely to have a low level of competence in using computer keyboards. Touch screens are sometimes used in restaurants, building societies and travel information systems.

Advantages

No computer literacy is required. It is simple to select your requirements by just touching a picture on the screen.

Disadvantages

Touch screens are not robust and can become faulty.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

SOURCE DATA ENTRY DEVICES

I. Video Digitizer

A 'Video-digitizer' captures television pictures from devices such as a TV- set, video camera or video recorder-and converts them into a format that a computer can use for display, storage or general manipulation. The format a computer can store and use is known as digital format because computers store and use digital (number) data. A video digitizer is the ideal tool for capturing images from real-life scenes to incorporate into computer work. Taking video pictures or frames is known as 'frame grabbing'.

Uses

Converting still or moving images into a digital format for use in a computerized presentation. A video digitizer may be required for video conferencing. (If a digital video camera is used there will be no need for a video digitizer.) This is when the video camera captures sound and images which are converted to a computerized format and transmitted over communication lines It can also be used to make television advertisements and pop videos.

Advantages

- Video digitizers allow us to capture real-life images which are often more appropriate then drawings.
- Captured images can then be transferred to paper.

Disadvantages

A fast computer with a large memory capacity is required to cope with the large amount of data involved.

II. Digital Camera

Digital cameras look like ordinary cameras but have sufficient memory to store images rather than using film. Each digital picture is made up of thousands of tiny dots, called pixels, and the camera stores data about the colour of each dot. Each image therefore takes up as much memory as is required to store the number of dots which make up the picture. If the picture is made up of a large number of small dots the picture is clear. If the picture is made up of a small number of large dots the picture will not be very clear.

The number of Dots Per Inch (DPI) determine the quality of the picture. Dots Per Inch (DPI) is called the resolution of the image. Because high- resolution images have more dots per inch, more computer memory will be needed. Most cameras allow the user to choose the resolution needed for a picture. Storage capacity varies but some cameras store approximately 100 images. Once the photograph is stored in the Camera it needs to be transferred to a computer where it can be edited, printed or more permanently stored. Some cameras store images on a floppy disk so that they can be easily transferred

Uses

Taking photographs and transferring them to a computer where they can be edited, used in documents and/ or printed out.

Advantages

- There are no expensive developing costs, no film is needed and you can insert images directly into a document on your computer.
- You can also edit, enlarge or enhance the images.

Disadvantages

- Digital cameras are generally more expensive than ordinary cameras.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

SCANNING DEVICES

a) Scanner

A scanner is another way through which we can capture still images or even text to be stored and used on a computer. The scanner shines a light onto the paper and measures the amount of light reflected back each part of the page. Like a digital camera the page is split into tiny pixels (dots) and a number representing the colour of each dot is sent to the computer. Scanner software usually allows the user to choose between a high and low resolution very high quality images taking up a lot of memory. Most scanners need a piece of paper to be scanned inside them when they are being used.

Uses

Transforming pictures or text from paper into documents that can be stored on the computer. For example, a scanner is often required to convert photographs which may be used in a desktop published school magazine.

Advantages

Any image can be converted from paper into digital format and later enhanced and used in other computer documents.

Disadvantages

Images take up a lot of memory space, but it is possible to reduce the size of the data file by reducing the resolution (number of dots per inch) or by using different ways of storing the data i.e., different file formats.

b) Optical Mark Readers (OMR)

Optical Mark Readers are able to detect marks on paper. A pre-printed document is prepared for users to select certain areas to insert a line of a mark. A document like this is scanned using reflected light to detect dark shadows.

USES

Universities and further education colleges often use pre-printed enrolment forms where students will enter a mark with a medium to soft pencil.

Lottery tickets also work this way. We put a line through an area on the ticket to select a number. Also multiple choice questions on examination papers often require a candidate to make a mark to indicate their answer.

The student will pencil a line through the option of their choice. Questionnaires and surveys may also use this technique.

Advantages

- Simply selecting and marking a choice in the right location on an OMR form is easier than typing in data.
- Documents can be scanned very quickly with only a few errors (only 2 or 3% of documents with modern OMR machines).

Disadvantages

- Badly damaged, creased or folded forms may be rejected.
- Because every option must be catered for the forms are sometimes not very easy for the user to understand. This may introduce errors.
- They can only be used when the data to be input can be selected by making with lines your choices.

c) Optical Character Readers (OCR)

Optical Character Readers (OCR) also detect patterns of marks which are in character format (numbers, letters, punctuation and some special characters like '-,@' etc.). The characters are more easily recognized if they are printed, not hand-written, because it is not easy to cater for every different style of handwriting. The characters are converted by a scanner from picture format into coded characters that the computer can work with. This is normally ASCII format (American Standard Code for Information Interchange). OCR software can be purchased separately for use with a standard A4 scanner.

USES:

OCR is often used to scan text for use in a word processor.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Advantages

When a computerized document has not been saved and only printed text is available, it is possible to use OCR to scan in the text for later editing and re-arrangement.

Disadvantages

Some characters cannot easily be interpreted and errors sometimes occur during conversion. This is especially true if there are a lot of diagrams and characters laid out in a way the software cannot interpret.

d) Magnetic Ink Character Recognition (MICR)

Characters which have been printed using ink which contains iron and can be magnetised may be read using a Magnetic Ink Character Reader device. This type of data collection has limited use and it mainly restricted to banking systems because of the expensive equipment involved. Cheques are pre-printed with the some code and account number and when a cheque has been written the data input clerk has to also mark the cheque amount in magnetic ink.

USES MICR is sued for Bank cheque processing.

Advantages

- Fast processing of large volumes of data.
- It is a fairly secure method of data collection as the shape of the characters cannot be changed without expensive equipment.
- This method of input is very reliable as number of errors generated during the reading of documents are negligible.

Disadvantages

MICR is expensive to use because special equipment is needed to create the characters as well as to read them.

e) Bar-Code Readers : Bar codes are basically a series to thick and thin vertical lines grouped together. The reader itself uses a laser beam to read the series of thick and thin lines which represent the bar code number. The bar code is 13 digits long and there are four main divisions. The first part of the bar code (two digits) represents the country, the second part represents the manufacturer's code (five digits), the third part represents the produce code (five digits) and there is a final digit, which represents the check digit. This last digit is a calculated digit to ensure that the bar code is read properly.

Bar code use: A bar code contains the inventory code. A typical inventory record stored on computer might contain: Inventory code, Product Description, Unit-cost, Unit-selling price, No. of items Re-order level; Re-order amount, Supplier-code.

USES: Bar codes are used on library tickets, airport baggage labels, supermarket products, clothes and on many other retail items.

Advantages

Fast, accurate data entry.

It is possible to store all the details about the country of origin and manufacturer as well as the product code itself. These are recorded in a standard format within the bar code.

Disadvantages

If the bar code is damaged the bar code reader cannot read it. It is then time consuming to enter all the separate digits using a keypad.

VOICE INPUT

We can now input data by simply talking into a microphone and using special software to recognize the voice. Software must be taught how each user pronounces the words before the speech can be interpreted and transferred to the computer. Some computer systems can respond to voice commands and can carry out tasks because the spoken word is interpreted by software and converted to instructions.

USES

- Inputting text to word processing software.
- Controlling devices such as electronically controlled doors and ma- chines.

Advantages

- Voice input is extremely useful for those users unable to use keyboards or mice.
- Software is now affordable by the home user

Disadvantages

- The system must be able to recognize the voice each user. Teaching the software can be tedious and time consuming.
- Speech recognition software is still not very accurate.

REMOTE CONTROL

Remote control devices are used for transmitting data when the user is some distance from the processor. Devices such as video recorders can accept data which is programmed into the system from a remote handset. Special keys for certain tasks or selections are usually found on the handset. An infra-red sensor on the main processing unit picks up the signals when keys are pressed.

Uses To select or enter data from a distance.

Advantages

- The device provides the user with the convenience of inputting and working away from the processing unit.

Disadvantages

- Small remote devices can easily be misplaced.
- The unit and remote need to be quite near to each other and have no other objects between them which may interfere with the signals sent.

MAGNETIC STRIPE READER

Magnetic stripes are the dark-looking stripe that may be seen on the back of many plastic cards for example, bankcards. The magnetic stripe holds data about the owner of the card, a bankcard will contain details such as the bank account number and sort code. When the card is used, for example to pay for goods in a store, the shop assistant inserts the card into the magnetic stripe reader and it is 'swiped' through a slot. The reader looks at the data on the magnetic stripe and money can be taken from the correct account to pay the bill.

Uses Magnetic stripe readers are often seen at supermarket and in many different types of shops, in fact anywhere where there is a 'point of sale'. The data is read electronically the point of sale is called EPOS (Electronic Point of Sale).

Advantages

- The card can be read many times without damaging the stripe.
- The data is not visible to people because a machine is needed to read the data (although with debit cards the bank details are actually displayed on the card as well).
- Putting magnetic stripes onto cards is not expensive so the cards are quite cheap to produce.

Disadvantages

- Magnetic stripes may get damaged or the magnetic stripe reader could break down.
- The biggest disadvantage of magnetic stripes is that ordinary equipment can read but not change the data. That is why Smart Cards are gaining more popularity.

SOUND SENSOR - MICROPHONE

A sound sensor, which is a microphone, detects sound and inputs it to the computer, and then the sound which is analog in nature is converted into digital format. We do this by 'sampling' the sound. Sound is produced as sound waves and not as numbers. Sound waves are in analogus format. To change them to digital format so that the data can be stored in the computer's memory we use a process called sound sampling. This process is carried out by software called a Sound Sampler.

The sound sampler measures a sound wave at fixed time intervals (deter- mined by the sound sampling software). Each measurement is given a digital value. Many different kinds of sound can be recorded,

including music and speech. This means that a large number of digital values must be stored to represent a sound if a good quality reproduction is required

Uses Microphones are necessary for speech recognition systems, which are now commonly used with word processors. Speech recognition software on the computer converts that the person has said into text. The text is displayed on the screen and can be saved as a word-processed file.

Advantages

- People with a range of disabilities can be benefited from the use of microphones and speech recognition, for example those who cannot use a keyboard. They can create word-processed documents by speaking instead of typing. Moreover, many devices in the home are now be controlled by voice.
- Recording our own sounds means that we can add spoken messages or music to our own presentations or even e-mails.

Disadvantages

- There should be no background noise while recording, otherwise sound may become distorted. Cutting down the background noise is sometimes difficult and the sound recorded using a microphone is often not very goods.
- Sound sampling (converting the analogue sound to digital format using software) often produces very large data files.

MIDI Instrument

MIDI means Musical Instrument Digital Interface. A serial interface standard that allow for connection of music synthesizers musical instrument and computers.

There are a variety of musical instruments such as keyboards, guitars and drums, which send and receive electronic messages. MIDI is also used to allow different instruments, e.g. an electronic piano and a synthesizer to communicate and work together.

Uses The music industry uses MIDI to input music directly into a computer so that it can be edited and developed, often by mixing it with other sounds which are input from a microphone.

Advantages

- Once the tune has been played on the musical instrument, all the details are held on the computer. These details may then be changed. It is possible to speed up the tune, slow it down or even make it sound like a completely different instrument.
- The data stored is very compact and takes up about 1/20th of the space taken up by even the lowest quality recorded sound samples.

Disadvantages

- A musician is required to play the instrument to acquire the input, so some musical knowledge is required.
- To take advantage of the digitally recorded sound, knowledge of software and music is required to edit recorded tunes, so the main disadvantage to using MIDI is that it has limited use.

OUTPUT DEVICES

Typically, the output operation comprises of activities such as, recalling results from memory and then passing it on to the output device in a readable form. The choice of output device depends on the need of the application and availability of output devices. Output devices fall into two main categories:

- Devices that produce output that is easily read by human beings, e.g. printers and display monitors.
- Devices that hold the data, which may be used for further processing. These are secondary/auxiliary storage devices, e.g. magnetic disks.

The output generated by the output devices may be of two types:

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Soft Copy Output: Such outputs are temporary, thus they are available as long as they are required. They may be displayed on screen or produced from voice response system.

Hard Copy Output: Such outputs are permanent, thus can be used at later date when they are required. They are usually on paper and are generally used to produce reports. The devices used are printers, graph plotters, computer output microfilm, etc.

1. **Soft Copy Output: Display Devices**

The display devices are terminals that may be used for entering data on-line from remote location and retrieving data.

a) Terminals

Terminal is a device through which end user receives information from a remote computer. Depending upon the processing capability, terminals may be classified as intelligent or dumb terminals. Dumb terminals are meant to enter the data and to receive output from main computer, as they cannot process data on their own. These terminals comprise of keyboard with monitor to display only alphanumeric characters. They are connected to the main computer through a communication link. Whereas intelligent terminals are usually personal computers, having their own local data processing capability, data storage and I/Ocapability. For a PC to be a terminal, a communication adapter (may be modem) is essential. If terminals are communicating to one or several hosts and peers, then a network program is also needed.

There are two types of intelligent terminals, namely, General purpose and Job-oriented.

Uses

VDUs or monitors are used to display output to the user as they carry out tasks. For example, when typing text into a word processor using keyboard, the VDU displays each character (or letter) as the keys on the keyboard are pressed.

Advantages

- As each task is processed the results can be displayed immediately on the screen.
- Output can be scrolled backwards and forwards easily, if results are too large to display on one screen. It is also possible, using certain software, to enlarge the display results.

Disadvantages

- The results displayed on the screen will disappear when the computer is switched off.
- Each new screen of information will replace the displayed output from the previous screen.
- Only a limited amount for information can be displayed at anyonetime.
- Screens are made of glass and can be very fragile.

b) Video System

The video output may be viewed on a video system i.e. VDU of a computer. Video system provides visual link between user and computer. Increasingly, personal computers (PCs) have replaced free-standing video display monitors. PCs have their own display monitor plus the added advantage of local CPU processing power and the capability of saving data, which supplements that of the host computer with which it communicates through a network.

The videos subsystem of a PC consists of two main components:

(i) Monitor

(ii) Video adapter (also called video card or graphic adapter)

(i) Monitor: A monitor is a display device which can produce text and graphic as output. A monitor requires an input. Basically the signals, that run to the user monitor come from a video adapter, which is inside the computer plugged into expansion slot. The expansion cards that produce video signals are called video cards or graphic cards. *A Cathode Ray Tube (CRT), contains a vacuum tube which is used to display on*

screen in a computer. Liquid Crystal Display (LCD) is a technology in which molecules of a liquid crystalline up in such amanner, that it create images on the screen by blocking or transmitting light on the screen.
This quantity is expressed in the number of horizontal and vertical picture elements called pixels on the screen. Good quality of a picture means greater number of pixels, usually resolution vary with screen size 640 X 480 or 800 X 600 resolution is generally for 14-15 inch monitor, whereas 1,024 X 768 or 1,280 X 1,024 resolutions bigger monitor such as 17-inch is preferable. It may be noted that desktop screen size is measured along its diagonal in inches.

RESOLUTIONS FOR DIFFERENT SIZES OF MONITORS

Monitor size (inches)	**Resolution**
14	640 X 480
15	640 X 480, 800 X 600
17	1024 X 768
19	1280 X 1024
21	1660 X 1200

Resolution = (maximum number of pixels in horizontal direction) X (maximum number of pixels in vertical direction)
Size of the monitor is measured along its diagonal in inches
Pixel refers to "picture element" and is the smallest unit on the screen which can be turned on or off
CRT versus Flat-panel display: *Flat panel display is a video display with a shallow physical depth based on technology other than CRT. Flat panel displays are generally electroluminescent light, thin, require low voltage and do not consume much power. In comparison to CRT displays, they are expensive and screen size is smaller and less commonly used.*
(ii) Videos Display Adaptor: - This provides interface between computer and monitor. It transmits the signals so that images are displayed. The PC video display technologies existed given below:

- **Monochrome Display Adaptor (MDA) -** This was the first adaptor introduced by IBM in 1981. MDA was capable of only one video mode i.e. a character mode with 25 lines of 80 characters each with underlining, blinking and high intensity characters.
- **Colour Graphic Adaptor (CGA)** enables the system to produce both text and graphic in colour and monochrome mode with various resolutions.

This was also introduced by IBM in 1981, having two combinations of resolutions in graphic mode. These are: (a) 640 X 200 pixels on screen with 16 colours

(b) 320 X 200 pixels with 4 palettes where each palette has 4 different colours and one palette is used at a time; and Character mode with 80 horizontal characters by 25 vertical lines with 16 colours.

- **Enhanced Graphic Adaptor (EGA) -** It supports all the features of MDA and CGA with higher resolution on digital monitor. An EGA supports a higher resolution with 640 X 350 may be upto 16 colours at a time. It was introduced by IBM in 1985.
- **Video Graphic Adaptor (VGA)** provides a high quality resolution to graphic or images and improved text display facility. VGA supports following resolution and colour combination, these include:

(a) 640 X 480 pixels with 16 colours
(b) 320 X 200 pixels with 256 colours.

- **Super Video Graphic Adaptor (SVGA)** is an enhancement over VGA as SVGA supports various resolutions and colours to graphic/images and text-these include:

(a) 640 X 480 pixels with 265 colours (or more up to 65536 colours)
(b) 1024 X 768 pixels with 256 colours
(c) 1280 X 1024 pixels with 16 colours.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- **Extended Graphic Array (XGA)** was introduced by IBM in 1987. It has enhanced feature over EGA having higher resolution and colour i.e. 640 X 480 pixel with 256 colours and text mode in 80 X 25 with 16 colours.

c) Audio Response Unit

As a voice recognition system allows the user to talk to a computer, a voice response system permits a computer to talk back to the user. All the sounds needed to process inquiries are prerecorded on a storage medium. Each sound is given a code. When inquiries are received, the processor follows a set of rules to create a reply message in a coded form. This coded message is then transmitted to an audio-response device, which assembles the sounds in a proper sequence and transmits the audio message back to the station requesting the information.

Advantages

- A user cannot see the result on the computer screen, but can know the result even when they are located far away from it, or when they are visually handicapped.

Disadvantages

- The quality of voice output is getting better but it still sounds like a set of small sounds strung together and is therefore obviously electronic and rather unnatural.
- Sound files used for voice output occupy a large amount of memory.

PRINTERS

A printer is an output device, that prints characters, images and symbols on paper. Printers may be classified on the basis of printing technique. Itmay be impact or nonimpact.

Impact printers are like typewriters, as they use some sort of striking action to press carbon or fabric ribbon against paper to create a character. Impact printers generally print black and white, are slow and produce sound while printing. Moreover, high resolution graphics cannot be produced by them. The most common impact printers are character printers (e.g. daisy wheel, dot matrix) and line printers (e.g. chain printers, drum printers).

Non-impact printers overcome the limitations of impact printers. They use thermal, electrostatic, chemical, or inkjet technology to produce printed output. They are fast and quite in operation. The most common non-impact printers are inkjet printers and laser printers. They generally produce colourful and high resolution output.

The quality and speed of colour printers is increasing as their cost is decreasing. Colour printers are more and more becoming the only kind used in home and office alike.

Non-Impact printers form characters or images without having physical contact between the paper and printing mechanism.

Difference between Impact printer and Non-Impact printer

Impact Printer	Non-Impact Printer
Works like typewriter and use striking action	They use thermal, electrostatic chemical and inkjet technologies.
Prints in black	Produces colourful output
Produces sound	Quite in operation
Works slow	Fast in operation
High resolution graphics cannot be produced.	High resolution graphics can be produced.
Examples are line printer, Character printer, Chain printer, Drum Printer	Examples are laser printer, Inject printer, Plotter, etc.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Printers (laser, ink-jet and dot-matrix)

Printers provide the hard copy output. There are generally three main types of printer in use today: Laser, ink-jet and dot-matrix.

A. Laser Printers (Non Impact Printer)

Laser printers produce high quality printed documents at fast speeds (on average at 10 pages per minute). Laser printers use cartridges of fine black (or co loured) powder which is transferred to the paper to produce the printed document. This powder is called toner. Toner cartridges are expensive but they last a long time, making a page printed by a laser printer much cheaper than a page printed by an inkjet-printer.

Uses

Laser printers were first used in offices where it was important to produce high quality printouts quickly and quietly. Now prices are lower and they are used anywhere where a large amount of printing is done including schools and many businesses.

Advantages

- Laser printers are fast and almost silent, producing high quality output.
- A large number of printouts can be done using a single toner cartridge, so frequent purchasers of replacements are not necessary.

Disadvantage

- Laser printers are usually the most expensive to buy.
- Laser printers are also often quite large and are heavier than other types of printer.
- Repairs are generally expensive.
- The cost of toner cartridge is much higher than the cost of ink cartridges or ink ribbons used in other types of printer.

B. Ink-jet printers (Non Impact Printer)

Ink-jet printers are very popular because they are relatively cheap to buy and usually offer colour printing. They are also very quite in their operation. However, they tend to be slower than most laser printers.

These printers also have cartridges but the cartridges contain ink and are quite small. They operate by heating the ink as it flows through a nozzle. The heating process causes ink to form a small droplet. This is then released as a single dot which forms part of a letter or image. There are many dots so the quality is usually very good but producing each droplet of ink can be slow. Ink-jet printers vary in price with the slowest ones being the cheapest printers available and the faster ones costing a similar amount to a small laser printer. The main disadvantage of an ink-jet printer is the cost of buying the cartridges. They are much cheaper to buy than laser toner cartridges but they do not last long. If lot of printing is not required then, an ink-jet printer is probably a better buy than a laser printer, especially if colour printouts are required.

Uses

Home users who do not need to do a lot of printing, but occasionally may need to print out in colour, often choose to buy an ink-jet printer.

Advantages

- Ink-jet printers are relatively inexpensive to buy, especially if colour is needed.
- Ink-jet printers are also lighter in weight than other printers and some are small enough to be carried around in the pocket of a notebook/ portable carry case.
- Ink-jet printers are very quite while they are printing except only a faint hiss and the movement of the paper being heard.

Disadvantages

- Ink-jet printers are slower when they are printing than laser printers.
- If let unused for long periods of time the cartridges can easily dry out.
- Replacing colour cartridges can be a particularly expensive running cost.

C. Dot-matrix printers (Impact Printer)

Dot-matrix printers are not used much today. They were the first type of printer to be used extensively in homes and schools. These pictures are very cheap alternative to ink-jet and laser printers, with much lower running costs. Dot-matrix printers print by striking tiny pins against a carbon ribbon. The print head contains a grid (or matrix) of these pins and different combinations of pins are pushed out to form different characters. This makes them much slower than ink-jet printers and they are very noisy because we can hear the pins hitting the paper. Any printer that works by striking something against the paper in this way is called in impact printer.

Uses

Dot-matrix printers are ideal when carbon copies are needed. This is because they are impact printers and the print head bangs on the top copy of paper the carbon paper transfers a copy to the paper beneath the carbon paper. They are often found in factories and garages where invoices and delivery notes are produced. Often the factory wants to keep a copy and the quality is not important.

Advantages

- The main advantage of dot-matrix printers now is the ability to produce carbon copies. They provide a cheap, reliable method of getting a second copy of printouts.
- The other advantage is the low cost of printouts, because printers are cheap to buy, and running costs are low.

Disadvantages

- Dot-matrix printers are noisy, slow and produce a lower quality output.

D. Plotter

Plotter is a special output device that is designed to produce sharp-qualities and large sized graphics in different colour schemes.

A plotter (sometimes called a graph plotter) is an output device, which produces high quality like diagram on paper. Pen are used to draw lines on the paper, which is placed in the plotter. Some plotter have a flat area to put the paper onto, others use a large roll of paper. Usually plotter can be larger sized sheets of paper than printers.

Uses : Graph plotters are used for drawing building plans, graphs and three- dimensional drawings. They are often used by architects and by engineers designing machines, bridges etc.

Advantages

- Drawings are of the same quality as if an expert drew them.
- Large size sheets of paper can be used than would be found on most printers.

Disadvantages

- Plotters are slower than printers, drawing each line separately.
- They are often more expensive to buy than printers.
- Although drawings are completed to the highest quality they are not suitable for text although text can be produced.
- There is a limit to the amount of detail these plotters can produce, although there are plotters which are pen-less and are used for high- density drawings like printed circuit boards.

CONTROL DEVICES (lights, buzzers, robotic arms, motors)

Many other devices may be connected to the computer to produce output. For example, lights may be connected to a computer which will respond to electrical signals that tell them to switch on and off. Buzzers may also be used for sound as the computer sends an electrical signal to switch them on and off.

Also, a computer can control all sorts of devices which are operated by motors. Robotic arms are an example of a device which has its movement controlled by the computer. When an output device moves something it is called an actuator.

Uses : Control devices like lights, buzzers and motors are used when the computer is controlling a situation such as traffic lights or cars assemble.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Advantage

- Devices which need controlling, because they use electrical signals from a computer, can be controlled very accurately.
- Using a computer to control devices means that few peoples are required for manual jobs and less money is needed to pay wages.

Disadvantage

- If the computer or control device goes wrong, the system comes to a halt.

COMPUTER EQUIPMENT MAINTENANCE

Computer maintenance entails a four-fold attack against deterioration of hardware, security threats, software bugs and performance degradation. Each can seem overwhelming or even insignificant, but considering the importance of computer systems operations, the cost and time are trivial compared to the hassle of getting a system up and running after a crash. In fact, we should upgrade the entire computer every three years or so, but we will have to transfer all the data and software from the original computer, whereas a hard-drive swap is a transparent upgrade.

Preventative Maintenance

The basic preventive maintenance feature are:

- Excessive dust buildup inside your computer can cause overheating, which degrades performance as well accelerates wear and tear on interior parts.
- The keyboard can be cleaned by turning the keyboard upside down and gently tapping the bottom of the keyboard.
- Monitors can be cleaned with a damp cloth.
- The mouse requires regular maintenance as the tracking ball under the mouse collects dust as it rolls inside the mouse. To clean the mouse, turn it upside, remove the panel, clean the tracking ball and the rollers inside the mouse.
- Printers require routine maintenance as well. Clearing out paper jams, changing cartridges and cleaning the printer will lengthen the life of your printer.
- Consumables like floppy disks, CD's and paper require correct storage to ensure that they will be usable when required. They should be stored in a clean and dry area.
- To ensure that the moisture content of the paper is kept it an optimum level, paper should be stored in a sealed container. Moist paper tends to curl and jam in the paper path of a printer.
- It generally has a number of layers on the lid for strength and protection. A number of vents allow for a controlled airflow from the power supply and processor fan to keep the components at a satisfactory operating temperature.

Effect of Negligence to Computer Equipment

- **Loss of Money:** The risk of loss to computer equipment caused by mechanical failure of hardware components. When losses occur, the cost of repairs is absorbed by the affected organizations budget.
- **Waste of time and reputation:** Gross negligence, result in physical damage, defacement, or destruction of computer equipment; theft; unauthorized access or use; unauthorized alteration, substitution or deletion of programs, command files, data files, documentation, or other material; divulgence of confidential information or computer access methods to unauthorized persons.
- **Performance Degradation:** Computers often have too many temporary files cluttering the system, a fragmented file system, viruses/spyware/ adware, and unnecessary programs running on startup. Windows XP has several software packages to help clean up systems, including Disk Cleanup, Disk Defragmenter and Check Disk. Each speeds up computers and is easily used by novices.

DATA REPRESENTATION: COMPUTER ARITHMETIC AND CODES

A computer can understand the positional number system where there are only a few symbols called digits and these symbols represent different values depending on the position they occupy in the number.

The value of each digit in a number can be determined using –

- The digit
- The position of the digit in the number

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- The base of the number system (where the base is defined as the total number of digits available in the number system)

REPRESENTING DATA

The different data types can be represented in the computer using bit patterns.

Text and Numbers: Text is a sequence of symbols used to represent thought or expressions in a language. For example, Human made reports contain symbols from the set of alphanumeric characters (such as alphabets A-Z, a-z, Numbers 0-9, +, -, ?, $, --). But computer cannot understand these, because computer is an electronic machine. Therefore, it is a bistable device means with two states "on" and "off". Thus computer deals with data converted into simplest form which can be processed electronically, that is binary form, where it substitutes "on" with 1I" and "off" with "0". To store and process data in a binary form a coding scheme had been devised to represent alphanumeric and special characters as standardization methods. These are:

1. ASCII (American Standard Code for Information Interchange):

This code is most widely used with micro-computers. The code represents English characters as numbers. ASCII was established to achieve compatibility between various types of data processing equipment and is now used as the common code for microcomputer equipment. The highlights of this code are:-

(1) It is a 7-bit code, it can accommodate 128 (=27) different characters.

(2) In this, first three bits are zone bits and last 4-bits indicate the digit.

(3) This code is generally used in microcomputers, as they use 8-bit byte and ASCII is 7-bit code, therefore, it leaves leftmost first bit of each byte as zero.

(4) 31 Control non-printable characters.

(5) 0-9 numeric Characters.

(6) 7 printable Characters.

(7) A - Z upper case letters, a-: lower case letters.

Example: Write binary coding for the word LOVE in ASCII-7. Also find the number of bytes required to represent the same.

Solution:

L = 1001100 in ASCII-7 binary notation.

0= 1001111 in ASCII-7 binary notation.

V= 1010110 in ASCII-7 binary notation.

E= 1000101 in ASCII-7 binary notation.

Hence binary coding for LOVE in ASCII-7 is

1001100	1001111	1010110	1000101
L	O	V	E

There are 4 characters in the string LOVE. Since each character requires one byte in ASCII-7 representation, thus 4 bytes (= 2S-bits) are required for this representation.

(2) ASCII-8

It is extended version of ASCII-7, because it is 8-bit code, thus it allows 256 (=28) different characters. In this code first four bits indicate zone bits and last four bits are for digits.

Example : Write binary coding for the word LOVE in ASCII-S. Also find the number of bytes required to represent the same.

Solution:

L = 10101100 in ASCII-S binary notation.

0= 10101111 in ASCII-S binary notation.

V= 10110110 in ASCII-S binary notation.

E= 10100101 in ASCII-S binary notation.

Hence binary coding for LOVE in ASCII-S is

10101100 10101111 10110110 10100101

L O V E

There are 4 characters in the string LOVE. Since each character requires one byte in ASCII-S representation, thus 4 bytes (= 32-bits) are required for this representation.

(3) EBCDIC

This coding scheme was developed by IBM, extended from 6-bit BCD (Binary-Coded Decimal) code to an 8-bit code by adding 2-bits in zone bit, expanding zone to 4-bits. This binary coding scheme is known as Extended Binary-Coded Decimal Interchange Code (EBCDIC) generally, used with large computers such as mainframes. In this coding scheme, it is possible to represent 256 (=28) different characters. It allows a large printable and non-printable control characters (e.g. Cursor movement, spacing, etc.) along with 64(=26) characters. However, this code is not used in any computer other than mainframes.

IMAGES: Images are represented in a computer using anyone of the two methods: **bitmap graphic or vector graphic**

Bitrnap graphic is comprised of matrix of pixels (picture element) where each pixel represent a small dot. The size and number of pixels determines resolution. For example, if an image containing 1000 pixels will have better resolution and better representation than the one containing 100 pixels. Further, each pixel corresponds to a bit pattern. A pattern of 0 represents a black pixel and a pattern of 1 represents white pixel. Thus when such patterns are recorded and stored in the computer, these form a black and white image. Images can be represented on a gray scales using following combinations: 00 represent block pixel; 01 a dark gray pixel, 10 a light gray pixel and 11 by white pixel. In similar manner, colour images formed from coloured pixel is decomposed into three primary colours red (R), green (G) and blue (B), forming a bit pattern of 8 bits. Where each pixel consists of three bit patterns to represent intensity of red colour, blue colour and green colour.

For example,	R	G	B
Red with 100% intensity	11111111	00000000	00000000
Green with 100% intensity	00000000	11111111	00000000
Blue with 100% intensity	00000000	00000000	11111111

This method has a drawback that it stores exact bit pattern for an image later if we resize that image, the size of the pixel also gets changed which creates a ragged look. An alternative to this solution is vector graphic.

Vector graphic forms an image by contribution of curves and lines. Each curve or line is represented by a mathematical formula. These formulas are stored in a computer. When we resize or redesign the image then system use th ame formula to draw the image. In this case, each time the image is redrawn, then corresponding formula is re-evaluated and fineness of image system retains.

AUDIO DATA: Both sound or music represent audio data. Audio is analog data, by nature it is continuous, not discrete (digital). There are no standards to sound or music, therefore audio is converted into digital data using bit pattern to them. It involves the following steps:

(i) The analog signals are sampled by measuring the value of signal at equal intervals

(ii) The samples are quantized by assigning a value to a sample

(iii) The quantized values are changed to binary patterns and stored in the computer memory .

VIDEO DATA: Video is a representation of movie in a series of frame shown one after another in a given time. Thus, each image or frame is converted to a set of bit patterns and stored. The video is thus a combination of images. They are stored as compressed videos using compression techniques.

COMPUTER SOFTWARE

Hardware

Hardware refers to the physical elements of a computer. This is also sometime called the machinery or the equipment of the computer. A computer's hardware is comprised of many different parts, but perhaps the most important of these is the motherboard. The motherboard is made up of even more parts that power and control the computer.

In contrast to software, hardware is a physical entity. Hardware and software are interconnected, without software; the hardware of a computer would have no function. However, without the creation of hardware to perform tasks directed by software via the central processing unit, software would be useless.

Hardware is limited to specifically designed tasks that are, taken independently, very simple. Software implements algorithms (problem solutions) that allow the computer to complete much more complex tasks.

Firmware

Firmware is a very specific, low-level program for the hardware that allows it to accomplish some specific task. Firmware programs are (relatively) permanent, i.e., difficult or impossible to change. From the higher-level view of software, firmware is just part of the hardware, although it provides some functionality beyond that of simple hardware.

Firmware is part of devices (or device components) such as a video card, sound card, disk drive and even the motherboard. The AMIBIOS image to the right is from a Baby AT Motherboard.

Liveware

Liveware a slang term for people, which is used to distinguish them from hardware, software and firmware.

Liveware may be the end users, programmers, application programs. They are basically the people who use hardware and software (or computer system) to develop, maintain and use the information system residing in the computer memory.

Software

The word software originated in 1960 to describe non-hardware component of the computer system. Primarily, it is the program that user needed to make the computers perform their intended job. The software is generally used in connection with all kinds of programs that are not limited to one particular task or application. The software may be of system software, of software system, of mathematical software, of business application software.

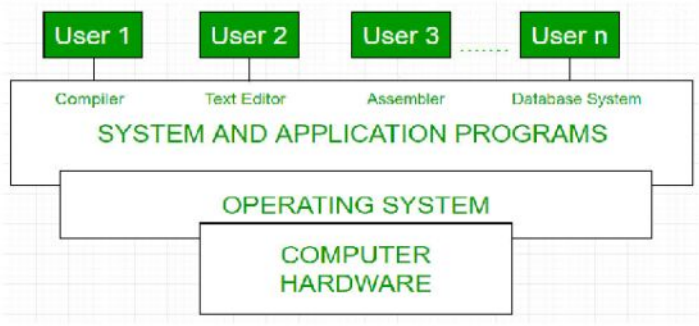
Software, commonly known as programs or apps, consists of all the instructions that tell the hardware how to perform a task. These instructions come from a software developer in the form that will be accepted by the platform (operating system + CPU) that they are based on. For example, a program that is designed for the Windows operating system will only work for that specific operating system. Compatibility of software will vary as the design of the software and the operating system differ. Software that is designed for Windows XP may experience a compatibility issue when running under Windows 2000 or NT.

Software is capable of performing many tasks, as opposed to hardware which can only perform mechanical tasks that they are designed for. Software provides the means for accomplishing many different tasks with the same basic hardware. Practical computer systems divide software systems into two major classes:

- **System software:** Helps run the computer hardware and computer system itself. System software includes operating systems, device drivers, diagnostic tools and more. System software is almost always pre-installed on your computer.
- **Application software:** Allows users to accomplish one or more tasks. It includes word processing, web browsing and almost any other task for which you might install software. (Some application software is pre-installed on most computer systems.)

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

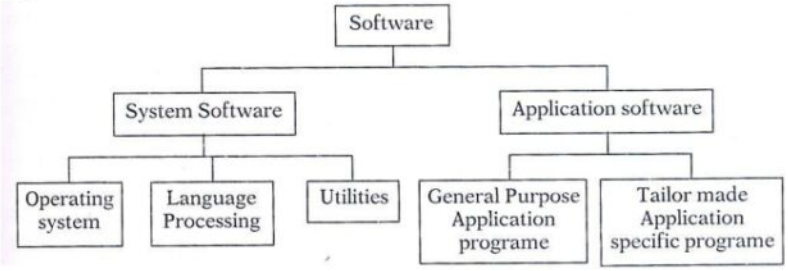
RELATIONSHIP BETWEEN HARDWARE AND SOFTWARE



TYPES OF SOFTWARE SYSTEMS

Software are usually divided into two major categories,

- System software
- Application software



SOFTWARE HIERARCHY

The application software may be split to obtain a third category known as General Purpose Software. But, both application software and system software are written in some coding schemes, known as **programming languages**. The programming language is the basic entity of computer software.

SYSTEM SOFTWARE

System software is a set of one or more programs, which are designed to control the operation and extend the processing capability of the computer system. It also acts as intermediary between computer hardware and application program.

The system software primarily controls and supports the computer hardware and its information processing activities. In general, computers system software perform the following functions as well :-

(a) It provides a platform to develop and execute application software (such as Word, Excel, Tally etc.)

(b) Monitor and control the effective use of hardware resources such as memory, CPU, I/O devices etc.

(c) It not only controls the operation of I/O devices but also establish a communication between devices.

The System Software are primarily of three types, given below:

(a) Operating System/user interface.

(b) Language processors (compilers interpreties, assemblers, linkers).

(c) Utilities.

OPERATING SYSTEM

An operating system is a set of integrated programs that manages overall performance and functioning of the computer system by controlling the resources such as CPU, memory, I/O devices, and overall flow of information within the system. It provides an interface between machine and its users.

Operating systems are designed and developed to optimize the man-machine capabilities. Programs are held in the computer memory, it makes operator free from inputing a program for each application. Operating systems are also known as monitor system or supervisor program.

The operating system is generally the most complex software in a computer system with two basic purposes:

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(1) Efficient use of hardware.

(2) Easy use of resources.

Location and Functions of an Operating System

Location : Operating system being an integrated system of program is usually stored on the computer's hard disk/ external memory. These days modern computer operating system consists of millions of lines of codes. Thus because of its size developers have found it necessary to structure OSs, so they do not swamp the computers. Therefore, OSs are engineered in such a way that they operate in sections and are loaded in computers main memory as and when needed.

