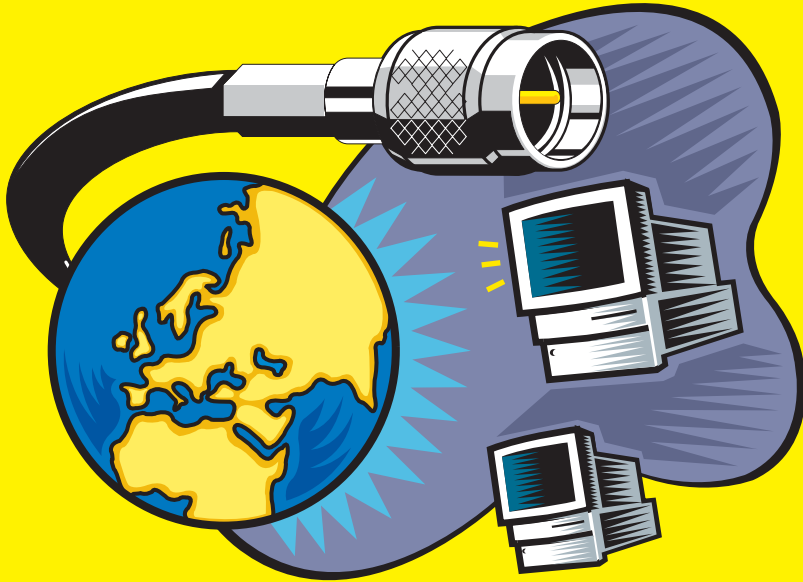


Everything you wanted to know about.....

Computer Hardware & Peripherals



**Munishwar Gulati
Mini Gulati**

Learn

about Computer Hardware

Apply

easy steps for fast results!

Build

foundation of Hardware

Master

assembling & troubleshooting

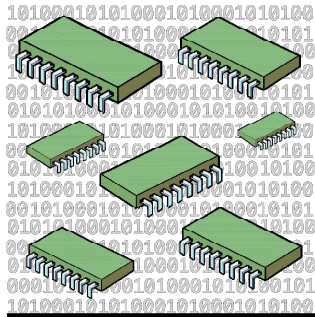


Computer Hardware & Peripherals

MUNISHWAR GULATI
Mini GULATI



SILICON MEDIA PRESS
NEW DELHI



CHAPTER 1

INSIDE PC

SOFTWARE
HARDWARE
CENTRAL PROCESSING UNIT (CPU)
THE PC FAMILY
SERVERS
DISK ARRAYS AND DATA STRIPING

Inside PC

Computers have become an integral part of man's everyday life. They continue to grow and open new horizons of discovery and application. It must be remembered that computers are machines created and managed by man. A computer has no brain of its own. Anything it does is the result of human instructions - supermarket scanners calculate our grocery bill while keeping store inventory; computerized telephone switching centers play traffic cop to millions of calls and keep lines of communication untangled; and automatic teller machines (ATM) let us conduct banking transactions from virtually anywhere in the world. **It is an obedient slave which carries out the master's orders as long as it can understand them no matter whether they are right or wrong.** In other words, computers lack common sense.

Modern computers possess certain characteristics and abilities peculiar to them. The application areas of computers are unlimited. We find a computer in every aspect of our life. From a simple operation as playing a video game to more complicated applications as weather forecasting, computers are found everywhere. Let us take a simple example of a person who needs to purchase a can of juice from a super market. He walks inside the super market, picks up a can of juice and proceeds to the cash counter. The counter person scans the code that is present in the label to generate a bill. This scanning of the code is computerized. The man pays his bills with his credit card and walks off the super market. He just used a computer, which will transfer the cost of the can of juice from his bank account to the super market. The Computers can:

- Perform complex and repetitive calculations rapidly and accurately.
- Store large amounts of data and information for subsequent manipulations.
- Make decisions.
- Provide information to the user.
- Automatically correct or modify-by providing signals -certain parameters of a system under control.

All the equipment and personnel connected with computers are divided in four parts.

- Hardware
- Software
- Firm ware
- Human ware

Here, in the book, the stress will be on Hardware of the Computer. The other three terms are being discussed in brief:

SOFTWARE

Computer need clear-cut instruction to tell them what to do, how to do, and when to do. A set of instructions to carry out these functions is called a program. A group of such programs, that are put into a computer to operate and control its activities, are called Software. These programs must reside in the internal storage (memory) to execute their instructions. For example, if we want to delete some data stored in the memory, the system uses one set of program instructions. Similarly, if we want to sort a list of names, it uses another set of instructions designed to perform the task.

Software is defined as set of instructions stored as programs that govern the operations of a computer system and make the hardware run. This is not something that can be touched, but which can be seen on computer when computer is operated. Software can be classified as :

- a) System Software
- b) Application Software

Software is an essential requirement of computer systems. As a car can't run without fuel, a computer can't work without Software. There are four kinds of software that are implemented -

- Operating system
- Utility program,
- Language processors,
- Application programs

Operating System

A Computer is a collection of parts, some mechanical, some electrical and some electronics. Getting all these parts to work together is a major problem, and that is basically what an operating system does. It manages, the resources of a Computer System and schedules its operation. It provides a link between the user and the computer hardware and software.

Operating system controls nearly every aspect of the computer system. It transfers the control of computer to user and supplies the way of communicating with the computer. Thus it provides the interface between user, and Computer. The principal functions of operating system include :

- To control and coordinate peripheral devices such as printers, display screen and disk drives.
- To monitor the use of machine's resources.
- To help the application programs and execute instructions.
- To help the user to develop programs.
- To deal with any faults that may occur in the computer and inform the user/operator.

Utility Program

There are many tasks common to a variety of applications. Examples of such tasks are :

- Sorting a list in a desired sequence
- Merging of two programs.
- Copying a program from one place to another.

One need not write programs for these task. They are standard, and normally handled by utility programs.

Ex. SORT.EXE is a utility of DOS
VSAFE.EXE is virus safe utility of DOS.

Language Processors

Computers can understand instructions only when they are written in their own language, called the machine language. Therefore, program written in any other language should be translated into machine language. Special programs called language processors are available to do this job. It could be Assembler Interpreter or Compiler.

These special programs accept the user programs and check each statements and, if it is grammatically correct, produce a corresponding set of machine code instructions.

Assembler - A program which translates an assembly language program into a machine language program is called an assembler.

Interpreter - Converts a high level language (like BASIC FORTRAN etc.) program to machine language line by line and executes each line simultaneously.

Compiler - It converts the High level language program to machine language in one go and reports all errors along with the line number. All the errors must be removed to successfully run the program.

Application Program

While an operating system makes the hardware run properly, application programs make the hardware do useful work. Application programs are specially prepared to do certain specific tasks. They can be classified into two categories :

- Standard Application.
- Unique Application.

Some applications are common for many organizations. These are standard software packages and are available from hardware or software vendors.

There are situations when one may develop one's own programs to suit the requirements. Once developed, they are considered as unique applications.

MS-Word, MS-Excel, WordStar, Lotus, dBASE are few application which are being widely used.

FIRMWARE

Firmware is the name given to a non-erasable program permanently in-built in computer's memory (ROM) that controls internal functioning of the computer.

HUMANWARE

is group name of all the personnel concerned with installation, operation, maintenance and testing of entire computer equipment. This includes all programmers, system analysts, data typists, hardware maintenance engineer, computer operator etc.

HARDWARE

It is general term used to represent the physical and tangible components of the computer i.e. those components which can be seen and touched. It includes :

- i) Input devices
- ii) Output Devices
- iii) Central Processing Unit
- iv) Backing storage & auxiliary memory devices.

The electronic circuits consisting of resistors, capacitors, ICs main computer equipment (visually the keyboard, visual display unit and central processing unit) and its peripherals like floppy drives, joystick, printer, tracker ball, microfiche, graphic pad, plotters, digitizer, card reader, paper and magnetic tapes and disks, modems, interfaces terminals etc. are all examples of computer hardware. All input and output devices connected to computer are collectively known as peripherals.

Basically, any computer is supposed to carry out four functions:

- Accepting the data as input.
- Storing the data and instructions in its memory and recalling the same as and when required (Retrieving).
- Processing the data as per instructions, to convert it into useful information.
- Communicating the Information as output.

The input data may be in the form of :-

Typed data and instructions in computer's language. Raw data can be in form of (a) Numeric, (b) Alphabetic, (c) Alphanumeric. Then this is coded into data acceptable to computer in its language.

It may be audible, visual or audiovisual. In these cases an interfacing device is required to convert it into signals which are acceptable to computer.

All data is first converted into binary signals before it is accepted by computer. The first function is performed in computer by its input unit, second and third functions are carried out by its central processing unit and last function is done by output unit. Since CPU is not able to keep everything in memory, this function is more better done by auxiliary memory. Thus a computer has got four main units.

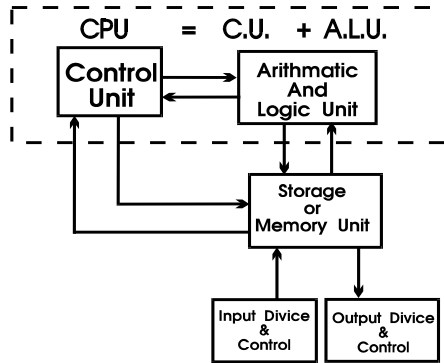


Fig 1.1 Diagrammatic representation of Computer

- Central Processing unit
- Input devices
- Output devices
- Primary memory (Immediate Access storage)

In addition to this, there is one more group of devices known as auxiliary memory or backing storage, in which bulk of information can be stored while inputting or outputting the data into computer or out of it.

Above four main units combined with auxiliary memory systems are known as Peripheral devices.

CENTRAL PROCESSING UNIT (CPU)

CPU is totally electronic based and its function are to:

- receive and store data into main storage memory
- control sequence of operations,
- give commands to all parts of the computer system,
- carry out data processing and to output the results.

The CPU assumes many different sizes in different computers. The CPUs of mainframe computers are like large cupboards and carry out

millions of instructions in second. The CPU of a microcomputer is much less powerful and is contained in a single chip. We shall learn about chip little later.

This is actually the brain cum heart of the computer. Data is processed here. It takes information from the input unit and processes it according to instructions. The instructions are given by programs. Programs are written in the languages which computer can understand. These are actually coded languages.

By data processing we mean:

- making arithmetic calculations like addition subtraction, multiplication, division and exponentiation etc.
- making logical decisions like comparing two values to find out which one is greater.
- manipulating alphabetic or alphanumeric data like word processing, letter writing, sorting in alphabetic or alphanumeric orders, editing, making catalogues etc.

Basically the CPU does everything in binary language, i.e., with the help of the digits 0 and 1. Actually the digit 1 means presence of electronic pulse and 0 means its absence.

CPU has three main units

- Arithmetic and Logic Unit (ALU)
- Control Unit,
- Memory Storage Unit.

In addition to above it has many other auxiliary units

- Registers
 - (a) Data Register
 - (b) Instruction Register
 - (c) Address register
- Instruction (program) counter,
- Instruction decoder,
- Clock

ARITHMETIC AND LOGIC UNIT (ALU)

The ALU (Arithmetic and Logic Unit) is a section where all the arithmetic and logical functions are carried out. We know that every arithmetic operation requires two operands. The operation then produces a result. This ALU, therefore, must be able to handle two data words and the result. This concept is dependent on the storage and processing techniques, which may vary, from model to model. Logical operations usually involve comparisons. Circuits in the ALU compare two values by subtracting one from another. For example, to determine whether number *A* is greater than number *B*, *B* is subtracted from *A* and the sign of the result is determined. Depending on the sign

of the result, the ALU does further operations. Alphabetic data can also be compared according to their assigned order sequence.

ALU comprises of a number of accumulators and registers. It derives all the data from the main memory as directed by control unit based on program given to it. This data gets loaded into accumulators in the ALU. For example if numbers A and B are to be added, the control unit selects number A from its address in memory and loads it into arithmetic unit. Then the value of B is selected and added to A in the arithmetic unit. The results are placed in the main storage (again under the direction of control unit) or may be retained in the accumulators themselves for further calculations. Some logical operations of the ALU enable the computer to have the decision making quality.

CONTROL UNIT

The Control unit directs all operations inside the computer. It is known as nerve centre of the computer, because it controls and coordinates all hardware operation, i.e., those of the CPU and input-output devices. It performs following actions:

It gives command to transfer data from the input device to the memory (IAS) of arithmetic and logic unit. It also transfers the results from ALU to the memory and on to the output device for printing.

It stores the program in the memory, takes instructions one by one, understands them and issues appropriate commands to the other units.

It fetches the required instructions from the main storage via memory data register (MDR) and places it in the control instruction register (CIR). Then it interprets the instructions in CIR and gets it executed by sending a command signal to the concerned hardware device. Memory data register is a special register which holds all data and data instructions temporarily, as they pass in and out of the main memory. Control instruction register is also a special register which holds machine instructions currently being interpreted by control unit.

MEMORY

Digital computers use the binary number system, which has two digits: 0 and 1. A binary digit is called a bit. Information is represented in digital computers in groups of bits.

The bit is an elementary unit of Memory.

A group of 8 bit is called a byte and a group of 4 bit is called a nibble.

One byte is the smallest unit which represents data. Other bigger units of memory are KB, MB and GB which has following relation:

1 KB = 1024 Byte (2^{10} byte)

1 MB = 1024 KB = 1024 x 1024 bytes

1 GB = 1024 MB = 1024 x 1024 x 1024 byte

The function of the memory is to store information. It stores program, data, results or any other kind of information. Two or three levels of memories such as main memory, secondary memory and cache memory are provided in a digital computer.

Main Memory or Primary Memory

Main memory or primary memory is a fast memory. It stores programs along with data, which are to be executed. It also stores necessary programs of system software, which are required to execute the users program. The main memory is directly addressed by the CPU. Semiconductor memories, RAMs are used as main memory. It possess random access property, and has smaller access time, 80-100 nsec.

Secondary Memory or Auxiliary Memory

It stores operating system, data files, compilers, assemblers, application programs etc. It is important to note here that CPU does not read information (residing in the secondary memory) directly from the secondary memory. The programs and data (residing in secondary memory), if needed by CPU, are first transferred from the secondary memory to the primary memory. Then the CPU reads them from the primary memory. The secondary memory is a mass storage memory. It is slow but cheap and permanent while the main memory (RAM) is volatile. Hard disks are used as secondary memory. Their access time is about 20 milliseconds.

Cache Memory

The word Cache is pronounced as cash. The cache memory is placed in between CPU and main memory. It is much faster than main memory. Its access time is much less compared to that of main memory. The access time of a Cache memory is 15-25 nanoseconds whereas that of the main memory is 80 ns. One nanosecond is $=10^{-9}$ sec.

The cache memory is not accessible to users. It stores instructions and data which are to be immediately executed. It is used to reduce the average access time for address, instructions or data which are normally stored in the main memory. Thus the cache memory increases the operating speed of the system. Cache memory is much costlier than main memory.

