

CHAPTER 1

INTRODUCTION TO COMPUTERS

Learning Objectives of this chapter are-

Unit 1:

By the end of this Chapter, Student will be able to know:

- Different Generations of Computers and their evolution.
- Classification of computers.
- Features of computers, their advantages and limitations
- Basic components of a computer system.
- Types of Storage devices, their use and capacity.
- Types of RAM and their working.

Unit 2:

- Different types of input devices available and how to use them.
- Where to use the right kind of input device.
- Add-ons required for the input Devices.

Unit 3:

- Meaning of software
- Different classifications of software.
- Available system software in the market with their usage and limitations
- Utility of the software.

In this Chapter we shall discuss what we understand by the term 'computer', its functions and various generations through which computer technology has advanced. Various categorizations of computers according to their purpose and size etc. shall also be discussed in this study paper. We will also overview hardware and software requirements. Hardware

© **The Institute of Chartered Accountants of India**

Get Free SMS Updates on Mobile. Type ON<space>SIDRUTIYA and send it on 9870807070.



Information Technology

consists of the mechanical and electronic components, which one can see and touch. Computer hardware falls into two categories: processing hardware, which consists of the central processing unit, and the peripheral devices. The software comprises of system and application programs, the operating systems and various other general purpose software.

1.1 HISTORICAL DEVELOPMENT OF COMPUTERS

The modern computer with the power and speed of today was not a solitary invention that sprang completed from the mind of a single individual. It is the end result of countless inventions, ideas, and developments contributed by many people throughout the last several decades. The history of automatic data processing begins with Charles Babbage's attempt to build an automatic mechanical calculator at Cambridge, England, in 1830. By the 1930's punched cards were in wide use in large business, and various types of punched card handling machines were available. In 1937, Howard Aiken, at Harvard, proposed to IBM that a machine could be constructed which would automatically sequence the operations and calculations performed. This machine used a combination of Electro-mechanical devices, including relays.

First Generation computers : UNIVAC (Universal Automatic Computer) was the first general purpose electrical computer to be available and marks the beginning of the first generation of electrical computers. The first generation electrical computers employed vacuum tubes. These computers were large in size and required air conditioning. The input and output units were the punched card reader and the card punches. Because of the inherently slow speed of these input/output units, the power of the CPU was subjugated to their speed. IBM-650 was however, the most popular first generation computer and was introduced in 1950 with magnetic drum memory and punched cards for input and output. It was intended for both business and scientific applications.



Fig-1.1.1 Transistor

Second generation computers: These computers employed transistors (see figure 1.1.1) and other solid state devices. Their circuits were smaller than the vacuum tubes, and



generated less heat. Hence the second-generation computers required less power, were faster and more reliable. IBM 1401 was the most popular second-generation computer. There were two distinct categories of the second-generation computers for business and scientific applications. They employed magnetic tape as the input/output media. Second generation computers successfully displaced the unit record equipment on cost benefit grounds in many installations.

Third generation computers: These employed integrated circuits in which all the elements of an electronic circuit are contained in a tiny silicon wafer. The third generation computers are much cheaper and more reliable than the second-generation computers. They are speedier with much vaster capacity and admit connection of a wide variety of peripherals particularly magnetic disk units. They are based on the principles of standardisation and compatibility. The core storage of a given model of a computer can be expanded by adding modules and it still permits the use of order program. The third generation computers can be used for both scientific and business applications.

The third generation computers permit multi-programming which is interleaved processing of several programmes to enhance the productivity of the computer, time-sharing which is the use of the computer by several customers at a time, operating systems which optimise the man-machine capabilities and such data communications facilities as remote terminals. They also permit use of such high level languages as FORTRAN and COBOL. The mini computers are also one of the developments in the third generation computers.

Each generation of computers has an effect on the MIS centralization and decentralization issue. The first generation computers were high in costs and large in size; therefore information systems were sought to be centralized to serve benefits of hard ware economies. The second-generation computers were substantially cheaper and the trend was towards MIS decentralization. Third generation computers however, offered communication capabilities and the use of remote terminals and the trend was reversed to centralization.

Fourth Generation Information Systems : Fourth generation machines appeared in 1970's utilizing still newer electronic technology which enabled them to be even smaller and faster than those of the third generation. Many new types of terminals and means of computer access were also developed at this time.

One of the major inventions, which led to the fourth generation, was the large scale Integrated Circuit (LSI) The LSI is a small "chip" which contains thousands of small electronic components which function as a complete system. In effect an entire computer can be manufactured on a single chip of size less than 1/3 inch square. A single chip may perform the functions of the entire computer, calculator or control device. Research into future developments promises the manufacture of large computer systems with enormous memory



Information Technology

capacity on small chips. This will reduce the cost and increase the speed of new systems still further.

Micro computers : In July, 1977, at National Computer Conference in Dallas, Commodore Ltd. startled the computing world by announcing a fully assembled microcomputer in a single housing called the Personal electronic Transactor (PET) The machine consisted of keyboard, processor unit, CRT and built in cassette tape recorder for \$595. The programming language BASIC was built into the system. Thus, for less than \$600, a fully programmable, powerful computer system was available for home or personal use. Later in 1977, Radio Shack Corporation announced the TRS 80 computer.

The IBM family of personal computers: In [1981] International Business Machines (IBM) made its first appearance in the field of microcomputer with the announcement of the IBM Personal Computers. The term personal computer captured the notion that an individual can have her or his own computer. With the advent of IBM PC, computers had stepped out of large organisations and entered into the home. However, instead of adopting 8-bit microprocessor, IBM selected Intel 8088 - a 16 - bit microprocessor which made the IBM PC “an overnight success”. In [1983], IBM's first addition to the PC-family - XT model was introduced, which added a high capacity hard disk storage facility to the PC. In [1984], IBM introduced two new high powered models of PC viz Compaq Desk Pro, the first member of the PC family to have more basic computing power than the original PC and the IBM PC



Fig 1.1.3 IBM PC

AT model, which had a much greater computing speed than the PC and XT or even the new Desk Pro. When software vendors began to orient their products to the IBM PC, many microcomputer manufacturers created and sold clones of it. These clones called IBM PC compatibles, run most or all the software designed for the IBM PC. Therefore, whatever IBM does in the personal computer era has immediate and far-reaching effects on PC market. The successor to the IBM PC, the IBM personal system/2, or IBM PS/2 (introduced in 1987), have almost certainly become a milestone in PC history. With IBM's products, the



Introduction To Computers

microcomputer took its place as an important tool for use in solving the information processing needs of both large and small businesses.

Other Significant Contributions: Several other personal computers have established their place in PC history. Introduced in [1982], the Commodore-64 was significant because it signaled the buying public that powerful micros could be manufactured and sold at a reasonable cost \$599. In the same year, Compaq Computer Corporation bundled the equivalent of an IBM PC in a transportable case and named it the Compaq Portable. Thus began the era of the portable computer. In [1984], Apple Computer introduced the Macintosh with a very “friendly” graphical user interface - proof that computers can be easy and fun to use.

Microcomputers have many of the features and capabilities of the larger system. The cost of microcomputers has dropped substantially since their introduction. Many now sell a microcomputer for as low as Rs. 15,000. This reduction in cost will bring about a significant increase in the number of microcomputers in use. The major application for microcomputer lies in the field of industrial automation, where they are used to monitor and control various manufacturing processes. Their low cost and lightweight make it feasible to carry them on site or into a field or to package them with other portable equipments as part of larger system.

The second decade (1986- present) of the fourth generation observed a great increase in the speed of microprocessors and the size of main memory. The speed of microprocessors and the size of main memory and hard disk went up by a factor of 4 every 3 years. Many of the mainframe CPU features became part of the microprocessor architecture in 90s. In 1995 the most popular CPUs were Pentium, Power PC etc. Also RISC (Reduced Instruction Set Computers) microprocessors are preferred in powerful servers for numeric computing and file services.

The hard disks are also available of the sizes up to 80 GB. For larger disks RAID technology (Redundant Array of Inexpensive Disks) gives storage up to hundreds of GB. The CDRoms (Compact Disk-Read Only Memory)and DVDs(Digital Video Disks) are have become popular day by day. The DVDs of today can store up to 17 Giga bytes of information.

The computer networks came of age and are one of the most popular ways of interacting with computer chains of millions of users. The computers are being applied in various areas like simulation, visualization, Parallel computing, virtual reality, Multimedia etc.

Fifth Generation : Defining the fifth generation of computers is somewhat difficult because the field is in its infancy. The most famous example of a fifth generation computer is the fictional HAL9000 from Arthur C. Clarke’s novel, 2001: A Space Odyssey. HAL performed all of the functions currently envisioned for real-life fifth generation computers. With artificial intelligence, HAL could reason well enough to hold conversations with its human operators,



Information Technology

use visual input, and learn from its own experiences. (Unfortunately, HAL was a little too human and had a psychotic breakdown, commandeering a spaceship and killing most humans on board.)

Though the wayward HAL9000 may be far from the reach of real-life computer designers, many of its functions are not. Using recent engineering advances, computers are able to accept spoken word instructions (voice recognition) and imitate human reasoning. The ability to translate a foreign language is also moderately possible with fifth generation computers. This feat seemed a simple objective at first, but appeared much more difficult when programmers realized that human understanding relies as much on context and meaning as it does on the simple translation of words.

Many advances in the science of computer design and technology are coming together to enable the creation of fifth-generation computers. Two such engineering advances are parallel processing, which replaces von Neumann's single central processing unit design with a system harnessing the power of many CPUs to work as one. Another advance is superconductor technology, which allows the flow of electricity with little or no resistance, greatly improving the speed of information flow. Computers today have some attributes of fifth generation computers. For example, expert systems assist doctors in making diagnoses by applying the problem-solving steps a doctor might use in assessing a patient's needs. It will take several more years of development before expert systems are in widespread use.

1.2 SIZE OF COMPUTERS

Computer systems are often categorized into super computers, mainframes, minis, and micros. These days computers are also categorized as servers and workstations.

1.2.1 Super computers - These are the largest and fastest computers available but are typically not used for commercial data processing. Instead they are used in specialized areas such as in Defence, aircraft design and computer generated movies, weather research etc. Predicting the weather involves analyzing thousands of variables gathered by satellites, aircrafts and other meteorological stations on the ground. This analysis has to be done in a very short time. A super computer can handle such situations efficiently. In the medical field, super computers are used to study the structure of viruses, such as those causing AIDS. Designing an aircraft involves simulating and analyzing the airflow around the aircraft. This again requires a super computer. The first super computer was the ILLIAC IV made by



Figure 1.2.1.1 :Super Computer(NEC)

Burroughs. Other suppliers of super computers are CRAY, CDC, Fujitsu, Intel Corporation, Thinking Machine Corporation, NEC, SGI, Hitachi, IBM and Sun Microsystems, etc.

In past, a high clock rate was one of the characteristics that distinguished super-computers from ordinary machines. For instance, the high clock rate of Cray processors made them the fastest available during the 1980s. However, microprocessor clock rates have now matched, and even in some cases surpassed the clock rates of super-computers. What distinguishes the super-computer of today from ordinary computers is their high degree of parallelism, i.e., their ability to perform a large number of operations simultaneously. All modern super-computers contain several processors, which can cooperate in the execution of a single program. Each processor can execute instructions following a program path independently of the others. Parallelism is achieved by decomposing programs into components, known as tasks or threads, which can be executed simultaneously on separate processors. Cray SVI super-computer introduced in 1998 can support 1,024 microprocessors, the cycle time is 4 nano seconds and has maximum memory size of 1024 gigabytes. On the other hand, Intel ASCI Red, introduced in 1997 which is a microprocessor-based super-computer, can support upto 9216 processors, Pentium Pro CPU and 584 MB of memory.

Super computers can process 64 bits or more at a time. Their processing speed ranges from 10,000 million instructions per second (MIPS) to 1.2 billion instructions per second. They can support up to 10,000 terminals at a time.

1.2.2 Mainframe computers: Mainframes are less powerful and cheaper than Super computers. However, they are big general-purpose computers capable of handling all kinds of scientific and business applications. Mainframes can process at several million instructions per second. A Mainframe can support more than 1,000 remote terminals.



Fig:1.2.2.1 Mainframe

Mainframes have large on-line secondary storage capacity. A number of different types of peripheral devices like magnetic tape drive, hard disk drive, visual display units, plotters, printers and telecommunication terminals can be attached with main-frame computers. They have high-speed cache memory which enables them to process applications at a faster rate than mini or microcomputers. They also offer the facility of multiprogramming and time-sharing.

Prices of Mainframe computers range between 1 crore to 5 crores depending upon the configuration. It is customary of Mainframe computer manufacturers to produce models ranging in size from small to very large, under a family designation. Computers belonging to a family are compatible *i.e.*, program prepared for one model of a family can run on another bigger model of the family. Major suppliers of Mainframe computers are IBM, Honey well, Burroughs, NCR, CDC and Sperry etc. Mainframes can be used for a variety of applications. A typical application of these computers is airline reservation or railway reservation system. The airlines have a mainframe computer at their head office where information of all flights is stored. Various terminals located at the booking offices are attached to the central date bank and up-to-date information of all flights can be obtained at any terminal.

1.2.3 Mini Computer - This type of computer performs data processing activities in the same way as the mainframe but on a smaller scale. The cost of minis is lower. Data is usually input by means of a keyboard. As the name implies, a minicomputer is small compared with a mainframe and may be called a scaled-down mainframe as the processor and the peripherals are physically smaller.

Minicomputers cost about Rs. 5 lacs to Rs. 50 lacs. The most popular minicomputer or minis, are the Data General Nova, DEC, PDP-11 and the IBM series/1. These systems can serve as information processors in small-to-medium sized firms or as processors in computer networks for large firms. Primary storage capacity starts at about 640K and can go as high as few mega bytes (MB) A minicomputer system consists of a CPU, several disk drives, a high-speed



Introduction To Computers

printer, perhaps a few magnetic tape units, and number of terminals. Programming languages include BASIC, PASCAL COBOL, C and FORTRAN. Much prewritten application software is also available.

Originally minicomputers were developed for process control and system monitoring etc. They were complicated to program and had minimal input/output capabilities as they were mainly concerned with “number crunching” rather than handling large amounts of data relating to business transactions. However, they are now fully developed, powerful computers with a wide range of peripherals to perform a wide range of data processing and computing activities. Minicomputer systems can be equipped with most of the input/output (I/O) devices and secondary storage devices that the large mainframe systems can handle, such as terminals and rigid disks. They are also making possible the installation of distributed data processing systems. Instead of a company having one large mainframe computer, it may have mini-computer at each of its remote locations and connect them to each other through telecommunications.

Minis certainly overlap mainframes. As minis become more powerful, they tend to perform with equal efficiency the jobs for which mainframes were used in the very near past, and the same is true for micros in relation to minis. Therefore, there is no definite delineation among the three types of computer systems, and the lines of demarcation are constantly changing.

1.2.4 Microcomputers: A microcomputer is a full-fledged computer system that uses a microprocessor as its CPU, these are also called personal computer systems. Microcomputers were first available for widespread use in the 1970s, when it became possible to put the entire circuitry of computers (CPU) onto a small silicon chip called microprocessor. A microprocessor is a product of the microminiaturization of electronic circuitry; it is literally a “computer on a chip”. Chip refers to any self-contained integrated circuit. The size of chips, which are about 30 thousandths of an inch thick, vary in area from fingernail size (about 1/4 inch square) to postage-stamp size (about 1-inch square) These days, relatively inexpensive microprocessors have been integrated into thousands of mechanical and electronic devices-even elevators, band saw, and ski-boot bindings. In a few years, virtually everything mechanical or electronic will incorporate microprocessor technology into its design.

The microprocessor is sometimes confused with its famous offspring, the microcomputer. A microprocessor, however, is not a computer. It only provides a part of CPU circuitry. This chip must be mounted together with memory, input and output chips on a single circuit board to make it a microcomputer. Thus, a microcomputer often called a micro is a small computer consisting of a processor on a single silicon chip which is mounted on a circuit board with other chips containing the computer's internal memory in the form of read-only-memory (ROM) and random-access-memory (RAM) It has a keyboard for the entry of data and instructions



Information Technology

and a screen for display purposes. It has interface for the connection of peripherals in the form of mouse, plotters, printers, cassette units, disk drives and light pens etc. IBM PC, APPLE II, TENDY TRS-80 are some of the popular microcomputers.

When people use the terms personal computers and microcomputers, they mean the small computer that are commonly found in offices, classrooms, and homes. Personal computers come in all shapes, and sizes. Although most models reside on desktops, others stand on the floor, and some are even portable.

The terms *microcomputer* and *personal computer* are interchangeable; however, PC - which stands for personal computer has a more specific meaning. In 1981, IBM called its first microcomputer the IBM PC. Within a few years, many companies were copying the IBM design, creating "clones" or "compatible computers" that aimed at functioning just like the original. For this reason, the term PC has come to mean that family of computers that includes IBM and compatibles. The Apple Macintosh computer, however, is neither an IBM nor a compatible. It is another family of microcomputers made by Apple computers. Apple computers are mainly used in Multimedia.

The earliest microcomputers were capable of supporting only a *single user at a time*. Now-a-days, multi-user microcomputer systems are also available and are becoming more prevalent. In multi-user systems, a powerful microcomputer may be used to substitute for Mainframe or minicomputer. single-user personal computers are also being connected to one another to form network. Multi-user microcomputers play key roles in some of the networks that are developed.

Currently IBM and Apple are the two most prominent manufacturers of microcomputers. A typical microcomputer consists of a processor on a single silicon chip mounted on a circuit board together with memory chips, ROM and RAM chips etc. It has a keyboard for the entry of data and instructions and a screen for display purposes. It has interfaces for connecting peripherals such as plotters, cassette units, disc drives, light pens, a mouse and joysticks.

A microcomputer including optional peripherals and other add-on-units may consist of the elements listed below.

- (a) 8, 16, or 32 bit processor,
- (b) Internal memory 256 MB expandable to 512 MB and more;
- (c) Backing storage-cassette, floppy disc, microfloppy discs, micro-drive, silicon disc or hard disc, CD-ROMS, DVDs, pen drives etc.;
- (d) Keyboard and screen (input and output);



Introduction To Computers

- (e) Interface (for the connection of peripherals);
- (f) Bus (communication and control channels);
- (g) Printer and/or plotter (multicolour text and graphics);
- (h) Pulse generator (clock);
- (i) Light pens, mouse, paddles/joysticks, Multimedia (graphics and games);
- (j) Software (programs)

Microcomputer systems are used by even the smallest of business, however their primary market is the personal home computer market. In the home, these computers can be used for a wide variety of tasks-from keeping track of the family budget to storing recipes to monitoring the home burglar alarm system. Currently, a small microcomputer system can be purchased for approximately Rs. 30,000. A more sophisticated microcomputer system with a 80 Giga bytes hard disk and 256 MB of primary storage can be purchased for approximately Rs. 25,000 to Rs. 40,000. With high-quality printer and additional memory (up to 512 MB), these microcomputer systems can cost in the vicinity of Rs. 50,000 to Rs. 75,000.

Examples of microcomputers are IBM PCs, PS/2 and Apple's Macintosh

1.2.5 Workstations : Between minicomputer and microcomputers - in terms of processing power - is a class of computers known as WORKSTATIONS. A workstation looks like a personal computer and is typically used by one person. Although workstations are still more powerful than the average personal computer, - the differences in the capabilities of these types of machines are growing smaller.

Workstations differ significantly from microcomputer in two areas. Internally, workstations are constructed differently than microcomputers. They are based on different architecture of CPU called *reduced instruction set computing* (RISC), which results in faster processing of instructions.

The other difference between workstations and microcomputers in that most microcomputers can run any of the four major operating systems* - DOS, Unix, OS/2, and Microsoft Windows NT), but workstations generally run the Unix operating systems or a variation of it. The biggest manufacturer of workstations is Sun Microsystems. Other manufacturers include IBM, DEC, Hewlette Packard and Silicon Graphics.

Many people use the term workstation to refer to any computer or terminal that is connected to another computer. Although this was once a common meaning of the term, it has become outdated. These days, a workstation is powerful RISC - based computer that runs the Unix Operating System and is generally used by scientists and engineers.

1.11



Information Technology

1.2.6. Server : A **server** is a computer system that provides services to other computing systems—called clients—over a network. The term is most commonly applied to a complete computer system today, but it is also used occasionally to refer only to the hardware or software portions of such a system. Servers occupy a place in computing similar to that occupied by minicomputers in the past, which they have largely replaced. The typical server is a computer system that operates continuously on a network and waits for requests for services from other computers on the network. Many servers are dedicated to this role, but some may also be used simultaneously for other purposes, particularly when the demands placed upon them as servers are modest. For example, in a small office, a large desktop computer may act as both a desktop workstation for one person in the office and as a server for all the other computers in the office.

Servers today are physically similar to most other general-purpose computers, although their hardware configurations may be particularly optimized to fit their server roles, if they are dedicated to that role. Many use hardware identical or nearly identical to that found in standard desktop PCs. However, servers run software that is often very different from that used on desktop computers and workstations.

Servers should not be confused with mainframes, which are very large computers that centralize certain information-processing activities in large organizations and may or may not act as servers in addition to their other activities. Many large organizations have both mainframes and servers, although servers usually are smaller and much more numerous and decentralized than mainframes.

Servers frequently host hardware resources that they make available on a controlled and shared basis to client computers, such as printers (*print servers*) and file systems (*file servers*). This sharing permits better access control (and thus better security) and can reduce costs by reducing duplication of hardware.

1.3 ADVANTAGES AND LIMITATIONS OF COMPUTERS

1.3.1 Advantages of Computer System : In a nutshell, computers are fast, accurate, and reliable; they don't forget anything; and they don't complain. We will now describe them in detail.

Speed: The smallest unit of time in the human experience is, realistically, the second. Computer operations (for example, the execution of an instruction, such as multiplying the hours worked times the rate of pay) are measured in **milliseconds, microseconds, nanoseconds, and picoseconds** (one thousandth, one millionth, one billionth, and one trillionth of a second, respectively)



Accuracy: is the degree to which information on a map or in a digital database matches true or known values. Accuracy is an issue pertaining to the quality of data and the number of errors contained in a dataset or map. In discussing a GIS database, it is possible to consider horizontal and vertical accuracy with respect to geographic position, as well as attribute, conceptual, and logical accuracy.

- The level of accuracy required for particular applications varies greatly.
- Highly accurate data can be very difficult and costly to produce and compile.

Reliability: Reliability can be defined as the ability of a person or system to perform and maintain its functions in routine circumstances, as well as in hostile or unexpected circumstances.

The IEEE defines it as ". . . the ability of a system or component to perform its required functions under stated conditions for a specified period of time."

Reliability may refer to:

Reliability (statistics), of a set of data and experiments

High reliability is informally reported in "nines"

Data reliability, a property of some disk arrays in computer storage

Reliability engineering ensures a system will be reliable when operated in a specified manner

Reliability theory, as a theoretical concept, to explain biological aging and species longevity

Reliability (computer networking), a category used to describe protocols

Memory Capability: Computer systems have total and instant recall of data and an almost unlimited capacity to store these data. A typical mainframe computer system will have many billions of characters stored and available for instant recall. High-end PCs have access to about a billion characters of data. To give you a benchmark for comparison, a 15-page report contains about 50,000 characters.

1.3.2 Limitations of Computer systems : The computer is one of the most powerful tools ever developed. But we've all read articles similar to the one about the man who was treated for pneumonia and then charged by the hospital's computer for the use of the delivery room and nursery. Such "computer failures" may be amusing, but most such foul-ups happen because people fail to consider some basic computer limitations. Without reliable programs and sound logic, no computer system will perform adequately.



Information Technology

Program must be reliable: The computer does what it's programmed to do *and nothing else*. A clever program can be written to direct the computer to store the results of previous decisions. Then, by using the program's branching ability, the computer may be able to modify its behavior according to the success or failure of past decisions. But a program that has operated flawlessly for months can suddenly produce nonsense. Perhaps some rare combination of events has presented the system with a situation for which there's no programmed course of action. Or perhaps the course of action provided by the programmer contains an error that's just being discovered. Of course, a reliable program that's supplied with incorrect data may also produce nonsense.

Application logic must be understood: The computer can only process jobs which can be expressed in a finite number of steps leading to a specify goal. Each step must be clearly defined. If the steps in the solution cannot be precisely stated, the job cannot be done. This is why the computer may not be helpful to people in areas where subjective evaluations are important. For example, it may not be able to tell a sales manager if a new product will be successful. The market decision may hinge on educated guesses about future social, political, technological and economic changes. But the computer can tell the manager how the product will fare under assumed price, cost, and sales volume conditions. These assumed values could be fed into the computer. An analysis program can then manipulate them in response to a series of "what if" questions to project the effects that the manager's questions will have on profits.

Even if program steps are finite and understood, there are still some tasks whose execution could take millions of years, even on a supercomputer. Joseph Weizenbaum, a computer scientist at MIT observed that a program could be written to try every legal chess move in a given situation. Every response to a move could then be evaluated, and the subsequent moves and countermoves could all be identified until the computer found a move, which, if suitably pursued, would guarantee a win. Weizenbaum notes that this program would certainly be finite, but the time needed to execute it would be unimaginably large. Although in principle the computer could do the job, in practice it cannot. The term **combinatorial explosion** is used for this type of problem where a finite number of steps generate an impossibly large number of computer operations.

1.4 COMPONENTS OF A COMPUTER SYSTEM- CPU

The hardware are the parts of computer itself including the Central Processing Unit (CPU) and related microchips and micro-circuitry, keyboards, monitors, case and drives (floppy, hard, CD, DVD, optical, tape, etc.) Other extra parts called peripheral components or devices include mouse, printers, modems, scanners, digital cameras and cards (sound, colour, video)



Introduction To Computers

etc... Together they are often referred to as a personal computers or PCs.

The schematic diagram of a computer is given below :



Fig. 1.4.1 Processor of Computer

We will now briefly discuss each of the above components.

1.4.1 Central Processing Unit: The Central Processing Unit (CPU)—also known as the processor—is the heart, soul and brain of the computer. In a microcomputer, the entire CPU is contained on a tiny chip called a microprocessor. Though the term relates to a specific chip or the processor a CPU's performance is determined by the rest of the computer's circuitry and chips. Currently the Pentium chip or processor, made by Intel, is the most common CPU though there are many other companies that produce processors for personal computers. One example is the CPU made by Motorola which is used in Apple computers. It is the most important component on the system's motherboard. The processor computes and processes data and delivers the results based on the instructions that are fed to the PC. Every CPU has at least two basic parts, the control unit and the arithmetic logic unit.

(i) The Control Unit

All the computer's resources are managed from the control unit. One can think of the control unit as a traffic cop directing the flow of data. It is the logical hub of the computer.

The CPU's instructions for carrying out commands are built into the control unit. The instructions, or instruction set, list all the operations that the CPU can perform. Each instruction in the instruction set is expressed in microcode- a series of basic directions that tell the CPU how to execute more complex operations. Before a program can be executed, every command in it must be broken down into instructions that correspond to the ones in the CPU's



Information Technology

instruction set. When the program is executed, the CPU carries out the instructions, in order, by converting them into microcode. Although the process is complex, the computer can accomplish it at an incredible speed, translating millions of instructions every second.

Different CPUs have different instruction sets. Manufacturers, however, tend to group their CPUs into “families” that have similar instruction sets. Usually, when a new CPU is developed, the instruction set has all the same commands as its predecessor plus some new ones. This allows software written for a particular CPU to work on computers with newer processors of the same family – a design strategy known as **upward compatibility**. Upward compatibility saves consumers from having to buy a whole new system every time a part of their existing system is upgraded. The reverse is also true. When a new hardware device or piece of software can interact with all the same equipment and software its predecessor could, it is said to have **downward**, or backward, compatibility.

(ii) The Arithmetic Logic Unit

Because computers store all the data as numbers, a lot of the processing that takes place involves comparing numbers or carrying out mathematical operations. In addition to establishing ordered sequences and changing those sequences, the computer can perform only two types of operations: arithmetic operations and logical operations. Arithmetic operations include addition, subtraction, multiplication, and division. Logical operations include comparisons, such as determining whether one number is equal to, greater than, or less than another number. Also, every logical operation has an opposite. For example, in addition to “equal to” there is “not equal operation has an opposite. For example, in addition to “equal to” there is “not equal to”.

Some of the logical operations can be carried out on text data. For example, a word is required to be searched in a document, the CPU carries out a rapid succession of “equals” operations to find a match for the sequence of ASCII codes that make up the word being searched.

Many instructions carried out by the control unit involve simply moving data from one place to another – from memory to storage, from memory to the printer, and so forth. However, when the control unit encounters an instruction that involves arithmetic or logical operation, it passes that instruction to the second component of the CPU, the arithmetic logical unit, or ALU. The ALU includes a group of registers – high-speed memory locations built directly into the CPU that are used to hold the data currently being processed. For example, the control unit might load two numbers from memory into the registers in the ALU. Then, it might tell the ALU to divide the two numbers (an arithmetic operation) or to see whether the numbers are equal (a logical operation)



1.4.2 Various features of the Central Processing Unit : Over a period of time, the processor has evolved from slow 286s or 386s running at speeds as low as 20 MHz to present day Pentium III and IV running at a whopping 3 GHz (3000 MHz.) Now we take a closer look at the various features that the Central Processing Unit of a PC offers.

Clock Speed: The clock speed is the speed at which the processor executes instructions. Clock speed is measured in megahertz (MHz)—which is a million cycles per second. Therefore, a 450 MHz processor performs 450 million instructions per second. Higher the clock's speed, the faster the processor, the better the system performance. Also, some microprocessors are super scalar, which means that they can execute more than one instruction per clock cycle.

Cache: Processors incorporate their own internal cache memory. The cache acts as temporary memory and boosts processing power significantly. The cache that comes with the processor is called Level One (L1) cache. This cache runs at the processor's clock speeds, and therefore is very fast. The L1 cache is divided into 2 sections—one for data, the other for instructions. Generally, more the L1 cache, faster the processor.

Additionally, PCs also include a much slower secondary, or Level Two (L2) cache. This cache resides on the motherboard and delivers slower performance when compared with the L1 cache. To overcome this limitation, newer chips (Pentium II and Pentium III) house the L2 cache in a cartridge along with the CPU.

Architecture: The CPUs architecture determines the manner in which it processes data. New CPUs employ multi-staged pipelines for transmitting data. To ensure proper data flow through these lines, the CPU includes a kind of prediction and error correction mechanism.

Slot: Different processors use different sockets or slots to fit onto the motherboard. Based on the type of processors, there are two main types of slots for connecting to the motherboard—Socket 7 and Slot 1.

Socket 7 is a 321-pin socket for Pentium class CPUs—Pentium MMX, K5, and K6—ranging from 75 MHz to 200 MHz processors. However, the Pentium II/III CPUs use Slot 1 for connecting to the motherboard.

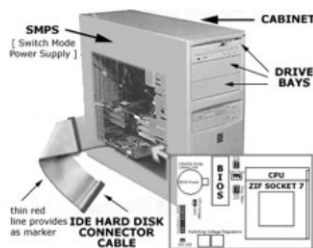


Fig 1.4.2.1 CPU

Instead of the usual manner in which a CPU fits onto the motherboard, Slot 1 CPUs fit onto the motherboard as a daughter card, allowing for faster communication between the CPU and the L2 cache.

Density: A CPU is made up of millions of small transistors. A CPU performs all the calculation and manipulation operations by synchronising between the transistors. Therefore, the shorter



Information Technology

the distance between two transistors on a CPU, the faster the performance. Older CPUs had a distance of one micron between the transistors. But, newer CPUs have a distance as small as 0.35 micron between two transistors, delivering faster performance.

MMX: MMX stands for Multimedia Extensions—a set of instructions built in to the CPU, specifically intended for improving the performance of multimedia or graphic applications—mainly games. However, one needs to have applications specifically designed to take advantage of MMX.

CPU generates lots of heat when in operation. If the CPU is not cooled properly, then it might lead to all sort of errors, including system crashes. Therefore, the CPU is usually covered by a heat sink and a small cooling fan to dissipate the heat generated by the processor.

The microprocessor, is not made by the manufacturers of micro computers but by companies, such as Motorola and Intel, that specialise in the development and manufacture of microprocessors. All the Apple's Macintosh-series micros use Motorola chips : the Motorola 68000 in earlier models, the Motorola 68020 in the Machintosh II, and the Motorola 68030 in recent models. The system board for IBM Personal Computer uses Intel Processors.

When someone talks about a “286”, “386”, “486” or Pentium machine, he or she is referring to a micro that uses an Intel 80286, 80386, 80486 or Pentium chip.

1.4.3 Types of Microprocessors - Currently three classes of microprocessors are used for personal computers: 8-bit, 16-bit and 32-bit. Basically, an 8-bit machine can process 8-bits (1 character) of data at a time, and each instruction will be represented by an 8-bit code. A 16-bit machine can process two bytes (or 16 bits) of data at a time, and the number of instructions is increased over that 8-bit machine.

All the microprocessors use a bus-type design to transfer bits within the computer and to input/output devices. The electric path or lines that transfer these bits are called **buses**. An 8-bit machine usually has 8-data buses that transfer 8-bits at a time between components of a computer.

A personal computer transfers data to its I/O devices through input/output ports connected to a bus. A **port** is a hardware device that allows a series of bits to be transferred to a bus for data input or, inversely, the transfer of data from a bus to the port for data output.

The 8-bit personal computers were based on two types of 8-bit microprocessors viz 8080/Z80 and the 6502. The industry standard operating system CP/M ran on 8080/Z80, and therefore



Introduction To Computers

many personal computers including Zenith, Tandy, TRS80, Morrow, Northstar, etc. were based on this microprocessor. Apple II, and Commodore computers were based on the 6502 microprocessor, and each used its own proprietary operating system.

The 16-bit personal computers were based on two classes of microprocessors: the 8086/8088 (Intel) and MC 68000 (Motorola). The industry standard MS-DOS operating system for IBM - PC is built around the 8088 microprocessor. Apple's Macintosh is built around the MC 68000 as is the AT&T UNIX PC. IBM AT used Intel 80286 microprocessor. Most of the PCs made during the 1980s had 16-bit processors. The 16-bit systems have provided many sophisticated functions such as colour graphics, database features etc. once limited to mini computers.

Personal computers based on the Intel 80386 and 80486 DX, SX are 32-bit processors and can process four 8-bit bytes at a time. The 32-bits microprocessors have tied the personal computers into more sophisticated information handling functions. The 386 processor used a new mode of operation called virtual 86 mode. This allowed operating systems such as Unix and OS/2, and special programs such as Microsoft Windows, to run several DOS programs at the same time. This feature is specially important for DOS-based control programs such as Microsoft Windows, because it allows software to simulate multitasking with DOS operating system which otherwise cannot perform true multitasking.

The 486 processor combined a 386 DX processor, a maths coprocessor and a cache memory controller onto a single chip. This increased the speed of the processor drastically. The SX versions of Intel chips, such as the 386SX and 486SX, are less expensive and less powerful than the processors upon which they are based.

For the moment, the most powerful member of the Intel family of microprocessors is the Pentium. With the Pentium processor, Intel broke its tradition of numeric model names. The speed and powers of the Pentium dwarf those of all its predecessors in the Intel family. The 486 has approximately 1.2 million transistors, the Pentium has over 3 million and can process 100 million instructions to 200 million instructions per second.

Introduced in 1993, the Pentium processor allowed computers to more easily incorporate "real world" data such as speech, sound, handwriting and photographic images. The name Pentium, mentioned in the comics and on television talk shows, became a household word soon after introduction.

Released in the fall of 1995 the Pentium Pro processor is designed to fuel 32-bit server and workstation-level applications, enabling fast computer-aided design, mechanical engineering and scientific computation. Each Pentium Pro processor is packaged together with a second speed-enhancing cache memory chip. The powerful Pentium pro processor boasts 5.5 million transistors.



Information Technology

The 7.5 million-transistor Pentium II processor launched in 1997 incorporates Intel MMX technology, which is designed specifically to process video, audio and graphics data efficiently. It was introduced in innovative Single Edge Contact (S.E.C) Cartridge that also incorporated a high-speed cache memory chip. With this chip, PC users can capture, edit and share digital photos with friends and family via the Internet; edit and add text, music or between-scene transitions to home movies; and, with a video phone, send video over standard phone lines and the Internet.

In 1998, Intel introduced Pentium II Xeon processors designed to meet the performance requirements of mid-range and higher servers and workstations. Consistent with Intel's strategy to deliver unique processor products targeted for specific markets segments, the Pentium II Xeon processors feature technical innovations specifically designed for workstations and servers that utilize demanding business applications such as Internet services, corporate data warehousing, digital content creation, and electronic and mechanical design automation. Systems based on the processor can be configured to scale to four or eight processors and beyond.

Continuing Intel's strategy of developing processors for specific market segments, the Intel Celeron processor (1999) is designed for the Value PC market segment. It provides consumers great performance at an exceptional value, and it delivers excellent performance for uses such as gaming and educational software.

The Pentium III processor (1999) features 70 new instructions—Internet Streaming SIMD extensions—that dramatically enhance the performance of advanced imaging, 3-D, streaming audio, video and speech recognition applications. It was designed to significantly enhance Internet experiences, allowing users to do such things as browse through realistic online museums and stores and download high-quality video. The processor incorporates 9.5 million transistors, and was introduced using 0.25-micron technology.

The Pentium III Xeon processor (1999) extends Intel's offerings to the workstation and server market segments, providing additional performance for e-Commerce applications and advanced business computing. The processors incorporate the Pentium III processor's 70 SIMD instructions, which enhance multimedia and streaming video applications. The Pentium III Xeon processor's advance cache technology speeds information from the system bus to the processor, significantly boosting performance. It is designed for systems with multiprocessor configurations.

The Intel Pentium 4 Processor is designed to deliver performance across usages—such as image processing, video content creation, games and multimedia—where end-users can truly appreciate the performance. With a PC based on the Intel Pentium 4 Processor with HT



Introduction To Computers

Technology, one gets advanced performance and multitasking capabilities for today's digital home and digital office applications. Hyper-threading enables multi-threaded software applications to execute two software threads in parallel, thereby improving system responsiveness. Intel Pentium 4 Processors enabled with HT Technology deliver performance and multitasking gains that result in increased productivity and efficiency. It allows the operating system to adjust the processor clock down when running applications that require less power. Increased power efficiency brings savings.

Intel Extended Memory 64 Technology can improve performance by allowing the system to address more than 4 GB of both virtual and physical memory. Intel EM64T also provides support for 64 bit computing to help handle the applications of tomorrow.

1.4.4 Processor Speed - As mentioned earlier, a crystal oscillator paces the execution of instructions within the processor of a microcomputer. A micro's processor speed is rated by its frequency of oscillation, or the number of clock cycles per second. Earlier personal computers rated between 5 and 50 megahertz, or MHz (millions of clock cycles) . Normally several clock cycles are required to retrieve, decode, and execute a single program instruction. The shorter the clock cycle, the faster the processor.

To properly evaluate the processing capability of a micro, one must consider both the processor speed and the word length. A 32-bit micro with a 25 - MHz processor has more processing capability than a 16-bit micro with a 25 - MHz processor.

The Pentium II processors can process in the range of 233 MHz to 300 MHz. The latest Pentium-III and Pentium 4 processors can run at a speed of 2.1 GHz and even higher.

1.5 MOTHERBOARDS

The motherboard or the system board is the main circuit board on the computer. It acts as a direct channel for the various components to interact and communicate with each other. There are various types of motherboards available (depending on the processors that are used) We now provide with an overview of the system motherboard, and about the various components that fit on it.

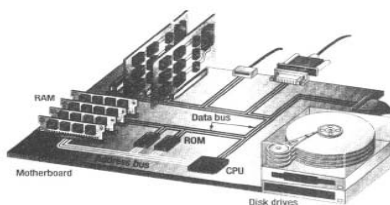




Fig. 1.5.1 Motherboard

1.5.1 Processor slot: The processor slot houses the processor. Based on the type of processors used, there are two main types of slots—Socket-7 and Slot-1.

BIOS: BIOS stands for Basic Input Output System—a small chip on the motherboard that loads the hardware settings required to load various devices like keyboards, monitors, or disk drives. Most new PCs come with a Flash BIOS—these BIOS can be software upgraded to support new devices.

CMOS: The PC uses the CMOS memory to store the date, time and system setup parameters. These parameters are loaded every time the computer is started. A small Lithium Ion battery located on the motherboard powers the CMOS as well as the BIOS.

Power supply connectors: The power supply connectors allow the user to connect the power supply unit to the motherboard and provide power for the functioning of the various components that fit on to the motherboard.

1.5.2 Expansion Slots and Boards : PCs are designed so that users can adapt, or configure the machines to their own particular needs. PC motherboards have two or more expansion slots, which are extensions of the computer's bus that provide a way to add new components to the computer. The slots accept circuit board, also called cards, adapters, or sometimes-just boards. Modern notebook computers are too small to accept the same type of cards that fit into desktop models. Instead, new components for notebooks come in the form of PC cards, small devices – about the size of credit cards – that fit into a slot on the back or side of the notebook. Figure –1.5.2.1 shows a PC expansion board being installed. The board is attached to the motherboard – the main system board to which the CPU, memory, and other components are attached.

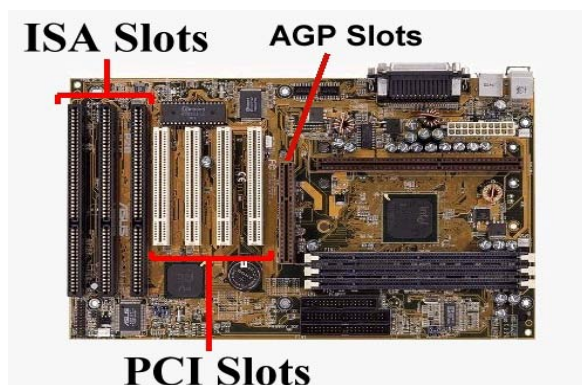




Fig. 1.5.2.1 : Expansion Slots

The expansion slots on the motherboard are used for three purposes:

1. To give built-in devices such as hard disks and diskette drives access to the computer's bus via controller cards.
2. To provide I/O (input/output) ports on the back of the computer for external devices such as monitors, external modems, printers, and the mouse (for computers that do not have a built-in mouse port)
3. To give special-purpose devices access to the computer. For example, a computer can be enhanced with an accelerator card, a self contained device that enhances processing speed through access to the computer's CPU and memory by way of the bus.