Functions: As we know that an operating system provides services to all application software and to the users of those programs. The functioning may vary from one operating systems to another, but generally, there are some common types of functions performed by all operating systems.

These main functions provided by usually most operating systems are as follows:

(i) User Interface is the part of the operating system that allows user to communicate with the computer so that we can load programs, access files and accomplish other tasks. It provides two main types of interfaces:

(a) Graphical-user Interface (GUI)

(b) Command line interfaces

Today the most common interface is Graphical-User Interface. For such interfaces pointing devices like mouse is used to make selections.

A GUI is also known as a WIMP interface (Windows, Icons, Menus, Pointer) and this is because each task is displayed in an individual window. Windows can also display the contents of folders (directories), which may contain software like a word processor or documents like a letter. A picture (icon) represents each item such as a document or piece of software. An icon is just a small graphical picture. (Note that a folder usually contains a selection of documents, programs/software and perhaps, other folders).

In addition at the top of some windows some words may be observed which, when we click on them, it displays a series of options to click on. These are called 'Pull down Menus.

An alternative user interface might be a command line interface. There are no windows or graphics just command line prompt such as : 'C>', This indicates that the computer is waiting for an instruction. The user must learn the command necessary to carry out a particular task. For example to display the contents of a folder, which is equivalent to a directory, we may have to type 'DIR' and press ENTER or FAT and press ENTER.

Comparing Command Line Interfaces and GUIs

The main advantage that GUI have over command line interfaces is that they are easy to use even for inexperienced users because they are very 'user friendly:

They are however slow in some ways because sometimes a sequence of actions is necessary in order to find one item or do one task. For example if we wish to find a document which is stored in folder it may be necessary to operate several mouse clicks and open several windows in order to find it. One of the biggest problems that users have is finding documents. Also if we use a pull down menu there may be several other sub-menus to negotiate before we can select the option of chosen menu.

The main advantage of Command Line Interfaces is that experienced users find them very quick to work as compared with GUIs. For any new software, user need not required to knows all the commands to run the system in CUI interface.

(ii) Resource Management: An operating system uses a variety of programs to manage the hardware resources of a computer system, including its CPU, memory, secondary storage devices and input/output peripherals. Thus, it include the following tasks management:

(a) Input/output management - Operating systems manage the input to and output from a computer system. It also manages the flow of data/information from storage devices to memory and vice versa, also to output devices such as printers; terminals and other computers.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(b) Memory Management - Memory management program of OS keep track of memory where data and program are stored. This also subdivide memory into a number of sections and perform swapping between parts of large programs and data between memory and magnetic disks or other secondary storage devices. This provides a computer system with a virtual memory capability which is significantly larger than the real memory capacity of its primary storage unit. In order to improve the CPU utilization and to provide critical response time to users both program and data it accesses are loaded together in the main memory. For this the memory management program takes care of allocation and deal location of memory space.

(iii) File Management : An operating system contains file management module it's primary job is to control the creation, deletion storage and access of files data and program. File management also keep track of physical locations of files on primary and secondary storage devices operating. Moreover operating system maintain information about the location and characteristics (naming, organization, protection etc.) of the files stored on secondary storage devices (i:e. floppy, disk, CDs etc.).

(iv) Task Management: The task management module of an operating system manage the computing tasks of end users. The tasks are managed primarily by CPU management task management involve giving each task a slice of a CPU time so that several computing tasks can occur at the same time where i.e. an operating system must possess multitasking capability where the computing task of several users can be processed at the same time in time-sharing mode. The efficiency of multitasking operations depends on the processing power of a CPU and the virtual memory and multitasking capabilities of the operating system it uses.

(v) Support Services Management : System support service module of operating system manage the following functions:

(a) Security - As modern computer system can store large amount of information and this information is an assets to an enterprise (business value of the information may be computed). Security feature of OS ensures the protection of information against destruction, disclosure and unauthorized access.

(b) Command Interpretation - As we know OS provides user interface, thus the commands issued by the user must be interpreted. The command interpretation module of an operating system takes care of interpreting user commands and instruct the resources to handle the request.

Other than this it also maintain the following:

- maintain the account of usage for each user
- maintain the log of system i.e. files accessed or updated
- maintain system date and time
- Schedule jobs using job control language.

VARIOUS OPERATING SYSTEM

MS/PC-DOS - The first version of Microsoft Windows OS was launched in 1983 like Operating System 2, the original release was not very successful, but Microsoft continued to develop the program and then in 1980 it launched Windows 3.0. The program became bestselling 16-bit GUI-OS. Windows 3 enabled true multitasking, allowed users to access programs written for MS/PC-DOS as well. It allowed user to employ increased RAM. Windows 3 resulted in an increased processing capability of the PC, To optimize its efficiency, the minimum configuration of the PC required to be 4MB of RAM along with 386 processor.

Windows 95 - In a series of windows up gradation 3½ years persistence resulted in Windows 95, a 32-bit OS was released in August 1995. It was extremely heavy task as far as computer projects go and approximately 75 million hours of testing was performed prior to its release. A 32- bit OS differ to a 16- bit OS in terms of amount of internal memory that can be directly accessed by the user/program. For example, with a 16-bit version of MS-DOS, the maximum amount of internal memory which can be directly accessed is 1MB, whereas with a 32-bit OS user can access up to 4GB of main memory. Thus, in order to run Windows 95, the minimum configuration required is :a computer equipped with 386 DX or higher processor with minimum of 4MB or 8MB memory along with hard disk 50 Mb, 3.5 inch disk drive and a CD ROM.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

WINDOWs NT : Windows new technologies also possess all the above features as Windows 95, but it has increased memory, security features, Network Utilities and port abilities because it can operate on other types of proces-sors as well. Windows NT are usually expensive.

Windows Vista : Another major release of Windows is Windows Vista. Previews of Windows Vista reveal a new graphical interface that has the core functionality of the Windows XP interface but a slightly different appearance. Windows Vista is expected to include more visual features and to have intuitive file management capabilities that allow documents to be created or organized based on characteristics, such as author, subject, keyword, category, artist, song title, and so forth. Vista also include improved search capabilities and better support for peer-to-peer searches and file sharing. In addition, Vista is expected to directly support DVD burning and more secure, easy to manage and update, and it will support advanced document technologies, such as XML (extensible markup language).

Mac OS : Mac as is the proprietary operating system for computers made by Apple Corporation. It is based on the UNIX operating system and originally set the standard for graphical user interfaces.

There have been a number of different versions of Mac OS since the original Apple Macintosh Computer in 1984. Mac as allows multithreading and multitasking; it also supports dual 64-bit processors and a high level of multimedia functions and connectivity. New features include a new search feature, called Spotlight, which allows users to search through documents, folders e-mail messages, and programs based on a variety of criteria, and then continues to update those searches and add related files as the user works. A new Dashboard feature hosts a variety of mini applications called wigets that can be used to receive timely information from the Internet.

Linux : Linux is a flavour of UNIX originally developed by Linus Torvalds in 1991. The operating system was released to the public as open source software; that is, a program whose source code is available to the public and can be modified to improve it or to customize it to a particular application. Some versions of Linux are available as free downloads from the Internet; users are also permitted to customize Linux and sell it as a retail product.

Although Linux originally used a command line interface, most recent versions of Linux programs use a graphical user interface. Engineers, and other technical workers who are already familiar with Linux or UNIX, agree that personal computers running Linux crash less often, are more secure and less prone to viruses and other security hazards, and are easier to run on older equipment. In addition to the growing use of Linux on servers and desktop PCs, it has been suggested that the next biggest growth area for the Linux operating system might be in lower-end personal computers, Internet applications, and point-of-sale terminals.

FILES AND DIRECTORIES

A File is named collection of a program, a data set used by a program, a user created document that exists in computer's secondary storage.

A Directory is a way of organising and grouping files, it works as catalogue for filenames and other directories stored on a disk.

In order to access a file, it must have a name which is recognised by the system. File name consists of two parts: name/ prefix and suffix/ extension, The prefix is supplied by the user and extension is supplied by the computer application currently being used. Extension is usually separated by a dot or a (full stop). File name plays a significant roll as it indicates the purpose, type of data stored in the file, Some examples of file name along with their extensions indicate the type of data stored, are listed below :-

File Name	Data type
STU.BAS	Basic Program (.BAS - Basic)
MARKS.DAT	Contains data only (.DAT - Data file)
REPORT.DOC	Document file (,DOC - Document)
PIC.IMG	Image file (.IMG - Image)
PRESENTATION.PPT	Powerpoint file or presentation file

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

APPLICATION SOFTWARE

Application software is a sequence of instructions coded in a programming language that directs computer hardware to perform specific data that provides functionality to user, Example of such programs are payroll, general accounting, inventory control. The application program may be customized or tailor-made according to their specific purpose. The special purpose programs are also known as "packages". There are general purpose application programs which support generally used information processing instead of specific task.

Some of the general purpose tools are development tools, that are used to build application programs. The most widely used general-purpose application packages are EXCEL, WORD, ORACLE, DTP packages etc. The general-purpose application for example EXCEL can be used to build decision support applications such as resource allocation, scheduling, inventory control.

Similarly, these packages are used for statistical analysis for conducting financial analysis and market research. Thus application programs can be summarized as follows :

(a) Program written by users for a particular purpose, such as payroll, inventory control, scientific or engineering calculations, etc.

(b) "Packaged" programs written for a wide and restricted purpose, are supplied by vendors for end users. For example: Word processors for creating text documents; Excel or Tally for financial analysis work; database packages for creating maintaining and searching large structured data collections; statistical analysis package (SPSS).

(c) Installation libraries, containing programs of types (a) and (b) and databases that are useful to several users at an enterprise. "Data dictionary" is a complex software used to manage access to these objects.

(i) General Purpose Application Software

General purpose software refers to pre-written software that is designed for a variety of applications and uses. Such as visual Foxpro, Excel, etc.

The packaged software are pre-written software this alternative is least expensive. But drawback of buying a packaged software is that the Software may not be exactly the same as per the requirements of the organisation.

These software are not linked to specific business tasks but support the framework for a number of personal, business and scientific applications. Spread sheet, Computer Aided Design (CAD), Word Processing software fall in this category. For example an accountant can use General purpose software such as EXCEL to create balance sheet of the company. MS Word can be used by general manager to draft a letter. Graphic software CAD/CAM enables, its uses to create store, edit, view architectural designs. Multimedia Software. Flash/Macro media allows its users to create animated movies with audio and video effects.

(ii) Tailored Made Software

Custom Software/Tailor made Software refers to any type of program developed for a particular client to address a special need. Certain general purpose software such as dBASE, Foxpro and Lotus 1-2-3 are designed to provide tools and flexibility required for producing tailor made applications.

In order to develop tailor madq software following steps need to be performed, these are:

- Identify the user requirements and list them out.
- Analyse the functional modules and decide for in-house development or select a vendor to develop that software that meets the user requirement.
- Process the tailor made software development in various stages such as coding, testing, debugging, documentation, implementation and training for the planned program.
- Skilled staff needed to repair and maintain the software. If software is developed by outside party. It is easier to carry out changes in the software, if it is developed in-house.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PREWRITTEN SOFTWARE

Ready-made application packages are available in the market for many business functions, including accounting (e.g., payroll and inventory control), general ledger, financial planning and numerous other applications.

The major source of ready-made information systems software packages are computer manufacturers, computer retailers, computer service centres, time-sharing companies, system integrators and independent consultants. These and other types of firms in the computer industry apart from providing ready-made software, offer a variety of other services as well. For example, computer service centres provide off-premises computer processing of customer jobs and time-sharing companies provide real time mainframe computing via wide area networks to subscribers.

Advantages and Disadvantages of Ready-made Prewritten Application Software

Main advantages of using ready-made application software are summarised below:-

(a) Low Cost - As ready-made software packages are developed for general market, development costs are distributed across a large customer base.

(b) Faster Procurement - Software package developed in-house takes a very long time to develop, whereas ready-made software package can simply be purchased off-the-shelf.

(c) Rapid implementation - Ready-made software package can be implemented immediately after their purchase.

(d) High quality - As the ready-made software package typically has a large customer-base, developers continuously invest in refinement and testing. This results in production of high quality software.

(e) Low risk - It is relatively easy to evaluate ready-made software package through in-house testing, customer feedback and software reviews in the press.

(f) Fast updation - Ready-made software packages are more general and thus can cope up with growth.

(g) Longer life cycle - Ready-made software packages have longer life cycle because there is usually far more user experience behind the software.

(h) Better control - Ready-made software packages are better from the point of control and audit because people associated with the in-house development of software may have interest in keeping separate exit routes in the software .

But ready-made software packages have certain **disadvantages as well, which are:**-

(a) The ready-made software package may not fit into the specific requirements of the applications.

(b) Frequent changes in a ready-made software are either expensive or not possible.

(c) Non-availability of software package that meets the specific requirements of the business process.

Public Domain Software A program donated for public use by its owner or developer and freely available for copying and distribution such software can be downloaded and used immediately.

The limitation of using such software is that they may not be tested properly; may fail during execution; may contain viruses thus cannot be trusted.

Shareware Shareware software is a copyrighted software that is distributed on a try before you buy basis by paying shareware fee. Users who want to continue using the program after the trial period such as one month are encouraged to send a payment for program to the program developer. Shareware is geperally less expensive because it is often developed by a single programmer.

Freeware Freeware program are software program that are given away free of charge and often made available on the Internet. The programmer offer programs as freeware either for personal satisfaction or to assess its reception among interested users. Program developers often retain all rights to their freeware and users are not free to copy or distribute or sell it further.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PREVIOUS YEAR QUESTIONS

Short Answer type Questions

1. The ALU of a computer normally contains a number of high speed storage elements called _____.
2. _____ performs arithmetic & logical operation in computer.
3. Converssion of binary number 1010101000010111 to hexadecimal number is _____.
4. The memory which is ultraviolet light erasable and electronically programmable is _____.
5. Mentionn the advantages of EPROM over PROM
6. What is joystick ? How it is different from trackball ?
7. Define the terms hardware and software.
8. _____ method is used for obtaining a record from a cassette tape.
9. The most common input device used today is _____.
10. Differentiate between CRT Monitor and LED Monitor.
11. What are salient features of Laser Printers ?
12. What are the limitations of MS-Dos ?
13. Windows is _____ software.
14. The interface between computer hardware and user is _____.
15. A common boundary between two systems is called _____ .
16. The symbol used in assembly language are _____.
17. _____ converts assembly language to machine language.
18. Why is middleware so called ?
19. What is the necessity of an operating system for a computer system ?

Focused-short Answer type Questions

1. What is meant by generations of computers ? List out various generations in computers technologies along with their characterstics.
2. Compare microcomputers, minicomputers and main-frames in terms of size and cost.
3. Draw and explain the basic strecture of CPU.
4. What is Operating System?
5. Differentiate between EPROM and EEPROM ?
6. What is Flash Memory?

Long Answer type Questions

1. What are the various types of memory which are usally used with modern digital computers ? Give their important features in brief.
2. Explain the components of CPU and their functionality
3. What is Operating System? Explain various services provided by Operating System.
4. What do you mean by computer generations ? Explain the salient features of Fourth and Fifth Generations of Computers.
5. What are the different types of storage devices highlighting their advantages and disadvantages.
6. What is OS. Discuss main functions of an OS. Write about different types of OS.
7. Explain briefly the use of computers in the following areas- (a) Educations (b) Advertising, (c) Banking Sectors, (d) Hospital Management.
8. What do you mean by System Utility programs ? Differentiate between system software and application software.
9. Draw the block diagram of computer and explain all the compinents of computers.
10. What is the advantage of secondary storage ? Explain various secondary storage devices.
11. Compare five generation of computers on the basis of the software with one example for each.
12. Differentiate between System Software and Application Software with one example for each.
13. What are the Registers ? Listout the commonly found registers in the CPU and explain their functions.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Biju Patnaik Institute of IT & Management Studies

Lecture Notes

on

IT FOR MANAGEMENT

Module-II

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMPUTER PROCESSING TECHNIQUES

Process Concept

Informally, a process is a program in execution. A process is more than the program code, which is sometimes known as the text section. It also includes the current activity, as represented by the value of the program counter and the contents of the processor's registers. In addition, a process generally includes the process stack, which contains temporary data (such as method parameters, return addresses, and local variables), and a data section, which contains global variables.

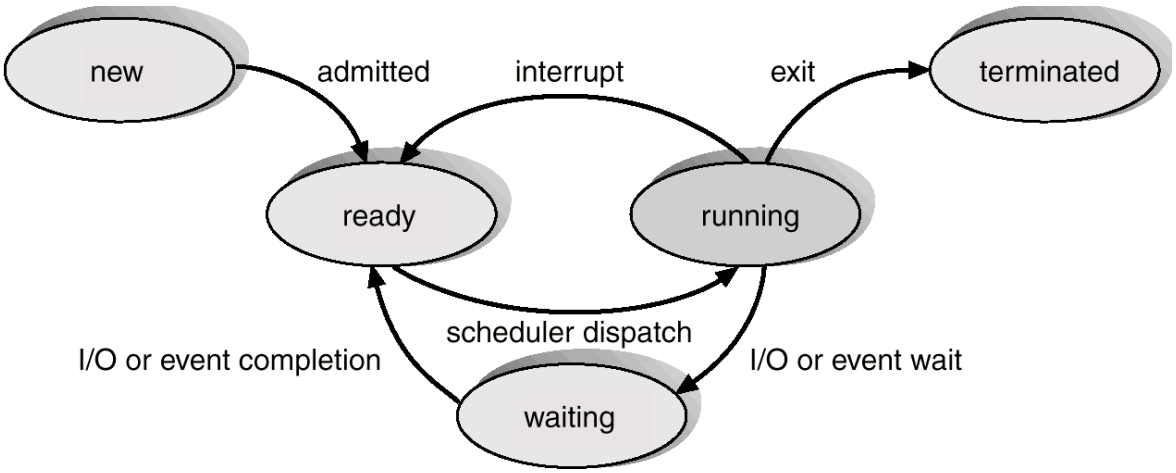
Process State: As a process executes, it changes state

- ⇨ **New State:** The process is being created.
- ⇨ **Running State:** A process is said to be running if it has the CPU, that is, process actually using the CPU at that particular instant.
- ⇨ **Blocked (or waiting) State:** A process is said to be blocked if it is waiting for some event to happen such that as an I/O completion before it can proceed. Note that a process is unable to run until some external eventhappens.
- ⇨ **Ready State:** A process is said to be ready if it needs a CPU to execute. A ready state process is runnable but temporarily stopped running to let another process run.
- ⇨ **Terminated state:** The process has finished execution.

What is the difference between process and program?

1. Both are same beast with different name or when this beast is sleeping (not executing) it is called program and when it is executing becomes process.
2. Program is a static object whereas a process is a dynamic object.
3. A program resides in secondary storage whereas a process resides in main memory.
4. The span time of a program is unlimited but the span time of a process is limited.
5. A process is an 'active' entity whereas a program is a 'passive' entity.
6. A program is an algorithm expressed in programming language whereas a process is expressed in assembly language or machine language.

Diagram of Process State

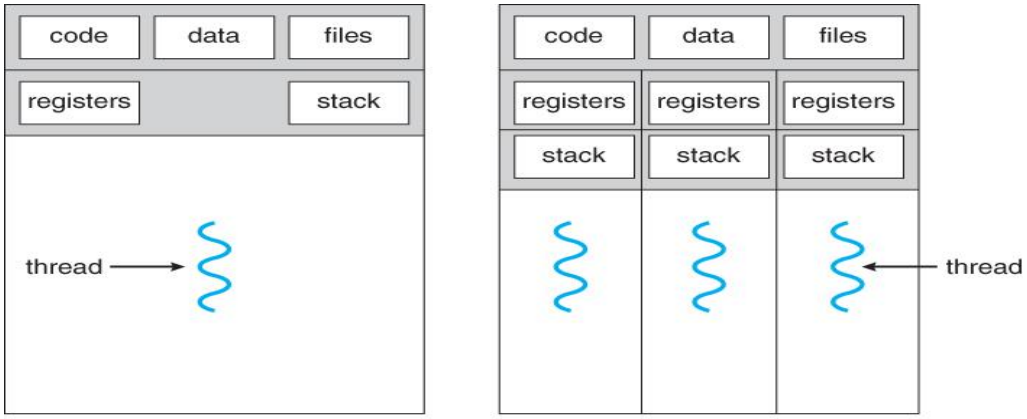


Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Thread

A thread, sometimes called a lightweight process (LWP), is a basic unit of CPU utilization; It comprises a thread ID, a program counter, a register set, and a stack. It shares with other threads belonging to the same process its code section, data section, and other operating-system resources, such as open files and signals. It shares with other threads belonging to the same process its code section, data section and other operating system resources, such as open files and signals. A traditional (or heavyweight) process has a single thread of control. If the process has multiple threads of control, it can do more than one task at a time.

Single-threaded and multithreaded



single-threaded process multithreaded process

For example, a web server accepts client requests for web pages, images, sound, and so forth. A busy web server may have several (perhaps hundreds) of clients concurrently accessing it. If the web server ran as a traditional single-threaded process, it would be able to service only one client at a time. One solution is to have the server run as a single process that accepts requests. When the server receives a request, it creates a separate process to service that request. In fact, this process-creation method was in common use before threads became popular. Process creation is very heavyweight. If the new process will perform the same tasks as the existing process, why incur all that overhead? It is generally more efficient for one process that contains multiple threads to serve the same purpose. This approach would multithread the web-server process. The server would create a separate thread that would listen for client requests; when a request was made, rather than creating another process, it would create another thread to service the request.

Process vs. Thread

Process	Thread
Process cannot share the same memory area(address space)	Threads can share memory and files.
It takes more time to create a process	It takes less time to create a thread.
It takes more time to complete the execution and terminate.	Less time to terminate.
Execution is very slow.	Execution is very fast.
It takes more time to switch between two processes.	It takes less time to switch between two threads.
System calls are required to communicate each other	System calls are not required.
It requires more resources to execute.	Requires fewer resources.
Implementing the communication between processes is bit more difficult.	Communication between two threads are very easy to implement because threads share the memory

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Evolution of OS:

BATCH PROCESSING

In a batch processing mode, job-to-job transition is not automatic. In this method, more than one jobs and these jobs are batched/ grouped together according to their similar requirements. This is done by computer operator. Whenever the computer becomes available, the batched jobs are sent for execution and gradually the output is sent back to the user. It allowed only one program at a time. This OS is responsible for scheduling the jobs according to priority and the resource required.

In this method lots of computer time was wasted, because computer remained idle while operator loaded and unloaded one job to another manually. An automatic method of transition of job was devised to reduce the computer idle time. To facilitate this, job control language was introduced. The control statements used by operating system are used to identify the next job to be processed and resources required to execute the job.

A classic business example of batch processing is payroll. In this, there are scheduled jobs which are done in a batch to produce cheques, registers and audits at a periodic intervals. Other examples of batch processing is the collection of bills or invoices and payment at the end of the month, where each bill requests for an action (payment) to be performed on the related data.

Batch processing is appropriate where large amount of data is needed to be processed, timeliness of the data is not critical, and user interaction is not required. Batch processing of transaction is useful where anum ber of static reports are produced.

MULTI-PROGRAMMING

Multi-programming refers to a form of processing in which more than one jobs are simultaneously executes by a single processor. It increases CPU utilization by organizing jobs so that the CPU always has one job to execute.

The concept of multiprogramming is described as follows:

- *All the jobs that enter the system are stored in the job pool(in disc). The operating system loads a set of jobs from job pool into main memory and begins to execute.*
- *During execution, the job may have to wait for some task, such as an I/O operation, to complete. In a multiprogramming system, the operating system simply switches to another job and executes.*
- *When that job needs to wait, the CPU is switched to another job, and so on.*
- *When the first job finishes waiting and it gets the CPU back.*
- *As long as at least one job needs to execute, the CPU is never idle.*

Another objective of multi-programming is to complete what may be a lengthy, complex start-up phase of execution, leaving the program suspended in a state of readiness with relatively little time required for real-time reactivation.

For this, the main memory is partitioned and each of the programs to be executed is resident in one of these partitions. The CPU can execute only one instruction at one time rather than executing two instructions simultaneously. Therefore, when a number of jobs are available in the main memory, CPU execution progress by executing one instruction of a user program, takes leave and then executes another instruction of other program, then again takes leave from there, returns to previous program and so on. This type of processing is called '**Concurrent Processing' or 'Interleaving'**. The concept of multi-programming was conceived in 1962 at the University of Manchester and first implemented on an Atlas computer by a team led by Tom Kilburn.

Multi-programming has better throughput because CPU's idle time is significantly minimized. It also requires additional software and hardware. For multi-programming, to accommodate more than one user program in the memory, it requires large main memory and a compatible operating system.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

MULTI-TASKING

Multi-tasking refers to form of processing supported by most of current operating systems in which computer works on multiple tasks of one user with one central processor.

Multi-tasking refers to an operating system's ability to support multiple processes simultaneously. Support for multiple processes is necessary in application where several computations must proceed in parallel. On a PC, a user may edit a file while another file is being printed and electronic mail is received. These three activities are best supported by three processes running simultaneously. Multi-tasking is also needed on servers and time-sharing systems where multiple users share a single computer system and all processes created by them should, at least in principle, execute simultaneously. Real-time systems that control multiple devices also need to support multiple processes. For instance, an Avionics computer on board an aircraft runs processes for monitoring the engines, updating the flight instruments, processing radar signals, and keeping the airplane on course. Batch operating systems depend on multi- tasking for overlapping computation with I/O operations. When a process performs I/O, the operating system runs another process to avoid CPU being idle for long period of time.

COMPARISON BETWEEN MULTI-PROGRAMMING AND MULTI-TASKING

Multi-programming	Multi-tasking
More than one program residing in the main memory.	Multi-tasking refers to the ability of the operating system to execute more than one job of a user e.g., Typing in MS WORD, printing on a printer, execution of a program in C may be executed simultaneously.
Only one program is executed at one time but independent programs are executed in same time span by interleaving their execution.	Multiple tasks are executed by the CPU switching between them.
Mainframes and mini computers work in multiuser-programming environment.	Personal computers generally work in multi-tasking environment.
The overall performance of the system can be enhanced by multi programming.	It leads to efficient CPU utilization as CPU time is shared by different processors.
CPU does not remain idle, as CPU time is used in processing not in wait state during 110 operations.	The tasks to be performed are not processing of the programs but to different tasks, that are chosen by CPU scheduler and switches the CPU to that task.

MULTI-THREADING

Another form of Multi-tasking is known as Multi-threading. Multi-threading allows different parts of a single program, called as threads, to run concurrently. It improves the performance of application program. Threads share a CPU time in the same way as the operating system executes each thread simultaneously, as processes do. In multi-threading the program is very complex, and processing provides for different threads at the same time without interfering with each other. A common example of multi-threading is new word processing software which allows editing of the document as well as spell checking at the same time.

TIME SHARING

Time sharing refers to the use of computer system by more than one user at the same time. Time sharing executes separate programs concurrently by interleaving portions of processing time allotted to each program (i.e. in round-robin fashion).

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Time sharing is an extension of multi-programming. Time sharing means many users are provided simultaneous interactive use of a computer system in such a way that each user has the feel as if it is his/her own computer. This can be achieved by a special CPU scheduling algorithm and use of multi-programming. In time sharing, users and programs make use of the given devices by a number of other devices, one at a time and in rapid succession. It is a processing technique for furnishing computing services to multiple users simultaneously and providing rapid response to each of the users. Time sharing computer systems employ multi-programming and/ or multi-processing techniques and are often capable of serving users at remote locations in a data communication network. The use of multi-programming and speed of the system allows the CPU to attend each user program for a given 'time slice'(may be 10 to 100 millisecond). It continuously switches from one user job to another and does a part of it in allocated 'time slice' for each job, until the work is completed. The speed changes so frequently such that the user has the illusion of having a computer for one's sole use. Therefore, it allows every user program, share of the CPU's time or 'time slice' turn by turn.

Time-sharing requires the additional hardware and software features such as:

(a) Different users on different Terminals i.e. multiple users require multiple terminals to work in interactive mode.

(b) Time sharing supervisor program in main memory.

(c) A large memory to support multi-programming.

(d) A special CPU scheduling algorithm to allocate time slice for each user program in circular manner.

(e) An alarm clock is required as CPU receives interrupt signals after every time slice.

Advantages of time sharing system:

(a) Throughput time of the system is increased as switching from one user program to another is in rapid succession. Thus, it reduces CPU idle time significantly.

(b) Small users can access to sophisticated computing resources.

(c) Time sharing provides the advantage of quick response time. This leads to optimum utilization of resources.

(d) As system software and applications share the same memory or are stored in system library, it saves space and is easier to maintain.

However, a time sharing system posses security threats, problems of reliability, communication problems and overloading issues because multiple users are given access right and anyone could be with bad intention of doing computer crime, may be for the sake of fun.

Difference Between Time Sharing and Multitasking

The main difference between time sharing and multitasking is that **time sharing allows multiple users to share a computer resource simultaneously using multiprogramming and multitasking while multitasking allows a system to execute multiple tasks or processes simultaneously.**

MULTI-PROCESSING

Multi-processing refers to the computer having more than one CPU. Hence there is concurrent execution of a number of programs as many CPUs are there in the computer system.

Multi-processing is a mode of operation in which two or more processing units are connected together, each carry out one or more job simultaneously.

The objective is to achieve increased speed or computing power, the same as in parallel processing and in the use of special units called co-processors.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

INTERACTIVE PROCESSING

Interactive processing mode permits the user to have continuous dialogue with the operating program in the computer. The computer keeps responding not only to the user's requests, but also replies to mistakes. The user can process the program step-by-step, testing portions of his procedure or trying out various solutions. The interactive processing approach is usually used in programs written in BASIC language, where when each statement of the program is entered, it is interpreted and errors are detected. Then finally, when these errors are edited, then program is executed. The benefit of interactive system is that errors in the program can be easily encountered, corrected and processing can continue immediately. Since transactions are dealt on individual basis with complete relevant processing operations, therefore, it is also called 'transaction driven processing system'. It does not begin next transaction until the previous transaction is completed.

ONLINE PROCESSING

Online processing refers to processing of individual transactions as they occur from their point of origin rather accumulating them into batches. This is possible by using direct access devices such as magnetic disks and terminals which are controlled by central processor.

In contrast to batch/ transaction processing, online processing is termed as event-driven. In online processing, transactions are executed as soon as they are entered into computer system in an interactive mode (where not operator but user interacts with computer system in the same time frame) and the user instantly gets the result of processed transaction on the monitor. For this the user must have online access to the computer, the computer may be a personal computer or a network computer. Computers of distant companies are connected by means of telegraph, telephone lines or satellite. This type of processing provides multi-access to information files by terminal users and also enables them to update files with transaction data. Online applications are the following :-

Stock Exchanges - Terminals located in major stock exchanges throughout the country and the offices of brokers enable the speedy processing of share dealings.

Tour Operators - Reservation offices accept telephone inquiries from travel agents regarding the availability of hotel accommodation and travelling services with respect to client inquiries. By means of terminals, the availability of the hotel accommodation can be checked and booked immediately.

Stock Control - Terminals located in warehouses provide the means for automatic recording of stocks, updating of stock records, reservations, follow up of outstanding orders and the printing of packing list etc.

Internet Data base - Information from the website about products and offering/ ordering information such as Amazon.com,ebay.com.

Advantages of using online systems are quite significant in the business environment, some of them are listed below :-

(i) Great computer usage - Online system generally serve the customer enquiries and order processing, such business system requires skilled clerical staff who can access to information stored in the computer by the use of terminals. Thus, clerical staff is integrated with the computer.

(ii) Great operatabilities - As usage of computer in business leads to sophistication in work environment, greater degree of job interest, operating efficiency and job satisfaction.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(iii) Reduces paper work - In batch processing environment, the volume of paper generated even for small routine clerical task is very high. Whereas online system reduces the hard copy of output because information is available online on the terminal screen as per demand.
(iv) Higher accuracy - The messages before being displayed on screen are checked for accuracy by validation programs, since input errors are reduced, thus reliability of information increases.
(v) Easy file updation - In an online system, updating a file is easy than in batch processing mode because no special runs are required. Simply using keyboard and terminal, a master file can be updated.
(vi) Availability of management information - It is easy for the managers to have access on the required information from the direct access online devices, resulting in a greater degree of control to the operations
for which they are responsible.
(vii) Improved business services - Online system can better serve the needs of todays business environment. For example, banks using computer have improved customer services than in early system.
(viii) Reduced time and cost in data preparation - As on-line system directly keys the data to the system, punching and verifying operations are eliminated. This saves time and money associated with such operations.

REAL TIME SYSTEM
Real-time system refers to a technique in which computer and software system that responds to events (transaction data) before the events become obsolete. Real time systems are online enquiry systems.
Real time systems are those systems in which correctness of the system depends not only on the logical computations, but also on the time at which results are produced. Real time systems are usually operated in multiprogramming and multi-processing, as this increases availability and reliability of the system. A real time system's CPU must possess the capability in handling interrups of program. Examples of real time system are Intensive Care Monitoring, Air Reservation System. Typically, a real time system consists of a controlling system and a controlled system, where controlling system interacts with its environment based on the information available about the environment from various sensors.
In real time system, the tasks might be scheduled statistically as dynamically, but as real time systems are large and complex and function in environment, that is uncertain and are also physically distributed. Therefore, real time system need to be fast, predictable, reliable and adoptive. More important, they need to be maintainable and extensible due to their evolving nature and projected long life times. It may be noted that all online systems are real time applications. Real time applications-

Sales Order Processing - Real-time systems offer the possibility of a greatly prompt service to the customer. The customer order can be processed instantly, within virtually no time. The finished goods inventory file may have data terminals throughout the sales territory. The CRT's can be made to display the invoice on a terminal for the salesman who is in the customer's factory and can input the particulars there and then. The customer would get to know the availability instantaneously. All appropriate journal entries would be made immediately in the general ledger. The invoice can also be posted to the accounts receivable file instantaneously. The selection of warehouse that is closet to the customer can also be made and a copy of the invoice could be displayed at the warehouse terminals. Besides, the following functions can be performed rapidly :-

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- The inquiries about the customer's accounts status can be answered by the salesman.
- Credit appraisal can be carried out on-line, real time.
- Sales analysis master file can be updated on-line and real time. This file provides a ready information to the sales manager regarding the sales trends etc.
- Inventory replenishment orders can be determined as a part of the sales order processing.

Transportation and Travel Reservation - These activities pertain notably to hotel chains and airline or railway reservation. Data terminals are generally located at each reservation point so that the customer reservation can be confirmed quickly. The various accounting calculations can also be performed and the relevant files can be updated.

REAL-TIME OPERATING SYSTEMS (RTOS)

A real-time operating system (RTOS) is a multitasking operating system intended for applications with fixed deadlines (real-time computing). Such applications include some small embedded systems, automobile engine controllers, industrial robots, spacecraft, industrial control, and some large-scale computing systems.

The real time operating system can be classified into two categories:

1. Hard real time system
2. Soft real time system

1. Hard real time system

A hard real-time system guarantees that critical tasks be completed on time. This goal requires that all delays in the system be bounded, from the retrieval of stored data to the time that it takes the operating system to finish any request made of it. Such time constraints dictate the facilities that are available in hardreal-time systems.

2. Soft real time system

A soft real-time system is a less restrictive type of real-time system. Here, a critical real-time task gets priority over other tasks and retains that priority until it completes. Soft real time system can be mixed withother types of systems. Due to less restriction, they are risky to use for industrial control and robotics.

DISTRIBUTED OPERATING SYSTEMS

- In distributed system, the different machines are connected in a network and each machine has its own processor and own local memory.
- In this system, the operating systems on all the machines work together to manage the collective network resource.
- It can be classified into two categories:

1. Client-Server systems
2. Peer-to-Peer systems

Advantages of distributed systems.

- Resources Sharing
- Computation speed up – load sharing
- Reliability
- Communications
- Requires networking infrastructure.
- Local area networks (LAN) or Wide area networks (WAN)

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

SUMMARY OF VARIOUS PROCESSING TECHNIQUES

Batch Processing	Multi-programming
One user program in main memory at one time- a mono-programming system.	Two or more user programs in main memory at one time.
May use any secondary storage device such as magnetic tape, disc or drum.	Only magnetic discs or drums suitable because of random access.
Fairly simple operating system required.	Complex operating system required.

Real-time	Time-sharing
Usually devoted to one application	Many different applications.
User makes enquiry only and cannot be allowed to write or modify programs.	User can write and modify programs.
User must get a response within a given time period, depending upon the application.	Users should get a response within 3-5 seconds but not disastrous, if response is not in desired limit.
Transaction files must be updated before another enquiry using those transaction files can be serviced.	Transaction files are updated but the time in which this is done is not so critical.

COMPARISONS OF VARIOUS PROCESSING TECHNIQUES

Processing Technique	Number of users	Number of processors	Processing order
Multitasking	One	One	Concurrent
Multiprogramming	Multiple	One	Concurrent
Time sharing	Multiple	One	Round Robin
Multiprocessing	One or more	More than one	Simultaneously

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

INTERPROCESS COMMUNICATION

Inter-Process Communication (IPC) is a set of techniques for the exchange of data among two or more threads in one or more processes. Processes may be running on one or more computers connected by a network. IPC techniques are divided into methods for message passing, synchronization, shared memory, and remote procedure calls (RPC). The method of IPC used may vary based on the bandwidth and latency of communication between the threads, and the type of data being communicated. Operating system provides facilities for IPC.

Inter Process communication may be act as **Unicast or Multi-Cast.**

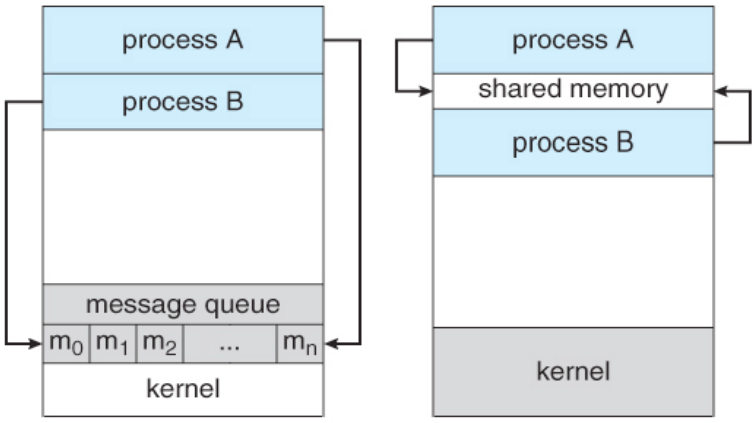
Unicast IPC:- Communication is from one process to single other process is called Unicast IPC. Example is Socket Communiation.