The types of memory that are used in a computer is summarized in the figure below:-

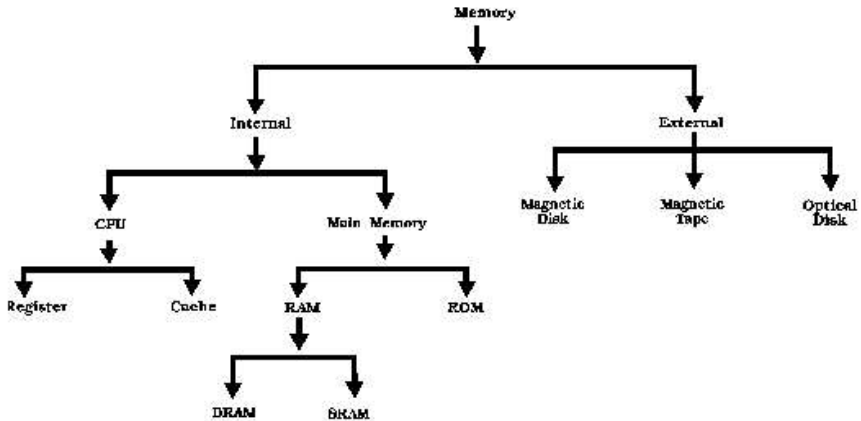


Fig 1.2 Various Type of Memory

INPUT/OUTPUT BUS

The IO bus is the interconnection path between the processor and input/output devices (including memory). The bus is divided into THREE main sections

Address

The address bus is used by the processor to select a specific memory location. This memory location may be in the memory subsystem (either RAM or ROM), or a peripheral device. The address bus is one way only (unidirectional).

Data

The data bus is used to transfer data between the processor and memory or peripheral devices. The data bus is two-way (read/write, bi-directional).

Control

The control bus is like the traffic signals. It provides timing, clock, and directional signals for each operation. Most of these signals are generated by the processor, as the processor generally controls the read or write operation.

In more complex systems, the memory subsystem or peripheral devices also provide timing signals to complete data transfers, or initiate requests that the processor responds to (called interrupts).

INPUT/OUTPUT PERIPHERAL DEVICES

Peripheral devices allow input and output to occur. Examples of peripheral devices are

- disk drive controllers
- keyboards
- mouse
- video cards
- parallel and serial cards
- real-time clocks

The processor is involved in the initialisation and servicing of these peripheral devices.

Input/Output Processors

An input output processor is a special processor dedicated to handling peripheral devices like terminals, tape and disk units, and printers.

Mainframe systems use I/O processors to off load work from the system processor. This lets the system processor get more work done executing user programs without having to worry about handling data input and output to terminals or printing documents.

The PC has an I/O processor in the keyboard, which handles the complex operations of scanning the keys.

In addition, it is now becoming common to have I/O processors on graphics cards. The S3 graphics card is a good example of this, which supports hardware support for scrolling, sizing and moving windows. This removes these tasks from the system processor, and performs them at a much higher rate (up to 30 times faster).

Many video cards offer added functions, such as accelerated rendering of 3D scenes, video capture, TV tuner adapter, MPEG-2 and MPEG-4 decoding, FireWire, light pen, TV output, or the ability to connect multiple monitors. The memory capacity of most modern video cards range from 128 MB to 4 GB, though very few cards actually go over 1 GB.

MOORE'S LAW FOR MICROPROCESSORS

Moore's law describes a long-term trend in the history of computing hardware. Since the invention of the integrated circuit in 1958, the number of transistors that can be placed inexpensively on an integrated circuit has increased exponentially, doubling approximately every two years. The trend was first observed by Intel co-founder Gordon E. Moore in a 1965 paper. It has continued for almost half a century and in 2005 was not expected to stop for another decade at least.

Almost every measure of the capabilities of digital electronic devices is strongly linked to Moore's law: processing speed, memory capacity, sensors and even the number and size of pixels in digital cameras. All of these are improving at (roughly) exponential rates as well. This has dramatically increased the usefulness of digital electronics in nearly every segment of the world economy. Moore's law describes this

driving force of technological and social change in the late 20th and early 21st centuries.

Moore's law has been discussed in detail in Chapter 3 : Microprocessors.

EVOLUTION OF COMPUTERS

There are three basic computer platforms (type of computing systems) that you will be working with. There is the IBM PC and Clones, the Apple Macintosh and the DEC VAX system.

The following is the History of Computer evolution. This history is indicative rather than accurate-

1978-79: First widely available Personal Computers - Apple II & Tandy TRS80 - about 2 MHz clock speed, cassette or 40 - 80K 5.25" floppy, Basic in ROM. Monochrome, no graphics. Custom operating system.

CP/M - portable 8 bit operating system, used by a wide variety of early personal computers (1979 - 1984), indirect forerunner of DOS. Limited to a maximum of 64K RAM, typically 8 - 16K in earlier machines rising to the 64K limit later.

Apple Lisa, based on Xerox Parc research and Motorola 16 bit chip, was forerunner of Macintosh line but far too expensive to compete with PCs because it required lots of (expensive) RAM to handle graphical user interface.

IBM PC 1982 - 16 bit (sort of) 4.7 MHz, 128K - 256K RAM, maximum 640K RAM. 360K floppy, Microsoft DOS. Very limited graphics (320x240x4 colours). Open specification, soon generated clones. The term PC came to be synonymous with IBM and IBM-compatible computers.

Initial release of Macintosh with 128K RAM (inadequate), 256K just adequate, which brought Apple into the market as a possible contender. The technology was far ahead of the IBM PC and clones, but the need for adequate RAM and Apple's proprietary stance led to a smaller market share and a majority of software development occurred on PCs. This led in turn to a much wider range of software development tools for the PC market, and many specialised software packages are only available for PCs (e.g. Idrisi, BASP, MV-Arch).

Apple also faced the problem of overcoming its image as a manufacturer of games machines - Apple survived, while two other contenders with very good machines - Atari and another company - failed to break into the business market even though they were also technically far ahead of IBM.

IBM AT 1984 - 286 chip, 1.2Mb 5.25" floppy. 8MHz clock speed, EGA (640x480x16 colours). Could handle extra memory to 16MBytes

through extended and expanded memory schemes. Still using character-based DOS operating system.

386 (SX & DX), 486 (SX & DX), Pentium, Pentium Pro, Pentium II = rapid improvements in performance, memory addressing and capabilities. By 1988 chips were pushing into 20-33 MHz 386 and Windows (Version 3) was becoming the standard operating system - very primitive compared with the well-developed Mac operating system.

Widespread networking of PCs developed in the early 1990s - hard to get working under Windows Version 3, where it was pretty much an afterthought, but much easier with Windows 95 and NT which have good support built in. Rapid development of Internet connectivity and dial-up connections.

In early 1998, entry level systems are 133 MHz Pentium and faster systems 266 - 300 MHz Pentium II, with Windows 95/98 as the standard end-user operating system and Windows NT for power users and servers.

It was followed by Pentium II, 233 -333 MHz/ AMD K6-2, Memory - 32 mb of RAM, Hard Drive of 1 -10 GB, VGA display with 28.8 or higher baud rate modem with Windows ME as end-user operating system.

Next came Pentium III 500 MHz - 733 Mhz/ AMD K6-3/ Intel Celeron processor, 64 -256 megabytes (MB) of RAM, 10-20 GB of Harddisk, DVD drives, optical mouse, SVGA display and sound cards with Windows ME/2000.

In year 2002, Itanium 64 bit 733 Mhz/ AMD Athelon with 1 MB internal Cache Memory, 512 MB Ram, 20-40 GB Harddisk, DVD Writers, Optical Mouse, Graphics and Sound Cards with Windows XP became the end-user operating system.

From 2003 to 2005, Pentium 900 MHz - 2.2GHz/ AMD Athelon 64 with 1 MB internal Cache Memory, 40-80 GB Harddisk, Sound and Graphic Cards, Webcams with Windows XP OS was the general configuration of the Computers.

In Year 2000, Computer were having the eighth generation microprocessor Intel Core Duo, Core 2 Duo and Core 2 Quad / AMD Opteron, AMD Phenom Technology with RAM hovering around 8-16 GB, Hard disks upto 500 GB, DVD Writers and Blue Ray Disc and options to attach multiple gadgets from Cell phones to Video Cameras.

Thus, from a humble 1 MHz (Mega Hertz or Millions of Cycles per second) Personal Computers PC 8088, we have reached to almost 2.5 GHz 64 Bit processors.

The various Processor and speed of Processor is given below :

	Processor	Speed of Processor
■	PC 8088	1-4 MHz

■	PC XT 8088	8-16 MHz
■	PC AT 286	16-20 MHz
■	PC AT 386	33-40 MHz
■	PC AT 486	33-66 MHz
■	Pentium -I /AMD PRO	66- 233 MHz
■	Pentium -II/AMD - K6-2/3	233-600 MHz
■	Pentium - III / AMD Athlon	400MHz - 1GHz
■	Pentium - IV/ AMD Athlon XP	1.1 GHz - 2.2 GHz
■	Itanium/ AMD Athlon 64	1.3 GHz - 1.5 GHz
■	Pentium M	900 MHz - 2.26 GHz
■	Pentium D / AMD Opteron	2.8 GHz - 3.4 GHz
■	Core Duo/ AMD Athlon 64 X2	1.5 GHz - 2.33 GHz
■	Core 2 Duo / AMD Opteron FX	1.8 GHz - 3 GHz

For the last 20 years, performance increase in PCs have been totally driven by hardware, with the software doing it's best to slow it down. Under DOS, most PC hardware was overkill. The bus was terrible, but it didn't matter because the CPU did all of the work - there was no task switching. OS/2 and Windows changed that.

The trend over the last few years has been to put good PC hardware into midrange machines, while PCs always opt for lesser hardware that is marginally less expensive. I/O has traditionally been a place to cut corners. It is normal for machines in the PC world to sell based solely on the CPU clock, which is almost always the most expensive component in the machine (after the software).

In most cases, these machines would benefit from improved I/O in lieu of the higher CPU clock. There is a higher margin on CPU chips, apparently.

PCI

Peripheral Component Interconnect (PCI) has become the standard system bus almost overnight. It can be found in everything from high end UNIX workstations down to routers and embedded control systems. It is documented, new, and at up to 64 bit at 64 MHz, it is quite fast.

It is an edge connected card, a trait that was undesirable in the past because of vibration. The interfaces are far easier to configure than their ISA counterparts, as there is either a BIOS or software routine that configures them at boot time.

IDE

The standard disk in PCs is the Integrated Drive Electronics, or IDE disk. It is cheap, and that is the reason for its popularity. It will never move out of the low end client, though other architectures are beginning to support it.

IDE moves the logic previously found in MFM drive controllers and moves it into the disk itself. The drives originally found in PCs had no intelligence, and had to be forced to seek to the appropriate head, cylinder and track. IDE disk could solve this geometry themselves, and do a few more advanced things such as bad block relocation.

IDE disks cannot have multiple pending requests, nor can more than one disk use the bus, nor can more than two disks be on the same bus. This is a huge performance hit. Each read request must be completely done before the next one can even leave the CPU.

There have been efforts to alleviate these problems, mainly involving speeding up the transfers. This is arguably a flawed approach; doubling the speed of transfer will at best double throughput, and possible a small improvement in latency. Using the bus to its capacity would improve both latency and throughput dramatically.

Ultra DMA IDE disks doubled the transfer rate of ATA disks. But tests have shown that the realized performance improvement is below 2% for these new drives.

SATA

Serial ATA (Serial Advanced Technology Attachment or SATA) is a new standard for connecting hard drives into computer systems. As its name implies, SATA is based on serial signaling technology, unlike current IDE (Integrated Drive Electronics) hard drives that use parallel signaling.

SATA has several practical advantages over the parallel signaling (also called Parallel ATA or PATA) that has been used in hard drives since the 1980s. SATA cables are more flexible, thinner, and less massive than the ribbon cables required for conventional PATA hard drives. SATA cables can be considerably longer than PATA ribbon cables, allowing the designer more latitude in the physical layout of a system. Because there are fewer conductors (only 7 in SATA as compared with 40 in PATA), crosstalk and electromagnetic interference (EMI) are less likely to be troublesome. The signal voltage is much lower as well (250 mV for SATA as compared with 5 V for PATA).

INTERRUPTS

PCs traditionally have had interrupts hard wired to devices. In general, it is difficult to change them when adding devices. There were originally 8 interrupts in the IBM PC. They were all taken by devices in the system. When the AT was developed, a second interrupt controller was added, cascaded from the first one. There were then 15 interrupts available, but they were notoriously difficult to assign.

A PC interrupt is vectored, that is, the interrupt controller knows where the interrupt came from, so there is no CPU time wasted polling devices. There is added complexity to support these interrupts, and in the end they probably increase costs while decreasing performance.

An interrupt is asserted over the ISA bus where it is masked by the interrupt controller. The interrupt controller asserts a pin on the CPU, which then causes the OS to trap to the appropriate interrupt handler. With DOS, this means disabling interrupts, pushing registers, doing everything, and then popping registers enabling interrupts and returning.

Newer operating systems on PC hardware make use of bottom half handlers, so that not as much has to be done while interrupts are disabled. This is a good idea, since another interrupt could happen while we are processing one interrupt. It would be undesirable to miss it.

MULTIPROCESSORS

Multiprocessors on x86 hardware are a complicated science. The hardware has evolved from a simple chip with complicated addressing to a huge chip with even more complicated addressing. Making matters worse, its functionality isn't documented incredibly well (making processor bugs easy to sneak in).

Interrupts again are a big problem here. When an interrupt happens, which CPU should handle it? Putting several hundred mutex locks in the OS code is necessary to have an MT system that will scale. This is not a trivial undertaking, and could easily take *years* to complete. Multiprocessor support is by no means done for x86 CPUs.

Putting a lock around the whole OS is another way to have multiple processors. One CPU will handle an interrupt at a time. This means that compositionally expensive jobs will run with linear improvement, but I/O bound ones will not scale past 1 processor.

DMA

DMA on PCs is also available, but is used primarily for audio hardware. DMA has been seen on PC network cards before, but it was expensive and fairly unnecessary. Most PCs don't tax the network sufficiently, and finding a free DMA channel is not very easy in a PC. For machines that do abuse their network connection, the CPU is usually doing it to get code to run a program. The CPU is probably faster than what would go onto the network card, so having it do all the work isn't a big loss.

Audio hardware does benefit from DMA. The audio hardware should be able to play with as little CPU intervention as possible. Those cool multimedia games aren't as cool when the CPU spends its time delivering the sound to the sound card.

There are disk controllers for PCs that use DMA, and there tend to be faster SCSI ones also, but these put far less stress on the CPU.

THE FUTURE OF I/O

As long as computers are used, I/O will have a future. I/O is dependent on mechanical devices such as disks, which will always be slow, until new technology replaces the mechanical devices, or major improvements are made to the existing mechanical interface.

Disks

Disks improve by about 5% per year in speed, while they grow in size by approximately 20%. With 5% growth, they will never catch up with CPUs that speed up faster and have a head start of at least 5 orders of magnitude. The current trend obviously is a growth of hard disk size, especially in the competitive pc market. However, there is bound to be a limit, with respect to size that will be marketable. Currently, 9GB drives, and more are available for real cheap prices, in comparison to the going rate for a hard drive of far lesser size just a few years ago. Not to mention that these newer drives are faster, and more reliable.

Another area of performance increase with respect to disks, include the possibility of newer materials which allow for more compact, as well as faster data access. Perhaps even more intelligent devices will be created, where the cpu is not needed to pass data amongst the various devices at all, and can simply update main memory on their own, like an even better SCSI.

More than likely, systems will try to implement a couple different hard disks, rather than a single large drive, so that more parallelism may occur. With parallelism, information may be written to or from two different drives more efficiently. For example, when reading from one disk, you could be positioning the read/write head on the other disk, in order to read data. The benefit is extremely easy to see in a virtual memory type of application. Suppose that a single smaller disk, that was optimized for faster access, due to it's smaller capacity, and special file system, was used solely for paging. If this was the case, it could be paging information as necessary, and reading data from the disk where all programs and data are stored.