The first and second of these are input/output (I/O) functions. Adapters that serve these purposes provide a port to which devices can be attached and serve as a translator between the bus and the device itself. Some adapters also do a significant amount of data processing. For example, a video controller is a card that provides a port on the back of the PC into which one can plug the monitor. It also contains and manages the video memory and does the processing required to display images on the monitor. Other I/O devices that commonly require the installation of a card into an expansion slot include sound cards, internal modems or fax/modems, network interface cards, and scanners. The third type, the accelerator cards, are often installed to speed up the CPU or the display of video. Some of the slots and connectors are briefly discussed below:

SIMM/DIMM slots: SIMM stands for Single Inline Memory Modules, while DIMM stands for Dual Inline Memory Module. SIMM/DIMM slots are used to house RAM modules.

PCI slots: The PCI (Peripheral Component Interface) slots are used for connecting PCI-based devices like graphics accelerator cards, sound cards, internal modems or SCSI cards.

AGP slot: All Celeron and Pentium-III motherboards come with an AGP (Accelerated Graphics Port) slot. AGP is a dedicated slot meant to provide faster access to AGP-based graphic accelerator cards, thus enhancing the visual experience for the user.

SCSI : It is a device interface that is used to solve the problem of a finite and possibly insufficient number of expansion slots. It is called *small computer system interface (SCSI pronounced "scuzzy")*. Instead of plugging interface cards into the computer's bus via the expansion slots, SCSI extends the bus outside the computer by way of a cable. In other



Information Technology

words, SCSI is like an extension cord for computer bus. IBM developed SCSI in 1970s. The current standard is SCSI - 3, which allows upto seven devices to be chained on a single SCSI port. Now-a-days many devices support the SCSI interface. Fast, high-speed hard disk drives often have SCSI interfaces, so do scanners, tape drives and optical storage devices.

1.5.3 Cards: Cards are components added to computers to increase their capability. When adding a peripheral device one should ensure that the computer has a slot of the type needed by the device. **Sound cards** allow computers to produce sound like music and voice. The older sound cards were 8 bit then 16 bit then 32 bit. Though human ear can't distinguish the fine difference between sounds produced by the more powerful sound card they allow for more complex music and music production. Colour cards allow computers to produce colour (with a colour monitor of course) The first colour cards were 2 bit which produced 4 colours [CGA]. It was amazing what could be done with those 4 colours. Next came 4 bit allowing for 16 [EGA and VGA] colours Then came 16 bit allowing for 1064 colours and then 24 bit which allows for almost 17 million colours and now 32 bit is standard allowing monitors to display almost a billion separate colours. **Video cards** allow computers to display video and animation. Some video cards allow computers to display television as well as capture frames from video. A video card with a digital video camera allows computers users to produce live video. A high speed or network connection is needed for effective video transmission. **Network cards** allow computers to connect together to communicate with each other. Network cards have connections for cable, thin wire or wireless networks.

1.5.4 Ports and connectors : Ports and connectors let the user connect external devices like printers, keyboards or scanners and let them interface with the PC. The physical interfaces for the ports and connectors are located on the outside—typically at the back of the PC, but they are directly or indirectly (using a connector card) connected to the motherboard. There are various types of ports or connectors, each providing different data transfer speeds to connect various external peripherals.

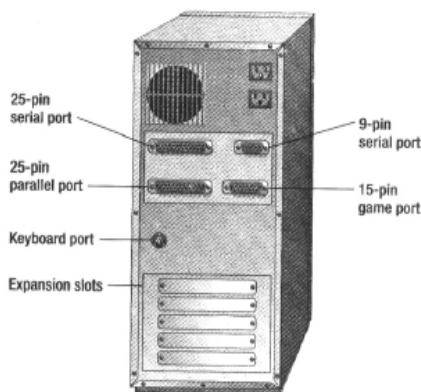


Fig. 1.5.4.1 (Ports and Connectors)

Parallel ports: Parallel ports are used to connect external input/output devices like scanners or printers. Parallel ports facilitate the parallel transmission of data, usually one byte (8 bits) at a time. Parallel ports use 25 pin RS-



232C.

Com/Serial ports: They are used for connecting communication devices like modems or other serial devices like mice. There are two varieties of Com ports—the 9-pin ports and 25-pin ports. Serial Ports facilitate the serial transmission of data, i.e. one bit at a time.

IDE drive connector: IDE devices like CD-ROM drives or hard disk drives are connected to the motherboard through the IDE connector.

Floppy drive connector: The floppy drive connectors are used for connecting the floppy drive to the motherboard, to facilitate data exchange.

USB connectors: USB stands for Universal Serial Bus. These ports provide the user with higher data transfer speeds for different USB devices like keyboards, mice, scanners or digital cameras.

PS/2 Connectors: PS/2 stands for Personal System/2. PS/2 connectors are used to connect PS/2 based input devices like PS/2 keyboards or mice.

In addition to the common components that are found on the motherboard, newer motherboards also come with integrated graphics accelerator cards or sound cards—there is no need to install a separate card to get the work done.

1.5.5 The bus : If one takes a close look at the system motherboard, one will notice a maze of golden electric circuits etched on both sides of the motherboard. This very maze of circuits etched on the motherboard forms the bus of the PC. A bus acts as the system's expressway - it transmits data between the various components on the motherboard. Theoretically, a bus is a collection of wires through which data is transmitted between the various components of a PC. A bus connects the various components of the PC with the CPU and the main memory (RAM) Logically, a bus consists of two parts—an address bus and a data bus.

The Data Bus : The Data Bus is an electrical path that connects the CPU, memory, and the other hardware devices on the motherboard. Actually, the bus is a group of parallel wires. The number of wires in the bus affects the speed at which data can travel between hardware components, just as the number of lanes on a highway affects how long it takes people to get to their destinations. Because each wire can transfer one bit at a time, an eight-wire bus can move eight bits at a time, which is a full byte. A 16-bit bus can transfer two bytes, and a 32-bit bus can transfer four bytes at a time.

PC buses are designed to match the capabilities of the devices attached to them. When CPUs could send and receive only one byte of data at a time, there was no point in connecting them



Information Technology

to a bus that could move more data. As microprocessor technology improved, however, chips were built that could send and receive more data at once, and improved bus designs created wider paths through which the data could flow.

When IBM introduced the PC-AT in 1984, the most dramatic improvement was an enhanced data bus that was matched with the capabilities of a new microprocessor, the Intel 80286. The data bus of the AT was 16 bits wide and became the de facto standard in the industry. It is still used for PC devices that do not require more than a 16 -bit bus. The AT bus is commonly known as the Industry Standard Architecture, or ISA, bus.

Two years later, however, when the first 80386 chips (commonly abbreviated as the 386) began shipping, a new standard was needed for the 386's 32-bit bus. The first contender was Micro Channel Architecture, or the MCA bus, from IBM. Then came the Extended Industry Standard Architecture (EISA) bus from a consortium of hardware developers who opposed IBM's new standard because it was not backward compatible. The winner of the bus wars was neither MCA nor EISA. It was the Peripheral Component Interconnect, or PCI, bus. Intel designed the PCI bus specifically to make it easier to integrate new data types, such as audio, video, and graphics.

The Address Bus : The second bus that is found in every microcomputer is the address bus. The address bus is a set of wires similar to the data bus that connects the CPU and RAM and carries the memory addresses. (Remember, each byte in RAM is associated with a number, which is the memory address)

The reason the address bus is important is that the number of wires in it determines the maximum number of memory addresses. For example, recall that one byte of data is enough to represent 256 different values. If the address bus could carry only eight bits at a time, the CPU could address only 256 bytes of RAM. Actually, most of the early PCs had 20-bit address buses, so the CPU could address 220 bytes, or 1 MB, of data. Today, most CPUs have 32-bit address buses that can address 4 GB (over 4 million bytes) of RAM. Some of the latest models can address even more.

One of the biggest hurdles in the evolution of PCs was that DOS, the operating system used in the vast majority of PCs for more than a decade, was designed for machines that could address only 1 MB of RAM. When PCs began to contain more RAM, special software had to be devised to address it. Programmers came up with two devices called expanded memory and extended memory. Windows 95 largely did away with these, although extended memory still exists in the operating system for purposes of backward compatibility.

1.6 STORAGE DEVICES



Introduction To Computers

The CPU contains the basic instructions needed to operate the computer, but it does not have the capability to store programs or large sets of data permanently. Just like the human brain, which helps to determine what to do and when, computers need blocks of space that it can address from time to time to help in processing arithmetical and logical operations and also hold programs and data being manipulated. This area is called memory or storage.

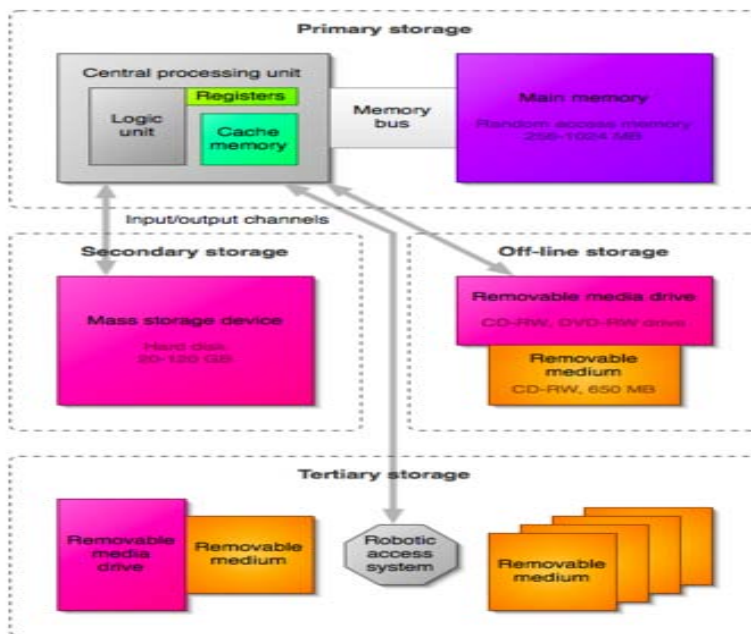


Fig. 1.6.1 Types of Storage

1.6.1 Types of storage : Various forms of storage, based on various natural phenomenon, have been invented. So far, no practical universal storage medium exists, and all forms of storage have some drawbacks. Therefore a computer system usually contains several kinds of storage, each with an individual purpose, as shown in Fig. 1.6.1.

- (i) **Primary storage :** Primary storage is directly connected to the central processing unit of the computer. It must be present for the CPU to function correctly, just as in a biological analogy the lungs must be present (for oxygen storage) for the heart to function (to pump and oxygenate the blood) As shown in the figure, primary storage typically consists of three kinds of storage:

Processor registers are internal to the central processing unit. Registers contain information that the arithmetic and logic unit needs to carry out the current instruction.



Information Technology

They are technically the fastest of all forms of computer storage, being switching transistors integrated on the CPU's silicon chip, and functioning as electronic "flip-flops".

Main memory contains the programs that are currently being run and the data on which the programs are operating. The arithmetic and logic unit can very quickly transfer information between a processor register and locations in main storage, also known as a "memory addresses". In modern computers, electronic solid-state random access memory is used for main storage, and is directly connected to the CPU via a "memory bus" (shown in the diagram) and a "data bus". The memory bus is also called an address bus or front side bus and both buses are high-speed digital "superhighways". Access methods and speed are two of the fundamental technical differences between memory and mass storage devices. (Note that all memory sizes and storage capacities shown in the diagram will inevitably be exceeded with advances in technology over time)



Fig 1.6.1.1 : Main Memory

Cache memory is a special type of internal memory used by many central processing units to increase their performance or "throughput". Some of the information in the main memory is duplicated in the cache memory, which is slightly slower but of much greater capacity than the processor registers, and faster but much smaller than main memory. Multi-level cache memory is also commonly used - "primary cache" being smallest, fastest and closest to the processing device; "secondary cache" being larger and slower, but still faster and much smaller than main memory.

- (ii) *Secondary, tertiary and off-line storage* : **Secondary storage** requires the computer to use its **input/output** channels to access the information, and is used for long-term storage of persistent information. Nowadays most computer **operating systems** also use secondary storage devices as **virtual memory** - to artificially increase the apparent amount of main memory in the computer. Secondary storage is also known as "mass storage", as shown in the **figure 1.6.1**. Secondary or mass storage is typically of much greater capacity than primary storage (main memory), but it is also very much slower. In modern computers, **hard disks** are usually used for mass storage. The time taken to



access a given byte of information stored on a hard disk is typically a few thousandths of a second, or milliseconds. By contrast, the time taken to access a given byte of information stored in random access memory is measured in thousand-millionths of a second, or nanoseconds. This illustrates the very significant speed difference which distinguishes solid-state memory from rotating magnetic storage devices: hard disks are typically about a million times slower than memory. Rotating optical storage devices (such as CD and DVD drives) are typically even slower than hard disks, although their access speeds are likely to improve with advances in technology. Therefore the use of virtual memory, which is about million times slower than "real" memory, significantly degrades the performance of any computer.

Tertiary storage is a system where a robotic arm will "mount" (connect) or "dismount" off-line mass storage media (see the next item) according to the computer operating system's demands. Tertiary storage is used in the realms of enterprise storage and scientific computing on large computer systems and business computer networks, and is something a typical personal computer user never sees firsthand.

Off-line storage is a system where the storage medium can be easily removed from the storage device. Off-line storage is used for data transfer and archival purposes. In modern computers, floppy disks, optical discs and flash memory devices including "USB drives" are commonly used for off-line mass storage purposes. "Hot-pluggable" USB hard disks are also available. Off-line storage devices used in the past include magnetic tapes in many different sizes and formats, and removeable Winchester disk /drums.

- (iii) **Network storage** : Network storage is any type of computer storage that involves accessing information over a **computer network**. Network storage arguably allows to centralize the **information management** in an organization, and to reduce the duplication of information. Network storage includes:

Network-attached storage is secondary or tertiary storage attached to a computer which another computer can access over a local-area network, a private wide-area network, or in the case of online file storage, over the Internet.

Network computers are computers that do not contain internal secondary storage devices. Instead, documents and other data are stored on a network-attached storage.

1.6.2 Characteristics of storage : The division to primary, secondary, tertiary and off-line storage is based on **memory hierarchy**, or *distance from the central processing unit*. There are also other ways to characterize various types of storage.

- (i) *Volatility of information*



Information Technology

Volatile memory requires constant power to maintain the stored information. Volatile memory is typically used only for primary storage.

Non-volatile memory will retain the stored information even if it is not constantly supplied with electric power. It is suitable for long-term storage of information, and therefore used for secondary, tertiary, and off-line storage.

Dynamic memory is volatile memory which also requires that stored information is periodically *refreshed*, or read and rewritten without modifications.

(ii) *Ability to access non-contiguous information*

Random access means that any location in storage can be accessed at any moment in the same, usually small, amount of time. This makes random access memory well suited for primary storage.

Sequential access means that the accessing a piece of information will take a varying amount of time, depending on which piece of information was accessed last. The device may need to *seek* (e.g. to position the read/write head correctly), or *cycle* (e.g. to wait for the correct location in a constantly revolving medium to appear below the read/write head)

(iii) *Ability to change information*

Read/write storage, or **mutable storage**, allows information to be overwritten at any time. A computer without some amount of read/write storage for primary storage purposes would be useless for many tasks. Modern computers typically use read/write storage also for secondary storage.

Read only storage retains the information stored at the time of manufacture, and **write once storage** (WORM) allows the information to be written only once at some point after manufacture. These are called **immutable storage**. Immutable storage is used for tertiary and off-line storage. Examples include CD-R.

Slow write, fast read storage is read/write storage which allows information to be overwritten multiple times, but with the write operation being much slower than the read operation. Examples include CD-RW.

(iv) *Addressability of information*

In **location-addressable storage**, each individually accessible unit of information in storage is selected with its numerical memory address. In modern computers, location-addressable storage usually limits to primary storage, accessed internally by computer programs, since



location-addressability is very efficient, but burdensome for humans.

In **file system storage**, information is divided into *files* of variable length, and a particular file is selected with human-readable directory and file names. The underlying device is still location-addressable, but the operating system of a computer provides the file system abstraction to make the operation more understandable. In modern computers, secondary, tertiary and off-line storage use file systems.

In **content-addressable storage**, each individually accessible unit of information is selected with a hash value, or a short identifier with no pertaining to the memory address the information is stored on. Content-addressable storage can be implemented using software (computer program) or hardware (computer device), with hardware being faster but more expensive option.

(v) Capacity and performance

Storage capacity is the total amount of stored information that a storage device or medium can hold. It is expressed as a quantity of bits or bytes (e.g. 10.4 megabytes)

Storage density refers to the compactness of stored information. It is the storage capacity of a medium divided with a unit of length, area or volume (e.g. 1.2 megabytes per square centimeter)

Latency is the time it takes to access a particular location in storage. The relevant is typically nanosecond for primary storage, millisecond for secondary storage, and second for tertiary storage. It may make sense to separate *read latency* and *write latency*, and in case of sequential access storage, *minimum*, *maximum* and *average latency*.

Throughput is the rate at which information can read from or written to the storage. In computer storage, throughput is usually expressed in terms of *megabytes per second* or *MB/s*, though bit rate may also be used. As with latency, *read rate* and *write rate* may need to be differentiated.

1.6.3 Primary Storage

(a) Semi-conductor memories or integrated circuits : As they are often called, are based on the principle of storage chips. The very thin silicon chip contains a number of small storage cells that can hold data. Instead of being made up of a series of discrete components, these units are constructed as integrated circuits, meaning that a number of transistors are integrated or combined together on a thin silicon wafer to form a complete set of circuits. The faster and more expensive bipolar semi conductor chips are often used in the arithmetic-logic unit and high-speed buffer storage sections of the CPU, while the slower and less expensive chips that employ metal-oxide semi-conductor (MOS) technology are used in the main memory



Information Technology

section.

Both *volatile* and *non-volatile* forms of semiconductor memory exist. In modern computers, primary storage almost exclusively consists of dynamic volatile semiconductor memory or dynamic random access memory. A back-up uninterruptible power system is thus desirable in installations with volatile semi-conductor storage. In spite of the volatile storage characteristic, these memory chips have found their way into the newer model of most computers due to several very good reasons.

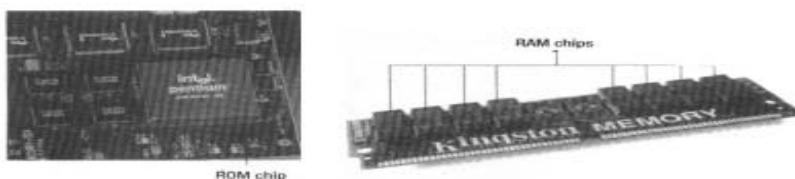


Fig. 1.6.3.1 RAM Chips

(ii) Random-Access-Memory (RAM) :The memory system constructed with metal-oxide semi conductor storage elements that can be changed is called a random access memory (RAM) When people talk about computer memory in connection with microcomputer, they usually mean the volatile RAM memory. The purpose of RAM is to hold programs and data while they are in use. It is called random access memory since access time in RAM is independent of the address of the word, that is, each storage location (address) inside the memory is as easy to reach as any other location and takes the same amount of time. One can reach into the memory at random and insert or remove numbers in any location at any time. A random access memory is extremely fast but can also be quite expensive.

RAMs can be further divided according to the way in which the data is stored, into dynamic RAMs and static RAMs. The computer designer's decision which to use where depends on what their function is to be, and on their speed and cost.

Dynamic RAM: Dynamic RAM (DRAM) is the most common type of main memory. It is dynamic because each memory cell quickly loses its charge so it must be refreshed hundreds of times each second to prevent data from being lost. Here are some of the types of DRAM that have been or will be popular in most desktop systems (listed from oldest to newest):

- Fast Page Mode (FPM) DRAM was used in most computers until EDO RAM came along.



Introduction To Computers

- Extended Data Out (EDO) DRAM is slightly faster than FPM. One variation called burst EDO (BEDO) DRAM assumes that the next data - address to be requested by the CPU follows the current one so it sends that also.
- Synchronous DRAM (SDRAM) can synchronize itself with the clock that controls the CPU. This makes data transfers more reliable and faster because timing delays are eliminated. It is anticipated that this form of memory will replace EDO as the most common form of memory.
- Rambus DRAM (RDRAM) is the latest design and Intel has announced that all of their future systems will require it. RDRAM is very fast, but the system must be slightly redesigned to use it. RDRAM sends data down a high-band width “channel” 10 times faster than standard DRAM.

Static RAM: Static RAM (SRAM) is like DRAM but it's a lot faster, larger, and more expensive. It's static because it doesn't need to be continually refreshed. Because of its speed, SRAM is used mainly in a special area of memory called a cache.

The Static RAM retains the stored data as long as the power remains in, whereas with dynamic RAM, the stored information disappears after a few milliseconds have elapsed. The data must, therefore be repeatedly refreshed before it disappears. The power consumption of a dynamic RAM is less than that of a static RAM, which has the advantage of making a higher degree of integration possible. The computer does the refreshing process itself, taking time out from other chores every few milliseconds. It will read all the RAM memory positions while they are still readable and put appropriate new charge on each capacitor. Some dynamic RAM memory circuits include built-in “refresh circuits” to relieve the computer.

(iii) Read-Only-Memory (ROM) : Another type of computer memory is the read-only-memory (ROM) It is used for microprograms not available to normal programmers. The term read-only means that the storage cannot be altered by regular program instructions. The information is stored permanently in such memory during manufacture. The information from the memory may be read out but fresh information cannot be written into it. The microprograms in read-only-memory may be used for a variety of purposes, but a common use is to hold a set of instructions that are needed frequently, for executing small, extremely basic operations, which are not otherwise available in the computer circuitry. One set of instructions found in ROM is called the ROM-BIOS which stands for Read-only Memory Basic Input Output services. These programs perform the basic control and supervisory operations for the computer. For example, it ensures that if a user pressed ‘one’ on the keyboard, the digit 1 appears on the, screen. ROM may be used for code converter, function generator (e.g. sine, cosine, Arctangent etc.) and character generators (e.g. characters displayed in dot matrix form) It also handles the basic needs of the hardware involved, which include all I/O devices.



Information Technology

PROM: Programmable Read Only Memory is a non-volatile memory which allows the user to program the chip with a PROM write. The chip can be programmed once, there after, it can not be altered.

EPROM: EPROM stands for Erasable Programmable Read Only Memory. EPROM chips can be electrically programmed. Unlike ROM and PROM chips, EPROM chips can be erased and reprogrammed. Erasure is performed by exposing the chip to Ultra-violet light.

EEPROM: Electrically Erasable Programmable Read Only Memory is EPROM. However, the data can be erased by applying electrical charges.

(iv) Bubble Memory : Bubble memory is composed of small magnetic domains (bubbles) formed on a thin single-crystal film of synthetic garnet. These magnetic bubbles, which are actually magnetically charged cylinders, only a few thousandths of a centimeter in size, can be moved across the garnet film by electric charges. The presence or absence of a bubble can be used to indicate whether a bit is “on” or “off”.

Since data stored in bubble memory is retained when power to the memory is turned off, it can be used for auxiliary storage. Bubble memory has high potential because of its low production costs and its direct access capabilities, thus it may become widely employed as a main memory technology. Since it is small, lightweight, and does not use very much power, bubble memory is finding a great deal of use as an auxiliary storage in portable computers. It is expected that as more portable computers are developed, bubble memory will become more widely used.

(v) Flash memory: Flash memory chips are one of the latest storage devices. These chips, a form of static RAM (SRAM) chips, store data much like those used in the computer’s primary storage. However, the data stays recorded even when the power is turned off-flash memory is non-volatile. Since flash memory devices have no moving parts, and are therefore very fast, they may eventually replace slower, mechanical hard disk drives.

(vi) Video RAM: Video RAM (VRAM) is used to accelerate the display of graphics on the screen. It does this by using two “ports,” one connected to the CPU and the other to the screen. Data flows in one port and out the other very smoothly. A variation of this is Window RAM (WRAM) that supports memory.

1.7 SECONDARY STORAGE DEVICES

As discussed in section 1.6.1, there are different types of computer storage. Primary storage is built into the CPU whereas secondary storage or auxiliary storage is usually housed in a



separate unit or units. Primary storage is very fast-its contents can be accessed in millionth or billionths of a second. But primary storage has a limited capacity. Although the cost per byte has continued to decrease, there is not enough capacity to store all of firms' files. Secondary storage supplements the capacity of primary storage. Secondary storage has an almost infinite capacity measured in millions and billions of bytes. Some secondary storage media (such as magnetic storage) offer direct access to data whereas tape devices offer sequential access. The access speed of secondary storage is slower than that of primary storage. It may be noted that these auxiliary storage media such as floppy disk, magnetic disk etc. can be used for both input and output purpose. We will now discuss these auxiliary storage media and devices.

1.7.1 FLOPPY DISKETTES

In the early 1970's IBM introduced a new medium for storing data. This medium consisted of a circular piece of thin plastic material, approximately eight inches in diameter, that was coated with an oxide material. The circular piece of plastic, called a disk, is enclosed in a square protective jacket with a cut out so that the magnetic surface is exposed. When inserted in the appropriate hardware device, the disk is rotated inside the protective jacket, allowing keyed data or data from main computer memory to be stored on the rotating disk. Once data is stored on the disk, it can be read from the disk into main computer. This medium for input and auxiliary storage is called a *floppy disk or diskette* (see figure 1.7.1.1 and 1.7.1.2)

Diskettes are available in a number of different sizes. The original diskette was of the size of 8 inches. During the 1980, most PCs used 5.25-inch diskettes. Today, the 3.5-inch diskette has largely replaced its 5.25-inch cousin. The size refers to the diameter of the disk, not to the capacity.



Information Technology

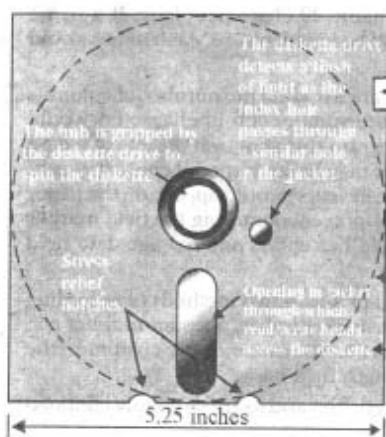


Fig. 1.7.1.1 5.25" Floppy

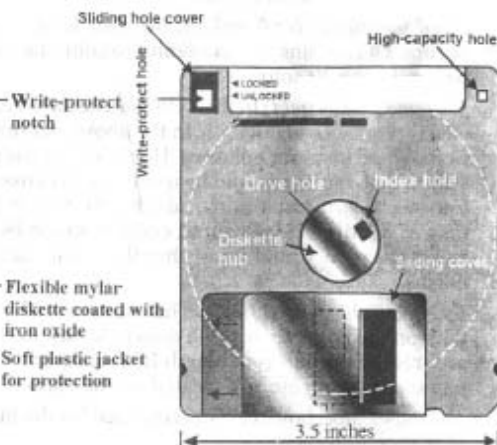


Fig. 1.7.1.2 3.5" Floppy

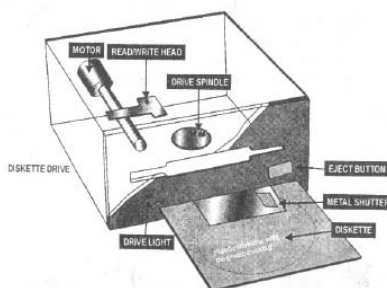


Fig. 1.7.1.3 Floppy Drive

The 5.25-inch type diskette is encased in a flexible vinyl envelope with an oval cutout that allows the read/write head to access the disk. The 3.5-inch type diskette is encased in a hard plastic shell with a sliding metal cover. When the disk is inserted into the drive, the cover slides back to expose the diskette to the read/write head.

The surfaces of diskettes (or disks/tapes discussed later) are coated with millions of tiny iron particles so that the data can be stored on them. Each of these particles can act as a magnet, taking on a magnetic field when subjected to an electromagnet. The read/write heads of a diskette drive (or a hard disk/ tape drive) contain electromagnets, which generate magnetic field in the iron on the storage medium as the head passes over the diskette (or disk or tape)



The diskette drive (Fig. 1.7.1.3) includes a motor that rotates the disk on a spindle and read/write heads that can move to any spot on the disk's surface as the disk spins. This capability is important, because it allows the heads to access data randomly, rather than sequentially. In other words, the heads can skip from one spot to another, without having to scan through everything in between.

Floppy diskettes spin at around 300 revolutions per minute. Therefore, the longest it can take to position a point on the diskette under the read/write heads is the amount of time required for one revolution - about 0.2 second. The farthest the head would ever have to move is from the centre of the diskette to the outside edge. The heads can move from the center to the outside edge in even less time - about 0.17 second. Since both operations (rotating the disk and moving the heads) take place simultaneously, the maximum time to position the heads over a given location on the diskette - known as the maximum access time remains the greater of the two times or 0.2 second.

1.7.1.1 How Data is organised on a disk : When the new diskettes (or a new hard drive) are purchased, the disks inside are nothing more than simple, coated disks encased in plastic. Before the computer can use them to store data, they must be magnetically mapped so that the computer can go directly to a specific point on the diskette without searching through data. The process of mapping a diskette is called **formatting** or **initializing**. Today, many diskettes come preformatted for either PCs or Macs. If unformatted diskettes are purchased, one must format them before they can be used. The computer will warn the user if this is the case, and will format the diskette for him if he wishes so.

The first thing a disk drive does when formatting a disk is to create a set of magnetic concentric circles called *tracks*. The number of tracks on a disk varies with the type (most high-density diskettes have 80) The tracks on a disk do not form a continuous spiral like those on a phonograph record - each one is a separate circle. Most tracks are numbered from the outermost circle to the innermost, starting from zero.

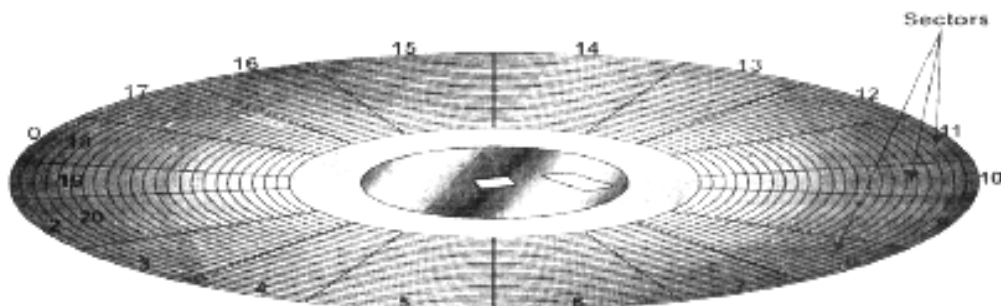


Fig. 1.7.1.1.1 Tracks & Sector

Each track on a disk is also split into smaller parts. Imagine slicing up a disk the way a pie is cut. As shown in Figure-1.7.1.1.1, each slice would cut across all the disk's tracks, resulting in short segments, or *sectors*. All the sectors on the disk are numbered in one long sequence, so the computer can access each small area on the disk with a unique number. This scheme effectively simplifies what would be a set of two-dimensional coordinates into a single numeric address.

When people refer to the number of sectors a disk has, the unit they use is sectors per track - not just sectors. If a diskette has 80 tracks and 18 sectors per track, it has 1440 sectors (80×18) - not 18 sectors.

Like any flat object, a disk has two sides. Some early drives could read data on only one side, but today, all disk drives can read and write data on both sides of a disk. To the computer, the second side is just a continuation of the sequence of sectors. For example, the 3.5 inch, 1.44-MB diskette has a total of 2880 sectors ($80 \text{ tracks per side} \times 2 \text{ sides} \times 18 \text{ sectors per track}$)

On most diskettes, a sector contains 512 bytes, or 0.5 KB

A sector is the smallest unit with which any disk drive (diskette drive or hard drive) can work. Each bit and byte within a sector can have different values, but the drive can read or write only whole sector at a time. If the computer needs to change just one byte out of 512, it must rewrite the entire sector.

The number of characters that can be stored on a diskette by a disk drive is dependent on following three basic factors:



Introduction To Computers

- 1 *The number of sides of the diskette used* : The earlier diskettes and drives were designed so that data could be recorded on only one side of the diskette. These drives were called single-sided drives. Now-a-days diskette drives are manufactured that can read and write data on both sides of the diskette. Such drives are called *double-sided drives*. The use of double-sided drives and diskettes approximately doubles the number of characters that can be stored on the diskette.
- 2 *The recording density of the bits on a track* : The recording density refers to the number of bits that can be recorded on a diskette in one inch circumference of the innermost track on the diskette. This measurement is referred to as bits per inch (bpi) For the user, the diskettes are identified as being either *single density (SD)* or *double density (DD)* A single density drive can store 2,768 bits per inch on the innermost track. Double density can store 5,876 bits per inch. With improved technology, it is anticipated that recording densities in excess of 10,000 bits per inch will be possible.
- 3 *The number of tracks on the diskette* : The number of tracks is dependent upon the drive being used. Many drives record 40 tracks on the surface of the diskette. Other drives, however, can record 80 tracks on the diskette. These drives are sometimes called double track drives.

Table given below shows how the capacities of diskettes relate to the dimensions.

Table : Formatting specifications for various Disks

Diameter (inches)	Sectors/ Sides		Bytes/ Tracks		Sector	Bytes	KB	MB
		Tracks		Sectors				
5.25	2	40	9	720	512	368,640	360	.36
5.25	2	40	18	1440	512	737,280	720	1.2
3.5	2	80	15	2400	512	1,228,800	1,200	.7
3.5	2	80	18	2880	512	1,474,560	1,440	1.44
3.5	2	80	36	5760	512	2,949,150	2,880	2.88

Because files are not usually a size that is an even multiple of 512 bytes, some sectors contain unused space after the end of the file. In addition, the DOS and Windows operating systems allocate groups of sectors, called clusters, to each of the files they store on a disk. Cluster sizes vary, depending on the size and type of the disk, but they can range from 4 sectors for diskettes, to 64 sectors for some hard disks. A small file that contains only 50 bytes will use only a portion of the first sector of a cluster assigned to it, leaving the remainder of the



Information Technology

first sector, and the remainder of the cluster, allocated but unused.

1.7.1.2 How the Operating System Finds Data on a Disk : A computer's operating system is able to locate data on a disk (diskette or hard drive) because each track and sector is labeled, and the location of all data is kept in a special log on the disk. The labeling of tracks and sectors is called performing a logical or soft format. A commonly used logical format performed by DOS or Windows creates these four disk areas:

The boot record

The file-allocation table (FAT)

The root folder or directory

The data area

The boot record: It is a small program that runs when the computer is started. This program determines whether the disk has the basic components of DOS or Windows that are necessary to run the operating system successfully. If it determines that the required files are present and the disk has a valid format, it transfers control to one of the operating system programs that continues the process of starting up. This process is called booting because the boot program makes the computer "pull itself up by its bootstraps."

The boot record also describes other disk characteristics, such as the number of bytes per sector and the number of sectors per track. The information is required by the operating system to access the data area of the disk

The file-allocation table (FAT): It is a log that records the location of each file and the status of each sector. When a file is written to a disk, the operating system checks the FAT for an open area, stores the file, and then identifies the file and its location in the FAT.

The FAT solves a common filing problem: What happens when a user loads a file, increases its size by adding text to it, and then save it again? For example, say, one needs to add 5000 bytes to a 10,000-byte file that has no open space around it. The disk drive could move the surrounding files to make room for the 5000 bytes, but that would be time consuming. Instead, the operating system checks the FAT for free areas, and then places pointers in it that link together the nonadjacent parts of the file. In other words, it splits the file up by allocating new space for the overflow.

When the operating system saves a file in this way, the file becomes fragmented. Its parts are located in nonadjacent sectors. Fragmented files do cause undesirable side effects, the most significant being that it takes longer to save and load them.



Introduction To Computers

Users do not normally need to see the information in the FAT, but they often use the folder information. A **folder**, also called a directory, is a tool for organizing files on a disk. Folders can contain files or other folders, so it is possible to set up a hierarchical system of folders on the computer, just as there are folders within other folders in a file cabinet. The top folder on any disk is known as the root. When the user uses the operating system to view the contents of a folder, the operating system lists specific information about each file in the folder, such as the file's name, its size, the time and date that it was created or last modified, and so on.

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

1.7.1.3 Care required in using and storing a diskette : On receiving a new diskette, it should be inspected for sign of obvious damage. The surface of the diskette should not be touched with hand or some sharp object. Write-protect precaution should be observed by peeling off or sticking on (as applicable) the aluminum square on the notch.

Correct insertion of disk in the disk drive is essential, otherwise some data stored on the disk is likely to be destroyed or the disk itself may get damaged. The diskette should be inserted slowly in the disk drive only when power to the entire computer system is on. It should be removed prior to turning the system off.

As a defensive measure, it is advisable that a back-up copy of the information stored on each diskette be prepared and stored separately at a safe location. The diskette should be properly labeled for right identification. While storing a diskette, both physical and environmental factors should be considered. Diskette should not be stored in such a way that may sag, slump or compress it. The main enemies of a diskette are temperature and direct sunlight, dust, liquids and vapors and electromagnetic interference. Diskette should be protected from them. Care should be taken to clean the disk drive head to remove dust regularly.

Floppy diskettes are very cheap and offer both sequential and direct access to data at a substantially high speed. Typically, data may be transferred at the rate of 30,000 to 1,50,000 bytes per second. Records and files on a flexible disk are organised and processed in the same way as with rigid disk systems.

Floppy disk drives are generally smaller and more economical to manufacture than rigid disk systems. That is why these are used as auxiliary storage and I/O media with mini and microcomputer installations. In Mainframes also, these are being used as input medium.

1.7.2 MAGNETIC DISC

Magnetic discs are the most popular *direct access medium*. By direct access we mean that a



Information Technology

record can be accessed without having to plod through the preceding records. The other direct access medium is floppy disc (discussed earlier)

Although a shift toward optical technology is occurring, the hard disk is still the most common storage device for all computers. Much of what was discussed about floppy diskettes and disk drives apply to hard disks as well. Like diskettes, hard disks store data in tracks that are divided into sectors. Physically, however, hard disks look quite different from diskettes.

A hard disk is a stack of one or more metal platters that spin on one spindle like a stack of rigid diskettes. Each platter is coated with iron oxides, and the entire unit is encased in a sealed chamber. Unlike diskettes, where the disk and drive are separate, the hard disk and drive is a single unit. It includes the hard disk, the motor that spins the platters, and a set of read/write heads.

Since the disk can not be removed from its drive (unless it is a removable hard disk, which will be discussed later), the terms hard disk and hard drive are used interchangeably.

Hard disks have become the primary storage device for PCs because they are convenient and cost-efficient. In both speed and capacity, they far outperform diskettes. A high-density 3.5-inch diskette can store 1.44 MB of data. Hard disks, in contrast, range in capacity from about 20 GB onward. Most PCs now come with hard disks of at least 80 GB and more.

Two important physical differences between hard disks and diskettes account for the differences in performance. First, hard disks are sealed in a vacuum chamber, and second, the hard disk consists of a rigid metal platter (usually aluminum), rather than flexible mylar.

The rigidity of the hard disk allows it to spin much faster - typically more than ten times faster than diskettes. Thus, a hard disk spins between 3,600 rpm and 7,200 rpm, instead of a diskette's 300 rpm. The speed at which the disk spins is a major factor in the overall performance of the drive.

The rigidity of the hard disk and the high speed at which it rotates allow a lot of data to be recorded on the disk's surface. It may be recalled, waving a magnet past an electric coil causes a current to flow through the coil. The faster the magnet is waved, and the closer the magnet is to the coil, the larger the current it generates in the coil. Similarly, a disk that spins faster can use smaller magnetic charges to make current flow in the read/write head. The drive's heads can also use a lower intensity current to record data on the disk.

1.7.2.1 Data Storage : Not only do hard disks pack data more closely together, they also hold more data, because they often include several platters, stacked one on top of another. To the computer system, this configuration just means that the disk has more than two sides; in addition to a side 0 and side 1, there are sides 2, 3, 4 and so on. Some hard disk drives hold as



many as 12 sides, but both sides of the disks are not always used

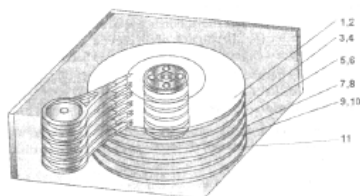


Fig. 1.7.2.1.1 Hard Disk

With hard disks, the number of read/write heads specifies the number of sides that the disk uses. For example, a particular hard disk drive might have six disk platters (that is, 12 sides), but only eleven heads, indicating that one side is not used to store data. Often, this is the bottom side of the bottom disk.

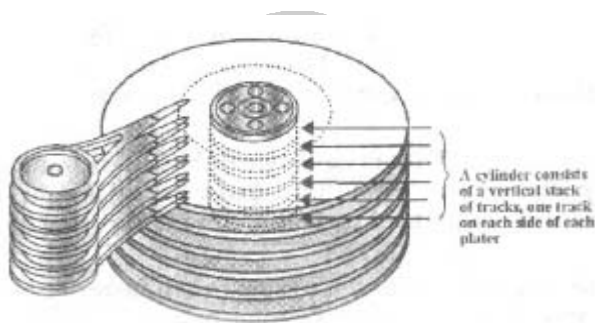


Fig. 1.7.2.1.2 Hard Disk Cylinder

The other point of importance is the fact that read/write heads move in and out simultaneously. Referring to figure 1.7.2.1.2, there are 11 magnetisable faces and there are therefore, 11 read/write heads. Even though, a record on the first disc face were to be accessed, not only the first read/write head would move but the other ten read/write heads would also move in unison. However, only the first read head would be activated, others remaining inactive. As a consequence of this, if the read/write heads have once been moved, all the eleven tracks vertically above and below each other should be read or written before any further movement of the heads take place. This eliminates first component of the seek time *i.e.*, horizontal movement of the read/write heads. This has led to the concept of cylinders (synonymous: seek areas) Any eleven tracks vertically above and below each other constitute more or less a hollow cylinder (See Figure 1.7.2.1.2) Therefore, there are 200 cylinders. Because of the simultaneous movement of the read/write heads, it is to be desired that the



Information Technology

records are arranged sequentially in cylinders so that when the first cylinder (*i.e.*, first track of all eleven faces) has been read, the heads move to the next cylinder *i.e.*, reading or writing is performed cylinder wise.

Like diskettes, hard disks generally store 512 bytes of data in a sector, but because of their higher tolerances, hard disks can have more sectors per track-54, 63, or even more sectors per track are not uncommon.

