

CHAPTER 1

COMPUTERS IN GENERAL

OBJECTIVES

At the end of this chapter the trainee will learn to

- Understand the characteristics of computers.
- List the five generations of computers.
- Understand the basic structure of a computer.

1.1 Introduction

The present century has seen a revolution in Information Technology all over the world. Computers are playing an increasingly vital role, touching upon every aspect of our day-to-day life. Information processing has become a major social activity. More than 70% of a typical executive's time is spent in processing some type of information.

A computer is an electronic device designed to perform arithmetic operation. It can also perform several non-arithmetic operations on the alphabetic or numeric data used. These operations are performed at very high speeds with a high level of accuracy.

The computer by itself is incapable of manipulations. It has to be instructed to perform a task in a language, which it can understand. The computer manipulates data by following a set of instructions given by the user. This set of instructions is called a program. A program directs the computer to perform a series of functions to achieve a predefined desired output.

Today the computer affects every sphere of human activity. It has already brought about revolutionary changes in industry, science, scientific research and education. Its impact on society has been tremendous; its areas of application virtually unlimited.

These applications are evident in almost every sector of society and business, ranging from travel (airlines, railways), manufacturing, banking, insurance, communications (including satellite communications for television and radio) to telephones, weather forecasting, space research, oil exploration, agriculture, defence, etc. In medicine, computers have helped to diagnose and treat several diseases hitherto believed to be incurable.

1.2 Characteristics of Computers

Some of the important characteristics of computer systems, which make computer based processing very useful in its practical applications, are described here.

Volume

Large volumes of data can be stored and processed by a computer system with a high degree of efficiency. The use of computers has thus made the processing of information economical by reducing unnecessary paper work.

Speed

A computer can process large volumes of data at very high speeds. The time required to execute a basic operation in a computer system varies between a few microseconds, nanoseconds or Pico seconds. With the advances in electronic technology, the processing speeds are increasing even further.

Accuracy

Coupled with the capacity to handle large volumes of data at high speeds, is the fact that the computer is extremely accurate. A human being commits approximately one error for every 500-1000 calculations using a calculator, while processing done by a computer is completely error free. The calculation can be repeated any number of times without worrying about errors. The accuracy of the computer output is wholly directly dependent on the accuracy of the input data and the correctness of the program, i.e., the set of instructions given by the user to process the data. The phrase garbage-in-garbage-out (GIGO) is generally used to describe inaccurate output caused by incorrect input data. This implies that errors seemingly caused by computers are in fact caused by incorrect input of data or incorrect programs, and are not errors incurred during the processing.

Complexity

In the past, computations with complex algorithms could not be done, as a single process, if done manually, was time consuming and also had a high probability of errors. Mathematical optimisation using operation research techniques like linear programming, simulation, etc., have been successfully used in modelling real-life applications only because of computers.

Quick/Easy Access

Information, by definition, is of no relevance if it cannot be made available at the right time. Computers, with their inherent characteristics of fast processing, have made it possible to retrieve and analyse data within a stipulated time. This has improved the timely availability of information.

Diligence

Unlike a human being, a computer does not suffer from fatigue, boredom, or lack of concentration while performing repetitive tasks. If a billion calculations are given to the computer, the billionth calculation will be as accurate as the first. Thus, for work that is repetitive in nature and where consistency needs to be maintained, computers have proved to be highly economical.

To summarize, the computer is a versatile, diligent machine that operates on a set of instructions and is capable of handling large volumes of data at tremendous speeds with a high degree of accuracy, ensuring timely, accurate and easy-to-access information.

1.2 A Brief History of Computers

In 1833, Charles Babbage began his work on the Analytical Machine. It was this machine which proved to be the prototype of the modern computer. The requirements for precision engineering were found impossible to achieve at the time and Babbage was unable to produce a working model. But this concept opened up avenues of the future development of computers.

1.2.1 Stages in the Evolution of the Computer

It has been more than 3000 years, from the days of the early computing machines to the existing generation of computers. The first stage in this evolution was the Abacus, a mechanical device; this followed by punched cards and finally came the revolution in computers and microprocessors. The impact of the evolution has been tremendous. A brief description of the various stages in the evolutionary process is given below:

1.2.1.1 Abacus

The Abacus was the first mechanical device developed approximately 3000 years ago to help process data. It consisted of a frame in which beads could slide on wires. The wires were arranged in columns. The column on the extreme right was for units, the one to its left for tens and so on. Pushing the appropriate beads could represent various numbers. This device could perform simple addition and subtraction operations.

1.2.1.2 Punched Card

The history of the punched card dates back to about the end of the American Revolution. At that time a French weaver, Joseph Marie Jacquard, used punched cards to control looms. The Jacquard-card is a card with punched holes to guide the warp threads on a loom. It was used primarily to automate weaving patterns into the fabric. It was this concept that gave Dr. Herman Hollerith the idea to build the census machine. i.e., a machine into which data was fed by means of card which had information punched on them. He devised a coding scheme, which could be used to punch data on cards. This machine was first used to tabulate census data in 1890. Hollerith later formed the Tabulating Machine Company to sell his invention. It is this company which later merged with others to form what is today the world's largest computer manufacturing company, the International Business Machines Corporation (IBM).

1.2.1.3 Analytical Engine

In 1833, Charles Babbage, a professor of mathematics at Cambridge University developed the concept for an analytical engine which could perform arithmetic functions on data, which was read in from punched cards. Babbage's requirements of precision engineering were impossible to achieve at that time and therefore a working model could not be produced.

Mark 1

In 1944, Howard Eiken, a Harvard professor, built an automatic calculating machine called the Mark 1 digital computer. It was an electromechanical device consisting of telephone relays and rotating mechanical wheels. It was controlled by a sequence of instructions punched into a roll of paper tape. In many ways, Mark 1 was a realization of Babbage's dream.

Eniac

While Mark 1 was the first electromechanical computer, the first electronic general purpose computer, called ENIAC (Electronic Numerical Integrator and Calculator), was built in 1946 by a team of professors at Moore School, University of Pennsylvania. As many as 18,000 vacuum tubes were used in building this machine. It weighed about 30 tons, occupied 1500 sq. ft of space and had a small memory (as compared to the memory capabilities of today's computers) of about 20 accumulators. It could perform 300 multiplication per second, making it 300 times faster than any other machine of that era.

Univac

The world's first commercial data processing machine UNIVAC (Universal Automatic Computer) was installed at the U.S. Bureau of Census in 1951. John Von Neumann had developed the concept of an "internally stored program", which was used in designing the UNIVAC system. By 1963, just 12 years after its installation, it retired as a historic relic.

1.3 Computer Generations

Computers are often grouped into categories known as computer generations. A brief description of the various generations of computers and their predominant features is given below:

1.3.1 First Generation (1946-1954)

These computers used vacuum tubes as the main electronic element and magnetic drums as the main memory. Programming had to be done in machine/assembly language only. These computers were physically large, generated significant heat and the speed of operations was in terms of milliseconds (i.e., thousandth of a second).

1.3.2 Second Generation (1955-1964)

In the second-generation computers, transistors replaced vacuum tubes as the basic electronic element. These transistorised computers were faster, more compact and reliable. The speed of operations was in terms of microseconds (i.e., millionth of a second). The magnetic core was introduced as the main memory. Magnetic tapes and disks were used for storage of bulk data. In the area of software, compilers were introduced thereby enabling the programmers to write programs in high-level languages. Some examples of such computers are: IBM-1401, Honeywell-800, IBM-1620.

1.3.3 Third Generation (1965-1974)

The introduction of the IBM-360 series of computers in 1965 marked the beginning of this generation. Transistors were replaced by monolithic integrated circuits. This reduced the size further and improved upon the speed, although the speed of computation was

still in terms of microseconds. In terms of software, it was in this generation that operating systems were introduced.

1.3.4 Fourth Generation (1975 onwards)

In this generation of computers, significant developments emerged, both in hardware and software. Semiconductor memory replaced the core memory, which enable the achievement of higher speeds at a lower cost. The introduction of micro programmed logic permitted the machine logic to be altered to suit the user's needs. The facility of operating systems with "virtual memory" enabled a programmer to write programs requiring greater memory space than actually available in the main memory. The processing speeds were of the order of nanoseconds.

1.3.5 Fifth Generation

This generation represents more technological developments both in hardware and software areas. Fifth-generation computers are aimed specifically at Artificial Intelligence (AI) applications.

1.4 Classification of Computers

Computers are divided into two categories on the basis of the type of data they are designed to process. Data may be obtained either as a result of counting, in which case it is called discrete data, or by using some continuous signal measuring instruments, in which case it is called continuous data.

The classification of computers, based on the type of data, is as follows:

1.4.1 Digital Computers

These computers operate on discrete data and are commonly used in business applications. Digital computers can be further classified into general-purpose and special-purpose computers.

1.4.2 General-Purpose Computers

A general-purpose computer is one that can be used for a variety of applications. Its versatility enables, execution of programs of almost any type. These are used for business applications.

1.4.3 Special-Purpose Computers

These are designed to perform specific tasks. Such computers lack versatility. However, they perform the task for which they are designed very efficiently. Aircraft control systems and missile guidance systems are some examples of special purpose computers.

1.4.4 Analog Computers

This type of computer works on continuous data measured along a continuous scale. For example, a speedometer is a mechanical device that works on continuous data. Analog computers are commonly used in process control systems, which monitor pressure, temperature, flow etc.

1.4.5 Hybrid Computers

This type of computer combines the properties of both analog and digital computers. A typical example is found in medicine, where analog properties are used to record the patient's data, while digital properties help in the analysis of the data as well as in monitoring the patient's health. Computers are also classified on the basis of their physical size, memory and processing speeds. Five of these categories are described below:

(I) Microcomputers

Low-Cost small digital computers are known as Microcomputers. Portable computers, personal computers, (PCs) (single user desk top computers), computers for dedicated applications such as industrial control, instrumentation, appliances control, etc. come under the category of Microcomputers. Microcomputers are system based on the use of microprocessors. A microprocessor is a programmable large-scale integrated circuit chip containing all the elements required to process binary encoded data. In simpler terms, a microcomputer performs all arithmetic and logical functions of a computer. Microcomputers are also called personal computers (PCs).

(II) Minicomputers

These systems are more powerful than microcomputers and are also more expensive. The word length is generally 16 bits or more. The processing speed of a computer is often measured in terms of MIPS, i.e., millions of instructions per second. These computers have a processing power of the order of 1.5 MIPS. They are multi-user system.

(III) Supermini Computers

Supermini computers are faster than minicomputers. The word length is generally 32 bits and the processing speed is more than 1.5 MIPS. These systems have efficient time-sharing operating systems with multi-programming features.