PC OR MAC?

Apple Macintosh computers have followed a parallel trend to faster chips, more RAM and HD space, larger screens etc. Prices started out much more expensive than equivalent PCs, owing to the requirements of the graphical operating system and the monopoly on supply, but have become much more comparable. PCs still have the edge owing to a much larger market and competition between clone makers.

Macs are still the system of choice for desktop publishing and graphics, and are widely used in education, but PCs have 90% of the market for business and home use. Macs still have a marginally better operating system and connectivity. However the difference has been largely eroded with the arrival of Plug-and-Play hardware for PCs and Windows 95.

Typical PC Systems

Intel® Core™ 2 Quad processor Q8200 (4MB L2, 2.33GHz, 1333FSB), 3GB2 DDR2 SDRAM3 at 800MHz- 2 DIMMs, 640GB Serial ATA Hard Drive (7200RPM), Integrated Intel® GMA X4500HD Graphics, Integrated 7.1 Channel Audio, 19" LCD Display, Windows Vista® Home Premium Edition SP1, 64-Bit.

AMD Phenom X4 9750 (2.4GHz, 1066MHz, 4MB), 4GB4 Dual Channel DDR2 SDRAM3 at 800MHz- 4 DIMMs, 640GB Serial ATA Hard Drive (7200RPM), Integrated ATI Radeon HD3200 Graphics, Integrated 5.1, 19" LCD Display, Windows Vista® Home Premium Edition SP1, 64-Bit.

SERVERS

Networking is the sharing of information and services. Networking is possible when individual or groups have information or abilities that they wish to share with others. Computer networking provides the communication tools to allow computers to share information and abilities.

At the heart of any networked computer based system is the server. There are of course many permutations and choosing the correct one to suit your business needs is crucial. An Intel based server or PC server is defined as a server based upon the Intel microprocessor architecture which supports 'industry standard' hardware components and 'industry standard' operating system and application software packages. While mainframes, minicomputers and RISC systems all function as servers, today more than 90% of all servers purchased are Intel processor based.

There are two main types or "personalities" of server - file servers and application servers. The majority of Intel-based servers installed today are used as file servers (approximately 60%) but the application server segment is growing more rapidly as many companies deploy business applications on Intel-based server platforms.

Application server platforms generally demand richer system configurations than file servers, with more processors, memory and disk drives, and with high availability and fault-tolerant features.

Novell NetWare holds the largest operating system share within existing customer installations. Windows NT, however, has been the fastest growing platform for new purchases, particularly in the application server segment. The Linux operating system is now also generating much interest.

File server examples

A file server, as its name suggests, describes a server which is used to share or "serve" files to other client PCs on a network. Examples include:-

- Shared data file storage (e.g. spreadsheet or text documents) on a departmental server.
- Shared (network) copies of office productivity software, such as Lotus SmartSuite or Microsoft Office.

Application server examples

An application server describes a server which runs one or more applications locally on the server itself, sending data or other processing output to client PCs or to other servers. Examples include:-

- A database server
- A groupware or Internet/intranet server
- An e-mail or messaging server
- A business application server
- A server hosting applications for network computers

IBM Servers

IBM® server platforms offer optimum availability, scalability, performance, and cost benefits to enable the industry's most comprehensive Enterprise-level solutions. Whether it's serving high-volume web pages, computing-intensive data, or critical transactions, there's a new IBM EServer or "old" favorite to do the job. Covering every platform from Windows 2000® to UNIX® to Linux to OS/400, the IBM line of servers, which include IBM AS/400, IBM iSeries, IBM pSeries and RS/6000, IBM xSeries and Netfinity, deliver superior performance and cost-effectiveness, backed by outstanding support and services.

PC makers such as Compaq, Data General, Dell Computer, Hewlett-Packard, and IBM want to grow into new markets, and give their salespeople something to talk about. New features offering modest benefits at a reasonable price include some I/O enhancements, such as faster and more reliable PCI, and "rack-dense" servers, which conserve precious space in the computer room. Other features, such as eight-way multiprocessing, may function primarily to establish a vendor's credentials rather than meet a real market need.

EIGHT WAYS SERVER

Eight-way multiprocessing for PC servers depends on the Profusion chipset, which is based on a technology Intel acquired from Corollary when it purchased the company in 1997. After being delayed by Intel several times, the chipset is finally starting to appear in servers from Compaq, Data General, Dell, HP, IBM, Toshiba America, and others.

Manufacturers say they're targeting eight-way servers at computer-intensive applications, including data mining and executive information systems; data center applications; server consolidation; and Internet applications. However, Intel servers currently don't offer the reliability, availability, scalability, and manageability needed to support these mission-critical applications.

This problem is largely due to popular PC server operating systems such as Linux and Windows NT. Currently, Linux does not scale well. While programmers are laboring to overcome that limitation, Linux remains, for the time being, a predominantly single-processor OS.

In an effort to make the Intel server platform as reliable, fault tolerant, and serviceable as possible, many vendors are bringing features previously associated with mainframes, minicomputers, and RISC/Unix machines to their PC servers. IBM, for instance, is borrowing mature technology from its higher-end machines and including it in its PC servers. The company's highly reliable "**chip kill correction**" memory, which allows a computer to keep functioning if it loses a memory chip in an inline module, was originally developed for NASA's Mars Rover unit. It's now an option for the Netfinity 7000 M10 and will become the standard memory technology on other high-end Netfinity products.

Another high-tech hand-me-down that increases reliability is the **SP Switch**, which was introduced for IBM's Netfinity PC servers earlier this year. This component, which comes from IBM's RS/6000, offers point-to-point I/O so that, for instance, a failure in a network card does not affect other network cards or disk controllers. This is a significant advantage over shared bus architectures, such as SCSI.

PC server vendors claim that at the hardware level, they can achieve 99.9 percent uptime using technologies such as error correcting memory; redundant power supplies, fans, and disks; and hot-swappable components such as disks and network cards. That adds up to nine hours of downtime a year, assuming 24-by-7 operation-not including downtime associated with OS and software failures.

DISK ARRAYS AND DATA STRIPING

A disk array organizes multiple independent disks into one large, high performance disk. In addition, some array configurations write backup parity information at the same time. (This renders them fault-tolerant, in that they can continue working even if a disk fails while in use.) Data blocks are split up and written in parallel to the disks, which speeds access.

The length of time taken to execute a read or write on a disk is determined by the time taken for the data area on the disk surface to pass under the read/write heads of the drive. Reading or writing an 8KB block takes eight times as long as a 1KB block. However, if the 8KB block is written to a disk array with eight disks, the data is split into eight stripes of 1KB, which are written to individual disks in parallel. In this way, disk arrays achieve a higher data transfer rate than non-parallel drives.

In practice, the expected linear scaling of throughput from using multiple disks is not achieved. This is because of seek-time, on-board disk caches and parity generation.

Data striping also results in uniform load balancing across all the disks on a system, eliminating disk hot spots. These arise when one disk is saturated with I/O requests while the rest lie idle.

However, when multiple disks are organized into arrays, the potential for data loss from disk failures is higher because the probability of a disk failure occurring in a given period is higher. For example, for a disk drive with a rated mean time before failure of 100,000 hours, there is a 50% probability of the drive failing in that length of time. However, for a disk array of 10 such drives, there is a 50% probability that one of them will fail in 10,000 hours.

RAID

A redundant array of inexpensive disks (RAID) uses multiple fixed-disk drives, high-speed disk controllers, and special software drivers. RAID was designed to remove the bottleneck of low performance of Disk drives. It used multiple-fixed disk and spread the data in all of them in parallel, and allowing the same data to be retrieved from different locations. This increases the performance of hard disk subsystem.

RAID has also tools to increase the safety of your data by spreading it over multiple disk drives and then calculating and storing parity information. This redundancy allows any one drive to fail without causing the array itself to lose any data.

All commercial RAID subsystems use the Small Computer System Interface (SCSI). The latest Ultra-wide SCSI host adapters for the PCI bus can deliver up to 40M per second (40M/s) of data to and from the PC's RAM.

The basic concept behind RAID is a technique called striping. When the system tries to write a block of information to the array, the array controller (NT) breaks the information into smaller chunks of a predetermined size and writes these chunks in parallel across all drives in the array.

There are six kinds of RAID implementations, each of which works in a different way. The NT server can handle 0, 1 and 5 levels, also known as the striping without parity, Disk Mirroring, and striping with parity respectively.

RAID 0 - Stripping :

It is not considered as proper RAID, as it provides high efficiency (fast writing and reading) but no redundancy. It stripes data across various drives thus allowing simultaneous I/O to the drives. This improves the performance of the system a lot but if one drive fails in a RAID 0 array, the data on all drives on the array becomes inaccessible.

RAID 1 : Disk Mirroring:

This provides hundred percent redundancy at the cost of performance of the system. It makes two complete copies of everything to mirrored or duplexed pairs of disk drives. Therefore if you loose data on one drive, you have its clone available with you. But in writing the contents to other drive, performance is lost. Also it requires double the investment on the storage.

But at the same time, the Read performance is greatly enhanced, as reading can be done from any of the drive having its head closest to the data.

RAID 1, therefore, provides a high level of data security by replicating all data, an increase in read performance by allowing either physical drive to fulfill the read request, and a lower level of write performance due to the necessity of writing the same information to both drives.

RAID 2

RAID 2 is a proprietary RAID architecture and it distributes the data across multiple dedicated disks to store parity information. It stripes data bit by bit across these drives in parallel and records parity information to the dedicated parity drives for complete redundancy. RAID 2 isn't a good choice for random-access applications, which require frequent, small reads and writes.

RAID 3

RAID 3 stripes data across drives, usually at the byte level, although bit-level implementations are possible. Unlike RAID 2, RAID 3 dedicated only one parity disk. A single parity disk creates bottlenecks for writing because the parity information must be written before the next write can take place. Similar to RAID 2, RAID 3 is optimized for long sequential disk access in applications such as imaging and digital video storage, and is inappropriate for random-access

RAID 4

RAID 4 is improved form of RAID 3, and stripes data at the block or sector level rather than at the byte level. This helps in better read performance than RAID 3 for small random reads but hardly address the bottleneck caused by the fact that write updates often have to wait to access the parity drive.

RAID 5 : Stripping with Parity

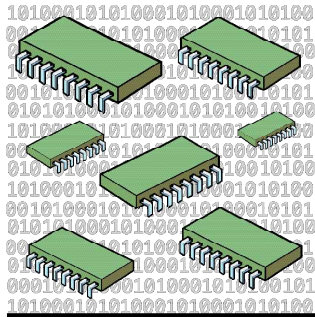
RAID 5 stripes both user and parity data across all the drives in the array, consuming the equivalent of one drive for parity information. With RAID 5, all drives are of the same size, and one drive is unavailable to the operating system. The parity information spreads equally across all the drives, thereby allowing both parallel reads and writes.

Windows NT includes software support for RAID levels 0 - Disk Stripping, 1 - Disk Mirroring and 5 - Disk Stripping with Parity. These are the most common RAID implementations. All these options are included in Disk Administrator. Besides these, Disk Administrator also provides another disk management function called Volume Set.

EXERCISE

1. Fill in the Blanks -
 - a) Results are obtained from computer through its _____ unit.
 - b) _____ unit provides communication from man to machine.
 - c) _____ is the name given to main computer equipment and its peripheral devices.
 - d) Instructions to computer are given through _____.
 - e) Calculations are made in computer with the help of its _____.
 - f) The Central Processing Unit of a computer consists of _____ Control unit and Arithmetic and Logic unit.
 - g) Central processing unit (C.P.U.) is called _____ of the computer.
 - h) Main storage is also called main _____.
 - i) RAID level 0 is called as _____
 - j) RAID stands for _____
 - k) The process by which server maintains a copy of other server, data is called as _____.
 - l) Windows NT allows upto _____ partition on the hard disk.
 - m) Mirroring is RAID _____ level.
 - n) RAID 0 is _____% redundancy where as RAID 1 is _____% redundancy.
2. Match the column -

(1) Input unit	(a) Retrieves data from the CPU.
(2) Output unit.	(b) Performs the arithmetic and logical operations
(3) Memory unit	(c) Directs the other units.
(4) Arithmetic and Logic unit	(d) Stores the computer program and the data which are to be immediately processed.
(5) Control unit.	(e) Feeds data into the CPU.
3. What are the differences between hardware and software?
4. What are the basic elements of a computer?
5. What are the four basic elements of computer management system?
6. What functions are performed by the control unit ?
7. What do you understand by RAID? Discuss various RAID levels.
8. Distinguish between following -
 - (a) Input and Output unit
 - (b) Primary and auxiliary memory
 - (c) CPU & ALU



CHAPTER 2

COMPONENTS OF COMPUTER

SYSTEM UNIT

KEY BOARD

MOUSE

VISUAL DISPLAY UNIT (VDU)

OTHER PERIPHERAL DEVICES

PRINTERS

Components of Computer

Modern computers are electronic and digital. The actual machinery - wires, transistors, and circuits - is called **hardware**; the instructions and data are called **software**.

All general-purpose computers require the following hardware components:

Memory	Enables a computer to store, at least temporarily, data and programs.
Mass Storage Device	Allows a computer to permanently retain large amounts of data. Common mass storage devices include disk drives and tape drives.
Input Device	Usually a keyboard and mouse, the input device is the conduit through which data and instructions enter a computer.
Output Device	A display screen, printer, or other device that lets you see what the computer has accomplished.
CPU	Central Processing Unit - the heart of the computer, this is the component that actually executes instructions.

In addition to these components, many others make it possible for the basic components to work together efficiently. For example, every computer requires a bus that transmits data from one part of the computer to another.

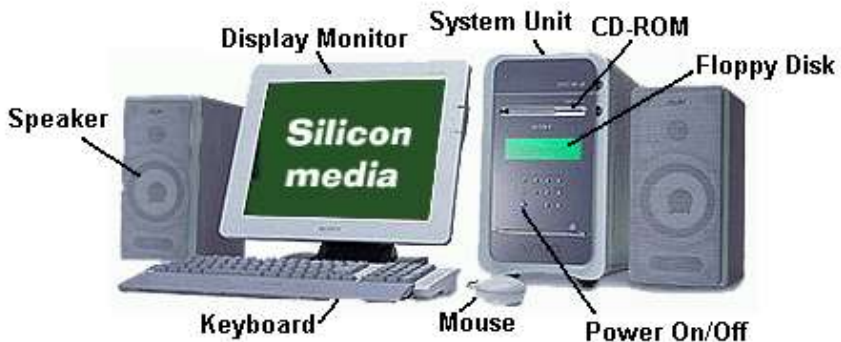


Fig 2.1 Components of a Personal Computer

SYSTEM UNIT

The System Unit contains all the essential parts for running the computer. It includes Motherboard, CPU, RAM, Storage (Floppy & Hard Disk), SMPS and all the ports for printer, sound, modem etc. Fig 2.2. displays the two views of System Unit from Inside and from back.