The computation of a hard disk's capacity is identical to that for diskettes-but the numbers are larger. The breakdown of the storage capacity for a disk that is sold as a 541-MB disk is given below:

$$\begin{aligned} 1,632 \text{ cylinders} \times 12 \text{ heads (sides)} &= 19,584 \text{ tracks} \\ 19,584 \text{ tracks} \times 54 \text{ sectors / track} &= 1,057,536 \text{ sectors} \\ 1,057,536 \text{ sectors} \times 512 \text{ bytes/sector} &= 541,458,432 \text{ bytes} \end{aligned}$$

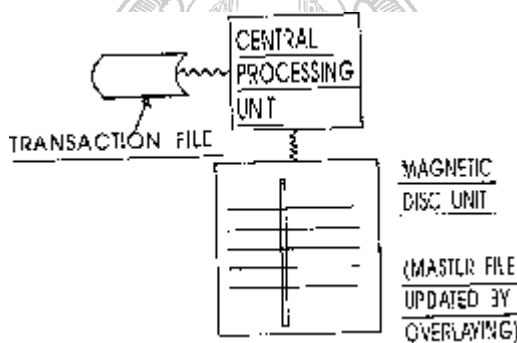


Fig. 1. 7.2.1.3

As depicted in the Figure 1. 7.2.1.3, only one magnetic pack is connected to the CPU and serves both as the input and output unit. The implication of this being that a record picked up from a sector and updated in the CPU is deposited back in the same place in that sector. *This mode of storing updated data is known as overlaying i.e.*, original data is automatically erased when the new updated record is being deposited in its place.

This leads to economies in the sense that fewer discs are needed but as a disadvantage it becomes difficult to trace any errors as it is not possible to reconstruct the latest file from the previous version.

In spite of all the capacity and speed advantages, hard disks have one major drawback. To



achieve optimum performance, the read/write head must be extremely close to the surface of the disk. In fact, the heads of hard disks fly so close to the surface of the disk that if a dust particle, a human hair, or even a fingerprint were placed on the disk it would bridge the gap between the head and the disk, causing the heads to crash. A head crash, in which the head touches the disk, destroys the data stored in the area of the crash and can destroy a read/write head, as well.

The time required in accessing a record has generally three components.

- (i) *Seek Time* : This is the time required to position a movable read-write head over the recording track to be used. If the read-write head is fixed, this time will be zero.
- (ii) *Rotational time* : This is the rotational delay also termed latency, to move the storage medium underneath the read-write head.
- (iii) *Data transfer time* : This is the time taken to activate the read/write head, read the requested data, and transmit them to primary memory for processing.

The total of these three components is known as the access time and typically ranges from, 8 to 12 milliseconds.

1.7.2.2 Advantages and disadvantages of magnetic disk

The advantages of magnetic disk include :

1. Magnetic rigid disk is a direct access storage medium; therefore, individual records can be retrieved without searching through the entire file.
2. The costs of disks are steadily declining.
3. For real-time systems where direct access is required, disks are currently the only practical means of file storage. Other new types of storage, such as bubble storage, are not widely used yet.
4. Records can be readily updated by writing the new information over the area where the old information was stored.
5. With removable disk packs, a single disk drive can store large quantities of data although all but one of the disks is offline at any given point in time. However, being offline is not a disadvantage for many applications, especially batch applications.
6. Interrelated files stored on magnetic disk can allow a single transaction to be processed against all of these files simultaneously.

The disadvantages of magnetic disk include :



Information Technology

Updating a master file stored on disk destroys the old information. Therefore, disk does not provide an automatic audit trail. When disk is used, back-up and audit trail require that each old master file records be copied to another medium prior to update.

1.7.2.3 Removable hard disks: Removable hard disks and drives attempt to combine the speed and capacity of a hard disk with the portability of a diskette. There are many different types of devices in this category. Choosing the best type is usually a matter of balancing the needs for speed, storage capacity, compatibility (will it work in different computers?), and price.

1.7.2.4 Hot-Swappable Hard Disks: At the high end, in terms of both price and performance, are hot-swappable hard disks. These are sometimes used on high-end workstations that require large amounts of storage. They allow the user to remove (swap out) a hard disk and insert (swap in) another while the computer is still on (hot)

Hot-swappable hard disks are like removable versions of normal hard disks: the removable box includes the disk, drive, and read/write heads in a sealed container.

1.7.3 OPTICAL LASER DISKS

Optical laser disk storage is capable of storing vast amount of data. Some industry analysts have predicted that optical laser disk technology may eventually make magnetic disk and tape storage obsolete. With this technology, the read/write head used in magnetic storage is replaced by two lasers. One laser beam writes to the recording surface by scoring macroscopic pits in the disk, and another laser reads the data from the light sensitive recording surface. A light beam is easily deflected to the desired place on the optical disk, so a mechanical access arm is not needed.

There are three main categories of optical laser disks.

1.7.3.1 CD-ROM disks : CD-ROM, a spinoff of audio CD technology stands for compact-disk-read-only memory. The name implies its applications, CD-ROM disks are created as a mastering facility. Most of the commercially produced read-only CD-ROM disks contain reference material. The master copy is duplicated or “pressed” at the factory and copies are distributed with their pre-recorded contents. Once inserted into the CD-ROM disk drive, the text, video images, and so on can be read into primary storage for processing or display. However, the data on the disk are fixed, they can not be altered.

The capacity of a single CD-ROM is over 650 MB which is equivalent to 250,000 pages of text, or 1500 floppy disks. The tremendous storage capacity has opened the door to a variety of multimedia applications. Multimedia can provide needed flexibility during a presentation. Unlike a video tape, CD-ROM gives the presenter instant random access to any sequence of



images on the disk.



Fig. 1.7.3.1.1 CD-ROM

CDs may soon take a new direction with the advent of DVD, digital video disk, a high density medium that is capable of storing a full-length movie on a single disk side of a CD (actually, it uses both sides of the disk) DVDs look like CDs and DVD - ROM drives are able to play current CD-ROMs. A slightly different player, the DVD movie player connects to the TV and plays movies like a VCR. The DVD movie player will also play audio CDs.

Each side of a DVD can hold up to 4.7 GB. Therefore, these two-sided disks can contain as much as 9.4 GB of data.

CD Rewritables : Hewlett Packard has introduced the next generation of CD Rewritable (CD-RW) drive. This is the third generation in CD technology which began with CD-ROM and was then followed by the CD-Recordable (CD-R) and CD-RW.

CD-R : It stands for compact disc, recordable. A person can only write once on this CD, though it can be read as many times as wished. It can be played in CD players and CD-ROM drives. In a normal CD, polycarbonate plastic substrate, a thin reflective metal coating, and protective outer coating layers are present. However, in a CD-R, an extra layer is present and is an organic polymer dye lying between the polycarbonate and metal layers and serves as a recording medium. A pregrooved spiral track guides the laser for recording data, which is encoded from the inside to the outside of a CD in a continuous spiral, much like the way it is read. The laser makes that are not dissimilar from the pits and lands of a normal CD. After the encoding process is completed, the data can be read as a normal CD. The CD recorders are sometimes referred to as CD burners. Modern recording devices can also read CDs as well. CD-Rs can be created in any CD-R or CD-RW drive.

CD-RW :The rewriteable compact disc is called a CD-RW. This disc allows for repeated recordings on a disc. It is relatively more expensive than the CD-R. However, in certain circumstances the benefits outweigh the cost.



Information Technology

While the CD-R may use a layer of organic dye, which can only be altered once, the CD-RW uses an alloy that can change to and fro from a crystalline form when exposed to a particular light. The technology of this process has a special name called phase changing. The patterns, however, are less distinct than that of other ordinary CD formats due to the greater difficulty of manipulating a metal instead of a dye. The alloy is usually made up of silver, indium, antimony and tellurium. After heating to a particular temperature, the alloy will crystallize when cooled. Heating that particular spot to a greater temperature results in the substance becoming amorphous when cooled. By controlling the temperature, some areas have crystals and others do not. The crystals will reflect the laser effectively while the non-crystalline areas would absorb most of it.

To rewrite on a CD-RW, the alloy is first made amorphous, then reshaped using the cooler laser. The CD-RW can be rewritten as many as 1000 times. With the rewritable drive, the CD has now got all the functionalities of the floppy drive. The drive is claimed to be simple to use and universally accepted as a conventional floppy disk drive, but with the qualities of a compact disc.

1.7.3.2 WORM disks : It stands for Write once, read many optical laser disks, or WORM disks. These are used by end user companies to store their own proprietary information. Once the data have been written to the medium, they only can be read, not updated or changed. The PC version of a WORM disks cartridge, which looks like a 5¼ inch version of the 3½ inch diskette, has a capacity of 200 MB.

Access times for CD-ROM and WORM drives tend to be quite slow by hard disk drive standards, ranging from 100 to 300 milliseconds.

The WORM disks cartridge is a feasible alternative to magnetic tape for archival storage, for example, a company might wish to keep a permanent record of all financial transactions during the last year. Another popular application of WORM disks is in information systems that require the merging of text and images that do not change for a period of time. A good example is an “electronic catalogue”. A customer can peruse retailer’s electronic catalogue on a VDT, or perhaps a PC, and see the item while he or she reads about it. And, with a few keystrokes, the customer can order the item as well. The Library of Congress is using WORM technology to alleviate a serious shelf-space problem.

1.7.3.3 Magneto-Optical Disks : Magneto-optical integrate optical and magnetic disk technology to enable read-write storage. The 5¼ inch disks store up to 1000 MB. However, the technology must be improved before the disks can experience widespread acceptance. At present, magneto-optical disks are too expensive and do not offer anywhere near the kind of



reliability that users have come to expect of magnetic media. In addition, the access times are relatively slow, about the same as a low-end Winchester disk.

As optical laser disk technology matures to offer reliable, cost - effective, read/write operation; it eventually may dominate secondary storage in the future as magnetic disks and tape do today.

640 MB Magneto Optical Drive : A **magneto-optical drive** is a kind of optical disc drive capable of writing and rewriting data upon a **magneto-optical disc**. Both 5.25" and 3.5" form factors exist. The technology was introduced at the end of the 1980s. Although optical, they appear as hard drives to the operating system and do not require a special filesystem (they can be formatted as FAT, HPFS, NTFS, etc.)

Initially the drives were 5.25" and had the size of full-height 5.25" hard-drives (like in IBM PC XT) Today a 3.5" drive(See fig. 1.7.3.3.1) has the size of 1.44 megabyte diskette drive. 5.25" media looks a lot like a CD-ROM enclosed in an old-style cartridge while 3.5" media is about the size of a regular 1.44MB floppy disc, but twice the thickness. The cases provide dust resistance, and the drives themselves have slots constructed in such a way that they always appear to be closed.



Fig 1.7.3.3.1 Magneto Optical Disk

The disc consists of a ferromagnetic material sealed beneath a plastic coating. There is never any physical contact during reading or recording. During reading, a laser projects a beam on the disk and according to the magnetic state of the surface, the reflected light varies due to the magneto-optical Kerr effect. During recording, the light becomes stronger so it can heat the material up to the Curie point in a single spot. This allows an electromagnet positioned on the opposite side of the disc to change the local magnetic polarization, and the polarization is



Information Technology

retained when temperature drops.

Each write cycle requires both a pass for the laser to erase the surface, and another pass for the magnet to write the information, and as a result it takes twice as long to write data as it does to read it. In 1996, a *Direct Overwrite* technology was introduced for 3.5" discs, to avoid the initial erase pass when writing. This requires special media.

Magneto-optical drives by default check information after writing it to the disc, and are able to immediately report any problems to the operating system. This means that writing can actually take three times longer than reading, but it makes the media extremely reliable, unlike the CD-R or DVD-R technologies upon which data is written to media without any concurrent data integrity checking.

1.7.3.4 Digital Video Disk: DVD (also known as "**Digital Versatile Disc**" or "**Digital Video Disc**") is an optical disc storage media format that can be used for data storage, including movies with high video and sound quality. DVDs resemble compact discs as their physical dimensions are the same- 120 mm (4.72 inches) or occasionally 80 mm (3.15 inches) in diameter, but they are encoded in a different format and at a much higher density. A video disk can store text, video, and audio data. Video disks can be accessed a frame at a time (to provide still information) or played like a phonograph record (to supply up to an hour of moving action) Any of the 54,000 tracks on the surface of typical video disk can be accessed in about three seconds.

A digital video disk (DVD) is a 5 inch plastic disk that uses a laser to encode microscopic pits in its substrate surface. But the pits on a DVD are much smaller and are encoded much closer together than those on a CD-ROM. Also, a DVD can have as many as two layers on each of its two sides (compared to the single-layered, single-sided CD-ROM) The end result is a medium that can hold as much as 17 gigabytes of data – over 25 times the capacity of a standard CD-ROM disk. The advantages of DVDs are therefore self-evident – a huge storage capacity that enables users to archive large amounts of data on a single, lightweight, removable, reliable, easily-transportable medium. Although VDDs are now used mostly for entertainment – for example, storing video movies or large amounts of prerecorded music-experts predict that DVDs will become the medium of choice for distributing software or archiving large amounts of accounting data.

Video disks were first introduced in 1983, as a video game product. Today, however, they can provide companies with a competitive advantage.

Video disk systems were developed to help real estate agents conduct better searches for homes and properties for their clients. For example, the client describes the type of home desired – perhaps three bedrooms, a garage, and priced below Rs.200,000. When these data



are entered into the video disk system, photographs and even “ video tours” of existing homes meeting the description can be summoned to the display screen.

Video disks are widely used for training applications. At a growing number of companies – Ford, Chrysler, Xerox, Pfizer, and Massachusetts Mutual Life Insurance, to name just a few- video disk systems take on such training tasks as showing how to boost factory performance, helping service technicians do a safer and better job, and training clerks to analyze insurance applications. The U.S. Army has also made extensive use of video disks for training purposes.

Video disks are also used by automobile manufacturers to show their lines and by travel agents to interest clients in resorts. In the future, some industry observers predict that many businesses will develop automatic customer service centres equipped with video disk components so that consumers do not have to depend only on clerks and showrooms. When a desired item flashes on the display screen, the customer can insert a credit card in a device that resembles a bank’s automatic teller machine and order that item immediately. Sears-chain of departmental store introduced systems like this in many of its department stores. Even the U.S. Postal Service is spending close to Rs.5 million to develop an automated video disk system that will allow its patrons to do many of the activities that human postal clerks now perform.

1.7.4 TAPE DEVICES

Magnetic tape is probably the oldest secondary storage technology still in wide use. Its biggest drawback is that it can only access data sequentially. However, many data processing operations are sequential or batch oriented in nature, and tape is still economical. Here we will look at the two most popular forms of magnetic tape for large system MIS applications: detachable reel magnetic tapes and tape cartridges.

1.7.4.1 Detachable Reel Magnetic Tapes : Many of the tapes used with mainframes and minicomputers are stored on detachable reels. These plastic tapes are, like disks, coated with a magnetizable surface (often iron oxide) that can be encoded with 0 and 1 bits . Tapes come in various widths, lengths, and data densities. A common specification is a 2400 feet reel of ½ inch diameter tape that packs data at 6250 bytes per inch. Recording densities of tapes are often cited as bytes per inch (bpi) because in most instances, a character (byte) is represented as a vertical slice of bits across the tracks of tape surfaces.

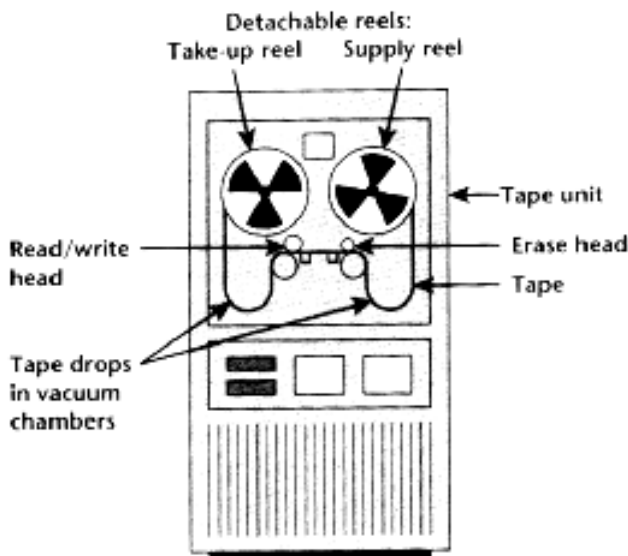


Fig 1.7.4.1.1

Tapes are read on a hardware device called a tape unit (Figure 1.7.4.1.1). Basically, this unit works the same way as the reel-to-reel tape units that were once popular on home stereo systems. An empty take-up reel, running at the same speed as the supply reel on which the tape is initially wound, accepts the tape as it is being processed. The tape is processed by passing it under read/write heads located between the two reels. Depending on the instructions given to the computer system, data can then either be read from the tape or written to it.

1.7.4.2 Tape Cartridge Systems : Cartridge tapes represent the leading edge of tape technology. Tape cartridges are available for both large and small computer systems.

Tape cartridges for microcomputer systems, which resemble cassette tapes in appearance, are frequently used to back up hard disks. These tapes, which are not designed for processing purposes, are sometimes called streaming tapes. The capacities of these tapes vary, but several megabytes of storage are typical. Streaming tapes can usually back up the contents of a hard disk in a few minutes. Among the leading tape cartridge system vendors are Colorado Memory Systems, Everex Systems, Micro Solutions Summit Memory Systems, and Tallgrass Technologies Corporation.

In 1986, IBM introduced a tape cartridge system, the IBM 3480, for its 3090 line of



Introduction To Computers

mainframes. Each of these cartridges has a capacity of 200 MB and a data-transfer rate of 3 MB/Sec. Unlike conventional detachable-reel tapes, which uses 9 parallel tracks, these ½ inch tapes store data in 18 tracks. In 1992, IBM released 36 track tape cartridges for use in its mainframes and AS/400 family of midrange computers.

Self-examination questions

1. Describe in detail various generations of computers.
2. Write short notes on the following types of computers :-
 - (i) Super computer (ii) Mainframe computer.
 - (iii) Mini computer. (iv) Micro computer.
 - (v) Workstations (vi) Server
3. Draw the schematic diagram of a computer. Briefly discuss each of the components covered in it.
4. What are the features of the Central Processing Unit?
5. Discuss, in brief, various types of microprocessors.
6. Discuss various components of a motherboard.
7. What do you understand by the term 'Bus'? Discuss two types of bus available on a computer.
8. Write short notes on the following :
 - (i) RAM
 - (ii) ROM
 - (iii) Bubble memory
 - (iv) Flash memory
9. Write short note on floppy diskette as an input medium.
10. What are the factors that determine the number of characters that can be stored in a floppy diskette?
11. Explain the following terms :
 - (i) The boot record



Information Technology

- (ii) File allocation table
12. What care is required for using and storage of a diskette?
13. Differentiate between floppy diskettes and hard disks.
14. Briefly explain the various characteristics of a hard disk.
15. What are the advantages and disadvantages of direct access storage?
16. Explain the following terms :
- (i) CD-ROM
 - (ii) WORM disk
 - (iii) Magneto Optical Disks
 - (iv) Video Disk
 - (v) Detachable Reel Magnetic Tape
 - (vi) Tape Cartridge Systems

Multiples choice:

1. First Generation of computer have used the following Circuitry option
- a) Vacuum Tubes b) transistor c) IC d) VLSI
2. Second generation computer have used the following options Circuitry
- a) Vacuum Tubes b) Transistor c) IC d) VLSI
3. Third Generation Computer have used the following circuitry options
- a) Vacuum Tubes b) transistor c) C d) VLSI
4. Fourth Generation computer have used the following Circuitry options
- a) Vacuum Tubes b) transistor c) IC d) VLSI
5. First Generation have used the following language
- a) Object oriented programmer b) Artificial Interlig
- c) Assembly Language d) Assembly & M/C Language
- 5) Which is not the attribute of file :

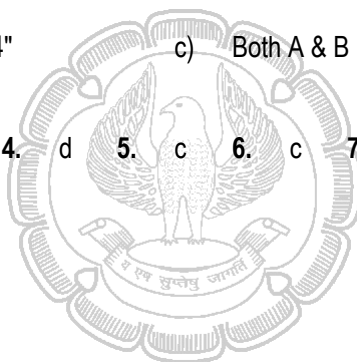


Introduction To Computers

- a) Hidden b) System c) Big d) Archive
- 6) Which is the fastest storage device :
- a) Floppy Disk b) Zip Disk c) CD - Rom d) Hard Disk
7. Storage Device which uses light to store data :
- a) Floppy Disk b) Hard Disk c) CD - ROM d) Tape Drive
8. In CD 'X' Denotes:
- a) Speed b) Storage Capacity c) Time d) None
9. In CD - R user can write data.
- a) Once b) More Than Once c) Both A & B d) None of the above
10. Size of Floppy Disk is:
- a) 3.5" b) 5.1/4" c) Both A & B d) None

Answer:

1. a 2. b 3. c 4. d 5. c 6. c 7. d 8. a 9. a
10. c





Information Technology

Unit 2 : Input and Output Devices

I/O devices (short for input/output devices) is a general term for devices that send computers information from the outside world and that return the results of computations. These results can either be viewed directly by a user, or they can be sent to another machine, whose control has been assigned to the computer: The first generation of computers were equipped with a fairly limited range of input devices. A punch card reader, or something similar, was used to enter instructions and data into the computer's memory, and some kind of printer, usually a modified teletype, was used to record the results. Over the years, other devices have been added. For the personal computer, for instance, keyboards and mice are the primary ways people directly enter information into the computer; and monitors are the primary way in which information from the computer is presented back to the user, though printers, speakers, and headphones are common, too. There is a huge variety of other devices for obtaining other types of input. One example is the digital camera, which can be used to input visual information.

We will now discuss some of these I/O devices in detail.

1.1 ON-LINE ENTRY

1.1.1. Keyboard : A microcomputer's keyboard is normally its primary input and control device. One can enter data and issue commands via the keyboard. The keyboard is used to type information into the computer or input information. There are many different keyboard layouts and sizes with the most common for Latin based languages being the QWERTY layout (named for the first 6 keys) The standard keyboard has 101 keys. Notebooks have embedded keys accessible by special keys or by pressing key combinations (CTRL or Command and P for example) Ergonomically designed keyboards are designed to make typing easier. Some of the keys have a special use. These are referred to as command keys. Three most common are the Control or CTRL, Alternate or Alt and the Shift keys though there can be more (the Windows key for example or the Command key) Each key on a standard keyboard has one or two characters. Press the key to get the lower character and hold Shift to get the upper. Besides the standard typewriter keyboard, most micro keyboards have function keys, also called soft keys. When tapped, these function keys trigger the execution of software, thus the name "soft key." For example, tapping one function key might call up a displayed list of user options commonly referred to as a **menu**. Another function key might cause a word processing document to be printed. Function keys are numbered and assigned different functions in different software packages. For example., HELP (context-sensitive user assistance) is often assigned to F1, or Function key 1. Most keyboards are equipped with key pad and cursor-



Introduction To Computers

control keys . The keypad permits rapid numeric data entry. It is normally positioned to the right of the standard alphanumeric keyboard.

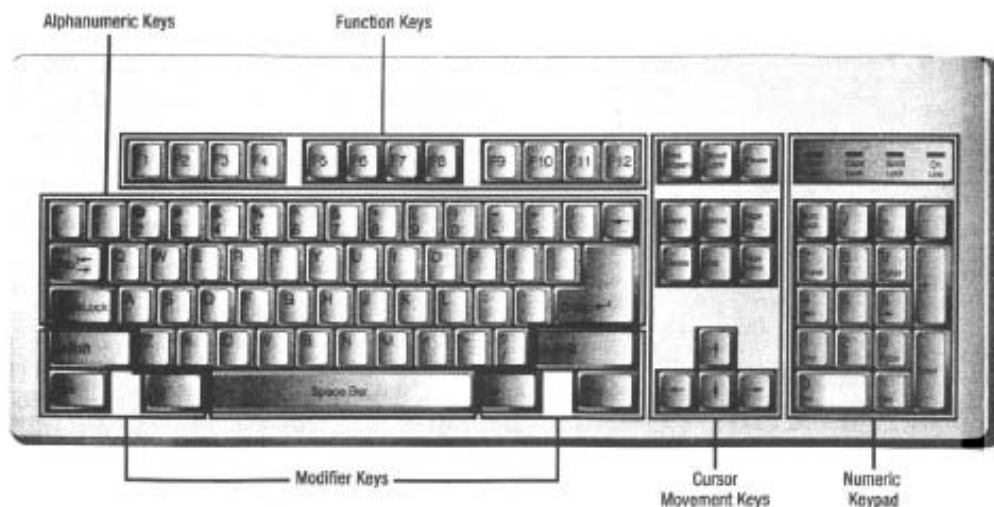


Fig. 1.1.1.1 Keyboard

The cursor-control keys, or “arrow” keys, allow the user to move the text cursor up (↑) and down (↓) usually a line at a time, and left (←) and right (→) usually a character at a time. The text cursor always indicates the location of the next keyed-in-character on the screen. The text cursor can appear as several shapes depending on the application, but frequently, one will encounter an underscore (_), a vertical line (|), or a rectangle (█). To move the text cursor rapidly about the screen, simply hold down the appropriate arrow key.

For many software packages, one can use the arrow keys to view parts of a document or worksheet that extend past the bottom, top, or sides of the screen. This is known as **scrolling**. The user can use the up and down arrow keys (↑/↓) to scroll vertically and the left and right keys (←/→) to scroll horizontally.

In summary, the keyboard provides three basic ways to enter commands:

- Key in the command using the alphanumeric portion of the keyboard.
- Tap a function key.
- Use the arrow keys to select a menu option from the displayed menu.

Other important keys common to most keyboards are the ENTER, HOME, END, PAGE UP



Information Technology

AND PAGE DOWN (abbreviated as PGUP and PGDN), DELETE (DEL), BACKSPACE (BKSP), Insert - typeover toggle (INS), ESCAPE (ESC), SPACEBAR, Shift Control (CTRL), Alternate (ALT), TAB, SCROLLLOCK, CAPSLOCK, NUMLOCK, and PRINT SCREEN keys (see Figure 1.1.1.1)

1.1.2 Mouse : The Mouse is a small box, from the bottom of which protrudes a small ball bearing. The ball bearing rotates when the user moves the Mouse across his desk and, as it is linked by cable to the microcomputer, this moves the cursor on the display screen. Another type of mouse uses an optical system to track the movement of the mouse. Most modern computers today are run using a mouse controlled pointer. Generally if the mouse has two buttons the left one is used to select objects and text and the right one is used to access menus. If the mouse has one button (Mac for instance) it controls all the activity and a mouse with a third buttons can be used by specific software programs.

Systems using Mouse have displays which include easily identified functions or programs, such as 'store', 'print', 'graphics', 'utilities' and so on. When the cursor alights on the facility required, the user presses a button on the top of the Mouse and it is activated. This is an ideal input medium for those who cannot use keyboards-or are reluctant to learn.

A mouse may have one, two or three buttons. The function of each button is determined by the program that uses the mouse. In its simplest form, a mouse has one button. Moving the mouse moves the cursor on the screen, and clicking the button results in selecting an option. A mouse normally has two or three buttons, but the software package used by the user may use one, two or all three of them.

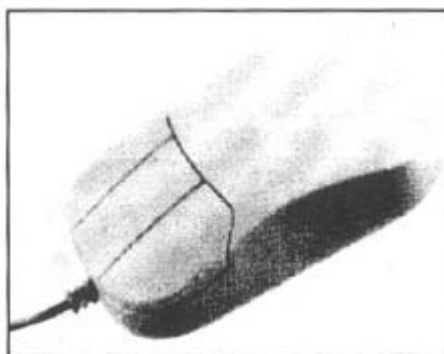


Fig. 1.1.2.1 Mouse

A mouse may be classified as a mechanical mouse or an optical mouse, depending on the



technology it uses. In a *mechanical mouse*, that projects through the bottom surface rotates as the mouse is moved along a flat surface. The direction of rotation is detected and relayed to the computer by the switches inside the mouse. Microsoft, IBM, and Logitech are some well-known makers of mechanical mouse. An *optical mouse* uses a light beam instead of a rotating ball to detect movement across a specially patterned mouse pad.

A *serial mouse* is connected to the PC through a serial port. A *bus mouse* is similar to a serial mouse except that it comes with a dedicated port and does not need a free serial port on the computer.

1.1.3 Touch Screen : The 'Touch Screen' is a Hewlett Packard innovation and was introduced on their 100 series microcomputers in 1984. An invisible microwave beam 'matrix' criss crosses the screen, emanating from holes along the bottom and sides of the display unit. By pressing the finger against a function or program displayed on the screen, the infrared beam is broken at that intersection and the system activated. In many ways, this is more effective than the 'Mouse' and very popular with users.

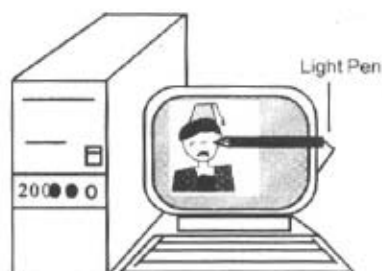


Fig. 1.1.3.1 Touch Screen

Two popular technologies exist for touch screens. In one, the screen is made sensitive to touch and the exact position is detected. In the other, the screen is lined with light emitting devices on its vertical sides, photo-detectors are placed on the horizontal sides. When the user's finger approaches the screen, the light beam is broken and is detected by the photo-detectors.

Touch screens are used in information-providing systems. For example, while performing an operation, if the doctor wants to see some test reports of the patient that have been stored in a computer, he can get the information just by touch of his finger. It is also used in airline and railway reservation counters. The user has to indicate the current place of stay and the destination by touching the screen (may be on a map), and all the possible routes with timings, rates, etc. are displayed. These interfaces are also used in travel agents offices to display the names and addresses of all hotels, restaurants, and other places of interest, at a desired destination. Touch screens are also used in stock exchanges where buying and selling of stock is done.

1.1.4 Light Pen : A light pen is a pointing device which can be used to select an option by simply pointing at it, or draw figures directly on the screen and move the figures around.





Information Technology

A light pen has a photo-detector at its tip. This detector can detect changes in brightness of the screen. When the pen is pointed at a particular point on the screen, it records the instant change in brightness that occurs and informs the computer about this. The computer can find out the exact spot

Fig.1.1.4.1 Light Pen

with this information. Thus, the computer can identify where the user is pointing on the screen.

Light-pens are useful for menu-based applications. Instead of moving the mouse around or using a keyboard, the user can select an option by pointing at it. A light pen is also useful for drawing graphics in CAD. An engineer, architect or a fashion designer can draw directly on the screen with the pen. Using a keyboard and a light pen, the designer can select colors and line thickness, reduce or enlarge drawings, and edit drawings.

These are also used to read the bar charts that are now appearing so frequently on the goods which are available in big departmental stores. By using a laser beam, computers are able to 'read' the information stored on the bar chart or on a thin strip of magnetic material and this is used to keep stock records and cheque costs, etc.

1.1.5 The Track Ball : A track ball is a pointing device that works like an upside-down mouse. The user rests his thumb on the exposed ball and his fingers on the buttons. To move the cursor around the screen, the ball is rolled with the thumb. Since the whole device is not moved, a track ball requires less space than a mouse. So when space is limited, a track ball can be a boom. Track balls are particularly popular among users of notebook computers, and are built into Apple Computer's Power Book and IBM ThinkPad notebooks.

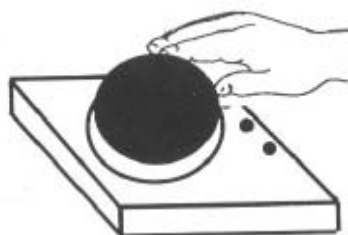


Fig. 1.1.5.1 Track Ball

1.1.6 Joystick : It is a screen pointing input device. It is a vertical lever usually placed in a ball socket, which can be fitted in any direction to control cursor movements for computer games and for some professional applications.

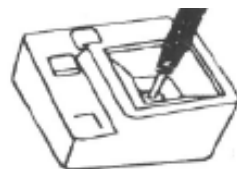


Fig 1.1.6.1 Joystick

1.1.7 Display Devices : Virtually everyone who interacts with a computer system today uses some type of display device. These peripheral hardware units consist of a television like viewing screen, to which computer output is sent. The two most common types of display devices found today are monitors and terminals.

Monitors are the devices found most commonly with microcomputer systems. As mentioned



Introduction To Computers

previously, a monitor is just a “box with a viewing screen”. On the screen, the user is able to see not only what is entered into the computer, but the computer output as well.

A **computer terminal** or video display terminal (VDT), generally combines input and output functions. It consists of a QWERTY keyboard for inputting information direct to the computer, and either a printer or a TV screen for displaying information from the computer. Terminals are most commonly found in settings that are remote from the main computer and they interact

with the computer through communications lines or networks. Airline agents are familiar examples of people who use communications terminals. Tellers in banks and cashiers in many retail settings also use terminals to perform their work duties.

There can be several types of terminals as discussed below :

A *dumb terminal* is an input/output (I/O) device that provides for data entry and information exit when connected to a computer but has no additional capability.

An *intelligent terminal* has an in-built processing capability. It is also user-programmable. It contains not only a storage area but also a microprocessor. The terminal can be programmed to communicate with and instruct the user who is entering data. It can also do some processing of the data internally such as sorting, summarizing, checking both input and computed values for reasonableness and so on, rather than relying on the mini-computer or the main-frame CPU. This feature can reduce the load on the central CPU. Thus, intelligent terminals can be used on a stand-alone basis or can be part of a distributed network of terminals.

Intelligent terminals cost several times more than non-intelligent terminals but the savings they provide for many companies is much more than their cost. Savings come about because the amount of data to be transmitted to the central CPU and the number of times it is interrupted are both reduced. Intelligent terminals also provide a type of back up to the main computer because the terminal can handle some of the processing.

Smart terminals additionally, contain a microprocessor and some internal storage. They have data editing capability and can consolidate input data before sending it to CPU. These terminals are non-programmable by users.

Remote job terminal (also referred to as Remote Job entry or RJE), groups data into blocks for transmission to a computer from a remote site. Some RJE terminals have the capability of receiving back and printing the results of the application program. Such a unit is in itself a small computer, which can be used either as job entry terminal or as a stand-alone computer.



Information Technology

A terminal may be situated at the computer site or situated at a remote place where the data to be input is more readily available. Terminals linked to the computer system by a direct cable are known as *hard-wired terminals*. However, for remote terminals, communication to the main system can be established via *telecommunication* lines such as ordinary telephone lines.

Keyboard printer terminal: The keyboard printer terminal or *teletypewriter* consists of a keyboard for sending information to the computer and a printer, for providing a copy of the input and for receiving information from the computer. The output is normally typed on a continuous roll of paper at speeds typically between 20 to 50 characters per second. A paper tape reader/punch is sometimes incorporated in the design of a terminal to enable information to be keyed in and punched on to paper tape for retention of data or for subsequent input to the computer. In place of the paper tape reader/punch, some more recently designed machines may have magnetic tape cassettes incorporated for the same purpose.

Hundreds of different display devices are now available. Although a number of important features distinguish one display device from another, the three that follow are among the most significant.

(a) Screen Resolution : One of the most important features used to differentiate display devices is the clarity, or resolution, of the images that are formed on –the screen. Most display devices form images from tiny dots - called pixels (a contraction of the two words “picture elements”) - that are arranged in a rectangular pattern. The more dots that are available to display any image on -screen, the sharper the image (the greater the resolution) is.

Images are formed on monitor's screen by a card called the display adaptor card. If a user wishes to change the kind of display, e.g, from black and white to colour, the display adaptor card must be changed. The key elements of display adaptor cards the video controller and the memory. A variety of display adaptor cards exist in the market, each with its own special features. Some of the popular display adaptors reported by personal computers are discussed below:

1. MGA- MGA or Monochrome Graphics Adapter is one of first adapters. It is a text -only adapter which generates very clear, easy-to-read characters. It works only with monochrome monitor.
2. CGA - CGA or Colour Graphics Adapter works in both text and graphics mode. It supports both colour and monochrome modes with various resolutions. However, it has relatively poor display quality in text mode. A CGA adapter provides following two combinations of resolutions
 - (i) 640 x 200 pixels with 16 colours.



(ii) 320 x 200 pixels with 4 palettes.

Each of these palettes has 4 different colours. Only one palette can be used at a given time.

3. EGA- An EGA or Enhanced Graphics Adapter combines all features of a MGA & CGA with higher resolutions. It supports up to 16 colours at a time. An EGA usually has a high resolution of either 640 x 200 pixels or 640 x 350 pixels.
4. VGA - VGA or Video Graphics Adapter is a high quality graphics adapter which provides upto 256 colours and also a high resolution. Following are the two typical combinations of resolutions and colours that a VGA provides.
 - (i) 640 x 480 pixels with 16 colours.
 - (ii) 320 x 200 pixels with 256 colours.
5. SVGA- SVGA or Super Video Graphics adapter is an improvement on the VGA. The two combinations of resolutions and colours provided by SVGA are
 - (i) 640 x 480 pixels with 256 colours.
 - (ii) 1024 x 480 pixels with 16 colours

(b) Text and Graphics : Many display devices made today (principal exceptions are inexpensive terminals such as those used in dedicated transaction processing applications) can produce both text and graphics output. Text output is composed entirely of alphabetic characters, digits, and special characters. Graphics output includes such images as drawings, charts, photographs, and maps.

Display devices that are capable of producing graphics output commonly employ a method called **bit mapping**. Bit-mapped devices allow each individual pixel on the screen to be controlled by the computer. Thus, any type of image that can be formed from the rectangular grid of dots on the screen (for example, a 640-by-480 grid) is possible. Character-addressable devices are not bit-mapped and partition the screen into standard character widths- for example, a series of 5-by-7 dot widths - to display text.

Perhaps the most important business-related use for graphics is presentation graphics. Presentation graphics enable managers to easily construct such information-intensive images as bar charts, pie charts, and line charts on their display devices and have these images sent to a printer, plotter, or slide-making machine so that they can be used later for presentations in meetings. Because these types of graphical images are relatively simple, a super-high-resolution workstation that can display photographs and sophisticated artwork is not needed.



Information Technology

Graphics display devices have been widely used for many years in the engineering and scientific disciplines. The display devices used for applications in these areas are extremely sophisticated and expensive.

(c) CRT Versus Flat-Panel : Most of the display devices used today are of the cathode ray tube (CRT) type. These devices use a large tube-type element that looks like the picture tube in a standard TV set. Inside the tube is a gun that lights up the phosphorescent pixels on the screen surface. Although CRT technology is relatively inexpensive and reliable, CRT-type display devices are rather bulky and limited in the resolution that they provide.

Currently challenging the CRT in the display device marketplace is the flat-panel display. The most common of these devices use either a liquid crystal display (LCD) or gas-plasma technology. To form images, LCD devices use crystalline materials sandwiched between two panes of glass. When heat or voltage is applied, the crystals line up. This prevents light from passing through certain areas and produces the display. Gas-plasma displays, which provide better resolution but are more expensive than liquid crystal displays, use gas trapped between glass to form images.

The biggest advantage of flat-panel displays is that they are lightweight and compact. This makes them especially useful for laptop, notebook, and pocket personal computers.

The Video Controller : As mentioned earlier, the quality of the images that a monitor can display is defined as much by the video controller as by the monitor itself. The video controller is an intermediary device between the CPU and the monitor. It contains the video-dedicated memory and other circuitry necessary to send information to the monitor for display on the screen. It consists of a circuit board, usually referred to simply as a card ("video card" and "video controller" mean the same thing), which is attached to the computer's motherboard. The processing power of the video controller determines, within the constraints of the monitor, the refresh rate, the resolution, and the number of colors that can be displayed.

During the 1980s, when most PCs were running DOS and not Windows, the screen displayed ASCII characters. Doing so took very little processing power, because there were only 256 possible characters and 2,000 text positions on the screen. Rendering each screen required only 4,000 bytes of data.

Windows, however, is a graphical interface, so the CPU must send information to the video controller about every pixel on the screen. At the minimum resolution of 640 × 480, there are 307,200 pixels to control. Most users run their monitors at 256 colors, so each pixel requires one byte of information. Thus, the computer must send 307,200 bytes to the monitor for each screen.



If the user wants more colors or a higher resolution, the amount of data can be much higher. For example, for the maximum amount of color (24 bits per pixel will render millions of colours) at 1024 × 768, the computer must send 2,359,296 bytes to the monitor for each screen.

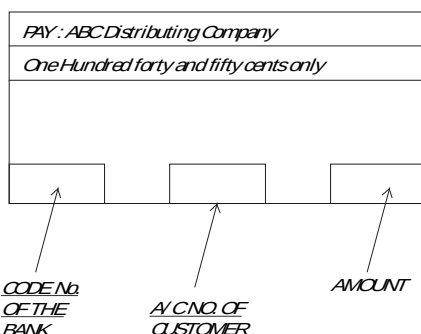
The result of these processing demands is that video controllers have increased dramatically in power and importance. There is a microprocessor on the video controller, and the speed of the chip limits the speed at which the monitor can be refreshed. Most video controllers today also include at least 2 MB of video RAM, or VRAM. (This is in addition to the RAM that is connected to the CPU.) VRAM is “dual-ported,” meaning that it can send a screenful of data to the monitor while at the same time receiving next screenful of data from the CPU. It's faster and more expensive than DRAM (Dynamic RAM). Users with larger monitors or with heavy graphics needs usually will want even more than 2 MB of VRAM.

1.2 DIRECT DATA ENTRY

Direct Data Entry (DDE) refers to entry of data directly into the computers through machine readable source documents. DDE does not require manual transcription of data from original paper documents. DDE devices can scan source documents magnetically or optically to capture data for direct entry into the computer. Magnetic ink character readers and optical character readers are examples of such devices..

We will now describe each of the above devices.

1.2.1 Magnetic ink character recognition (MICR) : MICR employs a system of printed characters which are easily decipherable by human beings as well as a machine reader. There is used special printing font to represent characters. In this font, each character is basically composed of vertical bars (See “2” of Figure 1.2.1.1) The characters are printed in special ink, which contains a magnetizable material. In “2” of Figure 1.2.1.1 there are four small gaps and two big gaps.



CHEQUE USING MICR



Fig. 1.2.1.1 Check using MICR

When a character is subsequently read it is passed beneath a reading head and big and small gaps send in different types of impulses represented by 1 bit and 0 bit respectively.

This method is primarily used in banking industry, and most cheques are now processed under the MICR approach. The data printed across the bottom of a blank cheque are recorded in MICR form ; the characters represent the bank on which the cheque is drawn, the customer's account number and the amount of the cheque. The cheques themselves are prepared off-line. When they are originally printed by a printing press, the MICR about the bank identification number, as well as the data about the customer's bank account number are printed simultaneously. The cheques then are turned over to the proper bank customer for use. Once the cheques have been cashed or deposited in a bank, an operator uses an off-line encoding machine to encode the amount on the cheque's bottom-right side.

MICR data are used for input purposes. Unlike other media (floppy disk and magnetic disk), MICR can not be used for output purposes.