Multi-Cast:- Communication is form one process to group of process is call Multi-Cast IPC. Example is Publish or Subscribe.

IPC Models (IPC Mechanism) are 1) Message Passing, 2. Shared Memory.

1) **Message Passing** - Message passing is a form of communication used in interprocess communication. Communication is made by the sending of messages to recipients. Each process should be able to name the other processes. The producer typically uses *send()* system call to send messages, and the consumer uses *receive()* system call to receive messages. These system calls can be either synchronous or asynchronous, and could either be between processes running on a single machine, or could be done over the network to coordinate machines in a distributed system. This allows the producer to transfer data to the consumer as it is created.
2) **Shared Memory** - Shared Memory is an OS provided abstraction which allows a memory region to be simultaneously accessed by multiple programs with an intent to provide communication among them. One process will create an area in RAM which other processes can access (this is typically done using system calls mmap, shmget etc). Normally the OS prevents processes from accessing the memory of another process, but the Shared Memory features in the OS can allow data to be shared. Since both processes can access the shared memory area like regular working memory, this is a very fast way of communication (as opposed to other mechanisms of IPC). On the other hand, it is less powerful, as for example the communicating processes must be running on the same machine (whereas other IPC methods can use a computer network), and care must be taken to avoid issues if processes sharing memory are running simultaneously and may try to edit the shared buffer at the same time.

The two communications models are contrasted in the figure below:



(a) Message Passing (b) Shared Memory

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

MEMORY MANAGEMENT

Memory management is the functionality of an operating system which handles or manages primary memory and moves processes back and forth between main memory and disk during execution. Memory management keeps track of each and every memory location, regardless of either it is allocated to some process or it is free. It checks how much memory is to be allocated to processes. It decides which process will get memory at what time. It tracks whenever some memory gets freed or unallocated and correspondingly it updates the status.

Process Address Space

The process address space is the set of logical addresses that a process references in its code.

The operating system takes care of mapping the logical addresses to physical addresses at the time of memory allocation to the program. There are three types of addresses used in a program before and after memory is allocated –

S.N.	Memory Addresses & Description
1	**Symbolic addresses** The addresses used in a source code. The variable names, constants, and instruction labels are the basic elements of the symbolic address space.
2	**Relative addresses** At the time of compilation, a compiler converts symbolic addresses into relative addresses.
3	**Physical addresses** The loader generates these addresses at the time when a program is loaded into main memory.

Virtual and physical addresses are the same in compile-time and load-time address-binding schemes. Virtual and physical addresses differ in execution-time address-binding scheme.

The set of all logical addresses generated by a program is referred to as a logical address space. The set of all physical addresses corresponding to these logical addresses is referred to as a physical address space.

The runtime mapping from virtual to physical address is done by the memory management unit (MMU) which is a hardware device. MMU uses following mechanism to convert virtual address to physical address.

Static vs Dynamic Loading

The choice between Static or Dynamic Loading is to be made at the time of computer program being developed. If you have to load your program statically, then at the time of compilation, the complete programs will be compiled and linked without leaving any external program or module dependency. The linker combines the object program with other necessary object modules into an absolute program, which also includes logical addresses.

If you are writing a Dynamically loaded program, then your compiler will compile the program and for all the modules which you want to include dynamically, only references will be provided and rest of the work will be done at the time of execution.

At the time of loading, with static loading, the absolute program (and data) is loaded into memory in order for execution to start.

If you are using dynamic loading, dynamic routines of the library are stored on a disk in relocatable form and are loaded into memory only when they are needed by the program.

Static vs Dynamic Linking

As explained above, when static linking is used, the linker combines all other modules needed by a program into a single executable program to avoid any runtime dependency.

When dynamic linking is used, it is not required to link the actual module or library with the program, rather a reference to the dynamic module is provided at the time of compilation and linking. Dynamic Link Libraries (DLL) in Windows and Shared Objects in Unix are good examples of dynamic libraries.

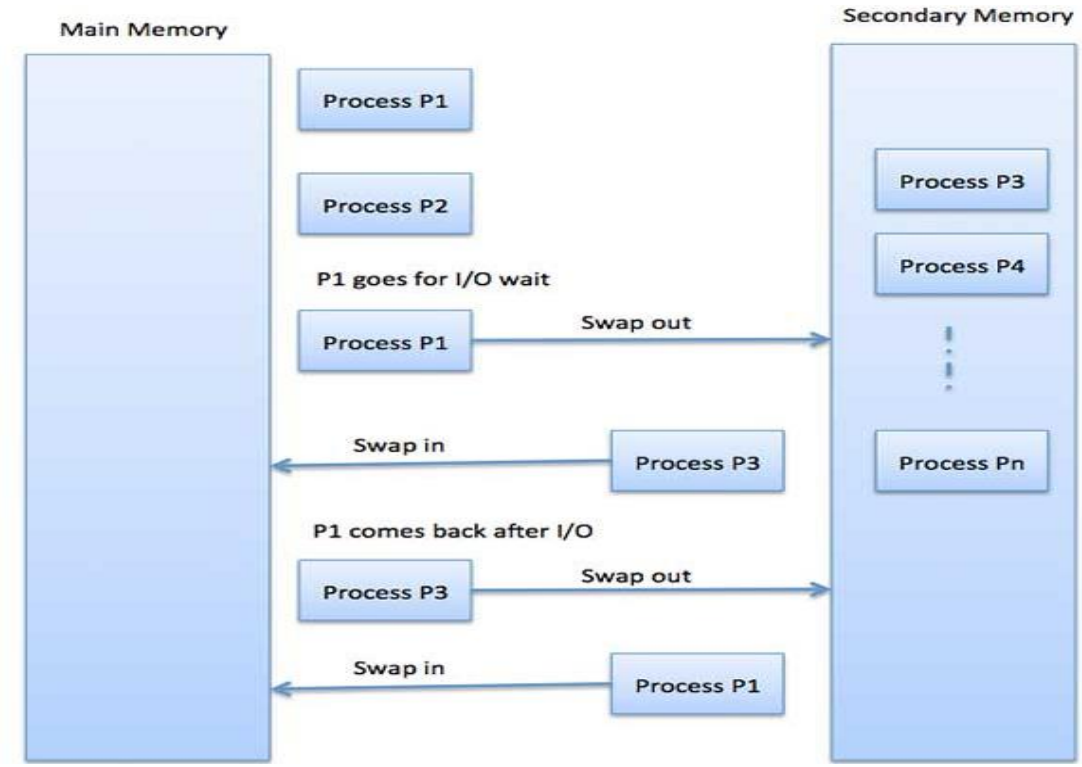
Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

SWAPPING

Swapping is a mechanism in which a process can be swapped temporarily out of main memory (or move) to secondary storage (disk) and make that memory available to other processes. At some later time, the system swaps back the process from the secondary storage to main memory.

Though performance is usually affected by swapping process but it helps in running multiple and big processes in parallel and that's the reason Swapping is also known as a **technique for memory compaction.**

The total time taken by swapping process includes the time it takes to move the entire process to a secondary disk and then to copy the process back to memory, as well as the time the process takes to regain main memory.



Memory Allocation

Main memory usually has two partitions −

- Low Memory − Operating system resides in this memory.
- High Memory − User processes are held in high memory.

Operating system uses the following memory allocation mechanism.

S.N.	Memory Allocation & Description
1	**Single-partition allocation** In this type of allocation, relocation-register scheme is used to protect user processes from each other, and from changing operating-system code and data. Relocation register contains value of smallest physical address whereas limit register contains range of logical addresses. Each logical address must be less than the limit register.
2	**Multiple-partition allocation** In this type of allocation, main memory is divided into a number of fixed-sized partitions where each partition should contain only one process. When a partition is free, a process is selected from the input queue and is loaded into the free partition. When the process terminates, the partition becomes available for another process.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

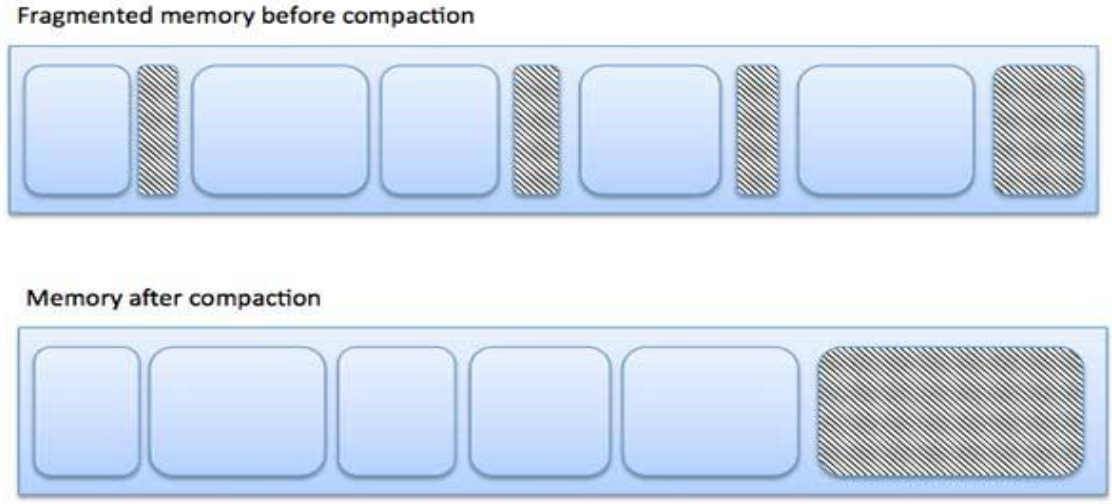
Fragmentation

As processes are loaded and removed from memory, the free memory space is broken into little pieces. It happens after sometimes that processes cannot be allocated to memory blocks considering their small size and memory blocks remains unused. This problem is known as Fragmentation.

Fragmentation is of two types –

1. **External fragmentation:** Total memory space is enough to satisfy a request or to reside a process in it, but it is not contiguous, so it cannot be used.
2. **Internal fragmentation** : Memory block assigned to process is bigger. Some portion of memory is left unused, as it cannot be used by another process.

The following diagram shows how fragmentation can cause waste of memory and a compaction technique can be used to create more free memory out of fragmented memory −



External fragmentation can be reduced by compaction or shuffle memory contents to place all free memory together in one large block. To make compaction feasible, relocation should be dynamic.

The internal fragmentation can be reduced by effectively assigning the smallest partition but large enough for the process.

SEGMENTATION

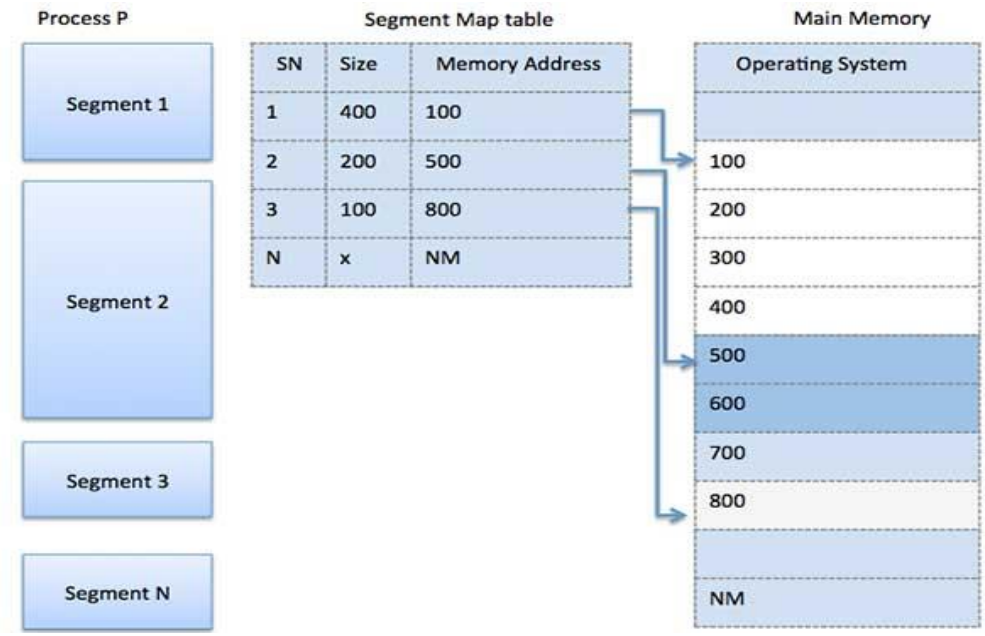
Segmentation is a memory management technique in which each job is divided into several segments of different sizes, one for each module that contains pieces that perform related functions. Each segment is actually a different logical address space of the program.

When a process is to be executed, its corresponding segmentation are loaded into non-contiguous memory though every segment is loaded into a contiguous block of available memory.

Segmentation memory management works very similar to paging but here segments are of variable-length where as in paging pages are of fixed size.

A program segment contains the program's main function, utility functions, data structures, and so on. The operating system maintains a segment map table for every process and a list of free memory blocks along with segment numbers, their size and corresponding memory locations in main memory. For each segment, the table stores the starting address of the segment and the length of the segment. A reference to a memory location includes a value that identifies a segment and an offset.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar



VIRTUAL MEMORY

Operating system manages allocation of main memory for various processes to specific job. These days, the application process requires huge amounts of main memory, which may not be present in the systems. This extra memory is actually called **virtual memory** (Non-contigious Allocation) and it is a section of a hard disk that's set up to emulate the computer's RAM.

OS giving illusion to the user such that : User can write a very big program and the user thinks that the entire program is present in RAM & all the space allocated to the user is present in RAM. But in reality only a small protion of the user program is in RAM (which may or may not be contigious) while the remaining program in secondary memory.

The main visible advantage of this scheme is that programs can be larger than physical memory. Virtual memory serves two purposes. First, it allows us to extend the use of physical memory by using disk. Second, it allows us to have memory protection, because each virtual address is translated to a physical address.

The operating system actually pulls data from secondary storage. The primary storage is extended into a secondary storage device allowing users to execute programs as if, primary storage were larger than it actually is. This improves the computer speed and also run programs efficiently even with large number of pages or instructions.

Difference between virtual memory and main memory.

Main memory	Virtual Memory
It is a part of the computer hardware	It is a processing technique.
It stores program and data that are in active use.	It allows execution of a process when logical address space requirement of the process is greater than the physical available main memory.
Access is relatively slower than CPU register as it is physically separated from the CPU and large in capacity.	It is disk storage, which is treated as logical extension of the main memory.
In main memory, storage locations are addressed directly by CPU's load and store instruction.	Virtual memory helps in efficient main memory utilization by dividing the program on disk into fixed length pages or into logical variables length segment.
Stores and runs one program at a time.	Virtual memory allows more programs to be loaded in the main memory, thus improves CPU utilization and throughput.
Cost per bit is very less.	Cost per bit is high.

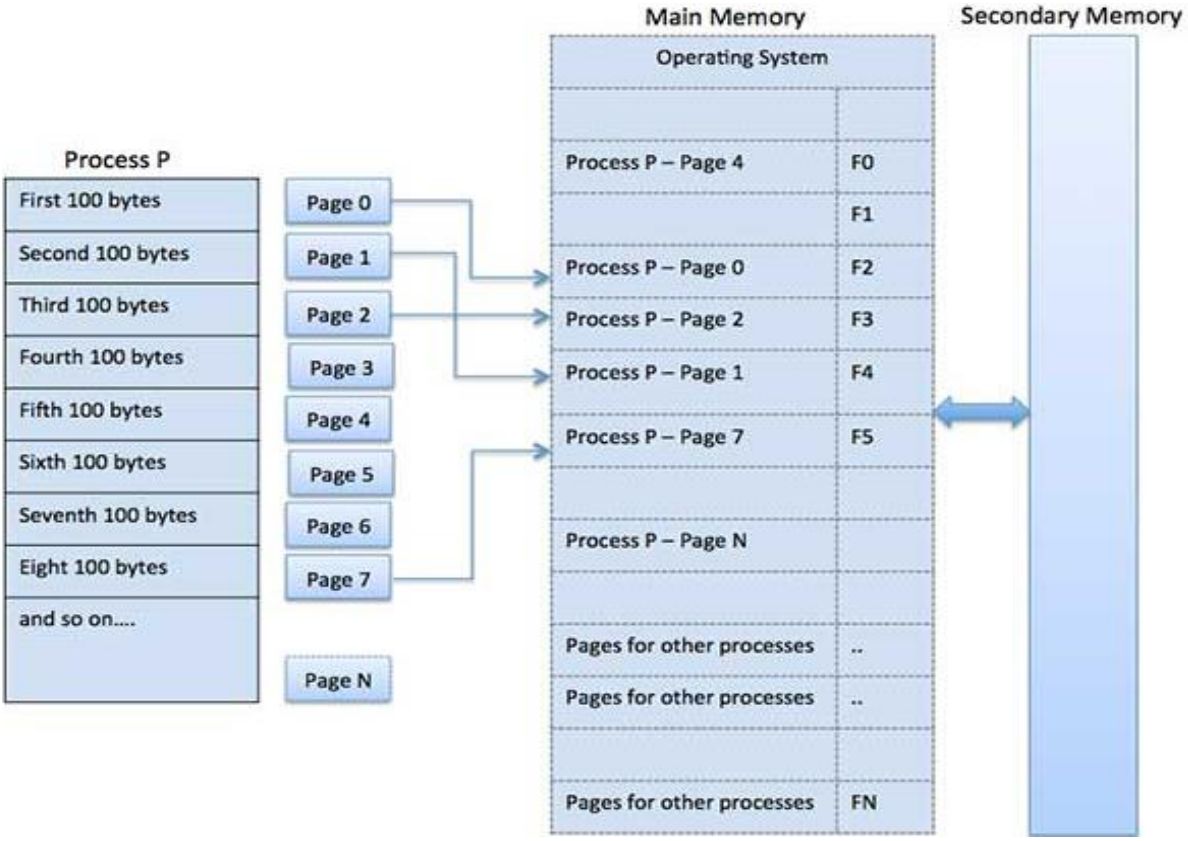
Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PAGING

A computer can address more memory than the amount physically installed on the system. This extra memory is actually called virtual memory and it is a section of a harddisk that's set up to emulate the computer's RAM. Paging technique plays an important role in implementing virtual memory.

Paging is a memory management technique in which process address space is broken into blocks of the same size called paging (size is power of 2, between 512 bytes and 8192 bytes). The size of the process is measured in the number of pages.

Similarly, main memory is divided into small fixed-sized blocks of (physical) memory called frames and the size of a frame is kept the same as that of a page to have optimum utilization of the main memory and to avoid external fragmentation.



Address Translation

Page address is called logical address and represented by page number and the offset.

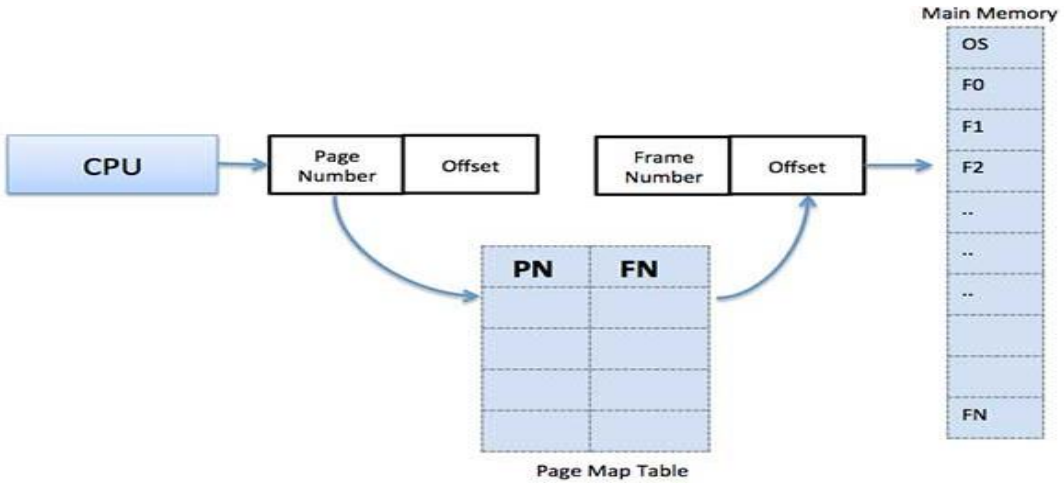
Logical Address = Page number + page offset

Frame address is called physical address and represented by a frame number and the offset.

Physical Address = Frame number + page offset

A data structure called page map table is used to keep track of the relation between a page of a process to a frame in physical memory.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar



When the system allocates a frame to any page, it translates this logical address into a physical address and create entry into the page table to be used throughout execution of the program.

When a process is to be executed, its corresponding pages are loaded into any available memory frames. Suppose you have a program of 8Kb but your memory can accommodate only 5Kb at a given point in time, then the paging concept will come into picture. When a computer runs out of RAM, the operating system (OS) will move idle or unwanted pages of memory to secondary memory to free up RAM for other processes and brings them back when needed by the program.

This process continues during the whole execution of the program where the OS keeps removing idle pages from the main memory and write them onto the secondary memory and bring them back when required by the program.

Advantages and Disadvantages of Paging

Here is a list of advantages and disadvantages of paging −

- Paging reduces external fragmentation, but still suffer from internal fragmentation.
- Paging is simple to implement and assumed as an efficient memory management technique.
- Due to equal size of the pages and frames, swapping becomes very easy.
- Page table requires extra memory space, so may not be good for a system having small RAM.

DEMAND PAGING

A demand paging system is quite similar to a paging system with swapping where processes reside in secondary memory and pages are loaded only on demand, not in advance. When a context switch occurs, the operating system does not copy any of the old program's pages out to the disk or any of the new program's pages into the main memory Instead, it just begins executing the new program after loading the first page and fetches that program's pages as they are referenced.

While executing a program, if the program references a page which is not available in the main memory because it was swapped out a little ago, the processor treats this invalid memory reference as a page fault and transfers control from the program to the operating system to demand the page back into the memory.

Advantages

Following are the advantages of Demand Paging −

- Large virtual memory.
- More efficient use of memory.
- There is no limit on degree of multiprogramming.

Disadvantages

- Number of tables and the amount of processor overhead for handling page interrupts are greater than in the case of the simple paged management techniques.

Page Replacement Algorithm

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Page replacement algorithms are the techniques using which an Operating System decides which memory pages to swap out, write to disk when a page of memory needs to be allocated. Paging happens whenever a page fault occurs and a free page cannot be used for allocation purpose accounting to reason that pages are not available or the number of free pages is lower than required pages.

When the page that was selected for replacement and was paged out, is referenced again, it has to read in from disk, and this requires for I/O completion. This process determines the quality of the page replacement algorithm: the lesser the time waiting for page-ins, the better is the algorithm.

A page replacement algorithm looks at the limited information about accessing the pages provided by hardware, and tries to select which pages should be replaced to minimize the total number of page misses, while balancing it with the costs of primary storage and processor time of the algorithm itself. There are many different page replacement algorithms. We evaluate an algorithm by running it on a particular string of memory reference and computing the number of page faults,

Reference String

The string of memory references is called reference string. Reference strings are generated artificially or by tracing a given system and recording the address of each memory reference. The latter choice produces a large number of data, where we note two things.

- For a given page size, we need to consider only the page number, not the entire address.
- If we have a reference to a page p, then any immediately following references to page p will never cause a page fault. Page p will be in memory after the first reference; the immediately following references will not fault.

First In First Out (FIFO) algorithm

- Oldest page in main memory is the one which will be selected for replacement.
- Easy to implement, keep a list, replace pages from the tail and add new pages at the head.

Optimal Page algorithm

- An optimal page-replacement algorithm has the lowest page-fault rate of all algorithms. An optimal page-replacement algorithm exists, and has been called OPT or MIN.
- Replace the page that will not be used for the longest period of time. Use the time when a page is to be used.

Least Recently Used (LRU) algorithm

- Page which has not been used for the longest time in main memory is the one which will be selected for replacement.
- Easy to implement, keep a list, replace pages by looking back into time.

Page Buffering algorithm

- To get a process start quickly, keep a pool of free frames.
- On page fault, select a page to be replaced.
- Write the new page in the frame of free pool, mark the page table and restart the process.
- Now write the dirty page out of disk and place the frame holding replaced page in free pool.

Least frequently Used(LFU) algorithm

- The page with the smallest count is the one which will be selected for replacement.
- This algorithm suffers from the situation in which a page is used heavily during the initial phase of a process, but then is never used again.

Most frequently Used(MFU) algorithm

- This algorithm is based on the argument that the page with the smallest count was probably just brought in and has yet to be used.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PROGRAMMING LANGUAGE

A computer or a system is operated by given a set of rules and instruction to perform some task. These sat of rules and instructions are able to control the working of computer or any automated and/or manipulated machine. To control the computer system using these sequential set of grammatical rules are known as programming language. In other words, 'A programming language is a vocabulary and a set of grammatical rules for instructing a computer system to perform any specific task'. Hence we can say that a programming language plays a very important role to control and operating a computer system.

Classification of programming languages:

Since the starting of programming language, it is classified into three categories. Sometime it's also known as level and/or generation of programming language. All the computer programming languages are broadly classified into the following;

- Machine level Language
- Assembly level Language
- High-level Language

MACHINE LEVEL LANGUAGES (First Generation of programming language):

It's the lowest level and named as first generation of programming language. Machine level language consist only two condition i.e. either true (1) or false (0); this type of language known as binary language. A computer system could understand only binary language i.e. all the instruction feed into the computer system must be in the form of 0 or 1. Machine level languages are very tough to understand by the humans.

Advantages of machine level language:

There have many advantage of machine level language; some of them are listed below;

- Machine level languages are directly interacting with computer system.
- There is no requirement of software of conversion like compiler or interpreters.
- It takes very less time to execute a program, because there is no conversion take place.

Disadvantages of machine language:

Some of the disadvantages of machine level language are listed below;

- Its machine dependent language i.e. individual program required for each machine.
- To develop a program in machine language, it's too hard to understand and program.
- Its time consuming to develop new programs.
- Debugging process is very hard because of finding errors process is typical.
- Machine language is not portable language.

ASSEMBLY LEVEL LANGUAGES (Second Generation programming language):

It's a middle level and named as second generation programming language. It contains the same instruction as machine level language, but the instructions and the variables have specific name or called commands instead of being just binary numbers. It also uses symbols to describe field of instructions. Every aspect of machine variable in program, assembly language uses one statement per machine instruction. It managed explicitly all instruction like register allocation, call, stack, timer, jump, loop etc.

To understand the assembly language computer must require an assembler which takes a part in between assembly language and computer system to convert the assembly instructions into the machine language. This assembler is software or a set of program which translates assemble language programming instruction into the machine language.

For example;

```
LOAD        BASEPAY
ADD         OVERPAY
STORE       GROSSPAY
```

Advantages of Assembly language:

There have many advantage of assembly level language; some of them are listed below;

- It is easily understood by human because it is uses statements instead of binary digits.
- To develop a program it takes less time.
- Debugging and troubleshoot is easy due to easily find error.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- It’s a portable language.

Disadvantages of Assembly language:

Some of the disadvantages of assembly level language are listed below;

- It’s a machine dependent language due to that program design for one machine no use of other machine.
- Sometime it’s hard to understand the statement or command use.

HIGH-LEVEL LANGUAGE (Third Generation):

High level language is the upper level language and also known as third generation programming language. It does consider as high level because, which language comes under this category are closer to human languages. Hence this is highly understood programming language by human. There have many examples of high level languages such as, FORTRAN, Pascal, C, C++, JAVA, ADA, COBOL, LISP, Prolog etc.

The first high level programming language was written in 1950s. Those programs written in high level language must require software or a set of program to translate that program into machine understandable. This software called compiler and/or interpreter. The main job of compiler and translator is to take the source code of the program and convert that code into the machine understood code.

Advantages of high level language:

There have many advantage of high level language; some of them are listed below;

- In this instructions and commands much easier to remember by programmer.
- Its logic and structure are much easier to understand.
- Debugging is easier compare to other languages.
- Less time consuming to writing new programs.
- HLL are described as being portable language.

Disadvantages of high level language:

Some of the disadvantages of high level language are listed below;

- HLL programming language take more space compare to other MLL (machine level language) and/or ALL (Assembly level language).
- This programming language execute slowly.

LANGUAGE PORTABILITY

Portability, in relation to software, is a measure of how easily an application can be transferred from one computer environment to another. A computer software application is considered portable to a new environment if the effort required to adapt it to the new environment is within reasonable limits. The meaning of the abstract term 'reasonable' depends upon the nature of the application and is often difficult to express in quantifiable units.

The phrase "to port" means to modify software and make it adaptable to work on a different computer system. For example, to port an application to Linux means to modify the program so that it can be run in a Linux environment.

Portability refers to the ability of an application to move across environments, not just across platforms. To clarify, a computer platform generally refers to the operating system and computer hardware only. A computer environment is much broader and may include the hardware, the operating system and the interfaces with other software, users and programmers

No software is perfectly portable because all softwares have limitations. Some programming languages are fairly portable, for example the C language. C compilers are readily available for the majority of operating systems, which in turn makes C programs very portable. This portability of C language programs has resulted in some programmers re-writing their programs and recompiling them in C to make them much more portable.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PREVIOUS YEAR QUESTIONS

Short Answer type Questions

1. Thread is a _____ process.
2. What are the classification of programming language?
3. Differentiate between Compiler and Interpreter.

Focused-short Answer type Questions

1. Differentiate between Process and Thread.
2. What is interprocess communication ?
3. Difference between hard and soft real time systems.
4. Explain the concept of virtual memory.
5. Explain the functions of virtual memory.
6. What is demand paging? Explain various page replacement techniques used in virtual memory.
7. Explain the concept of programming language ? Give its classification.
8. Explain the Evoluation of Programming Languages.
9. Describe language portability.
10. Write short notes on graphics and multimedia.

Long Answer type Questions

1. Differentiate among the following terms Multiprogramming, Multitasking, Multithreading and Multiprocessing.
2. What is assembly language ? Give advantages over machine language.
3. Differentiate between Process and Thread.
4. Explain the Evoluation and classification of Programming Languages.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Biju Patnaik Institute of IT & Management Studies

Lecture Notes

on

IT FOR MANAGEMENT

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Computer Graphics & Multimedia

Concept of Multimedia

- ⇨ Multi
  Many, Multiple,
- ⇨ Media
  - ✓ Tools that is used to represent or do a certain things, delivery medium, a form of mass communication – newspaper, magazine / tv.
  - ✓ Distribution tool & information presentation – text, graphic, voice, images, music and etc.

Multimedia is an interactive media and provides multiple ways to represent information to the user in a powerful manner. It provides an interaction between users and digital information. It is a medium of communication.Some of the sectors where multimedias is used extensively are education, training, reference material, business presentations, advertising and documentaries.

Multimedia is a representation of information in an attractive and interactive manner with the use of a combination of text, audio, video, graphics and animation. In other words we can say that Multimedia is a computerized method of presenting information combining textual data, audio, visuals (video), graphics and animations.

Components / Elements of Multimedia

Following are the common components of multimedia:

- **Text**- All multimedia productions contain some amount of text. The text can have various types of fonts and sizes to suit the profession presentation of the multimedia software.
- **Graphics**- Graphics make the multimedia application attractive. In many cases people do not like reading large amount of textual matter on the screen. Therefore, graphics are used more often than text to explain a concept, present background information etc. There are two types of Graphics:
  - **Bitmap images**- Bitmap images are real images that can be captured from devices such as digital cameras or scanners. Generally bitmap images are not editable. Bitmap images require a large amount of memory.
  - **Vector Graphics**- Vector graphics are drawn on the computer and only require a small amount of memory. These graphics are editable.
- **Audio**- A multimedia application may require the use of speech, music and sound effects. These are called audio or sound element of multimedia.Speech is also a perfect way for teaching. Audio are of analog and digital types. Analog audio or sound refers to the original sound signal. Computer stores the sound in digital form. Therefore, the sound used in multimedia application is digital audio.
- **Video**- The term video refers to the moving picture, accompanied by sound such as a picture in television. Video element of multimedia application gives a lot of information in small duration of time. Digital video is useful in multimedia application for showing real life objects. Video have highest performance demand on the computer memory and on the bandwidth if placed on the internet. Digital video files can be stored like any other files in the computer and the quality of the video can still be maintained. The digital video files can be transferred within a computer network. The digital video clips can be edited easily.
- **Animation**- Animation is a process of making a static image look like it is moving. An animation is just a continuous series of still images that are displayed in a sequence. The animation can be used effectively for attracting attention. Animation also makes a presentation light and attractive. Animation is very popular in multimedia application

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Applications of Multimedia

Following are the common areas of applications of multimedia.

- **Multimedia in Business**- Multimedia can be used in many applications in a business. The multimedia technology along with communication technology has opened the door for information of global wok groups. The multimedia network should support the following facilities:
    - Voice Mail
    - Electronic Mail
    - Multimedia based FAX
    - Office Needs
    - Employee Training
    - Sales and Other types of Group Presentation
    - Records Management
- **Multimedia in Marketing and Advertising**- By using multimedia marketing of new products can be greatly enhanced. Presentation that have flying banners, video transitions, animations, and sound effects are some of the elements used in composing a multimedia based advertisement to appeal to the consumer in a way never used before and promote the sale of the products.
- **Multimedia in Education**- Many computer games with focus on education are now available. Consider an example of various rhymes for kids. The child can paint the pictures, increase reduce size of various objects etc apart from just playing the rhymes.
- **Multimedia in Hospital**- Multimedia best use in hospitals is for real time monitoring of conditions of patients in critical illness or accident. The conditions are displayed continuously on a computer screen and can alert the doctor/nurse on duty if any changes are observed on the screen. Multimedia makes it possible to consult a surgeon or an expert who can watch an ongoing surgery line on his PC monitor and give online advice at any crucial juncture.
- **Multimedia Pedagogues**- Pedagogues are useful teaching aids only if they stimulate and motivate the students. The audio-visual support to a pedagogue can actually help in doing so. A multimedia tutor can provide multiple numbers of challenges to the student to stimulate his interest in a topic. The instruction provided by pedagogue have moved beyond providing only button level control to intelligent simulations, dynamic creation of links, composition and collaboration and system testing of the user interactions.
- **Communication Technology and Multimedia Services**- The advancement of high computing abilities, communication ways and relevant standards has started the beginning of an era where you will be provided with multimedia facilities at home. These services may include:
    - Basic Television Services
    - Interactive entertainment
    - Digital Audio
    - Video on demand
    - Home shopping
    - Financial Transactions
    - Interactive multiplayer or single player games
    - Digital multimedia libraries
    - E-Newspapers, e-magazines

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Biju Patnaik Institute of IT & Management Studies

Lecture Notes

on

IT FOR MANAGEMENT

Module-III

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Data:

- It is a collection of facts.
- The facts that can be recorded and which have implicit meaning known as 'data'.
- It can be a Number, Character, Image, or video.
  Example: Customer: 1. Customer Name
  2. Customer Number
  3. Customer City

Database: It is a collection of interrelated data. These can be stored in the form of tables. A database can be of any size and varying complexity. A database may be generated and manipulated manually or it may be computerized. Example: Customer database consists of the fields as Name, Number, and city

Database System: It is a computerized system, whose overall purpose is to maintain the information and to make that the information is available on demand.

CONCEPT OF COMPUTER FILES

File has been defined as a collection of related information known as records that may be accounts file, inventory file, payroll file, etc. Data files used in business have been classified into seven categories, these are:

(a) Master files - This file contains relatively permanent information which is used as a source of reference and is generally updated periodically. It contains relatively permanent records for identification and summarising statistical information. A stock file, customer file, employee file, product file, etc., are examples of the master file.

(b) Transaction file - The transaction files are created from a source document. Such files are generally used for recording events or transactions pertaining to the master file. Transaction files are used to update master files. If the processing is of the batch type, the transactions are collected for a period, and then at the end, the transaction file is created from collected transactions. For example, a customer file in a bank is a source document or master file. Transactions such as 'withdrawal', 'deposit', 'new account', 'close account' are events stored in the transaction file, and a new Masterfile is generated by updating the old Masterfile using a transaction file.

Transaction files are not needed after they are used once, but master files are permanent.

(c) Reference files - A file containing keys that have records in other files is known as a reference file because to retrieve a record from a file, referenced file is searched first to find the address of the record (or location).

(d) Table files - These files are tabular e.g. catalog or price lists.

(e) Report files - These files are concise, meaningful, and are generated from other data files. For example, the sales performance report of a sales person in a company may consist of yearly sales, zone sales, country sales, total-sale figures.

(f) Historical files - Files that contain information about the past are known as historical files. These files usually contain statistical information for trend analysis or compare values, etc.

(g) Back-up-files - These files are the copy of master files kept in the computer library as a measure of security, that can be recovered if a disaster happens.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

FILE MANAGEMENT: ORGANIZATION AND TYPES OF ACCESS

File Management refers to logically originating data into fields so that it is best understood by the users. File processing refers to an environment in which data are physically organized into files.

File processing environment varies for a file in application. The file management software programs used in an organization may be developed in-house or may be outsourced. Earlier both program files and data files were developed dependent on each other, this is known as a program - data dependence. For example, any program written for an application in any language COBOL or BASIC would require compatible COBOL or BASIC data files. Each programming language has its method for formatting data, and thus data file created for one language could not be used with programs written in any other language. This leads to the problem of data redundancy and integrity.

Data redundancy refers to the wastage of storage space due to inefficiency in data storage.

Data integrity refers to the accuracy of data and its conformity to its expected value, after being processed and transmitted.

Data Independence refers to the separation of data in a database from the programs that manipulate it. It makes stored data as accessible as passable.

File organization refers to the organization of
the data of a file into records, blocks, and access structure.

This includes the way records and blocks are placed on the storage medium and interlinked, whereas access methods are a group of operations that can be applied to a file. Some access methods can be applied only to files that are organized in a way. We have mainly focused on two types of access methods that are sequential access methods and direct access methods. These methods are influenced by how the data is organized on a file. Although there are few basic file organizations such as sequential, indexed, and direct, which are explained below:

1. **Sequential File Organization**

Sequential/ serial file access refers to a method of storing or retrieving information that requires the program to start reading at the beginning and continue until it finds the desired data.

In a sequential organization, records are arranged one after another in a predetermined order whereas in a serial organization records are arranged one after another in no particular order but in chronological order in which records are added to the file.

Advantages of sequential file organization are:

(i) Records are organized in a simple and contiguous fashion. Thus, such files are easy to design and simple to access as only a key field is required.

(ii) Sequential organizations are best to update master files when the number of transactions or file activity ratio is high and transactions are batched.

(iii) Sequential file organization is usually on magnetic tapes which is the cheapest medium.