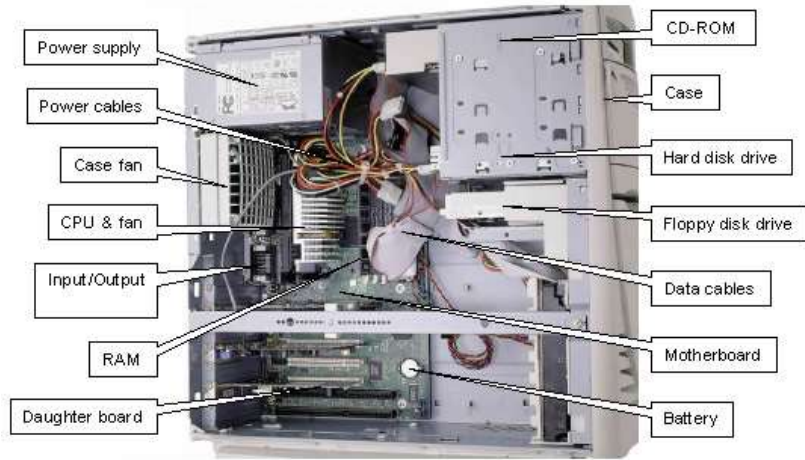


Fig 2.2 System Unit Inside View

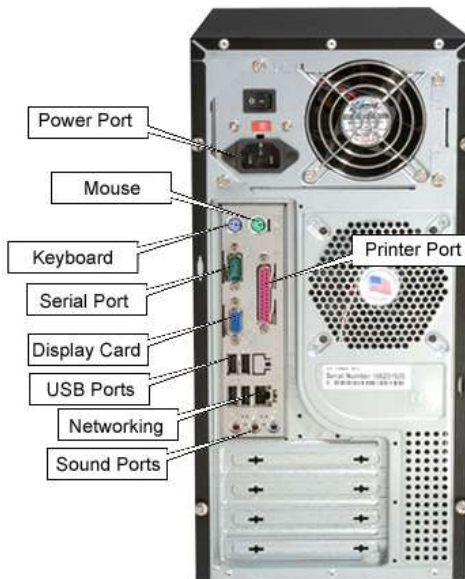


Fig 2.3 Back View of System Unit

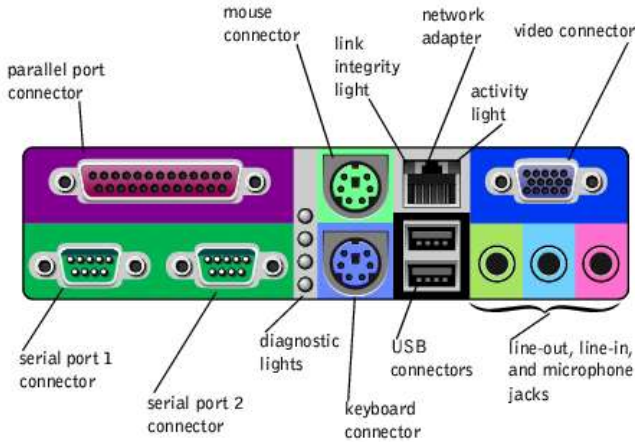


Fig 2.4 A typical Back Connector Panel

Connecting Devices

Parallel Port Connector

This is used to connect printers. Default designation is LPT1.

Mouse Connector

This is used to Connect PS/2 mouse to Computers.

USB Connectors

These are used to attach USB-compliant devices such as keyboards, mice, printers, and computer speakers to your computer.

Integrated Network Adapter Connector

This is used to connect your Computer to the network. It has two lights. A yellow activity light flashes when the computer is transmitting or receiving network data. A dual-colored link integrity light, which flashes when there is a good connection between a network and the network adapter.

Video Connector

This connector is used to attach a video graphics array (VGA)-compatible monitor to your computer.

Line-In Jack

This jack is used to attach record/playback devices such as cassette players, CD players, and VCRs.

Line-Out Jack

This jack is used to attach computer speakers.

Microphone Jack

This jack is used to attach a standard personal computer microphone.

Serial Port Connectors

Default port designations: COM1 for serial port 1 and COM2 for serial port 2. You can reassign the serial port's designation if you add an expansion card containing a serial port using this designation.

Keyboard Connector

This is used to attach the keyboard cable to the 6-pin connector on the back panel.

MOTHERBOARD

In any computer organization mother board is the primary component. It incorporates all the fundamental units required for a computer.

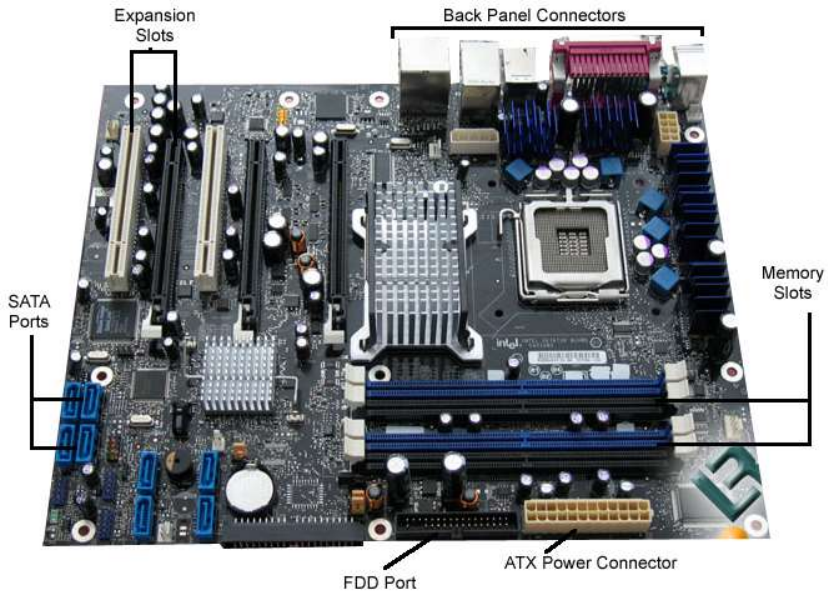


Fig 2.5 Intel D975XBX MotherBoard

The Main components of a motherboard are

- The CPU logic
- Memory (RAM / ROM)
- INPUT (PORTS)
- OUTPUT (PORTS)

Thus we see that a motherboard holds most of the primary logic required for a computer and thus justifies it's name.

In the motherboard, we need to connect the corresponding adapters. These adapters communicate with the CPU at the motherboard through address and data buses. Since different manufacturers are manufacturing different types of motherboard and adapters therefore it becomes necessary to have some standardized designs to make all the boards adaptable to each other. This is realized through the standardized designs of I/O slots called I/O buses. Motherboards will be discussed in detail in the coming chapters.

Motherboard Components

Apart from of the main CPU many other support chips called peripheral chips reside on the motherboard. The support chips are the programmable chips which support many important dedicated functions on the motherboard. The peripheral chips are designed to work for a particular CPU type. To list a few

- CLOCK GENERATOR
- PIT (Programmable interval timer)
- INTERRUPT CONTROLLER
- DMA CONTROLLER
- BUS CONTROLLER
- PPI (Programmable peripheral interface)

In the latest motherboards PC/AT, 386 or 486 no discrete chips are used. Instead specially designed CMOS VLSIs which incorporate one or more of the above mentioned circuits.

In addition to the CPU and peripheral chips other components available on the motherboards are

- Discrete registers, transceivers and buffers
- BIOS ROMs
- RAM banks
- Keyboard controllers
- CMOS chips and battery for storing SETUP
- Jumpers and connectors etc.

Many motherboards available today are also having some additional built-in controllers also. Therefore they have the controllers and circuitry mounted on them. To list a few following are the different types of controllers integrated into the motherboards these days.

- Video display controller
- Serial/Parallel port controller
- IDE controller
- SATA Controller
- Floppy Drive Controller
- Game port controller

Further to these motherboards are having I/O slots, jumpers, connectors (power, keyboard, etc.), some resistances and capacitors combinations, crystal oscillators.

THE PROCESSOR

Central Processing Unit (CPU) is control center for the computer. It carries out all instructions sent to it by the operating system or applications software.

As discussed earlier, the CPU contains two circuit elements that perform tasks and several memory locations where data/instructions are held temporarily while actions are performed. Here they are.

- A set of registers for holding binary information.
- An arithmetic and logic unit (ALU) for performing data manipulation and performing basic mathematical and logical operations.



Fig 2.6 An Intel Core 2 Extreme CPU

Heatsink and Fan on the motherboard



Fig 2.7 An Intel Core 2 Extreme CPU

- A control unit that coordinates and controls the various operations and initiates the appropriate sequence of micro-operations for each task.

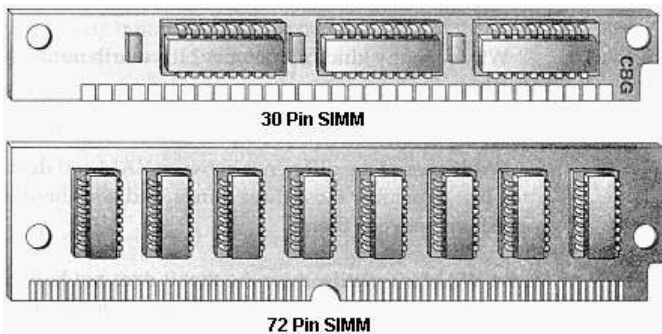
There have only been five common brands of CPU in the market in recent times: Intel, AMD, Cyrix, Centaur and IBM, and all of them are sold by vendors at one time or another. More about History of CPU and current trends will be discussed in the coming chapter dedicated to CPU only.

RAM

The PC has primary main memory called RAM. CPU can read from or write to " HIGH SPEED " memory. There are different kinds of RAM's which are used in PC's.

DYNAMIC RAM is a volatile RAM i.e. data is stored temporarily. Each bits of single transistor and capacitor. DRAM based system requires refresh circuitry. Each bit has row address and column address. DRAMs are measured by storage capability and access time. Storage is rated in megabytes (8 MB, 16 MB, etc). Access time is rated in nanoseconds (60ns, 70ns, 80ns, etc) and represents the amount of time to save or return information. With a 60ns DRAM, it would require 60 billionths of a second to save or return information. The lower the nanospeed, the faster the memory operates.

STATIC RAM is the simplest kind of memory. It is called " static " because entered data stays there in power up. They need six transistors for each bit storage location. SRAM is quite faster than DRAM. SRAM's are quite expensive and are very rarely used as main memory.



EDO RAM is a new type of memory being offered for Pentium systems today called EDO (extended data output).These are 72 pin SIMM's with specially manufactured chips that allow for timing overlap between successive accesses. This allows for tighter coupled access cycle and a performance improvement.

FPM (Fast Page Mode) was the most common at one time and was often just referred to as DRAM. It offered faster access to data located within the same row.

BEDO (Burst Extended Data Out) is a step up from the EDO chips. It requires zero wait states and provides at least another 13 percent increase in performance.

SDRAM (Synchronous DRAM) introduced in late 1996, retains memory and does not require refreshing. It synchronizes itself with the timing of the CPU. It also takes advantage of interleaving and burst mode functions. This type of SDRAM is slower than the DDR variants, because only one word of data is transmitted per clock cycle (single data rate).

DDR SDRAM (Double Data Rate Synchronous DRAM) allows transactions on both the rising and falling edges of the clock cycle. Also called DDR1 for greater clarity. Typical DDR SDRAM clock rates are 133, 166 and 200 MHz (7.5, 6, and 5 ns/cycle), generally described as DDR-266, DDR-333 and DDR-400 (3.75, 3, and 2.5 ns per beat).

DDR2 SDRAM is very similar to DDR SDRAM, but doubles the minimum read or write unit again, to 4 consecutive words. Typical DDR2 SDRAM clock rates are 200, 266, 333 or 400 MHz (periods of 5, 3.75, 3 and 2.5 ns), generally described as DDR2-400, DDR2-533, DDR2-667 and DDR2-800 (periods of 2.5, 1.875, 1.5 and 1.25 ns). Corresponding 240-pin DIMMS are known as PC2-3200 through PC2-6400.

DDR3 SDRAM continues the trend, doubling the minimum read or write unit to 8 consecutive words. Initial clock rates were 400 and 533 MHz, which would be described as DDR3-800 and DDR3-1066, but 667 and 800 MHz (DDR3-1333 and DDR3-1600) are now common.

DDR4 SDRAM will be the successor to DDR3 SDRAM. The new chips are expected to run at 1.2 V or less, versus the 1.5 V of DDR3 chips and have in excess of 2 billion data transfers per second. They are expected to be introduced at clock speeds of 1600 MHz, estimated to rise to a potential 3200 MHz (3.2 GHz) and lowered voltage of 1.0 V by 2013.

Direct Rambus is extraordinarily fast. By using doubled clocked provides a transfer rate up to 1.6GBs yielding a 800MHz speed over a narrow 16 bit bus.

Cache RAM is where SRAM is used for storing information required by the CPU. It is in kilobyte sizes of 128KB, 256KB, etc.

VRAM (Video RAM) is a video version of FPM and is most often used in video accelerator cards. Because it has two ports, It provides the extra benefit over DRAM of being able to execute simultaneous read/write operations at the same time. One channel is used to refresh the screen and the other manages image changes. VRAM tends to be more expensive.

Flash Memory is a solid-state, nonvolatile, rewritable memory that functions like RAM and a hard disk combined. If power is lost, all data remains in memory. Because of its high speed, durability, and low voltage requirements, it is ideal for digital cameras, cell phones, printers, handheld computers, pagers and audio recorders.

Shadow RAM is a RAM which helps in fast starting up of Computers. When your computer starts up (boots), minimal instructions for performing the startup procedures and video controls are stored in ROM (Read Only Memory) in what is commonly called BIOS. ROM executes slowly. Shadow RAM allows for the capability of moving selected parts of the BIOS code from ROM to the faster RAM memory.



Fig 2.8 SDRAM of 1 GB

THE EXPANSION SLOT

They are standard connectors. Expansion boards plug into these slots. More slots equals better flexibility to upgrade systems. Virtually, no PC systems have more than eight expansion slots. Some PC systems provide only a single expansion slot on the mother board. This slot is normally called RISER CARD. The RISER CARD in turn has expansion slots on its sides. Adapter cards are installed in the riser card, meaning that they are parallel to the mother board. RISER CARDS are normally used where the system height is shorter than normal.

The need for expansion

The original IBM PC in its most basic form was a very weak machine and it relied heavily on the idea of expansion options to make it into something useful. For example, one of the early models could only have 256 Kilobyte of memory fitted to its main board. If you wanted to expand this to the limit of 640 Kilobyte then you had to buy an extra printed circuit board to accommodate the RAM chips. This board plugged into a special connector or slot towards the back of the machine. The collection of expansion slots is generally referred to as an Expansion Bus. Eight such slots were generally available for expansion options that included, in addition to memory, video boards, disk controllers, printer ports, modem ports etc.. If it hadn't been for this remarkable expansion capability it is arguable that the IBM PC would have become extinct very rapidly.

Most owners ran their machines with nearly all of the expansion slots full and finding a spare slot for something new often proved difficult.

This experience has tended to drive home the idea that the provision of expansion slots in quantity is the mark of a good machine. Perhaps it should be remembered that a full set of expansion slots is a good indication that the original machine was severely lacking in basic features in the first place!

MAGNETIC DISKS

Magnetic disks can be of several types depending upon the nature of material of construction and its usage. Primarily they are of three types:

- Hard disk
- Floppy disk
- Winchester disk

Magnetic disks are widely used as input or output devices and for auxiliary computer's memory. Magnetic disks are getting very popular with modern computer. A magnetic disk is similar to gramophone's record but is coated with ferro magnetic materials and rotates at very high speeds of 1800 or 3600 rpm (revolutions per minute). With the help of both the devices, we can have direct access to any information, irrespective of its location on disk.

Hard Disks

A hard disk is divided into several concentric tracks in which information is stored magnetically. Both sides of the disk are used. A multi-disk system is shown in the figure.



Fig 2.9 Multi disk single head system in a Hard disk

In this figure a comb-type multi-disk memory device is shown, which consists of six number of disks mounted horizontally parallel to each other on a central spindle. The disk pack can be mounted or removed at will from a disk transport that is permanently connected to central processor. The data transfer in the disk memory can be serial or

parallel . In one head per system, the transfer is serial, but in case of multihead system, whole word may be accessed in parallel. In this case, bits of the same word are on different tracks.