Advantages of MICR :

- (i) MICR possesses a very high reading accuracy. Cheques may be smeared, stamped, roughly handled yet they are accurately read.
- (ii) Cheques can be handled directly without transcribing them on floppy disk, magnetic tape etc.
- (iii) Cheques can be read both by human beings and machines.

Disadvantages of MICR :

- (i) MICR has not found much favour from business.
- (ii) Damaged documents, cheques not encoded with amount etc. have still to be clerically processed.

1.2.2 Optical character reading (OCR) : OCR also employs a set of printing characters with standard font that can be read by both human and machine readers. The machine reading is done by light scanning techniques in which each character is illuminated by a light source and the reflected image of the character is analysed in terms of the light-dark pattern produced. Key-board devices are used to give the required print quality. OCR has the potential of reading even handwritten documents straightway.



Optical character readers can read upper and lower case letters, numerics and certain special characters from handwritten, typed and printed paper documents. The specific characters that can be read and whether the characters must be handwritten, typed or printed depends upon the type of OCR being used. Obviously OCR annihilates the time consuming step of transcription.

ABCDEF GHIJ KLMN
OPQR STUV WXYZ
1234 5678 90

Fig. 1.2.2.1 OCR

The optical character readers have slightly irregular type face as shown in the figure 1.2.2.1. They can read the characters printed by computer printers, cash registers, adding machines and typewriters. Some readers can also read hand-written documents. Figure 1.2.2.2 shows some handwritten characters that can be read by the recognition devices, most optical character readers can be used to sort out forms as well as to read data into computer storage.

<u>RULE</u>	<u>CORRECT</u>	<u>INCORRECT</u>
WRITE BIG	02831	02834
CLOSE LOOPS	06889	06889
USE SIMPLE		
SHAPES	02375	023375
CONNECT LINES	45T	45T
BLOCK PRINT	CSTXZ	CSTXZ

Fig. 1.2.2.2 Hand written characters

Large volume billing applications (e.g. the bills of utility companies, credit-card organisations, and magazine subscription outfits) increasingly are being adapted to OCR methods. The customer paying the bill returns the bill, which has OCR data (e.g. customer number and amount of the bill) recorded on it, along-with payment. Thus, the billing organisation's bill (or the part returned by the customer) becomes input for recording cash payments received from the customer. This procedure sometimes referred to as the use of "**turn-around documents**" has the advantage of minimizing or eliminating the keying process when cash receipts are received from customers.



Information Technology

Advantages of OCR

- (i) OCR eliminates the human effort of transcription.
- (ii) Paper work explosion can be handled because OCR is economical for a high rate of input.
- (iii) Since documents have only to be typed or handwritten, not very skilled staff (like the keypunch operators) is required.
- (iv) Furthermore, these input preparation devices (typewriters etc.) are much cheaper than the keypunch or the key-to-tape devices.

Limitations of OCR

- (i) Rigid input requirements - There are usually specific (and rather inflexible requirements for type font and size of characters to be used. In typing there is always the scope for strike-overs, uneven spacing, smudges and erasures; and the form design, ink specifications, paper quality, etc. become critical and have to be standardized.
- (ii) Most optical readers are not economically feasible unless the daily volume of transactions is relatively high. However, further developments in OCR are likely to make optical readers much cheaper.

OCR characters can be sent as input to the CPU. Also, printers can be used to generate OCR output. Optical character readers can read from both cut form and continuous sheets.

1.2.3. Optical mark recognition (OMR) : Optical marks are commonly used for scoring tests. The (Fig. 1.2.3.1) shows part (just the social security number in the U.S) of a typical test scoring sheet. It is marked by the person taking the test, and can be read by the optical mark page reader. The optical mark reader when on-line to the computer systems, can read up to 2,000 documents per hour. Seemingly this rate is slow but the fact that transcription has been eliminated, the overall time is less than those of the conventional file media.

OMR can also be used for such applications as order writing, payroll, inventory control, insurance, questionnaires, etc. However, it is to be noted that designing the documents for OMR is rather a tough task. They should be simple to understand otherwise errors may result more perhaps than would occur in using traditional source documents and keypunching from them.

SOCIAL SECURITY NO.

4 6 8 4 6 8 3 2 4

0 0 0 0 0 0 0 0

1.68



1 1 1 1 1 1 1 1 1
 2 2 2 2 2 2 2 2 2
 3 3 3 3 3 3 3 3 3
 4 4 4 4 4 4 4 4 4
 5 5 5 5 5 5 5 5 5
 6 6 6 6 6 6 6 6 6
 7 7 7 7 7 7 7 7 7
 8 8 8 8 8 8 8 8 8
 9 9 9 9 9 9 9 9 9

Fig. 1.2.3.1 Optical Mark Reorganization

Technical details of optical scanning: In all optical readers, the printed marks and/or characters must be scanned by some type of photo-electric device, which recognises characters by the absorption or reflectance of light on the document (characters to be read are non-reflective) Reflected light patterns are converted into electric impulses, which are transmitted to the recognition logic circuit — there they are compared with the characters the machine has been programmed to recognise, and, if valid, are then recorded for input to the CPU. If no suitable comparison is possible, the document may be rejected.

1.2.4 Smart Card Systems- Smart cards resemble credit cards in size and shape; however, they contain a microprocessor chip and memory, and some include a keypad as well. These were pioneered in France and many organizations are still just experimenting with them. In many instances, smart cards make it possible for organizations to provide new or enhanced services for customers.

So far, smart cards are used most frequently to make electronic purchases and to electronically transfer funds between accounts. However, the potential applications for these are abound. For example, in the health care industry, smart cards could be used to store the holder's identity, address, insurance data, relatives' details, allergies, and even a brief medical history. If the cardholder was disabled by an accident or illness, the card could be used immediately to assist with treatment. Smart cards could also be used for security applications. For example, a card could contain the digitized fingerprint of the cardholder, which could be compared at a security checkpoint to fingerprints of people who are authorized to enter a secured area.



Information Technology

1.2.5 Bar Code Reader : The most widely used input device after the keyboard and mouse is the flatbed or hand-held bar code reader commonly found in supermarkets and departmental stores. These devices convert the bar code, which is a pattern of printed bars on products, into a product number by emitting a beam of light – frequently a laser beam – which reflects off the bar code image. A light-sensitive detector identifies the bar code image by recognizing special bars at both ends of the image. Once the detector has identified the bar code, it converts the individual bar patterns into numeric digits. The special bars at each end of the image are different, so the reader can tell whether the bar code has been read right side up or upside down.

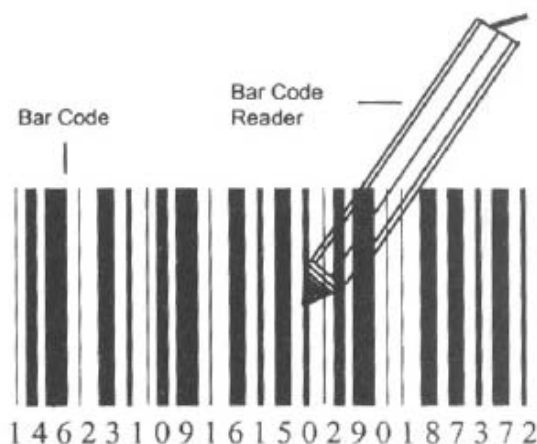


Fig. 1.2.5.1 Bar Code Reader

After the bar code reader has converted a bar code image into a number, it feeds that number to the computer, just as though the number had been typed on a keyboard.

Bar codes provide the advantages of improved accuracy of data entry, better customer service through faster checkout at the point of sale, and greater control and reliability of inventory records. They are used in industries and organizations that must count and track inventory, such as retail, medical, libraries, military and other government operations, transportation facilities, and the automobile industry.

Two-dimensional (2D) bar codes have been developed that store the equivalent of two text pages in the same amount of space as a traditional UPC. One of the first uses of 2D bar coding was handing barrels of hazardous toxic waste. Now it is commonly used in a variety of industries. For example, every shipping carton sent to one of Wal-Mart's distribution centers must have a 2D bar code. The bar code contains the purchase order, stock numbers, the



Introduction To Computers

contents of each box, a product's origin, its destination, and how it should be handled during shipping. These bar codes automate many of the mundane and time-consuming shipping tasks.

1.2.6 Image Processing : Image Processing captures an electronic image of data so that it can be stored and shared. Imaging systems can capture almost anything, including keystroked or handwritten documents (such as invoices or tax returns), flowcharts, drawings, and photographs. Many companies that use document imaging are making significant progress toward paperless offices. There are five distinct steps to document imaging:

- ☐ **Step1: Data capture.** The most common means of converting paper documents into electronic images is to scan them. The scanning device converts the text and pictures into digitized electronic code. This scanner can range from a simple hand held device to a high-end, high-speed scanner capable of scanning more than 2,500 pages an hour. *Hand held scanners* could transform text or graphical images into machine-readable data. Organisations such as publishers and law firms, that frequently receive documents, may use such scanners to convert the typed pages into word processing files. These can also be used for entering logos and other graphics for desk top publishing application. Fax modems are also used to receive electronic images of documents. Some of today's low speed printers and fax machines have removable print heads that can be replaced with a scanning head, enabling the printer to work as an image scanner.
- ☐ **Step2: Indexing.** Document images must be stored in a manner that facilitates their retrieval. Therefore, important document information, such as purchase order numbers or vendor numbers, is stored in an index. Great care is needed in designing the indexing scheme, as it affects the ease of subsequent retrieval of information.
- ☐ **Step 3: Storage.** Because images require a large amount of storage space, they are usually stored on an optical disk. One 5.25-inch optical platter can store 1.4 gigabytes, or about 25,000 documents (equivalent to 3 four-drawer filing cabinets) A 12-inch removable optical disk stores up to 60,000 documents, and up to 100 optical disks can be stored in devices called jukeboxes.
- ☐ **Step 4: Retrieval.** Keying in any information stored in an index can retrieve documents. The index tells the system which optical disk to search and the requested information can be quickly retrieved.
- ☐ **Step5: Output.** An exact replica of the original document is easily produced on the computer's monitor or on paper, or is transmitted electronically to another computer.

Advantages of Image Processing : It has been estimated that 90% of the work accountants and others do today is done using paper. It is also estimated that the volume of information



Information Technology

required by companies doubles every three or four years. As a result we are faced with being buried by paper. One solution is to make better use of document imaging. More companies are moving to this technology and it is estimated that by 2004 only 30% of our work will be paper-based; 70% will be electronic. The move to document imaging provides the following advantages:

- (i) *Accessibility* : Documents can be accessed and reviewed simultaneously by many people, even from remote locations.
- (ii) *Accuracy* : Accuracy is much higher because costly and error-prone manual data-entry processes are eliminated.
- (iii) *Availability* : There are no more lost or misfiled documents.
- (iv) *Capacity*: Vast amounts of data can be stored in very little space, which significantly reduces storage and office space.
- (v) *Cost* : When large volumes of data are stored and processed, the cost per document is quite inexpensive. As a result, the costs to input, file, retrieve, and refile documents are reduced significantly.
- (vi) *Customer satisfaction* : When waiting time is significantly reduced (due to lost or misfiled documents, queue time, etc.), customers can get the information almost immediately.
- (vii) *Security* : Various levels of passwords (network, data base, files, etc.) and clearances can be assigned to restrict document access.
- (viii) *Speed* : Data can be retrieved at fantastic speeds. Stored documents can be indexed using any number of identifying labels, attributes, or keywords.
- (ix) *Versatility* : Handwritten or types text can be added to an image, as can voice messages. Documents can be added to word processing files; the data can be included in a spreadsheet or data base.

1.3 TYPES OF COMPUTER OUTPUT

The various types of output from a computer system may be chosen from the following list according to specific requirements :

- Printed
- Visual display
- COM(Computer Output on Microfilm)



- Audio
- Magnetically encoded
- Graphical

1.3.1 PRINTED OUTPUT : PRINTERS

Printer is one of the most common output devices. It provides the user with a permanent visual record of the data output from the computer. Printers can print on ordinary paper or on specially prepared forms such as dispatch notes, invoices or packing slips. Printers have been developed that are capable of printing from 150 to 2,500 lines per minute, each line consisting of as many as 150 characters. Printers can broadly be subdivided into two categories: impact and non-impact printers. The former are the most common.

1.3.1.1 Impact printers : Impact printers can be described as printers which utilize some form of striking device to transfer ink from an inked ribbon onto the paper being printed to form images or characters. The characters printed are formed by one of two methods : (i) they are either distinct, whole alphanumeric images produced by a process known as full character or formed character printer or, (ii) they are formed by a dot matrix method which arranges a series of dots to assume the shape of each character being printed.

Impact printers fall into two basic categories-Serial or line printing.

(a) Serial Printers : Regard less of which character generation method is used, serial printers print one character at a time, usually from left to right. Some printers, however, can also print in a bidirectional format at an increased speed. In most business organisations two types of serial printers are used :

(i) Dot-matrix Printers : In the early 1970's a new print system called dot matrix printing was developed for use with data processing systems. These small, compact printers offered high speed, relatively low price and greater programmability for graphics and illustrations due to their method of character generation. They became the standard printers for many minicomputers and nearly all microcomputers and, consequently, were also used for word processing requirements. Dot matrix printers utilise wire needles or pins which strike the ribbon against the paper in the pattern necessary to produce each character image. The printing head is a matrix block which consists of rows and columns of holes through which pins appear. The characters being printed are formed by activating the printing head so that certain pins appear through the holes to form a pattern which resembles each character. The pins are formed into the shape of the character to be printed, then pressed against an inked ribbon and the pattern printed on



Information Technology

the paper. The character, whether they are letters, numbers or grammatical symbols are printed as a series of dots which merge together to form the character.

The matrices of these printers can vary in size from printer to printer depending on the print quality required. Some may have 11 rows of pins with 9 columns in each row; while others may have as many as 23 rows with 18 columns per row. The characters can, therefore, be formed from matrices using any combination from 99 to 414 dots printed by the pins. The greater the number of dots printed on the paper, the better the quality of the copy. An example of how a lower case letter 'd' is formed is shown in Figure 1.3.1.1.1.

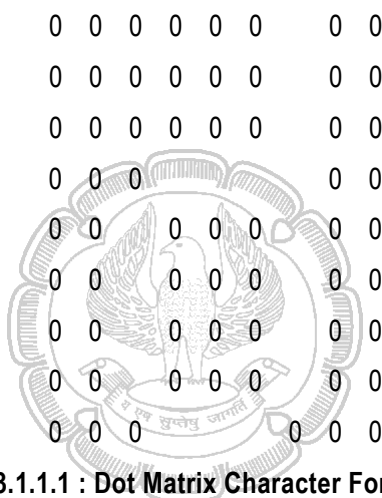


Figure 1.3.1.1.1 : Dot Matrix Character Formation.

Dot matrix printers are fast and cheap but their print quality is relatively poor. They are only really suitable for printing draft copies and documents, which are usually retained within an organisation. The printing also varies according to the type of matrix printer used. Most employ an ink ribbon, similar to conventional typewriter ribbons, to print documents. Although not of a high quality, when compared to letter quality printers, dot matrix printers do have the advantage of being able to print a much wider range and size of typeface and, in addition, can print graphs, illustrations and drawings. This type of printer can also be used to print colour.

Manufacturers of dot matrix printers include: Brother, Canon, Centronics, Commodore, Epson, Facit, General Electric, Hewlett Packard, Mannesheim Tally, NEC, OKI, Seikosha, Shinwa, Star, and TEC.



- (ii) **Daisywheel Printers** : 'Daisywheel' printers work in a similar manner to an electronic typewriter. The major difference is that they use a new type of printing element called a 'daisywheel'. This is a molded metal or plastic disc-shaped printing element which looks very much like a daisy - hence the name. (See Figure 1.3.1.2) It is about 65 m.m. in diameter and has a number of stalks or petals which radiate from a central core. On the end of each stalk is a type character set in a similar manner to the keys as on a typewriter. This type of printer works by rotating the print element until the required character is positioned in front of the sheet of paper at the point where it will be printed. A small hammer hits the back of the stalk forcing that character against an inked ribbon and through onto the paper. All this 'happens at anything from 10 to 80 characters per second, which is far faster than a typist can type. A similar device, shaped like and called a 'Thimble' printer is used by the Japanese company NEC on the 'Spinwriter' printer. These printers enable users to change the typeface elements very quickly giving far more scope to the typing of important documents.

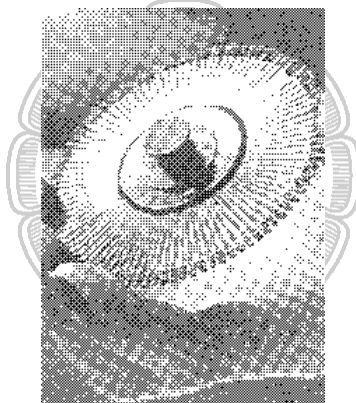


Figure 1.3.1.2 : The daisy wheel Printer

Quite recently cheaper daisywheel printers have appeared on the market. Major manufacturers of the daisywheel printers include: Brother, Data point, Diablo, Fujitsu, NEC, Olivetti, Olympia, TEC and Wang etc.

(b) Line Printers : A line printer operates at much higher speeds and prints what appears to be a full line at a time. Line printers are only used where high speed and volume is necessary and where quality is the lesser requirement. Two types of line printers are discussed belows :

- (i) **Chain Printers** : (Fig. 1.3.1.3) It has a chain that revolves at a constant speed in a horizontal plane. The complete chain has a complement of 48 numbers, alphabets and special symbols cast on 5 times over. It is confronted by a set of as many hammers as



Information Technology

the number of print position say, 160. These hammers are magnetically controlled. The continuous stationery and ribbon are inter posed between a segment of the chain and the set of hammers. When a required character on the chain faces its print position, the corresponding hammer is actuated.

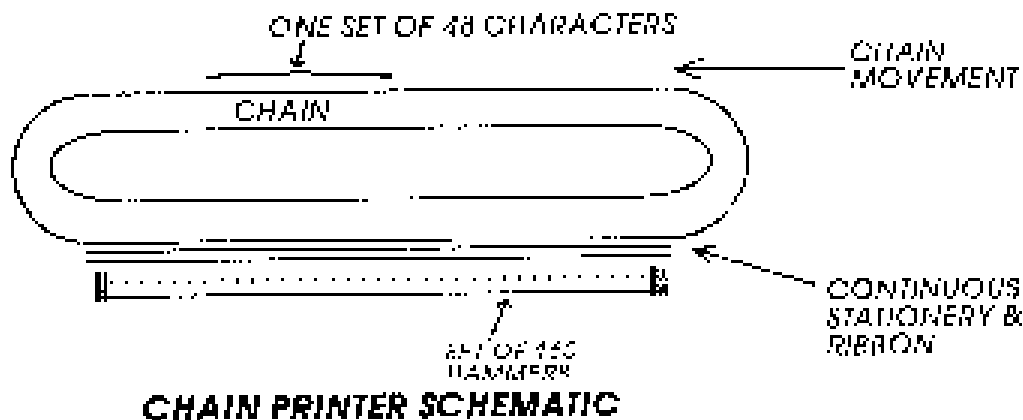
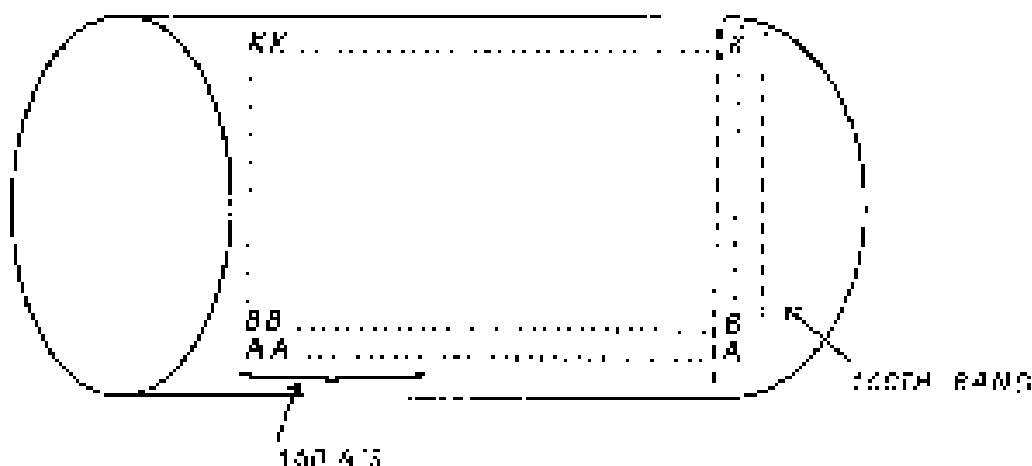


Fig. 1.3.1.3 : Chain Printer Schematic



DRUM PRINTER SCHEMATIC

Fig. 1.3.1.4 : Drum printer schematic



- (ii) **Drum Printers** : These printers use a solid cylinder. There are as many bands on it as the number of print positions. Each band has cast on it the complement of 48 numerals, alphabets and special characters. The drum rotates at a constant speed confronted by a set of as many hammers as the number of bands with the inked ribbon and continuous stationery interposed. In one rotation of the drum there would be appropriate strikes by the set of the hammers. In the first strike A's are printed in the appropriate print positions, followed by B,C,.....Z,0,1.....9 and special symbols one by one.

Various Characteristics of Printers

- (i) **Speed** : The speed of a printer is measured in terms of cps (characters per second), lps (lines per second) or PPM (pages per minute) The speed of a dot-matrix printer is measured in CPS. While the speed can vary widely, it is generally 200 CPS. A line printer prints a line at a time. Its speed can be any where from 5 to 50 lps.
- (ii) **Quality of output** : Depending on the type of characters formed, printers can be classified as draft, near letter quality (NLQ) or letter quality printers.

In a *draft quality printer*, a character is formed by arranging dots to resemble it. Although the characters can be distinguished, the output is not as good as that of near letter quality printouts. A dot-matrix printer is an example of a draft printer.

Near letter quality printers use a special character which resembles that of a typewriter. A daisy wheel printer is an example of a NLQ printer. Most dot-matrix printers can also be set to produce near - letter quality printouts.

Letter quality printers use a character set in which each letter or character is fully formed. The quality of output is the best in such printers. A *laser printer* discussed below is an example of a letter quality printer.

- (iii) **Direction** : Printers can be unidirectional or bi-directional. In a unidirectional printer, printing takes place in one direction only. After printing a line from left to right, the print head returns to the left without printing. A bi-directional printer prints both ways.

1.3.1.2 Non-impact printers : A non-impact printer forms characters by chemical or electronic means. The non-impact printers are, however, not commonly used for the following reasons:-

- (i) Special and more expensive paper is required.
- (ii) Only one copy can be printed at a time.
- (iii) The print is not as sharp or clear as with the impact printer.
- (iv) Output is difficult to copy on office machines.



Information Technology

However, three types of non-impact printers are worth mentioning because they are expected to become more important in the future, as the technology becomes cheaper. These are

- *Thermal printers*
- *Ink-jet printers*
- *Laser printers.*

They are fast in operation, printing a page, or even more in a second but currently they are too expensive to be widely used. The laser printer produces very high quality prints from a wide selection of character fonts. Not many business organisations can justify the present high cost of laser printing but costs are falling sharply and laser printing is likely to become more common place. We have discussed below each of these briefly :

- i) **Thermal printers** : These printers use thermal printing facility i.e. the printer forms each character on the matrix printing head in the normal way but, instead of the printing head being impacted physically against an inked ribbon to print the character, the pins are heated by the electric element and then pressed against the paper. When the pins touch the paper, the area heated by the pins changes colour, usually to black or brown to form the character.



Fig.1.3.1.2.1 Thermal Printers

- (ii) **Ink-Jet Printers** : This type of printers utilize the dot matrix principle but in stead of pins, set mechanically to form a character, it uses an array of nozzles which spray jets of ink onto the paper. When they first appeared; ink jet printers were very expensive but now a number of ink jet printers are



Fig. 1.3.1.2.2 Inkjet Printer



beginning to find their way on to the market at fairly reasonable prices. An excellent example of such a printer is the Hewlett Packard "Thinkjet" originally developed for the Hewlett Packard HP 150 "Touch-Screen" microcomputer. HP 670C and HP 810C have now been made available for use with the IBM PC and other microcomputers.

Fig. 1.3.1.2.2 : Inkjet Printer

Other examples include Canon's PJ 1080 A, the Diablo C-150 and the Olivetti Spark Ink Jet Printer, which can operate at 50 lines per minute.

There are two types of inkjet printers:

Inkjet printers are very quiet and provide laser-like quality at a much lower cost although supplies are relatively expensive. Although one can print on regular paper, better results are obtained by using special paper that doesn't allow the ink to soak in.

Liquid inkjet : Color inkjet printers use three separate inkjets, one for each of the primary colors (red, green, and blue)

Liquid inkjet printers use ink cartridges from which they spray ink onto the paper. A cartridge of ink attached to a print head with 50 or so nozzles, each thinner than a human hair, moves across the paper. The number of nozzles determines the printer's resolution. A digital signal from the computer tells each nozzle when to propel a drop of ink onto the paper. On some printers, this is done with mechanical vibrations.

Solid inkjet : Solid inkjet printers use solid ink sticks that are melted into a reservoir and sprayed in precisely controlled drops onto the page where the ink immediately hardens. High-pressure rollers flatten and fuse the ink to the paper to prevent smearing. This produces an exceptionally high-quality image with sharp edges and good colour reproduction. Solid inkjet printers are also the best for producing low-cost but high-quality transparencies. These printers are sometimes referred to as phase-change printers because the ink moves from a solid to a liquid phase to be printed then back to a solid phase on the page. As a final step, the paper moves between two rollers to cold-fuse the image and improves the surface texture.

- (iii) **Laser Printers** : Laser printer uses a combined system which utilizes laser and Xerographic photocopying technology. In a laser printer, a beam from a small laser scans horizontally across a charged xerographic selenium drum to build up an invisible electrostatic image of a whole page of text. Using standard photocopier techniques, the





Information Technology

formed image is then transferred to the paper being printed. Toner is attracted to the electrostatic image and it is then permanently fixed using heat. A wide variety of typefaces are available. Very competitively priced laser printers have been introduced by various companies such as Hewlett Packard. These printers, which can be used by most personal computers and the microcomputers, have an excellent print quality. Laser printer produces not only alphanumeric characters but also drawings, graphics and other requirements. Its output can be very close to professional printing quality. The dots making up the image are so closely spaced that they look like characters typed on a typewriter. The laser printer prints a page at a time. Its speed can be anywhere between 4 to 17 pages per minute (ppm) The resolution of laser printers is measured in dots per inch (DPI) The most common laser printers have resolutions of 600 DPI, both horizontally and vertically. Some laser printers have a resolution of 1200 DPI also.

Laser printer features : Even with the best printers, there are differences that affect their use. These include the number of sides the printer prints on, its memory, and its fonts.

Duplex printers print on both sides of a sheet of paper at the same time. This duplex printing is ideal for documents that will be stapled or bound.

Laser printers make up an entire page before printing it. The page is temporarily stored in the printer's memory while it is being processed and printed. If a page contains a graphics image, the memory required for it can be substantial. For example, it takes 1 megabyte to store a full-page black-and-white graphics image that is to be printed with a 300-dpi resolution. To fill the page, the printer has to address over 8 million dots (1 megabyte x 8 bits = 8 million bits) This is obviously a big chore.

All laser printers come with some built-in fonts; however thousands of others can be added. At one time, these were added using plug-in cartridges. Currently they are built-in or distributed on disks and installed onto the system's hard disk drive. From there they are fed to the printer when needed.

1.3.2 COMPUTER OUTPUT MICROFILM (MICROFICHE)

Computer output microfilm (COM) is an output technique that records output from a computer as microscopic images on roll or sheet film. The images stored on COM are the same as the images, which would be printed on paper. The COM recording process reduces characters 24, 42 or 48 times smaller than would be produced from a printer. The information is then recorded on sheet film called 16 mm, 35 mm microfilm or 105 mm microfiche.

The data to be recorded on the microfilm can come directly from the computer (online) or from



Introduction To Computers

magnetic tape, which is produced by the computer (off-line) The data is read into a recorder where, in most systems, it is displayed internally on a CRT. As the data is displayed on the CRT, a camera takes a picture of it and places it on the film. The film is then processed, either in the recorder unit or separately. After it is processed, it can be retrieved and viewed by the user.

COM has several advantages over printed reports or other storage media for certain applications. Some of these advantages are:

- (1) Data can be recorded on the film at up to 30,000 lines per minute-faster than all except very high-speed printers.
- (2) Costs for recording the data are less.
- (3) Less space is required to store microfilm than printed materials. A microfilm that weighs an ounce can store the equivalent of 10 pounds of paper.
- (4) Microfilm provides a less expensive way to store data than other media provide.

To access data stored on microfilm, a variety of readers are available which utilize indexing techniques to provide a ready reference to data. Some microfilm readers can perform automatic data lookup, called computer-assisted retrieval, under the control of an attached minicomputer. With powerful indexing software and hardware now available, a user can usually locate any piece of data within a 200,000,000 characters data base in less than 10 seconds at a far lower cost per inquiry than using an on-line inquiry system consisting of a CRT, hard disk, and computer.

Though both microfilm and microfiche are created on a continuous negative film and one can make as many copies of the film as one desires, there are certain basic differences between the two. The Physical difference between a microfilm and microfiche is that a microfilm stays in a continuous form while a microfiche is cut into pieces. A microfilm is 16 millimeters (mm) or 35 mm roll of film contained in cartridges. Each roll can hold 2,000 to 5,000 pages of information. A microfiche, on the other hand, is 105-mm film cut in 4 × 6 inch sheet, each sheet capable of reproducing upto 270 page sized images. The operational difference between the two is that a microfilm has to be read sequentially until the desired record is retrieved whereas a microfiche allows direct access to data through hunt and storage procedure.

For certain applications, COM is a viable way to store and retrieves data.

1.3.3 MICROPHONES AND VOICE RECOGNITION

Now that sound capabilities are a standard part of computers, microphones are becoming increasingly important as input devices. Sound is used most often in multimedia, where the



Information Technology

presentation can benefit from narration, music, or sound effects. In software, sounds are used to alert the user to a problem or to prompt the user for input.

For this type of sound input, basically a digitized recording is required. All that one needs to make such a recording are a microphone (or some other audio input device, such as a CD player) and a sound card that translates the electrical signal from the microphone into a digitized form that the computer can store and process. Sound cards can also translate digitized sounds back into analog signals that can then be sent to the speakers.

There is also a demand for translating spoken words into text, much as there is a demand for translating handwriting into text. Translating voice to text is a capability known as **voice recognition** (or speech recognition). With it, one can speak to the computer rather than having to type. The user can also control the computer with oral commands, such as “shut down” or “print status report”.

Voice recognition software takes the smallest individual sounds in a language, called phonemes, and translates them into text or commands. Even though English uses only about 40 phonemes, a sound can have several different meanings (“two” versus “too,” for example) making reliable translation difficult. The challenge for voice distinguish meaningful sounds from background noise.

Sound Systems : Just as microphones are now important input devices, speakers and their associated technology are key output systems. Today, when one buys a multimedia PC, one gets a machine that includes a CD-ROM drive, a high-quality video controller (with plenty of VRAM), speakers, and a sound card.

The speakers attached to these systems are similar to ones that are connected to a stereo. The only difference is that they are usually smaller and they contain their own small amplifiers. Otherwise, they do the same thing that any speaker does: they transfer a constantly changing electric current to a magnet, which pushes the speaker cone back and forth. The moving speaker cone creates pressure vibrations - in other words, sound.

The more complicated part of the sound output system is in the **sound card**. The sound card translates digital sounds into the electric current that is sent to the speakers. Sound is defined as air pressure varying over time. To digitize sound, the waves are converted to an electric current measured thousands of times per second and recorded as a number. When the sound is played back, the sound card reverses this process, translating the series of numbers into electric current that is sent to the speakers. The magnet moves back and fourth with the changing current, creating vibrations.

With the right software, one can do much more than simply record and play back digitized



sound. Utilities that are built into Windows 95 provide a miniature sound studio, allowing the user to view the sound wave and edit it. In the editing one can cut bits of sound, copy them, amplify the parts one wants to hear louder, cut out static, and create many exotic audio effects.

1.3.4 GRAPH PLOTTER

A graph plotter is a device capable of tracing out graphs, designs and maps into paper and even into plastic or metal plates. A high degree of accuracy can be achieved, even up to one thousandth of an inch. Plotters may be driven *on-line* or *off-line*. Computer systems dedicated to design work often operate plotter on-line but systems used for other applications as well as graphic applications operate them off-line.

There are two types of plotters:

Drum : A drum plotter plots on paper affixed to a drum. The drum revolves back and forth, and a pen suspended from a bar above moves from side-to-side taking up new plot positions or plotting as it moves. This device is suitable for routine graph plotting and also for fashion designs.

Flat-bed: On a flat-bed plotter, the paper lies flat. The bar on which the pen is suspended itself moves on a gantry to provide the necessary two-way movement. Colour plotting is usually possible.

Plotters are now increasingly being used in application like CAD, which requires high quality graphics on paper. A plotter can be connected to a PC through the parallel port. A plotter is more software dependent than any other peripheral, and needs much more instructions than the printer for producing output.

Many of the plotters now available in the market are desktop models that can be used with PCs. Business generally use plotters to present an analysis in visual terms (bar charts, graphs, diagrams) etc. as well as for engineering drawings.



Information Technology

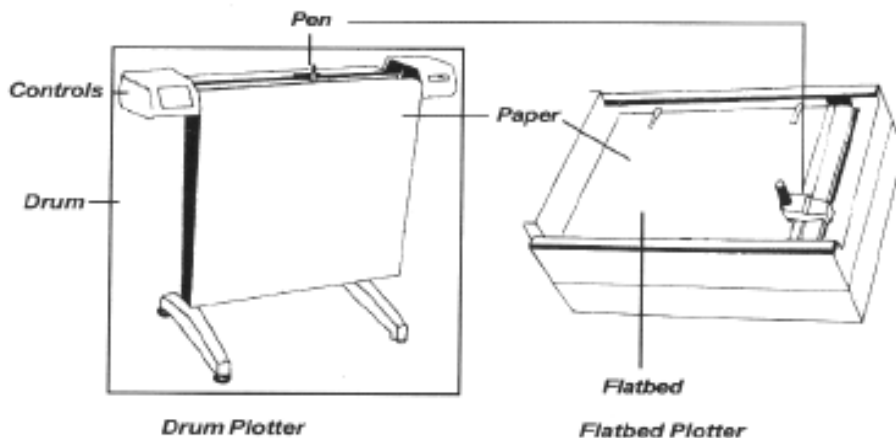


Fig. 1.3.4.1 : Graph Plotter

Self-examination questions

1. Write short notes relating to computer input and data capture.
2. Write brief notes on the following techniques and describe the situation in which each technique would be used :—
 - (i) Optical Character recognition (OCR)
 - (ii) Magnetic Ink Character Recognition (MICR)
 - (iii) Optical Mark Reader (OMR)
 - (iv) Image Scanners.
 - (v) Bar Codes
3. List the types of terminals which may be used for computer input.
4. What purpose do terminals serve ?
5. VDU (Visual display unit) and key-board-printer units are widely used as terminals to multiaccess systems. Give one example for each device where it is preferable to use it. Give reasons for your choice.
6. Explain the differences between a dumb terminal and an intelligent terminal.
7. Summarise the different methods of producing computer output.



Introduction To Computers

8. Distinguish between impact printers and non-Impact printers.
9. List the main features of a dot matrix and a daisy wheel printer.
10. Write short notes on laser printer and an ink-jet printer.
12. Write short notes on chain printer and a drum printer.
13. Explain the following terms:-
 - (i) Computer Output Microfilm.
 - (ii) Graph Plotter.
 - (iii) Voice recognition.

Multiple Choice Questions:

1. Mouse are of following types:
 - a) Mechanical b) Optomechanical c) Optical d) All above
2. Dot - Pitch is :
 - a) Vertical distance between pixel on a display screen
 - b) Distance between characters on printer
 - c) Horizontal distance between character in scanner
 - d) Rate at which screen is drawn
3. If the monitor has more resolution then:
 - a) Colour will not be good b) Picture quality is good c) Both A & B
 - d) None of the above
4. For graphics printing we use
 - a) Dot Matrix Printer b) Scanner c) Plotter d) None of the above
5. Which is impact printer :
 - a) Laser Printer b) Thermal Printer c) Ink Jet printer d) Dot Matrix Printer

Answer:

1. d 2. a 3. b 4. c 5. d



Unit 3 : Software

WHAT IS SOFTWARE ?

Software is a term used for the various kinds of programs, used to operate computers and related devices. It is different from the term hardware which describes the physical aspects of computers and related devices.

Software can be thought of as the variable part of a computer and **hardware** the invariable part. Software is often divided into **application software** (programs that do work users are directly interested in) and **System Software** (which includes operating systems and any program that supports application software). The term middleware is sometimes used to describe programming that mediates between application and system software or between two different kinds of application software (for example, sending a remote work request from an application in a computer that has one kind of operating system to an application in a computer with a different operating system).

There are basically three types of software : Systems software, applications software and general-purpose software. We will now discuss each of these in detail.

1.1 SYSTEMS SOFTWARE

It comprises of those programs that control and support the computer system and its data processing applications. It includes the following:

- Programming languages.
- Operating systems.
- Subroutines.
- Utility programs.
- Diagnostic routines.
- Language translators.

A broad spectrum of the above software is usually available from the manufacturer, but some software is also available from software concerns. There is generally an extra charge for software but some manufacturers furnish software without extra cost for those who purchase or lease their equipment. We will now describe each of the aforesaid software in detail.

1.1.1 PROGRAMMING LANGUAGES

The programming languages are part of the software or programming aids provided by the



manufacturer. A programming language is a language used in writing programs to direct processing steps to be carried out by a computer.

A wide variety of computer programming languages are available. The programmer must be familiar with one or more of them in order to write a computer program. Each programming language has very specific standards and rules of use, which the programmer must know and adhere to when developing a program. The programming languages can be hierarchically divided as follows.

1.2 OPERATING OR (EXECUTIVE) SYSTEMS

Operating system is system software, act as an interface between hardware and user. It has two goals:

- ◆ Resource allocator
- ◆ Control Program

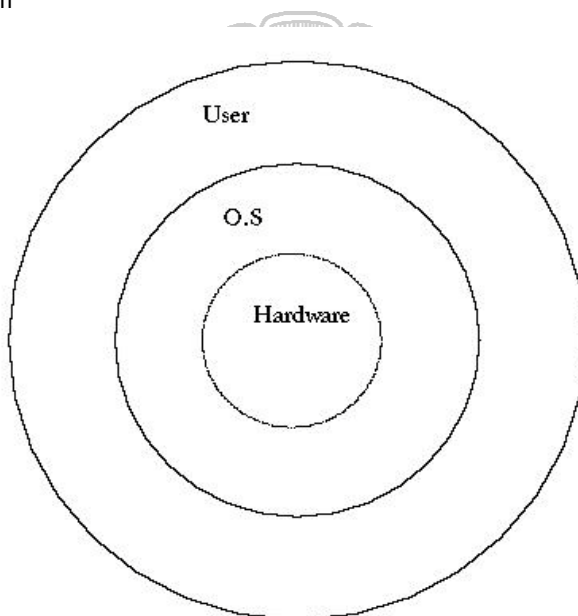


Fig. 1.2.1 Operating System as an Interface

The resources like Processor, Hard Disk, Memory etc are allocated by the Operating system.

As control program it provides the environment for the execution of the program and prevents improper use of computer.



Information Technology

There are six basic functions that an operating system can perform

- (i) *Schedule Jobs*: They can determine the sequence in which jobs are executed, using priorities established.
- (ii) *Manage Hardware and Software Resources*: They can first cause the user's application program to be executed by loading it into primary storage and then cause the various hardware units to perform as specified by the application.
- (iii) *Maintain System Security*: They may require users to enter a password - a group of characters that identifies users as being authorised to have access to the system.
- (iv) *Enable Multiple User Resource Sharing*: They can handle the scheduling and execution of the application programs for many users at the same time, a feature called multiprogramming.
- (v) *Handle Interrupts* : An interrupt is a technique used by the operating system to temporarily suspend the processing of one program in order to allow another program to be executed. Interrupts are issued when a program requests an operation that does not require the CPU, such as input or output, or when the program exceeds some predetermined time limit.
- (vi) *Maintain Usage Records* : They can keep track of the amount of time used by each user for each system unit - the CPU, secondary storage, and input and output devices. Such information is usually maintained for the purpose of charging users' departments for their use of the organisation's computing resources.

1.2.1 Files and Directories : OSs function at two different levels, the hardware level and the user level. The interaction of the OS with the computer's hardware level is, for the most part, hidden from the user. For example, signals relayed from the keyboard to the computer's microprocessor are done automatically, under the control of the OS, and without further recourse to instructions from the user.

Even so, the user will, at some stage or another, wish to install programs/ applications, copy files, perform search routines or execute any one of a wide and ever-increasing range of tasks. In order to carry out these procedures, the user has to access and operate an appropriate user interface. However, before doing so, the user should have an appreciation of the significance of files and directories at the operating system level.

File names : A file is a collection of related data, saved on a specific area of the surface of whatever storage medium the system is using. To be accessed, the file must have a name which is recognised by the system. The name is also significant in that it is usually an



Introduction To Computers

indication of purpose of the file and the type of data stored in the file. (It should be pointed out that OSs such as Windows 95 and Windows NT place greater emphasis on small icons as a means of identifying the nature and type of a particular file.) For most OSs the file usually has two names. These are known as the prefix and the suffix/ extensions. The prefix is supplied by the user. The suffix is supplied by either the user or the computer application currently being used.

The prefix is frequently separated from the suffix by a full stop or a slash. For example, in a DOS or Windows environment, the file LETTER.TXT would suggest that it is a letter file containing text data. Below is a listing of some typical DOS/ Window file names with appropriate suffixes indicating their contents:

File Name	Data Type
PAY.BAS	Basic program (.BAS - Basic)
COMP.WKS	Spreadsheet data file (.WKS -Worksheet)
ACC.CBL	Cobol program(.CBL - Cobol)
FRED.DAT	Simple data file (.DAT-Data)
SAVE.BAT	Batch file (.BAT -Batch file)
NOTE.TXT	Text file (.TXT - Text)
LETTER.DOC	Document file (.DOC - Document)
MENU.EXE	Executable file (.EXE - Executable)
PROG.COM	Command file (.COM - Command)
REG.DBS	Database file (.DBS - Database)
PIC.IMG	Image file (.IMG - IMAGE)

Directories : A directory is like a drawer in a filing cabinet. It contains files. In addition, just as a filing cabinet usually consists of multiple drawers, similarly a computer disk usually contains multiple directories. A directory can, if necessary, also contain other directories. When the directory contains another directory, it is known as sub directory. The purpose of directories is to enable the user to partition the computer's disk/storage devices in an organised and meaningful manner.