(iv) Easy to maintain a backup.

(v) Less prone to errors, this type of file organization is efficient and economical.

Limitations of using sequential file organization are:

(i) To update the master-file, the transactions have to be batched and the whole master-file has to be read even if few records are to be updated.

(ii) This is costly and time-consuming if the file activity ratio is low.

(iii) Since the response from sequential files is low, thus online inquiry system cannot be generated using sequential files.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

2. Direct (or Random) Access File Organization

Direct access, file records are placed in a manner that the desired record is found directly without searching sequentially. As records are stored in random order, records are also retrieved in any sequence irrespective of their physical order on the storage medium. This is directly performed by the CPU to access the desired record using randomizing procedure without searching all other records in the file. This is known as self (direct addressing).

Random access file organization requires direct access storage devices, although primary storage in CPU provides direct access, there are other devices which provide direct access feature such as, magnetic disk, floppy, drums, etc. This organization is useful for an online inquiry system where computer response is required to be immediate.

Advantages of direct file organization are:

(1) Records can be accessed both randomly and sequentially.

(2) Access to record is quick as it is not required to search the whole file.

(3) Transactions need not be batched as any single transaction is served as it is entered.

(4) Records need not be sorted, it may be in any random order.

(5) It is useful in interactive, real-time online processing environments such as Railway / Airline reservation, banks, etc.

Limitations of Using direct file organizations are:

(1) It uses relatively expensive media such as magnetic disks.

(2) Storage space is not efficiently used because the location of the stored records is randomly arranged.

(3) It involves extra computation to locate the address of the records.

The most commonly used direct access techniques are:

(i) Direct sequential access

- The self-direct addressing method
- Index sequential addressing method

(ii) Random access

- Address generation method
- Indexed random.

- **Self (Direct) Addressing** refers to addressing the record by its record key. Thus, it is easy to compute the records address from the record key and physical address of the first record in the sequential file. This method assists in computing the bucket address of the fixed-length record, where keys are arranged in a consecutive number. For example, if 10,000 records of college students need to be stored in the magnetic disc pack of 6 disks, cylinder-wise, let the first cylinder carry the first 500 records, the second cylinder carries the next 500 records, and so on. During retrieval of these records, read/write heads move cylinder by cylinder in which records have been sequentially arranged. Now if ten faces of the first cylinder carry the first 500 records in the following manner i.e.,

  cylinder 1, 1 1 to 50.
  cylinder 1, 2 51 to 100
  cylinder 1, 10 451 to 500

To have direct access to such file organization, we can demonstrate it in the following way:

The total No. of buckets = 1000, No. of records per bucket = 10

Let buckets address range from 101 to 1100

Keys of the records range from 1 to 10,000

To find where is the record key 9892 located i.e., in which bucket it is stored, follow the steps given below:

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- Divide the record key (to be searched)/No. of records per 9892/10 = 989 and a remainder 2 bucket
- Add first bucket number to the quotient of the record key (from step 1) 101 + 989 = 1090
- The remainder (2) is the record's position within the bucket and remainder 0 indicates the last record in the preceding bucket.

Advantages of Self-Addressing are:

- It does not require an index.
- A random inquiry can be made because it is easy to locate the position of a record.

Disadvantages of Self-Addressing:

- This method is impractical because gaps in key values would lead to empty buckets.
- The records must be of fixed length.
- If some records are deleted their storage space remains empty. This results in wastage of storage space.

➢ **Indexed sequential file organization:** In an indexed file organization, the records are stored either in a sequential or non-sequential manner. In business data processing system, a common file organization is used in an ordered file with a multilevel primary index on its ordering key field, such an organization is called an indexed sequential file. These files have indexed sequential access methods (ISAM). In an indexed file organization, an index is created. This index is similar to the index given at the end of the book. To locate a record, the cylinder index is searched to find the cylinder address. Then, for that cylinder track index is searched to find the track address of the wanted record. Each index entry matches a key value of the record and points to a unique record. As index table usually contains the key-value (unique) and address of that record (track number, cylinder number, surface numbers). The accessing is faster in such file organization because no sorting and additional block access are required. However, the disadvantage is that one cannot apply an indexed access method to a file that is not indexed.

Advantages of using index sequential access file organization are:

- This includes the features of a sequential and direct organization. Direct access is achieved through generating an index.
- It can be used in any kind of application.
- If the file activity ratio is high, then this file organization is efficient and economical.

Limitations of using index sequential file organization are:

- It increases the processing time as a reading of indexes is required before accessing the record.
- It requires extra storage as record insertion causes an overflow condition.
- Compared to random access, it requires more time because indexes are searched before record access.
- Requires more space.
- Requires expensive storage devices such as magnetic disk and software as well.

THE BEST FILE ORGANIZATION

The factors that influence the selection of a file organization may vary from application to application. Some general factors include the following:

(i) File Activity - This refers to the ratio of the number of records accessed in execution by the total number of records in a master file. If in any application this ratio is less, then files should be on the disk so that they can be addressed directly.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

This ratio is extremely small for real-time files, where only one master record is accessed and processed at one time. In such cases, direct access organizations are best suited. On the other hand, if this ratio is high, then indexed sequential processing is preferable.

The extreme of this ratio is one where all the records are accessed during the execution of the master file. In such cases, sequential file organizations are preferred. Examples of such organizations are payroll files, updating the master file with a transaction file at the end of the day in a bank.

(ii) File Volatility - Refers to the number of updates i.e. addition or deletion performed to the file in a given time. Thus, high volatile files are of those applications which require a large number of additions/ deletions. For high volatile applications direct access file organizations are best suited because, in indexed sequential access methods (ISAM) file organizations, the addition of records would have to be placed in the overflow area. Sequential file organizations would be appropriate if there were no interrogation requirements.

(iii) File Interrogation - It means the retrieval of information from a file. For real-time processing retrieval of records is usually faster, thus, direct file organization is preferred. In a batch-processing environment where the time of retrieval of records is not all that critical, sequential file organization may be considered. Moreover, the direct access method is preferred for large files over small files, because searching for records sequentially or with binary search in small files is easier than maintaining complex indexes.

(iv) File density - Refers to the number of bits stored per inch. File density determines the actual utilization of the storage space allocated to a file. The greater density implies low-cost storage and can be used for archival and backup storage purposes. In the case of greatly dense files, sequential file organization is preferred, as in such file organization 100% density can be achieved.

DRAWBACKS IN FILE BASED SYSTEM

As discussed above the file-based approach was difficult to manage information system because it suffered the following disadvantages;-

(i) Data redundancy: The same piece of information is stored in two or more files. Thus, there was a need to avoid this data redundancy and the keeping of multiple copies of the same information and replace it with a system where the address is stored at just one place physically and is accessible to all applications from this itself.

(ii) Program/Data dependency: In a traditional file system if the data field is to be added to a master file, all programs that access this master file, required to be changed.

(iii) Lack of flexibility: Because of the strong coupling between the program and the data, most information retrievals was limited to predetermined request for data.

(iv) Data availability: This file processing system caused serious inconsistency in the development of the programs which access such data because data is scattered in many files. Due to non-uniformity in file design, similar data may be stored in different files with different identification thus obtaining data is usually difficult.

(v) Data Integrity: Integrity (accuracy and completeness) of the data is suspected because there is no control over their use and maintenance by an authorized user and the same data is found in different forms in different files. Thus, data validation, when implemented may result in data in one file is correct and the same data in another file is wrong.

(vi) Data management control: As the same data is available in different files, it is difficult to relate them and implement uniform policies in an organization. Thus, because of a lack of central control over data, security standards cannot be enforced.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

WHAT ARE DATABASES?

The database is a collection of related data. By related data we mean, known facts that can be recorded and have implicit meaning. In other words, the database is derived from the real-world having some degree of interaction and a targeted audience that is actively interested in the contents of a database. Now databases and database systems have become an essential component of everyday life these days. As Journal on Information Technology quotes about databases that "Databases have evolved over centuries to cater to the needs of changing times and businesses. Today, they are integral to the success and progress of any organization". Our daily activities involve the same interaction with a database, for example, transactions to the bank account may be for deposit or withdrawal; Railway/airline/hotel reservation. These examples are traditional database applications. We may also consider the collection of words that make up this page of text to be related data and hence to constitute a database.

The database has the following properties:

- The database represents a part of the real world and any changes in this are reflected in the database.
- A database is a logically coherent collection of data with some inherent meaning. This means that data in the database belongs to one application.
- A database can always be designed, built, and populated with data for a specific purpose. The users of such a database are preconceived with the application.
- A database can be of any size and varying complexity.
- A database may be generated and maintained manually or it may be computerized. For example, the Library card catalog is a good example of this, which can be created manually and it can be automated as well.

A database is a logically coherent collection of data with some inherent meaning to facilitate rapid updating of individual records, concurrent updating of related records, easy access to all records by all application programs, generating routine or special purpose reports.

Types of DataBases based on processing

A database may be characterized as a warehouse which is "a subject-oriented, integrated, non-variant collection of data in support of management decisions". Thus, databases can be managed dynamically and can be processed on-line. Primarily database processing is of two types:-

(a) Analytical databases
(b) Operational databases

a) **Analytical Databases** (OLAP - On-Line Analytical Processing) are static, read-only databases. In this, archived and historical data is stored and used for analysis. For example, to analyze demographic marketing sales strategies, a company may store data for more than five years. On the web, the analytic database may be seen in the form of inventory catalog e.g. in Amazon.com such web pages are dynamic in nature and query can be made as it contains descriptive information about all the available products.

b) **Operational Databases:** Operational databases (OLTP - On-Line Transaction Processing) are used to manage more dynamic data. These types of databases allow the user to modify the data (add, change or delete). These types of databases are used to retrieve real-time information. Operational data is usually used by the organization to keep track of stock quantities.

Types of Databases: Based on Organization Requirement

The databases have been developed in numerous organizations to meet a particular need. Thus, the growth of distributed processing data warehousing, data mining, end-user computing, DSS (decision

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

support system) also known as EIS (Executive Information System) have caused the development of several types of databases. The six types of databases found in computers using organizations are the following:-

i. **Operational Databases:** Operational databases store detailed data with anticipated processing requirements programmed in conventional computer languages. The basic objective of an operational database is to support routine operations. They are also called supervisory systems. The response time in these databases is very fast, in seconds. But these databases take a long time to respond to a new request for information. Typical examples of operational databases are airline reservation and bank teller systems. As these databases serve a particular need, thus they are Subject Area Databases (SADB).

ii. **Management Databases:** Management databases are those which can respond to ad hoc requests for information that require quick answering. These databases store selective data extracted from the operational database. As the details of the type of information sought are not known before it being requested, a special capability is required in the system to process ad hoc requests by quickly generating a program that will respond to the request. The data exist but the request is new, so a new program is generated rapidly to answer the request. In management databases, a data request is spontaneous and not preprogrammed. The basic objective of a management database is to support planning and urgent information needs. The response time in these databases is relatively higher than in operational databases, but the time required to respond to a new request is relatively short. The typical users of management databases are information staff, middle management, and top management. A typical example of a management database is a sales analysis system and personnel information system.

iii. **Information Warehouse Databases:** An information warehouse stores a multidimensional data model, which supports OLAP and decision support technologies. In comparison to traditional databases, a data warehouse generally contains very large amounts of data from multiple sources and multiple databases. These data warehouses provide storage, functionality, and responsiveness to queries beyond the capabilities of transaction-oriented databases. An information warehouse primarily stores data from current professionals throughout an organization. The advantage of using an information warehouse database is pattern processing as they are designed for random searching and simple operations. However, the operational database is processed to identify key factors and trends in the historical pattern of the business activity.

iv. **Distributed Databases [DDB]:** These databases are defined as a collection of multiple logically interrelated databases distributed over a computer network. These databases may belong to local workgroups and departments at regional offices, branch offices, manufacturing plants, worksites, etc. The distributed databases may be located within the same building via a local area network or geographically distributed over large distances and connected via WAN - Wide Area Network. Thus, these databases include a segment of both common operational and common user databases, also data generated and used only at the user's site. However, distribution leads to increased complexity in system design and implementation, but major considerations of data resource management are to ensure that all data in an organization's distributed databases are consistently and concurrently updated. Moreover, distributed databases can access remote sites and transmit queries maintaining the integrity of the overall database, keeping track of replicated data, etc.

v. **End-User Database:** End users are the people whose jobs require access to the database for querying, updating, and generating reports. The database primarily exists for their use, thus these databases consist of a variety of data files developed by end-users at their workstations. Here, the end-user

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

database may contain documents generated by MSWORD or data files generated from spreadsheet package or DBMS package.

vi. **External Databases:** It is the part of the database that a particular user group is interested in and hides the rest of the database from another user group. End-users can access the external database from commercial information services (organizations) by providing a fee. The data of external databases may be available from statistical data banks (i.e. statistics on economic and demographic activity), bibliographic data banks (abstracts of newspapers, magazines, and other periodicals).

vii. **Text Databases:** The text databases support storage, search, and retrieval for text objects. Text databases are an outcome of the advancement in the use of computers in creating and storing documents electronically. The large text databases provide online database services of bibliographic information. Text databases are in use with big corporations, government agencies, and small companies using microcomputers. They use text DBMS software to help create, store, retrieve, and store other information as text in such databases. The Microcomputer version of this software has been developed to assist users to manage their text databases on CD ROM disks.

viii. **Image Databases:** Traditional databases store alphanumeric records in files or text databases, but these days databases can store images in a variety of formats such as TIFF, GIF, JPEG, commonly used image databases are electronic encyclopedia available on CD ROM disks that store numerous photographs, various animated video sequences, a large number of pages of text. Image databases are used in business for document image processing. In business, one may be required to process customer correspondence, invoices and purchase orders, sales catalogs, etc. these can be optically scanned through a scanner, which stores the document on an optical disk. Image database management software allows its user to store millions of pages of document images, view and modify document images at their work stations. It allows the user to transfer its data to other workstations in the organization.

COMPONENTS OF DATABASE SYSTEM

The database system comprises four components.

(a) Database files: The database systems are designed and developed around centralized, integrated shared files, that emphasize the independence of program and data.

These files have data elements stored in database file organization formats. As database system varies from application to application, the thus different database system has different ways of representing data physically on disk. These physical representations may be of varying complexities that are usually hidden from the end-user. The user is concerned with the logical view of data to view relationships among data from different files. The way data files stored at the physical level are managed by DBMS facilitates faster access and economically saves the storage area. Physical data descriptions are concerned to the MIS department because they select the database system that meets the end-user needs at a minimum cost.

(b) Hardware: Four hardware components are required, first, for support of the execution of the database system i.e. hardware processor and associated main memory. Second, for storage, secondary storage devices e.g., magnetic discs. Third for end-users on-line terminals are required for accessing and updating database files.

These terminals may be adjacent to the computer or even thousands of miles away. They may be dumb terminals, smart terminals, or micro-computers.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

The fourth hardware component required is the output system or report generator. It allows the user, programmer, or manager to design output without writing an application program in a programming language and it produces routine reports and special reports.

(c) Software: To handle all requests made to database by-user are facilitated by a software known as Database Management System. It also provides insulation to physical database hardware-level detail of database) from its users.

DBMS allows the user to access the files, update the records, insert or delete new records and retrieve data as per request. DBMS ensures data security, integrity because the database is accessed by many users. The other software that playa vital role and are submodules of DBMS are:

(i) A host language interface system: This is one of the modules of DBMS that communicates with application programs. Its basic functions are (a) To interpret instructions of an application program written in a high-level language such as COBOL or BASIC program (b) to retrieve requested data from the database files. However, application programs do not contain information about the file because program files are independent of the data files, but the operating system interacts with DBMS during the retrieval of the data, the above two-staged process works smoothly.

(ii) The application program: These programs are independent of the data files in a database environment. To perform a query, it uses standard data definitions. The program/ data independence and standardization make program development faster and easier. These programs are developed by skilled programmers using a host language interface.

(iii) A natural language interface system: The English-like, simple, easy-to-understand and write, command-oriented language SQL (structured query language) is used to retrieve and perform an online update on to the database by the user. The language constructs names consist of minimal syntax. It can be used by managers, programmers, and end-users to access databases and retrieve information about record status and content.

(iv) Data Dictionary: It is often seen that when a program becomes somewhat large, keeping a track of all the variable names that are used in the program and the purpose for which they were used becomes extremely difficult. It may be possible for a programmer to remember the variable names, but if the same author comes back to his program after a significant time or another programmer need to modify the program, it would be extremely difficult to make a reliable account of the purpose of the data files used in the program.

For this purpose, a data dictionary/directory is emerging as a major tool. As inventory provides definitions of things, a directory tells us where to find them. A data dictionary/directory contains information (or data) about the data.

The data dictionary holds detailed information about different structures and data types, the details of logical structures that are mapped into different structures, details of the relationship between data items, details of user privileges, and access rights.

(d) User: Users can be categorized based on their level of interaction and association with the databases, these are:

- **End-user:** This class of user interacts with the system where the database application is residing, through an on-line terminal using a query language. Here the user is provided with a built-in query language processor, by which the end-user can query the database by selecting commands such as SELECT, INSERT, DELETE, etc., and issue it to DBMS.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- **Application programmer:** Application programmers are conventional programmers responsible for writing programs in 4GLs (fourth-generation languages, i.e. COBOL, C++, Java) for database applications. These programs are independent of the data files. Such programs provide access to the end-user to the database from an on-line terminal by entering the desired request, generally in SQL statement to the DBMS.
- **Database administrator (DBA):** DBA is responsible for the overall control and security of the database system. The responsibilities of the database administrator include: (a) He decides the information contents of the database; (b) Implementation of technical controls to ensure data security and technical services for the effective performance of operations; (c) He creates new accounts and passwords for a user or group of users to enable them to access the DBMS (d) DBA grants certain privileges to certain accounts (e) He defines authorization checks and validation procedure (f) He defines the strategy for backup and recovery. (g) Implementation of various security policy decisions regarding, allowing access to the database only to legitimate users.

(e) Procedures: The final component of a database system is procedures. Both users and the operational staff need documented procedures for normal conditions. The users need to know how to sign in to the system, how to use the terminals, how to provide data, and so forth.

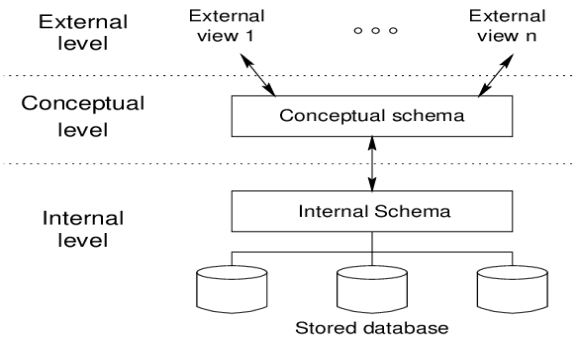
They also need to know procedures that ensure that they do not interfere with one another. Procedures are also needed for computer operations personnel to know how to start and stop the database applications and how to perform a backup. Procedures are needed both for users and operations personnel that describe the actions to be taken in the case of database system failure.

DATABASE SYSTEM CONCEPTS ARCHITECTURE

An early proposal for a standard terminology and general architecture for database systems was produced in 1971 by the DBTG (Data Base Task Group) appointed by the Conference on Data Systems and Languages (CODASYL, 1971). The DBTG recognized the need for a two-level approach with a system view called the schema and user views called sub-schemas.

The data as perceived at each level is described by a schema. A sub-schema is the specific set of data from the database, that may be required by each application. Therefore, in the case of the external level, several schemas exist.

Here is the figure showing the ANSI_SPARC Architecture of the database system:



The levels form a three-level architecture that includes an external, a conceptual, and an internal level.

The internal level: This describes the physical storage structure of the database along with the details of data storage and access paths for the database.

The conceptual level: The conceptual level gives the complete structure of the whole of the database for all users. In this description, the physical storage structure is hidden, but the emphasis is more on

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

describing entries, data types, relationships, user operations, and constraints. Data-mode implementation takes place at this level.

The external level: At this level, the user views part of the database as per the access rights given to the user.

The way users recognize the data is called the external level.

These three levels of database architecture separate the user application and the physical database: i.e. how it is stored on physical devices. Thus, there is always a correspondence at each level and its subsequent level, known as mapping. Thus, the above-stated architecture contains two types of mappings, these are:

External/ Conceptual mapping: This mapping defines that to a particular external view there is a corresponding conceptual view.

Conceptual/Internal mapping: This mapping basically, defines that to all records at the conceptual level there is a corresponding representation at an internal level.

SCHEMA AND INSTANCE

The data which is stored in the database at a particular moment is called an instance of the database. The overall design of a database is called schema.

• A database schema is the skeleton structure of the database. It represents the logical view of the entire database.

• A schema contains schema objects like table, foreign key, primary key, views, columns, data types, stored procedure, etc.

• A database schema can be represented by using the visual diagram. That diagram shows the database objects and their relationship with each other.

• A database schema is designed by the database designers to help programmers whose software will interact with the database. The process of database creation is called data modeling.

• A schema diagram can display only some aspects of a schema like the name of record type, data type, and constraints. Other aspects can't be specified through the schema diagram.

WHAT IS DATABASE MANAGEMENT SYSTEM (DBMS)?

The DBMS is a general-purpose software system that facilitates the processes of defining, constructing, and maintaining databases for various applications.

A database management system is a collection of programs that enables users to create and maintain a database. Defining a database involves specifying the data types, structures, and constraints for the data to be stored in a database. Constructing the database is the process of storing the data itself on some storage medium that is controlled by the DBMS. Manipulating a database includes some functions as querying the database to retrieve specific data, updating the database to reflect changes in the mini world, and generating reports from the data.

A DBMS may typically perform the following functions:

i. It loads the existing data files, text files, or sequential files into the database as per desired database structure.
ii. The database and DBMS catalog are usually stored on disk. Thus, the module of the DBMS controls access to information stored on disk. It also handles the data transfers in the main memory from secondary storage.
iii. When a command coded in data manipulation language reaches to run time database processor, then control passes from the application program to the DBMS. The control unit of the CPU causes each instruction of the application program to be executed in sequence.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

iv. The query it receives may be retrieval or update operations to be carried out onto the database. It parses, analyzes, and compiles or interprets a query by creating database access code and executes that code.
v. DBMS requests the operating system to schedule Input/output operations.
vi. DBMS interacts with the operating system when the disk access the database or to the catalog. DBMS also interfaces with compilers for general-purpose host programming language.
vii. DBMS backup utility module creates a backup copy of the database that can be used to restore the database in case of catastrophic failure.
viii. DBMS utility module reorganizes a database file into different file organizations to improve performance.
ix. DBMS provides status information to the application program regarding records found or not found status.
x. DBMS may be available for sorting files, handling data compression, monitoring access by user, and provide statistics to the DBA.
xi. Such responsibilities are performed by the database manager program module of the database management system, which is described below.

COMPONENTS OF A DBMS

The major software modules or components of DBMS are as follows:

(i) Query processor: The query processor transforms user queries into a series of low-level instructions. It is used to interpret the online user's query and convert it into an efficient series of operations in a form capable of being sent to the run-time data manager for execution. The query processor uses the data dictionary to find the structure of the relevant portion of the database and uses this information in modifying the query and preparing an optimal plan to access the database.

(ii) Run time database manager: Run time database manager is the central software component of the DBMS, which interfaces with user-submitted application programs and queries. It handles database access at run time. It converts operations in user's queries coming. It accepts queries and examines the external and conceptual schemas to determine what conceptual records are required to satisfy the user's request. The run time database manager is sometimes referred to as the database control system and has the following components:

- **Authorization control**: The authorization control module checks the authorization of users in terms of various privileges to users.
- **Command processor**: The command processor processes the queries passed by the authorization control module.
- **Integrity checker:** It .checks the integrity constraints so that only valid data can be entered into the database.
- **Query optimizer:** The query optimizers determine an optimal strategy for the query execution.
- **Transaction manager:** The transaction manager ensures that the transaction properties should be maintained by the system.
- **Scheduler:** It provides an environment in which multiple users can work on the same piece of data at the same time in other words it supports concurrency.

(iii) Data Manager: The data manager is responsible for the actual handling of data in the database. It provides recovery to the system which that system should be able to recover the data after some failure. The buffer manager is responsible for the transfer of data between the main memory and secondary storage. It is also referred to as the cache manager.

Execution Process of a DBMS

As a show, conceptually, the following logical steps are followed while executing users to request to access the database system:

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(i) Users issue a query using a particular database language, for example, SQL commands.
(ii) The passes query is presented to a query optimizer, which uses information about how the data is stored to produce an efficient execution plan for evaluating the query.
(iii) The DBMS accepts the user's SQL commands and analyses them.
(iv) The DBMS produces query evaluation plans, that is, the external schema for the user, the corresponding external/conceptual mapping, the conceptual schema, the conceptual/internal mapping, and the storage structure definition. Thus, an evaluation plan is a blueprint for evaluating a query.
(v) The DBMS executes these plans against the physical database and returns the answers to the user

DBMS LANGUAGES

For defining each level or component of the database i.e; internal level, conceptual level, and external level, DBMS provides languages and these are:-

i. **Data Definition Language (DDL) :**

Data Definition Language (DDL) statements are used to classify the database structure or schema. It is a type of language that allows the DBA or user to depict and name those entities, attributes, and relationships that are required for the application along with any associated integrity and security constraints. Here are the lists of tasks that come under DDL:

- CREATE - used to create objects in the database
- ALTER - used to alters the structure of the database
- DROP - used to delete objects from the database
- TRUNCATE - used to remove all records from a table, including all spaces allocated for the records, are removed
- COMMENT - used to add comments to the data dictionary
- RENAME - used to rename an object

DDL performs the following functions:

- DDL is used by DBA and by database designers to define the conceptual and internal schema.
- DDL specifies the fields in each record and the record's logical name.
- It specifies the data type and name of each field.
- It defines the structure and format of the data.
- It provides data independence at the logical and physical levels.
- It specifies and provides the means to establish the relationship among various records.
- It identifies the records by unique data items known as KEY.
- It sets up data security access and change/update restrictions.

ii. **Data Manipulation Language (DML):**

A language that offers a set of operations to support the fundamental data manipulation operations on the data held in the database. Data Manipulation Language (DML) statements are used to manage data within schema objects. Here are the lists of tasks that come under DML:

- SELECT - It retrieves data from a database
- INSERT - It inserts data into a table
- UPDATE - It updates existing data within a table
- DELETE - It deletes all records from a table, the space for the records remain
- MERGE - UPSERT operation (insert or update)
- CALL - It calls a PL/SQL or Java subprogram
- EXPLAIN PLAN - It explains the access path to data
- LOCK TABLE - It controls concurrency

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

The summary of functions and features of DML/SQL is following:

- (SQL - Structured Query Language is a data manipulation language used in querying, updating, and managing relational databases. It is English-like command such as SELECT ... FROM .... WHERE. SQL is a simple English-like language.
- DML/SQL allows users to perform typical manipulations that include retrieval, insertion, deletion, and modification of data.
- It is a comprehensive integrated language.
- DML statements can be entered interactively from a terminal or can be embedded in a general-purpose programming language, such as COBOL, PL/I, and C++_ Thus it provides, independence of programming language.
- It allows retrieving many records.
- It provides a user-friendly interface for end-users.
- It provides a record relationship.
- Its syntax is easy that it can be used by non-programmers

iii. **Data Control Language** There are another two forms of database sub-languages. The Data Control Language (DCL) is used to control privilege in the Database. To perform any operation in the database, such as for creating tables, sequences or views we need privileges. Privileges are of two types,

- System - creating a session, table, etc. are all types of system privilege.
- Object - any command or query to work on tables comes under object privilege. DCL is used to define two commands. These are:
- Grant - It gives users access privileges to a database.
- Revoke - It takes back permissions from the user.

iv. **Transaction Control Language (TCL)** Transaction Control statements are used to run the changes made by DML statements. It allows statements to be grouped into logical transactions.

- COMMIT - It saves the work done
- SAVEPOINT - It identifies a point in a transaction to which you can later rollback
- ROLLBACK - It restores the database to its original since the last COMMIT
- SET TRANSACTION - It changes the transaction options like isolation level and what rollback segment to use

COMPARISON BETWEEN DDL AND DML

DDL	DML(SQL)
DDL stands for Data Definition language.	Data manipulation language.
It defines the physical characteristics of each record; describes subschemas and schemas.	It provides a way to process the data that is already existing.
It is used to define the content and structure, format, of the database. It also specifies the relationship between records.	It manipulates the data by various techniques such as retrieval, insertion, deletion, replacement, sorting, etc
It establishes a link between conceptual and internal views.	It establishes a link between user view and conceptual view.
Describes the data type and name of each element in the Database.	It accesses the data by the name specified in DDL.
Indicates the keys in the relations.	It provides the means to use the keys to establish a relationship.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Provides data security restriction.	Provides data retrieval, updates deletion, insertion of data.
Provides means for associating related records or fields.	DML provides the use of record relationships.
Provides independence from storage structure and access strategy.	Provides independence from programming, also supports high-level languages.

ADVANTAGES OF DATABASE APPROACH USING DBMS

The database is a collection of related information stored so that it is available to many users for different purposes. Therefore, the contents of the database belong to the different departments of the organization so that data is available to all users and redundancy in the data be minimized. Besides these, there are several other advantages listed below:

(i) Standards can be enforced: In a centralized database environment, industry, State, country standards can be enforced more easily than in an environment where each group of users controls their files and software.

(ii) Development time of an application may be reduced: Designing and implementing a new database may take more time but once the database is set up and running, then developing any new application using DBMS requires 1/ 6th time of the traditional file processing system.

(iii) Flexible: DBMS allows changes to the structure if required, of the database affecting the stored data and existing application.

(iv) Availability of currently updated information: In transaction processing applications such as reservation system or banking database, it is essentially required to retrieve the information which is most currently updated. Moreover, the information or data may be updated just by one user according to the priority of the request.

(v) Reduces overall cost: The DBMS approach reduces the wasteful overlap of activity of data processing personnel in different departments. This reduces the overall cost of operation and management as well.

(vi) Redundancy can be controlled: Being central control over data removes, duplication of data also reduces the application of effort of recording and so is the updates of data. Secondly, saving of the storage space and thirdly, the inconsistency of the data may be avoided by using a database approach.

(vii) Unauthorized access may be restricted: Database is used when multiple users share the database. Financial data is considered to be confidential. Thus, only authorized user is permitted to make changes in the existing database, whereas another user category is allowed only to retrieve the database. This modern DBMS provides a security and authorization system for each class of user.

(viii) Provide multiple user interfaces and support 4GL (or non-procedural language) for manipulation of data: DBMS provides a variety of user interfaces. The most common is the graphical user interface (GUI). In addition to this users do not have to learn programming techniques to access the database. The DBMS supports non-procedural language such as SQL or QBE to access the database.

(ix) Support multiple files and complex relationships among data: DBMS generally can represent a variety of complex relationships among data (As Database is an integration of several files), which may be retrieved and updated efficiently.

(x) Database can maintain Integrity: A DBMS generally possesses the capability for defining and enforcing the constraints (restrictions) that must hold for the data, which results in the correctness and accuracy of data. For example, a student database contains information about marks scored by a

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

student in an examination, the maximum marks are 100 and a student cannot score more than 100. Integrity constraints can be specified by the user while designing the system, so that database will accept numbers only in the range of a to 100.

(xi) Provision of backup and recovery system: If there is any failure of hardware or software, then DBMS provides the facility of recovery by maintaining its regular backup. Moreover, it also ensures that the program may be resumed from the point at which it was interrupted.

LIMITATIONS OF DATABASES

Despite the advantages of using DBMS, there are certain limitations of using DBMS:

Overhead Cost: when very large data need to be managed in a largescale organization, then the database approach is motivated. This also requires a powerful hardware platform and software for database management, which are quite expensive. Another cost incurred will be the hiring of system analysts, database designers, database administrators, programmers, and data processing personnel and the cost of training. This means to adopt this approach, a significant extra cost has to be borne by the organization.

Security problem: Another disadvantage of this approach is that sharing of data also carries the risk of the data being accessed by an unauthorized user. Thus, the organization needs to cope with this problem by taking security measures, concurrency control, recovery, and integrity.

The problem of Resources: To run an on-line, real-time system to answer on-line queries, requires a large amount of data to be stored. As a result, more terminals may be needed to put managers and other users online. Communication devices are also required to connect extra terminals to the database. It may require resources such as a multiprocessor system and software to run a DBMS. Therefore, DBMS may require extra computing resources depending upon the application.

Ownership problem: In a file-based system, the programmer/user is the owner of the data and program whereas a database consisting of such files is owned by the entire company. For any change or read or insertion of data in the database, a user needs to seek permission from managers of the company. For a database to be successful the database must be viewed and updated as a corporate resource, not as an individual resource.

Concurrency problem: Several problems can occur when concurrent transactions execute in an uncontrolled manner. There may be a lost update problem when two transactions that access the same database items have their operations inter-leaved in a way that makes the value of some database item incorrect.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

DATABASE MODELS

Databases can be differentiated according to function and also according to how they model the data. Technically data models are an abstraction, usually mathematical algorithms and concepts. The evolution of databases is due to the analysis and design of data models. The advancement in data models has led to database efficiency.

As defined above, the database is the integration of multiple related files, each file contains records and each record comprises certain attributes. An association exists between different attributes of an entity. An association between two attributes indicates the interdependence between the values of the associated attributes. This association is used in modeling the database because these files are linked together in a manner known as database models/ structure. The databases can be modeled in three ways, these are:

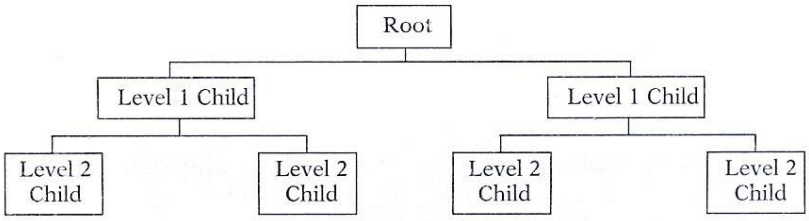
1) Hierarchical
2) Network
3) Relational

1) Hierarchical DataBase Model

As the name implies the hierarchical database model defines hierarchically arranged data. The hierarchical database represents a one-to-many (1:N) relationship. It can be visualized as a parent-child relationship. A parent data element has one or more child/sub-ordinate data elements. It may also be represented as an inverted tree structure having roots at the top and branches below this.

The hierarchical model is regarded as an inverted tree in which each record may be 'parent' of one or more child records. Thus when database records are modeled in the above-said structure, it is known as a Hierarchical database.

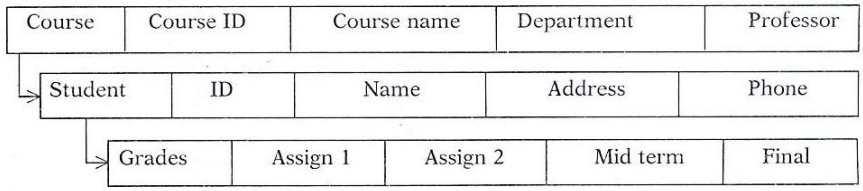
Thus, the hierarchical schema defines the tree data structure. As per hierarchical convention, every node except the root has exactly one parent node and a node can have several child nodes and they may be ordered from left to right. Parents and children links are called pointers. Hierarchical database modeling is not a popular technique because it does not support flexible data access. The data elements can be accessed only by following the path down in the structure.



EXAMPLE OF HIERARCHICAL DATABASE MODEL

In the above figure of an upside downflow of data, a single table (or file) acts as the root of the database, from which other tables (or files) branch out. The main advantage of using a hierarchical model is its efficiency in that it does not retain redundant data. If a change in the data is necessary, the change needs to be processed once.

For Example, consider the student file implemented in the hierarchical database model



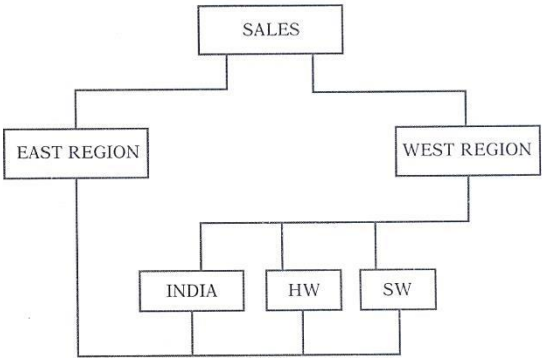
HIERARCHICAL DATABASE SCHEMA

The major disadvantage of the hierarchical database model is that we cannot add a record to a child table until it has been incorporated into the parent table. But the major limitation is that the hierarchical database model can handle the 'One: many' relationships and not ‘many: many'. This is because the child may have only one parent. However, in a realistic situation, a student can take more than one subject and a subject may also be taken by many students.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

2) Network DataBase

Network database represents many many relationships (M: N). A network data model can be defined as an extension of the hierarchical database structure. In this modeling technique, a child data element can have more than one parent or no parent at all. Therefore, when a database is initially designed, the relationship among the data elements must be determined. The basic data structures in the network model are records and sets. As we know data is stored in records, each record consists of a group of related data values. Records are classified into record types. Each record type specifies the structure of a group of records that store the same kind of information. Thus a typical database application has several record types. A network model is similar to a hierarchical model, such that the hierarchical model is the subset of a network model. Network model uses a set theory to provide a tree-like hierarchy. The network model represents the relationship between records using the construct called set type. A set type describes the 1:N relationship between two record types. Each set type is accomplished by its name, an owner record type, and a member record type. For example, a set type having name Major-Department, the owner record type is Department with Student as the member record type. A record of a member record type can appear only in one set occurrence. Thus, M: N relationship between two record types cannot be represented as a single set type and another link has to be defined as given in the figure. The network data modeling, however, supports flexible data access, because the desired data element fetching is not required to route from the root data element. It forms a complex structure and, therefore, network databases are not used quite often.



NETWORK DATABASE MODEL

3) Relational DataBase

A relational database is built and operated following the Relational Model of Data proposed by E.F.Codd (1970). This model has now gained wide acceptance and has engendered a great deal of additional study covering numerous aspects of database theory and practice.

Primarily, the Relational Model provides a simple and intuitive method for defining a database, storing and updating data in it, and submitting queries of arbitrary complexity to it. More important, it provides a firm, sound, and consistent foundation for all the other topics that database management systems must commonly embrace, such as security and authorization, database integrity, transaction management, recoverability, and distribution of data.