Fig 2.10 (a) A Hard Disk

(b) Inside View of The Hard Disk

The timing gap is used to reset the timing and addressing the circuits once in every revolution. The timing tracks provide clock pulses.

One Hard disk pack can store from 2 millions to 100 million characters. These disks are sealed into their boxes and run faster than floppy disks. But for the same storage capacity, a hard disk is bigger and more expensive than the magnetic tape.

Hard disks are of several types :

- Fixed disk which cannot be removed from disk drive.
- Fixed head disk, a fixed disk having one read/write head for each track.
- Exchangeable disks.

Winchester Disks

Winchester disks fixed in airtight units are intended for use in dusty or humid environments. These disks are fast and highly reliable and yet low priced as compared with other hard disks. No preventive maintenance is required. Their storage capacity is 10 to 50 Megabytes.

SCSI (SMALL COMPUTER SYSTEM INTERFACE)

SCSI is not a simple Interface. It is a stand alone, external bus capable of transmitting 8-bit of parallel data at speeds of up to 4 Mega Bytes per second. The SCSI bus treats devices connected to the bus as logical devices. It communicates with these devices through a defined Protocol, designed specifically to eliminate hardware in compatibilities.

Each device contains its own intelligent controller to control its mechanism that results in higher performance from each drive. In SCSI drive controller and disk controller are present in the drive itself. This way it improves the performance. This configuration allows all the control functions including motor control and data encoding, to be closely coupled with the drive's electronics, thereby decreasing transmission time and overhead.

FLOPPY DISK

Floppy disk looks like a post card size envelope. A floppy disk is made of plastic material, 20 cm in diameter and coated with minute coating of ferrite of 0.015 millimeter thickness. It is called 'Floppy' because it is made of flexible plastic and when swung like a hand fan it flops. Since the disk is very delicate and very fragile, it is always kept inside a plastic or cardboard cover to increase its stability. The cleaning and lubrication can be done without removing the plastic cover. There is a hole in the centre through which the spindle of drive unit rotates the disk. A window is also provided so that read and write head makes contact with storage surface of the disk. In addition, there is a third hole by which beginning of a track is determined by optical means. Floppy disks are called minidisks or diskettes also.

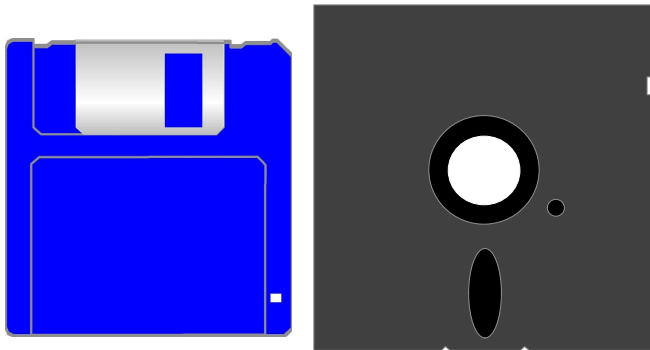


Fig 2.11 Floppy Disk

A standard and typical floppy disk has 73 data tracks each having 26 sectors, each accommodating 64 words. Thus, standards disk can store 121472 words. A floppy disk of size of 45 rpm record can hold the contents of several thousand typed pages. The biggest floppy can store 360 kilo-bytes of information. Even today such a floppy is also available which can store entire Encyclopedia Britanica twice. The disk drive unit may come as integral part of computer or it may be altogether separate unit than magnetic disk. As far as the performance is concerned it is inferior to hard disk but is very much handy. The disk is rotated at a speed of 360 rpm.

There are two standard sizes of floppy disks, 8 inch and 5.25 inch frequently referred to as diskette and mini-floppy respectively. The storage capacity of former is typically between 250K-bytes and 1.5 M-

bytes and that of latter is between 125 K-bytes to 500 K-bytes. Floppy disk is a low cost device particularly suited to supporting personal systems based on micro computers.

VIDEO OR OPTICAL DISK (CDROM)

The optical disc was invented in 1958. In 1961 and 1969, David Paul Gregg registered a patent for the analog optical disc for video recording. Initially, optical discs were for storing music and computer software. The laser disc format stored analog video signals, but, commercially, lost to the VHS videotape cassette, mainly its high cost and non-recordability; other first-generation disc formats are designed solely to store digital data.

Because optical disk systems can store as much as 100 times the contents of current magnetic disk drives, they seem to afford a promising solution to the storage problem. An optical disk storage system is similar in principle to the laser-optical video disks sold for home use. As applied to computer-related information systems, optical disk store digitized data instead of the digitized elements of a picture.



Fig 2.12 CD-ROM

In optical-disk recording system an intense laser light source is focused on a rotating disk, which is coated with a reflective surface. Data recording is accomplished by turning the laser beam.

But CD-ROM is just one form of optical storage of the first generation. The other First generation discs include Laser Disc and Magenta Optical Disc. Most first-generation disc devices had an infrared laser reading head. The minimum size of the laser spot is proportional to its wavelength, thus wavelength is a limiting factor against great information density, too little data can be stored so. The infrared range is beyond the long-wavelength end of the visible light spectrum, so, supports less density than any visible light colour. One example of high-density data storage capacity, achieved with an infrared laser, is 700MB of net user data for a 12 cm compact disc.

In addition to this prerecorded read only technology there are laser-based writable optical storage devices also appearing in the market. With Write ones, Read Many, or WORM systems, users record their own data on blank optical disks. A laser recording device writes bytes

by melting, bubbling, or otherwise deforming a thin sensitive layer of material on the disk surface. Unmodified areas represent 0 bits. The deformed WORM disk can't be restored to its original condition, so writing is inedible. Reading of stored data occurs when a lower power laser beam passes over the disk and detects differences in the reflections coming from 0 and 1 bit areas. There are no reading heads to rub against or crash into the disk surface.

Second-generation Optical Disc

Second-generation optical discs were for storing great amounts of data, including broadcast-quality digital video. Such discs usually are read with a visible-light laser (usually red); the shorter wavelength and greater numerical aperture allow a narrower light beam, permitting smaller pits and lands in the disc. In the DVD format, this allows 4.7GB storage on a standard 12 cm, single-sided, single-layer disc; alternately, smaller media, such as the MiniDisc and the DataPlay formats, can have capacity comparable to that of the larger, standard compact 12 cm disc.

DVD, also known as "Digital Versatile Disc" or "Digital Video Disc," is an optical disc storage media format. Its main uses are video and data storage.

GD-ROM (giga disk read-only memory) is the proprietary optical disc format used by the Sega Dreamcast. It is similar to the standard CD-ROM except that the pits on the disc are packed more closely together, resulting in a higher storage capacity: around 1.2 gigabytes.

Third-generation Optical Disc

Third-generation optical discs are in development, meant for distributing high-definition video and support greater data storage capacities, accomplished with short-wavelength visible-light lasers and greater numerical apertures. The Blu-ray disc uses blue-violet lasers of greater aperture, for use with discs with smaller pits and lands, thereby greater data storage capacity per layer.

Blu-ray Disc (also known as Blu-ray or BD) is an optical disc storage medium designed to supersede the standard DVD format. Its main uses are for storing PlayStation 3 games, high-definition video, and data storage, with up to 25 GB per single layered, and 50 GB per dual layered disc. The disc has the same physical dimensions as standard DVDs and CDs.

Versatile Multilayer Disc (VMD or HD VMD) is a high-capacity red laser optical disc technology designed by New Medium Enterprises, Inc.. VMD is intended to compete with the blue laser Blu-ray Disc format and has an initial capacity of up to 30GB per side.

Next Generation

The Holographic Versatile Disc (HVD) is an optical disc technology that, in the future, may hold up to 6TB (terabytes) of information,

although the current maximum is 500GB. It employs a technique known as collinear holography, whereby two lasers, one Blue and one green, are combined into a single beam. The green laser reads data encoded as laser interference fringes from a holographic layer near the top of the disc while the Blue laser is used as the reference beam and to read servoinformation from a regular CD-style aluminum layer near the bottom. Servoinformation is used to monitor the position of the read head over the disc, similar to the head, track, and sector information on a conventional hard disk drive. On a CD or DVD this servoinformation is interspersed amongst the data.

SMPS

One of the most important components in PC is the power supply. In PC systems power supply converts AC voltage into different DC voltages required by the system. The power supply in a PC is designed to convert the 220 V AC into something (DC) that the computer can use, which is specifically both 5V and 12V DC. Usually, the digital electronics components and circuits of the system (MOTHER BOARDS, ADAPTER CARDS, DRIVES LOGIC CARDS) Use 5V while the motors (DRIVES FAN) use 12V supply. One thing that the power supply does to ensure that the system does not run without proper power levels is to prevent the computer from starting up until all the correct power levels are present.



Fig 2.13 A Typical SMPS

Each power supply completes internal checks and tests before allowing the PC to start-up. The power supply sends to the motherboard a special signal called " POWER GOOD " signal. If the signal is not there the computer will not run. The effect of this setup is that when the AC voltage dips, the power supply becomes overstressed or overheated, the power good signal goes down, forcing a system to reset or complete shut down. The power good line is wired to the 8284 or 82284 clock generator chips, which control the clock and reset lines to the microprocessor.

OTHER PERIPHERAL DEVICES

The peripheral devices are defined as the devices that are connected to the computer in order to get most of the advantage out of it. The Motherboard, CPU (Central Processing Unit) and Memory (RAM (random access memory) & ROM (read only memory)) are the main components of the computer. A peripheral is any other device additionally attached to the computer behind the chipset whose primary functionality is dependent upon the host, and can therefore be considered as expanding the hosts capabilities, while not forming part of the system's core architecture.

KEY BOARD

Each micro computer consists of a key board as input unit. In addition, the key board contains several command keys and function keys. Beneath each key, there is a set of electrical contacts, which are closed when the key is pressed. By closing the electronic circuit, a signal (corresponding to the character contained by the key is produced which is sensed by computer. This computer appears on the screen. Multi-lingual key boards have been also devised.

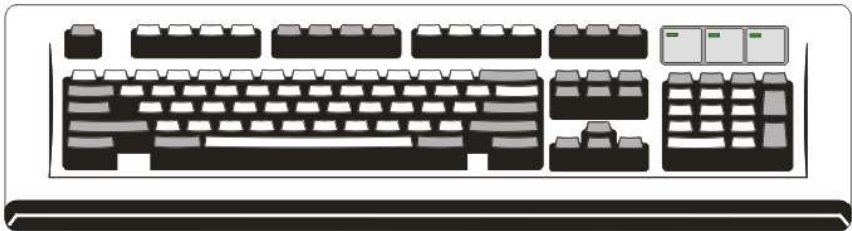


Fig 2.14 Keyboard

Keyboards of special design are used for musical computer. In such key boards, piano type keys are provided through which any tune can be played. The standard IBM PC keyboard can be divided into three general areas :

- Typewriter area, having the standard set of alphabet letters and number keys.
- Function keys - F1 through F10 - which have different functions depending on the software in use. F1 is mostly used for getting help.
- Numeric Keypad, having cursor control and number keys. Some keys on this pad have dual roles and the required number is obtained by either pressing along with the shift key or by first pressing the NUMLOCK key. The numeric keypad is used for entering large amount of numeric data.

Mostly the special function keys perform the following functions :



Moves the cursor to the upper left corner of the screen.



Moves the cursor up by one line.



Moves the cursor down by one line.



Moves the cursor one space to the left, i.e., one column to the left.



Moves the cursor one space to the right, i.e. one column to the right.



Moves the cursor to the bottom right of the screen.



Deletes the character at the current position of the cursor.



Set the insert mode ON/OFF.



Deletes the character on the left of the cursor, and, also the moves cursor one column to the left.



Helps to exit from the present situation and return, to the previous mode.



Moves the cursor to the next tab stop.



Prints the contents of the screen on the printer.



Helps to view the previous page.



Helps to view the next page.



By pressing this, the numeric keypad goes into the 'numeric only' mode, i.e., the function keys will go on displaying the numbers from one to nine.



Used for converting the lettercase to Capital Letter. Once pressed, all the letter would appear in capital.



Always pressed with another key. These two keys have different meanings in different software packages. When pressed together with the DEL key, the system resets and re-boot.



Causes the computer to obey commands. When typing matter, pressing the or key will make the cursor move down one line to the extreme left position.



When Pressed with another key, converts its lower case character to the upper case character. For example, the key D, when pressed alone, will show the small letter on the screen. How ever if you press it along with the shift key, it will show the letter which you know as Capital D.

If you carefully look at each key, you will find that a number of keys have two characters shown on each of them. If you press any of these keys alone, you will

get the symbol or the character shown on the lower part. But, if you press these keys while also pressing the shift key, you will get the character displayed on the top part of the key. This is why the key showing 4, when pressed alone, will give us 4.

However, when pressed along, with the shift key, it will show US\$, which is the sign of the dollar.



Inputs a blank character or space.

MOUSE

A mouse is also a pointing device. It is held in one hand and moved across a flat surface. About the size of the normal cake of bath soap, it usually rolls on a small rubber ball and has two or three buttons on the top. When the user rolls the mouse across the flat surface, the cursor on the screen moves in the direction of mouse movement, giving the user a sense of pointing at something on screen.



Fig 2.15 Mouse

Positioning a cursor with a mouse is easier for non-typists than pressing various key combinations. With the proper software, a mouse can also be used to 'draw' pictures on the screen and edit text, as is done with software like Ventura, Coreldraw, etc. in DTP systems.

VISUAL DISPLAY UNIT (VDU)

Till third generation of computer it was not possible to verify whether the program being typed into computer's memory is correct or correctly entered. It was possible only when microcomputer was introduced. The micro computer consisted of one visual display unit as its essential part on which the user can directly see, whatever he she is typing on its keyboard. When a key is pressed, corresponding character is displayed on the VDU. The result of the calculations or manipulation can also be seen directly on the visual display unit and that too instantaneously. By addition of this unit you can be interactive with computer, i.e, can talk to computer by typing a query and computer will answer your query.



Fig 2.16 A VDU

Visual display unit looks like a colour television and is fabricated as integral part of microcomputer. Sometimes it may be an independent unit and connected to keyboard and CPU by flexible chords. Quite often the VDU houses the floppy or Winchester disk drive also. The home computer can be attached to the colour television itself but through a interfacing device. Some televisions have this interfacing device in built with them small.

VDU's screen is made up of Cathode Ray Tube (CRT) due to which it is also called CRT terminal. When the data input is fed through the keyboard of the computer, electrical signals are generated. The CPU converts these signals into alphanumeric or graphic display. So the output is obtained as optical display. It contains about 20 to 40 rows of horizontal lines with each line divided into 80 spaces.

Some computers are provided with the facility of changing the modes of operation. In that case the number of columns and rows will change. The mode of dividing the screen into several small square grid areas for visual display is known as resolution. One square area is meant for displaying one dot. A group of several dots makes up a pixel. Every character is made up of several pixels. When the entire screen is divided into high number of square areas, then that particular mode or VDU is said to have high resolution. With high resolution screen the information is more accurately represented.

Visual display unit has separate storage of about 1 Kb capacity. This memory relieves the computer during display by storing the data. The most common technique for forming characters on video screen is the dot matrix consisting of 35 dots (5x7) as shown in Fig. A screen usually displays between 500 to 2000 characters. Visual display units lend themselves to graphic outputs also. VDU is particularly suitable for looking up a record held on the computer, for instance, whether some particular person has got the reservation or not.