(IV) Mainframe Computers

Mainframes are very powerful large computers. They are rarely used and are being replaced by computer network (distributed computer system). These are machines with word lengths of 32 bits or more. The processing speed is of the order of 10 MIPS. These computers support a large main memory. In the time-sharing mode, due to fast processing speeds, they can support hundreds of terminals. The number of processors in mainframe varies from one to six.

1.5. Computers Today

Today computers are classified as notebook computers (laptop computer), Personal Digital Assistant (PDA, also known as Palmtop computer), desktop computers (PCs). Work stations, Servers and Supercomputers.

1.5.1 Laptop or notebook computers

Laptop or notebook computers are personal portable computers. They are used for word processing and spreadsheet computing while a person is travelling. Their power supply is from batteries. They consume less power. They use hard disks, floppy disks and flat LCD (Liquid Crystal Display). They can be connected to computer network. Wireless connection can be provided to laptops so that they can get information from large stationary computers. Notebooks use 32-bit processor specially designed for such computers.

1.5.2 Desktop Computers

These are single-user PCs (Personal Computers). 32-bit processors like Pentium III, Celeron, PowerPC etc. They use hard disks of capacity 8 GB - 20 GB. RAM capacity is 16 or 32 MB. They use 3.5-inch floppy disks. They use MS-DOS or WINDOWS-98 operating system.

1.5.3 Workstations

Workstations are also desktop computers. They are more powerful than desktop machines. They are provided with graphics capability. They are used for numeric and graphic intensive applications. Their hard disk and RAM capacity is more than desktop computers. They use hard disks of about 20 GB to 75 GB and RAM 64 MB. They use RISC processors such as DEC's Alpha processor, SUN's SPARC processor, HP's PA-RISC, etc. Operating systems used are UNIX, SUN's Solaris, HP's HP-UX, etc.

Another definition of workstation is "When a computer is connected to a network, it becomes a node on the network and is referred to as a workstation". Diskless Workstations boot directly from the server using special boot routines on the network interface card.

1.5.4 Servers

Servers are power computers. A number of PCs and terminals are connected to a server through a communicating network. Server has large disk and RAM capacity. A low-end server contains one microprocessor whereas a high-end server may contain more than one microprocessor. Microprocessors within the computer operate in parallel. A person working on PC connected to server, makes simple computation on the PC. For more complex computation he can connect his PC to server through LAN. He can utilise computing power, facilities and database available with server. He can also use the facilities available at other PC connected to the server.

1.5.5 Supercomputers

Supercomputers are the fastest computers and can be used to solve a wide range of large-scale problems, which require extensive numeric computations like complex molecular structural analysis, weather forecasting, etc. The processing speeds are of the order of 100 MIPS. These machines have word lengths of 64 bits or more.

1.6 Basic Computer Concepts

A computer takes in data, processes it and gives out information. So when we speak of the parts of a computer, there should be parts to:

- Put in data
- Process this data and
- Get the information out.

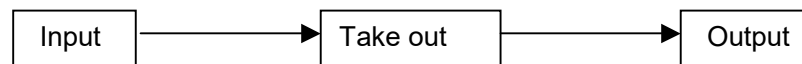


Fig. 1 Basic structure of a data processing system

The computer parts therefore, have been named exactly according to the function they are supposed to do. Look at the table below:

The Function	The Part
1. To put in the data	1. The input device
2. To process the data	2. The processor
3. To get the information out	3. The output device

The word 'input' means something that the user of the computer will 'put in'. Similarly the word 'output' means something that the computer will 'put out' after doing some work or 'processing' on the data. The word 'device' simply means a 'thing' or a 'gadget' that helps to input and output.

1.6.1 Parts of the computer therefore can be broadly listed as:

- ◆ Input devices
- ◆ The microprocessor (central processing unit)
- ◆ Video Display Unit or The Monitor
- ◆ Output devices
- ◆ Motherboard
- ◆ Power supply unit
- ◆ Memory chips
- ◆ Sound card (a few system build this into the motherboard)
- ◆ Power supply
- ◆ Floppy disk drive(s)
- ◆ Hard disk drive(s)
- ◆ CD-ROM drive(s)
- ◆ Cooling fan(s)
- ◆ A system board or motherboard; this usually contains:
 - ◆ Central Processing Unit (CPU)
 - ◆ Main memory (DRAM, EDO, SDRAM)
 - ◆ Expansion slots attached to PCI, ISA bus connectors

- ◆ System clock/calendar
- ◆ Keyboard adapter (interface)
- ◆ Mouse adapter (interface)
- ◆ Floppy disk controller
- ◆ Primary Enhanced Integrated Drive Electronics (EIDE)
- ◆ Interface, mainly for hard disks and CD-ROMs
- ◆ Secondary EIDE interface
- ◆ Many motherboards also contain these:
 - Heat sink unit
 - Serial port (RS-232C) or “COM” port
 - Parallel (printer) port
 - Static cache memory
 - Universal Serial Bus (USB) connections (usually two some times just one)

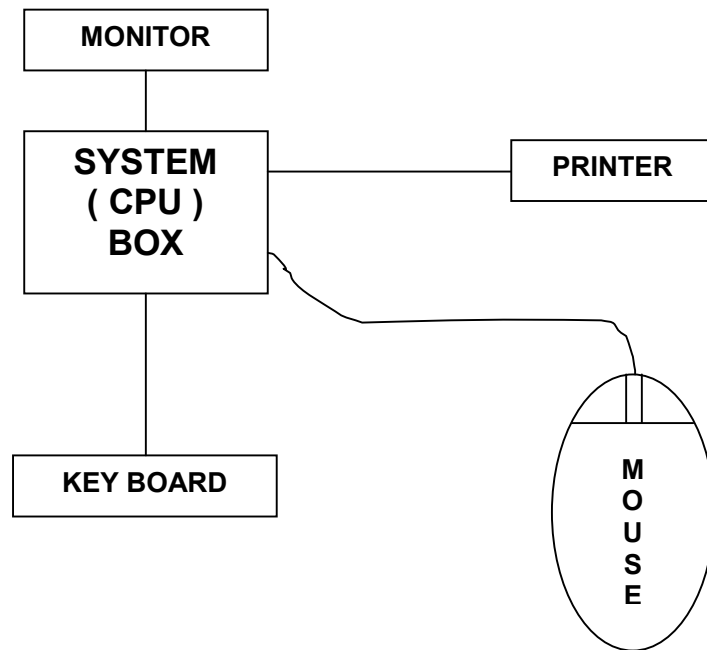


Fig. 2 Desktop Personal Computer

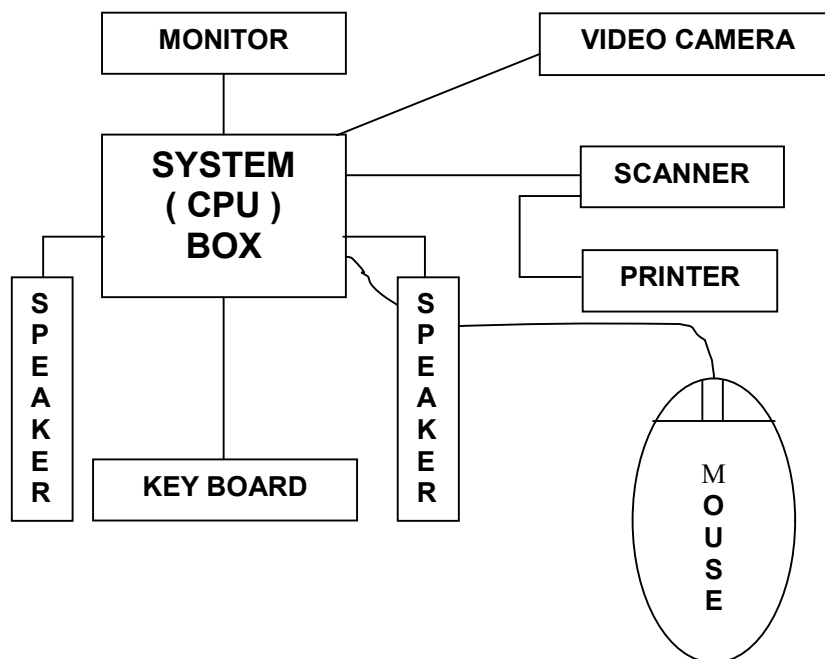


Fig. 3 Multimedia PC with Scanner

The Desk top Personal Computer, Multimedia PC with Scanner and PC – Functional block diagrams are shown in the figures

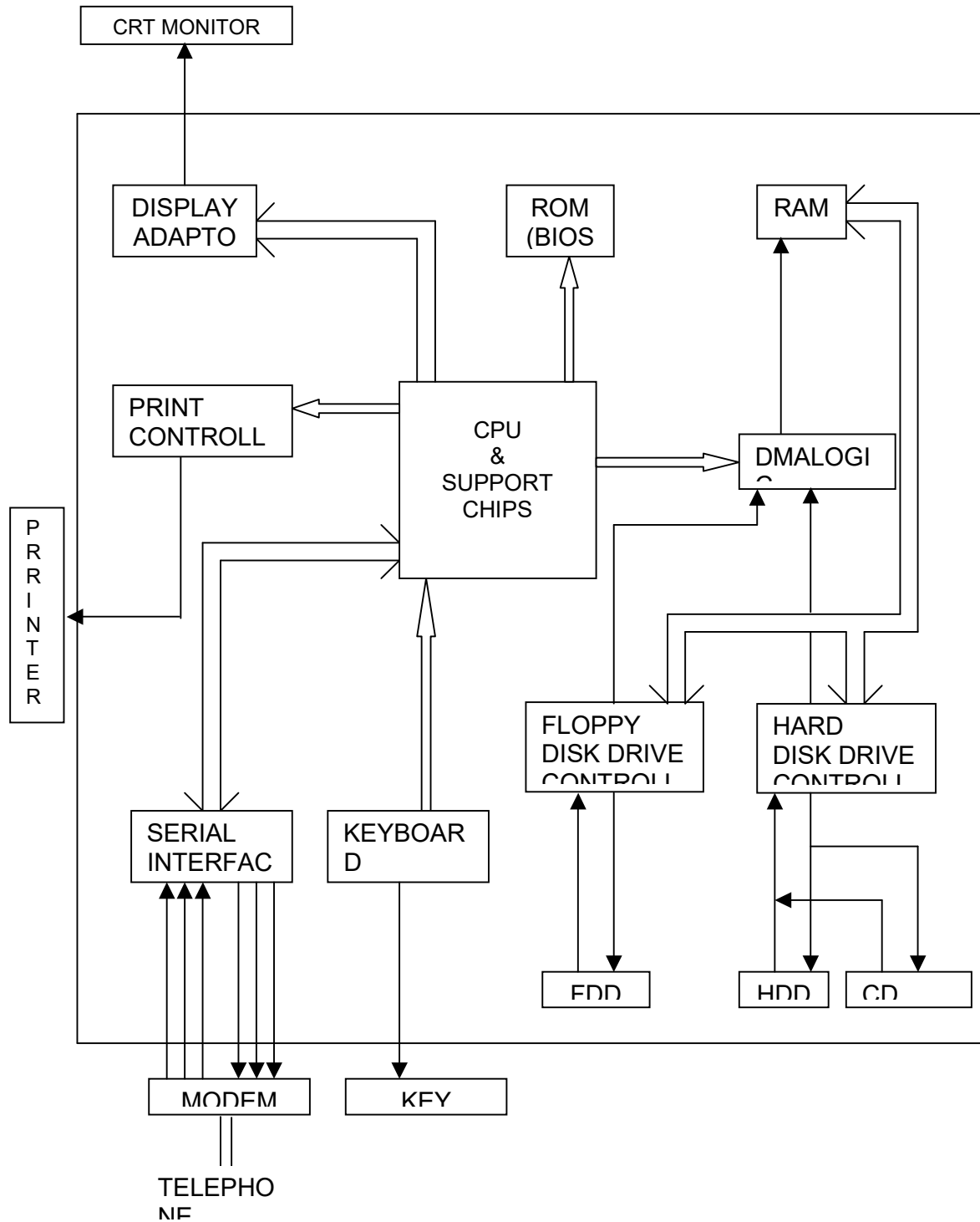


FIG. 4

PC – Functional block diagram

CHAPTER 2

C1B

COMPUTER AND ITS PERIPHERALS

OBJECTIVES

At the end of this chapter the trainee will learn to

- List the basic components of a computer.
- List various I/O devices.
- Understand the hardware terms.