1.2.2 Graphical User Interfaces - Through the 1980s, micro computer's operating system was strictly text-based, and syntax sensitive command *i.e.* the users were required to issue commands to the operating system by entering them on the key board, one character at a



Information Technology

time. For example, to copy a word processing document from one disk to another, he might enter “copy c:\myfile a:” via the key board.

The trend now-a-days is away from command-driven interfaces to an user-friendly graphics oriented environment called Graphical User Interfaces or GUI.

GUIs provide an alternative to cryptic text commands. With the GUI, the user can interact with operating system and other software packages by selecting options from menus that are temporarily super-imposed over whatever is currently on the screen by using mouse to position the graphics cursors over the appropriate icons. Graphical user interfaces have effectively eliminated the need for users to memorise and enter cumbersome commands.

When 386 and 486 — based computers were launched, these systems supported complex graphic displays, hence graphical user interface became necessary for almost all applications. Responding in a timely manner, Microsoft produced MS-Windows. This was actually not an operating system but an operating environment which internally exploited MS-DOS capabilities to execute various system calls. Officially introduced in November 1983, Windows provided an effective GUI cover for MS-DOS. In November 1985, Microsoft released Windows 2.0, but it met with a lukewarm response. Windows 3.0, which became widely popular, was launched in May 1990.

1.2.3 Various Operating Systems

MS/ PC-DOS : The origins of Microsoft's Disk Operating System (MS-DOS) lie in the prelaunch era of the IBM PC. As IBM was about to finalise the PC, it started to negotiate with other computer companies to secure a suitable OS for the product. Initial contracts between IBM and Digital Research, the company which happened to own the rights to CP/M (Control Program for a Micro computer) which was, at that time, the market leader in microcomputer OSs, proved unsuccessful.

Microsoft Windows:

Microsoft windows is one of the most popular operating system, its success is often attributed because of Graphical user interface(GUI) feature. In this system we click on icons with mouse button rather than writing the command as in MS-DOS. Microsoft windows have the following features:

- ◆ Simple and user friendly.
- ◆ No need to write the command.
- ◆ Screen quite animated and attractive.



Introduction To Computers

- ◆ It has the multitasking feature because of that more than one program can be executed concurrently.

Flavors of windows Operating System:

Windows 95 : Windows 95, a 32 bit OS was released in August 1995. It took Microsoft three and a half years to develop. It was a gigantic task as far as computer projects go and was estimated to have taken 75 million hours of testing prior to its release. It was greeted enthusiastically by the computer industry, which saw it as a significant launch platform which would enable it to sell even more sophisticated computers.

The significance of a 32-bit OS as opposed to a 16-bit OS can be measured by the amount of internal main memory that can be directly access by the user/program. For example, with a 16-bit version of MS-DOS, the maximum amount of directly accessible memory is 1 MB. However, with a 32 bit OS, the user has direct access to 4 GB of main memory. To run Windows 95 users need a computer equipped with a 386DX or higher processor with a minimum of 4 Mb of memory (8Mb is recommended) along with a hard disk of 50 Mb as well as 3.5 inch disk drive or a CD-ROM.

Windows 95 was designed to have certain critical features over and above what was already supplied by Windows 3.1 or Windows for Workgroups. These included:

- (a) A 32-bit architecture which provides for a multitasking environment allowing the user to run multiple programs or execute multiple tasks concurrently. This architecture also enables faster data / file access as well as an improvement in printing delivery.
- (b) A friendlier interface fitted with what is described as 'one click' access. One click access refers to the fact that users didn't have to double click on the mouse every time that they wanted to activate an application. Other congenial attributes include the ability to employ long file names, easy navigation routes and 'plug and play technology' enabling users to connect various peripheral devices or add-ons with the minimum of fuss.
- (c) Windows 95 is also network ready. In other words the OS is designed for easy access to network resources. The OS also facilitates gateways to e-mail and fax facilities and access to the Internet via the Microsoft Network. In addition Windows 95 is backwardly compatible with most 3.1 Windows / DOS applications so enabling users to migrate from previous systems / applications.

Windows NT: Unlike Windows 3 and Windows 95, Windows New Technology (NT) is what is known as an industry standard mission critical OS. As a 32 bit OS Windows NT represents the preferred platform for Intel's more powerful Pentium range of processors. Although not exactly



Information Technology

the same, Windows NT 4.0 is, as might be expected, very similar in appearance to Windows 95. Critical features that allow the program to context the commercial OS market include:

- ☐ A stable multitasking environment
- ☐ Enhanced security features
- ☐ Increased memory
- ☐ Network utilities
- ☐ Portability: NT can operate on microprocessors other than those designed for the PC.

Windows NT is, as might be expected, more expensive than the other Windows OS and makes greater processing demands. However, it should be pointed out that Windows NT is making massive inroads into the corporate computing market and is fully recognised as being a competent useful OS.

OS/2: In 1987 IBM and Microsoft announced a new PC OS called OS/2 (Operating System Two) Unfortunately, the original OS/2 was not very successful. Hindsight suggests that, as with the early versions of Windows, one of the reasons for the slow uptake of OS/2 was the then considerable hardware demand of this particular application.

Another more serious problem with the original OS/2 that it was unable to support many existing PC applications. So users faced problems due to lack of compatibility between their original applications and OS/2. Predictably, the initial lack of interest in the original OS/2 resulted in a considerable strain on the IBM - Microsoft alliance. Not long after the launch of OS/2, IBM and Microsoft began to go their separate ways. Microsoft effectively abandoned OS/2 to IBM and chose instead to concentrate on MS-DOS and Windows.

1.3 OPERATING SYSTEMS FOR LARGER SYSTEMS

- Operating systems for mid-range and mainframe systems are often more complex than those for microcomputers.
- MVS is the most common operating system used on IBM mainframe. OS/400, an operating system for the IBM AS/400 line of midrange computers, is used at most of the sites where AS/400 is installed.
- VMS is the operating system used most frequently on DEC midrange and mainframe systems.

Interleaving Techniques : Large centralized systems often support multiple simultaneous users. The users' terminals may have limited processing capabilities and actual processing



may be done entirely on the large computer that is connected to the terminals. Hence, this computing configuration requires an operating system that enables many users to concurrently share the central processor. To do this, the operating systems on large computer systems often combine (interleave) the processing work of multiple simultaneous users or applications in a manner that achieves the highest possible resource efficiency. Among the interleaving techniques commonly used are multiprogramming, foreground/background processing, multi-tasking, virtual memory, and multi-processing.

Feature of Operating System:

1.3.1 Multiprogramming - The purpose of multiprogramming is to increase the utilization of the computer system as a whole. You might have noted when a program issues an input/output command, the program and hence the CPU is placed in a wait state until the execution of the command has been completed. When the transfer of data between main memory and the input/output devices has been completed, the device generates an *interrupt*, which is a signal that the data has been transferred. Till then the CPU remains idle and only after it receives the interrupt signal, it continues processing. Hence, in a way, the speed of the CPU is restricted by the speed of I/O devices and most of the time the CPU keeps waiting for the I/O operations to be completed. In order to utilize the computer more effectively, a technique known as *multiprogramming* has been developed. It is a module that is available in an operating system.

Multiprogramming is defined as execution of two or more programs that all reside in primary storage. Since the CPU can execute only one instruction at a time, it cannot simultaneously execute instructions from two or more programs. However, it can execute instructions from one program then from second program then from first again, and so on. This type of processing is referred to as "concurrent execution". Using the concept of concurrent execution, multiprogramming operates in the following way:

When processing is interrupted on one program, perhaps to attend an input or output transfer, the processor switches to another program. This enables all parts of the system, the processor, input and output peripherals to be operated concurrently thereby utilizing the whole system more fully. When operating on one program at a time, the processor or peripherals would be idle for a large proportion of the total processing time even though this would be reduced to some extent by buffering. **Buffering** enables the processor to execute another instruction while input or output is taking place rather than being idle while transfer was completed. Even then, when one program is being executed at a time, basic input and output peripherals such as floppy disk drive and line printers are slow compared with the electronic speed of the processor and this causes an imbalance in the system as a whole. However, in a multi-programming environment, the CPU can execute one program's instructions while a



Information Technology

second program is waiting for I/O operations to take place.

In a system of multiprogramming, storage is allocated for each program. The areas of primary storage allocated for individual programs are called “partitions”. Each partition must have some form of storage protection and priority protection to ensure that a program in one portion will not accidentally write over and destroy the instructions of another partition and priority (when two or more programs are residing in primary storage) because both programs will need access to the CPU’s facilities (e.g., the arithmetic and logic section) A system of priority-a method that will determine which program will have first call on the computer’s facilities-is normally determined by locating the programs in specific partitions.

1.3.2 Foreground/background processing: Usually, it is possible to partition main memory into logically separate areas. This enables, for instance, two different operating systems to work on the same machine because each will have its own memory to manage in its own way. Partitioning also allows separate “job streams” to be set up. A common procedure is to set up a partition for high-priority tasks (called a foreground partition) and one for low-priority tasks (called a background partition) With foreground/background processing, foreground jobs are usually handled first. When no foreground task awaits processing, the computer goes to the background partition and starts processing tasks there. As other foreground tasks come into the job queue, the computer leaves the background partition and resumes working in the foreground

1.3.3 Multi-tasking : Multi-tasking refers to the operating system’s ability to execute two or more of a single user’s tasks concurrently. Multitasking operating systems are often contrasted with single-user operating systems. *Single-user operating systems* have traditionally been the most common type of operating system for microcomputers. These only allow the user to work on one task at a time. For example, with many single-user operating systems for microcomputer systems, a word-processing user cannot effectively type in a document while another document is being printed out on an attached printer. For microcomputers, multi-tasking operating systems provide single users with multiprogramming capabilities. This is often accomplished through foreground/background processing. Multitasking operating systems for microcomputers-Such as Windows, OS/2, UNIX, Xenix, and Macintosh System 7 – only run on the more powerful microprocessors that were developed; older machines with less powerful microprocessors typically have single-user operating systems.

1.3.4 Virtual Memory - A programmer has to take into account the size of the memory to fit all his instructions and the data to be operated in the primary storage. If the program is large, then the programmer has to use the concept of virtual memory. Virtual memory systems, sometimes called virtual storage systems, extend primary memory by treating disk storage as



Introduction To Computers

a logical extension of RAM. The technique works by dividing a program on disk into fixed-length pages or into logical, variable-length segments.

Virtual memory is typically implemented as follows. Programs stored on disk are broken up into fixed-length pages. When a program needs to be processed, the first few pages of it are brought into primary memory. Then, the computer system starts processing the program. If the computer needs a page it does not have, it brings that page in from secondary storage and overwrites it onto the memory locations occupied by a page it no longer needs. Processing continues in this manner until the program finishes. This is known as *overlaying*.

By allowing programs to be broken up into smaller parts, and by allowing only certain parts to be in main memory at any one time, virtual memory enables computers to get by with less main memory than usual. Of course, during page swapping in multiprogramming environments, the system may switch to other programs and tasks.

Thus, virtual memory is primary storage-that does not actually exist. It gives the programmers the illusion of a primary storage that is for all practical purposes never ending. It uses the hardware and software features, which provide for automatic segmentation of the program and for moving the segments from secondary storage to primary storage when needed. The segments of the program are thus spread through the primary and secondary (on-line) storage, and track of these segments is kept by using tables and indices. So far as the programmer is concerned, the virtual memory feature allows him to consider unlimited memory size, though not in physical term.

Virtual memory is tricky because writing software that determines which pages or segments are to be swapped in and out of real storage is a system programming art. A disadvantage of virtual memory systems is that time is lost when page swapping occurs. In some cases, the same pages will be brought in and out of primary memory and unusually large number of times, which is an undesirable condition known as thrashing. Machines that do not offer a virtual memory feature usually compensate for this by having larger primary memories.

1.3.5 Multiprocessing - The term multiprogramming is sometimes loosely interchanged with the term multiprocessing, but they are not the same. Multiprogramming involves concurrent execution of instructions from two or more programs sharing the CPU and controlled by one supervisor. Multiprocessing (or parallel processing) refers to the use of two or more central processing units, linked together, to perform coordinated work simultaneously.

Instructions are executed simultaneously because the available CPUs can execute different instructions of the same program or of different programs at any given time.

Multiprocessing offers data-processing capabilities that are not present when only one CPU is used. Many complex operations can be performed at the same time. CPU can function on



Information Technology

complementary units to provide data and control for one another. Multiprocessing is used for nation's major control applications such as rail road control, traffic control, or airways etc.

Although parallel processing is not widespread yet, multiprocessing should be the wave of the future. Because of the availability of cheaper but more powerful processors, many computer manufacturers are now designing hardware and software systems to do multiprocessing. Since several machines can work as a team and operate in parallel, jobs can be processed much more rapidly than on a single machine..

1.4 OTHER SYSTEMS SOFTWARE

1.4.1 Utility Programs or Service Programs: Utility programs are systems programs that perform general system –support tasks. These programs are provided by the computer manufacturers to perform tasks that are common to all data processing installations. Some of them may either be programs in their own right or subordinates to be assembled/compiled in the application programs. The following tasks are performed by the utility programs.

- (i) Sorting the data.
- (ii) Editing the output data.
- (iii) Converting data from one recording medium to another, viz., floppy disk to hard disc, tape to printer, etc.
- (iv) Dumping of data to disc or tape.
- (v) Tracing the operation of program.

In many instances, it is unclear what differentiates an operating system routine from a utility program. Some programs that one vendor bundles into an operating system might be offered by another vendor as separately priced and packaged utility programs.

A wide variety of utilities are available to carry out special tasks. Three types of utility programs found in most computer systems: sort utilities, spooling software, and text editors are discussed below:

(a) Sort utilities: Sort utility programs are those that sort data. For example, suppose we have a file of student records. We could declare “name” the *primary sort key* and arrange the file alphabetically on the name field. This would be useful for, perhaps, producing a student directory. Alternatively, we could sort the file by name, and then within name, by date-of-birth. Hence, we would declare name as the primary sort key and date-of-birth as the *secondary sort key*. Although the examples described here use only one or two sort keys, many sorting packages enable the user to identify 12 or more sort keys and to arrange outputted records in



either ascending or descending order on each declared key.

Sort utilities are often found in mainframe and minicomputer environments. In the micro-computing world, it is typical for sort routines to be bundled into application packages; for example, sort routines are commonly found in spreadsheet and database management software.

(b) Spooling software: The purpose of spooling software is to compensate for the speed differences between the computer and its peripheral devices. Spooling software is usually encountered in large system and network computing environments. For instance, during the time it takes to type in or print out all the words on this page, the computer could begin and finish processing dozens of programs. The computer would be horribly bottlenecked if it had to wait for slow input and output devices before it could resume processing. It just does not make sense for a large computer, which may be worth lacs of rupees, to spend any time sitting idle because main memory is full of processed but unprinted jobs and the printer attached to the system cannot move fast enough.

To preclude the computer from being slowed down by input and output devices, many computer systems employ **spooling** software. These programs take the results of computer programs and move them from primary memory to disk. The area on the disk where the program results are sent is commonly called the output spooling area. Thus, the output device can be left to interact primarily with the disk unit, not the CPU. Spooling utilities can also be used on the input side, so that programs and data to be processed are temporarily stored in an input spooling area on disk.

Assume for example that a floppy disc, a line printer and a disk are used in a spooling operation on a computer system to process the pay-roll and prepare invoices by loading both programs into the main memory. While the line printer is printing an invoice line, the processor switches to the pay roll application and transfers input data from floppy disc to magnetic disk. Afterwards the processor reverts back to the invoice application. As the printer is being used for printing invoices, pay roll application will be executed and output data would be recorded on the Magnetic disk for later conversion when the printer becomes available. As a result, the CPU can give the output at the maximum speed, while several relatively slow input and output units operate simultaneously to process it.

(c) Text editors: Text editors are programs that allow text in a file to be created and modified. These utilities are probably most useful to professional programmers, who constantly face the problems of cutting and pasting programs together, changing data files by eliminating certain data fields, changing the order of certain data fields, adding new data fields, and changing the format of data. Although text editors closely resemble word processors, they are not the same. Word processors are specifically designed to prepare such



Information Technology

“document” materials as letters and reports, where text editors are specifically designed to manipulate “non-document” instructions in computer programs or data in files. Text editors lack the extensive text-formatting and document-printing capabilities found on most word processors.

Some of the other commonly used utilities for a microcomputer operating systems are discussed below :

- (i) *Disk copy program* - This program allows an user to copy the entire contents of one diskette to another diskette. It is generally used to make a backup or archive copy of a data diskette or an application program. The diskcopy program can also be used to transfer data stored from one size or capacity diskette to another. For example, it can be used to transfer data from a 360 KB diskette to a 1.2MB diskette or from a 5¼ inch diskette to a 3½ inch diskette.
- (ii) *File copy program* - This program allows an user to copy just one file or a group of files, rather than the entire contents of the diskette, to be copied to another diskette. It has the same functions as a diskcopy utility except that it allows an individual file or group of files to be copied.
- (iii) *Disk formatting program* - This program allows an user to prepare a new, blank diskette to receive data from the computer system. The data can not be stored on a diskette until it is formatted or initialized. The formatting process writes the sectors on the diskette so that the operating system is able to place data in these locations.
- (iv) *File deletion program* - It allows an user to delete a file stored on a diskette.
- (v) *File viewing program* - This program is used to view the contents of a file on the display screen of the microcomputer.
- (vi) *Directory program* - This program allows an user to view the names of the data and program files which are stored on a disk/diskette. It will not only list the files, but also will show the amount of kilobytes of memory these files occupy, the time and day they were last revised and the amount of unused storage space available on the floppy.

1.4.2 Diagnostic Routines : These programs are usually written and provided by the computer manufacturers. They assist in program debugging. They usually trace the processing of the program being debugged. In personal computer, if an user wants to know anything regarding the processing equipments in his computers, he/she can consult the Microsoft Diagnostic Program, a utility built into the DOS version 6.0 operating system. Using the program, one can get answer to following questions e.g. :—



Introduction To Computers

1. What type of processor does the computer use?
2. Is there a math coprocessor in the computer ?
3. Who is the BIOS manufacturer? (BIOS stands for Basic Input/Output System. It is a set of instructions, contained on a ROM chip, that are loaded in the computer memory before the operating system. The BIOS manufacturer is the maker of the ROM chip)
4. What is the total amount of conventional memory in the computer?
5. What type of keyboard is attached to the computer?
6. What is the display type?
7. If there is a mouse attached to the computer, what type is it and who made it?

The diagnostic routines are however also often treated as a category of the utility or service programs

1.4.3 Language translators : A language translator or language processor is a general term used for any assembler, compiler or other routine that accepts statements in one language and produces equivalent statements in another language. The language processor reads the source language statements one at a time and prepares a number of machine instructions to perform the operations specified or implied by each source statement. Most computer installations have several language processors available, one for each programming language the computer can accept.

The three most widely used types of language translators are compilers, interpreters, and assemblers.

Compilers : A compiler translates the entire program into machine language before the program is executed. Compilers are most commonly used to translate high-level languages such as COBOL, FORTRAN, and Pascal. Compilers typically result in programs that can be executed much more swiftly than those handled by interpreters. Since either a compiler or an interpreter can be developed to translate most languages, compilers would be preferred in environments where execution speed is important.

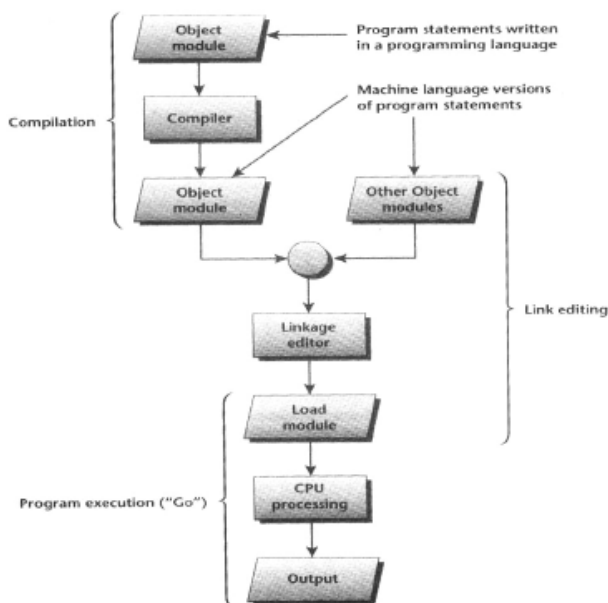


Fig. 1.4.2.1 Compiling Process

Compilers work in the manner illustrated in Figure 1.4.2.1. Program is entered into the computer system and submitted to the appropriate compiler. For instance, A COBOL program is input to a COBOL compiler; a Pascal program, to a Pascal compiler. The program submitted for compilation is called a *source program* (or source module) The compiler then translates the program into machine language, producing an *object program* (or object module) Then, another software program called a **linkage editor** binds the object module of this program to object modules of any subprograms that must be used to complete processing. The resultant program, which is ready for computer execution, is called a *load program* (or load module) It is the load program that the computer actually executes.

The entire process is sometimes referred to as “compile/link-edit/go,” corresponding to the compilation, link-editing, and execution stages that the user must go through to get a program processed. Programs can be saved on disk for processing in either source, object, or load-module form. Frequently run applications will often be saved in load module form to avoid repeated compilation and link-editing.



Interpreters: Whereas compilers translates programs into machine language all at once before programs are run, *interpreters* translate programs a line at a time as they are being run. For instance, if a user has a program in which a single statement is executed a thousand times during the course of the program's run, the interpreter would translate that statement a thousand different times into machine language. With an interpreter, each statement is translated into machine language just before it is executed. No object module or storable load module is ever produced.

Although interpreters have the glaring weakness of inefficiency because they translate statements over and over, they do have some advantages over compilers. First, they are usually easier and faster to use, since the user is not bothered with distinct and time-consuming compile, link-edit, and execution stages. Second, they typically provide users with superior error messages. When a program contains an error and "blows up," the interpreter knows exactly which statement triggered the error – the one it last translated. Because interpreters stop when errors are encountered, they help programmers debug their programs. This boosts programmer productivity and reduces program development time. Syntax errors encountered by compilers during the program translation process are counted, but the diagnostic routines and error messages associated with most compilers do not help programmers locate errors as readily as interpreters. Third, an interpreter for a 3GL typically requires less storage space in primary memory than a compiler for that language. So they may be ideal for programming environments in which main memory is limited, such as on low-end microcomputers. Fourth, interpreters are usually less expensive than compilers.

Assemblers : Assemblers are used exclusively with assembly languages. They work similarly to compilers, translating an assembly language program into object code. Because assembly language programs are usually more machine efficient than those written in high-level languages, a two-step translation process may take place. First, the high-level language is translated to assembly language; then, using an assembler, it is converted to machine language.

1.4.5 Firmware : Firmware or micro programs refer to a series of special program instructions. The most basic operations such as addition, multiplication etc. in a computer are carried out by hardwired circuits. These fundamental tasks are then combined in the form of micro programs to produce higher level operations such as move data, make comparisons etc. These micro-programs are called firmware because they deal with very low-level machine operations and thus essentially substitute for additional hardware. Firmware are held in the CPU in a special control storage device.

1.5 GENERAL PURPOSE SOFTWARE/UTILITIES



Information Technology

This software provides the framework for a great number of business, scientific, and personal applications. Spreadsheet, data bases, Computer-Aided Design (CAD) and word processing software etc. fall into this category. Most general-purpose software is sold as a package. The software is accompanied by user-oriented documentation such as reference manuals, keyboard templates, and so on. It is then upto the user of the software to create the application. For example, an accountant can use spreadsheet software to create a template for preparing a balance sheet of a company. An aeronautical engineer can use CAD software to design an airplane or an airport. A personnel manager can use word processing software to create a letter and so on.

The three basic types of software are; commercial, shareware and open source software. Some software is also release into the public domain without a license. **Commercial** software comes prepackaged and is available from software stores and through the Internet. **Shareware** is software developed by individual and small companies that cannot afford to market their software world wide or by a company that wants to release a demonstration version of their commercial product. You will have an evaluation period in which you can decide whether to purchase the product or not. Shareware software often is disabled in some way and has a notice attached to explain the legal requirements for using the product. **Open Source software** is created by generous programmers and released into the public domain for public use. There is usually a copyright notice that must remain with the software product. Open Source software is not public domain in that the company or individual that develops the software retains ownership of the program but the software can be used freely. Many popular Open Source applications are being developed and upgraded regularly by individuals and companies that believe in the Open Source concept.

1.5.1 Word Processor : A **word processor** (more formally a **document preparation system**) is a computer application used for the production (including composition, editing, formatting, and possibly printing) of any sort of printable material.

Word processors are descended from early **text formatting** tools (sometimes called **text justification** tools, from their only real capability) Word processing was one of the earliest applications for the personal computer in office productivity.

Although early word processors used tag-based markup for document formatting, most modern word processors take advantage of a graphical user interface. Most are powerful systems consisting of one or more programs that can produce any arbitrary combination of images, graphics and text, the latter handled with type-setting capability.

Microsoft Word is the most widely used computer word processing system; Microsoft estimates over five million people use the Office suite. There are also many other commercial



word processing applications, such as WordPerfect. Open-source applications such as OpenOffice's Writer and KWord are rapidly gaining in popularity.

1.5.2 Spreadsheet Program : A **spreadsheet** is a rectangular table (or grid) of information, often financial information. The word came from "spread" in its sense of a newspaper or magazine item (text and/or graphics) that covers two facing pages, extending across the center fold and treating the two pages as one large one. The compound word "spread-sheet" came to mean the format used to present bookkeeping ledgers -- with columns for categories of expenditures across the top, invoices listed down the left margin, and the amount of each payment in the cell where its row and column intersect -- which were traditionally a "spread" across facing pages of a bound ledger (book for keeping accounting records) or on oversized sheets of paper ruled into rows and columns in that format and approximately twice as wide as ordinary paper.

1.5.3 Database Management System: A **database management system (DBMS)** is a system or software designed to manage a database, and run operations on the data requested by numerous clients. Typical examples of DBMS use include accounting, human resources and customer support systems. DBMSs have more recently emerged as a fairly standard part of any company back office.

1.5.4 Internet Browser : An Internet Browser or a **web browser** is a software application that enables a user to display and interact with text, images, and other information typically located on a web page at a website on the World Wide Web or a local area network. Text and images on a web page can contain hyperlinks to other web pages at the same or different websites. Web browsers allow a user to quickly and easily access information provided on many web pages at many websites by traversing these links.

Web browsers available for personal computers include Microsoft Internet Explorer, Mozilla Firefox, Apple Safari, Netscape, and Opera, in order of descending popularity (July 2006) Web browsers are the most commonly used type of HTTP user agent. Although browsers are typically used to access the World Wide Web, they can also be used to access information provided by web servers in private networks or content in file systems.

1.5.5 Electronic mail, abbreviated as e-mail or e-Mail or email, is a method of composing, sending, storing, and receiving messages over electronic communication systems. The term e-mail applies both to the Internet e-mail system based on the Simple Mail Transfer Protocol (SMTP) and to intranet systems allowing users within one company to e-mail each other. Often these workgroup collaboration organizations may use the Internet protocols for internal e-mail service.



1.6 APPLICATION SOFTWARE

Application software is a loosely defined subclass of computer software that employs the capabilities of a computer directly to a task that the user wishes to perform. This should be contrasted with system software which is involved in integrating the computer's various capabilities, but typically does not directly apply them in the performance of tasks that benefit the user. The term application refers to both the application software and its implementation. A simple, if imperfect, analogy in the world of hardware would be the relationship of an electric light—an application—to an electric power generation plant—the system. The power plant merely generates electricity, itself not really of any use until harnessed to an application like the electric light which performs a service that the user desires.

Multiple applications bundled together as a package are sometimes referred to as an application suite.. The separate applications in a suite usually have a user interface that has some commonality making it easier for the user to learn and use each application. And often they may have some capability to interact with each other in ways beneficial to the user.

User-written software tailors systems to meet the user's specific needs. User-written software include spreadsheet templates, word processor macros, scientific simulations, graphics and animation scripts. Even email filters are a kind of user software. Users create this software themselves and often overlook how important it is.

The program usually solves a particular application or problem. Examples of such programs are payroll, General accounting, sales statistics and inventory control etc. Usually different organisations require different programs for similar application and hence it is difficult to write standardized programs. However, tailor-made application software can be written by software houses on *modular* design to cater to the needs of different users.

1.6.1 Enterprise Resource Planning systems (ERPs) integrate (or attempt to integrate) all data and processes of an organization into a single unified system. A typical ERP system will use multiple components of computer software and hardware to achieve the integration. A key ingredient of most ERP systems is the use of a single, unified database to store data for the various system modules.

The term ERP originally implied systems designed to plan the utilization of enterprise-wide resources. Although the acronym ERP originated in the manufacturing environment, today's use of the term ERP systems has much broader scope. ERP systems typically attempt to cover all basic functions of an organization, regardless of the organization's business or charter. Business, non-profit organizations, governments, and other large entities utilize ERP systems.



Introduction To Computers

Additionally, it may be noted that to be considered an ERP system, a software package generally would only need to provide functionality in a single package that would normally be covered by two or more systems. Technically, a software package that provides both Payroll and Accounting functions (such as QuickBooks) would be considered an ERP software package.

However; the term is typically reserved for larger, more broadbased applications. The introduction of an ERP system to replace two or more independent applications eliminates the need for interfaces previously required between systems, and provides additional benefits that range from standardization and lower maintenance (one system instead of two or more) to easier and/or greater reporting capabilities (as all data is typically kept in one database)

Examples of modules in an ERP which formerly would have been stand-alone applications include: Manufacturing, Supply Chain, Financials, CRM, Human Resources, and Warehouse Management.

1.6.2 DECISION SUPPORT SYSTEMS

Decision support systems are information processing systems frequently used by accountants, managers, and auditors to assist them in the decision-making process. The concept of decision support systems evolved in the 1960s from studies of decision making in organisations. These studies noted that managers required flexible systems to respond to less well-defined questions than those addressed by operational employees. Advances in hardware technology, interactive computing design, graphics capabilities, and programming languages contributed to this evolution. Decision support systems have achieved broad use in accounting and auditing today.

Characteristics of Decision Support Systems : Although decision support system applications vary widely in their level of sophistication and specific purpose, they possess several characteristics in common.

(1) *Decision support system support management decision making* – Although most heavily used for management planning decisions, operational managers can use them (e.g., to solve scheduling problems), as can top managers (e.g., to decide whether to drop a product line) Decision support systems enhance decision quality. While the system might point to a particular decision, it is the user who ultimately makes the final choice.

(2) *Decision support systems solve relatively unstructured problems* – problems that do not have easy solution procedures and therefore problems in which some managerial judgment is necessary in addition to structured analysis. Thus, in contrast to transaction processing systems, decision support systems typically use non-routine data as input. These data are not easy to gather and might require estimates. For example, imagine that the manager is selecting accounting software



Information Technology

for his company's use. This problem is unstructured because there is no available listing of all the features that are desirable in accounting software for his particular company. Furthermore, he will need to use his judgment to determine what features are important.

Because managers must plan for future activities, they rely heavily on assumptions of future interest rates, inventory prices, consumer demand, and similar variables. But what if managers' assumptions are wrong? A key characteristic of many decision support systems is that they allow users to ask what-if questions and to examine the results of these questions. For instance, a manager may build an electronic spreadsheet model that attempts to forecast future departmental expenditures. The manager cannot know in advance how inflation rates might affect his or her projection figures, but can examine the consequences of alternate assumptions by changing the parameters (here, growth rates) influenced by these rates. Decision support systems are useful in supporting this type of analysis.

Although systems designers may develop decision support systems for one-time use, managers use them to solve a particular type of problem on a regular basis. The same is true of expert systems. However, decision support systems are much more flexible and may handle many different types of problems. Accountants might use a spreadsheet model developed to calculate depreciation only for depreciation problems, but many more general decision support system tools such as Expert Choice (discussed later as an example of a decision support system) are sufficiently flexible and adaptive for ongoing use. Another example is decision support systems that perform data mining tasks.

(3) Finally, a "friendly" computer interface is also a characteristic of a decision support system – Because managers and other decision makers who are non programmers frequently use decision support systems, these systems must be easy to use. The availability of nonprocedural modeling languages, eases communication between the user and the decision support system.

Components of Decision Support Systems : A decision support system has four basic components: (1) the user, (2) one or more databases, (3) a planning language, and (4) the model base (see Figure 1.6.2.1)

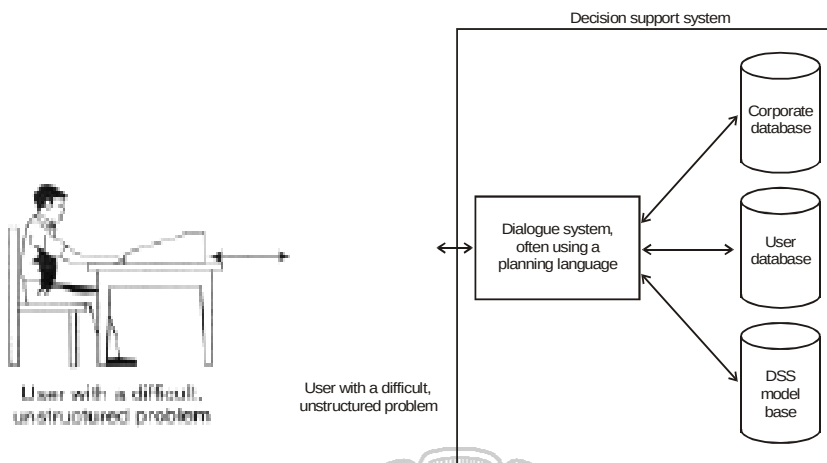


Fig. 1.6.2.1 Components decision support system

(i) **The users** : The user of a decision support system is usually a manager with an unstructured or semi-structured problem to solve. The manager may be at any level of authority in the organisation (e.g., either top management or operating management). Typically, users do not need a computer background to use a decision support system for problem solving. The most important knowledge is a thorough understanding of the problem and the factors to be considered in finding a solution. A user does not need extensive education in computer programming in part because a special planning language performs the communication function within the decision support system. Often, the planning language is nonprocedural, meaning that the user can concentrate on what should be accomplished rather than on how the computer should perform each step.

(ii) **Databases** : Decision support systems include one or more databases. These databases contain both routine and non-routine data from both internal and external sources. The data from external sources include data about the operating environment surrounding an organisation – for example, data about economic conditions, market demand for the organisation's goods or services, and industry competition.

Decision support system users may construct additional databases themselves. Some of the data may come from internal sources. An organisation often generates this type of data in the normal course of operations – for example, data from the financial and managerial accounting systems such as account, transaction, and planning data. The database may also capture data from other subsystems such as marketing, production, and personnel. External data



Information Technology

include assumptions about such variables as interest rates, vacancy rates, market prices, and levels of competition.

(iii) Planning languages : Two types of planning languages that are commonly used in decision support systems are: (1) general –purpose planning languages and (2) special-purpose planning languages. General-purpose planning languages allow users to perform many routine tasks – for example, retrieving various data from a database or performing statistical analyses. The languages in most electronic spreadsheets are good examples of general-purpose planning languages. These languages enable user to tackle a broad range of budgeting, forecasting, and other worksheet-oriented problems. Special-purpose planning languages are more limited in what they can do, but they usually do certain jobs better than the general-purpose planning languages. Some statistical languages, such as SAS, SPSS, and Minitab, are examples of special purpose planning languages.

(iv) Model base : The planning language in a decision support system allows the user to maintain a dialogue with the model base. **The model base** is the “brain” of the decision support system because it performs data manipulations and computations with the data provided to it by the user and the database. There are many types of model bases, but most of them are custom-developed models that do some types of mathematical functions-for example, cross tabulation, regression analysis, time series analysis, linear programming and financial computations. The analysis provided by the routines in the model base is the key to supporting the user's decision. The model base may dictate the type of data included in the database and the type of data provided by the user. Even where the quantitative analysis is simple, a system that requires users to concentrate on certain kinds of data can improve the effectiveness of decision making.

Examples of Decision Support Systems in Accounting : Decision support systems are widely used as part of an organisation's AIS. The complexity and nature of decision support systems vary. Many are developed in-house using either a general type of decision support program or a spreadsheet program to solve specific problems. Below are several illustrations of these systems.

Cost Accounting system: The health care industry is well known for its cost complexity. Managing costs in this industry requires controlling costs of supplies, expensive machinery, technology, and a variety of personnel. Cost accounting applications help health care organisations calculate product costs for individual procedures or services. Decision support systems can accumulate these product costs to calculate total costs per patient. Health care managers many combine cost accounting decision support systems with other applications,



such as productivity systems. Combining these applications allows managers to measure the effectiveness of specific operating processes. One health care organisation, for example, combines a variety of decision support system applications in productivity, cost accounting, case mix, and nursing staff scheduling to improve its management decision making.

Capital Budgeting System: Companies require new tools to evaluate high-technology investment decisions. Decision makers need to supplement analytical techniques, such as net present value and internal rate of return, with decision support tools that consider some benefits of new technology not captured in strict financial analysis. One decision support system designed to support decisions about investments in automated manufacturing technology is AutoMan, which allows decision makers to consider financial, nonfinancial, quantitative, and qualitative factors in their decision-making processes. Using this decision support system, accountants, managers, and engineers identify and prioritize these factors. They can then evaluate up to seven investment alternatives at once.

Budget Variance Analysis System: Financial institutions rely heavily on their budgeting systems for controlling costs and evaluating managerial performance. One institution uses a computerized decision support system to generate monthly variance reports for division comptrollers. The system allows these comptrollers to graph, view, analyse, and annotate budget variances, as well as create additional one-and five-year budget projections using the forecasting tools provided in the system. The decision support system thus helps the controllers create and control budgets for the cost-center managers reporting to them.

General Decision Support System: As mentioned earlier, some planning languages used in decision support systems are general purpose and therefore have the ability to analyze many different types of problems. In a sense, these types of decision support systems are a decision maker's tools. The user needs to input data and answer questions about a specific problem domain to make use of this type of decision support system. An example is a program called **Expert Choice**. This program supports a variety of problems requiring decisions. The user works interactively with the computer to develop a hierarchical model of the decision problem. The decision support system then asks the user to compare decision variables with each other. For instance, the system might ask the user how important cash inflows are versus initial investment amount to a capital budgeting decision. The decision maker also makes judgments about which investment is best with respect to these cash flows and which requires the smallest initial investment. Expert Choice analyzes these judgments and presents the decision maker with the best alternative.

1.6.3 ARTIFICIAL INTELLIGENCE



Information Technology

Artificial intelligence (AI) is software that tries to emulate aspects of human behavior, such as reasoning, communicating, seeing, and hearing. AI software can use its accumulated knowledge to reason and, in some instances, learn from experience and thereby modify its subsequent reasoning. There are several types of AI, including natural language, voice and visual recognition, robotics, neural networks, and expert systems.

Natural language and voice and visual recognition both focus on enabling computers to interact more easily and naturally with users. Robotics focuses on teaching machines to replace human labour. Both neural networks and expert systems aim to improve decision-making.

1.6.4 EXPERT SYSTEMS

An expert system (ES) is a computerized information system that allows non-experts to make decisions comparable to those of an expert. Expert systems are used for complex or ill-structured tasks that require experience and specialised knowledge in narrow, specific subject areas, as shown in Figure 1.6.4.1, expert systems typically contain the following components:





Introduction To Computers

Major Components of an Expert System

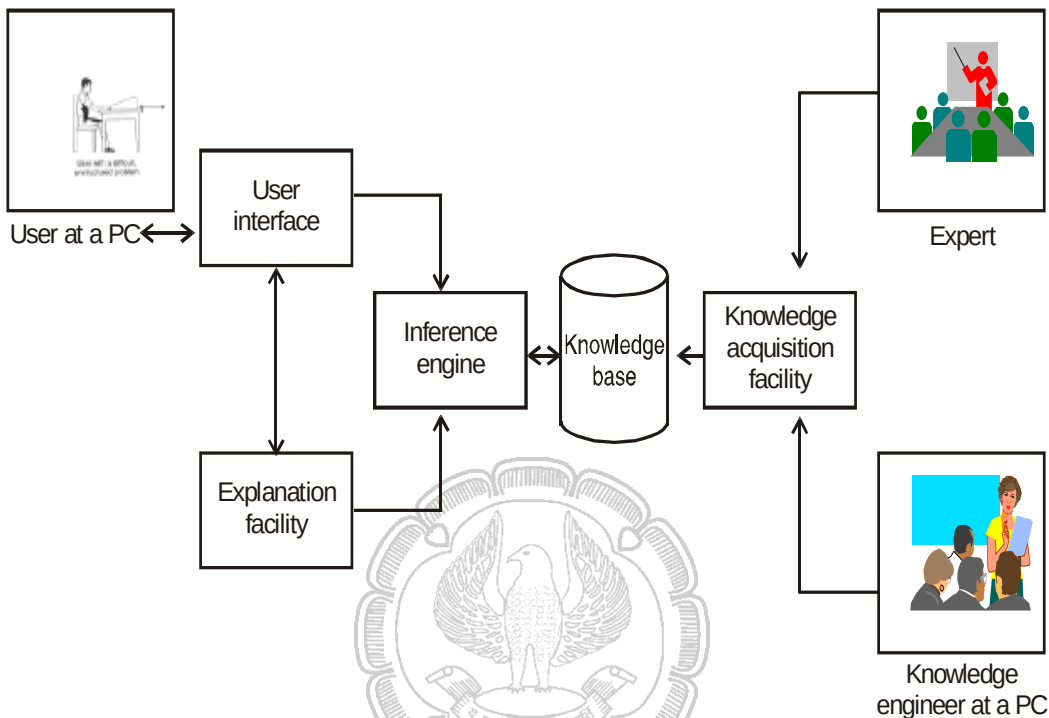


Fig 1.6.4.1

1. **Knowledge base** : This includes the data, knowledge, relationships, rules of thumb (heuristics), and decision rules used by experts to solve a particular type of problem. A knowledge base is the computer equivalent of all the knowledge and insight that an expert or a group of experts develop through years of experience in their field.
2. **Inference engine** : This program contains the logic and reasoning mechanisms that simulate the expert logic process and deliver advice. It uses data obtained from both the knowledge base and the user to make associations and inferences, form its conclusions, and recommend a course of action.
3. **User interface** : This program allows the user to design, create, update, use, and communicate with the expert system.
4. **Explanation facility** : This facility provides the user with an explanation of the logic the ES used to arrive at its conclusion.