The Relation Model is based on the mathematics of n-ary relation, which is in turn founded on the disciplines of predicate calculus and set theory. The relational model is the concept of a table, also called 'relation'. Each table is made up of records. Horizontal rows, known as tuples or records, and vertical columns are also known as attributes. The number of attributes determines the degree of the relation. The number of tuples in a relation is known as the cardinality of the relation. Each table can be identified by a unique name. In relational databases, each tuple is uniquely identified by one or more field(s) known as the

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

primary key. For example student roll numbers, social-security-number, customer account numbers are unique for each record, thus referred to as primary keys.

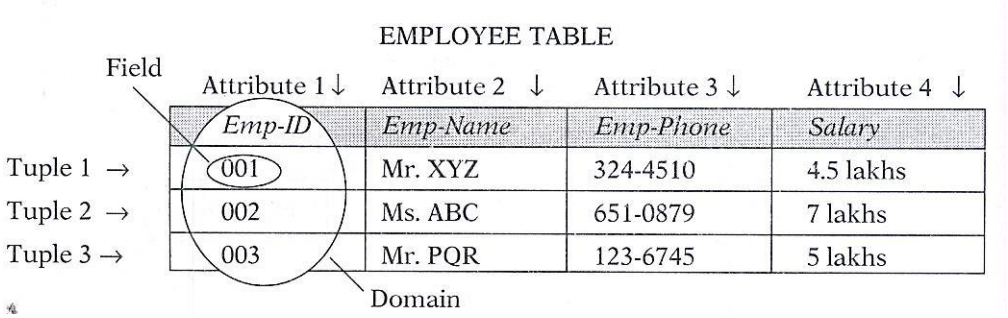


TABLE: RELATIONAL DATABASE MODEL

In this example EMPLOYEE TABLE is nothing but a file, where data is represented in a tabular form. This can also be called an Employee Relation with each row being a tuple and four attributes. Therefore, this relation is having three tuples and the degree of this relation is four. Moreover, Emp- ID is the primary key in the Employee table.

Constraints in Database

The difference between the database and file approach is the use of constraints in a database. Data values stored in a computerized database are often subjected to a validation condition. The condition may be expressed as a range of values within which the data should be accepted as an input. In a relational database, such a condition is called a constraint. There are two types of constraints:

i) **Uniqueness constraint:** This is the simplest kind of constraint, in which a data value in a column of a table must be different from the values already stored in the same column in the same table. The 'Uniqueness constraint' check is automatically performed by DBMS to ensure that new value is not already existing in the primary key column of the table.

ii) **Referential Constraint:** The primary key in one table appears as a foreign key in another table are equal (or matches) then it is named as a referential constraint. If the check fails, then the table is not updated.

The connection between referential constraints and uniqueness constraints refers to a column in another table that follows uniqueness constraint.

KEY CONSTRAINTS

Keys are nothing but attributes that are used to uniquely identify the rows in a relation.

1. Primary Key: A primary key is a column or columns in a table whose values uniquely identify each row in a table.

In the Employee table, we might use ENAME to find a particular entry. But that column does not make a good key. Just imagine what will happen if more than one employee has the same name "Ram". In most cases, we shall create our keys to ensure that they are unique. For example, we have created EMPNO as an identification number.

The relationship between the primary key and the rest of the data is one-to-one that is each entry for key points to exactly one employee row.

2. Composite Key or Concatenated Keys: In many cases, when we use more than one column as a part of the primary key. These are called Composite Keys or Concatenated Keys. We use composite keys when the table contains one to many or many-to-many relationships.

3. Foreign Key: A foreign key is a column or column whose values are the same at the primary key of another table. We can think of a foreign key as a copy of the primary key from another relational table. The relationship is made between two relational tables by matching the values of the foreign key in one table with the values of the primary key in another.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

ADVANTAGES OF RELATIONAL APPROACH

The popularity of the relational database approach has been apart from access of availability of a large variety of products also because it has certain inherent advantages:

(1) Ease of use: The information is represented in tables consisting of rows and columns, which is quite easy to understand, therefore, even first-time users find it attractive.

(2) Flexibility: Different tables from which information needs to be extracted, can be easily manipulated by using link operators, such as project and join to provide information in the desired form.

(3) Precision: As RDBMS is based on mathematical set theory, the usage of relational algebra and relational calculus for the manipulation of the relations/tables ensures that there is no ambiguity, which may otherwise, arise in establishing the linkages in a complicated network type database.

(4) Security: Security control and authorization can also be implemented more easily by moving sensitive attributes in a given table into a separate relationship with its authorization controls. If the authorization requirement permits, a particular attribute could be joined back with others to enable full information retrieval.

(5) Data Independence: Data independence is achieved more easily with the normalization structure used in a relational database than in the more complicated tree or network structure.

(6) Data Manipulation Language: The possibility of re ponding to ad-hoc query using a language based on relational algebra and relational calculus is easy in the relational database approach because data is organized in other structures. The query language either becomes complex or extremely limited in its capabilities.

DISADVANTAGES OF RELATIONAL APPROACH

RDBMS possess a major constraint and, therefore, a disadvantage in the use of relational database system is machine performance. If the number of tables in a relation is large and the tables themselves are voluminous, then the performance of responding queries to RDBMS is degraded although, it is appreciated that the simplicity in the relational database approach would arise in the logical view. For example, in an interactive system, an operation like join would depend upon the physical storage also. It is, therefore, common in a relational database to tune the database and in such a case the physical data layout would be chosen to give good performance in the most frequently run operations. Hence, it would result that the frequently run operations that would tend to become even more shared.

DIFFERENCE BETWEEN RELATIONAL AND OTHER MODELS

1) **Implementation independence:** The relational model logically represents all relationships implicitly. Therefore, the physical associations represented by an efficient access path among various files, one does not know without looking at the internal data model. However, the complexities of Network and Hierarchical models are defined at a logical level. Thus, the relational model provides data independence over other models.
2) **Logical key pointers:** The relational data model uses primary and secondary keys in records to represent the association between two records. This model's implementation is independent of the data manipulation language as only the physical database use address pointers and the physical database is masked from the user of a relational database.
3) **Normalization theory:** Properties of a database that make it free of certain problems of updation and redundancy have been developed within the context of the relational model. However, these properties can also be designed into a network data model database.
4) **Set theory:** As RDBMS is based on mathematical set theory, the usage of relational algebra and relational calculus for the manipulation of the relations/tables ensures that there is no ambiguity, which may otherwise, arise in establishing the linkages in a complicated network type database.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

ADVANCED DATABASES

1) Distributed Databases

Earlier independent or decentralized systems were used for data processing. There was a duplication of hardware because of a lack of management control. In a centralized database system, the database management system and the data reside at a single location, and, therefore, control over data and processing was limited to that particular location. However, organizations usually have geographically scattered operations. The current trends are towards the distributive system. In this a Central Host System is connected to the intelligent remote devices, thus, the distribution of Processing Power creates a feasible environment for data distribution which can be easily accessed from each site.

Distributed Database can be defined as a collection of data with different parts under the control of separate DBMS running on an independent computer system.

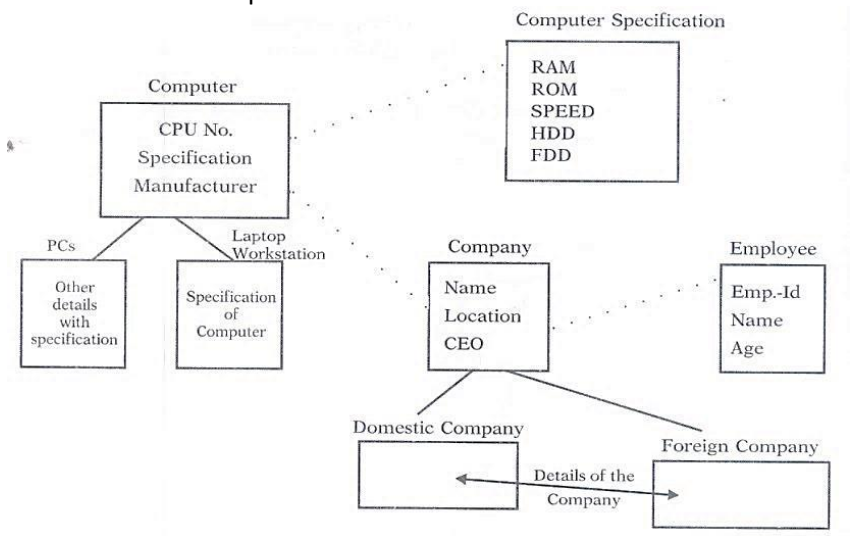
Disadvantages: Primarily distributed systems are costly and complex. But distributed databases are designed in such a way that if a certain failure occurs in the system, even then distributed databases are operational. There could be three types of failures: (a) node failures, (b) loss of message, and (c) communication link failure.

2) Object-Oriented Databases (OODB)

"Object" means "everything that exists in the real world." Thus, in the object-oriented approach "everything is an object". The object-oriented database is a collection of objects, then relationship behavior and other features are defined according to the object-oriented DB model. The need for OODB arose because everything could not be modeled in conventional databases. The conventional database models such as hierarchical network and relational were developed for the conventional business database applications. These are financial accounting, employee information system, banking information system. However, these traditional database models have certain limitations for implementing complex database applications such as databases for engineering design and manufacturing (CAD/CAM), geographic information systems, multimedia, etc. These new applications differ from traditional business applications in terms of their requirement and characteristics, these include mainly storing images, longer duration of transactions, complex structure for objects, etc.

The object-oriented databases, support the needs of these complex database applications. The object-oriented approach offers flexibility to handle these requirements with the support of object-oriented programming languages and database languages. Moreover, the object-oriented concept consists of several data modeling concepts such as aggregation, generalization, and membership relation, that one needs in modeling complex applications. The object-oriented approach supports these because of the concepts such as object, class, inheritance that are in object-oriented languages. The object-oriented data model could incorporate these shortcomings of traditional database models by modeling not only real-world objects but also the relationship among these objects and constraints on them. The use of object-oriented databases is on increasing trend because of the frequent and increasing use of object-oriented programming languages in developing software applications. The object-oriented database management system allows the various users to define and manipulate object-oriented databases.

Example of OBJECT ORIENTED DATABASE



Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

In the above object-oriented database structure, the computer is the root, this includes classes, computer specifications, company, and employee. It is also a root of PCs and laptop, work-stations.
Similarly, a company is a root of a class with sub-classes such as Domestic Company, Foreign Company, and of Employee. "_____" represent class/ sub-class link, "……." represent attribute-domain link. The query to such database may be "generate a list of CEO's and company names in Delhi, manufacturing Laptops".

3) Multimedia Database Systems

Over the years multimedia information systems are dominating our daily life. So the advancement in technology is an added support to it. The multimedia database may be a collection of any of the data types such as text, graphics, images, animations (sequence of images) video (photographic data), audio, etc. Multimedia applications are usually meant for presentation to execute complex design tasks, intelligent healthcare network, knowledge dissemination, education and training, marketing, advertising, retailing, entertainment, travel, etc. As conventional databases were designed to store and manage simple data systems such as strings and numbers (real integer), requirements in multimedia data made traditional databases unsuitable. Thus, the limitation of traditional database systems motivated researchers to build technologies for the multimedia database system.
A multimedia database management system requires an enormous amount of memory or data storage as by nature multimedia data are characterized by huge values. Thus, multimedia database management systems are very complex, because the problems arise in modeling the complex objects made up of a wide range of data types: text, numeric, graphic, animated graphic image, audio stream video sequence. Another problem lies in the retrieval of text data as text retrieval has always been the key feature in business applications. This is because the text information system has low precision as there may be ambiguity in the word meaning. To resolve ambiguity online, the dictionary may be used. Moreover, it is difficult to perform indexing on text data.
There are no DBMSs designed especially for multimedia data management or to support full functionality to multimedia data management. However, several other DBMSs support multimedia data types; these include Universal Database (UDB) of IBM, ORACLE 9.0, etc.

4) Mobile Databases

The new dimension in data communication and processing is mobile computing which has been made possible due to recent advances in wireless technology. The features of mobile computing allow the users to manage their tasks while they are mobile or communicating with other users. Mobile computing is useful in geographically dispersed organizations, for example, stockbroking information, traffic police, reporting
services, financial marketing reporting. Mobile data management has its origin in the distributed database system. However, the problem of managing a large amount of data is difficult because of the narrow bandwidth of wireless communication channels and the power supply battery of mobile units generally have a short active life. These are certain peculiar characteristics in a mobile environment. In the mobile database environment data changes vary frequently, therefore database updation is critical. In such an environment, users are also mobile and randomly enter and exit from the cell and the user generally accesses a certain part of the database very frequently. Thus, it is difficult for DBMS to maintain user's entry, exit, or specific data request information.
But mobile applications offer instantaneous information services to the users, where timeliness of the information is critical, for example, users can have information about doctors in case of emergency or availability of parking at public places i.e., airport. Data used in Mobile Environment may be of three types as given below :
(a) Private data: This data is used and managed by a single user i.e., it cannot be accessed by any other user.
(b) Public data: Can be used by any user for retrieval! reading purposes such as stock prices, weather information. However, this data is updated only by an authorized user.
(c) Shared data: This data may be read and written by several users e.g., inventory data.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Regarding the database management issues, in a mobile environment, these are generally the same as in distributed data management with some additional considerations.

ER (ENTITY RELATIONSHIP MODEL)

The ER model defines the conceptual view of a database. It works around real-world entities and the associations among them. At the view level, the ER model is considered a good option for designing databases.

Entity: An entity can be a real-world object, either animate or inanimate, that can be easily identifiable. For example, in a school database, students, teachers, classes, and courses offered can be considered entities. All these entities have some attributes or properties that give them their identity.

An entity set is a collection of similar types of entities. An entity set may contain entities with attributes sharing similar values. For example, a Student set may contain all the students of a school; likewise, a Teacher set may contain all the teachers of a school from all faculties. Entity sets need not be disjoint.

Attributes: Entities are represented using their properties, called attributes. All attributes have values. For example, a student entity may have a name, class, and age as attributes.

There exists a domain or range of values that can be assigned to attributes. For example, a student's name cannot be a numeric value. It has to be alphabetic. A student's age cannot be negative, etc.

Types of Attributes

- **Simple attribute** − Simple attributes are atomic values, which cannot be divided further. For example, a student's phone number is an atomic value of 10 digits.
- **Composite attribute** − Composite attributes are made of more than one simple attribute. For example, a student's complete name may have first_name and last_name.
- **Derived attribute** − Derived attributes are the attributes that do not exist in the physical database, but their values are derived from other attributes present in the database.
- **Single-value attribute** − single-value attributes contain a single value. For example − Social_Security_Number.
- **Multi-value attribute** − Multi-value attributes may contain more than one value. For example, a person can have more than one phone number, email_address, etc.

Entity-Set and Keys

Key is an attribute or collection of attributes that uniquely identifies an entity among an entity set.

For example, the roll_number of a student makes him/her identifiable among students.

- **Super Key –** A set of attributes (one or more) that collectively identifies an entity in an entity set.
- **Candidate Key –** A minimal super key is called a candidate key. An entity set may have more than one candidate key.
- **Primary Key –** A primary key is one of the candidate keys chosen by the database designer to uniquely identify the entity set.

Relationship: The association among entities is called a relationship. For example, an employee works_at a department, a student enrolls in a course. Here, Works_at and Enrolls are called relationships.

Relationship Set: A set of relationships of similar type is called a relationship set. Like entities, a relationship too can have attributes. These attributes are called descriptive attributes.

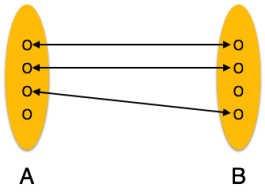
Degree of Relationship: The number of participating entities in a relationship defines the degree of the relationship.

- Binary = degree 2
- Ternary = degree 3
- n-ary = degree

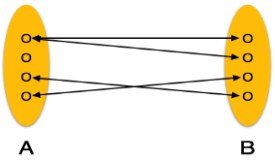
Mapping Cardinalities: Cardinality defines the number of entities in one entity set, which can be associated with the number of entities of other sets via a relationship set.

- **One-to-one –** One entity from entity set A can be associated with at most one entity of entity set B and vice versa.

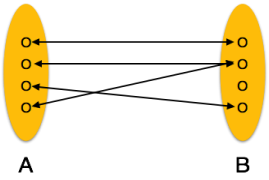
Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar



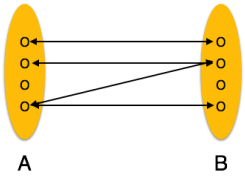
- **One-to-many –** One entity from entity set A can be associated with more than one entity of entity set B however an entity from entity set B, can be associated with at most one entity.



- **Many-to-one** – More than one entity from entity set A can be associated with at most one entity of entity set B, however an entity from entity set B can be associated with more than one entity from entity set A.

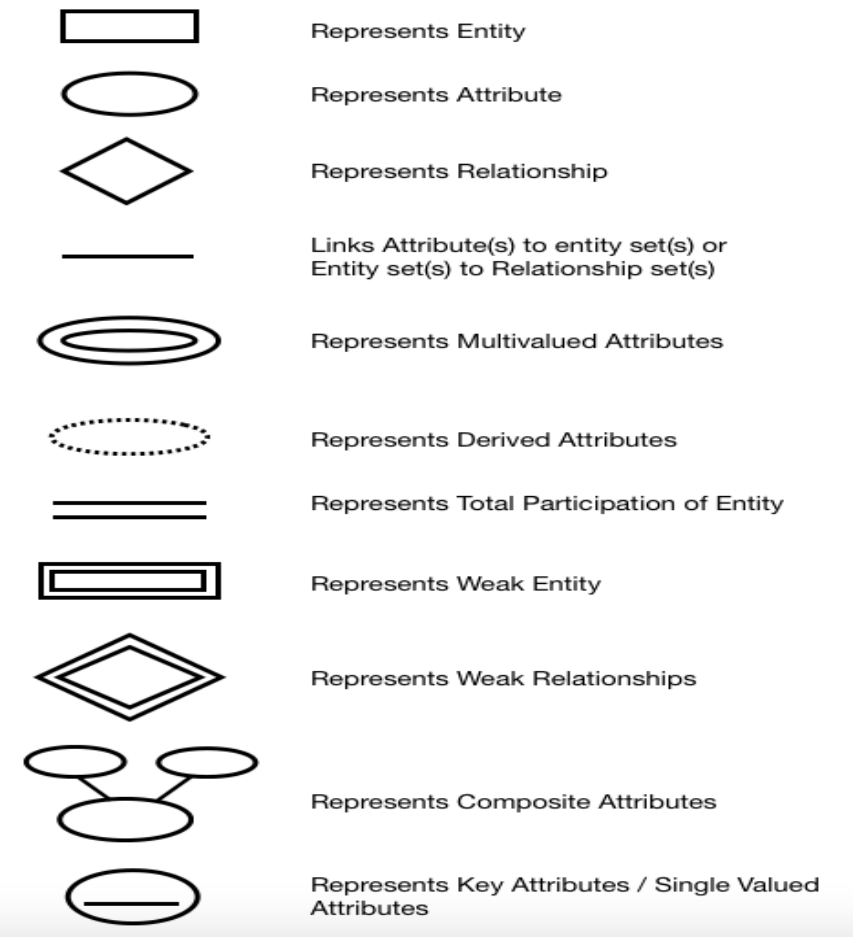


- **Many-to-many –** One entity from A can be associated with more than one entity from B and vice versa.



NOTATION FOR ER DIAGRAM

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar



CHOOSING A DATABASE MANAGEMENT SYSTEM

The process of selecting, evaluating, and finally acquiring a DBMS package requires substantial time and effort. The tasks begin when users strongly feel the need and subject its requirement for an organization. Designing such a package in-house is not a realistic alternative when reasonably good and effective commercial systems are available in the market.

The important criteria for the selection of DBMS are technical and administrative. The key technical criteria relate to the type of system required, balancing the competing objectives of efficiency and functionality. Administrative criteria include vendor characteristics, maintenance support, documentation, training, and ease of learning and use, cost, etc.

An organization usually lives with chosen DBMS for several years. If the initial study and selection is done with a broad view of the organizational needs of today and keeping in mind the need of the future, the choice can enhance data processing responsiveness to user needs, managerial decision effectiveness, and organization profitability.

DATABASE LIFE CYCLE

Database application development is the process of obtaining real-world requirements, analyzing requirements, designing the data and functions of the system, and then implementing the operations in the system.

1. **Requirements Gathering:** The first step is requirements gathering. During this step, the database designers have to interview the customers (database users) to understand the proposed system and obtain and document the data and functional requirements. The result of this step is a document that includes the detailed requirements provided by the users.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

2. **Analysis:** Data analysis begins with the statement of data requirements and then produces a conceptual data model. The analysis aims to obtain a detailed description of the data that will suit user requirements so that both high and low-level properties of data and their use are dealt with. These include properties such as the possible range of values that can be permitted for attributes (e.g., in the school database example, the student course code, course title, and credit points).
3. **Logical Design:** Database design starts with a conceptual data model and produces a specification of a logical schema; this will determine the specific type of database system (network, relational, object-oriented) that is required. The relational representation is still independent of any specific DBMS; it is another conceptual data model.
4. **Implementation:** Implementation involves the construction of a database according to the specification of a logical schema. This will include the specification of an appropriate storage schema, security enforcement, external schema, and so on. Implementation is heavily influenced by the choice of available DBMSs, database tools, and operating environments. There are additional tasks beyond simply creating a database schema and implementing the constraints – data must be entered into the tables, issues relating to the users and user processes need to be addressed, and the management activities associated with wider aspects of corporate data management need to be supported. In keeping with the DBMS approach, we want as many of these concerns as possible to be addressed within the DBMS. We look at some of these concerns briefly now.
5. **Realizing the Design:** After the logical design has been created, we need our database to be created according to the definitions we have produced. For an implementation with a relational DBMS, this will probably involve the use of SQL to create tables and constraints that satisfy the logical schema description and the choice of appropriate storage schema (if the DBMS permits that level of control).
6. **Populating the Database:** After a database has been created, there are two ways of populating the tables – either from existing data or through the use of the user applications developed for the database. For some tables, there may be existing data from another database or data files.

NORMALIZATION

If a database design is not perfect, it may contain anomalies, which are like a bad dream for any database administrator. Managing a database with anomalies is next to impossible.

• **Update anomalies –** If data items are scattered and are not linked to each other properly, then it could lead to strange situations. For example, when we try to update one data item having its copies scattered over several places, a few instances get updated properly while a few others are left with old values. Such instances leave the database in an inconsistent state.

• **Deletion anomalies –** We tried to delete a record, but parts of it were left undeleted because of unawareness, the data is also saved somewhere else.

• **Insert anomalies –** We tried to insert data in a record that does not exist at all.

Normalization is a method to remove all these anomalies and bring the database to a consistent state.

FIRST NORMAL FORM

First Normal Form is defined in the definition of relations (tables) itself. This rule defines that all the attributes in a relation must have atomic domains. The values in an atomic domain are indivisible units.

Course	Content
Programming	Java, c++
Web	HTML, PHP, ASP

We re-arrange the relation (table) as below, to convert it to First Normal Form.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Course	Content
Programming	Java
Programming	c++
Web	HTML
Web	PHP
Web	ASP

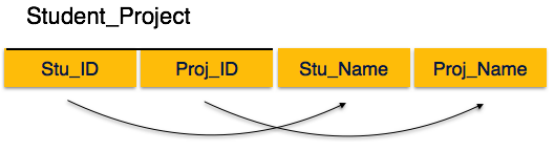
Each attribute must contain only a single value from its pre-defined domain.

SECOND NORMAL FORM

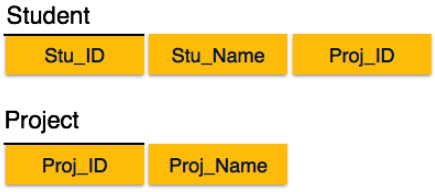
Before we learn about the second normal form, we need to understand the following −

- **Prime attribute –** An attribute, which is a part of the candidate-key, is known as a prime attribute.
- **Non-prime attribute –** An attribute, which is not a part of the prime-key, is said to be a non-prime attribute.

If we follow the second normal form, then every non-prime attribute should be fully functionally dependent on the prime key attribute. That is, if X → A holds, then there should not be any proper subset Y of X, for which Y → A also holds true.



We see here in Student_Project relation that the prime key attributes are Stu_ID and Proj_ID. According to the rule, non-key attributes, i.e. Stu_Name and Proj_Name must be dependent upon both and not on any of the prime key attributes individually. But we find that Stu_Name can be identified by Stu_ID and Proj_Name can be identified by Proj_ID independently. This is called partial dependency, which is not allowed in the Second Normal Form.



We broke the relation in two as depicted in the above picture. So there exists no partial dependency.

THIRD NORMAL FORM

For a relation to be in Third Normal Form, it must be in Second Normal form and the following must satisfy –

- No non-prime attribute is transitively dependent on the prime key attribute.
- For any non-trivial functional dependency, X → A, then either −

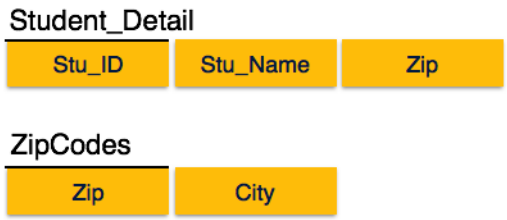
X is a superkey or,

A is a prime attribute.

We find that in the above Student_detail relation, Stu_ID is the key and only prime key attribute. We find that City can be identified by Stu_ID as well as Zip itself. Neither Zip is a superkey nor is City a prime attribute. Additionally, Stu_ID → Zip → City, so there exists transitive dependency.

To bring this relation into the third normal form, we break the relation into two relations as follows −

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar



Boyce-Codd Normal Form

Boyce-Codd Normal Form (BCNF) is an extension of the Third Normal Form on strict terms. BCNF states that

• For any non-trivial functional dependency, X → A, X must be a super-key.

In the above image, Stu_ID is the super-key in the relation Student_Detail, and Zip is the super-key in the relation ZipCodes. So,

Stu_ID → Stu_Name, Zip

and

Zip → City

This confirms that both the relations are in BCNF.

DBMS USERS

1. Naive Users: Naive users are end-users of the database who work through a menu-driven application program where the type and range of response are always indicated to the user.

2. Online Users: Online users are those who may communicate with the database directly via an online terminal or indirectly via a user interface and application program.

3. Application Programmers: Professional programmers are those who are responsible for developing application programs or user interfaces.

4. DBA: One of the main reasons for using DBMS is to have central control over both the data and the programs that access those data. The person who has such control over the system is called the Database Administrator (DBA).

The **functional of DBA** include the following:

(1) Schema Definition: The DBA create an original Database schema by writing a set of definitions.

(2) Storage Structure and Method Definition: The DBA creates an appropriate storage structure and access method by writing a set of definitions.

(3) Granting of Authorization for Data Access: The authorization information is kept in a special system structure.

(4) Integrity Constraint Specification: The number of hours an employee may work in one week may not exceed a specified limit. Such a constraint must be specified explicitly by the Database Administration (DBA).

DATABSASE ADMINISTRATOR

A database administrator (DBA) is a person who is responsible for the environmental aspects of a database. The role of a database administrator has changed according to the technology of database management systems (DBMSs) as well as the needs of the owners of the databases. For example, although logical and physical database design is traditionally the duties of a database analyst or database designer, a DBA may be tasked to perform those duties.

ROLE OF DATABASE ADMINISTRATOR

Some of the roles of the DBA may include:

Installation of new software - It is primarily the job of the DBA to install new versions of DBMS software, application software, and other software related to DBMS administration. The DBA or other must be staff members test this new software before it is moved into a production environment.

Configuration of hardware and software with the system administrator - In many cases, the system software can only be accessed by the system administrator. In this case, the DBA must work closely with

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

the system administrator to perform software installations, and to configure hardware and software so that it functions optimally with the DBMS.

Security administration - One of the main duties of the DBA is to monitor and administer DBMS security. This involves adding and removing users, administering quotas, auditing, and checking for security problems.

Data analysis - The DBA will frequently be called on to analyze the data stored in the database and to make recommendations relating to the performance and efficiency of that data storage. This might relate to the more effective use of indexes, enabling "Parallel Query" execution, or other DBMS-specific features.

Database design (preliminary) - The DBA is often involved at the preliminary database design stages. Through the involvement of the DBA, many problems that might occur can be eliminated. The DBA knows the DBMS and system, can point out potential problems and can help the development team with special performance considerations.

Data modeling and optimization - By modeling the data, it is possible to optimize the system layouts to take the most advantage of the VO subsystem.

- Responsible for the administration of existing enterprise databases and the analysis, design, and creation of new databases.
- Data modeling, database optimization, understanding and implementation of schemas, and the ability to interpret and write complex Structured Query Language (SQL) queries
- Proactively monitor systems for optimum performance and capacity constraints
- Establish standards and best practices for SQL
- Interact with and coach developers in SQL scripting

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PREVIOUS YEAR QUESTIONS

Short Answer Type Questions

1. What is data abstraction?
2. Differentiate between data & information.
3. What are the types of keys that are used in the relational model?
4. Oracle is a _____ DBMS.

Focused-short Answer type Questions

1. Discuss the concept of data independence and explain its importance in a database environment.
2. What are different languages and their purpose that are associated with DBMS?
3. Explain the roles, duties, and responsibilities of the different users of a database.
4. What is a Relational Database?
5. Which language is used to define the structure of the relation, deleting relations and relating schemas?
6. Explain the difference between composite data types and derived data types in the database with examples.

Long Answer Type Questions

1. Explain different components of a Database System.
2. What are various database models? Describe the features of the relational database model.
3. Explain different types of Database Languages.
4. Describe the advantages and limitations of the Database.
5. What do you mean by data model? Compare and contrast different data models.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Biju Patnaik Institute of IT & Management Studies

Lecture Notes

on

IT FOR MANAGEMENT

Module-IV

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMPUTERS NETWORKS COMPONENTS

Computer Network involves a group of computers and associated devices that are connected by communications facilities. Telecommunication systems can transmit digitally coded information or data including text, voice, pictures, maybe in motion. Communication is the process of transferring information from a source to a destination through some media.

In computer networking architecture, compatible software and hardware are arranged in a manner, so that information can be communicated from one location to another. A networking system consists of five basic elements arranged in the following manner and follows the process of communication in that order, these are:

- The originating/source computer that transmits data. Data can be in the form of files on the disk or entered via a keyboard.
- A data communication device is attached to the sending computer. It converts the data in the form it can be transmitted.
- A communication channel, which carries the data from point to point. This can be a telephone line or microwave relay system.
- Data communication device attached to the receiving system, capable of converting the data in the form which receiving system can read.
- The receiving computer, which receives the information and displays it on the screen, prints them and stores them in a file.

OBJECTIVES OF NETWORKING

There are various types of networks that share a common set of objectives. They include:

a) To provide connectivity between various hardware and software products, the connected component can communicate seamlessly.
b) To permit simple and easy installation of the software and operation of all network components.
c) To enable the modular building of a wide variety of network devices to reduce the cost.
d) To allow the network to grow in all dimensions, as and when needed.
e) To permit reliable and error-free transmission by providing appropriate error-correcting and detecting capabilities.
f) The availability of the network is fully operational 24 hours/day, 7 days/week.
g) Flexibility to permits the availability of the new technologies in the network, as new needs evolve.
h) To provide services to the user, that can be easily used, yet isolate users from the technical details of network structure and implementation.
i) Manageability to detect and isolate problems, and to take appropriate corrective action.

ADVANTAGES OF NETWORKING

Computer networks have many advantages. Some of them are listed below:

Sharing of Hardware: Networking allows the user to share the hardware in departments of a company/Institution. The company often wants expensive peripheral devices, only if, they are shared by several computers. For example, an expensive printer may not be affordable, if it is used by a single department. But, if it is shared among various departments, then it comes out to be very economical and hence affordable.

Sharing of Information: It allows the sharing of information, which may be used constantly throughout the company. For example, the rate of any product of a company is used by all the offices in different parts of the country. In such cases, the company can have only one copy instead of multiple copies and can be shared by different branches connected to a network.

Transfer of Text: Networking facilitates the electronic transfer of text. Organizations, always transfer data from one place to another.

Decentralization of Data Processing: A complicated job can be divided into multiple modules and can be distributed to different departments for processing.

Easy Communication: Various organizations cooperating in a task can communicate with each other on a network and exchange information.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMMUNICATION INTERFACE DEVICES

(i) Network Interface Cards (NIC) - Network Interface Card allows the transmission of data over a cable network which connects various computer and other devices such as printers. NIC has additional memory for buffering, incoming and outgoing data packets, thus it increases the network throughput. Since it is an expansion card on the mother-board, thus, circuit on the board provides protocols and commands required to support this kind of network card. NICs are available in 8-bit or in 16-bit bus standards.

(ii) Modem stands for Modulator/demodulator. Thus, the modem refers to a device that performs two processes modulation i.e. conversion from digital to analog and demodulation i.e. reverse the conversion from analog to digital. It is a communication device that converts digital data from a computer of terminal and analog audio signals that can pass through standard telephone lines. The modem is a device that carries voice or sound signals like a telephone system. To carry digital information, that information must be converted into an analog pattern. Modems transmission speed is measured in bits per second. Typically, modems' speed ranges from 14,000 bps to 56,000 bps.

(iii) Multiplexer is an electronic device that divides a physical channel into a logical channel to carry data transmission simultaneously from many sources. Multiplexing makes optimal use of the existing communication channel by merging the transmission of several sources at one end of the channel. Similarly, a multiplexer is used to separate the single transmission into multiple transmissions at the receiving end. The multiplexer is a device that merges several low-speed transmissions into one high-speed transmission.

(iv) Front-end Communication Processor: The most sophisticated communication management device is the front-end processor. In communications, a front-end processor is a small computer that is connected to the main (host) or larger computer and is used to transmit and receive messages over the communication channels, detect and correct errors, encoding of messages, and relieve the host computer from routine computational tasks.

(v) Protocol converter - Protocol is a set of conventions/rules which govern the exchange of data between sender and receiver (computer) with as little error as possible. Dissimilar devices cannot communicate with each other.

(vi) Connecting Devices: Routers, Bridges, Gateways, Repeaters, and Switches:

Repeaters: A repeater is a device used on communication circuits that decreases distortion by amplifying or regenerating a signal so that it can be transmitted in its original form. A repeater generally connects two networks at the physical layer.

Bridges: It is a device that connects the same type of networks or LANs using the same communication protocols so that information can be passed from one to another. To transmit this data successfully the bridge magnifies the data transmission signal. This means that bridge can act as a repeater as well as a link.

Gateway: It is a device that connects dissimilar networks using different communication protocols so that information can flow from one to the other. A Gateway performs two functions, it transfers the information and converts information to a form compatible with the protocol used by the receiving network.

Router: A router is an intermediary device on a communication network that directs the communicating messages when several networks are connected. A router receives transmitted messages and forwards them to their correct destinations utilizing the most efficient available route. High-speed routers can serve as part of the Internet backbone or transmission path, handling the major data traffic.

The major difference between routers, gateways, and bridges is that gateway is comprised of hardware and software that enables communication among dissimilar networks whereas bridge can connect only those networks that employ similar protocols and topologies. Similar to the bridge, routers also connect similar networks.

Hub: In a network, a hub is a device joining communication lines at a central location, providing a common connection to all devices on the network. In this, each node is connected to the hub through simple twisted-pair wires.

Switch: In networking, a switch is a device capable of forwarding packets directly to the ports associated with a particular network address.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

ELECTRONIC SIGNALS AND TYPE OF COMMUNICATION

Data Communication refers to communication between the computers linked directly with the computer locally or worldwide through a telecommunication network. Data Communication can be classified on the type of electronic signals and type of communication.

Electronic Signals

Telecommunication channels carry two types of signals:

Analog signals are continuous, they carry information in the form of waves e.g., the way sound travels in the medium such as telephone lines. Analog communication uses general-purpose communication channels. These signals are characterized by two parameters like amplitude and frequency.

Digital signals are discrete. They transfer information in the form of 0's and 1's. In Digital communication, data are transmitted directly in binary form that is a sequence of 0's and 1's. To transmit binary data over these channels 0's and 1's should be converted to electrical signals. This is done by Modem (modulator-demodulator). Digital signals are preferred over analog signals, because of the discrete nature, these signals are not affected by noise or any other disturbances. Moreover, a computer at the receiving end does not perform any conversion as in analog signals i.e. analog to digital.

Serial and Parallel Transmission

Serial Transmission involves sending information over a single line one bit at a time, as discrete signals, one after another.

Parallel Transmission refers to the transmission of a group of bits (8 bits with PCs) over separate wires,

Generally, data are organized into 8-bit bytes. These bytes are further organized to form multi-byte words. In parallel transmission, these bytes are transferred using multiple wires, each wire carrying a byte. For example, it can be seen when data is transferred from CPU to printer through the LPT port. In serial transmission, these bytes are sequentially transferred one after the other. As discussed in Chapter 8 serial port RS-232 is used for data transmission between CPU and mouse or modem.

Advantages of serial transmission are:-

(i) It is an inexpensive mode of transferring data.

(ii) It is preferred for transmitting data over a long distance.

Its limitation is that in serial transmission; data transfer is significantly slow because it transfers data in series.

COMMUNICATION MEDIA (CHANNELS)

Transmission media are electronic roadways through which data travels from one place to other. Earlier, only telephone lines were used for data communication. But now there are several other communication channels available for transmission. The user needs to understand the cost and other basic characteristics such as data transmission rate, reliability, security of different communication channels, one of the most important characteristics is **bandwidth.**

Bandwidth refers to data transfer capacity or speed of transmission of a digital communication system measured in bps (bits per second).

Thus, to transmit more information in less time, higher bandwidth is preferable. The communication channel can be categorized as wired and wireless media as follows:

Wired Media - This comes in a variety of different types as given below:

- **Twisted Pair:** This is the most commonly used form of communication wire. It is generally used for almost all telephone wiring. The purpose of this twisting is to reduce electrical inference. Each wire consists of a copper core covered in a plastic sheet. This copper core is responsible for data transmission and the plastic sheet is for protection. Twisted pair has a disadvantage as it emits electromagnetic interference, thus it transmits data relatively at a slow speed, it can be easily tapped by an unauthorized user.
- **Coaxial Cable:** A coaxial cable consists of many small cables in a protective/insulated cover, thus is less susceptible to tapping as well as electromagnetic interference. Therefore, it carries much more data than twisted pairs. Coaxial cables are expensive and relatively inflexible. Also, as and when

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

equipment needs to be shifted, it requires rewiring. Because of its inflexibility, the cost of installation is increased.

- **Optical Fibre Transmission:** These are the most efficient and speedy communication channels. Optical fiber transmits information through glass fiber filament along which data travels at high speed as waves of light. They are less susceptible to electromagnetic interference and do not generate heat, as they are optical. Fiber optic cable, which has the highest bandwidth, provides higher speed and greater data-carrying capacity. This mode of transmission is more secured from interference and tapping. This cable is expensive but most efficient amongst all.