2.1 Input And Out Put Devices

We now communicate with computers using input and output device. We use input devices to enter data into the computer so it can be processed. We use output devices to get it out so we can see it or distribute it to others. But input and output devices are not limited to the common display screen and printer. There is a vast range of devices in use and more are being developed every day. Some devices such as modems or fax are two-way in nature, able to both input and output information.

2.1.1 Input devices

There are lots of common types of input devices. The most common are keyboards, pointing devices, and scanners. Coming fast are speech recognition devices.

- Keyboards are used to enter data by typing it in-a very slow manual process. Fax modems send and receive faxes.
- Pointing devices are used to move a pointer around on the screen to execute commands or to manipulate the data that has already been entered.
- Scanners are used to enter data into the computer automatically. To use this form of data entry, the data must already exist in a printed form-either as text or graphics. Scanning devices can then read the printed data and convert it into a digital form that the computer can process.
- Speech recognition devices are able to hear the spoken words and then either execute them as commands or enter them as text.

2.1.2 Keyboard

The computer keyboard is similar to the typewriter keyboard. In addition to the typewriter keys, the keyboard has function keys for communicating specific instructions to the internal processor. The function keys perform tasks automatically that would otherwise be tedious and time consuming. For example, the function key labelled delete is used to remove a particular character or word from the text. The computer keyboard also has cursor keys, which are used to move the cursor up, down, right or left. It is a flat piece of plastic board with several buttons arranged on it. These buttons are for punching alphabets (A-Z), numbers from 0 to 9 and special characters like:

+ For addition, - For subtraction, * For multiplication & / For division, & , . " ? / @ \$! % & * () etc. All these buttons are called keys. And the entire plastic board along with the keys is called a keyboard. Whenever a key is pressed on the typewriter, that alphabet appears on the monitor.

Most computers today have keyboards with 104 keys. All the data is given to the computer by typing out the keys on the keyboard.



Fig. 5 Typical Keyboard

The keyboard of the computer is the most important device for communication between the operator and the computer. The key board has got the following displays and controls.

2.1.3 Displays

- Num Lock Indicator – Lights When numeric keypad is in numeric mode.
- Caps Lock Indicator – Lights when keyboard is in Caps Lock mode.
- Scroll Lock Indicator – Lights when keyboard is in Scroll Lock mode.

2.1.4 Controls

The keyboard (controls) is divided into following four sections:

- Alphanumeric keys
- Numeric key pad/cursor movement keys

- Special keys
- Function keys

Space Bar (the longest key in the last row of the key board) shifts the cursor one position to the right. If any character exists at the current cursor position when the space bar is pressed, that character is erased. After pressing the Insert key, however, that character will be shifted to the right when the space bar is pressed.

Alphanumeric keys

The keys have alphanumeric characters (a combination of the words, alphabet and numeric) that have functions identical to those of an ordinary typewriter. When the right or left **Shift** key is pressed, any simultaneously pressed alphabetic key, symbol key or numeric keypad key is switched from Lowercase letters mode to Upper letters mode.

When the **Caps Lock** key is locked, the alphanumeric keys are locked into the upper mode but other keys remained unaffected. Unlike a standard typewriter, while **Caps Lock** key is locked, the **Shift keys** convert uppercase letters to lowercase. Pressing the **Caps Lock** key causes the **Caps lock Status Indicator** to light up, showing that the keyboard is in **Caps Lock** mode. This mode is cancelled by pressing the **Caps Lock** key again.

2.1.5 Numeric key pad/cursor movement keys

These keys, which are arranged like a calculator, have two different functions according to the status of the **Num Lock** key. While the **Num Lock** key is not locked, these keys are used for controlling cursor movement or as edit keys. There is a separate set of Cursor Movement Keys, which are independent of status of **Num Lock** key.

When **Num Lock** is locked, these keys serve to input numbers to the computer and the **Shift** keys work in reverse. Pressing the **Num Lock** key causes the **Num Lock Status Indicator** to light up, showing that the keyboard is in **Num Lock** mode. This is cancelled by pressing the **Num Lock** again.

2.1.6 The functions of the numeric key pad/cursor movement keys are described below:

- **Cursor movement keys** move the cursor in space units in the horizontal direction and in line units in the vertical direction.
- **Home Key** shifts the cursor to the topmost left corner (home position) of the screen.
- **End key** shifts the cursor to right of the last character on the same line.
- **PgUp Key** moves the cursor to the identical screen position with in the preceding page.
- **PgDn Key** moves the cursor to the identical screen position with in the following page.
- **Insert Key** when pressed once, switches the key board from Replace mode to Insert mode, and the cursor changes to a blinking block, permitting characters to be inserted into the desired position of any line. Pressing this key again, switches the keyboard back to the Replace mode.
- **Delete Key** Deletes a character at the current cursor position on the screen, and all following data to the right of that position are shifted to left by one position.

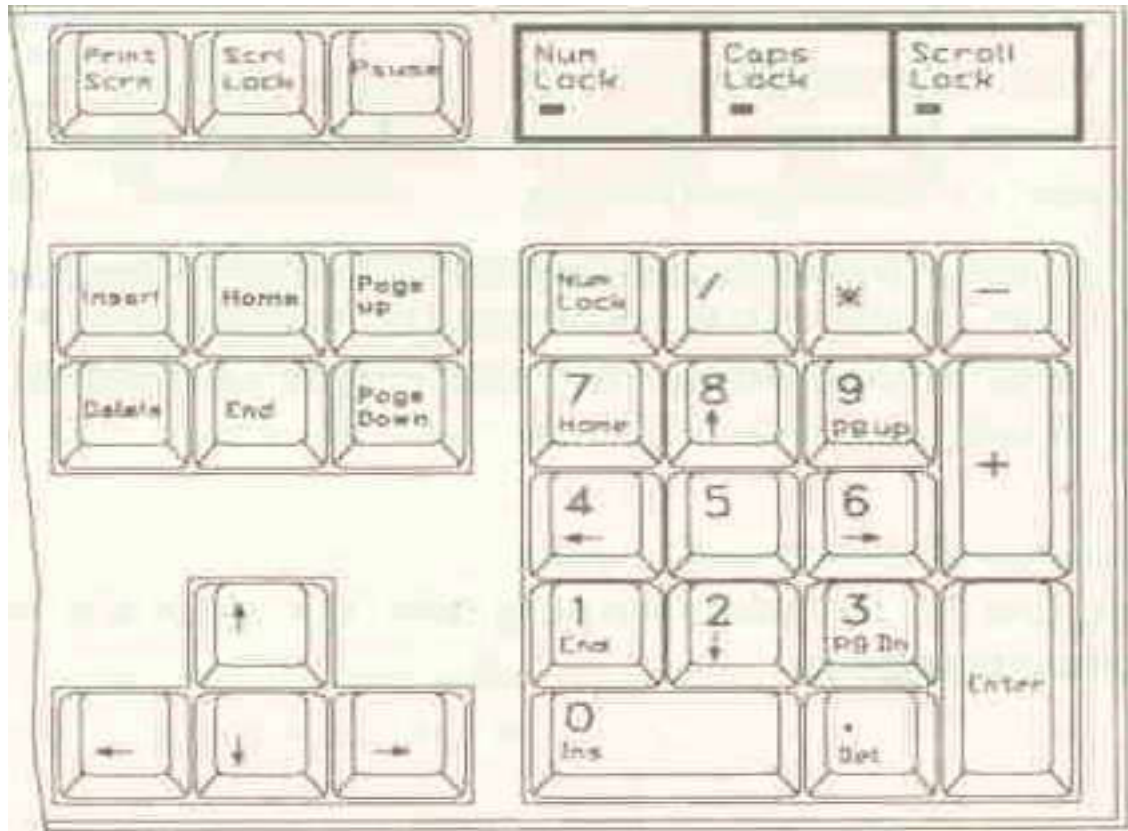


Fig.6 Numeric key pad/cursor movement keys

NOTE Pressing the **Del** key keeping both the **Ctrl** and **Alt** keys depressed causes the system to reboot.

2.5 Mouse

Keyboards were and are still very popular for typing data. For giving short instructions to the computer, the mouse has become very convenient to use today. With the recent software, most of the commands are picture based (GUI). Therefore the need of a point and click input device was felt and the mouse is an ideal pointing device. Generally 3-button mouse are in use. The left side button and the right buttons are for the point and click operations. The mouse is kept resting on a flat surface, and it is moved by the user on that surface. There is a pointer, which is displayed on the monitor screen, and the physical movement of the mouse is correspondingly shown on the screen through the pointer.

- The standard Mouse components are rollers ball that signals movement to the system.
- Buttons for making selections.
- A cable for connecting the mouse to the system.

- An interface connector to attach the mouse to the system.
- The Mouse works on OPTOMECHANICAL mechanism.
- The Mouse can be connected to the computer through Serial Interface or Dedicated Motherboard Mouse port.

2.5.1 Using the Mouse is shown in the following figures



Fig.7 Using the Mouse

If you are running the Microsoft Windows Operating system or another graphics based programmes, the pointer is often shaped like an arrow. If you are running a text-based programme, such as Microsoft Works, the pointer is often shaped like a box.

You move the pointer on the screen by moving the mouse in the direction you want the pointer to move. Moving the mouse slowly results in small movements of the pointer. Moving the pointer quickly results produces larger pointer movements.