The screen is coated with dots of phosphor called pixel from inside. The more is the number of pixels from inside. The more is the number of pixels or dots in a screen, the better will be the capability of

computer to display the information. The phosphors can be made to glow red, green or blue by firing electrons at them from behind.

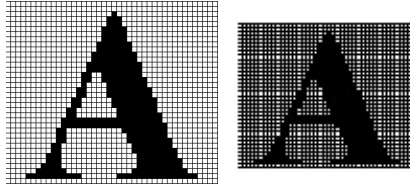


Fig 2.17 Pixels

The electrons have to be guided to the right spot through the vertical and horizontal deflection systems. This ensures that the beam lights up the right phosphors to produce the correct image on the screen.

All the time, the screen deals with tiny points of light, grouped together, to form letters or numbers. It moves so fast that you never notice all the steps involved in putting together a screenful of information. To make it even faster, microcomputers often have a ROM (Read only memory) programmed with the exact format of each better symbol and number.

The VDU can be broadly classified as :

- Monochrome monitors
- Colored Monitors.
- Thin film transistor liquid crystal display (TFT - LCD)
- Projectors

Monochrome Monitors

These are generally black & white monitors or green phosphor monitors. They do not receive TV signals.

Coloured Monitors

The colors are the human perception of the different wavelengths of light. It is possible to create almost all visible colors using two systems of primary colors. Transmitted colors use red, green and blue and reflected colors use cyan (light blue), magenta (purple red) and yellow. Color monitor use clusters of red, green and blue phosphor dots.

On a color monitor, colors are displayed as varying intensities of red, green and blue dots. When red, green and blue are turned on high, white is produced. As the intensities are equally lowered, shades of gray are produced. When all dots are turned off, the base color of the screen appears.

The picture tube of a color monitor is coated with three different phosphors, one of the each of the chosen red, green and blue primaries. The three phosphors are physically separate from one

another and each is energized by an electron beam of intensity that is proportional to the respective color voltage signal produced in the computer. The object is to produce three coincidental raster, which produce the red, green and blue contents of the transmitted picture. While seeing from a normal viewing distance the eye integrates the three color information to convey the sensation of the hue at each part of the picture.

Liquid Crystal Display

In LCD liquid crystalline material is sandwiched between two glass or plastic plates. The front plate is transparent and the back plate is reflective. There is a coating of thin film on the front plate. The coating is transparent and conductive. Its sections (segments) are in the shape of desired characters. LCDs do not emit their own light. Therefore, a light source is to be used. LCDs simply change the reflection of available light.



Fig 2.18 Light Pen

Projectors

A projector takes a video signal and projects the corresponding image on a projection screen using a lens system. All projectors use a very bright light to project the image, and most modern ones can correct any curves, blurriness, and other inconsistencies through manual settings.

Projectors come with many features, strengths, or weaknesses, but most will run on either LCD (Liquid Crystal Display) or DLP (Digital Light Processing) technology. LCD is the older technology, but this does not mean that it is becoming obsolete by any means.

LCD is the simplest system, making it one of the most common and affordable for home theaters and business use. Its most common problem is a visible “screen door” or pixelation effect, although recent advances have minimized this.



Fig 2.19 LCD Projector

DLP uses one, two, or three micro fabricated light valves called digital micro mirror devices (DMDs). The single- and double-DMD versions use rotating color wheels in time with the mirror refreshes to modulate color. The most common problem with the single- or two-DMD varieties is a visible “rainbow” which some people perceive when moving their eyes.

Touch Screen

A touchscreen is a display that can detect the presence and location of a touch within the display area. The term generally refers to touch or contact to the display of the device by a finger or hand. Touchscreens can also sense other passive objects, such as a stylus on a smaller area. However, if the object sensed is active, as with a light pen, the term touchscreen is generally not applicable. The ability to interact directly with a display typically indicates the presence of a touchscreen.



Fig 2.20 Touch Screen

The touchscreen has two main attributes. First, it enables one to interact with what is displayed directly on the screen, where it is displayed, rather than indirectly with a mouse or touchpad. Secondly, it lets one do so without requiring any intermediate device, again, such as a stylus that needs to be held in the hand. Such displays can be attached to computers or, as terminals, to networks. They also play a prominent role in the design of digital appliances such as the personal digital assistant (PDA), satellite navigation devices, mobile phones, and video games.

LIGHT PEN

Light pen is a device which enables us to write anything directly on the screen of computer. We can draw sketches and diagrams also. The light pen has got a tip which is very much sensitive to light. Actually it is a photosensitive detector.

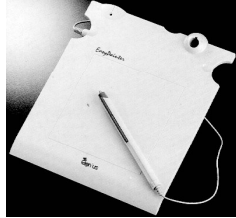


Fig 2.21 Light Pen

A bright spot of light known as raster spot moves in a orderly way across the back of the computer's screen which alerts the tip of light pen. Because of this, the micro-computer can work out where the spot was, when the tip of the pen was altered. Thus an image is formed on screen in the same manner, as the pen is subjected to move. Now you can write on to the screen of computer which will alter the computer's memory accordingly.

WEBCAMS

Webcams are video capture devices connected to computers or computer networks, often using a USB port or, if they connect to networks, via ethernet or Wi-Fi.

Their most popular use is for video telephony, permitting a computer to act as a videophone or video conferencing station. Other popular uses, which include the recording of video files or even still-images, are accessible via numerous software programs, applications and devices.



Fig 2.22 A Typical WebCam

JOYSTICK

A joystick is also used to move the cursor position on screen. Its function is similar to that of a mouse. Children can play with computers in a simple way by the use of joystick. While playing video games like SPACE-INVADERS, the user need to move the space craft quickly across the screen. This can be done by pressing a key on the key board but it is neither convenient nor comfortable for small children to use the key board. This type of actions can be better done with the help of joystick. As you move the stick around in its socket the movements are translated into binary instructions with the help of electrical contacts.



Fig 2.23 A Typical Joystick

DIGITIZERS

A digitizer is an input device that converts graphic and pictorial data to digital (binary) form which can be direct fed and stored inside a computer. There are two types of digitizers : rectangular-coordinate or flatbed digitizer and image-scan digitizer. In case of a flatbed digitizer, the drawing to be digitized is spread and fixed over a rectangular flatbed table. A mechanism is now moved over the surface of the drawing that scans the drawing and produces signals related to the X and Y coordinates of the table. Image-scan digitizers scan and reproduce entire drawing and photographs automatically. They are costlier and more powerful than the flatbed digitizers and are capable of digitizing not only the shape and size of the drawings but also the varying intensities on a gray-to black scale at different points of the drawings. Thus, flatbed digitizers are mainly use to digitize simple drawings, graphs, charts, etc. and image-scan digitizers are used to digitize more complex pictures and photographs.

OPTICAL MARKS READER (OMR)

It is a device for checking the documents and question papers. Now-a-days it is quite a common practice to judge the knowledge and I.Q. level of a candidate by multiple choice objective type questions. The candidates are required to write their answers by putting a standard dot in front of the correct answer with the help of specially designed

soft pencils. While evaluating the answer papers, OMR directs a thin beam of light on the surface of the paper. Lesser amount of light is transmitted through the dot as compared to nearby surface due to which answer is recognised by the computer. The OMR can evaluate only those documents which are printed with marked positions in the specified areas. OMR can read about 10,000 documents per hour, and thus can print the result within the same day without any favoritism or mistakes. The UPSC, New Delhi is equipped with such a computer based device.

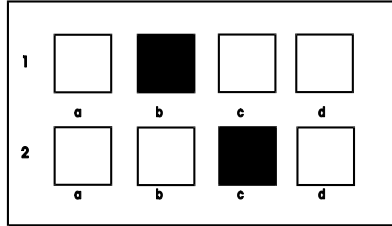


Fig 2.24 Optical Mark Reader

OPTICAL CHARACTER READER (OCR)

This is also a device for reading the document, It can read alphanumeric characters created in a particular style or good quality paper using special ink.

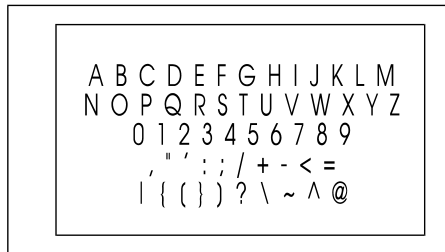


Fig 2.25 Optical Character Reader

The shape of character is recognised by the OCR with the help of a light source in conjunction with photo electric cells. Optical characters are being extensively used for billing purposes, i.e, in preparation of electricity bills and Insurance premium renewals.

BAR CODE READERS

This method uses a number of bars (lines) of varying thickness and spacing between them to indicate the desired information. Bar code are used mostly on grocery items. An optical bar code reader can read such bars and convert them into electrical pulses to be processed by a computer. The most commonly used bar code is universal product code (UPC). The UPC code uses a series of vertical bars of varying width. These bars are detected as ten digits. The first five

digits identify the supplier or manufacturer of the item. The second five digits identify individual product code . The code also contains a check digit to ensure that the information read is correct or not.

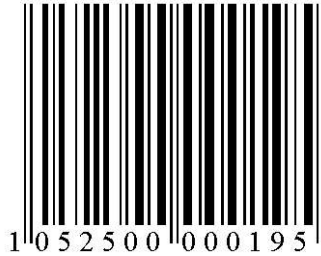


Fig 2.26 Bar Code

MAGNETIC INK CHARACTER READER (MICR)

It uses a special ink that can be magnetized so that the individual characters can be recognised by the reader from the shape of the wave form of the electrical pulse induced in the magnetism.

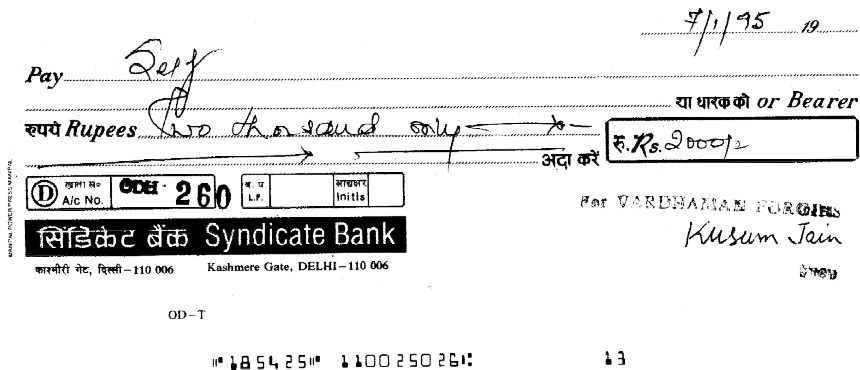


Fig 2.27 Magnetic Ink Character Reader

Such a system uses special inks containing magnetic particles of iron oxide. The characters are printed in a very distinctive style. When such documents are passed through a strong magnetic field, the Ink coded characters become magnetised. MICRs have got reading speeds of about 24000 documents per minute. Such a character recognition system is now being used by all commercial banks even in India. A MICR coded cheque.

OPTICAL SCANNER

An **optical scanner** is an input device that captures text or illustrations and translates them into a form that the computer recognizes. The optical scanner reads customized response forms and creates a data field directly, thus avoiding the expensive and tedious hand-keying of

data. The subjects themselves record the responses directly on the forms and clerical staff check for errors before entry.

Since the scanner does not differentiate between illustrations and text, all images are formatted the same, to edit the text requires OCR software.

OCR , optical character recognition, is software that reads text from paper and translates them into digital formats. The OCR software also enables the user to further edit the text with a word processor. Material that has been scanned into the computer can be faxed, sent by e-mail, printed, or edited.



Fig 2.28 A typical Flatbed Scanner

The scanner sends a light beam across an image and reads the image as a series of dots. Each dot is represented as a 1 or 0 and is stored in a file called a **bit map file**. The scanning process reduces the image that has been scanned to digital or numeric data that is stored on the computer as bits. This data is transferred to the CPU and processed by the CPU in the same manner as CD-ROM, sound, and video bits. The scanner requires an IRQ number, DMA channel, an I/O address, and room in upper memory for its BIOS, just as other peripherals do.

There are three major types of scanners. A **flatbed scanner** resembles a photocopier in appearance with a flat glass-scanning surface on which a document is placed for scanning. With a **sheet fed scanner**, the hard copy is inserted into a mechanism (just like Fax machine), which guides it through the scanner. **Handheld scanners** require the user to move the scanner across the page and usually scan areas only a few inches in width. They often include “stitching” software that will join two halves of a page together.

FINGERPRINT SCANNERS

Today fingerprint devices are by far the most popular form of biometric security used, with a variety of systems on the market intended for

general and mass market usage. Long gone are the huge bulky fingerprint scanners; now a fingerprint scanning device can be small enough to be incorporated into a laptop for security.

A fingerprint is made up of a pattern of ridges and furrows as well as characteristics that occur at Minutiae points (ridge bifurcation or a ridge ending). Fingerprint scanning essentially provides an identification of a person based on the acquisition and recognition of those unique patterns and ridges in a fingerprint.

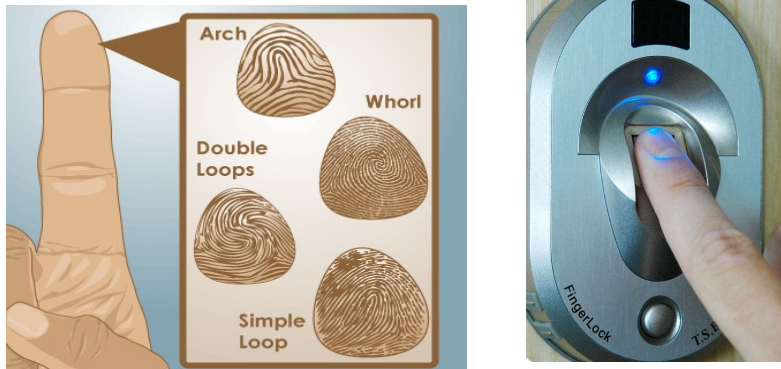


Fig. 2.29 Finger features and a Scanner.

PRINTERS

Printer is an output device which provides a hard copy on piece of paper and may resemble like a typewriter. The printers of different types have different speeds and qualities of printing. Their basic classification is:

- Character printers which print one character at a time.
- Line printers which print whole line at a time.
- Page printers (or image printer) printer) print a complete page or photograph at a time.

According to speed the printers are called low or high speed printers. Low speed printer can print 10 character per second to 300 lines per minute whereas high speed printers can print upto 3000 lines per minute. Character printers are low speed printers whereas line and page printers are always high speed printers.

The printers have been classified above as to how they print. There is one more classification which depends on the technology used in their manufacture. According to this consideration the printers are classified into the following two broad categories :

- (a) Impact printers
- (b) Non-Impact printers

Impact printers use electromechanical mechanism that causes hammers or pins to strike against a ribbon and paper to print the text. Non-Impact printers do not use electromechanical printing head to strike against ribbon and paper. They use thermal, chemical, electrostatic, laser beam or inkjet technology for printing the text. Usually, a non-impact type printer is faster than an impact type printer. The disadvantage of non-impact type printer is that they produce single copy of the text whereas impact printers can produce multiple copies of the text.

Thermal printers use a thermal (heating) process. Heat applied by the type head on to special paper of which surface turns into a different colour, creating characters where the heat was applied. Others use an electrostatic process in which top layer of a special type of paper is actually burned away, allowing a darker second layer to show through. Both processes require special paper that generally costs more than standard typing paper.