Wireless Channel

Wire media possess some problems such as cost of installation, limited capacity of transmission of information, subject to electromagnetic interference, etc. although fiber optic cables are an exception. Thus, the other option exists in wireless communication.

Microwave Transmission - Like radio or television signals, microwave signals are used for transmitting data without the use of cables. Microwave systems are generally used for high volume, long-distance, point-to-point communication. Such transmissions were used to transmit high-frequency radio signals in a line-of-sight path. But the transmission is limited to 30 miles. This is because the relay stations are spaced approximately 30 miles. To remove the line-of-sight problem, microwave antennas are usually placed on the peaks, like the top of the buildings or mountains through which microwave signals are transmitted. It is a high-speed data transmission media. It uses the upper atmosphere as a reflective surface, which subjects the microwave transmission to environmental interference during bad weather conditions i.e. heavy rain, snowstorms. It is widely used for broadband communication and telephone services.

- **Infrared Transmission -** Infrared light is a red light that carries information and can be modulated. This beam of light is invisible to the naked eye. The most common application of infrared light is with television or videocassette recorder remote control units. They are generally used for short-distance connections. The advantage of using this is that it does not require any wiring and electrical interference. Moreover, it is unaffected by artificial light and weather conditions.
- **Radio Transmission -** In this radio beam is used for transmission. Radio is used to connect computers and peripheral equipment may be in LAN to transmit and receive data at both ends. Data can be transmitted in a specified frequency range only. There is no security in such communication links. FM is the biggest example of such a transmission. Radio waves can be propagated easily. The devices used are inexpensive and easy to install.
- **Laser Transmission -** Laser beams, which are a very low-powered narrowly focussed beam of light, are used for communication. This light is invisible to the naked eye. Special laser transmitting and receiving equipment are needed at both ends. Repeaters can be used to boot up the signals for long distances.
- **Satellite Transmission -** The major advancement in telecommunication technology is the use of satellites for digital transmission. These satellites act as communication stations. The satellite receives data signals from one earth station, amplifies them, and sends them to the other earth station. This entire process takes place in one step. The satellites have multiple independent receptions and transmission devices known as trans-receivers. In these types of communications, the cost of transmission remains the same regardless of the distances and the number of stations receiving the transmission simultaneously. They can carry a large amount of data. Many factors limit the use of satellites. First of all, it is a very expensive affair unless the organization's need is huge. Secondly, the satellite has to be synchronized with the speed of the earth as most of the satellites are placed at the geo-synchronous orbit above the equator. Signal delays are caused by the extreme length of the transmission path between the sender or the receiver. This can be annoying in the case of voice transmission.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Advantages and Disadvantages of Communication Media

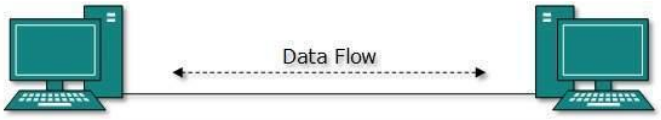
A comparative study has been expressed in the table given below:

Communication Channel/Media	Advantages	Disadvantages
Twisted pair	- Cheaper - Easily available, easy to work with	- Electromagnetic interference is emitted. - Low Bandwidth, slows transmission. - Not very secure media as it can be easily tapped.
Coaxial cable	- Higher bandwidth than twisted pair - Being insulated, thus less susceptible to electromagnetic interference	- More expensive and relatively inflexible - Generally difficult to work with - Easily tapped
Fibre Optic Cable	- Very high Bandwidth higher than coaxial cable - Difficult to be tapped	- Expensive medium - Relatively inflexible
Microwave	- Very high bandwidth - Relatively cheap	- Essentially, line of sight should be unobstructed - Weather condition (i.e., rain) may cause interference - Encryption should be used while transmitting confidential data.
Satellite	- Very high bandwidth - Cover large geographical region on earth	- Very costly - Signal propagation may be delayed. - Encryption of confidential information is essential

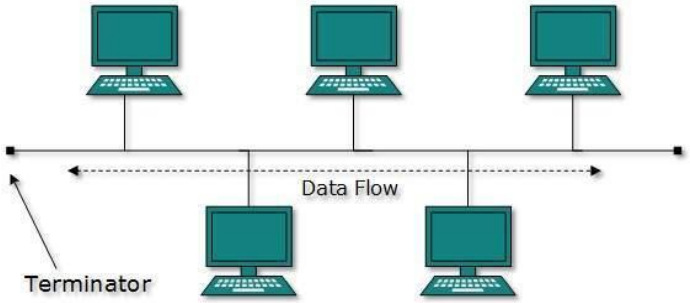
NETWORK TOPOLOGIES

A Network Topology is the arrangement with which computer systems or network devices are connected. Topologies may define both physical and logical aspects of the network. Both logical and physical topologies could be the same or different in the same network. Topology is the layout of the connection between the computers. Network topology generally, determines the reliability and efficiency of computer networks. Depending upon different topologies, the network can be divided into the following types.

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



2. Bus Topology (Linear bus): In a bus structure of the network, there is a single cable and all the workstations are attached to it. Its structure is like the Spinal Cord of the human body. In this structure, all network nodes have unique addresses and are connected to a common transmission medium that is called a broadcasting bus. When a device transmits data onto the bus, it is received by all devices and is ignored, except by the one that is addressed. Bus networks refer to topology for a LAN in which all nodes are connected to the main connection line (bus), on a bus network each node monitors activity on the line, messages are detected by all nodes but are accepted only by the nodes to which they are addressed.



Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Advantages of bus topology are:

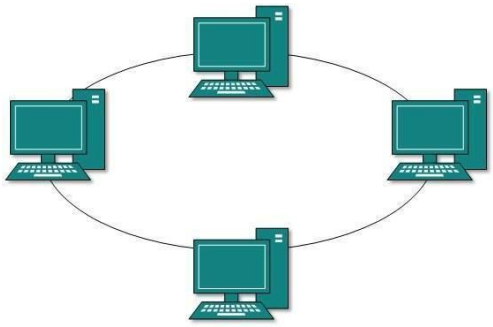
(a) It is easy to use and understand, reliability is higher in very small networks.
(b) It is cheaper than other topologies because it requires less cable to connect the computers than other cabling arrangements.
(c) It is easy to extend and more computers can be connected to the network as cables can be joined with a connector.
(d) A repeater can also be used to extend bus configuration.
(e) Bus networks commonly rely on collision detection or token passing to regulate traffic.

Disadvantages of bus topology are:

(a) Bus network works slowly in heavy traffic because any computer can transmit at any time and can interrupt each other by using a lot of bandwidth.
(b) Each connection between two cables weakens the electrical signals.
(c) A network is disabled if someone accidentally cuts the main cable.

3. Ring Topology (Ring connection)

Ring topology is similar to the bus, as it is also having a single cable, but the cable (media) forms a complete loop. In this type of network, the computer communicates with others by sending a signal around the ring and each workstation takes an active part in the transmission. Information is passed from node to node around the ring until it arrives at the node that is addressed. The ring network is a LAN in which all the workstations along with the servers are connected in the form of a ring, in which, consecutive nodes are connected by point-to-point, links arranged to form a closed path (ring). As the message proceeds through the ring, each node checks the address, if it is for that particular node. If not, the message goes to the next computer. This process goes on till the message reaches the destination. Token ring networks use this topology.



The advantages of the Ring network are following:

(a) Ring network offers reliable performance for small and larger distances. This is so because when a node receives a message it examines the destination address attached to the message. If the address is the same as the nodes, the node accepts the message, otherwise, it regenerates the signal and passes the message along to the next node in the ring. Such regeneration allows ring networks to cover larger distances than star or bus networks.
(b) It can be easily extended.
(c) A ring topology makes it easy for computers to co-ordinate access and to detect whether the network is operating correctly.

Disadvantages of ring network :

(a) Compared to other networks it is difficult to install and relatively expensive.
(b) The entire network is disabled if one of the cables is cut.
(c) Because of the closed-loop, adding a new node can be difficult and can disrupt the network.
(d) It is difficult to maintain and troubleshoot a network. Often extensive wiring has to be done to maintain the ring structure when a computer is added to or removed from the network.
(e) Ring networks are highly vulnerable, if a single computer fails, then at least one part of the network is not working.

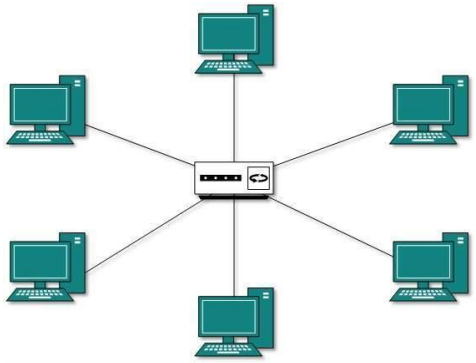
Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

4. Star Topology (Star connection)

In some networks when communication goes through a center point, then star topology is preferred. A star-shaped network resembles the spoke of a wheel, the center of a star network is often called a hub. The wiring between the nodes is not shared as in the case of bus topology. The computer in the center acts as a server. This topology is used in applications where a central computer communicates with remote terminals or workstations. The central node can be active or passive. If it is active, it is usually used to control the entire network and performs all the routing. When two computers of the network communicate with each other, the message must pass through this central server, which retransmits it to the destination. Also in a star network, when communication takes place between two nodes (workstation), a complete circuit is dedicated to the connection for the duration of the call.

The biggest disadvantage of a star network is that if the central server fails, the whole network doesn't work. Secondly, it is costlier than other topologies.

A Star network is a LAN in which each node is connected to a central computer in a star-shaped configuration.

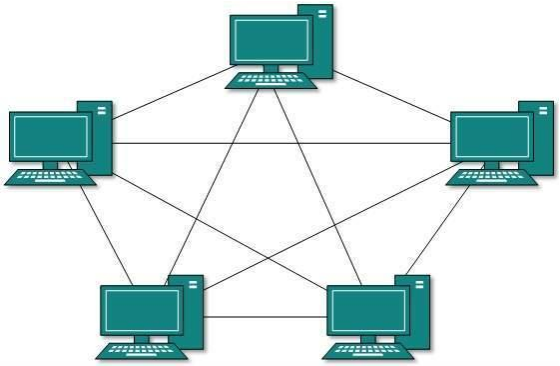


Advantages of a star network are :

(a) It is easy to add and remove nodes
(b) It is easy to troubleshoot the network through a central hub.
(c) If a node fails, it does not affect the network.
(d) A star topology helps protect the network from damage to a single table.

5. Mesh Topology

In this type of topology, a host is connected to one or multiple hosts. This topology has hosted in point-to-point connection with every other host or may also have hosts which are in point-to-point connection to few hosts only.



Hosts in Mesh topology also work as a relay for other hosts which do not have direct point-to-point links. Mesh technology comes into two types:

Full Mesh: All hosts have a point-to-point connection to every other host in the network. Thus for every new host n(n-1)/2 connections are required. It provides the most reliable network structure among all network topologies.

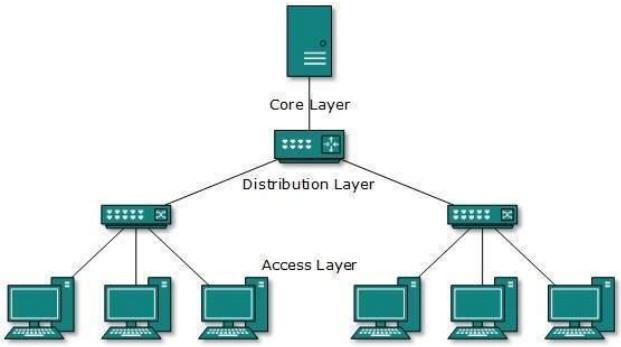
Partially Mesh: Not all hosts have a point-to-point connection to every other host. Hosts connect in some arbitrary fashion. This topology exists where we need to provide reliability to some hosts out of all.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

6. Tree Topology

Also known as Hierarchical Topology, this is the most common form of network topology in use presently. This topology imitates as extended Star topology and inherits properties of bus topology.

This topology divides the network into multiple levels/layers of the network. Mainly in LANs, a network is bifurcated into three types of network devices. The lowermost is the access layer where computers are attached. The middle layer is known as the distribution layer, which works as a mediator between the upper layer and lower layer. The highest layer is known as the core layer and is the central point of the network, i.e. root of the tree from which all nodes fork.



All neighboring hosts have a point-to-point connection between them. Similar to the Bus topology, if the root goes down, then the entire network suffers even.though it is not the single point of failure. Every connection serves as a point of failure, the failing of which divides the network into an unreachable segment.

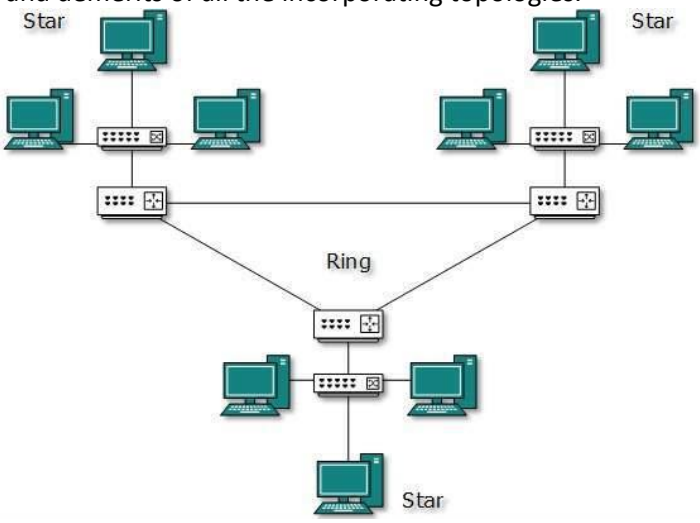
7. Daisy Chain

This topology linearly connects all the hosts. Similar to a Ring topology, all hosts are connected to two hosts only, except the end hosts. This means, if the end hosts in the daisy chain are connected then it represents Ring topology.

Each link in the daisy chain topology represents a single point of failure. Every link failure splits the network into two segments. Every intermediate host works as a relay for its immediate hosts.

8. Hybrid Topology

A network structure whose design contains more than one topology is said to be a hybrid topology. Hybrid topology inherits merits and demerits of all the incorporating topologies.



The above picture represents an arbitrarily hybrid topology. The combining topologies may contain attributes of Star, Ring, Bus, and Daisy-chain topologies. Most WANs are connected through Dual-Ring topology and networks connected to them are mostly Star topology networks. Internet is the best example of the largest Hybrid topology

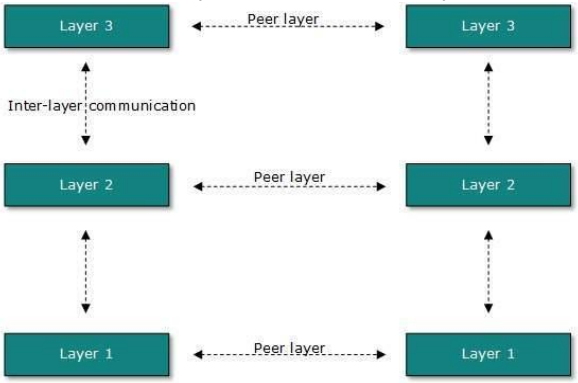
Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMPUTER NETWORK MODELS

The whole networking concept is divided into multiple layers. Each layer is involved in some particular task and is independent of all other layers. But as a whole, almost all networking tasks depend on all of these layers. Layers share data between them and they depend on each other only to take input and send output.

LAYERED TASKS

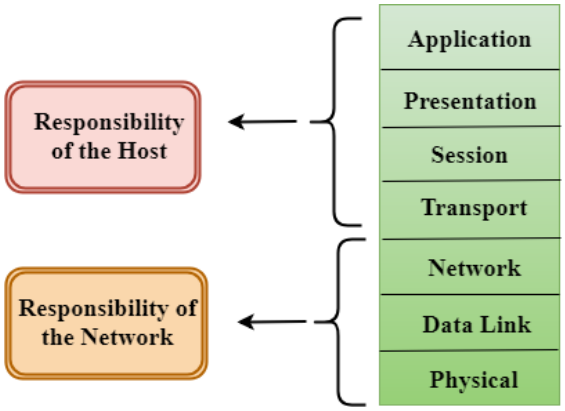
In the layered architecture of the Network Model, one whole network process is divided into small tasks. Each small task is then assigned to a particular layer that works dedicatedly to process the task only. Every layer does only specific work. In a layered communication system, one layer of a host deals with the task done by or to be done by its peer layer at the same level on the remote host. The task is either initiated by layer at the lowest level or the topmost level. If the task is initiated by the-top most layer, it is passed on to the layer below it for further processing. The lower layer does the same thing, it processes the task and passes it on to the lower layer. If the task is initiated by the lower most layer, then the reverse path is taken.



Every layer clubs together all procedures, protocols, and methods which it requires to execute its piece of task. All layers identify their counterparts using encapsulation header and tail.

OSI MODEL

OSI stands for Open System Interconnection is a reference model. OSI consists of seven layers, and each layer performs a particular network function. OSI model was developed by the International Organization for Standardization (ISO) in 1984, and it is now considered as an architectural model for inter-computer communications.OSI model divides the whole task into seven smaller and manageable tasks. Each layer is assigned a particular task. Each layer is self-contained so that tasks assigned to each layer can be performed independently.



Characteristics of OSI Model:

- The OSI model is divided into two layers: upper layers and lower layers.
- The upper layer of the OSI model mainly deals with application-related issues, and they are implemented only in the software. The application layer is closest to the end-user. Both the end-user and the application layer interact with the software applications. An upper layer refers to the layer just above another layer.
- The lower layer of the OSI model deals with the data transport issues. The data link layer and the physical layer are implemented in hardware and software. The physical layer is the lowest layer of the OSI model and is closest to the physical medium. The physical layer is mainly responsible for placing the information on the physical medium.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Functions of the OSI Layers

1. Physical layer

- The main functionality of the physical layer is to transmit the individual bits from one node to another node.
- It is the lowest layer of the OSI model.
- It establishes, maintains, and deactivates the physical connection.
- It specifies the mechanical, electrical, and procedural network interface specifications.

Functions of a Physical layer:

- **Line Configuration:** It defines the way how two or more devices can be connected physically.
- **Data Transmission:** It defines the transmission mode whether it is a simplex, half-duplex, or full-duplex mode between the two devices on the network.
- **Topology:** It defines the way how network devices are arranged.
- **Signals:** It determines the type of signal used for transmitting the information.

2. Data-Link Layer

- This layer is responsible for the error-free transfer of data frames.
- It defines the format of the data on the network.
- It provides reliable and efficient communication between two or more devices.
- It is mainly responsible for the unique identification of each device that resides on a local network.
- It contains two sub-layers:
    - Logical Link Control Layer
    - Media Access Control Layer

Functions of the Data-link layer

- **Framing:** The data link layer translates the physical's raw bitstream into packets known as Frames. The Data link layer adds the header and trailer to the frame. The header which is added to the frame contains the hardware destination and source address.
- **Physical Addressing:** The Data link layer adds a header to the frame that contains a destination address. The frame is transmitted to the destination address mentioned in the header.
- **Flow Control:** Flow control is the main functionality of the Data-link layer. It is the technique through which the constant data rate is maintained on both sides so that no data get corrupted. It ensures that the transmitting station such as a server with a higher processing speed does not exceed the receiving station, with a lower processing speed.
- **Error Control:** Error control is achieved by adding a calculated value CRC (Cyclic Redundancy Check) that is placed to the Data link layer's trailer which is added to the message frame before it is sent to the physical layer. If any error seems to occur, then the receiver sends the acknowledgment for the retransmission of the corrupted frames.
- **Access Control:** When two or more devices are connected to the same communication channel, then the data link layer protocols are used to determine which device has control over the link at a given time.

3. Network Layer

- It is layer 3 that manages device addressing, tracks the location of devices on the network.
- It determines the best path to move data from source to destination based on the network conditions, the priority of service, and other factors. The Data link layer is responsible for routing and forwarding the packets.
- Routers are the layer 3 devices, they are specified in this layer and used to provide the routing services within an internetwork.
- The protocols used to route the network traffic are known as Network layer protocols. Examples of protocols are IP and Ipv6.

Functions of Network Layer:

- **Internetworking:** An internetworking is the main responsibility of the network layer. It provides a logical connection between different devices.
- **Addressing:** A Network layer adds the source and destination address to the header of the frame. Addressing is used to identify the device on the internet.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- **Routing:** Routing is the major component of the network layer, and it determines the best optimal path out of the multiple paths from source to destination.
- **Packetizing:** A Network Layer receives the packets from the upper layer and converts them into packets. This process is known as Packetizing. It is achieved by internet protocol (IP).

4. Transport Layer

- The Transport layer is a Layer 4 ensures that messages are transmitted in the order in which they are sent and there is no duplication of data.
- The main responsibility of the transport layer is to transfer the data completely.
- It receives the data from the upper layer and converts them into smaller units known as segments.
- This layer can be termed as an end-to-end layer as it provides a point-to-point connection between source and destination to deliver the data reliably.

The two protocols used in this layer are:

  - Transmission Control Protocol
  - User Datagram Protocol

Functions of Transport Layer:

- **Service-point addressing:** Computers run several programs simultaneously due to this reason, the transmission of data from source to destination not only from one computer to another computer but also from one process to another process. The transport layer adds the header that contains the address known as a service-point address or port address. The responsibility of the network layer is to transmit the data from one computer to another computer and the responsibility of the transport layer is to transmit the message to the correct process.
- **Segmentation and reassembly:** When the transport layer receives the message from the upper layer, it divides the message into multiple segments, and each segment is assigned with a sequence number that uniquely identifies each segment. When the message has arrived at the destination, then the transport layer reassembles the message based on their sequence numbers.
- **Connection control:** The transport layer provides two services Connection-oriented service and connectionless service. A connectionless service treats each segment as an individual packet, and they all travel in different routes to reach the destination. A connection-oriented service makes a connection with the transport layer at the destination machine before delivering the packets. In connection-oriented service, all the packets travel in a single route.
- **Flow control:** The transport layer also responsible for flow control but it is performed end-to-end rather than across a single link.
- **Error control:** The transport layer is also responsible for Error control. Error control is performed end-to-end rather than across a single link. The sender transport layer ensures that the message reaches the destination without any error.

5. Session Layer

- It is layer 3 in the OSI model.
- The Session layer is used to establish, maintain, and synchronizes the interaction between communicating devices.

Functions of Session layer:

- **Dialog control:** The session layer acts as a dialog controller that creates a dialog between two processes or we can say that it allows the communication between two processes which can be either half-duplex or full-duplex.
- **Synchronization:** The session layer adds some checkpoints when transmitting the data in a sequence. If some error occurs in the middle of the transmission of data, then the transmission will take place again from the checkpoint. This process is known as Synchronization and recovery.

6. Presentation Layer

- A Presentation layer is mainly concerned with the syntax and semantics of the information exchanged between the two systems.
- It acts as a data translator for a network.
- This layer is a part of the operating system that converts the data from one presentation format to another format.
- The Presentation layer is also known as the syntax layer.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Functions of Presentation layer:

- **Translation:** The processes in two systems exchange the information in the form of character strings, numbers, and so on. Different computers use different encoding methods, the presentation layer handles the interoperability between the different encoding methods. It converts the data from sender-dependent format into a common format and changes the common format into receiver-dependent format at the receiving end.
- **Encryption:** Encryption is needed to maintain privacy. Encryption is a process of converting the sender-transmitted information into another form and sends the resulting message over the network.
- **Compression:** Data compression is a process of compressing the data, i.e., it reduces the number of bits to be transmitted. Data compression is very important in multimedia such as text, audio, video.

7. Application Layer

- An application layer serves as a window for users and application processes to access network service.
- It handles issues such as network transparency, resource allocation, etc.
- An application layer is not an application, but it performs the application layer functions.
- This layer provides the network services to the end-users.

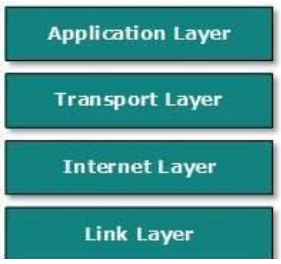
Functions of Application layer:

- **File transfer, access, and management (FTAM):** An application layer allows a user to access the files in a remote computer, retrieve the files from a computer, and manage the files in a remote computer.
- **Mail services:** An application layer provides the facility for email forwarding and storage.
- Directory services: An application provides the distributed database sources and is used to provide that global information about various objects.

TCP/IP Model

The OSI Model we just looked at is just a reference/logical model. It was designed to describe the functions of the communication system by dividing the communication procedure into smaller and simpler components. But when we talk about the TCP/IP model, it was designed and developed by the Department of Defense (DoD) in the 1960s and is based on standard protocols. It stands for Transmission Control Protocol/Internet Protocol. The TCP/IP model is a concise version of the OSI model. It contains four layers, unlike seven layers in the OSI model. The layers are:

1. Process/Application Layer
2. Host-to-Host/Transport Layer
3. Internet Layer
4. Network Access/Link Layer



1. Network Access Layer –

This layer corresponds to the combination of the Data Link Layer and Physical Layer of the OSI model. It looks out for hardware addressing and the protocols present in this layer allow for the physical transmission of data.
We just talked about ARP being a protocol of Internet layer, but there is a conflict about declaring it as a protocol of Internet Layer or Network access layer. It is described as residing in layer 3, being encapsulated by layer 2 protocols.

2. Internet Layer –

This layer parallels the functions of OSI's Network layer. It defines the protocols which are responsible for the logical transmission of data over the entire network. The main protocols residing at this layer are :

- **IP –** stands for Internet Protocol and it is responsible for delivering packets from the source host to the destination host by looking at the IP addresses in the packet headers. IP has 2 versions: IPv4 and IPv6. IPv4 is the one that most of the websites are using currently. But IPv6 is growing as the number of IPv4 addresses is limited in number when compared to the number of users.
- **ICMP –** stands for Internet Control Message Protocol. It is encapsulated within IP datagrams and is responsible for providing hosts with information about network problems.
- **ARP –** stands for Address Resolution Protocol. Its job is to find the hardware address of a host from a known IP address. ARP has several types: Reverse ARP, Proxy ARP, Gratuitous ARP, and Inverse ARP.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

3. Host-to-Host Layer –

This layer is analogous to the transport layer of the OSI model. It is responsible for end-to-end communication and error-free delivery of data. It shields the upper-layer applications from the complexities of data. The two main protocols present in this layer are :

- **Transmission Control Protocol (TCP) –** It is known to provide reliable and error-free communication between end systems. It performs sequencing and segmentation of data. It also has an acknowledgment feature and controls the flow of the data through a flow control mechanism. It is a very effective protocol but has a lot of overhead due to such features. Increased overhead leads to increased cost.
- **User Datagram Protocol (UDP) –** On the other hand does not provide any such features. It is the go-to protocol if your application does not require reliable transport as it is very cost-effective. Unlike TCP, which is a connection-oriented protocol, UDP is connectionless.

4. Process Layer –

This layer performs the functions of the top three layers of the OSI model: Application, Presentation, and Session Layer. It is responsible for node-to-node communication and controls user-interface specifications. Some of the protocols present in this layer are HTTP, HTTPS, FTP, TFTP, Telnet, SSH, SMTP, SNMP, NTP, DNS, DHCP, NFS, X Window, LPD. Have a look at Protocols in the Application Layer for some information about these protocols. Protocols other than those present in the linked article are :

- **HTTP and HTTPS –** HTTP stands for Hypertext transfer protocol. It is used by the World Wide Web to manage communications between web browsers and servers. HTTPS stands for HTTP-Secure. It is a combination of HTTP with SSL(Secure Socket Layer). It is efficient in cases where the browser needs to fill out forms, sign in, authenticate and carry out bank transactions.
- **SSH –** SSH stands for Secure Shell. It is a terminal emulations software similar to Telnet. The reason SSH is more preferred is because of its ability to maintain the encrypted connection. It sets up a secure session over a TCP/IP connection.
- **NTP –** NTP stands for Network Time Protocol. It is used to synchronize the clocks on our computer to one standard time source. It is very useful in situations like bank transactions. Assume the following situation without the presence of NTP. Suppose you carry out a transaction, where your computer reads the time at 2:30 PM while the server records it at 2:28 PM. The server can crash very badly if it's out of sync.

Difference between TCP/IP and OSI Model:

TCP/IP	OSI
TCP refers to Transmission Control Protocol.	OSI refers to Open Systems Interconnection.
TCP/IP has 4 layers.	OSI has 7 layers.
TCP/IP is more reliable	OSI is less reliable
TCP/IP does not have very strict boundaries.	OSI has strict boundaries
TCP/IP follows a horizontal approach.	OSI follows a vertical approach.
TCP/IP uses both the session and presentation layer in the application layer itself.	OSI uses different session and presentation layers.
TCP/IP developed protocols then model.	OSI developed the model then the protocol.

Difference between TCP & UDP

BASIS FOR COMPARISON	TCP	UDP
Meaning	TCP establishes a connection between the computers before transmitting the data	UDP sends the data directly to the destination computer without checking whether the system is ready to receive or not
Expands to	Transmission Control Protocol	User Datagram Protocol
Connection Type	Connection-Oriented	Connection Less
Speed	Slow	Fast
Reliability	Highly Reliable	Unreliable
Header Size	20 Bytes	8 Bytes

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Acknowledgment	It takes acknowledgment of data and can retransmit if the user requests.	It neither takes acknowledgment nor retransmits the lost data.
Protocol connection setup	Connection-oriented, the connection must be established before transmission	Connectionless, data is sent without setup
Data interface to the application	Stream-based	Message-based
Retransmissions	Delivery of all data is managed	Not performed
Overheads	Low but greater than the UDP	Very low
Data quantity suitability	Small to moderate amounts of data	Small to enormous amounts of the data
Implemented over	Applications where reliable transmission of data matters.	An application where data delivery speed matters.
Applications and protocols	FTP, Telnet, SMTP, IMAP etcetera.	DNS, BOOTP, DHCP, TFTP etcetera.

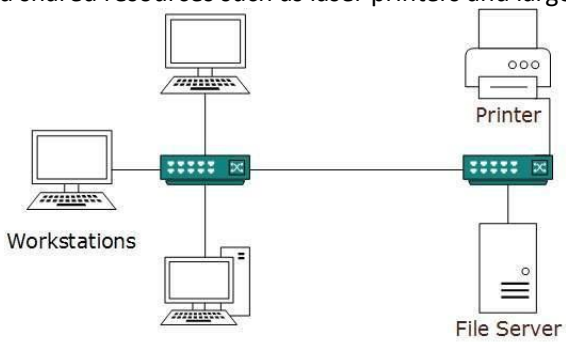
NETWORK SIZE

The network consists of various combinations of computers, storage devices, and communication devices. Based on the size of the geographical area a network can be divided into three following categories:

1. Local area network (LAN)
2. Wide area network (WAN)
3. Metropolitan area network (MAN).

1. Local Area Network (LAN)

Local Area Network (LAN) refers to a group of computers and other devices dispersed over a relatively limited area and connected by a communication link that enables any device to interact with any other on the network, LANs commonly include PCs and shared resources such as laser printers and large hard disks.



When the network is spread in a small geographical area say 0 to 10 km. then such network is called Local area network. Ring, Star, etc. are examples of LAN. LANs are usually intra-organizational and privately owned. For example, the network can be used to connect different terminals in the lab, to connect different offices of the same organization, or to connect different departments of a university, and so on.

Factors Contributed To The Growth of LAN

The primary reason for the users to go for LAN was distributed environment, which enabled them to have their independent processing stations while sharing computer resources like disk files, printers, and plotters. The other factors include:-

(i) Security for program and data: Both program and data reside on servers which can be made secure physically as well as logically (i.e., through software control). Also, diskless nodes are attached to the server that offers security by not allowing the user to use floppies or downloading software or data and servers are saved from the virus.

(ii) Expand PC usage: PCs in a LAN environment can be used as nodes and stand-alone PCs. For using PCs on nodes, Network Interface Cards (NIC) is required. This reduces the cost of setting LAN and training additional employees in such an environment.

(iii) Distributed processing: LAN provides the basic platform for interuser communication and information exchange in a distributed processing environment. This is usually the requirement of many companies.

(iv) Easy data management: Since data in a LAN environment is residing centrally on the server, it is easy to manage, easy to maintain back up, easy to enforce standards, easy to manage concurrent control, etc.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(v) Message broadcasting and e-mail: User can communicate in LAN environment by electronic mail, where each user can be assigned an e-mail-Id and mail-box on the server. From nodes, users can access their mailbox using their e-mail IDs after they log onto their networks.

(vi) Organizational benefits of LAN: These include an overall reduction in cost and time, faster communication between managers and executives, ability to access the data pertinent to them. LAN provides modularity, connectivity, superior performance, security, and reliability.

(vii) Reduced cost of software and up-gradation: Every organization is supposed to use the licensed software on every system. Being all the systems under LAN share with the software installed on the server, thus only a legal copy of the software needs to be bought instead of multiple copies. This reduces the cost significantly. Moreover, software upgrades are easy if they are centrally located on the server.

COMPONENTS OF LAN

LAN is comprised of Network, Operating System (NOS), Workstations, and shared resources. However, other components are associated with LAN like a file server, hubs, network interface card, etc. Every computer on LAN can be either a workstation or a file server (server). Other hardware components are also needed in the construction of LAN.

- **Workstation:** A workstation in LAN is a single user microcomputer with communication capabilities. It is a personal computer capable of supporting hardware and software necessary to connect to a LAN.
- **File Server:** A file server is a powerful computer that works as a server for the LAN. It is loaded with special software such as Network Operating System is loaded to control the functioning of LAN. It has all the capability for processing and has a huge disk space, which is shared by all workstations. All the shareable resources like printer, scanner are attached to the file server only. Database for the entire LAN is loaded on this computer only. It is the mother of the LAN system.
- **Network Interface Card:** Network interface cards mediate between the computer and the physical media such as cabling over which data transmission takes place. The network interface card is a microprocessor-based device containing hardware and software which supplies the intelligence to control the access and communication across the network and performs all communication process. It is the device through which each workstation is connected to the network functionally and physically. LAN network interface card specifies the data transmission rate, the size of the message unit, the addressing information attached to each message, and the network topology.
- **Network Operating System:** An operating system specifically designed to support networking. A server-based network operating system provides networking support for multiple simultaneous users as well as administrative, security, and management functions. On the desktop, it provides users with the ability to access network resources. It acknowledges and responds to requests from many workstations, managing details as network access and communication resource allocation and sharing, data protection, and error control. Examples of popular NOS software are Novell Netware, Microsoft NT /2000, Unix, and Linux. Peer-to-peer networking can also be accomplished with Microsoft windows 95/98/MF /XP.
- **Network cabling or cabling system -** Once the server, workstations, and network interface cards are in place then cable is required to connect the network. The network cables and network interfacing cards are chosen should match each other. To transmit data, the commonly used network cables are shielded twisted-pair, coaxial, and fiber optic cables.

2. WIDE AREA NETWORK

Wide area networks (WAN) are formed when a geographically distributed network composed of LANs joined into a single large network using services provided by common carriers. WANs are commonly implemented in enterprise networking environments in which company offices are in different cities, states, or countries, or on different continents. WANs are of two types. WANs are broadbands, usually public access networks.

The well-known examples of private and public networks are given below:

(a) Private Network: WANs are commercial and regulated networks, while others are privately owned usually by large-sized organizations. Some W ANs Private networks are regional networks for corporations spread over a large area such as telephone companies.

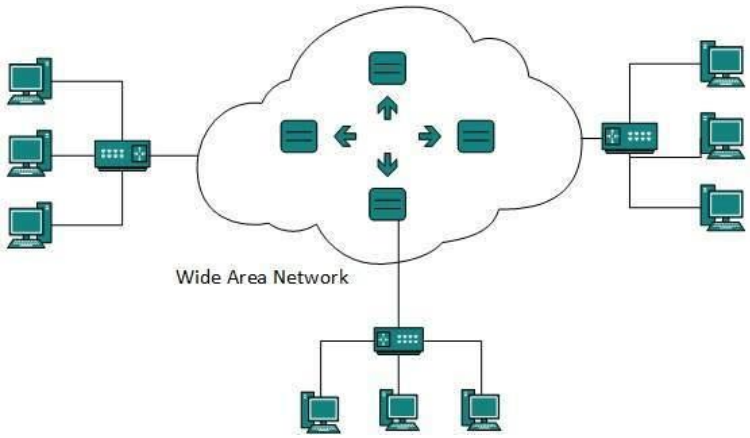
Examples:

ARPANET: This was the first Wide area network, built-in the United States of America. It was built for the Defence Departments to utilize some special facilities available in different computer centers, by the large community of users. It was a nationwide network supporting more than 50 different universities.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

SWIFT (Society of Wide Area Inter Bank Financial Transactions): It is a network of Banks. The objective of this network is to replace mail, cable, telex and work for 24 hours a day, 365 days a year. It handles millions of transactions every day.

SITA: It is another private network that passes the information between airline computers around the world.



(b) Public Network:

Some W ANs can be implemented by the public networks in terms of their management, resources, and access. One such public WAN is the Internet whereas public networks are, maybe international networks as well, built by the Government Telecommunication agencies e.g., global communication service providers. Generally, public network handles a bulk amount of data and is more economical.

Examples:

NICNET: National Informatics Centre Network was developed by the Government of India, in the public interest for various Government activities.

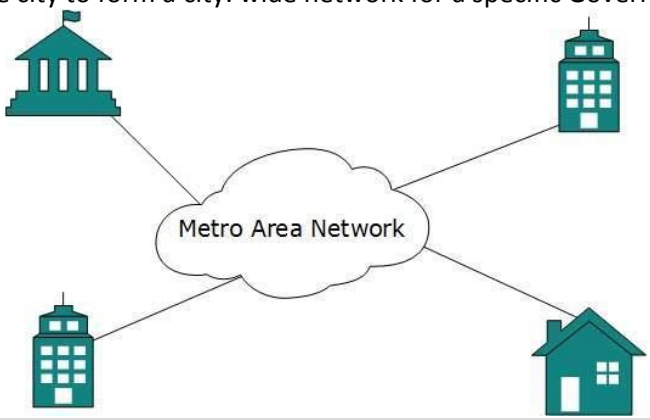
Viewdata Scheme: It was developed for British Postal services for operating public data Bank Accessible with television sets at homes or offices, initially via the telephone lines.