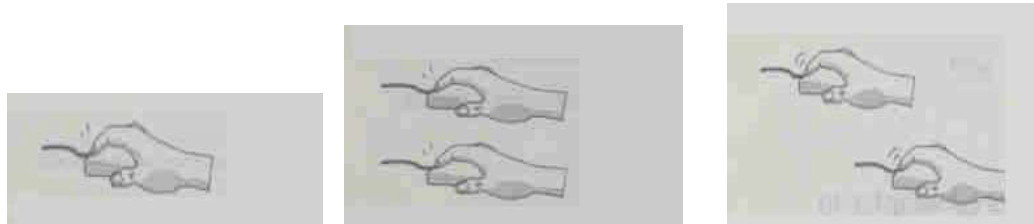


Fig. 8 Clicking

Double-clicking

Dragging

- **Clicking:** Clicking means pressing and releasing a mouse button. By pointing to something on the screen and clicking a mouse button, you select the object under the pointer.
- **Double-clicking:** Double-clicking means pointing to an object on the screen and quickly clicking a mouse button twice.
- **Dragging:** Dragging means Holding down a mouse button while moving the pointer.



Fig.9 Cleaning the Mouse

2.5.2 To clean the Mouse:

1. Turn the computer off, turn the mouse upside down, and remove the mouse-ball cover.
2. Turn the mouse right side up and let the ball and cover drop down.
3. Use adhesive tape to pick up any dust or lint on the surface of the ball. Wipe away dust or lint inside the mouse socket. Return the ball to its socket and replace the mouse-cover.



9-pin serial connector



PS/2 Connector

Fig.10 Mouse Connectors

Note: Always turn the computer **OFF** before connecting or disconnecting the Mouse.

2.6 Other Input Devices

Some other input devices, which are very widely used today, are:

- *MICR*: Magnetic ink character reader. Used in Banks for scanning the cheque numbers.
- *BCR*: Bar code reader. Used in supermarkets etc, to read the coded price tags on the items.
- *OCR*: Optical character recognizer. Used for validating examination papers and application forms, or for reading any printed matter.

2.7 Output devices

Once data has been input and processed you need access to it in a form you can use. The devices you use to do this are called output devices because data is sent or output

to them from the computer. The output can be in the form of words, numbers, graphics, sounds, videos, and animations.

- You may want to print your work on paper-called hard copy-so that you can take it with you, share it with others, or file it for later reference. Hard copy can be created with a stunning array printers that range from desktop units to units that paint an image on the wall.
- Output can be displayed on the screen-sometimes called soft copy-when creating, editing, and manipulating data. In this the soft copy is an intermediate stage towards producing hard copy. The screen, however increasingly the intended output device for multimedia out put.
- Musical output can come through a synthesizer or other musical system.
- Output can be in the form of instructions to another device, such as a robot.

2.8 Printers

The printer is the most commonly used output device to produce printouts of results, program and data. The printers are classified into the following three types:

(i) Character Printers

(ii) Line Printers

(iii) Page Printers

The above classification is according to the manner in which they print the text. There is one more classification depending upon the technology employed in manufacturing a printer i.e. depending on the mechanism used for printing, speed of printing, quality of output, direction of printing and the kind of interface they have with the computer.

According to this classification the printers fall into the following two categories:

(i) Impact Printers

(ii) Nonimpact Printers

In impact printers electromechanical mechanisms are employed to cause hammers or pins to strike against a ribbon and the paper to print the text. The most commonly used impact printer is the Dot-matrix printer (DMP). In a DMP an arrangement of tiny hammers or pins typically print in a matrix of 7 dots across and 9 dots down. Nonimpact printers do not employ mechanical print-head to strike against ribbon. They use thermal, chemical, electrostatic or inkjet technology for printing the text. Nonimpact printers are quiet during operation whereas impact printers are noisy. Usually, nonimpact printers are faster than impact printers. The disadvantage of nonimpact printers is that they produce single copy of the texts whereas impact printers can print multiple copies. To solve this problem of nonimpact printer's manufacturers have designed printers, which can be used as offline devices to produce additional copies of computer prepared output. The printer accepts data from magnetic tapes and produce out put.

2.8.1 Character Printers

Character printers print only one character at a time with the print head moving across a line. They are low-speed printers. Their printing speed varies in the range of 30 to 600 characters per second depending upon the type of the printer. They are used for low volume printing jobs. They are used with almost all microcomputers. Characters to be printed are sent serially to the printer. A modern printer contains a microprocessor and buffer memory. The microprocessor allows printer to receive the next set of characters to be printed from the computer while printing of the current set is going on. A character printer may be either an impact character printer or nonimpact character printer. Character printers can be further classified into two main categories, namely, dot matrix printers and letter quality printers.

2.8.2 Dot-Matrix Impact Character Printers

Characters are formed using a number of dots. To print a character selected dots, which are required to form the desired character, hit the ribbon against paper. Dot-matrix printers contain 9, 14, 18, or 24 printing dots in the print head. The large number of printing dots produces a good quality print. Some printers print a line twice positioning the dots for the second printing slightly offset. This technique produces characters as good as those printed by formed character printers. Dot matrix printers can print graphics. Colour graphics can also be printed using ribbons of different colour and making several passes across a line of the graph or figure. Most dot matrix printers contain microprocessor to control its functions. The print speed of dot matrix printers ranges up to 600 cps (character per second). Dot matrix printers can operate at two or three speeds. At lower speed printing quality is better. 80 columns and 132 column printers are available. Many dot-matrix printers are bi-directional which are capable of printing characters left to right as well as from right to left.

To print characters one character at a time is sent from the buffer memory to printing mechanism. The printer electronic decoder decodes the code of the character. Then the appropriate pins in the pin-head are actuated to print the desired character. Dot-matrix printers can be converted to print alphabets other than English. A dot-matrix printer is usually faster than letter quality impact printer such as a daisy-wheel printer. The impact type dot-matrix printers are inexpensive but noisy. Its print quality is not as good as that of a daisy wheel printer. Unlike the formed character such as a daisy-wheel printer, dot-matrix printers can also print graphics.

2.8.3 Letter Quality Impact Character Printers

Daisy-wheel type impact character printers are formed character printers. They were used with personal computers (PCs), word processor and electronic typewriters in the past. Now a days they are rarely used. In a daisy wheel printer each petal of the wheel has a character embossed on it. It operates like a typewriter. To print characters one character at a time is sent from the memory to printing mechanism. A motor is used to spin the wheel rapidly. To print a desired character the wheel is rotated until the desired character comes in the correct position over the paper. A solenoid driven hammer then strikes the petal against the ribbon to print the character.

The main advantage of a daisy-wheel printer is excellent print quality. For different types and styles of character such as italics, letters of different sizes etc different types of daisy-wheels are available. The daisy-wheel of the desired quality of letters can be fitted

with the printer. The speed of a daisy-wheel printer is 10 to 90 characters per second (cps). As a daisy-wheel printer employs impact mechanism it is noisy output device. Many daisy wheel printers are bi-directional.

2.8.4 Non-impact Character Printers

Non-impact character printers are also available. These printers use thermal, electrostatic, chemical and inkjet technologies.

2.8.5 Dot Matrix Thermal Character Printers

Most thermal printers use papers, which have a special heat sensitive coating. When a spot on the special paper is heated, the spot becomes dark. To print a character the print head consisting of 5 by 7 or 7 by 9 matrix of tiny heating elements is moved to a character position. Then the dot-sized heating elements for the desired character are turned on. The heating elements are turned off after a short time and the print-head is moved to the next character position.

Some of the newer thermal printers use special ribbon that contains the ink in a wax binder. When the hot pins of the print head press the ribbon against the paper, the wax melts and the ink is transferred to the paper. Colour graphics and text can be printed by changing ribbons. These printers use standard papers.

2.8.6 Ink- Jet Printers

An ink-jet printer employs a droplet generator to break special ink into tiny drops. Droplets are moved through a valve where each droplet receives a small electric charge. Thereafter the droplets are passed through deflecting plates. The deflected droplets fall on a sheet of paper. An ink-jet printer uses a dot-matrix approach to produce a character. The speed of this type of printer lies in the range of 40-300 CPS. Some printer can print upto 45000 lines per minute.