The main type of printers that are used widely are summarized in Fig below:

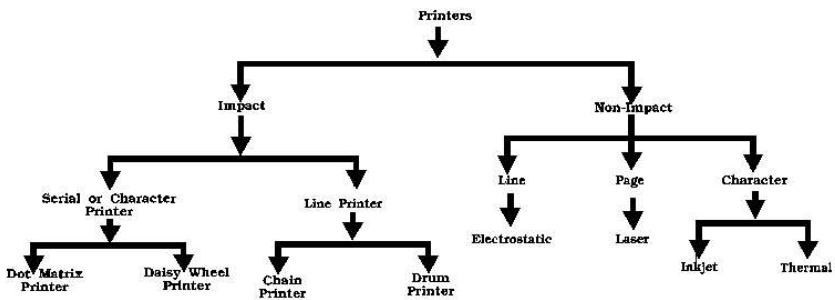


Fig 2.30 Classification of Printers

If you want a printed copy of a very long program or a report that is fifteen or twenty pages long, the normal procedure is to tell the computer to list or print the material on printer. The computer then sends data to printer at a speed, the printer can handle usually 18 characters per second to 200 characters per second. At this speed computer will have to wait for a long time. This problem is solved by interfacing a device known as printer buffer which is sometimes called printer spooler. The buffer generally takes only a second or two to accept all the material to be printed out. It then frees the computer for other work and sends the material in the buffer on to the printer at a speed the printer can handle.

SERIAL CHARACTER PRINTERS

These Printers are of following four types:

- Dot Matrix Printers

- Ink-jet electrostatic Printer.
- Daisy wheel Printer.
- Thermal Printer.

Dot Matrix Printers

Dot Matrix Printer produces letters and shapes by pushing a matrix of tiny wires into a ribbon which strikes the paper. Thus each letter consists of a pattern of tiny ink dots on the paper.

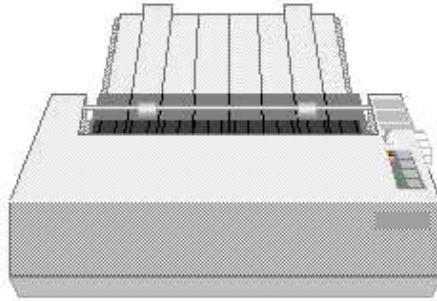


Fig 2.31 Dot Matrix Printer

Some printers form numbers letters and graphics with the help of a 5x7 matrix formed by dots. The printing head has seven needles arranged in a vertical line and five rows horizontally. Electronic pulses instruct the print head to choose the correct pattern of dots which are fired on to a typewriter ribbons, and then on to paper as the print head moves across.

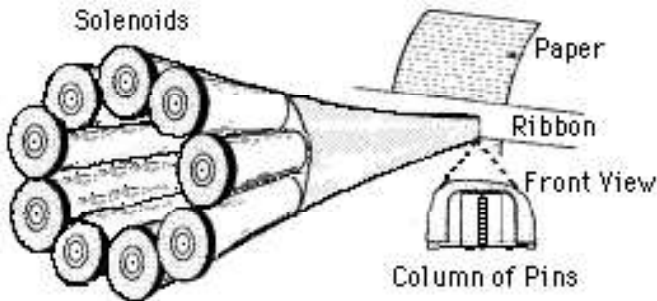


Fig 2.32 Printhead of Nine Pin Printer

The print out from 5x7 matrix printer is cruder than a print out from 9x9 matrix. A few of the dot matrix printer use an overlapping dot pattern to create higher quality output which is almost similar to electronic type writer. In that case the printer is called letter quality printer.

To print on 5x7 matrix, there are 9 pins contained in the print head. They are about one inch long and are driven by several hammers which force each pin into contact with the ink ribbon (and paper) at a certain time. The force on these hammers comes from the magnetic pull of small wire coils (solenoids) which are energized at a particular time, depending on the character to be printed. Timing of the signals sent to the solenoids is programmed into the printer for each character, and translated from information sent by the computer about which characters to print.

This type of printer is very versatile and is most often used, if the printing is wanted in the form of graphics. Some printers are capable of reproducing anything on paper that you can create on the computer's screen. This feature can be very useful if you regularly create reports that use a lot of charts and figures. The ability to print out graphic display is possible by dot matrix process. Dot matrix printer can be utilized to produce the photographs as well. Their speed is from 300 cps to 2000 cps.

Dot matrix printers can either be 80 column printer or 132 column printers depending on the maximum number of characters that can be printed in a line. A dot matrix printer may cost between Rs.5500 to Rs.1,00,000.

Ink-jet electrostatic printer

An inkjet printer system produces characters by spraying small ink droplets onto paper. A high frequency electrical generator causes a piezoelectric crystal to vibrate and throw off tiny ink droplets supplied from a reservoir. The droplets are then charged by an electrode and guided by horizontal and vertical deflection plates in the precise manner to form the desired characters on paper.

The speed of the inkjet printers lies in the range of 40-300 cps. The average life of an inkjet printer head is about 10 billion characters which is 5 times more than that of the print head of an impact type dot matrix printer. The inkjet printer uses ink cartridges containing a column of tiny heaters. The print quality of these printers is very near letter quality.

Daisy Wheel Printers

They have a spinning wheel having several radial arms. Each arm is meant for carrying a character at its end touching the periphery. As the wheel spins, a hammer punches them against the ribbon at the right moment. This wheel can be easily changed with a similar wheel within a moment. If you want different styles of letters or unusual symbols, you can change the daisy wheel. When the right letter hovers over the page a hammer comes down and slams the letter against the ribbon on to the paper.

Daisy wheel printers make sharper images on the page, but they cost more and print at a slower rate than dot matrix printers around 45

to 60 characters per second. The output in the form bits, that the computer produces, are turned into characters of different types by the printer.

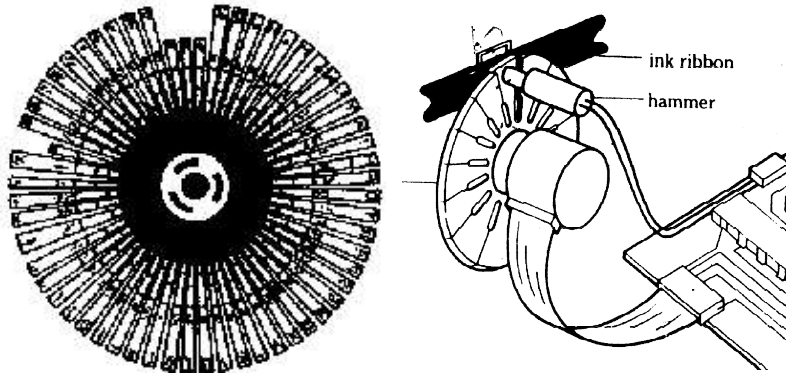


Fig 2.33 Daisy Wheel

Thermal Printer

A thermal printer uses a hot needles pressed against heat sensitive paper to form characters. It is not possible to produce multiple copies simultaneously, like impact printers.

The advantage of this type of printer over the dot matrix type is that the thermal unit is much quieter. The disadvantages are that a special type of paper must be used and it is not possible to produce multiple copies.

LINE PRINTERS

These are high speed printers which print a complete line at a time. The printing speed is at the rate of 20 lines to 80 lines per second. Typical Line printers can print about 150 lines per minute for numbers and 100 lines per minute for alphanumeric characters. Line printers are normally designed for heavy printing applications and can operate for few hours. The main disadvantage of using line printer is that the hammer movement and movement of drum, chain or band must be synchronised very accurately, otherwise lines get blurred.

Speed of such machines is limited by the speed of paper movement. A special type of paper is used which holds electro-static charge in the form of spot. The printing process involves placing of the spots to form the character. The paper then passes through a powdered ink bath whence powder-ink sticks to the position of spots. The ink is melted by high temperature so as to make a permanent record.

Line printers are three types :

- Drum printer
- Chain printer
- Band printers

Drum Printer

A drum printer has a cylindrical drum having the various characters embossed on its surface. Each character position along the text line contains a band of raised character set. There is a magnetically driven hammer in each character position of the line. The drum rotates rapidly, and one revolution is required to print each line a fast-acting hammer opposite each band strikes the paper against the proper character as it passes. Thus, in one rotation, hammers of several positions may "fire" when A appears. At the end of the rotation the line has been printed.

One complete set of character is embossed for each and every print position on a line. There are 132 characters generally printed in one line. The codes of all characters to be printed on one line are transmitted from the memory of computer to a storage unit in the printer. This unit is called printer buffer.

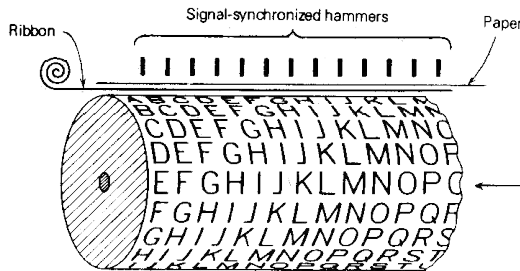


Fig 2.34 Drum Printer

The printer drum is costly and complete unit needs replacement even if one character is spoiled. The noise level of the drum printer is very high. Its speed varies from 200 to 2000 lines per minute.

Chain Printer

A chain printer has a steel band on which the various characters are embossed. The chain moves at a high but constant speed past the printing position.

Magnetically controlled hammers force the paper against the appropriate print head. Each chain may contain 4 sets of different styles for each of 64 (or 120) characters. For printing a line, all characters in the line are sent from the memory to the printer buffer register. As the band rotates, a hammer is activated when desired character as specified by buffer register comes in front of it.

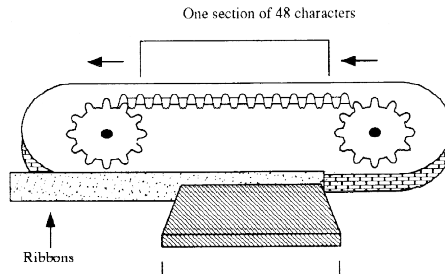


Fig 2.35 Chain Printer

Speeds of 50 lines per second can be achieved. Using continuous stationary with inter leaved carbons seven copies may be obtained. The noise level of the printer is high. Its speed lies in the range of 400-2400 lines per minute.

Band Printer

The band printer works in a similar manner to the chain printer, but here bands of steels are used instead of chains and can be easily replaced quite often. Thus different scripts may be used with the same printer.

PAGE PRINTERS

These print complete page itself at a time as it is done by Xerox like photocopying electrostatic process. These may be :

- Laser Printers.
- Inkjet printer
- Electrostatic Printer
- Photographic film outputs

Laser Printer

Laser printers are non-impact type printers. They print one page at a time. These printers use laser or other light source to produce an image on a photosensitive drum. The computer controls the laser beam to turn it on and off when it is cut back and forth across the drum.

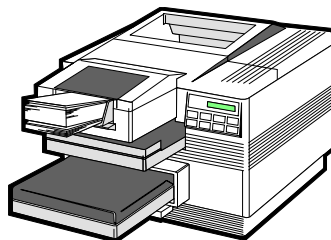


Fig 2.36 Laser Printer

An image is produced on the raster scan principal. The laser exposed areas attract toner (an ink powder). Thereafter the drum transfers the toner to the paper. The paper then moves to a fusing station where the toner is permanently fused on the paper with heat or pressure. After this the drum is discharged and cleaned. Now the drum is ready for processing the next information on the page.

Laser beam is the technology of transmitting powerful beams of light rays in a line without any reflection and refraction and losing its power. The full form of LASER is Light (frequency) Amplification of Stimulated Emission of Radiation.

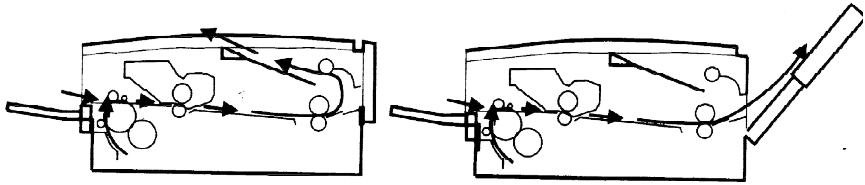


Fig 2.37 Working of a Laser Printer

The laser printers are quiet and they produce high quality output. These printers are expensive and require lot of maintenance. Low speed laser printers produce 10 pages or more per minute and are used with microcomputer. High speed laser-printers producing up to 300 pages per minute are manufactured for mini and large computers. The laser printers have become popular for large printing work. They are used for desk top publishing work also.

The difference between a Xerox machine and laser printer is that in a Xerox machine, the image of the original text is formed on the photosensitive drum with a Camera lens whereas in laser printer image is formed by scanning laser beamed using computer's control. The computer sends the characters of a page of the text to be printed in the printer's buffer memory.

INKJET PRINTER

An inkjet printer system produces characters by spraying small ink droplets onto paper. A high-frequency electrical generator causes a piezoelectric crystal to vibrate and throw off tiny ink droplets supplied from a reservoir. The droplets are then charged by an electrode and guided by horizontal and vertical deflection plates in the precise manner to form the desired characters on paper.

The speed of the inkjet printers lies in the range of 40-300 cps. The average life of an inkjet printer head is about 10 billion characters which is 5 times more than that of the print head of an impact type dot matrix printer. The inkjet printer uses ink cartridges containing a column of tiny heaters. The print quality of these printers is very near letter-quality.

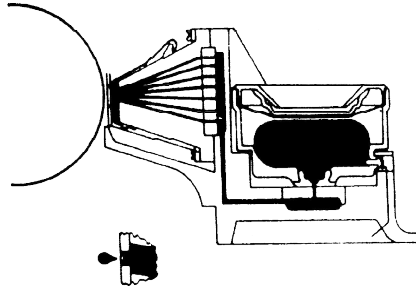


Fig 2.38 Inkjet printer

ELECTROSTATIC PRINTERS

A printer of this type forms characters by impressing static charges of electricity on special paper. The charges attract a toner to the charged spots for coloration. The needles apply charges on the paper in a serial fashion.

These printers are capable of very fast speed, up to 20,000 lines per minute. A variety of alphabets, graphics and very large drawings can be produced in a very short time. These devices can draw faster than a line plotter and even produce images similar to black and white photographs, with densities of several hundred dots per square inch.

PHOTOGRAPHIC FILM OUTPUTS

These means are used to store mass amount of data in a very small space. These are of two types:

- Micro film
- Microfiche

The output from the computer instead being stored or printed is displayed on high resolution CRT or VDU. Then it is photographed on a 16 mm or 35 mm film. A camera is controlled to make films of successive output pages on the screen of CRT. A special micro film reader is used to read the output. Some micro film readers also produce a hard copy using xerographic process.

Another method is known as COM (Computer Output on Microfilm/microfiche). In this first computer's output is recorded on to magnetic tape which is then fed into a machine called MICROFORM READER. It reads the magnetic tape and copies the data on to microforms. The copied data can be inspected by using a viewer which projects the copied information on to a screen. Full size copies can be printed.

Microfiche is a sheet of film of size 100 mmx 75mm (i.e., just little bigger than playing cards) on which information worth 80 pages can be recorded. The advantage of using microfiche is that it records information at the rate of 1 million characters per minute and that too in a small space. Microfiche is generally used to record display catalogues, archival documents.

PLOTTERS

Computers can be used to plot various figures and graphs with the help of plotter. Plotters are of two types.

Flat bed type

The pen moves up down, across or side to side.

Drum type

The pen moves up down and across. The paper provides sideway movement.

Plotters use pens of different colours to plot a graph as output after making desired calculations. The bar stretching across the paper moves horizontally, while the pens move along the bar. Coloured pens move over the surface of the paper according to the way, the computer has been programmed. Thus it saves the time which may be needed in plotting a graph manually.