The Salient Features of WAN are:

(i) Communication network covers a wide geographical area such as country or the world.
(ii) Multiple user's computers are connected.
(iii) Communication channels are usually provided by a third party. A W A may use a combination of wired and non-wired media, a telephone company, a public data network, or a non-wired media-satellite carrier.
(iv) Communication channels are of relatively low capacity, throughput is measured in kilobits per second (kbps).
(v) Communication channels are error-prone.
(vi) A variety of computers are linked, from mainframe to terminals.

3. METROPOLITAN AREA NETWORK (MAN)

MAN is a network bigger than LAN but smaller than a WAN it is also defined as multiple networks that is connected within the same city to form a city. wide network for a specific Government or industry.



Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

MAN is a high-speed network that can carry voice data and images at up to 200 Mbps or faster over distances of up to 75 km. The transmission speed can be higher for shorter distances, depending on the network architecture.

A MAN can include one or more LANs as well as telecommunication equipment such as a microwave or satellite relay station. It operates at a higher speed and is smaller than WAN.

CHARACTERISTICS OF LAN AND WAN

Characteristics	LAN	WAN
Geographic coverage	Restricted to limited geographic coverage(only a few kilometers)	WAN may extend over several thousand kilometers i.e., nationwide or worldwide
Data transmission Rate	Higher i.e., 2 Mbps to 1 Gbps(Mbps - Megabits per second)	Slower, i.e., 1200 bits per second to 1 Mbps (Gbps - Gigabits per second)
Transmission error rate	Few, data transmission errors i.e., 10-8 to 10-12	High data transmission error rate i:e; 10-5 to 10-7 bit error rate
Communication link	Common communication link are twisted pair, coaxial cable, and fiber optics	Relatively slow and unreliable, links used are telephone lines, microwave links are satellite channel
Ownership organization	Owned by a single organization	Belongs to different organizations
Communication Cost	Cheaper	Expensive
Maintenance and administration cost	Less	More

INTRODUCTION TO INTERNET

Internet

Internet is defined as an Information super Highway, to access information over the web. However, It can be defined as a worldwide global system of interconnected computer networks.

Internet refers to an internetwork which is a set of computer networks that may be dissimilar and is joined together through gateways that handle data transfer and conversion of messages from sending network protocols to those of the receiving network.

BRIEF HISTORY: GROWTH OF INTERNET (Evolution)

The concept of the Internet was originated in 1969 and has undergone several technological & Infrastructural changes as discussed below:

- The origin of the Internet is devised from the concept of Advanced Research Project Agency Network (ARPANET).
- ARPANET was developed by the United States Department of Defense.
- The basic purpose of ARPANET was to provide communication among the various bodies of government.
- Initially, there were only four nodes, formally called Hosts.
- In 1972, the ARPANET spread over the globe with 23 nodes located in different countries and thus became known as the Internet.
- By the time, with the invention of new technologies such as TCP/IP protocols, DNS, WWW, browsers, scripting languages, etc., the Internet provided a medium to publish and access information over the web.

In the 1980s Internet use spread among universities and in the 1990s its usage was among people and organizations worldwide. NSFNET backbone beginning in 1995 sparked a massive restructuring aimed at shaping the Internet into a faster and more productive tool for business. Internet followed a voluntary policy called Acceptable Usage Policy (AUP) that allowed only non-profit educational and Government agencies to use the Internet. The new policies allowed many commercial uses of the Internet. To serve the new customers, a new industry emerged called Internet Service Provider (ISP). ISP is a business that supplies Internet connectivity services to individuals, businesses, and other organizations.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

CHARACTERISTICS OF INTERNET

As we know the Internet is a network of separate distinct networks. It does not represent a single entity and is not owned by any single person or organization. Thus being "Wild West", nature of the Internet possess the following characteristics, these are:

(i) Complex Network: The Internet is a huge set of networks connected i.e., more than 100 separate Internet networks meet to exchange data at the Network Access Point (NAP).

(ii) Hierarchical in structure: At the top, there are very large national Internet Service Providers (ISPs). These national ISPs connect and exchange data at network access points (NAPs). Though at the end-user level, it appears to be unorganized and confusing. But it is a strict rigidly controlled club in which deviation from the rules is prohibited.

(iii) Open-membership professional society: Because the Internet is said to be a network of networks, not one organization operates and owns the Internet. The owner of the Internet is the Internet Society (ISOC: www.internetsociety.org) which is open and allows, anyone including students to join, Currently, ISOC is an open-membership professional society that has created the Internet and its technologies, which works on public policy, education, and Internet standards. It is comprised of nearly 200 organizations and corporations 8000-9000 individual members in over 100 countries.

(iv) Dynamic in nature: The Internet is changing. New applications and access technologies keep changing. The new version of the Internet is Internet 4. There are many different organizations currently working on the next generation of the Internet.

(v) Internet covers globe: Internet is accessed by people across the globe in over 155 countries.

(vi) Wider application accessibility: The Internet offers access to data, graphics, sound, software, text, and people through a variety of services and tools for communication and data exchange.

(vii) Growing exponentially: The Internet joins a set of one or more networks every day. The expansion rate of the Internet is 12% per month.

ANATOMY OF INTERNET: HOW INTERNET WORKS

Anatomy of the Internet refers to the study of structure. The Internet is hierarchical in structure. At the top, there are very large (ISPs) Internet Service Providers such as VSNL, DSL. These national ISPs connect and exchange data at Network Access Point (NAP).

As you know from the history of the Internet, in the early 1990, the Internet was primarily run by National Science Foundation (NSF-US). The four main NAP were established by NSF in the US to connect national-level ISPs. Since Internet work was growing exponentially than NSF stopped funding the Internet. The companies involved in running NAP began charging for connections to the national ISPs. Therefore presently, many international commercial companies are running various NAPs in various countries. In the hierarchy of the Internet, local ISP relies on regional ISPs and regional ISPs on national ISPs to transmit their messages in other countries. The local ISPs are responsible for providing connectivity to the individuals. Due to the exponential growth of ISPs, another form of NAP emerged -the Metropolitan Area Exchange (MAE). MAE is responsible to link regional SPs whose networks are connected in major cities. As they (MAE) act as NAP and work at the metropolitan level, thus they are considered to be a smaller version of NAP. To provide a backup connection, in case one Internet connection fails, regional and local ISP usually have several connections into national and other regional ISP. This way internetwork maintains its efficiency because they are not dependent on one higher level ISP.

INTERNET GOVERNANCE

There is no single organization that operates the Internet. It is not owned by a single owner. Internet Society (ISOC) is the closest as the owner of the Internet. ISOC is an open-membership professional society with over 175 organizations, corporations, government agencies, and nearly 9000 individual members.

Membership in ISOC is open to anyone, information regarding this is available on www.isoc.org. ISOC's critical mission is to work for Internet development, evolution, use, and benefit the people around the world. ISOC primarily addresses three issues:

1) **Public policy:** They frame policy regarding copyright, privacy access, and censorship.
2) **Education and Training:** To improve Internet infrastructure in developing countries, ISOC delivers education and training programs.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

3) **Maintaining standards:** ISOC works typically for the development and maintenance of Internet standards.

ISOC works through four worldwide known interrelated standard bodies, these are:

- **Internet Engineering Task Force (IETF):** It is a large open international task force concerned with the evolution of Internet structure and efficiency in the operation of the Internet. This team consists of network designers, operators, vendors, researchers. The information regarding IETF is available at www.ietf.org.
- **Internet Engineering Steering Group (IESG):** IESG is responsible for administering the rules and procedures laid down by ISOC. IESG is primarily responsible for technical management of IETF activities and processing the Internet standards. Thus a member of IESG chairs each the working group of IETF.
- **Internet Architecture Board (IAB):** IAB works for the strategic architecture of the Internet. The concerning strategic issues are passed and guidance to IESG, such as top-level portal names, domain names, use of international character set. IAB community is comprised of all IESG members and chairperson of IETF.
- **Internet Research Task Force (IRTF):** This community operates in the same manner as IETF. The only difference is that they focus on specific issues through research groups. IRTF primarily works on long-term issues such as Internet applications, architecture, protocols, and technology. IAB is responsible for framing and appointing the chairperson of IRTF.

VARIOUS COMMUNICATION SERVICES

The innovation of the technological key areas that have fulled e-commerce are the following:

(I) Telecommunication Companies: Providing simple services such as phone calls to all types of customers. They are developing new technologies XDSL, fast switches for higher bandwidth communication across existing networks.

(II) Satellite Technology: Vendors setting up a new broadband network with global reach.

(III) Wireless Network: Provides Internet facility at a faster pace, even on cell phones using wireless Internet protocol.

(IV) Cable Companies: Provide two-way Internet traffic on television by introducing set-top boxes which act as converters and separators for the inbound and outbound traffic.

Dialup Internet Service

Dialup internet service is a service that allows connectivity to the internet through a standard telephone line. By connecting the telephone line to the modem in your computer and inserting the other end into the phone jack, and configuring the computer to dial a specific number provided by your internet service provider (ISP) you can access the internet on your computer.

Dial-up internet service is provided through several ISP. The majority of internet service providers give you a set of telephone numbers either national or local that allows you to dial into a network that feeds into the internet. This allows you to receive and send an email, search the World Wide Web, participate in chat rooms, and plenty of other features the web has to offer.

DSL Internet

DSL stands for Digital Subscriber Line. Users get a high-speed bandwidth connection from a phone wall jack on an existing telephone network. DSL works within the frequencies that the telephone doesn't so you can use the Internet while making phone calls.

Consider DSL Internet the big brother to dial-up. Many local phone services provide DSL offering High-Speed Internet access delivered through existing telephone networks. Telephone wires carry hundreds of thousands of frequencies. Only a few thousand are for telephone communications so the telephone and DSL modem can work at the same time. There are two main types of DSL technology.

Symmetrical DSL– Symmetrical connections offer equal bandwidth for upload and download speeds. **Asymmetrical DSL**– This is the most popular type of DSL connection. Most people download more information than they are uploading. Because of this, an asymmetrical connection has more downstream bandwidth and less upstream bandwidth.

Broadband technologies

In telecommunications, a broadband signaling method handles a wide band of frequencies. "Broadband" is a relative term, understood according to its context. The wider (or broader) the bandwidth of a channel, the greater the data-carrying capacity, given the same channel quality.

In radio, for example, a very narrow band will carry Morse code, a broader band will carry speech, and a still broader band will carry music without losing the high audio frequencies required for realistic sound reproduction. This broadband is often divided into channels or "frequency bins" using passband techniques to allow frequency-division multiplexing instead of sending a higher-quality signal.

In data communications, a 56k modem will transmit a data rate of 56 kilobits per second (kbit/s) over a 4-kilohertz-wide telephone line (narrowband or voiceband). In the late 1980s, the Broadband Integrated Services Digital Network (B-ISDN) used the term to refer to a broad range of bit rates, independent of physical modulation details. The various forms of digital subscriber line (DSL) services are broadband in the sense that digital information is sent over multiple channels. Each channel is at a higher frequency than the baseband voice channel, so it can support plain old telephone service on a single pair of wires at the same time. However, when that same line is converted to a non-loaded twisted-pair wire (no telephone filters), it becomes hundreds of kilohertz wide (broadband) and can carry up to 100 megabits per second using very-high-bit-rate digital subscriber line (VDSL or VHDSL) techniques.

RESOURCES REQUIRED TO ACCESS INTERNET

A typical infrastructure set up consists of the following main components:

- **Hardware** These consist of desktops, workstations, servers, storage/ backup devices, printers, plotters, and pervasive computing devices like mobile phones, personal digital assistants, etc.
- **Network Infrastructure** these include data communication components like switches, routers, hubs, firewall, and Public Key Infrastructure (PKI) for transaction security, etc.
- **Core/System Software** consists of operating systems like Windows NT, Linux, UNIX, Sun Solaris, etc. databases like Oracle, DB2 and MS-SQL, mail servers, e-security software like firewalls, etc., and also the application software.
- **Common Business Services Infrastructure** This infrastructure includes design development, implementation, customization, integration, training, certificates, digital signatures, and payment services.

APPLICATION OF INTERNET

Internet applications are many and depend on the innovation of the user. The common applications of the Internet can be classified into three primary types namely: Communication, Data retrieval, and Data publishing.

(I) Communication: Communication on the Internet can be both online or offline. When some users connect to a single server at the same time, they can communicate to each other in an "online chat manner". This can be "many to many" as in a room full of people talking to each other on a peer to peer basis. Alternatively, the users send e-mail to each other which can be read by the receiver whenever he/ she finds the time. This is off-line communication, that can be "one to one" or "one to many". Similarly, users can get together electronically with those sharing common interests in "use net" groups (maybe an unknown group). The users post messages to be read and answered by others at their convenience, in turn, all of which can be read and replied to by others.

(II) Data Retrieval: On the Internet, a large number of databases exist. For data retrieval, the availability of data that has been compiled from various sources is put together in a usable form. These databases have been put together by commercially-run data providers as well as individuals or groups with a special interest in particular areas. To retrieve such data, any user needs to know the address/ s of such Internet servers. Then depending on the depth of information being sought, different databases have to be searched and required information complied. The Internet "library" is immense, dynamic, because of regular updating, and entirely electronic. While some skill is required for searching, the user will be able to access, search and check a large collection of servers. It may be concluded that the Internet will have the

latest information available because all data is transmitted electronically the world over. For corporate users who are looking for the latest information on products, markets, and strategies, such sources are of great help.

(III) Data Publishing: Data publishing is a new opportunity that Internet has made possible. Information that needs to be made available to others can be either forwarded to specific addresses, posted on a Usenet site, or kept on display on a special site. Internet discourages by social pressure, sending of unsolicited e-mail.

(IV) Business use of the Internet: The world is shrinking, thus, global communication facilities made the world appear even smaller. Now Internet is affecting human living more than any other change that was what everyone believed when steam engines and other transportation means changed the way people lived. Business methods and practices have changed the lifestyles of the people.

This is already happening in economically advanced countries. Telephone and fax are things of the past. No one can conduct any worthwhile business unless the business communication occurs on E-mail. An E-mail address is a basic facility that anyone essentially gains on the Internet. Communication is quicker and more convenient than any other mode ever used. It is reliable and provides easy identification of the sender. Electronic mail works similar to paper mail through the postal system but is entirely electronic (Electronic mail has been discussed in detail in the subsequent section)

Through the Internet, users can have online interactive communication. This is another major facilitator for any business. Most businesses need a lot of interaction internally among their executives and externally with business associates when many persons have to contribute to discussions. This is normally achieved by arranging meetings and conferences for which people travel from different locations to come together in one place. This "many to many" communications can be handled on the Internet quite effectively, without any need for people to travel. At pre-defined timings, those who have to communicate together can be online through the Internet and have electronic conferences. This is a "many to many" interaction. Similarly, many discussions can be done through a forum where people post messages for each other and respond to them periodically. This is similar to having a bulletin board on which everyone posts messages.

INTERNET COMPONENT

Internet is used for various applications by its users from different domains such as communication, files transfer Education Entertainment business, etc. because of the various services offered through the Internet these services are nothing but components of the Internet which are described below:

Electronic Mail (e-mail)

E-mail (electronic mail) refers to the exchange of text messages and computer files over a communication network such as LAN or Internet.

The Internet provides basic communication service to its user in any part of the world by-mail facility. E-mail allows messages, maybe in text, audio, video, or combination of all, accessed by multiple users, delivered exclusively on a computer network. When a message is sent from a source user, it reaches the destination user mailbox. The user at the receiving end has the option to read the e-mail, send a reply, edit the email, save it or forward it. It is the most popular application of the Internet, as a means of communication. E-mail messages are usually received a second after the transmission. E-mail can be sent by connecting to the network from any location that has a telephone line using a portable computer and a modem. It is also possible to connect to networks with wireless technologies. This job is performed by Simple Mail Transfer Protocol (SMTP) which runs over TCP /IP. Many organizations have adopted e-mail as an important tool of communication. The Internet has proved to be a faster and productive tool of communication because of its e-mail service.

The primary advantages/useful features of e-mail are:

- **Paperless communication:** It conducts paperless communication of messages very quickly.
- **Composing messages:** Copies of messages can be sent to many users simultaneously, the messages may include pictures, voice, video, audio, film clips, text, and animation.
- **Printing of messages:** The messages can be printed, prioritized, and stored.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- **Public bulletin boards:** It can be created where everyone in the organization can post and view messages. Moreover, it also shares text messages and application files across popular computer platforms.
- **Faxes and meetings:** It allows delivery and receiving of faxes and meetings can also be scheduled.
- **Forwarding of messages:** It is possible to forward any message without retyping the message.
- **Transfer of data/files:** One of the important features of e-mail is that a larger data documents file can be sent using Attach File Command to attach the document.
- **Address Book:** The user can store the e-mail addresses of people sending them messages in the user program's address book. This feature allows the users to organize their e-mail addresses according to the person's nickname.
- **Composing messages:** Once connected with ISP, it is possible to compose a message with the help of an Internet Browser. Like any other text editing tool, users can compose attractive mails by using stylish fonts, pictures and before finalizing, it is possible to spell check the message.
- **Replying to the mail received:** The user can save a lot of time in terms of remembering addresses and typing the subject matter by using the "Reply" option on the Internet Browser.
- **Offline Editing/Composing/Reading:** The user can save Internet time and money by keeping telephone lines free, provided the user downloads all the messages into the hard disk and performs reading, composing and editing of messages offline (i.e., disconnected from Internet).
- Once the user logs on then messages can be upload and sent to desired addresses.
- **Greeting Card:** Users can save a considerable amount of time, energy, money by using an electronic greeting card facility. Unlike traditional printed cards, these cards are extremely attractive, animated musical, interesting, easy-reachable.

However, there are certain **drawbacks in using e-mail**, such as the following:

- Face-to-face communication is not possible.
- All e-mail systems are not very user-friendly.
- The confidentiality and privacy of the messages are generally at risk.
- E-mail system is always prone to virus threat.

Uploading is the process of putting web pages, images, and files onto a web server. To make a file visible to everyone on the internet, you will need to **upload** it.

Downloading is the process of getting web pages, images, and files from a web server.
When users are copying this file to their computer, they are **downloading** it.

Downloading Software and" Files

Software packages can be downloaded, maybe free either as public domain (cheap ware), or copyrighted, but given away (freeware). However, others are called shareware, which can be used on a trial basis, and if, user likes, then it may be expected to be purchased by them. Internet users can also retrieve documents including graphics, voice, and other media for use on their PCs. Usually, with the use of Web or Gopher, it is fairly easy to do so. Many of the documents are transferable from the Internet using a member of the TCP lIP family called **File Transfer Protocol (FTP)**.

File Transfer Protocol (FTP): The file transfer protocol provides the facility to transfer files between machines over the Internet. With this, large files of a computer program containing text, images, sound, movie, etc., can be transmitted. It also allows the concerned access to the same Internet servers by multiple users.

By using FTP, a file transfer takes place in the following manner. Using special FTP software, you can easily download files stored in an FTP site, a server that archives such files.

Chat Programs

Chat programs allow Internet users to send and receive messages instantly free of cost. For this, the people must establish the connection in the same channel of communication. In chat programs, the user receives messages on the screen, even when the sender of the message is in the middle of typing. It is also possible

for the users to have voice chats like voice e-mail or telephonic conversation. As any number of users can join in the online conversation, it appears as if a global conference call has been scheduled on such systems. Two major types of chat programs exist:

(a) Webchat refers to a Web-based program, which allows Internet users to send and receive messages using a Web browser onto Web chat site. For example, the Budweiser Web site has a popular chat program.

(b) Internet Relay Chat (IRC) is an e-mail-based (text only) program, that provides the facility to interact with customers in business and online expert's answers to questions.

Newsgroups

The Usenet or just Newsgroup service of the Internet allows a group of users to focus on a topic of their interest by exchanging their information. To become a member of a newsgroup, the users are required to subscribe to it and they are organized in a directory and are divided into categories and sub-categories. There may be thousands of users that join such a group, each of which is dedicated to a specific topic. In this, a user can read all the messages sent by other members of that group and can respond to all members by posting a message, or e-mail to just one individual. However, the chat program allows the users to communicate in real-time, but it may be a waste of time if the chatters are not focused on a particular issue. Moreover, in contrast to a chat room, where all communication is done in real-time, here communication is done via messages in a different time/ different place mode.

All newsgroups may not be available at the user's site. As there are several users in the newsgroup, that may create Internet traffic, thus, some of the messages may get blocked. For example, AOL and similar services have their newsgroups.

Newsgroups are generally used for discussion and posting business opportunities, such as creating partnerships or selling items. In chat or newsgroup, the user can use a nickname to keep his/her identity secret as the user is usually not acquainted with all possible readers/members.

Electronic Bulletin Boards (EBBs)

EBBs refer to mailing lists on which users can leave messages for other people and receive massive amounts of information, including free software. They are similar to the Web newsgroup, but initially, they resided on the non-commercial Internet and private information providers' networks. To access an EBB, you need to be a member and sometimes pay a nominal fee. Special interest groups, like users of specific software, display messages on these boards and exchange experiences about the software. Universities and schools use the EBB to advertise classes and special programs.

Telnet (Terminal Network)

Telnet is a standard Internet application that provides the user the ability to log into the machine over TCP/IP. Using such a connection, users can execute standard programs from remote locations including e-mails. Some of the common uses of the telnet service are :

1. The computing power of the remote computer that may be a powerful supercomputer can be used by the local computer that may be an ordinary personal computer.
2. Software of the remote computer may be used that the user wants to use, which may not be available on the user computer.
3. The remote computer's database or archive can be accessed. As the information archive such as the public database or library resources may be available on the remote computer.
4. A user can log into one's computer from another computer. For example, if the user is attending a conference in another city, then telnet on the Internet allows the user to access his/her computer and read his/her electronic mails.

Internet Telephony

Internet Telephony refers to a combination of software and hardware that allows a user to approximate a telephone conversation over an Internet connection instead of using a traditional telephone system.

Internet telephony is a technology that lets us make a long-distance voice call through the Internet. This is also called IP telephony. A combination of hardware and software that allows the Internet to be used as a telephone carrier. Ideally, if the user's Internet access number is a local call, then the user can talk to anyone on the planet that is also set up for Internet Telephone for the price of a local call. However, the quality of the connection may not be considered as good as one gets over a regular phone line. After the

costs of initial set-up and access to an Internet Service Provider, long-distance voice calls can be made via the Internet free of charge. Any means of transmitting the human voice (real-time or close to real-time) over the internet. There are several components required to establish such connection:
(1) On the client-side, a multimedia-equipped PC with special client software will digitize the sender's voice. This can be done with a voice modem or other voice encoding method;
(2) A direct or dial-up connection to the internet allows your voice to be transmitted in packet form to its destination;
(3) Connection with the far side is achieved by IP address search, common servers, or beacons to identify the called party (and to "ring" that person's phone);
(4) A similar arrangement on the far end completes the call and allows both parties to speak. There are also Internet gateways that allow regular telephone callers to make Phone-to-Internet-to-Phone connections. There are PC-to-Phone connections and Phone-to-PC connections.

WAIS: Online Databases

WAIS is an acronym for Wide-Area Information Server. They are standardized databases with hypertext features. A Unix-based document search and retrieval system on the Internet that can be used to search over 400 WAIS libraries, such as Project Gutenberg, for indexed files that match keywords entered by the user. WAIS can also be used on an individual website such as a search engine. W AIS developed by Apple Computer, Thinking Machine Corporation, and Dow Jones, uses the Z39.50 standard to process natural language queries that means we can type the whole sentence like "find all papers on security risk management in e-commerce". The list of documents returned by WAIS often contains numerous false matches. Like WWW and Gopher, WAIS is a client-server arrangement, so users need a WAIS client to use a WAIS server. WAIS clients can retrieve text or multimedia documents stored on the servers. The kind of information resources we can find in the WAIS database are typically bibliographies, books, archive databases (FTP files) library catalogs, and many Internet-related databases.

INTERNET RELATED TERMINOLOGIES

World Wide Web (WWW)

World Wide Web (WWW) refers to an interconnected system of computers all over the World that store information in multimedia form.
The webpage is a document on WWW that can include text, pictures, audio, and video. Documents on WWW are called Web pages.
WWW or just the Web is the most popular method of accessing the Internet services and its information located on different servers, it resides on the top of the Internet. The web is an integrator of Internet applications with multimedia. The web is popular because of the use of the hypertext concept. The principal Web tool is the browser, a program that uses HTTP to retrieve information (or surfing) from a website containing web pages from HTTP servers worldwide. Hypertext provides a new way of storage and retrieval of information on the Internet, with this author can structure the information. An effectively designed hypertext facilitates the user to locate the desired information as fast as possible, as the Internet is an ocean of information. This is possible due to the series of links present in the hypertext document. A link connects a document to another document that may be anywhere on the Internet, basically to provide more information about the linked item. The linked item is simply clicked with the mouse to make the connection possible. The first page is called a Home Page which contains a sequence of linked pages. Browsers like Netscape Navigator and Microsoft Internet Explorer provide GUI for receiving or sending multimedia information. Electronic mail and surfing a site over the World Wide Web cause traffic in the network, as these are the important and common applications over the Internet.
Website refers to the group of related HTML documents and associated files, scripts, databases, etc. served by the HTTP server on the WWW. The HTML document generally covers one or more related topics that are interconnected through hyperlinks.
Website home page works on starting page which frequently functions as a table of contents for the site.
Web client refers to a computer that is connected to the Internet and used to Download Webpages.

IP Address

IP address stands for Internet Protocol address. An IP address is used to identify a computer on the Internet uniquely, when a user connects to an ISP, the ISP provides an IP address from the list so that user connection with an ISP can be uniquely identified. When a user breaks its connection with the Internet, then its IP address is added back to the list of the free IP address. So that ISP can now allocate an IP address to another user. IP addresses are assigned by Internet Assigned Number Authority (IANA:HTTP:/ / www.iana.org.). The ISP buys several IP addresses for a number of its subscriber. It is a 32-bit (or 4-bytes) binary number that uniquely identifies a host (computer) connected to the Internet to other Internet hosts for communication through a transfer of packets. An IP address is generally pressed in "dotted quad" format, consisting of 4 bytes decimal values, separated with periods i.e. 132.0.0.1. These first three bytes identify the host itself and the other represent class and network number. The 4 bytes consisting of 32 bits signify 232 (or 4 billion) hosts.

WEB Browser

A browser is a software application used to locate, retrieve and display content on the World Wide Web, including Web pages, images, video, and other files. As a client/server model, the browser is the client run on a computer that contacts the Web server and requests information. The Web server sends the information back to the Web browser which displays the results on the computer or other Internet-enabled device that supports a browser.

A browser is a group of structured codes which together perform a series of tasks to display a web page on the screen. According to the tasks they perform, these codes are made as different components.

There are four leading web browsers − Explorer, Firefox, Netscape, and Safari, but there are many others browsers available. You might be interested in knowing Complete Browser Statistics. Now we will see these browsers in a bit more detail.

While developing a site, we should try to make it compatible with as many browsers as possible. Especially sites should be compatible with major browsers like Explorer, Firefox, Chrome, Netscape, Opera, and Safari.

Internet Search Engines

A search engine is a tool that makes a browser more effective. It enables the Internet user to locate information by similarly using keywords or key phrases as one would search online library sources. The search engine matches the keywords, even if they appear anywhere in the abstracts, and presents the results to the user. The abstracts contain links to the source documents, so one can view them in full length. Search engines can be reached directly by entering their address in the location box of most browsers. For example, www.google.com.Altavista.com, Yahoo.com.

Internet Addresses

A Uniform Resource Locator (URL) indicates the location or address of a Website that the user wants to visit. The address of a Website consists of several parts, for example, Yahoo's address (http://www.yahoo.com)means that it uses a Hypertext Transfer Protocol (HTTP), which is a set of communication rules that allows a browser to connect with the Web Server for retrieving hypermedia and hypertext documents. "WWW" indicates that the user is requesting a Web server. "Yahoo" is the name of the company, and ".com" refers to a commercial organization. The URLs are classified into generic domains. In addition to .com and to an educational one .edu, there are .net, .gov, .firm, .store, .info, .art, .web, .rec, and .nom. Universal Resource Locator (URL) is a string of characters that points to a specific piece of information anywhere on the Web. A URL consists of **(1)** Web protocol; **(2)** the name of the Webserver; **(3)** the directory on the server and **(4)** file within the directory.

INTRANET

The Intranet is a type of information system that facilitates communication within the organization, among widely dispersed departments, divisions, and regional locations. Intranets connect people with Internet technology, using Web Browsers, Web Servers, and Data Warehouses in a single view. With an Intranet, access to all information, applications, and data can be made available through the same browser. The objective is to organize each individual's desktop with minimal cost, time, and effort to be more productive, cost-efficient, timely, and competitive. According to James Cimino, the challenge is to realize the following from focused Intranet work:

- Easily accessible information
- Reduced information searching time
- Sharing and reuse of tools and information

- Reduced set-up and update time
- Simplified reduced corporate licensing
- Reduced documentation costs
- Reduced support costs
- Reduced redundant page creation and maintenance
- Faster & cheaper creation
- One-time archive development costs
- Sharing of scarce resources of skills.

A properly planned Intranet implemented after a careful study of the business problems or issues can be a great help in the streamlining of a company. Some of the key benefits of using Intranet are:

- ✓ Reduced costs - printing, paper, software distribution, mailing, order processing, telephone
- ✓ Easier, faster access to information
- ✓ Easier, faster access to remote locations
- ✓ The latest, up-to-date research base
- ✓ Easier access to customers and partners
- ✓ Collaborative, group working

COMPARISON BETWEEN INTERNET AND INTRANET

Internet	Intranet
Internet refers to the network of various networks worldwide.	Intranet refers to an organizational network, where communication is established via the internet and web technology.
Information can be accessed from anywhere, everywhere, all-time (24 X 7) days.	Only organizational information is accessible from any location on the network.
Every organization or individual can keep information on their website.	Information is restricted to the particular organization only.
The Internet offers services, such as worldwide webs (www) e-mail, Gopher data publishing, Usenet, etc.	Using the Internet, organizations reduce the time of communication, searching information, business objectives, and effective management of information system i.e., supplier management, inventory management, distribution management, channel management.
Trading is with untrusted partners using the Internet.	Trading is with trusted partners
Security is a critical issue	It is more secured than the Internet

EXTRANET

Extranet refers to a group network that uses Internet technology to connect business partners, including suppliers, customers distribution service providers, or any other businesses engaging in collaborative ventures. Extranets may be designed to allow authorized trading partners to link together portions of their intranets because extranet is considered an extension of a corporate intranet using World Wide Web technology to, facilitate communication with the corporation's supplier and customers. An extranet allows customers and suppliers to gain limited access to a company's intranet to enhance the speed and efficiency of their business relationships. The Extranet is thus extended Intranet, which allows business communications from the Internet through secure solutions.

Does a question arise how extranets differ from Internet?

Answer Extranets may use Internet routes and ISPs to transmit data.

Extranets are wider in scope than Intranets but are only a subset of the Internet

The designs of the extranets need to have well-designed security controls. The connecting partners are connected via a leased line and they maintain data confidentiality using encryption method. Extranets may rely on firewalls. By implementing an extranet, the entire network becomes as vulnerable as the weakest link. End-to-end security measures are a crucial factor for extranet members, and these security issues and techniques are essential to implement.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Comparison between Intranet and Extranet

Intranet	Extranet
Intranet refers to an organizational network where communication is established via the Internet and web technology.	Extranet refers to a group of the network which uses Internet technology to connect business partners including suppliers, customers, distribution service providers, and other businesses engaging in collaborative ventures.
Only organizational information is accessible from any location on the network.	It is an extended Intranet that allows business communication among trading partners.
Due to the use of the Internet, organizations reduce the time of communication, searching information, effective management of information system.	An extranet allows customers and suppliers to gain limited access to a company's Intranet to enhance the speed and efficiency of their business relationships.
Security is concerned but not as critical as in Extranet.	Business communication on other Intranet is established from the Internet through a secure solution.

USAGE OF INTERNET TO SOCIETY

Usage of the Internet to society can be visualized with the usage of Internet by the people to perform certain things the Internet can be used for:-

(a) Finding businesses, products, and services (E-business): The Internet can be used to find businesses, products, and services. Yellow page directory services enable to search by the type of company. The Internet can be used for shopping hard-to-find special gift items.

(b) Research: The Internet can be used by scientists and researchers to download up-to-date results from around the world. Similarly, law firms can find a great deal of information on the net. Internet may also be used to assess the value of property owned by people.

(c) Education: The Internet may also be used by students for educational purposes. The concepts of distance learning and virtual classrooms have emerged due to the internet only.

(d) Travel: The Internet can be used for getting tourist and event information. Travelers can find weather conditions, maps, transportation schedules, and museum hours online.

(e) Marketing and sales: The Internet can also be used for selling products over the net. The emergence of e-commerce and virtual shops has been the outcome of the internet only.

(f) Love: People may find romance on the Net. The Internet may also be used for matrimonial purposes.

(g) Investing: The Internet may be used to buy and sell shares online.

(h) Patient-doctor relationship: The Internet may be used by doctors and patients to keep up-to-date with the Patent medical findings, share treatment experience, etc.

(i) Reduce discrimination: The person on the net is identified in the way in which he presents himself through the keyboard. If what he says sounds like an intelligent, interesting person, that's what he is. It does not matter how old he is, or how he looks like. Physical disability does not matter.

(j) Finding people: Directory services may be searched to find the phone numbers of people.

(k) E-mail: E-mail facilitates the electronic transfer of information including messages, texts, graphics, and pictures between two or more selected individuals. E-mail ensures the delivery of messages almost instantaneously.

(l) Games and gossip: The Internet can be used by play games. Also, it can be used to converse with people all over the world. Chating enables one to have an online conversation with two or more people at the same time.

(m) Exchanges and auctions: The Internet can be used to exchange goods and for auctioning of goods-directly to the consumers who need them. Moreover, the internet provides a perfect marketplace with no single buyer or seller can affect the price of a product because all buyers and sellers have access to the same information and can have the information at the same time.

(n) Advertising: The Internet may also be used for advertising a product or service.

(o) E-Government: The quality and efficiency of information and operations provided by the Government can be improved with the use of the internet. Citizens can receive information and forms online and can also file applications for renewal of licenses, etc. online.

(p) Entertainment: The internet may be used for entertainment purposes also e.g. listening to music or for watching a film.

(q) Conferencing: Computer conferencing can be established through the Internet. In this, person-to-person interaction can be held through the use of computers located in different places.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

COMPUTER SYSTEM SECURITY AND DATA PROTECTION

WHY IS INFORMATION SECURITY IMPORTANT - NEED FOR INFORMATION SECURITY

Although all the assets of an organization are subject to lose, damage, and destruction from various causes, information systems tend to be particularly susceptible to these dangers for several reasons, such as:-

- The components of an information system are comparatively more fragile.
- Computer hardware can be damaged more easily.
- Data files are also extremely fragile compared with most other assets of the organization.
- Computer systems are likely to be targets of workers, protestors, and even criminals.
- Decentralization of facilities and the use of distributed processing have increased the difficulty of protecting information and computers.

The failure of security of information system may cause financial losses or intangible losses. The unauthorized disclosure of sensitive information is an intangible loss. Threats to information systems may come from internal or external sources in the form of intentional or accidental threats. Knowing about major potential threats to information systems is important, but understanding ways to defend against these threats is also criticaL Security of information system is not a simple or inexpensive task because of the following reasons:-

- Numerous potential threats exist.
- Decentralization of facilities and the use of distributed processing.
- Many individuals control information systems.
- Computer networks may be outside the organization and difficult to protect.
- Rapid technological changes make some controls obsolete as soon as they are installed.
- Many computer crimes may go undetected for a long period, so it is difficult to learn from experience.
- People tend to violate security procedures because the procedures may be inconvenient.
- Punishment for computer crimes is not enough to deter them from committing crimes in the future.
- With the advent of the internet, many computer crimes can be committed from any part of the globe, without computer knowledge.
- Information security involves cost. It is difficult to conduct a cost-benefit justification for controls before an attack occurs since it is difficult to assess the risk of a hypothetical attack. Therefore, most organizations simply cannot afford to protect against all possible hazards.

WHAT IS INFORMATION SECURITY?

Information security is not simply computer security. Primarily, computer security relates to securing computing systems against unwanted access and use, whereas information security also includes issues such as information management, information privacy and data integrity. It includes personnel security and policies, steps taken for effective build ups, and physical integrity of computing facilities. Effective information security should include the following :-

- Staff assigned to information security tasks.
- Training all personnel in information security issues and procedures.
- Specific policies related with information privacy, physical security of equipment, and computer security procedures.
- Physical security plans.
- Data integrity measures.
- Level of access to data or equipment, and monitoring for different types of access.

The major objectives of information security are :-

- To reduce the risks of system and organization ceasing operation.
- To maintain information confidentiality.
- To ensure the reliability and integrity of data resources.
- To ensure the uninterrupted availability of data resources and online operations.
- To ensure compliance with national security laws and privacy policies and laws.
- To ensure that the system performs as intended.

CONTENTS OF INFORMATION SECURITY PLAN

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

When planning for system security, the following basic elements of security must be considered and included in the security plan:-

Information integrity: Integrity protects data from corruption. Integrity is concerned with the reliability, accuracy, and management of data.

Access-control / Authorization: Provides a means of enforcing authorization to use system resources.

Confidentiality: Ensures that a user or process other than those that are authorized cannot access information.

Authentication: A method of uniquely identifying a user, machine, or application and verifying its identity; includes a clearly stated policy on passwords.

Non-repudiation: The ability to prevent denial that a message has been sent or received or an action taken, the message or action could only have been produced by the sender.

Availability: Continuity, durability, recovery, and consistency of data.

PROTECTING COMPUTER-HELD INFORMATION

The value and importance of information may vary in an organization. For example, information about product prices may be less important than transactional information. The transaction information, customer identity, account detail of employees require greater protection than other information. Moreover, these days organizations keep valuable information on a computer, therefore, protecting the information which is held electronically is a major concern for the management. More specifically, in the process of designing a security framework, we need to seek the answers to the following broad questions (also known as ground Rules):-

- What is the information which needs to be protected and Where it is located?
- How it can be protected? We also need to know the business value of the information held and what difficulty would arise to recreate, if it were damaged or lost.
- Who is authorized to access the information and what privileges (scope of access rights) will be given to the user to do with the information?
- When it should be readily made available and when should it become unavailable.

The answer to the above set of questions (rules) assists the management in designing a security framework and implementing security procedures in an organization.

An organization can use two primitive types of protection as a security measure, these are: Preventive and Restorative.