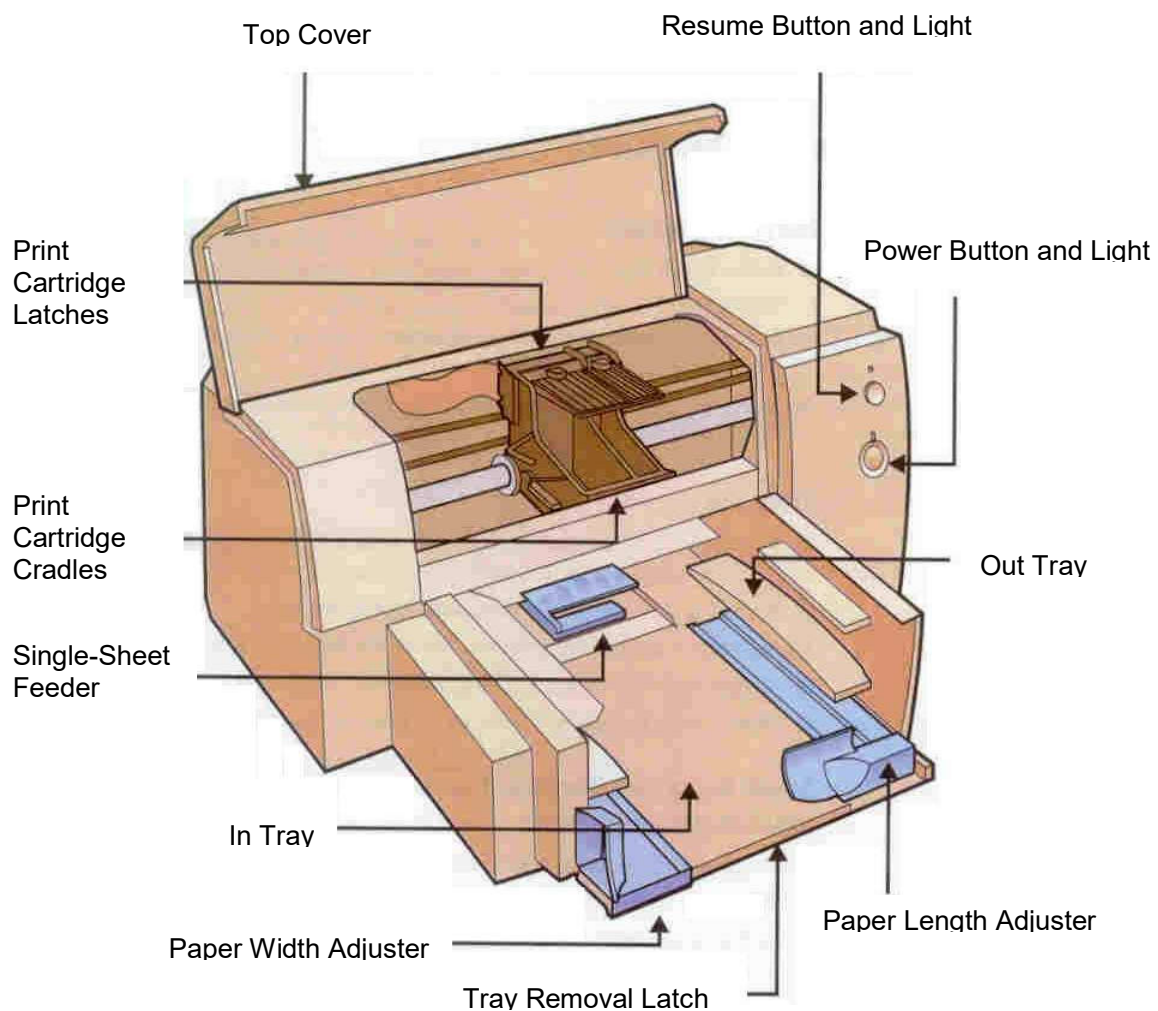


Fig. 11 HP DeskJet 870Cxi Printer

2.8.7 Electro static Printers

An electrostatic printer has a mechanism to move a sheet of paper over printing pins, which place small electric charge on the paper. Thereafter, the paper sheet is passed through a batch of oppositely charged toner particles. The toner particles are attracted to the spot charged on the paper by print pins. The paper sheet then moves and passes through a fusing process, which melts toner particles to form a character on the paper. An electrostatic printer can also print a graph.

2.8.8 Line Printers

A line printer prints one line at a time. The speed of line printers lies in the range of 300-3000 lines /minute. Line printers are used with mini and large computers. Such printers may be impact or nonimpact type. Impact type line printers are of three types: drum printers, chain printers and band printers. A drum printer employs a rapidly rotating drum that contains a complete set of raised characters in each band around the drum. There are a number of such bands on the drum. At each character position there is a printing head. When a character comes in the printing position a hammer strikes the paper and ribbon against the appropriate character on the drum. One revolution of the drum is necessary to print each line. All the characters of a line are not printed at a time. Since the drum rotates at a very high speed, the time required to print a line is short enough to call it a line printer.

A chain containing several character sets is employed in a chain printer. It is rotated at high speed. When the desired character comes in the correct position the hammer strikes against the ribbon and paper to print a character. Each character position has a magnetically driven hammer. A chain printer can print up to 2000 lines per minute.

A band printer is similar to a chain printer. It contains a fast rotating scalloped steel print band instead of a chain. Several raised character sets are made on a band. A band printer can print up to 3000 lines per minute.

2.8.9 Page Printers

Page printers print a page at a time. These are nonimpact printers. From speed point of view there are two types of page printers, namely, low speed page printers and high-speed page printers.

2.8.10 Low-Speed Page Printers

Such printers are used for low-volume printing jobs. They combine the printer and office-copier techniques, and use Laser or other light sources to produce the image of the desired output. Low speed printers can provide a much larger number of fonts. These printers are used with microcomputer. Their prices are falling rapidly. The speed of desk-size printer ranges between 6 to 12 pages per minute.

2.8.11 High Speed Page Printers

These printers are used for large-volume printing jobs. They can print at speeds over 200 pages per minute. High-speed page printers use xerography, lasers and other technologies for high-speed printing.

2.9 Operation of a Laser Printer

A page printer using laser beam employs a photosensitive drum to provide an image of the text on it. The bits of data sent by the computer to the printer act as triggers to turn the laser beam on and off. A negative electric charge is applied to the surface of the drum. An image is produced on the drum by turning the laser beam on and off when it is swept back and forth across the drum. The image is produced in about the same way as an image is produced on a raster scan CRT. The electric charge becomes positive on just those parts of the drum surface, which are exposed to the laser beam. To print page a modulated laser beam scans back and fourth across the surface of the drum. The

laser-exposed parts of the drum surface attract negatively charged toner (an ink powder) particles, which are transferred to the paper. Then the paper moves to a fusing station where toner is permanently fused on the paper with heat and pressure. Then the drum is cleaned and it is ready for the next page. When the drum rotates, the process of applying electric charge, writing texts using laser beam, applying ink and cleaning the drum, all are done sequentially.

The print quality of laser printers is excellent. Now-a days low-speed laser printers are also used with microcomputers. They are used for book printing. An output of the text is produced first by a computer-controlled printer. Then plates are prepared photographically from the output of the printer. The final printing of the book is done by conventional offset printing system. Laser can also print graphics. Texts and graphics can easily be printed on the same page. They are quiet in operation and use standard paper. HP Laser printer is shown in the figure.

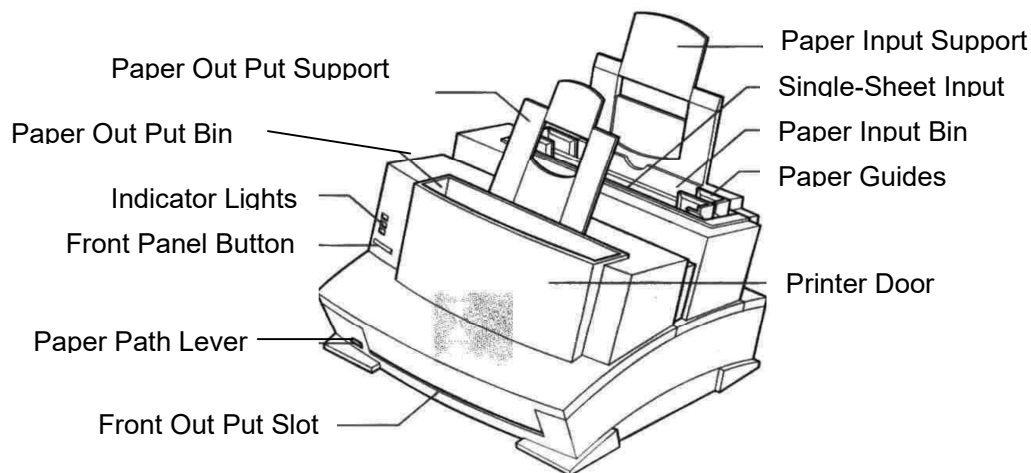


Fig.12 HP LaserJet Printer

2.10 Scanners

Data can be entered directly into computer with the help of scanners. The advantage of direct data entry is that one does not have to key data. This leads to faster and more accurate data entry. There are two major types of scanners: optical scanners and magnetic-ink character readers.

2.10.1 Optical Scanners

These are input devices, which read data recorded on paper. These are called optical devices because reading technique uses light sensors. The data to be scanned may be typewritten, data-coded as pencil marks or data-coded as bars. The commonly used scanning devices are: optical character reader, optical mark reader and bar-code readers.

2.10.2 Optical Character Readers (OCR)

These devices are capable of reading alphabetic and numeric characters printed on paper. The characters are of special designs, which are typed, using a special font known as OCR font. The device reads the characters and converts them into appropriate codes to be directly entered and stored in the computer. Earlier units were developed for large volume applications. But newer systems have been designed for use with personal computers. Several thousands printed characters can be read per second by an OCR.

2.10.2 Optical Mark Reader (OMR)

These devices can detect pencil or ink marks made on special paper forms using light sensors. The special forms are designed with or marks which can be filled with soft pencil or ink. This technique is suitable for areas where one out of small number of choices is to be marked. For example, objective type answer sheets, market survey, population survey etc. where responses can be restricted to one out of a few alternatives.

2.10.3 Optical Bar-code Readers

An optical bar-code reader detects printed bars and converts them into electrical pulses. The most widely used bar-code is the Universal Product Code (UPC), which is marked on retail packages. The UPC code consists of small vertical bars of varying thickness. These bars are read as ten digits. The first five digits identify the manufacturer or supplier of the product. The second five digits identify a specific product. An optical bar-code reader is often used with a point of sale terminal.

2.10.4 Magnetic-ink Characters Reader (MICR)

A magnetic-ink character reader can read characters printed on paper by using magnetic ink i.e. an ink that contains magnetizable particles of iron oxide. MICR is widely used by banks to process the tremendous volume of cheques received each day. On the lower left-side of the cheque the bank's identification number and the depositor's account number are printed with the special ink before the cheques are issued to the user. On the lower right-side of the cheque the amount of the cheque is printed by the bank staff with the same special ink after the cheque has been presented. An MICR can read these characters and transfer the data to the computer. As the cheque enters the reading unit, it passes through a magnetic field, which causes the iron oxide particles in the ink to become magnetised. Read heads can now interpret the characters printed on the cheque with the special ink. The data being read is entered directly into a computer. Sorting of cheques is also done according to their identification code numbers. MICR can process Upto 2500 cheques per minute.

2.11 Central Processing Unit (CPU)

The CPU is the heart of any computer system. This microprocessor is the heart of the computer. The microprocessor is the central processing. Each microprocessor carries a number on it like 8086, 80286, 80386, 80486, 'Pentium'. This number refers to the processor series number. The higher the series number, the higher is the power of the microprocessor. One can identify the microprocessor chip in the computer system very easily as, of all the other chips; this is the biggest one with respect to its size.

2.11.1 Arithmetic Logical Unit (ALU)

The ALU is the area where calculations and comparisons (decisions) are made. Once the data from the input device is fed into primary storage, it is transferred to the ALU, as and when required. The entire processing takes place here. Data moves back and forth from the primary storage area to the ALU, using the working storage area (of the primary storage section) until the processing is over. The results of processing are released to the output storage area where they are held until release.

The ALU contains one or more storage locations called registers. The registers provide temporary storage for one or both of the operands that take part in the arithmetic or logical operations. Generally, the result is also produced in a register.

2.11.2 The Control Unit

This is the central nervous system of the computer. It controls, maintains order and directs the operations of the entire system. The program instructions are stored in the main memory before execution. During execution they are transferred one by one to the control unit of execution. Here, the instructions are interpreted and appropriate action for its execution is taken.

2.11.3 Output Devices

Similar to the input devices, the output devices are instruments of communication between the user and the computer system. These devices are used to output results from the CPU. The most common output device is printer, which prints the result on paper. Other output devices are: magnetic tape drive, magnetic disk drive, floppy disk drive, plotter, visual display unit (VDU), etc.

2.11.4 Primary Storage Section

This is the main memory of the computer and is logically divided into the following four areas:

1. Input storage area: The area in which input data is held until it is ready to be processed.
2. Working storage area: The area that holds the data being processed and the intermediate results.
3. Output storage area: This is where finished results (output) are held until released to the output devices.
4. Program storage area: While the above three areas are data related storage areas, the program storage area holds the processing instructions.