Information Technology

5. **Knowledge acquisition facility** : Building a knowledge base, referred to as knowledge engineering, involves both a human expert and a knowledge engineer. The knowledge engineer is responsible for extracting an individual's expertise and using the knowledge acquisition facility to enter it into the knowledge base.

Expert systems can be example-based, rule based, or frame based. Using an example-based system, developers enter the case facts and results. Through induction the ES converts the examples to a decision tree that is used to match the case at hand with those previously entered in the knowledge base. Rule-based systems are created by storing data and decision rules as if- then -else rules. The systems asks the user questions and applied the if-then-else rules to the answers to draw conclusions and make recommendations. Rule -based systems are appropriate when a history of cases is unavailable or when a body of knowledge can be structured within a set of general rules. Frame -based systems organize all the information (data, descriptions, rules etc.) about a topic into logical units called frames, which are similar to linked records in data files. Rules are then established about how to assemble or interrelate the frames to meet the user's needs.

Expert systems provide several levels of expertise. Some function as assistants that perform routine analysis and call the user's attention to tasks that require human expertise. Others function as colleagues, and the user "discusses" a problem with the system until both agree on a solution. When a user can accept the system's solution without question, the expert system can be referred to as a true expert. Developers of expert systems are still striving to create a true expert; most current systems function at the assistant or colleague level.

Expert systems offer the following benefits:

- ☐ They provide a cost-effective alternative to human experts.
- ☐ They can outperform a single expert because their knowledge is representative of numerous experts. They are faster and more consistent and do not get distracted, overworked, or stressed out.
- ☐ They produce better-quality and more consistent decisions. Expert systems assist users in identifying potential decision making problems, which increases the probability that sound decisions will be made.
- ☐ They can increase productivity.
- ☐ They preserve the expertise of an expert leaving the organization.

Although expert systems have many advantages and great promise, they also have a



Introduction To Computers

significant number of limitations

- ☐ Development can be costly and time-consuming. Some large systems required upto 15 years and millions of dollars to develop.
- ☐ It can be difficult to obtain knowledge from experts who have difficulty specifying exactly how they make decisions.
- ☐ Designers have not been able to program what humans consider common sense into current systems. consequently, rule-based systems break down when presented with situations they are not programmed to handle.
- ☐ Until recently, developers encountered skepticism from businesses due to the poor quality of the early expert systems and the high expectations of users.

As technology advances, some of these problems will be overcome and expert systems will play an increasingly important role in accounting information systems. Here are specific examples of companies that have successfully used expert systems:

- ☐ The IRS analyzes tax returns to determine which should be passed on to tax fraud investigators.
- ☐ IBM designs and evaluates internal controls in both new and existing applications.
- ☐ American Express authorizes credit card purchases to minimize fraud and credit losses. Its ES replaced 700 authorization clerks and saved tens of millions of dollars.

Self-examination questions

Multiple Choice Questions

1. First Generation of computer have used the following
 - a) Vacuum Tubes
 - b) Transistor
 - c) IC
 - d) VLSI
2. Second generation computer have used the following.
 - a) Vacuum Tubes
 - b) Transistor
 - c) IC



Information Technology

- d) VLSI
- 3. Third Generation Computer have used the following circuitry options
 - a) Vacuum Tubes
 - b) Transistor
 - c) IC
 - d) VLSI
- 4. Forth Generation computer have used the following Circuitry options
 - a) Vacuum Tubes
 - b) Transistor
 - c) IC
 - d) VLSI
- 5. First Generation have used the following language
 - a) Object oriented programmer
 - b) Artificial Intelligence
 - c) Assembly Language
 - d) Assembly & M/C Language
- 6. Which is not the attribute of file :
 - a) Hidden
 - b) System
 - c) Big
 - d) Archive
- 7. Which is the fastest storage device :
 - a) Floppy Disk
 - b) Zip Disk
 - c) CD - Rom
 - d) Hard Disk
- 8. Storage Device which uses light to store data :



- a) Floppy Disk
- b) Hard Disk
- c) CD - ROM
- d) Tape Drive

Answers:

1 a 2 b 3 c 4 d 5 d 6 d 7 c 8 c

Short Term Questions

1. Write short notes on the following types of computers :-
 - (i) Super computer
 - (i) Mainframe computer.
 - (iii) Mini computer.
 - (iv) Micro computer.
 - (v) Workstations
 - (vi) Server
2. What are the features of the Central Processing Unit?
3. Discuss various components of a motherboard.
- 4.. What do you understand by the term 'Bus'? Discuss two types of bus available on a computer.
5. Write short notes on the following :
 - (i) RAM
 - (ii) ROM
 - (iii) Bubble memory
 - (iv) Flash memory
6. Write short note on floppy diskette as an input medium.
7. What are the factors that determine the number of characters that can be stored in a floppy diskette?
8. Explain the following terms :



Information Technology

- (i) The boot record
 - (ii) File allocation table
9. What care is required for using and storage of a diskette?
10. Differentiate between floppy diskettes and hard disks.
11. Briefly explain the various characteristics of a hard disk.
12. What are the advantages and disadvantages of direct access storage?
13. Explain the following terms :
- (i) CD-ROM
 - (ii) WORM disk
 - (iii) Magneto Optical Disks
 - (iv) Video Disk
 - (v) Detachable Reel Magnetic Tape
 - (vi) Tape Cartridge Systems

Long Term Questions

- 1. Describe in detail various generations of computers.
- 2. Draw the schematic diagram of a computer. Discuss each of the components covered in it.
- 3. Discuss, in detail, various types of microprocessors.
- 4. Describe in detail various generations of computers.
- 5. Draw the schematic diagram of a computer. Discuss each of the components covered in it.
- 6. Discuss, in detail, various types of microprocessors

CHAPTER 2

DATA STORAGE, RETRIEVAL AND DATA BASE MANAGEMENT SYSTEMS

Learning Objectives

- ◆ Understand number systems and file organizations.
- ◆ Understand the difference between relational database systems and others.
- ◆ Understand the need of data warehousing and data mining in business applications.

2.0 INTRODUCTION

We use the decimal numbers or the decimal number system for our day-to-day activities. As we all know, in the decimal number system there are ten digits – 0 through 9. But computers understand only 0s and 1s - the machine language. But using 0s and 1s to program a computer is a thing in the past. Now we can use the decimal numbers, the alphabets and special characters like +, -, *, ?, /, etc. for programming the computer. Inside the computer, these decimal numbers, alphabets and the special characters are converted into 0s and 1s, so that the computer can understand what we are instructing it to do. To understand the working of a computer, the knowledge of binary, octal and hexadecimal number systems is essential.

2.1 DECIMAL NUMBER SYSTEM

The base or radix of a number system is defined as the number of digits it uses to represent the numbers in the system. Since decimal number system uses 10 digits - 0 through 9 - its base or radix is 10. The decimal number system is also called *base-10 number system*. The weight of each digit of a decimal number system depends on its relative position within the number. For example, consider the number 3256.

$3256 = 3000 + 200 + 50 + 6$, or, in other words,

$$3256 = 3 \times 10^3 + 2 \times 10^2 + 5 \times 10^1 + 6 \times 10^0$$

From the above example, we can see that the weight of the n th digit of the number from the right hand side is equal to $n^{\text{th}} \text{ digit} \times 10^{n-1}$ which is again equal to $n^{\text{th}} \text{ digit} \times (\text{base})^{n-1}$. The number system, in which the weight of each digit depends on its relative position within the numbers, is called the *positional number system*.



Information Technology

2.1.1 Binary Number System

The base or radix of the binary number system is 2. It uses only two digits - 0 and 1. Data is represented in a computer system by either the presence or absence of electronic or magnetic signals in its circuitry or the media it uses. This is called a binary or two-state representation of data since the computer is indicating only two possible states or conditions. For example, transistors and other semiconductor circuits are either in a conducting or non-conducting state. Media such as magnetic disks and tapes indicate these two states by having magnetized spots whose magnetic fields can have two different directions or polarities. These binary characteristics of computer circuitry and media are the primary reasons why the binary number system is the basis for data representation in computers. Thus, for electronic circuits, the conduction state (ON) represents a ONE and the non-conducting state (OFF) represents a ZERO.

Therefore, as mentioned earlier, the binary number system has only two symbols, 0 and 1. The binary symbol 0 or 1 is commonly called a bit, which is a contraction of the term binary digit. In the binary system, all numbers are expressed as groups of binary digits (bits), that is, as groups of 0s and 1s. Just as in any other number system, the value of a binary number depends on the position or place of each digit in a grouping of binary digits. The values are based on the right to left position of digits in a binary number, using the power of 2 as position values. For example, consider the binary number 10100.

$$10100 = 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 = 16 + 0 + 4 + 0 + 0 = 20$$

Table 1 gives the binary equivalents of the decimal numbers from 0 to 20.

Table 1: Binary equivalents of the decimal numbers

Decimal	Binary	Decimal	Binary
0	0	11	1011
1	01	12	1100
2	10	13	1101
3	11	14	1110
4	100	15	1111
5	101	16	10000
6	110	17	10001
7	111	18	10010
8	1000	19	10011
9	1001	20	10100
10	1010		



Data Storage, Retrieval and Data Base Management Systems

Binary-decimal Conversion

To convert a binary number to its decimal equivalent, we use the following expression:

The weight of the n^{th} bit of a number from the right hand side = n^{th} bit $\times 2^{n-1}$

After calculating the weight of each bit, they are added to get the decimal value as shown in the following examples:

$$101 = 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 4 + 0 + 1 = 5$$

$$1010 = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 8 + 0 + 2 + 0 = 10$$

$$1111 = 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 8 + 4 + 2 + 1 = 15$$

$$1.001 = 1 \times 2^0 + 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 1 + 0 + 0 + .125 = 1.125$$

2.1.2 Decimal-binary Conversion: Decimal numbers are converted into binary by a method called *Double Dabble Method*. In this method, the mantissa part of the number is repeatedly divided by two and noting the remainders, which will be either 0 or 1. This division is continued till the mantissa becomes zero. The remainders, which are noted down during the division is read in the reverse order to get the binary equivalent. This can be better illustrated using the following example.

2	14	Remainder
2	7	0
2	3	1
2	1	1
2	0	1

The number is written from below, that is
1110. So the binary equivalent of 14 is
1110

If the decimal number has a fractional part, then the fractional part is converted into binary by multiplying it with 2. Only the integer of the result is noted and the fraction is repeatedly multiplied by 2 until the fractional part becomes 0. This can be explained using the following example.

$$\begin{array}{r} 0.125 \\ \times 2 \\ \hline 0.25 \quad 0 \\ \times 2 \\ \hline 0.5 \quad 0 \\ \times 2 \\ \hline 1.0 \quad 1 \end{array}$$

Here the number is written from top - 0.001. So the binary equivalent of 0.125 is 0.001

2.3



Information Technology

2.1.3 Fixed-Point Representation of Number

In the fixed-point number representation system, all numbers are represented as integers or fractions. Signed integer or BCD numbers are referred to as fixed-point numbers because they contain no information regarding the location of the decimal point or the binary point. The binary or decimal point is assumed to be at the extreme right or left of the number.

If the binary or decimal point is at the extreme right of the computer word, then all numbers are positive or negative integers. If the radix point is assumed to be at the extreme left, then all numbers are positive or negative fractions.

Consider that you have to multiply 23.15 and 33.45. This will be represented as 2315×3345 . The result will be 7743675. The decimal point has to be placed by the user to get the correct result, which is 774.3675. So in the fixed-point representation system, the user has to keep track of the radix point, which can be a tedious job.

2.1.4 Floating-Point Representation of Number

In most computing applications, fractions are used very frequently. So a system of number representation which automatically keeps track of the position of the binary or decimal point is better than the fixed-point representation. Such a system is the floating-point representation of numbers.

A number, which has both an integer part and a fractional part, is called a real number or a floating-point number. These numbers can be either positive or negative. Examples of real numbers (decimal) are 123.23, -56.899, 0.008, etc. The real number 123.23 can be written as 1.2323×10^2 or 0.12323×10^3 . Similarly the numbers 0.008 and 1345.66 can be represented as 0.8×10^{-2} and 1.34566×10^3 respectively. This kind of representation is called the *scientific representation*. Using this scientific form, any number can be expressed as a combination of a mantissa and an exponent, or in other words, the number 'n' can be expressed as 'n = m re' where 'm' is the **mantissa**, 'r' is the **radix** of the number system and 'e' is the **exponent**.

In a computer also the real or floating-point number is represented by two parts mantissa and exponent. Mantissa is a signed fixed point number and the exponent indicates the position of the binary or decimal point. For example, the number 123.23 is represented in the floating-point system as:

Sign
0 .12323
Mantissa

Sign
0 03
Exponent

The zero in the leftmost position of the mantissa and exponent indicates the plus sign. The



Data Storage, Retrieval and Data Base Management Systems

mantissa can be either a fraction or an integer, which is dependent on the computer manufacturer. Most computers use the fractional system of representation for mantissa. The decimal point shown above is an assumed decimal point and is not stored in the register.

The exponent of the above example, +3, indicates that the actual decimal point is 3 digits to the right of the assumed one. In the above example, the mantissa is shown as a fraction. As mentioned, we can use an integer as **the mantissa**. The following example shows how it is done.

Sign
0 .12323
Mantissa

Sign
1 02
Exponent

In the above representation, the sign of the exponent is negative and it indicates that the actual decimal point lies 2 decimal positions to the left of the assumed point (in this case, the assumed decimal point is placed at the extreme right of the integer or 12323.)

A negative number say -123.23 can be expressed as follows.

Sign
1 .12323
Mantissa

Sign
0 03
Exponent

A negative fraction, say -0.0012323 can be represented as follows.

Sign
1 .12323
Mantissa

Sign
1 02
Exponent

2.1.5 Binary Coded Decimal (BCD)

The BCD is the simplest binary code that is used to represent a decimal number. In the BCD code, 4 bits represent a decimal number.

2.1.6 ASCII CODE

ASCII stands for American Standard Code for Information Interchange. ASCII code is used extensively in small computers, peripherals, instruments and communications devices. It has replaced many of the special codes that were previously used. It is a seven-bit code.



Information Technology

Microcomputers using 8-bit word length use 7 bits to represent the basic code. The 8th bit is used for parity or it may be permanently 1 or 0.

With 7 bits, up to 128 characters can be coded. A letter, digit or special symbol is called a character. It includes upper and lower case alphabets, numbers, punctuation mark and special and control characters.

ASCII-8 Code: A newer version of ASCII is the ASCII-8 code, which is an 8-bit code. With 8 bits, the code capacity is extended to 256 characters.

2.1.7 EBCDIC Code

EBCDIC stands for Extended BCD Interchange Code. It is the standard character code for large computers. It is an 8-bit code without parity. A 9th bit can be used for parity. With 8 bits up to 256 characters can be coded.

In ASCII-8 and EBCDIC, the first 4 bits are known as zone bits and the remaining 4 bits represent digit values. In ASCII, the first 3 bits are zone bits and the remaining 4 bits represent digit values.

Some examples of ASCII and EBCDIC values are shown in the table 2.

Table 2: ASCII and EBCDIC Codes

Character	ASCII	EBCDIC
0	00110000	11110000
1	00110001	11110001
2	00110010	11110010
3	00110011	11110011
4	00110100	11110100
5	00110101	11110101
6	00110110	11110110
7	00110111	11110111
8	00111000	11111000
9	00111001	11111001
A	01000001	11000001
B	01000010	11000010



Data Storage, Retrieval and Data Base Management Systems

C	01000011	11000011
D	01000100	11000100
E	01000101	11000101
F	01000110	11000110
G	01000111	11000111
H	01001000	11001000
I	01001001	11001001
J	01001010	11010001
K	01001011	11010010
L	01001100	11010011
M	01001101	11010100
N	01001110	11010101
O	01001111	11010110
P	01010000	11010111
Q	01010001	11011000
R	01010010	11011001
S	01010011	11100010
T	01010100	11100011
U	01010101	11100100
V	01010110	11100101
W	01010111	11100110
X	01011000	11100111
Y	01011001	11101000
Z	01011010	11101001

2.2 BITS, BYTES AND WORDS

A **byte** is a basic grouping of bits (binary digits) that the computer operates on as a single unit. It consists of 8 bits and is used to represent a character by the ASCII and EBCDIC coding

2.7

© The Institute of Chartered Accountants of India

Get Free SMS Updates on Mobile. Type ON<space>SIDRUTIYA and send it on 9870807070.



Information Technology

systems. For example, each storage location of computers using EBCDIC or ASCII-8 codes consist of electronic circuit elements or magnetic or optical media positions that can represent at least 8 bits. Thus each storage location can hold one character. The capacity of a computer's primary storage and its secondary storage devices is usually expressed in terms of bytes.

A word is a grouping of bits (usually larger than a byte) that is transferred as a unit between primary storage and the registers of the ALU and control unit. Thus, a computer with a 32-bit word length might have registers with a capacity of 32 bits, and transfer data and instructions within the CPU in groupings of 32 bits. It should process data faster than computers with a 16-bit or 8-bit word length.

However, processing speed also depends on the size of the CPU's data path or data bus, which are circuits that interconnect the various **CPU components**. For example, a microprocessor like the Intel 80386 SX has 32-bit registers but only a 16-bit data bus. Thus, it only moves data and instructions 16 bits at a time. Hence, it is slower than Intel 80386 DX microprocessor, which has 32-bit registers and data **paths**.

Lots of Bites: When you start talking about lots of bytes, you get into **prefixes** like kilo, mega and giga, as in kilobyte, megabyte and gigabyte (also shortened to K, M and G, as in Kbytes, Mbytes and Gbytes or KB, MB and GB) The following table shows the multipliers:

Name	Abbr.	Size
Kilo	K	$2^{10} = 1,024$
Mega	M	$2^{20} = 1,048,576$
Giga	G	$2^{30} = 1,073,741,824$
Tera	T	$2^{40} = 1,099,511,627,776$
Peta	P	$2^{50} = 1,125,899,906,842,624$
Exa	E	$2^{60} = 1,152,921,504,606,846,976$
Zetta	Z	$2^{70} = 1,180,591,620,717,411,303,424$
Yotta	Y	$2^{80} = 1,208,925,819,614,629,174,706,176$

You can see in this chart that kilo is about a thousand, mega is about a million, giga is about a billion, and so on. So when someone says, "This computer has a 2 giga hard drive," what he



Data Storage, Retrieval and Data Base Management Systems

or she means is that the hard drive stores 2 gigabytes, or approximately 2 billion bytes, or exactly 2,147,483,648 bytes. How could you possibly need 2 gigabytes of space? When you consider that one CD holds 650 megabytes, you can see that just three CDs worth of data will fill the whole thing! Terabyte databases are fairly common these days, and there are probably a few petabyte databases floating around the world by now.

2.3 CONCEPTS RELATED TO DATA

2.3.1 Double Precision: Real data values are commonly called single precision data because each real constant is stored in a single memory location. This usually gives seven significant digits for each real value. In many calculations, particularly those involving iteration or long sequences of calculations, single precision is not adequate to express the precision required. To overcome this limitation, many programming languages provide the double precision data type. Each double precision is stored in two memory locations, thus providing twice as many significant digits.

2.3.2 Logical Data Type: Use the Logical data type when you want an efficient way to store data that has only two values. Logical data is stored as true (.T.) or false (.F.)

Data Type	Description	Size	Range
Logical	Boolean value of true or false	1 byte	True (.T.) or False (.F.)

2.3.3 Characters: Choose the Character data type when you want to include letters, numbers, spaces, symbols, and punctuation. Character fields or variables store text information such as names, addresses, and numbers that are not used in mathematical calculations. For example, phone numbers or zip codes, though they include mostly numbers, are actually best used as Character values.

Data Type	Description	Size	Range
Character	1 byte per character to 254	1 byte	Any characters

2.3.4 Strings: A data type consisting of a sequence of contiguous characters that represent the characters themselves rather than their numeric values. A **String** can include letters, numbers, spaces, and punctuation. The **String** data type can store fixed-length strings ranging in length from 0 to approximately 63K characters and dynamic strings ranging in length from 0 to approximately 2 billion characters. The dollar sign (\$) type-declaration character represents a **String**.



Information Technology

The codes for String characters range from 0–255. The first 128 characters (0–127) of the character set correspond to the letters and symbols on a standard U.S. keyboard. These first 128 characters are the same as those defined by the ASCII character set. The second 128 characters (128–255) represent special characters, such as letters in international alphabets, accents, currency symbols, and fractions.

2.3.5 Variables: A variable is something that may change in value. A variable might be the number of words on different pages of this booklet, the air temperature each day, or the exam marks given to a class of school children.

2.3.6 Memo Data Type: Use the Memo data type if you need to store more than 255 characters. A Memo field can store up to 65,536 characters. If you want to store formatted text or long documents, you should create an OLE Object field instead of a Memo field.

2.4. KEY

The word "**key**" is used in the context of relational database design. In pre-relational databases (hierarchical, networked) and file systems (ISAM, VSAM, etc.) "key" often referred to the specific structure and components of a linked list, chain of pointers, or other physical locator outside of the data. It is thus natural, that today people often associate "key" with a RDBMS "index". We will explain what a key is and how it differs from an index.

A key is a set of one or more columns whose combined values are *unique* among all occurrences in a given table. **A key is the relational means of specifying uniqueness.**

There are only three types of relational keys.

2.4.1 Candidate Key: A **Candidate key** is any set of one or more columns whose combined values are unique among all occurrences (i.e., tuples or rows) Since a null value is not guaranteed to be unique, no component of a candidate key is allowed to be null.

There can be any number of candidate keys in a table. Relational pundits are not in agreement whether *zero* candidate keys are acceptable, since that would contradict the (debatable) requirement that there must be a primary key.

2.4.2 Primary Key: The **primary key** of any table is any candidate key of that table which the database designer arbitrarily designates as "primary". The primary key may be selected for convenience, comprehension, performance, or any other reasons. It is entirely proper to change the selection of primary key to another candidate key.

Enrollment	
student	class
PK	

Enrollment		
student	class	row id
		PK



Data Storage, Retrieval and Data Base Management Systems

2.4.3 Alternate Key: The **alternate keys** of any table are simply those candidate keys which are not currently selected as the primary key. An alternate key is a *function* of all candidate keys *minus* the primary key.

Let's look at this simple association or join table below which holds student class enrollment:

This is the default form, and often the only form taught or recommended for a join table. The primary key of each contribution table is inherited so that this table has a compound primary key. Since the Data Definition Language (DDL) will create a unique index on the two columns, the designer knows that each student-class pair will be unique; i.e., each student may enroll in each class only once.

Several years later a new Data Base Administrator (DBA) decides that it is inefficient to use two columns for the primary key where one would do. She adds a "row id" column and makes it the primary key by loading it with a system counter. This is fine as far as an identity for each row. But now nothing prevents a student from enrolling in the same class multiple times!

Enrollment			Enrollment		
student	class	row id	student	class	row id
CK		CK	AK		PK

This happened because the data model did not retain a candidate key property on the two original columns when the primary key was changed. Therefore the new DBA had no direct way of knowing (other than text notes somewhere) that these two columns must still remain unique, even though they are no longer part of the primary key.

Notice here how this could have been handled automatically by the model, if it had captured candidate keys in the first place and then generated alternate keys as a function of those candidates not in the primary key. The original two columns remain unique even after they are no longer primary.

2.4.4 Secondary Key: Secondary keys can be defined for each table to optimize the data access. They can refer to any column combination and they help to prevent sequential scans over the table. Like the primary key, the secondary key can consist of multiple columns.

A candidate key which is not selected as a primary key is known as Secondary Key.

2.4.5 Referential Integrity: A feature provided by relational database management systems (RDBMS's) that prevents users or applications from entering inconsistent data. Most RDBMS's have various referential integrity rules that you can apply when you create a relationship between two tables.



Information Technology

For example, suppose Table B has a *foreign key* that points to a field in Table A. Referential integrity would prevent you from adding a record to Table B that cannot be linked to Table A. In addition, the referential integrity rules might also specify that whenever you delete a record from Table A, any records in Table B that are linked to the deleted record will also be deleted. This is called *cascading delete*. Finally, the referential integrity rules could specify that whenever you modify the value of a linked field in Table A, all records in Table B that are linked to it will also be modified accordingly. This is called *cascading update*.

A

DEPTNO	DNAME	LOC
20	RESEARCH	DELHI
30	TRAINING	KANPUR

Foreign Key
(values in dependent table must match a value in unique key or primary key of referenced table)

B

EMPNO	ENAME	DESIG	DTOFJOI	SALARY	DEPTNO
6584	SUNIL	MANAGER	10-DEC-99	40000	20
7451	SANTA	PROFESSOR	21-MAR-95	50000	30
6680	JONES	DIRECTOR	02-APR-90	60000	20
7550	RAVI	MANAGE	10-DEC-02	80000	40

This row violates the referential constraint because "40" is not present in the referenced table's primary key; therefore, the row is not allowed in the table.

Fig. 2.4.5.1 illustrates example of Referential Integrity in the database

2.4.6 Index Fields: Index fields are used to store relevant information along with a document. The data input to an Index Field is used to find those documents when needed. The program provides up to twenty-five user-definable Index Fields in an Index Set.

An index field can be one of three types:

Drop-Down Look-Up List,

Standard,

Auto-Complete History List.



Data Storage, Retrieval and Data Base Management Systems

2.4.7 Currency Fields—The currency field accepts data in dollar form by default.

2.4.8 Date Fields – The date field accepts data entered in date format.

2.4.9 Integer Fields – The integer field accepts data as a whole number.

2.4.10 Text Fields – The text field accepts data as an alpha-numeric text string.

2.5 WHAT IS DATA PROCESSING

Data are a collection of facts - unorganized but able to be organized into useful information. A collection of sales orders, time sheets, and class registration cards are a few examples. Data are manipulated to produce output, such as bills and pay cheques. When this output can be used to help people make decisions, it is called *information*.

Processing is a series of actions or operations that convert inputs into outputs. When we speak of data processing, the input is data, and the output is useful information. Hence, data processing is defined as series of actions or operations that converts data into useful information. *The data processing system* is used to include the resources such as people, procedures, and devices that are used to accomplish the processing of data for producing desirable output.

2.5.1 Data storage Hierarchy: The basic building block of data is a *character*, which consists of letters (A, B, C ... Z), numeric digits (0, 1, 2 ... 9) or special characters (+, -, /, *, ., \$...) these characters are put together to form a *field* (also called a fact, data item, or data element) A **field** is a meaningful collection of related characters. It is the smallest logical data entity that is treated as a single unit in data processing.

For example, if we are processing employees data of a company, we may have an employee code field, an employee name field, an hours worked field, a hourly-pay-rate field, a tax-rate-deduction field, etc. Fields are normally grouped together to form a record. A **record**, then, is a collection of related fields that are treated as a single unit. An employee record would be a collection of fields of one employee. These fields would include the employee's code, name, hours-worked, pay-rate, tax-rate-deduction, and so forth. Records are then grouped to form a file. A **file** is a number of related records that are treated as a unit. For example, a collection of all employee records for one company would be an employee file.

It is customary to set up a master file of permanent (and, usually, the latest) data, and to use **transaction files** containing data of a temporary nature. For example, the master payroll file will contain not only all the permanent details about each employee, his name and code, pay-rate, income tax rate and so forth, but it will also include the current gross-pay-to-date total and the tax paid-to-date total. The transaction payroll file will contain details of hours



Information Technology

worked this week, normal and overtime, and, if piecework is involved, the quantity of goods made. When the payroll program is processed, both files will have to be consulted to generate this week's payslips, and the master file updated in readiness for the following week.

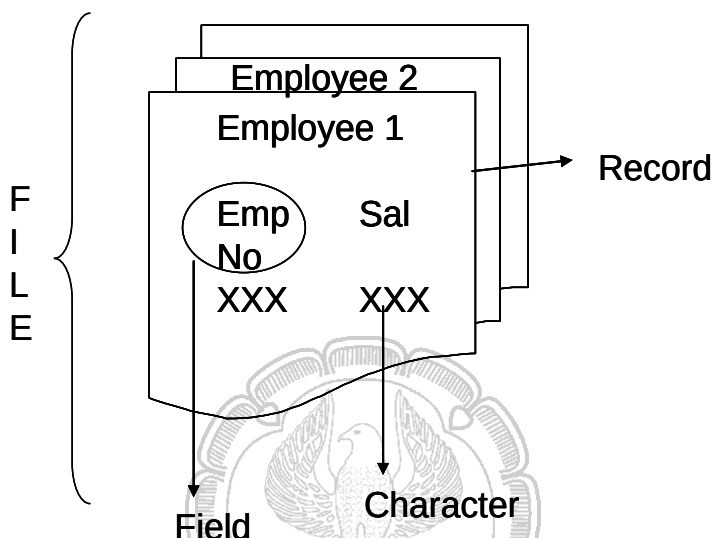


Fig 2.5.1 illustrates relationship between character, field, record, and file.

A *data base* is a collection of integrated and related master files. It is a collection of logically related data elements that may be structured in various ways to meet the multiple processing and retrieval needs of organizations and individuals. Characters, fields, records, files, and data bases form a hierarchy of data storage. Fig. 2.5.1 summarizes the data storage hierarchy used by computer based processing systems. Characters are combined to make a field, fields are combined to make a record, records are combined to make a file, and files are combined to make a data base.

2.6. FILE ORGANIZATIONS

System designers choose to organize, access, and process records and files in different ways depending on the type of application and the needs of users. The three commonly used file organizations used in business data processing applications are - sequential, direct and indexed sequential organizations. The selection of a particular file organization depends upon the type of application. The best organization to use in a given application is the one that happens to meet the user's needs in the most effective and economical manner. In making the choice for an application, designers must evaluate the distinct strengths and weaknesses of



Data Storage, Retrieval and Data Base Management Systems

each file organization. File organization requires the use of some key field or unique identifying value that is found in every record in the file. The key value must be unique for each record of the file because duplications would cause serious problems. In the payroll example, the employee code field may be used as the key field.

2.6.1 Serial File Organization: With serial file organization, records are arranged one after another, in no particular order- other than, the chronological order in which records are added to the file. Serial organization is commonly found with transaction data, where records are created in a file in the order in which transactions take place.

Records in a serially organized file are sometimes processed in the order in which they occur. For example, when such a file consists of daily purchase and payment transaction data, it is often used to update records in a master account file. Since transactions are in random order by key field, in order to perform this update, records must be accessed randomly from the master file.

Transaction data is not the only type of data found in serially organized files. In many businesses, customer account numbers are issued in a serial manner. In this scheme, a new customer is given the next highest account number that has not been issued and the data about the new customer (such as name, address, and phone number) are placed at the end of the existing customer account file.

2.6.2 Sequential files: In a sequential file, records are stored one after another in an ascending or descending order determined by the key field of the records. In payroll example, the records of the employee file may be organized sequentially by employee code sequence. Sequentially organized files that are processed by computer systems are normally stored on storage media such as magnetic tape, punched paper tape, punched cards, or magnetic disks. To access these records, the computer must read the file in sequence from the beginning. The first record is read and processed first, then the second record in the file sequence, and so on. To locate a particular record, the computer program must read in each record in sequence and compare its key field to the one that is needed. The retrieval search ends only when the desired key matches with the key field of the currently read record. On an average, about half the file has to be searched to retrieve the desired record from a sequential file.

Advantages of Sequential files

- Easy to organize, maintain, and understand.
- There is no overhead in address generation. Locating a particular record requires only the specification of the key field.
- Relatively inexpensive I/O media and devices can be used for the storage and processing of such files.



Information Technology

- It is the most efficient and economical file organization in case of applications in which there are a large number of file records to be updated at regularly scheduled intervals. That is, when the activity ratio (the ratio of the total number of records in transaction file and the total number of records in master file) is very high. Applications such as payroll processing, billing and statement preparation, and bank cheque processing meet these conditions.

Disadvantages of sequential Files

- It proves to be very inefficient and uneconomical for applications in which the activity ratio is very low.
- Since an entire sequential file may need to be read just to retrieve -and update few records, accumulation of transactions into batches is required before processing them.
- Transactions must be sorted and placed in sequence prior to processing.
- Timeliness of data in the file deteriorates while batches are being accumulated.
- Data redundancy is typically high since the same data may be stored in several files sequenced on different keys.

2.6.3 Direct Access File Organization: Direct file organization allows immediate direct access to individual records on the file. The most widely used direct access techniques are depicted in the chart below:-

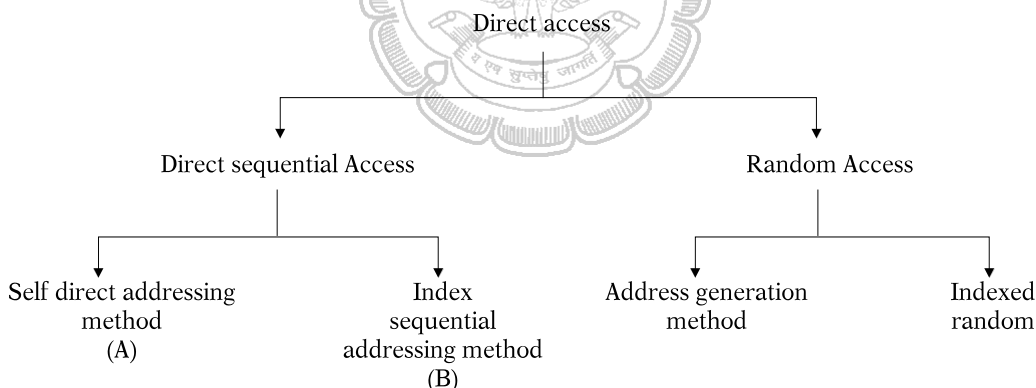


Fig. 2.6.3.1 Direct Access File Organization

The primary storage in a CPU truly provides for direct access. There are some devices outside the CPU which can provide the direct access feature; the direct access storage devices (DASD) have the capability of directly reaching any location. Although there are several types of direct storage devices including discs and other mass storage devices, discs are by far the



Data Storage, Retrieval and Data Base Management Systems

most widely used direct access storage devices. We will now describe the methods A-B mentioned above to show how data are stored on magnetic disks using these methods.

2.6.4 Direct Sequential Access Methods

(A) Self (Direct) Addressing: Under self direct addressing, a record key is used as its relative address. Therefore, we can compute the record's address directly from the record key and the physical address of the first record in the file.

Thus, this method is suitable for determining the bucket address of fixed length records in a sequential file, and in which the keys are from a complete or almost complete range of consecutive numbers. Suppose we want to store 1,60,000 payroll records cylinder-wise in the magnetic disc pack of 6 discs. The first cylinder carries the first 800 records; the 2nd cylinder carries the next 800 records, and so on. For periodic processing of the file, the read/write heads would move cylinder by cylinder in which the records have been sequentially arranged. For example, the ten faces in the first cylinder would carry the first 800 records as below:

$f_{1,1}$	1 to 80
$f_{1,2}$	81 to 160
:	
:	
:	
$f_{1,10}$	721 to 800



How do we have direct access then in such a file organization? There are a total of 16,000 buckets. Let the bucket address range from 10,001 to 26,000. And the keys of the records range from 1 to 1,60,000. We wish to know where record of the key 1,49,892 is to be found i.e., in which bucket it is stored. The following arithmetic computations would have to be performed towards this purpose.

1. Divide the wanted record's key by the number of records per bucket.
2. Add the first bucket number to the quotient to give the wanted record's bucket:
 $14989 + 10001 = 24990$.
3. The remainder (2) is the record's position within the bucket. The remainder 0 would indicate that it is the last record of the preceding bucket. Thus, if a manager wishes to know the qualification of a particular employee (say, no 149892) i.e., makes a random



Information Technology

inquiry, the above computations would be performed to derive the bucket number, command the read/write heads to move to that bucket and supply the wanted information.

But this method is highly impractical because files too have gaps in the keys and this would leave too many empty buckets *i.e.*, storage would not be compact.

The advantage of Self-addressing is that there is no need to store an index.

The disadvantages of Self-addressing are:

- (i) The records must be of fixed length.
- (ii) If some records are deleted their storage space remains empty.

(B) Indexed-Sequential File Organization: *The indexed sequential file organization or indexed sequential access method (ISAM)*, is a hybrid between sequential and direct access file organizations. The records within the file are stored sequentially but direct access to individual records is possible through an index. This index is analogous to a card catalog in a library.

Cylinder Index

Cylinder	Highest Record key in the Cylinder
1	84
2	250
3	398
4	479
5	590

Cylinder 1 Track index	
Track	Highest record key in the Track
1	15
2	40
3	55

Cylinder 2 Track index	
Track	Highest record key in the Track
1	94
2	110
3	175

Cylinder 3 Track index	
Track	Highest record key in the Track
1	280
2	301
3	330



Data Storage, Retrieval and Data Base Management Systems

4	75	4	225	4	365
5	84	5	250	5	398

Fig 2.6.4.1 illustrates a cylinder and track index for an ISAM file.

To locate a record, the cylinder index is searched to find the cylinder address, and then the track index for the cylinder is searched to locate the track address of the desired record. Using Fig . 2.6.4.1 to illustrate, we assume that the desired record has a key value of 225. The cylinder address is 2, since 225 is greater than 84 but less than 250. Then, we search the track index for cylinder 2 and find that 225 is greater than 175 and equal to 225, therefore, the track address is 4. With the cylinder address, control unit can then search through the records on track 4 within cylinder 2 to retrieve the desired records.

Advantages of indexed sequential files

- Permits the efficient and economical use of sequential processing techniques when the activity ratio is high.
- Permits direct access processing of records in a relatively efficient way when the activity ratio is low.

Disadvantages of indexed sequential files

- These files must be stored on a direct-access storage device. Hence, relatively expensive hardware and software resources are required.
- Access to records may be slower than direct files.
- Less efficient in the use of storage space than some other alternatives.

2.6.5 Random Access Organization: In this method, transactions can be processed in any order and written at any location through the stored file. The desired records can be directly accessed using randomizing procedure without accessing all other records in the file.

Randomizing Procedure is characterized by the fact that records are stored in such a way that there is no relationship between the keys of the adjacent records. The technique provides for converting the record key number to a physical location represented by a disk address through a computational procedure.

Advantages of Direct Files

- The access to, and retrieval of a record is quick and direct. Any record can be located and retrieved directly in a fraction of a second without the need for a sequential search of the file.



Information Technology

- Transactions need not be sorted and placed in sequence prior to processing.
- Accumulation of transactions into batches is not required before processing them. They may be processed as and when generated.
- It can also provide up-to-the minute information in response to inquiries from simultaneously usable online stations.
- If required, it is also possible to process direct file records sequentially in a record key sequence.
- A direct file organization is most suitable for interactive online applications such as airline or railway reservation systems, teller facility in banking applications, etc.

Disadvantages of direct files

- Address generation overhead is involved for accessing each record due to hashing function.
- May be less efficient in the use of storage space than sequentially organized files.
- Special security measures are necessary for online direct files that are accessible from several stations.

2.6.6 THE BEST FILE ORGANIZATION

Several factors must be considered in determining the best file organization for a particular application. These factors are file volatility, file activity, file size, and file interrogation requirements.

File volatility: It refers to the number of additions and deletions to the file in a given period of time. The payroll file for a construction company where the employee roster is constantly changing is a highly volatile file. An ISAM file would not be a good choice in this situation, since additions would have to be placed in the overflow area and constant reorganization of the file would have to occur. Other direct access methods would be better. Perhaps even sequential file organization would be appropriate if there were no interrogation requirements.

File activity: It is the proportion of master file records that are actually used or accessed in a given processing run. At one extreme is the real-time file where each transaction is processed immediately and hence at a time, only one master record is accessed. This situation obviously requires a direct access method. At the other extreme is a file, such as a payroll master file, where almost every record is accessed when the weekly payroll is processed. There, a sequentially ordered master file would be more efficient.

File interrogation: It refers to the retrieval of information from a file. If the retrieval of individual records must be fast to support a real-time operation such as airline reservation



Data Storage, Retrieval and Data Base Management Systems

then some kind of direct organization is required. If, on the other hand, requirements for data can be delayed, then all the individual requests or information can be batched and run in a single processing run with a sequential file organization.

Large files that require many individual references to records with immediate response must be organized under some type of direct access method. On the other hand, with small files, it may be more efficient to search the entire file sequentially or, with a more efficient binary search, to find an individual record than to maintain complex indexes or complex direct addressing scheme.

2.7 DATA BASE MANAGEMENT SYSTEMS

Traditional sequential and random files are designed to meet specific information and data processing requirements of a particular department such as accounting, sales, or purchasing etc. Different files are created to support these functions, but many of the fields on each of these files are common. For example, each of these functional areas needs to maintain customer data such as customer name, address and the person to be contracted at the customer location etc. In a traditional file environment, when information relating to any of the fields change, each relevant file must be updated separately.

Through the early 1980s, most information systems were implemented in an environment with a single functional objective (such as accounts receivable, purchase accounting, payroll etc.) in mind. The integration of information systems was not a priority. Today companies are using database management systems software (DBMS) as a tool to integrate information flow within an organization.

2.7.1 What is a DBMS? : A DBMS is a set of software programs that controls the organization, storage, management, and retrieval of data in a database. A data base is a repository for related collection of data. For example, an address book can be a data base where the names, address and telephone numbers of friends and business contacts are stored. A company data base might contain information about customers, vendors, employees, sales and inventory. Each piece of information can be added to a data base and extracted later in a meaningful way. DBMS is the program (or collection of programs) that allows users (and other programs) to access and work with a database.

Database programs for personal computers come in many shapes, sizes and variations. Examples of DBMS are tabulated below:



Information Technology

Database	Manufacturer	Computer Type
Access	Microsoft Corporation	Personal computer, server, PDA
Adabas	Software AG	Midrange server, mainframe
D³	Raining Data	Personal computer, midrange server
DB2	IBM Corporation	Personal computer, midrange server, mainframe
Essbase	Hyperion Solutions Corporation	Personal computer, server
FastObjects	FastObjects Inc.	Personal computer, midrange server
GemFire	GemStone Systems, Inc.	Midrange server
Informix	IBM Corporation	Personal computer, midrange server, mainframe
Ingres	Computer Associates International, Inc.	Personal computer, midrange server, mainframe
InterBase	Borland Software Corporation	Personal computer, server
JDataStore	Borland Software Corporation	Personal computer, server
KE Texpress	KE Software, Inc.	Personal computer, server
MySQL	MySQL AB	Personal computer, midrange server
ObjectStore	Progress Software Corporation	Personal computer, midrange server
Oracle	Oracle Corporation	Personal computer, midrange server, mainframe, PDA
SQL Server	Microsoft Corporation	Server, personal computer, PDA
Sybase	Sybase Inc.	Personal computer, midrange server, PDA
Versant	Versant Corporation	Personal computer, midrange server
Visual FoxPro	Microsoft Corporation	Personal computer, server

Fig. 2.7.1.1 Examples of DBMS Programs

2.7.2 An Example of File Processing Approach: An example to illustrate why organizations started using database processing as an alternative to traditional file processing:

A firm might have a customer credit file containing data such as:

- ◆ Customer number
- ◆ Customer name and address
- ◆ Credit code
- ◆ Credit limit

Another file, called a customer master file, contains:

- ◆ Customer number
- ◆ Customer name and address
- ◆ Sales region number
- ◆ Salesperson number
- ◆ Customer class
- ◆ Shipping code
- ◆ Year to date sales this year
- ◆ Year to date sales last year



Data Storage, Retrieval and Data Base Management Systems

A third file, for accounts receivable, contains:

- ◆ Customer number
- ◆ Customer name and address
- ◆ First invoice data
 - Invoice number
 - Invoice date
 - Invoice amount
- ◆ Second invoice data
 - Invoice number
 - Invoice date
 - Invoice amount
- ◆ n^{th} invoice data
 - Invoice number
 - Invoice date
 - Invoice amount

Each of these files has one or more purposes. The customer credit file is used for approving customer orders, the customer master file is used for invoicing customers, and the accounts receivable file represents the money which is to be recovered from customers on account of sales by the firm. All are master files.