PREVENTIVE INFORMATION PROTECTION

This type of protection is based on the use of preventive access controls. Preventive access controls are deployed to stop an unwanted or unauthorized activity from occurring. Preventive access controls can be grouped into:

Physical access control: Physical access control is a physical barrier deployed to prevent direct contact with the system. Examples of physical access controls include guards, fences, motion detectors, locked doors, sealed windows, lights, cable protection, laptop locks, wipe cards, clogs, CCTV, alarm, etc.

Logical/Technical access controls: A hardware or software mechanism used to manage access to resources and systems. It also protects these resources. They are the same as technical access controls. Examples of logical or technical access controls include encryption, smart cards, passwords, biometrics, protocols, firewalls.

Administrative access controls: Administrative access controls are the policies and procedures defined by an organization's security policy to implement and exercise overall access control. Examples of administrative access controls include hiring practices, background checks, data classification, security training, vacation, history reviews, work supervision, personnel controls, and testing.

The organization's information security policy through the associated informal security standards assigns the duties and responsibilities to the employees.

INTERNET SECURITY

Internet security refers to securing communication over the internet. It includes specific security protocols such as:

- Internet Security Protocol (IPSec)
- Secure Socket Layer (SSL)

Internet Security Protocol (IPSec): It consists of a set of protocols designed by the Internet Engineering Task Force (IETF). It provides security at a network level and helps to create authenticated and confidential packets for the IP layer.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Secure Socket Layer (SSL): It is a security protocol developed by Netscape Communications Corporation.). It provides security at the transport layer. It addresses the following security issues:

- Privacy
- Integrity
- Authentication
- Threats

Internet security threats impact the network, data security, and other internet-connected systems. Cybercriminals have evolved several techniques to threaten the privacy and integrity of bank accounts, businesses, and organizations.

Following are some of the internet security threats:

- Mobile worms
- Malware
- PC and Mobile ransomware
- Large scale attacks like Stuxnet that attempts to destroy infrastructure.
- Hacking as a Service
- Spam
- Phishing

Email Phishing

Email phishing is the activity of sending emails to a user claiming to be a legitimate enterprise. Its main purpose is to steal sensitive information such as usernames, passwords, and credit card details.

Such emails contain a link to websites that are infected with malware and direct the user to enter details at a fake website whose look and feels are the same as the legitimate one.

Cyber Law (IT Law) in India

Cyber Law also called IT Law is the law regarding Information-technology including computers and the internet. It is related to legal informatics and supervises the digital circulation of information, software, information security, and e-commerce.

IT law does not consist of a separate area of law rather it encloses aspects of contract, intellectual property, privacy, and data protection laws. Intellectual property is a key element of IT law. The area of software license is controversial and still evolving in Europe and elsewhere.

According to the Ministry of Electronics and Information Technology, Government of India :

"Cyber Laws yields legal recognition to electronic documents and a structure to support e-filing and e-commerce transactions and also provides a legal structure to reduce, check cyber crimes."

Importance of Cyber Law:

- It covers all transactions over the internet.
- It keeps eye on all activities over the internet.
- It touches every action and every reaction in cyberspace.

Area of Cyber Law:

Cyber laws contain different types of purposes. Some laws create rules for how individuals and companies may use computers and the internet while some laws protect people from becoming the victims of crime through unscrupulous activities on the internet. The major areas of cyber law include:

- Fraud
- Copyright
- Defamation
- Harassment and Stalking

COMPUTER FRAUDS

Computer frauds may be defined as intentional, illegal acts to cause damage to the computer. Computer frauds can be of the following types:-

(i) Theft of data

(ii) Inappropriate use of data (e.g., manipulating inputs)

(iii) Theft of mainframe computer time

(iv) Theft of equipment and/or software

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(v) Destruction from viruses and similar attacks
(vi) Deliberate manipulation in handling, entering, processing, transferring, or programming data
(vii) Theft of money by altering computer records
(viii) Damage to computer resources
(ix) Labour strikes, riots, or sabotage.

Internal Computer Threats

Computer frauds may be committed by outsiders by penetrating a computer system (frequently via communication lines) or by authorized insiders by misusing their authorization. Outside people who penetrate the computer system, are called hackers. A cracker is a malicious hacker who may represent a serious problem for a company. There is no single profile for computer criminals. They range from application programmers and clerical personnel to managers and accountants. In general, perpetrators of computer crime can be divided who two categories.

Internal threats - Employees use their knowledge of how a business operates to identify opportunities for theft or sabotage and to obtain easy access to the resource they need for their criminal activity.

Outside threats - Outsiders often have a more difficult task of learning how to penetrate a system without having easy access to information about how it works. Hackers commit computer fraud, not for their gains, but they commit computer crime for the 'fun' or intellectual challenge of breaking into a system.

The advent of computers has given rise to several frauds which were not possible earlier in the past. The common computer frauds and the measures that can be taken to prevent them have been explained below:

(a) Theft of money - Infrastructure of information technology can be used to steal money from the organization. For example, a bank employee may transfer amounts from inoperative accounts of various customers to an account opened in a fictitious name. These amounts are too small to raise any alarm, but over some time, the employee accumulates a substantial amount. To prevent such crimes, procedures must be developed to detect discrepancies in transactions and files. Financial records should be audited at regular intervals. Every effort should be made to hire only trustworthy employees. The performance of software should be independently checked to ensure its performance. Moreover, a programmer who develops the program should not be the same person who tests the program to ensure illegal modification of programs.

(b) Theft of data - Like any other financial asset, data is a valuable resource of an organization. Computer systems store data regarding the organization's operations, which is often of interest to competitors. To prevent the improper use of data, data security measures should be taken. One of such measures is the use of passwords to access the computer system. Retrieval or change of data should be allowed by the system to the authorized users only. Another security measure is to encrypt the data to an unintelligible form. Encryption ensures that the hacker will not be able to understand the encrypted data even if he has been able to illegally gain access to the system. Data encryption also ensures the secrecy of data transmitted over communication channels. A hacker may intercept the data for illegal purposes by tapping the communication channels, but would not be able to interpret the encrypted data. Theft of data can also take place by stealing the data stored on removable storage media such as floppy disks. To prevent such fraud, policy should be made that removable storage media may not be taken out from the organization's premises and that such media must be kept under lock and key, while not in use. Moreover, data stored on removable media can be encrypted to ensure that only the persons with the key may be able to view the data.

(c) Theft and destruction of hardware - Hardware can be physically stolen or damaged. With the downsizing of computers, it has become relatively easy to steal the computer itself. To prevent the theft or destruction of hardware, measures such as keeping the equipment in locked rooms, allowing access only to authorized persons in the computer room, placing fire extinguishers to protect the hardware in case of fire, etc., should be taken. To prevent the theft of personal computers, a security cable may be installed, which is often attached to a desk.

(**d) Illegal copying of software -** Just as books, most software is copyrighted. It is illegal to make a copy of the software without the permission of the owner of the copyright. But at times, people make copies of the licensed software. To prevent the illegal copying of software, a policy should be developed that illegal copying of software

by employees would be punishable. Moreover, employees should not be allowed to take software disks out of the premises of the organization.

(e) Destruction of data and software - Instead of stealing the data of software, a disgruntled employee might try to physically damage disks or tapes. Similarly, a hacker may try to erase programs or data electronically. Data

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

and program destruction may also be caused by a computer program called a virus. A virus is created by a hacker who puts the virus on floppy disks or a hard disk of a computer used by several people. When these viruses are activated, they destroy programs and data. To prevent data and program theft, backup copies should be made of all data and programs. These backup copies must be stored in a fireproof vault at a location away from the organization. To prevent the destruction of programs and viruses, it should be ensured that the disk is not copied from an unreliable source. Anti-virus software can be used to search disks for the presence of viruses.

(f) Theft of mainframe computer time - Another type of computer fraud is the theft of mainframe computer time. Employees may use the computer system for their purposes e.g., sending personal e-mail. This can be prevented by tracking the activities of the users.

External Threats

The main external threats may take the following form:-

- Interception of e-mail
- Removal of information
- Interception of e-payment
- Transmission of virus
- Destruction of system integrity
- Interference with web pages
- Tampering with computer system source code document
- Transmitting obscene material in electronic forms
- Breaching confidentiality of electronic documents.

Why should Business Take Computer Fraud Seriously

Adequate security measures should be taken to protect against computer frauds because:-

(a) Some of the information systems are critical in the day-to-day operations of an enterprise and any damage to them brings the entire enterprise to a halt.

(b) Extensive use of networks and data communications, especially the internet, has increased the scope of frauds. Data sent over communications channels can be intercepted, unauthorized access to a system can be gained through a terminal or a PC, and even the internet can be used for intercepting data or accessing a company's computer system.

(c) Electronic commerce systems have also increased the scope of the fraud, because, in e-commerce, payments are made through credit cards.

(d) IT infrastructure resources are more vulnerable to the risks of security and the implication of injury to these resources has a great impact on the organization because most of them cannot be restored to their original form.

Computer Fraud and Abuse Techniques

There are many devices for entering into computer systems and creating havoc. These devices involve programming tricks. Some of these are listed below:

Computer Fraud technique	Description
Data diddling	Changing of data going into or out of a computer. For example, an employee breaks into the payroll system and changes the number of days worked, thereby increasing his wage bill.
Cracking	Breaking into a computer system to do damage or commit a crime. Those individuals who are knowledgeable enough to gain access to the computer system without authorization are referred to as hackers.
Salami slicing	When a person transfers small amounts from financial accounts to a personal account. For example, a bank employee transfers small amounts from various customer accounts in an account opened in a fictitious name. These amounts are too small to raise any alarm but over a while, the employee accumulates a substantial amount.
Social Engineering	Employees are persuaded to reveal passwords and other information needed to access the system by posing to be an authorized user (e.g. by posing as a journalist, telephone company employee, or co-worker). The information is then used to break into a

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

	computer system or to steal company equipment, etc.
Spoofing	A scam is used to steal passwords by using a program that duplicates an organization's login screen. Spoofing is the act of pretending to be something other than what you are. It consists of replacing the valid source and/or destination IP address and node numbers with false ones. It grants attackers the ability to hide their identity through misdirection. Spoofing is employed when an intruder uses a stolen username and password to gain entry; when an attacker changes the source address of a malicious packet, or when a hacker assumes the identity of a client to fool a server into transmitting controlled data.
Virus	The destructive program disrupts the normal functioning of computer systems. They can reproduce themselves. Boot sector viruses attach themselves to that section of a hard or floppy disk which allows the user to start the computer or boot up. They are most often spread through the use of an infected floppy disk. File infector viruses attach themselves to files with certain extensions, such as .doc or .exe. Some viruses are a combination of boot sector and file sector viruses, and many of these can change to fool anti-virus programs. Viruses are transmitted through e-mail spread when an unsuspecting recipient of an e-mail message opens the message or an attachment to the message. Usually, such e-mail viruses can then send copies of themselves to everyone in the victim's address book, thus spreading through networked computers at an alarming rate.
Worm	Is designed to copy and send itself, spreading rapidly throughout networked computers. Eventually, it brings computers to a halt simply by clogging memory space with the outlaw code, thus preventing normal function.
Trojan horse	Trojan horse is a hidden malicious code that could alter or delete the information of a client computer or perform any other unauthorized function. When a Trojan Horse is planted in a computer, its instructions remain hidden. The computer appears to function normally but it is performing underlying functions dictated by the instructive code.
Logic or time bombs	They also do not disrupt normal computer function and lie in wait for unsuspecting computer users to perform a trigging operation. Time bombs are set off by specific dates, such as the birthday of a famous person. Logic bombs are set off by certain types of operations, such as entering a specific password, or adding or deleting names and information to and from certain computer files.
Piggybacking or shoulder surfing	The act of standing in line behind a card user at an ATM and memorizing the card's PIN. The stolen cards can then be placed on counterfeit access cards and used to draw cash from the account.
Carding	Involves stealing of credit card numbers online to be used to charge goods purchased against victim's account.
Spamming	The term spam describes unwanted e-mail. It may be in the form of advertisements from vendors. It may also be in the form of unrequested messages with viruses or Trojan horses attached. It is a type of denial of service attack. As the level of spam increases, locating or assessing legitimate messages can be difficult. In addition to the nuisance value, spam consumes a significant portion of internet resources (in the form of bandwidth and CPU processing), resulting in overall slower internet performance and lower bandwidth availability. Spams fill up the storage space and prevent legitimate messages from being delivered.
Sniffer/ Snooping	Any activity which results in a malicious user obtaining information about a network or the traffic over that network.
Brute force and dictionary attacks	These attacks are waged against passwords database files or an active login prompt. Brute force is an attempt to discover user accounts by systematically attempting every possible combination of letters, numbers, and symbols. A dictionary attack is an attempt to discover passwords by attempting to use every possible password from a predefined list of common or expected passwords.
(DoS attack) Denial of service attacks	Denial of service attacks prevent the system from processing or responding to legitimate traffic or requests for resources and objects. The most common of these attacks are

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

	transmitting so many data packets to a server that it cannot process them all.
Data Leakage	Removing copies of confidential data from within the system without any trace.
Trap-door	Using systems escapes gaining illegitimate access to a computer system.
Piggybacking	Using access permission belonging to someone else to gain entry to a computer system.
Superzapping	By-passing all security systems through specialized software packages.
Scavenging	Searching trash cans for printouts, memos, carbons, etc., which contain confidential information not intended for public distribution.
Ware & Trading	Exchanging or selling pirated software.
Software piracy	Copying computer software without the permission of the publisher.
Eavesdropping	Listening to private voice or data transmissions often using wiretap.
Internet misinformation	Spread of false or misleading information about companies with the use of the internet.

Preventing Computer Frauds

As the IT infrastructure resources are more vulnerable to security risks, any organization needs to take adequate security measures to prevent computer fraud. The following security measures can substantially decrease the potential for fraud and the resulting losses if any:-

(I) Users should be trained to be aware of security concerns and to understand how these concerns are related to rules and procedures.

(II) Every employee who uses a computer or is at all involved with transaction processing should be familiar with the security issues.

(III) Users should be trained to be aware of some of the signs of suspicious activity and the procedures for reporting that activity.

(IV) Training should be used to reinforce the value of security consciousness in all employees.

(V) Training and awareness programs should stress the individual employee's responsibility for information security, proper use of sensitive information, and procedures for reporting potential violations.

(VI) Training should cover the following issues:

- The organization's data backup policies
- The type of data that should be encrypted
- How the data encryption keys are managed
- What type of data may be shared with colleagues
- What type of data is available to the public
- Employees should be informed of the consequences of unethical behavior
- Employees should be educated on security issues
- Employees should be taught not to give confidential information over the telephone without establishing the authenticity of the caller
- Awareness of the employee's security responsibilities as well as those of his co-workers.

(VII) Develop a strong system of internal controls - Controlling access to any resource in a secure system involves two entities, firstly the user or process that requests to access a resource and secondly, the resource a user or process wants to access. Access can mean reading from or writing to a resource. Internal controls use access rules to limit the access by a user or process to a resource.

(VIII) Protect network cabling - Different types of network devices and technologies are used with different types of cabling. Each type of cable has its unique useful lengths, throughput rates, and connectivity requirements. Intruders may use telephone lines to transmit viruses or to steal access and destroy the data. Therefore, it is very essential to protect these telephone lines. These may be protected in various ways, such as attaching an electronic lock and key to the line.

(IX) Control sensitive data classification - Data classification is the primary means by which data is protected based on its secrecy, sensitivity, or confidentiality. All data should not be treated the same way. Some data needs more security than others. Securing all data in the same manner at a low-security level would make the sensitive data easily accessible. However, securing every data at a high-security level would be too expensive and access to non-critical data will be restricted. Therefore, it is important to classify the data according to its importance and confidentiality. The objective of the data classification process is to provide adequate security mechanisms for the storage, processing, and transfer of data based on assigned labels of importance and sensitivity. Data classification involves the evaluation of data.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Based on such evaluation, a data classification label is assigned to the data, such as top secret, secret, confidential, sensitive but unclassified and unclassified. The sensitive data must be controlled on a top priority basis. Unauthorized disclosure of sensitive data can cause severe damage to the organization. Allocation of resources to protect the data and control access to it can be made based on this classification. The objective of data classification is to ensure that the best set of security controls is applied to sensitive data. Moreover, data classification should be viewed continuously and should not be considered as a one-time process. Following measures may be taken to secure the sensitive data:-

(a) An appropriate clearance level should be assigned to all sensitive data. Only those personnel who are cleared for that level of sensitivity should have access to the sensitive data.

(b) The sensitive data should be regularly inspected regarding its storage and it should be audited.

(c) A chain of custody for highly sensitive data should be maintained. Each person that takes possession of the data should sign or otherwise verify that he has done so. Regular audits of a chain of custody logs ensure that employees are using them.

(d) Access codes and combinations of sensitive data should be stored in sealed envelopes in safes, or as encrypted data. The regular audit ensures that codes are not compromised and that the appropriate person is properly destroying codes when the situation dictates it.

(e) Sensitive data should be properly destroyed when not required. It should be ensured that documents marked for destruction are destructed.

(X) Protect the system from viruses - Viruses are destructive programs that disrupt the normal functioning of computer systems. They differ from malicious code in the sense that they can reproduce themselves. Viruses, often damage a computer system by erasing files on the hard drive or by slowing computer processing. All computers connected to networks are vulnerable to virus infections. To prevent viruses, virus detection and virus protection programs should be installed. Virus protection programs search for viruses trying to infiltrate the system. Virus protection gains greater significance primarily because the source of the virus is either unknown or is not under the control of the business enterprise. Virus detection software can be used for continual scanning of the hard disk(s), memory, and network interfaces for known viruses and to treat the system by removing the malicious code. Both proxy servers and e-mail systems should be running virus detection software. Since new viruses are continually being developed, it is important to regularly update virus detection software to the latest version to ensure detection of any recent additions to the list of known viruses.

(XI) Encrypt data and programs

Encryption is the process of encoding readable information into a scrambled of data that cannot be interpreted without the knowledge of a correct key.

Encryption is the technique for keeping sensitive data and programs securely locked up. By translating data into a secret code, encryption ensures that anyone who can access that data by some means illegally, would not be able to understand the contents of data.

(XII) Control laptop computers - Computer theft is second to auto theft. Since laptops are becoming lighter and smaller, thieves have more temptation to steal them. The data stored in a laptop can be more valuable than the laptop itself. To secure the laptop, the following measures may be taken:

- Keep a record of a serial number and detailed description, which helps the police in tracing a stolen laptop. It also helps in making an insurance claim.
- Name and phone number should be marked on laptops with the use of etching devices.
- Laptop must be insured.
- The data on the laptop should be backed up at frequent intervals to a separate source, especially before traveling.
- It should be ensured that the laptop does not boot up without a password.
- Data on the laptop may be encrypted to ensure its misuse, even if the laptop is stolen.

(XI) Hacking

Hacker is a person who gets into a computer system to look at files, which are private and do not belong to him or her. When a person does this it means that they probably have logged onto a computer system as though they are someone else.

Hacking is defined as 'unauthorized access to data held on a computer system: Once hackers break into a computer system this way they may:

- ✓ look at data;
- ✓ copy data;

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

- ✓ change data.

Usually, a hacker has worked out the password and user identification of a person who is authorized to use the computer system. Some people have special monitoring software, which can obtain your password as you type it on the keyboard of your computer. Even though your password is not displayed on your computer monitor as you type it in, it may still be displayed on the computer monitoring your input. If these people intercept your telephone link to a banking system, they can obtain your bank details, password, and other security information. Once they have all this information they can make transactions on your bank account by pretending to be you.

To fight against this type of hacking, banks have had to take more care. Some banks now have other security procedures as well as user ids and passwords. Sometimes data is coded specially so that it is impossible to read it as it is being transmitted. This is known as encryption. Once received the data is then decoded so that it can be read.

People who hack into computer systems do not always want to steal money. Sometimes the hacker may be a young computer enthusiast who hacks into a computer system to prove that it can be done.

Improve Detection Methods

Recent advances in Information Technology have removed many of the physical barriers that inhibited information flow in the past. Never has information been so readily available as is now. However, as information around the globe becomes more accessible, it has also become more vulnerable to misuse. A class of potential intruders characterized by the ability to mount sophisticated attacks, with grave potential risks, including numerous individuals of questionable ethics, would require strong, comprehensive security countermeasures. One of these countermeasures is to improve the fraud detection methods and detect the fraud at the earliest. Detection enables the identification of a security breach such as unauthorized access, corruption of data, and unauthorized actions. The objective of fraud detection is to determine illegal activities that may be occurring and provide sufficient information about the activities and their source so that corrective action can be taken. The following measures ensure the early detection of frauds:-

(a) Conduct frequent audits - By analyzing transaction data collected over long periods (e.g. a few months), from several sources, the cumulative effects of the intrusions may become apparent. If adequate records are kept, it may be possible both to detect the intrusion and to identify the intruder. A security audit trail is an event log that captures such security-relevant information. A security audit trail records all security alarm notifications it has issued and the events that triggered such alarms. It may also record events such as association setup attempts and access to certain types of data, which may reveal some unexpected patterns.

(b) Use a computer security officer - Most computer frauds cannot be detected by auditors. Therefore, a security officer may be appointed. He should work independently of the information system function.

(c) Use security services - Specialized security services may be hired to protect the computer from the attacks discussed in the earlier section. These agencies have varying coverage, they protect against different threats from intruders with different levels of technical sophistication and tenacity. But the drawback of this measure is that the weaknesses in the system may be exposed to outsiders.

(d) Use fraud detection software - A security audit trail is especially valuable when a persistent low-level intrusion is suspected. The security administrator may then configure the security audit trail to capture sensitive events, based on the type of suspected intrusion. On this basis, the pattern of fraud commitment practices can be made with the use of the software. Accordingly, preventive actions may be taken.

Reduce Fraud Losses

The various measures to prevent computer frauds, decrease the potential for fraud and the resulting losses. But these measures are not foolproof. Despite very strong security measures applied by the organization, IT infrastructure resources are still vulnerable to security risks. Therefore, an important strategy should be developed to minimize the potential fraud losses. This strategy should include the following:-

(a) Current back-up of all program and data files should be kept in a fireproof vault situated at a place away from the business location. This backup serves not only the purpose of disaster recovery but also helps in the detection of fraud.

(b) It is essential to develop a security policy. The objectives of the policy should be clearly defined. A well-defined security policy prevents the occurrence of fraud, helps in early detection of fraud, mitigates the amount of loss due to fraud, helps in restoring the normal operations after the disaster, and also helps in establishing accountability for the damage to the system.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(c) Adequate insurance coverage should be taken to cover the risk due to fraud. It should be ensured that the insurance policy is duly renewed on the due date. Moreover, insurance coverage should be reviewed at frequent intervals to ensure its adequacy.
(d) Viruses spread very quickly between computers that are connected to a network. Because they can cause so much damage to computer systems and files, it is important to take precautions against being 'infected' by a virus. There are several ways of doing this.

- ⇨ When opening the e-mail, be sure not to open messages from unknown senders. They may have a virus hidden in its attachment or text messages.
- ⇨ Check your computer regularly for viruses using anti-virus software.
- ⇨ Do not accept free copies of software.
- ⇨ Always virus check your floppy disk when you have been using a different machine.
- ⇨ Make sure you regularly back up your files just in case your antivirus software does not detect the latest virus.

(e) To protect a computer system from unauthorized access it is important to :
log off correctly every time you leave your machine;

- ⇨ choose a password that is not obvious;
- ⇨ keep your password private;
- ⇨ change your password frequently.

Prosecute and Imprison Fraud Perpetrators
Most the computers frauds are not reported due to the following reasons:-
(a) It is relatively difficult to detect computer fraud. Only a fraction of all computer crimes are detected.
(b) Companies affected by computer crimes generally do not report these crimes due to fear of adverse publicity.
(c) Laws governing computer crimes are not adequate. It is very difficult to produce the evidence in the Court of Law. Moreover, the penalties for various crimes are not large enough to discourage fraud perpetration. Moreover, enforcement officers and judges lack the computer skills required to investigate, prosecute and evaluate computer crimes.

LEGAL ISSUES RELATING TO DATA PROTECTION
As we know these are some of the problems caused by people who break the law by copying software and spreading viruses. There are also those people who break into computer systems and gain access to information that is private. These people are known as hackers. Therefore we will look at the data protection legislation regarding software copyright viruses and hacking.

Software copyright
It is a criminal offense to steal software or to copy software without the permission of the person who owns the copyright on a piece of software. It is also an offense to run copied software on your computer. The law, which makes this a criminal offense, is the Copyright, Designs and Patents Act, 1989. To ensure that software copies are legal, companies usually issue a license. Individual licenses like this usually come packaged with any software you buy. Sometimes companies also issue a special number to type in when you install your software. Businesses however may buy many licenses and only one copy of the software. These are called multiple licenses and allow more than one user to use the copy of the software. Multiple licenses are usually available at reduced cost and sometimes rather than specify how many licenses are needed it is possible to buy a 'site license: A site license means that anyone using a computer at a given site (such as a school or any business premises) can legally use a copy of the software. Users cannot make copies of software away from the site. It is only licensed for those machines at the site.
Copying software for your use without purchasing a license is known as 'software piracy'. Sometimes people copy the software and pass it on to others, sometimes even selling it. They may even issue a copy of the license agreement. All of these examples of copying software are against the law. Software costs a lot of money to develop and copying software costs software companies lost income.
Copying software illegally is so common that it is estimated that more than half of the software running, on computers in few countries is not licensed. It is very tempting for employees using the software at work to illegally copy the software and take it home for their use. It is also possible for employees to distribute illegal copies of the software to colleagues. Many people who do this are sometimes unaware that they are breaking the law.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

PREVIOUS YEAR QUESTIONS

Short Answer Type Questions

1. What purpose does a modem serve in a data communication system?
2. What do you mean by internet security?
3. Differentiate between downloading and uploading information.
4. An attempt to make a computer resource unavailable to its intended users is called _____.
5. The OSI model has _____ layers.
6. _____ is a one way communication.
7. Air is a _____ transmission media.
8. _____ attaches itself automatically with other files.
9. FTP is a _____ protocol.
10. Distinguish between LAN and MAN.
11. What do you mean by wild card character?
12. A telephone network is an example of which network.

Focused-short Answer type Questions

1. With the help of suitable diagrams, explain the concept of simplex, half-duplex, and full-duplex communication.
2. Define the computer virus. How do they infect and spread in an ac computer system? Explain at least three categories of viruses based on their mode of existence. Give three examples of computer viruses.
3. What is a Web Browser? Explain how a Web Browser works. Also, explain the types of Web Browsers.
4. Explain the various uses of the Internet.
5. Write about the various application of RFID technology.
6. In which mode FTP, the client initiates both the control and data connections.
7. What is the Home Page of a website?
8. Which layer of the OSI model performs data compression operation?
9. What is a DOS attack?
10. Explain the difference between TCP & UDP.
11. Describe snooping mechanism. How does it differ from sniffing?
12. What is WWW?

Long Answer Type Questions

1. What is a Repeater? Differentiate between repeater and hub.
2. Explain about Intranet and Extranet.
3. What is a web browser? List various features of a web browser. Explain how the Internet works?
4. Describe the features of several network topologies.
5. Explain the OSI architecture model.
6. Describe cyberlaw and internet security legislation.
7. List out various communication connectivity mechanisms on the Internet.
8. Describe the basics of working on the Internet.
9. Explain the difference between TCP-IP and OSI architecture model.
10. Describe various types of Attacks.
11. What is the meaning of the term “topology” in the context of computer networks? Explain different topologies giving their advantages and disadvantages.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Biju Patnaik Institute of IT & Management Studies

Lecture Notes

on

IT FOR MANAGEMENT

Prepared by,

Mr. Amiya Ranjan Das, Asst. Prof.(IT)

BIITM, Bhubaneswar

DATA REPRESENTATION: COMPUTER ARITHMETIC AND CODES

A computer can understand the positional number system where there are only a few symbols called digits and these symbols represent different values depending on the position they occupy in the number.

The value of each digit in a number can be determined using −

- The digit
- The position of the digit in the number
- The base of the number system (where the base is defined as the total number of digits available in the number system)

REPRESENTING DATA

The different data types can be represented in the computer using bit patterns.

Text and Numbers: Text is a sequence of symbols used to represent thought or expressions in a language. For example, Human made reports contain symbols from the set of alphanumeric characters (such as alphabets A-Z, a-z, Numbers 0-9, +, -, ?, $, --). But computer cannot understand these, because computer is an electronic machine. Therefore, it is a bistable device means with two states "on" and "off". Thus computer deals with data converted into simplest form which can be processed electronically, that is binary form, where it substitutes "on" with “1" and "off" with "0". To store and process data in a binary form a coding scheme had been devised to represent alphanumeric and special characters as standardization methods. These are:

1. ASCII (American Standard Code for Information Interchange):

This code is most widely used with micro-computers. The code represents English characters as numbers. ASCII was established to achieve compatibility between various types of data processing equipment and is now used as the common code for microcomputer equipment. The highlights of this code are:-

(1) It is a 7-bit code, it can accommodate 128 (=27) different characters.

(2) In this, first three bits are zone bits and last 4-bits indicate the digit.

(3) This code is generally used in microcomputers, as they use 8-bit byte and ASCII is 7-bit code, therefore, it leaves leftmost first bit of each byte as zero.

(4) 31 Control non-printable characters.

(5) 0-9 numeric Characters.

(6) 7 printable Characters.

(7) A - Z upper case letters, a-: lower case letters.

Example: Write binary coding for the word LOVE in ASCII-7. Also find the number of bytes required to represent the same.

Solution:

L = 1001100 in ASCII-7 binary notation.

0= 1001111 in ASCII-7 binary notation.

V= 1010110 in ASCII-7 binary notation.

E= 1000101 in ASCII-7 binary notation.

Hence binary coding for LOVE in ASCII-7 is

1001100	1001111	1010110	1000101
L	O	V	E

There are 4 characters in the string LOVE. Since each character requires one byte in ASCII-7 representation, thus 4 bytes (= 2S-bits) are required for this representation.

(2) ASCII-8

It is extended version of ASCII-7, because it is 8-bit code, thus it allows 256 (=28) different characters. In this code first four bits indicate zone bits and last four bits are for digits.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Example : Write binary coding for the word LOVE in ASCII-8. Also find the number of bytes required to represent the same.

Solution:

L = 10101100 in ASCII-S binary notation.

0= 10101111 in ASCII-S binary notation.

V= 10110110 in ASCII-S binary notation.

E= 10100101 in ASCII-S binary notation.

Hence binary coding for LOVE in ASCII-S is

10101100 10101111 10110110 10100101

L O V E

There are 4 characters in the string LOVE. Since each character requires one byte in ASCII-S representation, thus 4 bytes (= 32-bits) are required for this representation.

(3) EBCDIC

This coding scheme was developed by IBM, extended from 6-bit BCD (Binary-Coded Decimal) code to an 8-bit code by adding 2-bits in zone bit, expanding zone to 4-bits. This binary coding scheme is known as Extended Binary-Coded Decimal Interchange Code (EBCDIC) generally, used with large computers such as mainframes. In this coding scheme, it is possible to represent 256 (=28) different characters. It allows a large printable and non-printable control characters (e.g. Cursor movement, spacing, etc.) along with 64(=26) characters. However, this code is not used in any computer other than mainframes.

NUMBER REPRESENTATION IN COMPUTER

Decimal number system is the most commonly used system. Computer designers prefer the binary number system because computers can work only with binary data at the machine level. To fulfil the needs of both the programmer and user, some coding systems have been devised. With the help of these codes, the communication becomes easier. Various computer codes in use are :

1. NUMBER SYSTEM

Number system are of two types:

i. Non-Positional Number System

In this system there are symbols such as I for 1; II for 2; III for 3; ..... ; IIIII for 5; and so on, each symbol represents the same value regardless of its position in the number and symbol is simply added to find out the value of a particular number. This concept was used in early days, when human beings used to count on fingers, stones, sticks, etc. for each object. Since it was difficult to perform arithmetic with such numbers, therefore, this led to the positional number development.

ii. Positional Number System

In this system, a set of symbols called digits, represent values depending on the position they occupy in the number. The value of each digit in a number is determined by:

(a) Face value of the digit

(b) Place value of the digit

(c) The base of the 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. Its value can be written as

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

(1 x 1000)+ (2 x 100)+ (3 x 10)+ (4 x 1)
(1 x 103)+ (2 x 102)+ (3 x 101)+ (4 x l00)
1000 + 200 + 30 + 4
1234
The number systems which are frequently used in computers are.

S.No.	Number System and Description
1	Binary Number System Base 2. Digits used : 0, 1
2	Octal Number System Base 8. Digits used : 0 to 7
3	Hexa Decimal Number System Base 16. Digits used: 0 to 9, Letters used : A- F

Binary Number System

Characteristics of the binary number system are as follows −

- Uses two digits, 0 and 1
- Also called as 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}

Octal Number System

Characteristics of the octal number system are as follows −

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

Hexadecimal Number System

Characteristics of hexadecimal number system are as follows −

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

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Example : Hexadecimal Number: $19FDE_{16}$

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}

Binary Arithmetic

Like in decimal number systems, binary arithmetic involves mainly four operations viz: : Addition, subtraction, multiplication and division. The operations of each are described below:

1. Addition

While performing addition on binary numbers following rules may be applied:

0+0=0

0+1=1

1+0=1

1 + 1 = 10 or 0 with carryover 1

Example:

```
Carries  10011          25
          11001              +19
         +10011          44
         101100
```

2. Subtraction

While performing subtraction on binary numbers following steps may be followed:

Step-1: Both the numbers should be of the same size. If not, then make both the numbers of the same size by adding leading zeros.

Step-2: Complement the number to be subtracted, by changing 1's to 0's and 0's to 1's i.e. the complement of 1100 is 0011.

Step-3: Perform binary addition as above.

Step-4: If, the result of the addition is increased by one digit means there is a carry of 1, then remove this (leftmost digit)and add to the final result.

Step-5: If the size of the result is not increased, then complement the answer and place a negative(-) sign in front of the number.

Example: 11001 - 111

Step-1. Make 111 of the same size as of 11001, by adding leading zeros such as 00111.

Step-2. Complement 00111, as 11000

Step-3. Add 11001 + 11000 = 110001

Step-4. Remove the leftmost digit of the answer and add it to the answer again

10001 + 1 = 10010 is the final answer (25 - 7= 18)

3. Multiplication

The multiplication of binary numbers follows the same convention as of decimal number system. The result may be obtained by sequence of additions and shifts.

Example :Multiply 1101 with 101.

```
    1101
  X 101
    1101
```


Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

0000
1101
1000001

4. Division

The division in binary number system is also same as of in decimal number system. It may be performed as sequence of subtraction and shifts.

CONVERTING FROM DECIMAL TO ANOTHER BASE

As decimal number may contain only integer or integer part along with fractional part, thus calculation shall be done in two parts, requiring following steps:

Integer Part:

Step-1: Divide the decimal number by the value of the new base.
Step-2: Record the remainder
Step-3: Repeat the step-1 with the quotient and then step-2, until the quotient becomes 0 or less than the value of the new base.

Fractional Part:

Step-1: Multiply the fractional part by the value of the new base.
Step-2: Record the integer part, if it exists, else record 0.
Step-3: Repeat the step-1 with the result of the previous multiplication and then step-2, until the fractional part becomes O. In case of infinite calculations, generally, 6 digits are taken.

Example : $(50)_{10} = (?)_2$
Here the new base is 2.
50/2 = 25 Remainder 0
25/2 = 12 Remainder 1
12/2 = 6 Remainder 0
6/2 = 3 Remainder 0
3/2 = 1 Remainder 1
Now, start writing in the order from the last obtained till the first remainder.
Thus Binary Equivalent = 110010

Example: $(50.625)_{10} = (?)_2$
Convert integer part i.e. 50 into binary as above = 110010
Convert fractional part 0.625 into binary as below:
0.625 X 2 = 1.25 take away mteger part and record = 1
0.25 X 2 = 0.50 take away integer part and record = 0
0.50 X 2 = 1.00 take away integer part and record = 1
Thus, binary equivalent of fractional part is 101.
Hence (50.625)10 = (1010010.101)2

Example: $(952)_{10} = (?)_8$
Here the new base is 8.
952/8 = 119 Remainder 0
119/8 = 14 Remainder 7
14/8 = 1 Remainder 6
Now, start writing in the order from the last obtained till the first remainder.
Thus Octal Equivalent = (1670)8
$(952)_{10} = (1670)_8$

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar

Example : $(952)_{10} = (?)_{16}$
Here the new base 16
952/16 = 59 Remainder 8
59/16 = 3 Remainder 11
Now, start writing in the order from the last obtained till the first remainder.
Thus Hexadecimal Equivalent =3B8
$(952)_{10} = (3B8)_{16}$

CONVERSION FROM BINARY TO HEXADECIMAL AND VICE VERSA

A binary number is easily converted to a hexadecimal number by dividing the bits of the binary number into groups of 4-bits. This is because of the fact that the maximum value of one digit is equal to the maximum value of four digits in binary. Therefore, the value of one hexadecimal digit is equivalent to four bits of binary. This notation, therefore, reduces the space for storage and time required for memory dumps about four times. Now, for converting each group of four bits into one hexadecimal digit in binary number system see the example given below:

Example : $(11110100001)_2 = (?)_{16}$
Divide these bits into group of 4 bits such as : 111 / 1010 / 0001
Converting each group into one hexadecimal digit to make 111 in hexadecimal an implicit leading zero may be added (i.e. 0111). Thus 0111 corresponds to 7 and A corresponding to 1010 and 0001 denote 1. This hexadecimal number is : 7 A 1_{16}
Now for quick conversion of hexadecimal to binary, each digit of hexadeci-
mal number be converted into its 4-bits of its binary equivalent.

Example : $(F3D)_{16} = (?)_2$
Now
$F_{16} = 1111$
$3_{16} = 0011$
$D_{16} = 1101$
Hence $(F3D)_{16} = (1111\ 0011\ 1101)_2$

CONVERSION FROM BINARY TO OCTAL AND VICE VERSA

In the same manner as done above, binary numbers are easily converted to octal (base 8) by using groups of 3-bits.

Example : $(11110100001)_2 = (?)_8$
Divide the bits into group of 3 : 11/ 110/ 100/ 001
Converting each group into one octal digit, in octal with the 3 corresponding to 11 with an implicit leading zero, the 6 corresponds to 110, 4 correspond to 100 and 1 corresponding to 001, this becomes:$(11110100001)_2 = (3641)_8$
However, at one time the preponderance of computers in current use mainframes and all popular personal computers - have word lengths that are divisible by four but not by three (16,32, or 64 bits), octal is almost never used now.

Prepared by,
Mr. Amiya Ranjan Das, Asst. Prof.(IT)
BIITM, Bhubaneswar