It should be noted that these four storage areas are not created through physical boundaries in the primary storage section. They may vary from application to application.

The memory of the computer is of two types:

- i) Read Only Memory (ROM)
- ii) Random Access Memory (RAM)

2.11.5 Scrabble - An Analogy

The Scrabble brand crossword game is like a computer's memory. The instruction booklet is like ROM-you can read it, but you cannot change it. The board is like RAM-you can add, remove, or move data to and from, around on it.

2.12 MEMORY

Memory is an essential component of a digital computer. It is needed to store programs, data and results. A computer uses a number of memory devices of different technologies such as semiconductor memory, magnetic memory and optical memory. The CPU is a semiconductor device and it operates at a very high speed. The speed of the memory should also be very high and it must match the speed of the CPU. If memory is slow, the CPU has to wait for data and instructions. This will reduce the processing speed of the computer. To match the speed of the CPU a very fast semiconductor memory called **cache memory** is directly connected to the CPU. It stores currently needed instructions and data of a program being executed. Its access time is about 10 ns. It uses static RAMs. The next level of memory, which is used in a computer system, is the **main or primary memory**. It is also a fast semiconductor memory. It stores programs and data, which are being processed by the CPU. Its access time is about 50 ns. It is cheaper than cache memory. It uses dynamic RAMs. Since cache memory is very expensive its capacity is kept about 2 to 3 percent of the capacity of the main memory. The main memory is connected to the cache memory. The currently needed instructions and data of the program are loaded into the cache memory from the main memory.

Semiconductor memory is a static device. It is faster, smaller in size and lighter in weight and consumes less power compared to other types of memory devices. The main memory of the computer system is volatile (cache is also volatile). It holds information (programs, data and results) as long as power supply is on. Therefore, to store programs, data, results and supervisory programs (operating system) a computer must have a permanent type of memory at 3rd level. Permanent memories such as magnetic and optical disks are used at 3rd level of memory. The memory that stores information permanently is called secondary or auxiliary memory. Magnetic and optical memories are cheaper than semiconductor memories and hence they are quite suitable for secondary memory. The secondary memory is a mass storage device. It stores operating system, compilers, assemblers, application programs, etc. Programs, data and results are saved on secondary memory. The magnetic disks and optical disks are not static. They rotate during operation. They are slower than semiconductor memory. The access time of magnetic disks is about 5-10 ms. The access time of optical disks is about 80 ms. Magnetic disks are suitable for online secondary storage. Optical disks (CD ROMs) are suitable for archival memory.

Backup memory is also needed for a computer system. It is used to store copy of important programs such as operating system, compiler, application programs, etc. These programs are usually loaded into the secondary memory but their copies are kept in backup memory so that they can be reloaded into the secondary memory in case the programs stored in the secondary memory are lost or corrupted. Floppy disks, detachable hard disks, magnetic tapes, and optical disks are suitable for backup memory.

1.12.1 Semiconductor Memory

There are two main types of semiconductor memory: RAM (Random Access Memory) and ROM (Read Only Memory). RAMs are used as **read/write** memory of a computer. They are volatile memory whereas ROMs are permanent type of memory. The brief description of different types of RAMs and ROMs are given below:

2.12.1 Random-access memory (RAM)

Random Access Memory (RAM): Under this type of memory, programs/data are stored, only as long as the power is ON. It is a read/write memory, i.e., information may be written on to, or read from it. RAM is used to store an ever-changing parade of programs and data. When you load a program into the computer or create a word processing document, the data you enter using the keyboard is stored in random-access memory (RAM) also called main, primary, or user memory. Usually, if you turn off the computer, any programs or data stored in this memory are lost; thus RAM is said to be volatile memory.

The term random comes from the way the data in memory can be located or accessed by the computer. One way to understand random is to think of the differences between a tape cassette and an audio CD disc. To play the third song on a tape cassette, you must first advance the tape past the first two songs. This is called sequential access because you access each song in sequence. With an audio CD disc you can go directly to the track where the third song begins. This is called random access because you can randomly access songs without first advancing through the songs that precede them.

The access time is same for each and every memory location. Since information can be written into or read from RAMs, they are used as **read/write** memory of a computer system. RAM is not the correct name because ROM has also random access property. The correct name of RAM should be read/write memory. Since RAM is the established name for the read/write memory, and hence, it is usually used for it.

RAM is a volatile memory. It stores information as long as power is supplied to it. Its contents are lost when power supply is switched off or interrupted. The user has to reload his programs and data when power supply is resumed.

There are two types of RAM: **Static RAM** (SRAM) and **dynamic RAM** (DRAM). A SRAM retains the stored information as long as power supply is on. But a DRAM loses its stored information in a few milliseconds even though its power supply is on. A DRAM stores information in the form of charge on a capacitor, which leaks away in very short time. Therefore, its contents must be periodically (usually every 2 milliseconds) refreshed by restoring the capacitor charge to its full value. It requires a refreshing and control circuitry. A DRAM needs only one transistor per memory cell and hence its packing density is high. A memory cell is an electronic circuitry, which stores a binary bit 0 or 1. DRAM consumes less power. It is cheaper than static RAM. It is produced in large sizes. SRAM uses conventional flip-flop circuit consisting of two cross-coupled inverters, to store binary bit 0 or 1. A RAM memory cell must be associated with read and write facility. Six transistors are needed to form a memory cell of SRAM. Hence, its packing density is low. It consumes more power. It is faster than DRAM. It is manufactured in smaller sizes.

2.12.3 IRAM

It is an integrated RAM. With advances in VLSI technology it has become possible to integrate a dynamic RAM and its controller, and to build them on a single chip (IC).

2.12.4 SDRAM (Synchronous DRAM)

It uses the same clock rate as the CPU. As a result the memory chips remain ready to transfer data when the CPU expects them to be ready. They run at the processor memory bus without imposing wait states. DRAM is an asynchronous device.

2.12.5 SGRAM (Synchronous Graphics RAM)

It is synchronous RAM suitable for graphics applications. EDO (Extended Data Output) is a modified DRAM. In EDO memory, any memory access (including refresh) stores 256 bits into latches. The latches hold 256 bits of information, so in most programs, which are sequentially executed, the data are available without wait states.

2.12.6 Error Detection and Correction in RAMs

Error detection and correction circuitry are also used for RAMs. The simplest technique for error detection is parity checking. A parity bit is generated and written along with data. When data is read, parity bit is also read and checked. The parity checking can detect only single-bit errors. It does not detect multiple errors. To detect multibit errors Hamming codes are used. ICs such as Intel 8206 are available for detecting multibit errors using Hamming codes. The IC also corrects errors.

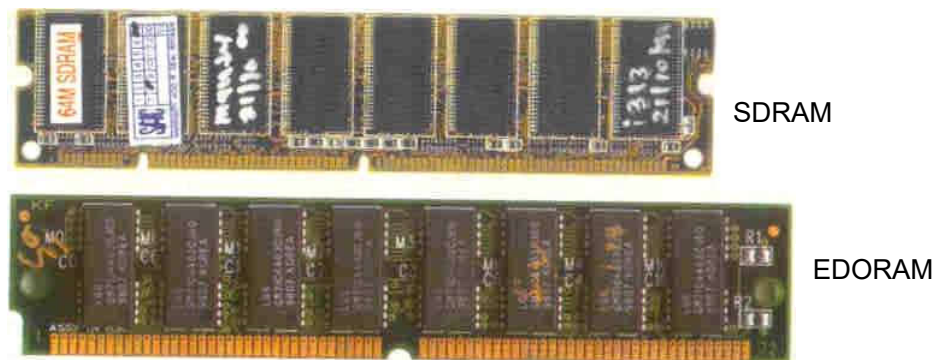


Fig. 13 EDO RAM and SDRAM

2.12.7 Memory modules

In most systems you don't add memory by adding individual chips. Instead, memory is plugged into the computer's system board using memory modules. These are circuit boards into which memory chips have been inserted

2.12.7.1 SIMM

Single In-Line Memory Module (SIMM) is a small circuit board on which DRAM memory chips are placed. This board plugs vertically into a single socket on the main system board of the computer (mother board). SIMMS of different sizes are designed to use the same size socket. For example, 1Mx8, 4Mx8 and 16Mx8 bit SIMIA all use the same 30-pin socket. The 72-pin SIMMs are often organized as 1Mx32 or 1Mx36 (with parity). Other sizes are: 2Mx32, 4Mx32 or 81Ax 32. These are also available with parity. The capacity of 4Mx36 SIMM is 16M bytes.

2.12.7.2 DIMM (Double In-Line Memory Module)

In DIMM circuit cards memory chips and contacts are placed on both sides of the printed circuit card. DIMMs are preferred in newer systems because they can be used singly instead of in pairs like the older SIMMs.

2.12.7.3 ROM (Read Only Memory)

A ROM is a static, unchanging and non-volatile memory. It stores information permanently. Its contents are not lost when its power supply is switched off. It is not accessible to user, and hence he cannot write anything into it. ROM is used to store permanent (fixed) programs. It is used to store initializing programs of a computer, microcodes of a CISC processor, supervisory programs (monitor) of a microprocessor-kit, fixed programs in microcontrollers, etc. Its contents are written at the time of its IC fabrication. It is cost-effective for high-volume production because of high cost of preparing a mask needed for writing the desired contents into the ROM. It is simple, cheap and dense. It uses one-transistor memory cell.

2.12.7.4 PROM (Programmable ROM)

A PROM is a programmable ROM. The user can write a program, data or any other kind of information permanently into a PROM. PROMS are only once programmable i.e. the user can write his program/ data into a PROM only once. A special equipment called **PROM programmer** is required for writing programs into a PROM. A PROM programmer is provided on standard microprocessor kit. **Universal PROM programmer** is also available which is more powerful and provides greater facilities, and accepts a variety of PROMS.

2.12.7.5 EPROM

An EPROM is an erasable PROM. Its contents can be erased and it can be reprogrammed more than once. To erase its contents it is exposed to high intensity short wave ultraviolet light for about 20 minutes. To facilitate exposure of ultraviolet light the EPROM chips are packed in a package, which has transparent window. The ultraviolet light having wavelength of 2537Å is used for erasing the contents of an EPROM. An

equipment with ultraviolet source called **EPROM eraser** is available for this purpose. The entire contents are erased in this process. Another disadvantage of this technique of erasure is that the EPROM chip has to be taken out of the system board and placed in an EPROM eraser's chamber. EPROMs are cheap and reliable. Hence, they are widely used. EPROMs are used to store programs, data or any other kind of information which are permanent but need updating. The programs, which are at development stage, are also stored in EPROM because they are to be modified many times. EPROMs are far more economical than PROMS because they can be reused. It uses one-transistor memory cell and hence it has high density

2.12.7.6 EEPROM or E²PROM

EEPROMs are electrically erasable PROMS. They are also called EAROM (Electrically Alterable ROM). They need not be removed from the system board for erasure. EEPROM is byte erasable. So, selective erasure of its contents is possible. Its contents can be erased and programmed on the system board itself very easily on byte-by-byte basis. Its disadvantage is that different voltages are required for erasing, writing and reading the stored information. It requires a pulse of 21 volts for writing or erasing, and 5 volts for read operation. It uses complex memory cell structure and therefore, it has low density. It has also high cost and lower reliability it is used in certain applications, where byte-alterability is needed.