Some redundancies are found in the data elements contained within the files. All three files include customer number, and customer name and address. This redundancy is necessary since each file is designed to provide all of the data needed by a particular program.

Let us assume that the sales manager wants a report showing the amount of receivables by a salesperson. The firm's customers have not been paying their bills promptly, and the sales manager wants to know which sales persons have neglected to follow up on past due receivables. He wants the report to include the data listed in Table 3. It can be seen that this special report will require data from four files. A salesperson master file is needed to provide the sales person name.

**Information Technology**

Report Data	Customer Credit File	Customer Master File	Accounts Receivable File	Salesperson Master File
Salesperson number		X		
Sales person Name				X
Customer Data				
Customer number		X		
Customer name		X		
Credit code	X			
Year to date sales this year		X		
Total accounts receivable			X	

Table 3 illustrates example of integration of report data from several files

The report will list each customer by salesperson, following the process illustrated in Fig. 2.7.2.1. In step 1 a program selects data from the three customer files that are maintained in customer number sequence. An intermediate file is created with the selected data (all the data elements listed in Table 3 except salesperson name) this intermediate file is stored into salesperson sequence in step 2. A sort is necessary since the salesperson master file is maintained in salesperson sequence. A second intermediate file is created and used with the salesperson master file to prepare the report in step 3. The programs for step 1 and step 3 would have to be specially written to satisfy this request.

Similarly a manager may require ad hoc reports for management information. For example, a manager might request a report showing sales for sales person 23. Assume that the firm assigns certain customers in a territory to a salesperson and that a customer file contains a record for each customer. The task is to select records for salesperson 23 only and print data on the report. Since the customer file is in sequence by customer, each record will have to be examined to determine if the sales person field contains a 23. This could be a time consuming process.



Data Storage, Retrieval and Data Base Management Systems

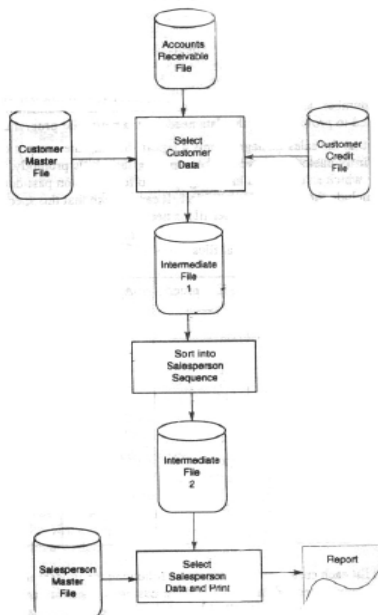


Fig. 2.7.2.1 illustrates example of file processing

2.7.3 Management Problems of File Processing: For many years, information systems had a file processing orientation, as illustrated in the previous example. Data needed for each user application was stored in independent data files. Processing consisted of using separate computer programs that updated these independent data files and used them to produce the documents and reports required by each separate user application. This file processing approach is still being used, but it has several problems that limit its efficiency and effectiveness for end user applications.

1. Data Duplication: Independent data files include a lot of duplicated data; the same data (such as a customer's name and address) is recorded and stored in several files. This data redundancy causes problems when data has to be updated, since separate file maintenance programs have to be developed and coordinated to ensure that each file is properly updated. Of course, this proves difficult in practice, so a lot of inconsistency occurs among data stored in separate files.

2. Lack of Data Integration: Having data in independent files makes it difficult to provide end users with information for ad hoc requests that require accessing data stored in several



Information Technology

different files. Special computer programs have to be written to retrieve data from each independent file. This is difficult, time consuming, and expensive for the organizations.

3. Data Dependence: In file processing systems, major components of a system i.e., the organization of files, their physical locations on storage, hardware and the application software used to access those files depend on one another in significant ways. For example, application programs typically contain references to the specific format of the data stored in the various files they use. Thus, if changes are made in the format and structure of data and records in a file, changes have to be made in all the programs that use this file. This program maintenance effort is a major burden of file processing systems. It is difficult to do it properly, and it results in a lot of inconsistency in the data files.

4. Other Problems: In file processing systems, data elements such as stock numbers and customer addresses are generally defined differently by different end users and applications. This causes serious inconsistency in the development of programs which access such data. In addition, the integrity (i.e. the accuracy and completeness) of the data is suspected because there is no control over their use and maintenance by authorized end users.

2.7.4 The Database Management Solution: The concepts of databases and database management were, therefore, developed to solve the problems of file processing systems. A database is an integrated collection of logically related records and files. It consolidates records previously stored in independent files so that it serves as a common pool of data to be accessed by many different application programs.

Benefits of DBMS Solution are:

- (a) Reduced data redundancy and inconsistency.
- (b) Enhanced data integrity
- (c) Logical and physical data independence
- (d) Application data independence
- (e) Reduced complexity of the organization's Information System environment.

2.8 WHAT IS A DATABASE?

Regardless of its file organization, a data base system includes several components that collectively give it certain distinct, specific characteristics. The following is a precise definition of data base as given by "G.M.Scott".

A data base is a computer file system that uses a particular *file organization* to facilitate *rapid updating of individual records, simultaneous updating of related records, easy access to all*



Data Storage, Retrieval and Data Base Management Systems

records, by all applications programs, and rapid access to all stored data which must be brought together for a particular *routine report or inquiry* or a *special purpose report or inquiry*.

Each of the italicized phrases in the preceding definition has a special meaning that helps define a database. "*File organization*" indicates that the database has one of the three file structures (discussed in the next section) that enable programs to establish associations between the records in the database.

A database facilities "*rapid updating of individual records*" and "*simultaneously updating of related records*" that is a data base permits the entry of an individual transaction to update all records affected by that transaction simultaneously. For example, consider a Rs.100,000 credit sale. In a data base system, the following accounts, along with others could be updated simultaneously with the input of one transaction.

- ◆ Sales record
- ◆ Salesperson's commissions record
- ◆ Division sales record
- ◆ Inventory item record
- ◆ Accounts receivable customer record
- ◆ Cost of sales of individual item record

If transactions are entered as they occur, records that are simultaneously updated are continuously up to date for managerial inquiry purposes. Simultaneous updating also means that the records have consistent contents. For example, the total of the sales record would be consistent with the salesperson's commission's record because the latter is based on the former and both are updated at the same time.

"*Easy access to all records by all applications programs*" means that the standard data definitions and record formats permit, for example, a payroll applications program to access the employee number, and data about them from the personnel section of the data base. It also implies that work force planning programs can access pay rates from the payroll section and employees skills from the personnel section of the database. Without a database, each application program would be able to access data only from its own file.

With respect to "*rapid access*" to all stored data needed of a "*routine report or inquiry*", routine reports can be provided quickly after the end of the accounting period and often whenever requested during the period if the processing of transactions is kept up to date. This is possible because transfer file processing is not required at the end of the period and because



Information Technology

data summarization of reports can be fully automated within a database. In other words, little period end processing is required. Similarly, inquiries can be routinely made into the files, for example, to see whether a particular product is available for immediately shipment.

Rapid access with respect to a “*special purpose report or inquiry*” means that records are kept continuously up to date for unanticipated inquiries into the files by managers and that the structure of the data base files facilitates the rapid development of special programs to prepare reports about unanticipated problems.

2.8.1 Architecture of a Database: It follows a three level architecture –

- (i) External or user view in different ways to say, the chairman, or the operation manager or the data entry operator,
- (ii) Conceptual or global view,
- (iii) Physical or internal view.

External or user view encircles the following –

- (i) It is at the highest level of the database abstraction,
- (ii) It includes only those portion of database or application programs which is of concern to the users,
- (iii) It is described by means of a scheme, called the external schema,
- (iv) It is defined by the users or written by the programmers.

For example an external view in its Logical Record 1 may indicate employee name, employee address and in its Logical Record 2 may indicate employee name, employee address and employee code and employee salary.

Global or conceptual view, which is viewed by the Data Base Administrator, encompasses the following –

- (i) All database entities and relationships among them are included,
- (ii) Single view represents the entire database,
- (iii) It is defined by the conceptual schema,
- (iv) It describes all records, relationships and constraints or boundaries,
- (v) Data description to render it independent of the physical representation.

For example a conceptual view may define employee code as a string of characters, employee address also as a string, employee code as a key and employee salary as an integer.



Data Storage, Retrieval and Data Base Management Systems

The physical or internal view contains the following –

- (i) It is at the lowest level of database abstraction,
- (ii) It is closest to the physical storage method,
- (iii) It indicates how data will be stored,
- (iv) It describes data structure,
- (v) It describes access methods,
- (vi) It is expressed by internal schema.

The internal view instead, may define employee name is comprised of 30 characters, employee address is also comprised of 30 characters, employee code is comprised of 5 characters and employee salary is comprised of 12 numbers.

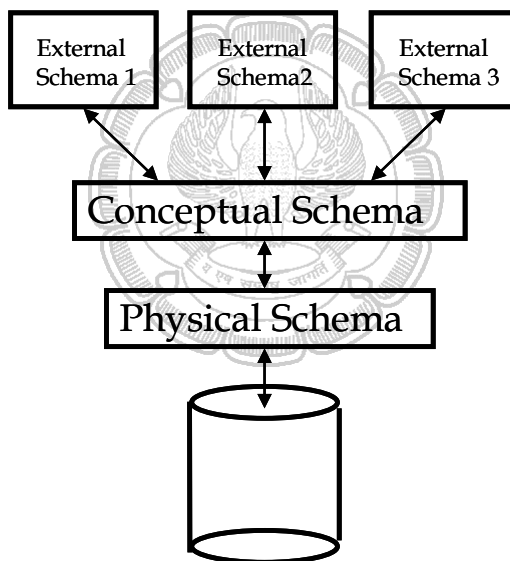


Fig. 2.8.1.1 illustrates three level architecture of a database

The first step in moving from ordinary file management to a data base system is to separate all data definitions from the applications programs and to consolidate them into a separate entity called a **schema**, as illustrated in the fig. 2.8.1.1. In addition to data definition, the schema also includes an indication of the logical relationships between various components of the data base.



Information Technology

The schema then becomes a component of the overall data base itself. From the schema, the installation can generate dictionaries containing a complete description of the data base. These will, in turn, be used by systems analysts in defining new applications.

Database systems have several schemas, partitioned according to the levels of abstraction that we have discussed. At the lowest level is the **physical schema**; at the intermediate level is the **logical schema**; and at the highest level is a **subschema**.

2.8.2 Data independence:

- (i) In a database an ability to modify a schema definition at one level is done without affecting a schema in the next higher level,
- (ii) It facilitates logical data independence,
- (iii) It assures physical data independence.

2.8.3 Classification of DBMS users:

- (i) Naive users who are not aware of the presence of the database system supporting the usage,
- (ii) Online users who may communicate with database either directly through online terminal or indirectly through user interface or application programs. Usually they acquire at least some skill and experience in communicating with the database,
- (iii) Application programmers who are responsible for developing the application programs and user interfaces,
- (iv) Data Base Administrator who can execute centralized control and is responsible for maintaining the database.

We will now discuss how data is stored in a database.

2.8.4 File pointers: File pointers establish linkage between records and are a basic part of the file organization of all the database models except the relational model. A pointer is placed in the last field of a record, if more than one pointer is used, then in the last fields. A pointer is the address of another; related record that is “pointed to” and the pointer directs the computer system to that related record. File pointers are used with many database organizations.

Linked List: A linked List is a group of data records arranged in an order, which is based on embedded pointers. An embedded pointer is a special data field that links one record to another by referring to the other record. The field is embedded in the first record, i.e. it is a data element within the record.



Data Storage, Retrieval and Data Base Management Systems

Linked lists often have a head, which is a pointer to the first record. It has a tail, which points to the last record. One can start at the head and follow the list to the tail, or one can start in the middle and follow the list to the tail. The user cannot, however, start in the middle and go back to the head. In another words, the linked list is a one way street.

Embedded pointer

Customer	Sales person number	Sales person link
22504		
23694	23	25410
24782		
25409		
25410	23	30102
26713		
28914		
30004		
30102	23	30111
30111	23	*
30417		
31715		

Fig. 2.8.4.1 illustrates example of a pointer application.

Fig. 2.8.4.1 shows a linked list of customer records. Each row is a record (only relevant fields are shown) the records are arranged sequentially using customer number as the key. Each record includes a data element, which identifies assigned salesperson. In the right most field of the record there is a pointer (a link) that chains together all customer records for a particular salesperson in the example it is salesperson 23. It can be assumed that customer 23694 is at the head of the list. The pointer links this record to a record for customer 25410 and so on until the tail for customer 30111 is encountered. The asterisk in the link field indicates the tail of the list.

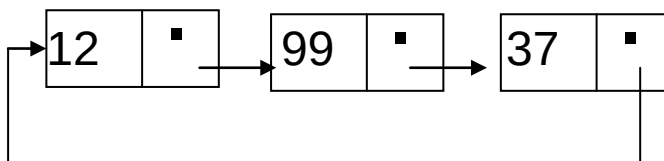


Fig. 2.8.4.2 Example of a linked list

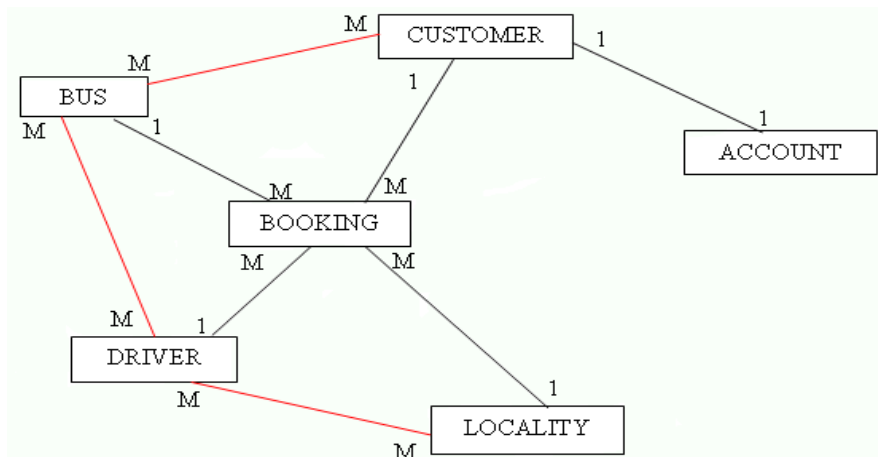
This chaining feature is very powerful. The application program can initiate a search at the beginning of the file looking for the first customer assigned to salesperson 23. When that record is found, the salesperson links enable the program to follow the chain and process records only for salesperson 23. It is a more convenient method than searching through the entire file.

2.8.5 Record relationship in Database: Organizing a large database logically into records and identifying the relationships among those records are complex and time-consuming tasks. There are large number of different records that are likely to be part of a corporate database and the numerous data elements constituting those records. Further, there can be several general types of record relationships that can be represented in a database. These are briefly discussed below:

1. *One –to-one relationship*, as in a single parent record to a single child record or as in a husband record and wife record in a monogamous society [see. fig. 2.8.5.1].
2. *One-to-many relationships*, as in a single parent record to two or more child records – for example, a teacher who teaches three single-section courses [see fig. 2.8.5.1].
3. *Many-to-one relationships*, as in two or more parent records to a single child record-for example, when three administrators in a small town share one minister [see fig. 2.8.5.1].
4. *Many-to-many relationships*, as in two or more parent records to two or more child records – for example, when two or more students are enrolled in two or more courses [See fig. 2.8.5.1].



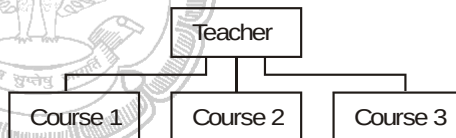
Data Storage, Retrieval and Data Base Management Systems



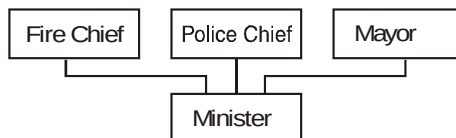
(a) One-to-one relationship



(b) One-to-many relationship



(c) Many-to-one relationship



(d) Many-to-many relationship

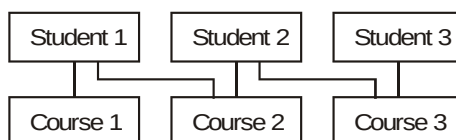


Fig. 2.8.5.1 Illustrates example of relationship in the database



2.9 DATABASE STRUCTURES

Three traditional approaches have been implemented commercially to organize records and their relationships logically. These logical organizational approaches are known as database structures. The three traditional database structures are the

1. Hierarchical database structure
2. Network database structure
3. Relational database structure

These models differ in the manner in which data elements (fields) can be logically related and accessed. Hierarchical models are often considered to be the most restrictive and relational models are the most flexible.

2.9.1 Hierarchical Database Structure: In a hierarchical database structure, records are logically organized into a hierarchy of relationships. A hierarchically structured database is arranged logically in an inverted tree pattern. For example, an equipment database, diagrammed in Fig. 2.9.1.1 may have building records, room records, equipment records, and repair records. The database structure reflects the fact that repairs are made to equipment located in rooms that are part of buildings.

All records in hierarchy are called **nodes**. Each node is related to the others in a parent-child relationship. Each parent record may have one or more child records, but no child record may have more than one parent record. Thus, the hierarchical data structure implements one-to-one and one-to-many relationships.

The top parent record in the hierarchy is called the *root record*. In this example, building records are the root to any sequence of room, equipment, and repair records. Entrance to this hierarchy by the database management system is made through the root record i.e., building. Records that “own” other records are called *parent records*. For example, room records are the parents of equipment records. Room records are also children of the parent record, building. There can be many levels of node records in a database.



Data Storage, Retrieval and Data Base Management Systems

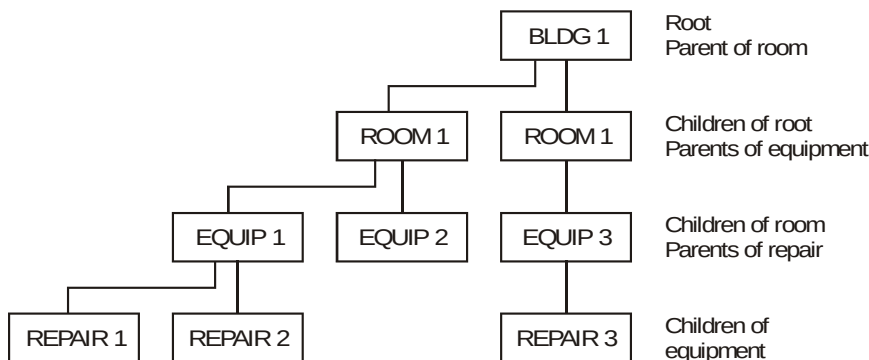


Fig. 2.9.1.1 Hierarchal database structures

Features of Hierarchical Database

- ◆ hierarchically structured database are less flexible than other database structures because the hierarchy of records must be determined and implemented before a search can be conducted. In other words, the relationships between records are relatively fixed by the structure.
- ◆ Ad hoc queries made by managers that require different relationships other than that are already implemented in the database may be difficult or time consuming to accomplish. For example, a manager may wish to identify vendors of equipment with a high frequency of repair. If the equipment record contains the name of the original vendor, such a query could be performed fairly directly. However, data describing the original vendor may be contained in a record that is a part of another hierarchy. As a result, there may not be any established relationship between vendor records and repair records. Providing reports based on this relationship in a large database is not a minor task and is not likely to be undertaken by the data processing staff for a one-time management query.
- ◆ Managerial use of query language to solve the problem may require multiple searches and prove to very time consuming. Thus, analysis and planning activities, which frequently involve ad hoc management queries of the database, may not be supported as effectively by a hierarchical DBMS as they are by other database structures.
- ◆ on the plus side, a hierarchical database management system usually processes structured, day-to-day operational data rapidly. In fact, the hierarchy of records is usually specifically organized to maximize the speed with which large batch operations such as payroll or sales invoices are processed.



Information Technology

- ◆ Any group of records with a natural, hierarchical relationship to one another fits nicely within the structure. However, many records have relationships that are not hierarchical. For example, many records relationships require that the logical data structure permit a child record to have more than one parent record. The query to isolate vendors of equipment with extensive repairs might be completed more easily if the equipment records were the children of both the room records and the vendor records
- ◆ Though a hierarchical database structure does not permit such a structure conceptually, a commercial hierarchical database management system must have ways to cope with these relationships. Unfortunately, they may not always be easy to implement.

2.9.2 Network Database Structure: A network database structure views all records in sets. Each set is composed of an owner record and one or more member records. This is analogous to the hierarchy's parent-children relationship. Thus, the network model implements the one-to-one and the one-to-many record structures. However, unlike the hierarchical mode, the network model also permits a record to be a member of more than one set at one time. The network model would permit the equipment records to be the children of both the room records and the vendor records. This feature allows the network model to implement the many-to-one and the many-to-many relationship types.

For example, suppose that in our database, it is decided to have the following records: repair vendor records for the companies that repair the equipment, equipment records for the various machines we have, and repair invoice records for the repair bills for the equipment. Suppose four repair vendors have completed repairs on equipment items 1,2,3,4,5, 7 and 8. These records might be logically organized into the sets shown in Fig. 2.9.2.1.

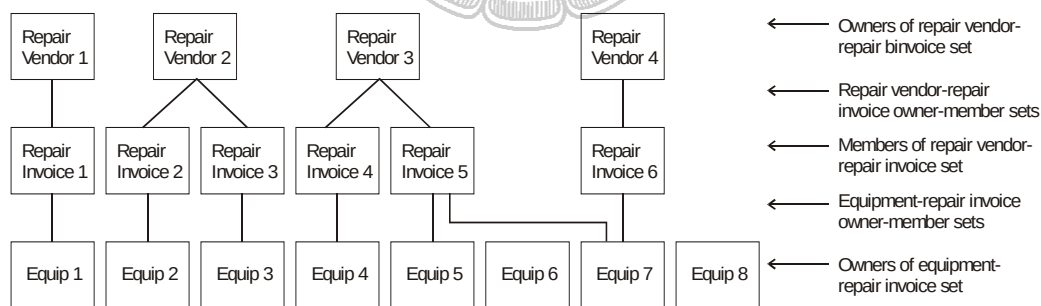


Fig.. 2.9.2.1 Example of Network Database Structure



Data Storage, Retrieval and Data Base Management Systems

Notice these relationships in the above:

1. Repair Vendor 1 record is the owner of the Repair Invoice 1 record. This is a one-to-one relationship.
2. Repair Vendor 2 record is the owner of the Repair Invoice 2 and 3 records. This is a one-to-many relationship.
3. Repair Vendor 3 record is the owner of Repair Invoice 4 and 5 records, and the Equipment 7 record owns both the Repair Invoice 5 and 6 records because it was fixed twice by different vendors. Because many equipment records can own many Repair Invoice records, these database records represent a many-to-many relationship.
4. Equipment 6 record does not own any records at this time because it is not required to be fixed yet.
5. Equipment 7 and 8 own Repair Invoice 6 because the repairs to both machines were listed on the same invoice by Repair Vendor 4. This illustrates the many-to-one relationship.

Thus, all the repair records are members of more than one owner-member set: the repair vendor-repair invoice set and the equipment-repair invoice set. The network model allows us to represent one-to-one, one-to-many and many-to-many relationships. The network model also allows us to create owner records without member records. Thus, we can create and store a record about a new piece of equipment even though no repairs have been made on the equipment yet.

Unlike hierarchical data structures that require specific entrance points to find records in a hierarchy, network data structures can be entered and traversed more flexibly. Access to repair records, for example, may be made through either the equipment record or the repair vendor record. However, like the hierarchical data model, the network model requires that record relationships be established in advance because these relationships are physically implemented by the DBMS when allocating storage space on disk. The requirement of established sets means that record processing for regular reports is likely to be swift. Record relationships are usually structured to fit the processing needs of large batch reports. However, ad hoc requests for data, requiring record relationships not established in the data model, may not be very swift at all, and in some cases, they may not be possible.



Information Technology

2.9.3 Relational Database Model: A third database structure is the relational database mode. Both the hierarchical and network data structures require explicit relationships, or links, between records in the database. Both structures also require that data be processed one record at a time. The relational database structure departs from both these requirements.

A relational database is structured into a series of two-dimensional tables. Because many managers often work with tabular financial data, it is easy for most of them to understand the structure used in a relational database. For example, consider the repair vendor relationship. The repair vendor table consists of the repair vendor master records, which contain the repair shown, in fig. 2.9.2.1 as network structure. This can be represented in a tabular form as shown in fig. 2.9.3.1. The repair vendor table consists of the repair vendor master records, which contain the repair vendor number, name and address. The table itself is really a file. Each row in the table is really a record and each column represents one type of data element.

Column 1	Column 2	Column 3
Repair Vendor Number	Repair Vendor Name	Repair Vendor Address
43623	Telo, Inc.	15 Parliament Street
42890	A-Repair Company	25 G.B.Road
43118	Beeline, Ltd.	498 Old Street
43079	Aspen, Inc.	12 Rouse Avenue
43920	Calso, Inc.	5 Janpath Road

Fig. 2.9.3.1 Repair Vendors Records

A similar table for equipment could look like the one in Fig 2.9.3.2. The table contains the records for each piece of equipment in the firm. Each record also contains the number of the repair vendor who has a contract to repair that piece of equipment.

Column 1	Column 2	Column 3	Column 4
Equipment Number	Equipment Name	Date Purchased	Repair Vendor No.
10893	Typewriter	12/02/1999	43623
49178	Microcomputer	01/31/2000	43920
10719	Telephone	03/12/2000	43079



Data Storage, Retrieval and Data Base Management Systems

18572	Copier	11/06/1998	43890
60875	Calculator	08/01/1997	43118

Fig. 2.9.3.2 Equipment Records

If the manager wished to create a report, showing the names of each repair vendor and the pieces of equipment that each vendor repairs, he could combine both tables into a third table. The manager might joint the two tables with a query statement such as this: JOIN REPAIR VENDOR AND EQUIPMENT ON REPAIR VENDOR NUMBER.

This would create a new table with six columns: Repair Vendor Number, Repair Vendor Name, Repair Vendor Address, Equipment Number, Equipment Name and Date Purchases. Now the manager could print out only the columns for vendor name and equipment name. A example of a report is shown in fig 2.9.3.3.

Equipment Repairs 1999	
Repair Vendor	Equipment
Modern Insurance	Telephone
B-line Ltd.	Calculator
Telco India	Type writer

Fig 2.9.3.3 Example of a report

The manager might also produce a report by selecting from both tables only the rows for specific equipment types or for equipment purchased in specific years. The important things to notice are that the relationships, or links, do not need to be specified in advance and that whole tables or files are manipulated.

Relational databases allow the manager flexibility in conducting database queries and creating reports. Queries can be made and new tables created using all or part of the data from one or more tables. The links between data elements in a relational database do not need to be made explicit at the time the database is created since new links can be structured at any time. The relational database structure is more flexible than hierarchical or network database structures and provides the manager with a rich opportunity for ad hoc reports and queries. However, because they do not specify the relationships among data elements in advance,



Information Technology

relational databases do not process large batch applications with the speed of hierarchical or network databases.

Many relational database management system products are available. For example, Oracle and IBM offer commercial relational database management systems, Oracle and DB2 respectively.

2.10 DATABASE COMPONENTS

- (i) **Data Definition Language (DDL)** that defines the conceptual schema providing a link between the logical (the way the user views the data) and physical (the way in which the data is stored physically) structures of the database. As discussed earlier, the logical structure of a database is a schema. A sub schema is the way a specific application views the data from the database.

Following are the functions of Data Definition Language (DDL) –

- (a) They define the physical characteristics of each record, field in the record, field's data type, field's length, field's logical name and also specify relationships among the records,
- (b) They describe the schema and subschema,
- (c) They indicate the keys of the record,
- (d) They provide means for associating related records or fields,
- (e) They provide for data security measures,
- (f) They provide for logical and physical data independence.

- (ii) **Data Manipulation Language (DML):** –

- (a) They provide the data manipulation techniques like deletion, modification, insertion, replacement, retrieval, sorting and display of data or records,
- (b) They facilitate use of relationships between the records,
- (c) They enable the user and application program to be independent of the physical data structures and database structures maintenance by allowing to process data on a logical and symbolic basis rather than on a physical location basis,
- (d) They provide for independence of programming languages by supporting several high-level procedural languages like, COBOL, PL/1 and C++.



Data Storage, Retrieval and Data Base Management Systems

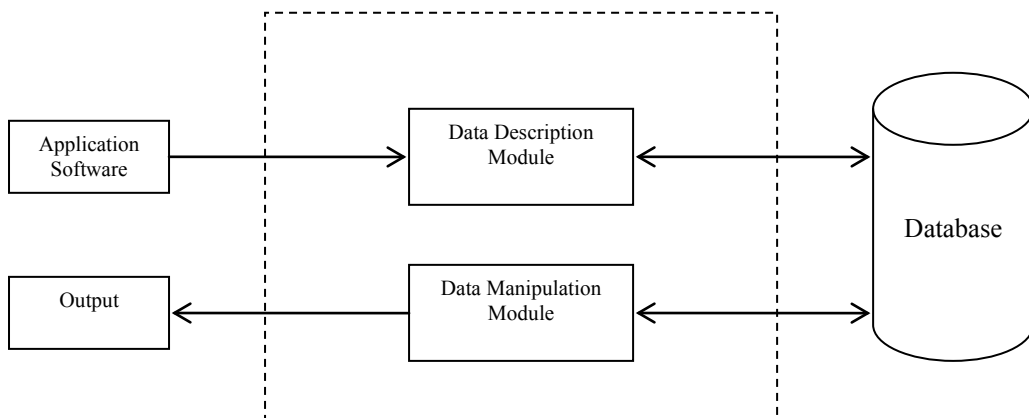


Fig 2.10.1 Data Base Management Systems Components

2.11 STRUCTURE OF DBMS: –

(i) DDL Compiler –

- It converts data definition statements into a set of tables,
- Tables contain meta data (data about the data) concerning the database,
- It gives rise to a format that can be used by other components of database.

(ii) Data Manager –

- It is the central software component,
- It is referred to as the database control system,
- It converts operations in users' queries to physical file system.

(iii) File Manager –

- It is responsible for file structure,
- It is responsible for managing the space,
- It is responsible for locating block containing required record,
- It is responsible for requesting block from disk manager,
- It is responsible for transmitting required record to data manager.



Information Technology

(iv) **Disk Manager –**

- a. It is a part of the Operating System,
- b. It carries out all physical input / output operations,
- c. It transfers block / page requested by file manager.

(v) **Query Manager –**

- a. It interprets user's online query,
- b. It converts to an efficient series of operations,
- c. In a form it is capable of being sent to data manager,
- d. It uses data dictionary to find structure of relevant portion of database,
- e. It uses information to modify query,
- f. It prepares an optimal plan to access database for efficient data retrieval.

(vi) **Data Dictionary –**

- a. It maintains information pertaining to structure and usage of data and meta data,
- b. It is consulted by the database users to learn what each piece of data and various synonyms of data field means.

Data base administrator: As mentioned earlier data base systems are typically installed and coordinated by an individual called the **data base administrator**. He has the overall authority to establish and control data definitions and standards. He is responsible for determining the relationships among data elements, and for designing the data base security system to guard against unauthorised use. He also trains and assists applications programmers in the use of data base. A data dictionary is developed and used in a data base to document and maintain the data definitions.

To design the database, the data base administrator must have a discussion with users to determine their data requirement. He can then decide the schedule and accuracy requirements, the way and frequency of data access, search strategies, physical storage requirements of data, level of security needed and the response time requirements. He may also identify the source of data and the person responsible for originating and updating of data. The database administrator then converts these requirements into a physical design that specifies hardware resources required for the purpose.

Defining the contents of the data base is an important part of data base creation and maintenance. The process of describing formats, relationships among various data elements



Data Storage, Retrieval and Data Base Management Systems

and their usage is called data definition and the DBA uses data definition language (DDL) for this purpose.

Maintaining standards and controlling access to data base are two other important functions that are handled by the DBA using DDL. The DBA specifies various rules which must be adhered to while describing data for a database. Data description not meeting these rules are rejected and not placed in the data dictionary. Invalid data values entered by users are also rejected. The DBA uses access controls to allow only specified users to access certain paths into the data base and thus prevent unauthorized access. For example, in an airline reservation system, an airline agent should be prevented from offering an expired rate to a passenger.

The DBA also prepares documentation which includes recording the procedures, standards guidelines and data descriptions necessary for the efficient and continued use of the data base environment. Documentation should be helpful to end users, application programmers, operating staff and data administration personnel. The DBA also educates these personnel about their duties.

It is also a duty of the DBA to ensure that the operating staff performs its database processing related responsibilities which include loading the database, following maintenance and security procedures, taking backups, scheduling the database for use and following restart and recovery procedures after some hardware or software failure, in a proper way.

The DBA also monitors the data base environments. He ensures that the standards for database performance are being met and the accuracy, integrity and security of data is being maintained. He also sets up procedures for identifying and correcting violation of standards, documents and corrects errors. This is accomplished by carrying out a periodic audit of the database environment.

The DBA is also responsible for incorporating any enhancements into the database environment which may include new utility programs or new systems releases, and changes into internal procedures for using data base etc.

2.12 TYPES OF DATABASES

The growth of distributed processing, end user computing, decision support and executive information systems has caused the development of several types of databases. Fig 2.12.1 illustrates six of the main databases that may be found in computer using organizations.

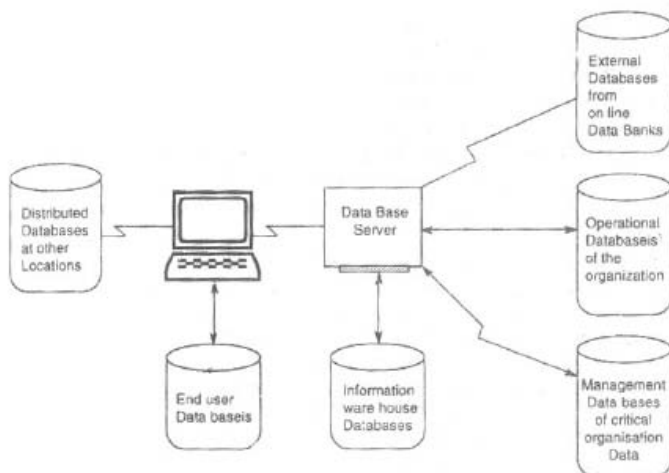


Fig 2.12.1 illustrates six of the main databases

Operational databases: These databases store detailed data needed to support the operations of the entire organization. They are also called subject area databases (SADB), transaction databases, and production databases. Examples are a customer database, personnel database, inventory database, and other databases containing data generated by business operations.

Management Database: These databases store data and information extracted from selected operational and external database. They consist of summarized data and information most needed by the organization's managers and other end users. Management databases are also called information databases. These are the databases accessed by executive end-users as part of decision support systems and executive information systems to support managerial decision making.

Information Warehouse Databases: An information warehouse stores data from current and previous years. This is usually data that has been extracted from the various operational and management databases of an organization. It is a central source of data that has been standardized and integrated so that it can be used by managers and other end-user professionals throughout an organization. For example, an important use of information



Data Storage, Retrieval and Data Base Management Systems

warehouse databases is pattern processing, where operational data is processed to identify key factors and trends in historical patterns of the business activity.

End User Databases: These databases consist of a variety of data files developed by end users at their workstations. For example, users may have their own electronic copies of documents they generated with word processing packages or received by electronic mail. Or they may have their own data files generated from spreadsheet and DBMS packages.

External Databases: Access to external, privately owned online databases or data banks is available, for a fee, to end users and organizations from commercial information services. Data is available in the form of statistics on economic and demographic activity from statistical data banks. One can receive abstracts from hundreds of newspapers, magazines, and other periodicals from bibliographic data banks.

Text Databases: Text databases are natural outgrowth of the use of computers to create and store documents electronically. Thus, online database services store bibliographic information such as publications in larger text databases. Text databases are also available on CD-ROM optical disks for use with microcomputer systems. Big corporations and government agencies have developed large text databases containing documents of all kinds. They use text database management systems software to help create, store, search, retrieve, modify, and assemble documents and other information stored as text data in such databases. Microcomputer versions of this software have been developed to help users manage their own text databases on CD-ROM disks.

Image Databases: Up to this point, we have discussed databases, which hold data in traditional alphanumeric records, and files or as documents in text databases. But a wide variety of images can also be stored electronically in image databases. For example, electronic encyclopedias are available on CD-ROM disks which store thousands of photographs and many animated video sequences as digitized images, along with thousands of pages of text. The main appeal of image database for business users are in document image processing. Thousands of pages of business documents, such as customer correspondence, purchase orders and invoices, as well as sales catalogues and service manuals, can be optically scanned and stored as document images on a single optical disk. Image database management software allows employees to hold millions of pages of document images.



Information Technology

Workers can view and modify documents at their own workstations and electronically transfer them to the workstations of other end users in the organization.

2.12.1 Other Database models

(i) **Distributed Database:** When an organization follows a centralized system, its database is confined to a single location under the management of a single group. Sometimes an organization may require decentralizing its database by scattering it with computing resources to several locations so that running of applications programs and data processing are performed at more than one site. This is known distributed data processing to facilitate savings in time and costs by concurrent running of application programs and data processing at various sites. When processing is distributed since the data to be processed should be located at the processing site, the database needs to be distributed fully or partly, depending on the organizational requirements. There are two methodologies of distribution of a database. In a *replicated database* duplicates of data are provided to the sites so that the sites can have frequent access to the same data concurrently. But this method of replication is costly in terms of system resources and also maintaining the consistency of the data elements. In a *partitioned database* the database is divided into parts or segments that are required and appropriate for the respective sites so that only those segments are distributed without costly replication of the entire data. A database can be partitioned along *functional lines* or *geographical lines* or *hierarchically*.



Data Storage, Retrieval and Data Base Management Systems

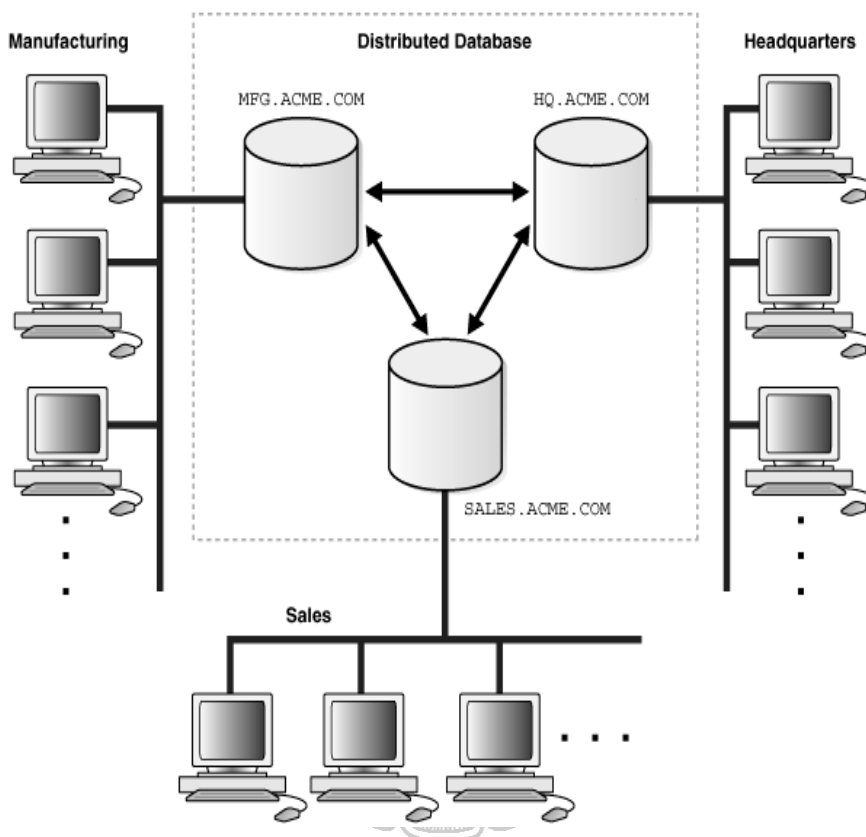


Fig 2.12.1.1 Example of distributed database

(ii) **Object Oriented Database:** It is based on the concept that the world can be modeled in terms of objects and their interactions. Objects are entities conveying some meaning for us and possess certain attributes to characterize them and interacting with each other. In the fig 2.12.1.2, the light rectangle indicates that 'engineer' is an object possessing attributes like 'date of birth', 'address', etc. which is interacting with another object known as 'civil jobs'. When a civil job is commenced, it updates the 'current job' attribute of the object known as 'engineer', because 'civil job' sends a message to the latter object.