2.12.7.7 Flash Memory

It is electrically erasable and reprogrammable. It is inherently non-volatile memory. Unlike EEPROM, flash memory uses one-transistor memory cell resulting in high packing density, lower cost and higher reliability. Its entire contents can be erased in *one* operation. It is not byte by byte erasable like EEPROM. Some new flash memory provides block erasure of its contents. It can be erased or reprogrammed on the system board itself. It possesses the feature of quick pulse programming. Flash memory name has been given due to fast reprogramming capability. Flash memory is suitable for storing firmware (codes) whereas conventional EEPROMs are suitable for storing parameters. When codes are updated the entire codes are rewritten to avoid errors. For updating firmware codes flash memory is quick and economical compared to an EPROM. In case of parameter updating, they are to be rewritten individually on-line byte by byte. Hence, EEPROMs are best suited for parameter updating.

2.12.7.8 Virtual memory

Very large programs and very large files of data can strain the capacity of even a large computer's memory. To solve this problem, modern operating systems and many application programs store parts of their programs or data outside RAM until needed. This type of storage is called virtual memory. On a system that uses virtual memory, only the parts of the program or data file currently needed are stored in the computer's memory; the rest are stored on a hard disk drive. When the computer needs a part of the file stored on the disk, it moves something in memory to the disk to make memory available. It then moves the programme or data from the disk into memory. For example, as you scroll through a long document, only part of it may be in RAM at any one time. The rest remains on the disk until you scroll toward it. Virtual memory reduces the overall cost of the system because it's cheaper to store data on a hard disk drive than it is to add additional memory chips to the computer.

2.12.7.9 Caches

When a computer is operating, it is frequently moving data between the CPU, disk, and memory. To process data, the data must first be moved to the CPU from wherever it is stored. Since this takes time, special techniques have been developed to speed up this operation. The basic technique is to store data in a special area of high-speed memory called a cache that is located between the CPU and main memory. When the processor needs data, it will check in high-speed cache first to see if the data is there. If not, then the processor will retrieve the data from slower main memory.

How does the computer know what to store in the cache? Basically it guesses. When you ask for something, the computer needs to go to RAM or the disk for it, it also gets the data that follows it. The assumption is that if you look at the first name in a database, the next one you want may be to look at the second name. It then moves all of this data into the cache so the CPU can get it from there rather than going to the disk for it. Unfortunately computer can't always anticipate correctly what should be in the cache. When the CPU finds the right data there, it's called a "cache hit." When it doesn't, it's called a "cache miss."

Cache memory uses special chips, often SRAM (Static RAM) chips. On some systems, these chips are four times as fast as regular memory. There are two levels of cache memory:

Level 1 cache (also called L1, internal, or primary cache) is physically a part of the microprocessor chip. This cache typically ranges up to 256 kilobytes on desktop systems.

Level 2, cache (also called L2, external, or secondary cache) is in the form of memory chips mounted in special sockets on the system board. This cache is larger than L1, ranging up to 1MB or more.

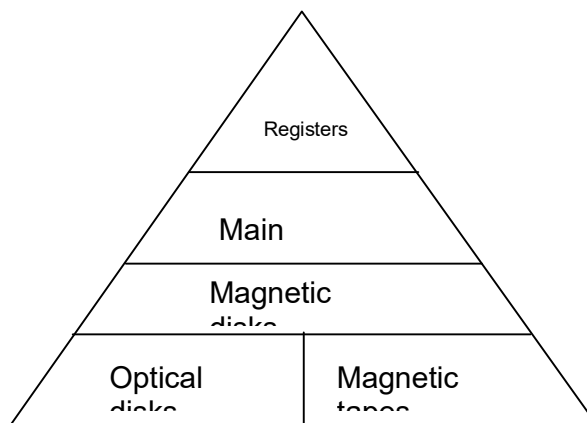


Fig.14 Memory hierarchy

2.13 Buffers

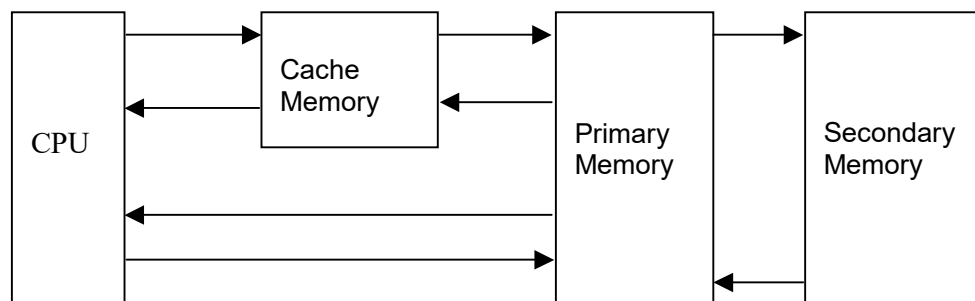


Fig.15 Connection of different types of memory

Some programs allocate a small portion of the computer's random access memory as a buffer. Buffers have several applications. For example, a buffer stores data that you enter from the keyboard until the CPU can accept it. (If the buffer becomes full, the computer beeps, and any keys you press are not stored, so you must re-enter them)

when the CPU empties the buffer). Another buffer temporarily saves a record of the most recently used commands and deleted text so you can "undo" any mistakes. Still another buffer (sometimes called a clipboard) is used as a storage area when you cut or copy data so that you can later paste it elsewhere. For example, you can create a graph in a spreadsheet program and copy it to the clipboard. You can then open a word processing program and paste the graph into a report.

2.14 Buses for Input And Output

The computer must have connections, or circuits, along which information travels among all the components. This communication path is called a bus. All communications, whether internal or external, are sent along a bus. A bus is a pathway across which data travel within computer. A Personal Computer has Processor Bus, Address Bus, Memory Bus & I/O Bus.

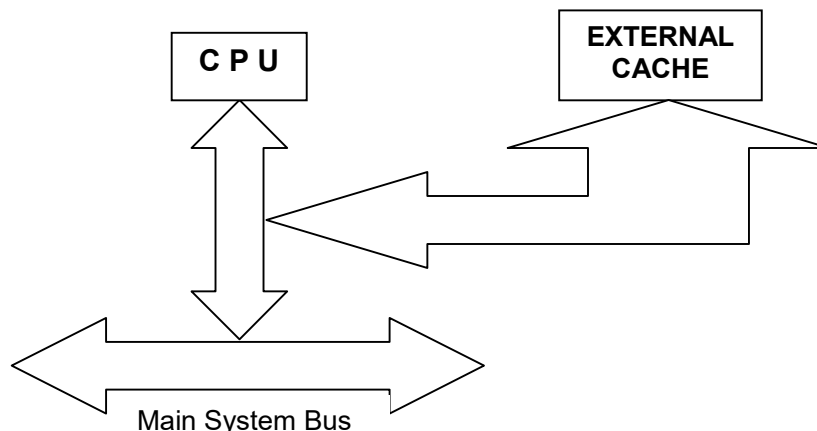


Fig.16 Processor Bus

- The Processor Bus is a communication pathway between the CPU and immediate support chips as shown in the fig. The support chips are called CHIPSET.
- The Memory Bus is used to transfer information between the CPU and main memory (RAM) as shown in the fig.

2.15 The bus

The bus in a computer is like a high way for the data. A bus is a set of copper connections on the computer's system board. The concept behind a bus is relatively simple. There are only two ways to connect all the components: Either run wires between each component in the system or connect them all to a common set of wires a bus. The bus simplifies the design, reduces the number of required circuits, and lowers the cost. A bus also allows new components to be added by simply plugging them into the bus.

2.15.1 System buses

The system bus is simply a set of connections that connect the CPU to other devices in the computer. For example, any peripheral connected to the bus with an adapter card is connected to the CPU. The system bus includes address line and data lines. The address lines carry the address of data in memory. The data lines carry the data bits themselves. The number of data lines determines the width of the bus. Think of the data bus as a set of parallel wires, each of which can send one bit at a time. The more wires in the bus, the more bits that can be sent at a time. The number has increased over the years from 8, to 16, and now 32 to keep up with the performance of microprocessors. Some high-end systems are already using the next step bus, one that carries 64 bits.

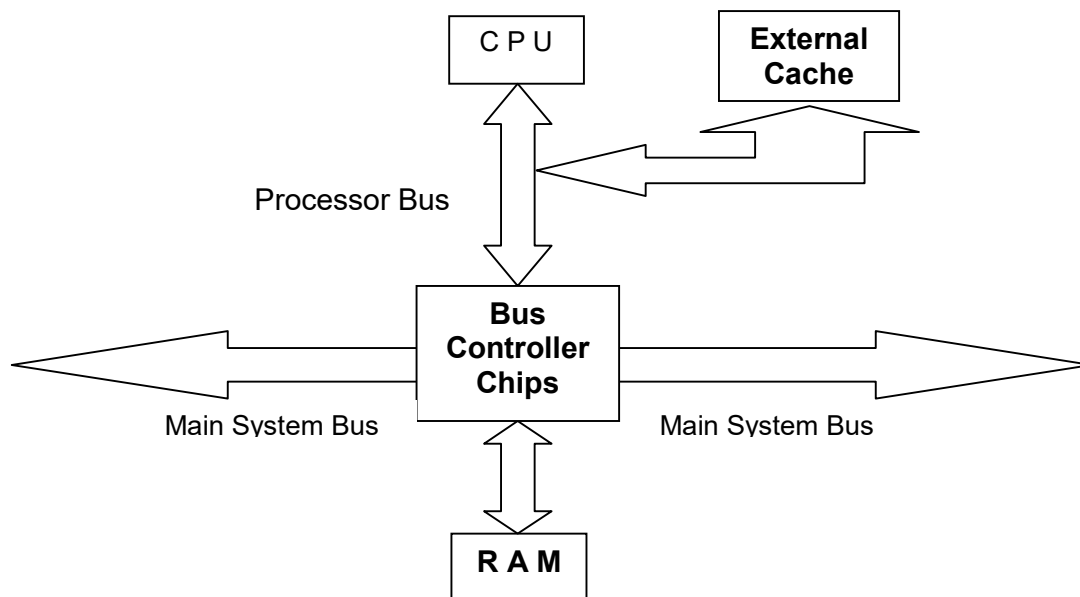


Fig.17 Memory Bus

2.15.2 Local buses

As graphical user interfaces such as Windows have become more popular, parts of the computer system have become bottlenecks to acceptable performance. One of these bottlenecks has been updating the screen display. This is no small task. For example, a high-end display may have 1024 pixels across the screen by 768 down, or a total of almost 800,000 pixels that have to be updated 60 or more times per second. This puts a tremendous load on the system and you notice it when it takes a long time for the screen to be redrawn when you page up or down. To solve the problem, local buses have been added to take peripherals off the I/O bus and connect them, together with the CPU and memory, to a wider data pathway. This speeds up how fast data can be sent from memory through the microprocessor to the screen or other devices.