Information Technology

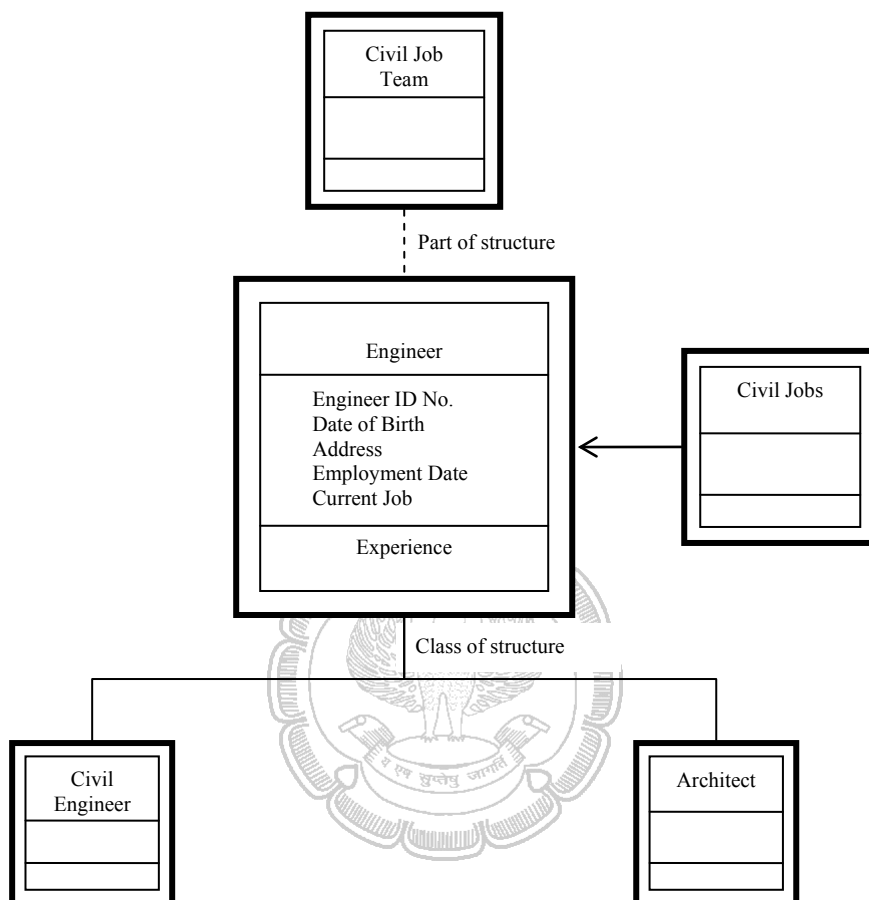


Fig 2.12.1.2 An object-oriented database design

Objects can be organized by first identifying them as a member of a class / subclass. Different objects of a particular class should possess at least one common attribute. In the fig 2.12.1.2 the dark rectangles indicate 'Engineer' as a class and 'Civil Engineer' and also 'Architect' as both subclasses of 'Engineer'. These subclasses possess all the attributes of 'Engineer' over and above each possessing at least one attribute not possessed by 'Engineer'. The line intersecting particular object classes represents the class of structure. Secondly, objects can be identified as a component of some other object. In the fig 2.12.1.2 'Engineers' are



Data Storage, Retrieval and Data Base Management Systems

components of a 'Civil Job Team' which may have one to more than one number of member(s). An 'Engineer' may not be a member of the 'Civil Job Team' and may not be a member of more than one team. The dotted line intersecting particular object classes represents the part of structure. Apart from possessing attributes, objects as well as possess methods or services that are responsible for changing their states. In the fig 2.12.1.2 for example, the service 'Experience' as a Civil Engineer or Architect for the object 'Engineer' calculates how much experience the engineers of these particular two subclasses have as professionals.

The motivation for development of object-oriented analysis and design of database are encapsulation and inheritance. Encapsulation indicates that the particulars of an object are hidden in capsule to keep it apart from the other objects. In the fig 2.12.1.2 for example, only minimum details about the attributes and services of an 'Engineer' is exposed to other objects. But the hiding technique weakens the coupling between the objects resulting in having fewer effects when there is a change to the system. Inheritance indicates that the object in a subclass automatically acquire or inherit the attributes and services of their class. In the fig 2.12.1.2 for example, the 'Civil Engineers' and the 'Architects' possess all the attributes and services of the class 'Engineers'. In fact inheritance develops reuse of objects and higher system reliability.

Now a day the database system is used increasingly to store –

- (i) Data about manufacturing designs in which focus is given to design objects that can be composed or decomposed into other design objects (like Telco resorts to CAD-CAM techniques),
- (ii) Images, graphics, audio, video which can be used to support multimedia applications.
- (iii) **Client-server Database:** It is designed in a structure in which one system can connect to another system to ask it question or instruct it to perform job. The system that asks the questions and issues the instructions is the client and the system answering the queries and responding to the instructions is the server. The client machine contains the user interface logic, business logic and the database logic and the server machine contains the database. Both are coupled with a network of high bandwidth.



Information Technology

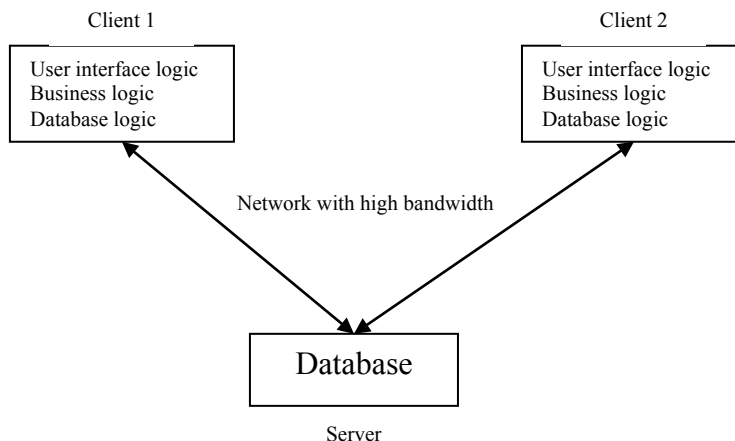


Fig 2.12.1.3 A Client-server Database Design (2-tier)

Whereas the user interface program or front end program is called the client, a back end program is called a server which interacts with share resources in an environment which can be based on heterogeneous hardware / software (Operating System) platforms of the client and the server and multi-vendor. The computational functions are shared in such a way that the server does all such higher level functions which it alone can do leaving the client to perform low level functions. The system scaleable in as much as clients may be added or removed and the shared resources may be relocated to a larger and faster server or to multiple servers.

The above is a 2-tier model implying a complicated software distribution procedure. Since all the application logic is executed on the personal computers, all these personal computers have to be updated in case of a new software release, which is bound to be very costly, time consuming, complicated and error prone. Once it reaches one of the users, the software first has to be installed and then tested for correct execution. Due to distributed character of such a procedure, it is not assured whether all clients work on the correct copy of the program.

3-tier and n-tier client-server database designs try to solve these problems by simply transferring the application logic from the client back to the server. This is accomplished by inserting an application server tier between the data server tier and the client tier. Client tier is responsible for data presentation, receiving user events and controlling the user interface. The actual business logic is not handled by the client tier; instead it is handled by the application server tier. In the light of object-oriented analysis (OOA), business objects which implement the business rules, reside in application server tier which forms the central key to solve the 2-tier problems and protects the data from direct access by the clients. Data server tier is



Data Storage, Retrieval and Data Base Management Systems

responsible for data storage. Besides, the relational database structure, legacy system structure is often used.

(iv) **Knowledge Database:** A database system provides functions to define, create, modify, delete and read data in a system. The type data maintained in a database system historically has been *declarative data* describing the static aspects of the real world objects and their associations. A pay roll file and a personnel file can share data about pay rates for each and every employee, their positions, names, etc.

A database system can also be used to maintain *procedural data* describing the dynamic aspects of the real world objects and their associations. The database can contain for example, several amended versions of enactments in the field of labour laws to facilitate management decisions in pay negotiations. When both the declarative and procedural data are stored in a database it constitutes a knowledge database with more *powerful* data maintenance.

The emergence of voluminous databases and higher use of decision support systems (DSS) and executive information systems (EIS) have led to increased interest regarding database structures which allow recognition of patterns among data and facilitate knowledge discovery by the decision makers.

A voluminous database which contains integrated data, detailed data, summarized data, historical data and metadata (data about data) is called a **data warehouse**. A database which contains selective data from a data warehouse meant for a specific function or department is called a **data mart**. The process of recognizing patterns among data contained in a data warehouse or a data mart is called a process of *data mining*.

2.13 STRUCTURED QUERY LANGUAGE AND OTHER QUERY LANGUAGES

A query language is a set of commands to create, update and access data from a database allowing users to raise adhoc queries / questions interactively without the help of programmers. A structured query language (SQL) is a set of thirty (30) English like commands which has since become an adoptable standard. The structured query language syntax use some set of commands regardless of the database management system software like 'Select', 'From', 'Where'.



Information Technology

SQL Statement

```
SELECT FIRST NAME, LAST NAME, MONTHLY FEE  
FROM MEMBER, MEMBERSHIP PLANS  
WHERE MEMBER.MEMBER ID = MEMBERSHIP PLANS.MEMBER ID  
ORDER BY LAST NAME
```

SQL Statement Result

First Name	Last Name	Monthly Fee
Marcus	Green	20.25
Shannon	Murray	39.50
Adrian	Valesquez	45.50
Donna	Vandenburg	55.50
Jonah	Weinberg	45.50

Fig 2.13.1 Example of SQL Statement

Some query languages have been designed in such a way that the command used is as close to Standard English text as possible. Query languages in a user-friendly way allow users to retrieve data from database without exposure to -

- (i) file / record structure,
- (ii) processes that the system performs,
- (iii) Languages like Common Business Oriented Language (COBOL), Beginner's All-purpose Symbolic Instruction Code (BASIC) or such other standard programming languages. Data retrieving efficiency may be improved by gaining knowledge about query shortcuts, query strategies and the types of data used in the database. A training and development program for this purpose may prove useful for the end users.

2.13.1 Natural Language: It is difficult for a system to understand natural language due to its ambiguity in sentence structure, syntax, construction and meaning. Accordingly it is not possible to design an interface given this problem. However systems can be designed to understand a subset of a language which implies the use of natural language in a restricted domain.

2.14 STORAGE

The storage structure design entails the decisions, which must be made on how to make it linear and to partition the data structure so that it can be stored on some device. The file library functions as an aspect of Operations Management Controls, takes responsibility of machine readable storage media for the management. In this context four functions must be undertaken: -

- (i) Ensuring that removable storage media are stored in a secure and clean environment,



Data Storage, Retrieval and Data Base Management Systems

- (ii) Ensuring that storage media are used for authorized purposes only,
- (iii) Ensuring maintenance of storage media in good working condition and
- (iv) Ensuring location of storage media at on-site / off-site facilities.

A storage media can contain critical files and hence should be housed securely. In case of mainframe operations, a separate room may be required just adjacent to computer room to house storage media required during daily operations, restricting access to the room and fixing the responsibility in the hands of a file librarian. Similar arrangements may be followed in the case of an off-site backup storage. Both on-site and off-site facilities should be maintained in a constant temperature, dust free environment. For managing a bulk of removable storage media effectively an automated library system is needed which records the following –

- (i) An identifier of each storage medium,
- (ii) Location where each storage medium is at present placed,
- (iii) Identity of the person responsible for storage medium,
- (iv) Identity of the present custodian of the storage medium,
- (v) List of files stored on each medium,
- (vi) List of authorized persons to access,
- (vii) Date of purchase and history of use with history of difficulties in using the storage medium,
- (viii) Date of expiry when contents of the storage medium can be deleted,
- (ix) Date of last release from and date last return to the file library,

Use of storage media is controlled with grate care. In mainframe environments, file librarians issue removable storage media as per authorized production schedule by transporting the required media themselves to computer room and collecting it after the production runs. In other cases file librarians issue it on the basis of an authorized requisition only, side by side recording the use of removable storage media. In micro computer environments, the extent of control over its use is dependent on the criticality of the data maintained on it. At an off-site location asset safeguarding and data integrity objectives can be seriously compromised. Off-site personnel usually works for a vendor who may specialize in providing backup facilities and the off-site personnel may not have the knowledge of the customer organizations to determine who is authorized to access the storage media. Care should also be taken when multiple files



Information Technology

are assigned to a single medium, for an application system reading one file may also read another file containing sensitive data. Actually, sensitive files should exist alone or separately on a medium. Otherwise the Operating System should be designed to exercise control over an application system to restrict access to authorized files only. To reduce the chance of exposure of sensitive data at some point of time, file should be expunged out of storage media as soon as their retention dates expire. Such media can then be reused to store other files, thereby reducing the inventory requirement.

Reliability of storage media should be monitored, especially which are used to store critical files, since the reliability decreases with age. For example, an increasing number of read / write errors are crept in the magnetic tapes and diskettes as these become older and used more. Storage media should not be kept unused for long, otherwise the possibility of read / writes error occurring with the media increases. With a magnetic tape for example, pressure builds up towards its center on its hub as it grows older. The pressure causes bonding and compacting of the ferromagnetic material from which the magnetic tape is made resulting in unread ability of the tape. Keeping this in view, a backup storage media need to be managed. Even if the backup has to be stored for long periods, it should be retrieved at regular intervals and backup files re-written to another fresh medium.

When storage media becomes unreliable, it is best to discard it; after ensuring that all sensitive data is removed from the discarded media. Simply deleting the files on the media does not serve the purpose. It may be necessary to carry out bulk erasure followed by writing random bits to the media followed by carrying out of bulk erasure once again. In some cases, unreliable storage media can be repaired. For example magnetic tapes can be cleaned to improve its reliability and to enhance its life. If storage media is required to be sent out side the organization for cleaning or repair, care should be taken to erase any sensitive data contained. The data on magnetic tapes for example, can be erased by degaussing or demagnetization.

Removable storage media should be located on-site when these are needed to support production running of application systems and should be located off-site when intended for backup and recovery. In a mainframe environment, it is the responsibility of the file librarian to manage transport of removable storage media to and from off-site locations. These movements should have compliance with the backup schedules. Backup schedules are prepared by a team comprised of security administrator, data administrator, application project manager, system development manager, operations manager and the file librarian. In a mini computer / micro computer environment, a person still performs file librarian duties for both on-site and off-site movements. Backup is prepared for minicomputer disk storage, local area network file server storage and critical microcomputer files. Good control is assured if these responsibilities are vested for example, in the file librarian who takes responsibility for backup



Data Storage, Retrieval and Data Base Management Systems

of the mainframe. One operation staff dealing with minicomputers and local area networks and the other operations staff dealing with microcomputer can provide backup to mainframe file librarian for transport to off-site storage. If the mainframe file librarian does not perform the above mentioned functions, the other two operation staff may perform it themselves, provided management formulates standards on it and propagates it in that way.

2.15 DOCUMENTATION AND PROGRAM LIBRARY

The documentations that are needed to support a system in an organization are–

- (i) Strategic and operational plans,
- (ii) Application systems and program documentation,
- (iii) System software and utility program documentation,
- (iv) Database documentation,
- (v) Operations manuals,
- (vi) User manuals,
- (vii) Standard manuals.

Besides, ancillary documents like memoranda, books and journals are also the required documents to support a system. These are kept in automated form for example, computer aided systems engineering (CASE) tools are used to provide machine readable formats of dataflow diagrams or entity relationship diagrams or some software that can provide documentation on optical disks (CD-ROM) However much of the documentation is still kept in hard copy formats since it has still some advantages over online documentation.

The difficulties in the management of systems documentation are as follows –

- (i) Responsibility for documentation is dispersed throughout the organization. For example, a librarian may be responsible for documentation supporting mainframe and mini computer systems, whereas documentation supporting micro computer system may be the responsibility of its users,
- (ii) Documentation is maintained in multiple forms and in multiple locations. For example, some part may exist in magnetic form, some other part in hard copy form and the remaining part in the micro form,
- (iii) Given the density and dispersion of documentation, proper updating, accessibility and adequate backup are not ensured.



Information Technology

The responsibilities of documentation librarians are ensuring that –

- (i) Documentation is stored securely,
- (ii) Only authorized user can have access to it,
- (iii) It is updated,
- (iv) Adequate backup exists for it,

Many organizations acquire a large number of software packages to support its microcomputers operations. In case the inventory of software is not managed by the documentation librarians properly, it may lead to problems like –

- (i) Purchase of too many copies of the software,
- (ii) Loss or theft of the software,
- (iii) Loss or theft of the documentation,
- (iv) Illegal copying of the software,
- (v) Use of software not complying with the terms and conditions of the software license,
- (vi) Absence of the software backup or improper backup.

Various types of software are available to mitigate the above hardship by taking the responsibility for maintaining the records of purchases, distributions and uses of the software and its related documentation and ensuring compliance with terms and conditions of the licensing agreements by the users. Some local area network Operation Systems for example, can provide a utility which generates a report listing all software located at workstations or file servers in the network, for review.

2.15.1 Program Library Management System Software:

- (i) It provides several functional capabilities to effectively and efficiently manage data center software inventory which includes –
 - a. Application Program Code,
 - b. System Software Code,
 - c. Job Control Statements
- (ii) It possesses integrity capabilities such that –
 - a. Each source program is assigned,
 - b. A modification number is assigned,



Data Storage, Retrieval and Data Base Management Systems

- c. A version number is assigned,
- d. It is associated with a creation date
- (iii) It uses –
 - a. Password,
 - b. Encryption,
 - c. Data Compression,
 - d. Automatic backup
- (iv) It possesses update capabilities with the facilities of –
 - a. Addition,
 - b. Modification,
 - c. Deletion,
 - d. Re-sequencing library numbers
- (v) It possesses reporting capabilities for being reviewed by the management and the end users by preparing lists of –
 - a. Additions,
 - b. Deletions,
 - c. Modifications,
 - d. Library catalogue,
 - e. Library members attributes
- (vi) It possesses interface capabilities with the –
 - a. Operating System,
 - b. Job scheduling system,
 - c. Access control system,
 - d. Online program management
- (vii) It controls movement of program from test to production status
- (viii) At last, it changes controls to application programs.



Information Technology

2.15.2 Design of User Interfaces: After discussing about storage media, we now turn to design of user interface. It is important since it involves the ways in which the users will interact with the system. Elements which have to be considered in designing of user interface are as follows –

- (i) Source documents to capture data,
- (ii) Hard copy output reports,
- (iii) Screen layout for source document input,
- (iv) Inquiry screens for database queries,
- (v) Command languages for decision support systems,
- (vi) Query languages for the database,
- (vii) Graphic display and color or non-monochromatic display,
- (viii) Voice output to answer users or answer queries,
- (ix) Screen layout manipulation by mouse or light pen,
- (x) Icons for representation of the output.

The interface design is developed as follows –

- (i) Identifying system users and classifying them into homogeneous groups,
- (ii) The user group characteristics are understood like whether the system will be handled by novices or experts,
- (iii) Eliciting the tasks which the users wish to perform using the system,
- (iv) Commencing a preliminary design of the form of interaction that will support these tasks. Prototyping tools are usually employed to refine the design aspect with the users.

2.16 BACKUP, AND RECOVERY

Generally 'backup and recovery' is treated as one topic and 'disaster recovery' as another. 'Backup' is a utility program used to make a copy of the contents of database files and log files. The database files consist of a database root file, log file, mirror log file, and other database files called dbspaces.

'Recovery' is a sequence of tasks performed to restore a database to some point-in-time. Recovery is performed when either a hardware or media failure occurs. Hardware failure is a physical component failure in the machine, such as, a disk drive, controller card, or power supply. Media failure is the result of unexpected database error when processing data.



Data Storage, Retrieval and Data Base Management Systems

Before one begins recovery, it is a good practice to back up the failing database. Backing up the failing database preserves the situation, provides a safe location so files are not accidentally overridden, and if unexpected errors occur during the recovery process, database Technical Support may request these files be forwarded to them.

'Disaster recovery' differs from a database recovery scenario because the operating system and all related software must be recovered before any database recovery can begin.

2.16.1 Database files that make up a database: Databases consist of disk files that store data. When you create a database either using any database software command-line utility, a main database file or root file is created. This main database file contains database tables, system tables, and indexes. Additional database files expand the size of the database and are called dbspaces. A dbspace contains tables and indexes, but not system tables.

A **transaction log** is a file that records database modifications. Database modifications consist of inserts, updates, deletes, commits, rollbacks, and database schema changes. A transaction log is not required but is recommended. The database engine uses a transaction log to apply any changes made between the most recent checkpoint and the system failure. The checkpoint ensures that all committed transactions are written to disk. During recovery the database engine must find the log file at specified location. When the transaction log file is not specifically identified then the database engine presumes that the log file is in the same directory as the database file.

A **mirror log** is an optional file and has a file extension of **.mlg**. It is a copy of a transaction log and provides additional protection against the loss of data in the event the transaction log becomes unusable.

2.16.2 Online backup, offline backup, and live backup : Database backups can be performed while the database is being actively accessed (online) or when the database is shutdown (offline) When a database goes through a normal shutdown process (the process is not being cancelled) the database engine commits the data to the database files.

An online database backup is performed by executing the command-line or from the 'Backup Database' utility. When an online backup process begins the database engine externalizes all cached data pages kept in memory to the database file(s) on disk. This process is called a checkpoint. The database engine continues recording activity in the transaction log file while the database is being backed up. The log file is backed up after the backup utility finishes backing up the database. The log file contains all of the transactions recorded since the last database backup. For this reason the log file from an online full backup must be 'applied' to



Information Technology

the database during recovery. The log file from an offline backup does not have to participate in recovery but it may be used in recovery if a prior database backup is used.

A **live backup** is carried out by using the BACKUP utility with the command-line option. A live backup provides a redundant copy of the transaction log for restart of your system on a secondary machine in the event the primary database server machine becomes unusable.

Full and incremental database backup: A database backup is either a full or incremental backup. For a full backup, the database backup utility copies the database and log. An incremental backup uses the DBBACKUP utility to copy the transaction log file since the most recent full backup. When you perform an incremental backup the mirror log is not backed up. When you backup and rename the log files the transaction and mirror log file is renamed and new log files are created. You must plan to manually back up the mirror log. Be aware of this when planning out your backup and recovery strategy.

2.16.3 Developing a backup and recovery strategy: The steps suggested in the development of a backup and recovery strategy consist of the following:

- Understand what backup and recovery means to your business
- Management commits time and resources for the project
- Develop, test, time, document, health check, deploy, and monitor
- Beware of any external factors that affect recovery
- Address secondary backup issues.

(i) Understand what backup and recovery means to your business: How long can my business survive without access to the corporate data? Express your answer in terms of minutes, hours, or days.

If your recovery time is in minutes then database backup and recovery is critical to your business needs and it is paramount that you implement some kind of backup and recovery strategy. If recovery can take hours then you have more time to perform the tasks. If recovery can be expressed in terms of days then the urgency to recover the database still exists, but time appears to be less of a factor.

(ii) Management commits time and resources for the project: Management must decide to commit financial resources towards the development and implementation of a backup and recovery strategy. The strategy can be basic or quite extensive depending upon the business needs of the company. After developing a backup and recovery strategy management should be informed of the expected backup and recovery times. Anticipate management countering the timings by preparing alternative solutions. These alternative solutions could be requesting



Data Storage, Retrieval and Data Base Management Systems

additional hardware, improved backup medium, altering backup schedule, accepting a longer recovery time versus backup time. Then it will be up to management to decide what solution fits their corporate needs.

(iii) Develop, test, time, document, health check, deploy, and monitor: These phases are the core in developing a backup and recovery strategy:

- Create backup and recovery commands. Verify these commands work as designed. Does your full or incremental online backup work? Verify that your commands product the desired results.
- Time estimates from executing backup and recovery commands help to get a feel for how long will these tasks take. Use this information to identify what commands will be executed and when.
- Document the backup commands and create written procedures outlining where your backups are kept and identify the naming convention used as well as the kind of backups performed. This information can be very important when an individual must check the backups or perform a database recovery and the data base administrator (DBA) is not available.
- Incorporate health checks into the backup procedures. You should check the database to ensure the database is not corrupt. You can perform a database health check prior to backing up a database or on a copy of the database from your backup.
- Deployment of your backup and recovery consists of setting up your backup procedures on the production server. Verify the necessary hardware is in place and any other supporting software necessary to perform these tasks. Modify procedures to reflect the change in environment. Change user id, password, server and database name to reflect the change in environment.
- Monitor backup procedures to avoid unexpected errors. Make sure any changes in the process are reflected in the documentation.

(iv) Beware of external factors that affect recovery: External factors that effect database recovery are time, hardware, and software. Allow additional recovery time for entering miscellaneous tasks that must be performed. These tasks could be as simple as entering recovery commands or retrieving and loading tapes. Factors that influence time are the size of database files, recovery medium, disk space, and unexpected errors. The more files you add into the recovery scenario, it increases the places where recovery can fail. As the backup and recovery strategy develops it may be necessary to check the performance of the equipment and software ensuring it meets your expectations.



Information Technology

(v) **Protect database backups by performing health checks:** Database health checks are run against the database and log files to ensure they are not corrupt. The database validity utility is used to scan every record in every table and looks up each record in each index on the table. If the database file is corrupt, you need to recovery from your previous database backup. A database can be validated before being backed up or against a copy of the database from your backup.

2.17 DATA WAREHOUSE

A **Data warehouse** is a repository of an organization's electronically stored data. Data warehouses are designed to facilitate reporting and supporting data analysis. The concept of data warehouses was introduced during the late 1980's. The concept was introduced to meet the growing demands for management information and analysis that could not be met by operational systems. Operational systems were unable to meet this need for a range of reasons:

- The processing load of reporting reduced the response time of the operational systems,
- The database designs of operational systems were not optimized for information analysis and reporting,
- Most organizations had more than one operational system, so company-wide reporting could not be supported from a single system, and
- Development of reports in operational systems often required writing specific computer programs which was slow and expensive

As a result, separate computer databases began to be built that were specifically designed to support management information and analysis purposes. These data warehouses were able to bring in data from a range of different data sources, such as mainframe computers, minicomputers, as well as personal computers and office automation software such as spreadsheets and integrate this information in a single place. This capability, coupled with user-friendly reporting tools, and freedom from operational impacts has led to a growth of this type of computer system. As technology improved (lower cost for more performance) and user requirements increased (faster data load cycle times and more functionality), data warehouses have evolved through several fundamental stages:

- **Offline Operational Databases** - Data warehouses in this initial stage are developed by simply copying the database of an operational system to an off-line server where the processing load of reporting does not impact on the operational system's performance.



Data Storage, Retrieval and Data Base Management Systems

- **Offline Data Warehouse** - Data warehouses in this stage of evolution are updated on a regular time cycle (usually daily, weekly or monthly) from the operational systems and the data is stored in an integrated reporting-oriented data structure
- **Real Time Data Warehouse** - Data warehouses at this stage are updated on a transaction or event basis, every time an operational system performs a transaction (e.g. an order or a delivery or a booking etc.)
- **Integrated Data Warehouse** - Data warehouses at this stage are used to generate activity or transactions that are passed back into the operational systems for use in the daily activity of the organization.

2.17.1 Components of a Data warehouse: The primary components of the majority of data warehouses are shown in the fig 2.17.1.1 and described in more detail below:

Data Sources: Data sources refer to any electronic repository of information that contains data of interest for management use or analytics. This definition covers mainframe databases (e.g. IBM DB2, ISAM, Adabas, etc.), client-server databases (e.g. Oracle database, Informix, Microsoft SQL Server etc.), PC databases (e.g. Microsoft Access), spreadsheets (e.g. Microsoft Excel) and any other electronic store of data. Data needs to be passed from these systems to the data warehouse either on a transaction-by-transaction basis for real-time data warehouses or on a regular cycle (e.g. daily or weekly) for offline data warehouses.

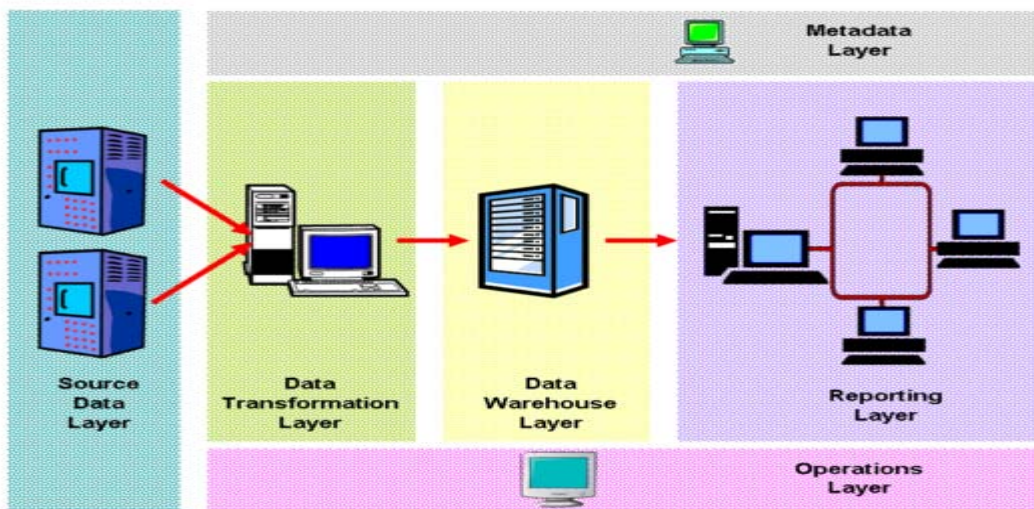


Fig 2.17.1.1 Components of Data warehouse.

2.63



Information Technology

Data Transformation: The Data Transformation layer receives data from the data sources, cleans and standardizes it, and loads it into the data repository. This is often called "staging" data as data often passes through a temporary database whilst it is being transformed. This activity of transforming data can be performed either by manually created code or a specific type of software could be used called an Extract, Transform and Load (ETL) tool. Regardless of the nature of the software used, the following types of activities occur during data transformation:

- Comparing data from different systems to improve data quality (e.g. Date of birth for a customer may be blank in one system but contain valid data in a second system. In this instance, the data warehouse would retain the date of birth field from the second system)
- standardizing data and codes (e.g. If one system refers to "Male" and "Female", but a second refers to only "M" and "F", these codes sets would need to be standardized)
- integrating data from different systems (e.g. if one system keeps orders and another stores customers, these data elements need to be linked)
- Performing other system housekeeping functions such as determining change (or "delta") files to reduce data load times, generating or finding surrogate keys for data etc.

Data Warehouse: The data warehouse is a relational database organized to hold information in a structure that best supports reporting and analysis.

Reporting: The data in the data warehouse must be available to the organization's staff if the data warehouse is to be useful. There are a very large number of software applications that perform this function, or reporting can be custom-developed. Examples of types of reporting tools include:

- Business intelligence tools: These are software applications that simplify the process of development and production of business reports based on data warehouse data.
- Executive information systems: These are software applications that are used to display complex business metrics and information in a graphical way to allow rapid understanding.
- Online Analytical Processing (OLAP) Tools: OLAP tools form data into logical multi-dimensional structures and allow users to select dimensions to view data.
- Data Mining: Data mining tools are software that allows users to perform detailed mathematical and statistical calculations on detailed data warehouse data to detect trends, identify patterns and analyze data.



Data Storage, Retrieval and Data Base Management Systems

Metadata: Metadata, or "data about data", is used to inform operators and users of the data warehouse about its status and the information held within the data warehouse. Examples of data warehouse metadata include the most recent data load date, the business meaning of a data item and the number of users that are logged in currently.

Operations: Data warehouse operations comprises of the processes of loading, manipulating and extracting data from the data warehouse. Operations also cover user management, security, capacity management and related functions.

Optional Components: In addition, the following components also exist in some data warehouses:

1. **Dependent Data Marts:** A dependent data mart is a physical database (either on the same hardware as the data warehouse or on a separate hardware platform) that receives all its information from the data warehouse. The purpose of a Data Mart is to provide a sub-set of the data warehouse's data for a specific purpose or to a specific sub-group of the organization.
2. **Logical Data Marts:** A logical data mart is a filtered view of the main data warehouse but does not physically exist as a separate data copy. This approach to data marts delivers the same benefits but has the additional advantages of not requiring additional (costly) disk space and it is always as current with data as the main data warehouse.
3. **Operational Data Store:** An ODS is an integrated database of operational data. Its sources include legacy systems and it contains current or near term data. An ODS may contain 30 to 60 days of information, while a data warehouse typically contains years of data. ODS's are used in some data warehouse architectures to provide near real time reporting capability in the event that the Data Warehouse's loading time or architecture prevents it being able to provide near real time reporting capability.

2.17.2 Different methods of storing data in a data warehouse: All data warehouses store their data grouped together by subject areas that reflect the general usage of the data (Customer, Product, Finance etc.) The general principle used in the majority of data warehouses is that data is stored at its most elemental level for use in reporting and information analysis. Within this generic intent, there are two primary approaches to organizing the data in a data warehouse.

The first is using a "dimensional" approach. In this style, information is stored as "facts" which are numeric or text data that capture specific data about a single transaction or event, and "dimensions" which contain reference information that allows each transaction or event to be classified in various ways. As an example, a sales transaction would be broken up into facts



Information Technology

such as the number of products ordered, and the price paid, and dimensions such as date, customer, product, geographical location and sales person. The main advantage of a dimensional approach is it is easier to understand and use. Also, because the data is pre-processed into the dimensional form, the Data Warehouse tends to operate very quickly.

The second approach uses database normalization. In this style, the data in the data warehouse is stored in third normal form. The main advantage of this approach is that it is quite straightforward to add new information into the database, whilst the primary disadvantage of this approach is that it can be quite slow to produce information and reports.

2.17.3 Advantages of using data warehouse: There are many advantages to using a data warehouse, some of them are:

- Enhances end-user access to reports and analysis of information.
- Increases data consistency.
- Increases productivity and decreases computing costs.
- Is able to combine data from different sources, in one place. (Provides a common data model).
- It provides an infrastructure that could support changes to data and replication of the changed data back into the operational systems.

2.17.4 Concerns in using data warehouse

- Extracting, cleaning and loading data could be time consuming.
- Data warehouses can get outdated relatively quickly.
- Problems with compatibility with systems already in place e.g. transaction processing system.
- Providing training to end-users.
- Security could develop into a serious issue, especially if the data warehouse is web accessible.
- A data warehouse is usually not static and maintenance costs are high.

2.18 DATA MINING

Data mining is concerned with the analysis of data and picking out relevant information. It is the computer, which is responsible for finding the patterns by identifying the underlying rules and features in the data.



Data Storage, Retrieval and Data Base Management Systems

Data mining analysis tends to work from the data up and the best techniques are those developed with an orientation towards large volumes of data, making use of as much of the collected data as possible to arrive at reliable conclusions and decisions. The analysis process starts with a set of data, uses a methodology to develop an optimal representation of the structure of the data during which time knowledge is acquired. Once knowledge has been acquired this can be extended to larger sets of data working on the assumption that the larger data set has a structure similar to the sample data. Again this is analogous to a mining operation where large amounts of low grade materials are sifted through in order to find something of value.

Examples of Data mining Software's are SPSS, SAS, Think Analytics and G-Stat etc.

The following fig 2.18.1 summarizes the common stages or processes identified in data mining and knowledge discovery by Usama Fayyad & Evangelos Simoudis, two of leading exponents of this area.

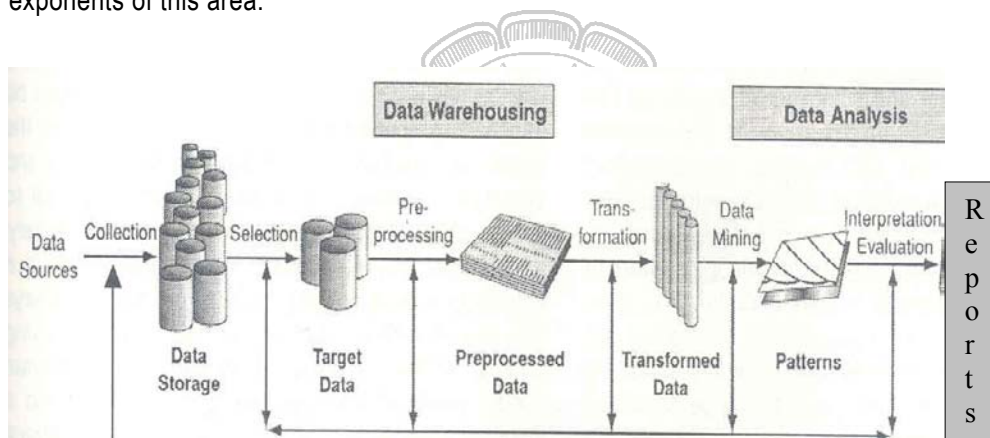


Fig 2.18.1 Components of Data mining.

The phases depicted start with the raw data and finish with the extracted knowledge which was acquired as a result of the following stages:

- **Selection** - selecting or segmenting the data according to some criteria e.g. all those people who own a car, in this way subsets of the data can be determined.
- **Preprocessing** - this is the data cleansing stage where certain information is removed which is deemed unnecessary and may slow down queries for e.g.: gender of the patient. The data is reconfigured to ensure a consistent format as there is a possibility of



Information Technology

inconsistent formats because the data is drawn from several sources e.g. gender may be recorded as F or M and also as 1 or 0.

- **Transformation** - the data is not merely transferred, but transformed. E.g.: demographic overlays commonly used in market research. The data is made useable and navigable.
- **Data mining** - this stage is concerned with the extraction of patterns from the data. A pattern can be defined as given a set of facts(data) F , a language L , and some measure of certainty C a pattern is a statement S in L that describes relationships among a subset F_s of F with a certainty c such that S is simpler in some sense than the enumeration of all the facts in F_s .
- **Interpretation and Evaluation** - the patterns identified by the system are interpreted into knowledge which can then be used to support human decision-making e.g. prediction and classification tasks, summarizing the contents of a database or explaining observed phenomena.

Self Examination Questions

Multiple Choice Questions

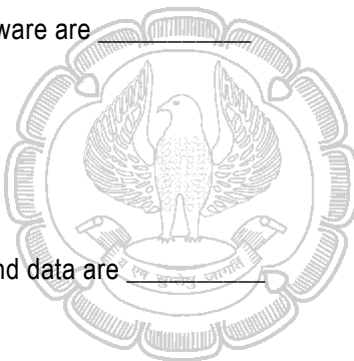
1. Hard disk is _____ disk on which you can store computer data
 - (a) Magnetic
 - (b) Removable
 - (c) Flash
 - (d) Temporary.
2. _____ in a database, which contains data about definition of the data element and their relationship.
 - (a) Data dictionary
 - (b) Data file
 - (c) Database
 - (d) None of these.
3. ER diagram helps to identify the database's _____ and their _____.
 - (a) Entities/ relationship
 - (b) Entities/ Models
 - (c) Entities/ data dependence
 - (d) None of these.





Data Storage, Retrieval and Data Base Management Systems

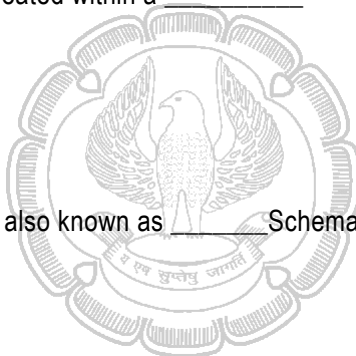
4. In a Network Database model _____ Structure is possible
 - (a) N:1
 - (b) M:N
 - (c) Both a and b
 - (d) None of these
5. All Candidate keys other than the primary keys are called _____
 - (a) Alternate keys
 - (b) Secondary keys
 - (c) Eligible keys
 - (d) None of these
6. Examples of RDBMS software are _____
 - (a) Oracle
 - (b) Sybase
 - (c) Informix
 - (d) All of these
7. In RDBMS the program and data are _____
 - (a) Data dependent
 - (b) Dependent
 - (c) Independent
 - (d) Program dependent
8. SQL is considered to be _____
 - (a) 1GL
 - (b) 2GL
 - (c) 3GL
 - (d) 4GL





Information Technology

9. DBMS comprises of following components, these are _____
- (a) Hardware and software
 - (b) People
 - (c) Procedures and data
 - (d) All of these
10. A DBMS that supports a database located at a single site is called a _____ DBMS
- (a) Client Server
 - (b) Distributed
 - (c) Centralized
 - (d) None of these
11. The data structures are created within a _____
- (a) File
 - (b) Wizard
 - (c) Database
 - (d) None of these
12. Physical level in DBMS is also known as _____ Schema
- (a) View
 - (b) External
 - (c) Logical
 - (d) Internal
13. View level in DBMS is also known as _____ Schema
- (a) View
 - (b) External
 - (c) Logical
 - (d) Internal





Data Storage, Retrieval and Data Base Management Systems

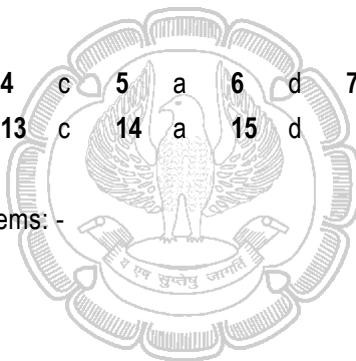
14. In the case of Physical data dependence, changes in _____ Schema do not require applications to be rewritten
- Physical
 - External
 - Logical
 - Internal
15. The phases in data mining include _____
- Selection
 - Transformation
 - Pre-processing
 - All of these

Answers

1 a 2 a 3 a 4 c 5 a 6 d 7 c 8 d 9 d
10 c 11 c 12 d 13 c 14 a 15 d

Short Term Questions

1. Explain the following systems: -
- Decimal
 - Binary
 - BCD
 - BCDIC
 - ASCII
2. Define the following: -
- Records
 - Fields
 - Date Field
 - Integer Filed
 - Double Precision Data,
 - Logical Data
 - Primary Key





Information Technology

- (h) Secondary Key
 - (i) Foreign Key
 - (j) Referential Integrity
3. In a database management system, describe the 3 level architecture with the following views: –
- (i) External or User View
 - (ii) Conceptual or Global View
 - (iii) Physical or Internal View
4. Define –
- (a) External Schema
 - (b) Conceptual Schema
 - (c) Internal Schema
5. In a database management system, there are two components –
- (a) Data Definition Language and
 - (b) Data Manipulation Language
- Do you agree? If yes, discuss the above two database management system facilities.
6. Elucidate the functions of the following which are parts of structure of a database management system: –
- (i) Data Definition Language Compiler
 - (ii) Data Manager
 - (iii) File Manager
 - (iv) Disk Manager
 - (v) Query Manager
 - (vi) Data Dictionary
7. Define the following models of Database: –
- (i) Distributed database (both replicated and partitioned)
 - (ii) Client-server database (2-tier architecture and 3-tier architecture)
 - (iii) Object-oriented database
 - (iv) Knowledge database.



Data Storage, Retrieval and Data Base Management Systems

8. Define the following high level languages and their functions: –
 - (a) Structured Query Language
 - (b) Natural Language
9. What are the documents that are needed to support a system in an organization?

Long Term Questions

1. Discuss the comparative advantages and disadvantages of the following models of a database management system: –
 - (i) Hierarchical Model,
 - (ii) Network Model,
 - (iii) Relational Model
2. The file Library function as an aspect of Operations Management Controls, takes responsibility of the machine readable storage media for the management by undertaking four functions. Discuss.
3. Describe how the use of storage media is controlled? How the reliability of storage media is monitored?
4. What are the difficulties in management of systems documentation? What are the responsibilities of documentation librarians?
5. Discuss the features of Program Library Management System Software.
6. What are the elements that are required to be considered in designing of user interface? How the interface design is developed?