Almost all new desktop computers use a fast and inexpensive Peripheral Component Interconnect (PCI) bus. PCI was invented by Intel but is now controlled by an industry

wide standards committee. Like everything else in computing, PCI is evolving quickly. As microprocessors speed up, it's essential that other components of the system keep up the pace so they aren't bottleneck. The PCI bus is constantly evolving to meet this challenge. It gets faster in one of two ways: its speed increases or its data path widens. This is like raising the speed limit on a freeway or adding additional lanes. The bus is now fast enough for computers to process multiple video streams and simultaneously support very high-speed networking.

Older computers used first ISA buses (Industry Standard Architecture) and then EISA buses (Extended Industry Standard Architecture). The first carried 16 bits at 8MHz and later 32 bits at 10MHz. Both were abandoned as microprocessor speeds increased and they could no longer keep up. For example, where PCI can send up to 256MBps, ISA can send only 5MBps. The local bus uses a new set of connections in line with the normal bus for a direct path to the CPU.

2.16 Different Expansion Cards

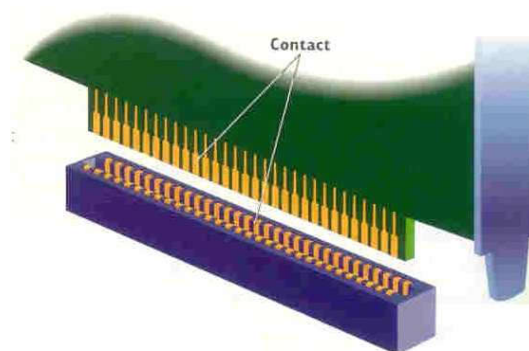


Fig. 17 8-BitExpansion Card

8-Bit Expansion card Data is transmitted among expansion slots and other components on the bus only along 8 parallel data lines. The data lines use only a fraction of the 31 pairs of connectors that fit into the expansion slot. As remains the same with never boards, the other connectors supply the boards, the other connectors supply the board with power, instructions, and addresses for data locations either on the expansion boards or in memory. This was the only type of expansion slot in the original IBM PC.

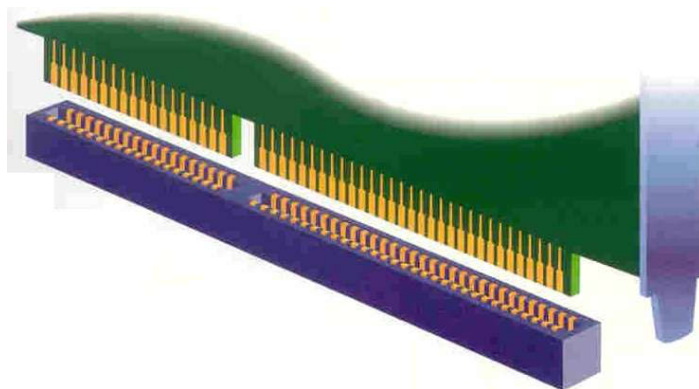


Fig. 18 16-Bit or ISA

16-Bit or ISA card With 18 more pairs of connectors, the ISA (Industry Standard Architecture) card transmits data over 16 data lines, doubling the amount of data it moves compared to an 8-bit card. Today, PCs with faster and never local bus slots still have ISA expansion slots. A 16-bit card is powerful enough for components, such as serial and parallel ports, and internal modems that don't handle the extreme amounts of data transmitted by video, net-work, and disk controller cards.

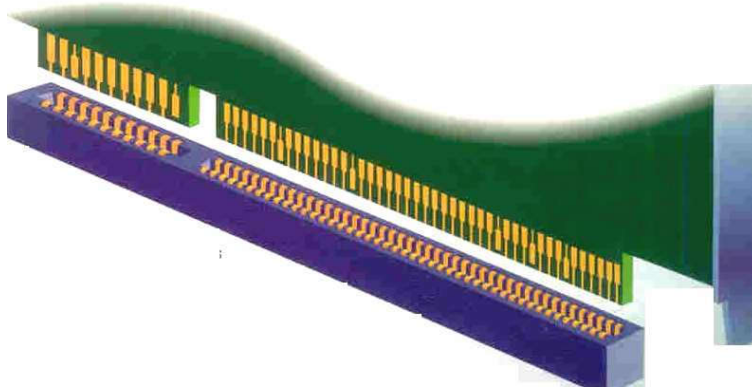


Fig.19 32-Bit MCA Card

32-Bit MCA Card The IBM Micro Channel Architecture (MCA) card uses 32 of its 93 lines to send and receive data. It also includes special circuitry that, like plug and play technology, makes the card easier to install. The MCA expansion slot, which IBM refused to let others clone for a long time, will not accept 8-bit or adapter cards.

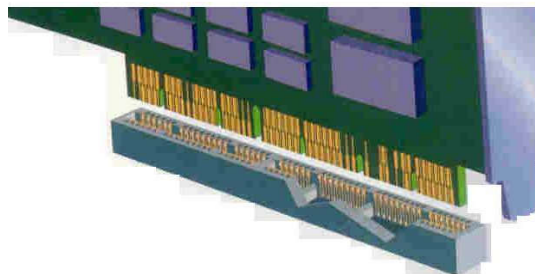


Fig. 20 32-Bit EISA Card

32- Bit EISA Card The design of Extended industry standard Architecture (EISA) can use expansion cards designed specifically to work with the slot's 97 connectors that are

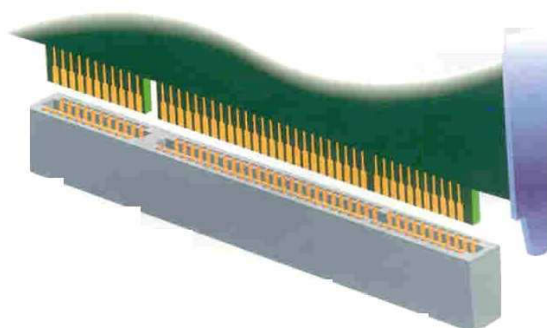


Fig.21 32- Bit PCI Local-Bus Card

divided between two levels. These EISA-specific cards transmit 32 bits of data at a time, and, like MCA and plug and play, are easier to set up. But EISA slots also accept 8-bit and 16-bit cards. Plastic tabs allow the older cards to fit only far enough into the slots to make contact with the first level of connectors, which work the same as ISA connections. But boards designed specifically for the EISA slot can enter farther and align their connectors with the lower row of connectors that handle signals based on EISA specifications.

32-Bit PCI Local-Bus Card The peripheral Component Interconnect (PCI) local-bus adapters have connectors similar to those on MCA and EISA cards. All handle 32 bits of data at a time, and are smaller and more tightly packed than ISA connections. PCI slots won't accept ISA or 8-bit cards. Closely integrated with plug and play, PCI slots currently are the most common in PCs.

32-Bit Accelerated Graphics Port Despite being called a port, the AGP is an expansion slot, technically an adaptation of the PCI slot design. Unlike other expansion slots that can handle a variety of add-in cards, the AGP is designed only for a special type of video card. The 44 pins on an AGP card are arranged in a fashion similar to those on an EISA board, interweaving every other connector so that some touch slot connectors along a bottom row and remaining pins come in contact with top row of slot connector.

2.17 Adapter cards And Multimedia Systems

Many users like to customize the computers to better serve the needs. To make this possible, most computers have expansion slot inside the cabinet into which you can plug adapter cards (also called add-in boards or cards). These expansion slots are connected directly to the computer's bus, so cards plugged into them perform just as if they were built into the computer. Adapter cards can serve variety of functions. Some allow the computer to display colours or graphics, send faxes, connect with the Internet, or add sound. Still others connect peripherals or are the peripherals themselves. For example, you can plug in an adapter card that controls a hard disk drive located elsewhere in the system, or you can plug in a card that contains the hard disk drive itself.

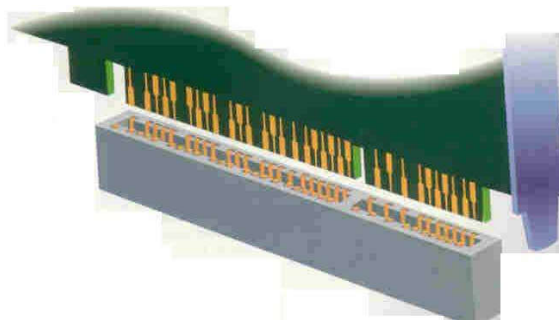


Fig. 22 32-Bit Accelerated Graphics Port

2.18 Plug and play

When you want to change equipment on your system, such as plugging a modem into a notebook computer, you often have to first turn it off. If you are adding a new component, such as a fax card, you might even have to configure it. To eliminate these

problems, some operating systems support plug and play. Plug and play means the operating system automatically detects new hardware and automatically configures it to run on your system. It does this by loading and unloading device drivers as you connect and disconnect peripherals such as printers, modems, and network cards. When you plug one of these into the computer while it's running, the operating system detects the new device. It then loads the required driver without rebooting the system. Plug and play requires three components that support the technology: an operating system, a BIOS, and devices.

2.19 Multimedia systems

At the current time, one widespread use of expansion slots is to make PCs into multimedia systems capable of displaying high-quality graphics, animations, and sound. These capabilities were originally used for games but are increasingly used for education, business presentations, and demonstrations. To upgrade a PC for multimedia, the system must have added to it a sound card, a CDROM drive, and an optional graphics accelerator. All of these components are added using cards that slide into expansion slots in the system's bus.

- **Graphics accelerator:** Multimedia requires a powerful system capable of showing high-quality illustrations, movies, and animations on the screen. This isn't an easy task because of the amount of information that is sent to the screen to achieve these effects. To speed up the screen display so it isn't a drag on the system, a graphics accelerator card can be inserted into a slot and the display screen is then plugged into that card. These cards have their own processors that take much of the graphics load off the system's CPU.
- **CD-ROM drive:** Multimedia uses lots of graphics, sound, animation, and movies. All of these take up huge amounts of storage space. To get all of these files onto the same disk, CD-ROM disks are used because they have much more capacity at lower costs than other media. To play these disks, the system must have a CD-ROM drive. To connect the drive to the system, you must plug an adapter card into one of the expansion slots.
- **Sound card:** To get high-quality-spoken sound from a system, you insert a sound card into one of the slots and then plug speakers into the card.