MODEM

The modem is a device that converts digital information to analog by MODulating it on the sending end , and DEModulating the analog information into digital information at the receiving end.

The need to communicate between distant computers led to the use of the existing phone network for data transmission. Most phone lines were designed to transmit analog information - voices, while the computers and their devices work in digital form - pulses. So, in order to use an analog medium, a converter between the two systems is needed. This converter is the MODEM which performs MODulation and DEModulation of transmitted data. It accepts serial binary pulses from a device, modulates some property (amplitude, frequency, or phase) of an analog signal in order to send the signal in an analog medium, and performs the opposite process, enabling the analog information to arrive as digital pulses at the computer or device on the other side of connection.



Fig 2.39 A typical 56 kbps modem

Modems, in the beginning, were used mainly to communicate between DATA TERMINALS and a HOST COMPUTER. Later, the use of modems was extended to communicate between END COMPUTERS. This required more speed and the data rates increased from 300 bps in early days to 56 Kbps today. Today, transmission involves data compression techniques which increase the rates, error detection and error correction for more reliability.

Today's modems are used for different functions. They act as textual and voice mail systems, facsimiles, and are connected or integrated into cellular phones and in notebook computers enabling sending data from anywhere. The future might lead to new applications. Modem speeds are not expected to be increased much over today's 28.8 kbps. Further dramatic speed increases will require digital phone technology such as ISDN and fiber optic lines.

The following characteristics distinguish one modem from another:

bps

How fast the modem can transmit and receive data. At slow rates, modems are measured in terms of baud rates. The slowest rate is 300 baud (about 25 cps). At higher speeds, modems are measured in terms of bits per second (bps). The fastest modems run at 57,600 bps, although they can achieve even higher data transfer rates by compressing the data. Obviously, the faster the transmission rate, the faster you can send and receive data. Note, however, that you cannot receive data any faster than it is being sent. If, for example, the device sending data to your computer is sending it at 2,400 bps, you must receive it at 2,400 bps. It does not always pay, therefore, to have a very fast modem. In addition, some telephone lines are unable to transmit data reliably at very high rates.

Voice/data

Many modems support a switch to change between voice and data modes. In data mode, the modem acts like a regular modem. In voice mode, the modem acts like a regular telephone. Modems that support a voice/data switch have a built-in loudspeaker and microphone for voice communication.

Auto-answer

An auto-answer modem enables your computer to receive calls in your absence. This is only necessary if you are offering some type of computer service that people can call in to use.

Data compression

Some modems perform data compression, which enables them to send data at faster rates. However, the modem at the receiving end must be able to decompress the data using the same compression technique.

Flash memory

Some modems come with flash memory rather than conventional ROM, which means that the communications protocols can be easily updated if necessary.

Fax capability

Most modern modems are fax modems, which means that they can send and receive faxes.

To get the most out of a modem, you should have a communications software package, a program that simplifies the task of transferring data.

PLOTTERS

Computers can be used to plot various figures and graphs with the help of plotter. Plotters are of two types.

Flat bed type

The pen moves up down, across or side to side.

Drum type

The pen moves up down and across. The paper provides sideways movement.

Plotters use pens of different colours to plot a graph as output after making desired calculations. The bar stretching across the paper moves horizontally, while the pens move along the bar.

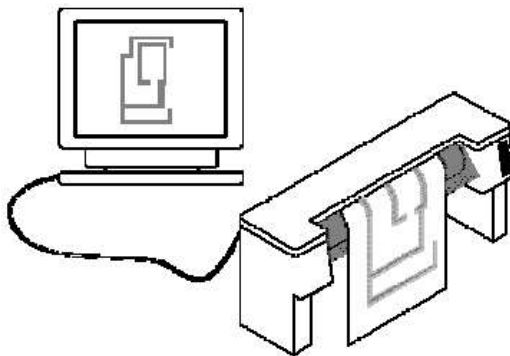


Fig 2.40 Plotter

Coloured pens move over the surface of the paper according to the way, the computer has been programmed. Thus it saves the time which may be needed in plotting a graph manually.

EXERCISE

1. Match the following -
 - i). Impact printer.
 - a. It prints by the impact of wheels of type against an inked ribbon and paper.
 - ii). Non-impact printer
 - b. It prints by chemical or photographic means.
 - iii). Chain printer
 - c. It prints by means of type striking against an inked ribbon and paper.
 - iv). Wheel printer
 - d. It prints by the impact of a chain of type against an inked ribbon and paper.
2. Fill in the blanks -
 - a. The visual display unit is also known as _____ or _____.
 - b. Daisy wheel printer costs _____ and print at slower rate than dot matrix printer.
 - c. _____ printer requires special paper.
 - d. A set of instruction is called a _____.
 - e. Magnetic disk is a storage medium that lends itself to _____ processing.
 - f. Memories which can be both read from and written into are called _____ memories.
3. Briefly describe the two types of printers, Impact & Non-impact.
4. What is a bit? What is a binary code?
5. Differentiate between a bit, a byte, and a word.
6. How many bytes can be stored in the primary storage unit of a 512 K-byte computer?
7. Distinguish between
 - (a) CPU and ALU
 - (b) internal and external memory.
 - (c) Accumulator and Register.
 - (d) Processor and Central Processing Unit.
8. Draw a block Diagram of the main units of a computer hardware system. Show the flow of data between the various peripheral devices attached to the CPU.
9. Explain the meaning of GIGO. Why is GIGO an important concept in data input to a computer?
10. What are optical bar readers? How are they each used for data input?
11. What dose MICR stand for? Who uses MICR? What are the advantages and disadvantages of MICR?

12. How is optical character recognition (OCR) used for data input? What are the advantages and disadvantages of OCR?
13. How is a light pen used? Is it an input or output device?
14. What are line printers? Why are they so called ?
15. What is the difference between a drum printer and a chain printer?
16. What is a page printer? Why is it so called? In what type of computer installations is it used?
17. How does an ink jet printer create characters on a print page? What are some of the advantages and disadvantages of ink jet printers?
18. Write brief notes on each of the following types of printer. Make clear the differences between them in terms of speed, cost and method of operation, and suggest suitable applications.
 - (a) Daisy Wheel printer
 - (b) Chain printer
 - (c) Dot Matrix printer.
 - (d) Laser printer.
19.
 - a. What's an impact printer?
 - b. Identify and discuss two types of impact character printers.
20.
 - a. What's a Non-impact printer?
 - b. Identify and discuss the ink jet approach to non-impact printing.
21. What is a dot-matrix printer? How is it different from a daisywheel printer?
22. Give the full form of the following abbreviations :
VDT, CRT, CAD, OMR, OCR and MICR.
23. Write the names of any six types of printers.
24. Distinguish between a character printer and a line printer.

© Copyright 2002 - Silicon Media Press

All rights reserved. No part of this publication, may be reproduced or distributed in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, or stored in a database or retrieval system without the prior permission in writing of the publisher.

Microsoft ® , Windows ® ,MS ® , and MS-DOS ® are registered trademarks of Microsoft Corporation. All other brand and product names are trademarks or registered trademarks of their respective companies.

Every effort has been made to supply complete and accurate information. Silicon Media Press does not guarantee the accuracy or completeness of any information and assumes no responsibility for its use.

I S B N - 8 1 - 8 7 8 7 0 - 3 5 - 4

First Edition 2002

Second Edition 2009

The export rights of this book are vested solely with the publisher.

Published by Silicon Media Press, Regd. Off. I-19, Lajpat Nagar - II, New Delhi. Works : KJ-75, Kavi Nagar, Ghaziabad. B.O B-24, ECOTECH III, Graeter Noida. 9810626977. www.siliconmedia.in, e: info@siliconmedia.in and **Printed by** : Siddharth Printers, B-23C, ECOTECH III, Graeter Noida.

**Dedicated to our lovely kids
Priyasha, Kashish & Krishiv**

Contents

INSIDE PC	2
SOFTWARE	3
Firmware	5
Humanware	5
HARDWARE	5
CENTRAL PROCESSING UNIT (CPU)	6
Arithmetic and Logic Unit (ALU)	7
Control Unit	8
Memory	8
Input/Output Bus	10
Input/Output Peripheral Devices	10
MOORE'S LAW FOR MICROPROCESSORS	11
EVOLUTION OF COMPUTERS	12
PCI	14
IDE	14
Interrupts	15
Multiprocessors	16
DMA	16
The Future Of I/O	17
PC or Mac?	17
SERVERS	18
Eight Ways Server	19
DISK ARRAYS AND DATA STRIPING	20
RAID	21
EXERCISE	24
COMPONENTS OF COMPUTER	26
SYSTEM UNIT	27
Motherboard	29
The Processor	31
RAM	32
THE EXPANSION SLOT	34
Magnetic Disks	35
SCSI (SMALL COMPUTER SYSTEM INTERFACE)	36
Floppy Disk	37
Video or Optical Disk (CDROM)	38
SMPS	40
OTHER PERIPHERAL DEVICES	41

Key board	41
Mouse	43
Visual Display Unit (VDU)	43
Light Pen	48
WebCams	48
Joystick	49
DIGITIZERS	49
Optical Marks Reader (OMR)	49
Optical Character Reader (OCR)	50
BAR CODE READERS	50
Magnetic Ink Character Reader (MICR)	51
Optical Scanner	51
FingerPrint Scanners	52
PRINTERS	53
Serial Character Printers	54
Line Printers	57
Page Printers	59
Inkjet Printer	60
Electrostatic Printers	61
Photographic Film Outputs	61
Plotters	62
Modem	62
Plotters	64
EXERCISE	65

THE MICROPROCESSOR 68

Arithmetic /Logic unit	68
Control Unit	69
COMPUTER REGISTERS	69
80x86 Registers	71
Common Bus System	74
The Fetch Execute Cycle	76
MEMORY	76
INPUT/OUTPUT (I/O)	77
CLASSIFICATION OF MICROPROCESSOR	77
CISC Processors	78
RISC Processors	79
EPIC Processor	80
TRENDS OF MICROPROCESSORS	81
First Generation Microprocessors	83
Second Generation Microprocessors	84

Third Generation Microprocessors	85
80486	86
Fourth Generation Microprocessors	88
Fifth Generation Microprocessors	91
Sixth Generation Microprocessors	93
Seventh Generation Microprocessors	96
Eighth Generation Microprocessors	97
AMD AND OTHER CHIPS	99
EXERCISE	110
MEMORY CONCEPTS	112
ROM-BIOS	112
How Does The BIOS Work?	113
How Do You Add to the ROM-BIOS?	113
RAM AND ROM	114
ROM ,PROM, EPROMs, EEPROMs, and FLASH	
EPROMs	115
RAM	115
Cache Memory	120
Hardware & Software Limits for Memory	120
CONVENTIONAL MEMORY: THE 640 K LIMIT	121
Expanded memory	124
EXTENDED MEMORY	126
EXERCISE	130
THE MOTHERBOARDS	132
SPECIFICATIONS	132
HISTORY OF MOTHERBOARDS	132
XT board	133
AT Motherboard	134
Computer Form factor	135
ATX Motherboards	135
Power Supply Connectors	137
PHYSICAL LAYOUT OF MOTHERBOARDS	138
Size	139
Chip Sets	140
I/O Buses	142
Memory Module	142
MOTHERBOARD COMPONENTS	142
THE MICROPROCESSORS	143
THE CO-PROCESSOR	144
MEMORY	145

BIOS	146
ORGANIZATION OF ROM-BIOS	146
BIOS FEATURES IN ADVANCED COMPUTERS	147
HOW DOES THE BIOS WORK	148
CMOS Setup	149
EXPANSION SLOTS	158
Processor Bus	159
The Memory Bus	160
Data Bus	160
Address Bus	161
Control Bus	161
TYPES OF I/O BUS	161
ISA Bus	161
The EISA BUS	162
The MCA bus	164
VESA Local Bus	166
AGP Bus	167
PCI Bus	167
PCI Express	168
SUPPORT CIRCUITRY	170
RESET LOGIC	170
CPU NUCLEUS LOGIC	170
EXERCISE	171

SECONDARY STORAGE - I 174

FLOPPY DISK DRIVES	174
THEORY AND OPERATION	174
DRIVE COMPONENTS	174
DRIVE CONFIGURATION DEVICES	175
TERMINATOR	176
DOS REQUIREMENT	176
BASIC FORMATTING	176
DRIVE CABLING	177
TYPES OF DISK DRIVE	177
FLOPPY DISK CONTROLLER	178
Input/Output Ports	179
DISK FORMAT	180
EXERCISE	182

SECONDLY STORAGE - II 184

HARD DISK DRIVE	184
-----------------------	-----

HARD DISK OPERATION	185
WRITE OPERATION	185
Physical Operation	185
Construction	186
Disk Platters	186
Air filters	189
Head actuator	190
Spindle motors	191
Logic boards	191
Hard disk Features	192
HARD DISK CONTROLLERS AND INTERFACES	195
SCSI (SMALL COMPUTER SYSTEM INTERFACE) ..	195
IDE & ATA	196
Drive Configuration	197
HISTORY OF HARD DISK	198
SATA Technology	205
EXERCISE	207

SECONDARY STORAGE - III 210

CARTRIDGE TAPE DRIVE	210
TAPE OPERATION	210
DAT	212
8 mm	213
INTERNAL AND EXTERNAL CTDs	215
CTD INSTALLATION	215
BACKUP USING SOFTWARE	216
CD- ROM (COMPACT-DISK-READ-ONLY-MEMORY)	217
Disk Structure	218
Data Organizations	218
Drive Mechanism	219
Reading the data	219
MEDIA FOR CD-RECORDING	220
Compact Disk Recordable (CD-R)	220
DEVELOPING INDUSTRY STANDARDS	224
INTERFACE TYPE	225
DVD	225
MAGNETO-OPTICAL TECHNOLOGY	227
SUPER-FLOPPIES (ZIP)	228
FLASH DRIVE	229
History of USB	229
EXERCISE	231

KEYBOARD & MOUSE	234
KEYBOARD	234
How Keyboards Work	234
Enhanced Keyboards	236
Compatible Keyboards	236
KEYBOARD TECHNOLOGIES	237
Capacitive keyboards	237
Hard-Contact Keyboards	238
KEYBOARD CABLING	239
Keyboard Organization	240
keyboard interface	242
MOUSE	244
Opto-Electronic Mouse	245
Optical Mouse	246
Wireless Mouse	247
EXERCISE	249
SMPS	252
SWITCHED MODE POWER SUPPLY	252
POWER MANAGEMENT	253
POWER SUPPLY CONNECTOR	255
POWER SUPPLY TECHNOLOGIES	255
Working Principle	256
Calculating PC Power requirements	256
Power Supply Fan	257
Power Supply Connections	257
Power Line Irregularities	257
Circuit Explanation	258
POWER SUPPLY	258
EXERCISE	261
DISPLAYS & VIDEO ADAPTER	264
COMPUTER DISPLAYS - MONITORS	264
Cathode Ray Tubes (CRTs)	264
COLOUR CRT	267
LCD (LIQUID CRYSTAL DISPLAY) MONITOR	268
LCD Technology	269
GRAPHICS CARD	271
Refresh Memory	271
Resolution	272
Colour depth	273

Dot pitch	273
SPEED	273
Video Memory	274
Video Graphics Array (VGA)	275
RAMDAC	275
VESA ADAPTER	276
EXERCISE	277
LAPTOPS	280
LAPTOP HISTORY	280
Laptop Today	281
Advantages of Laptop	282
INSIDE LAPTOP	284
Motherboard	284
Laptop Processors	284
Laptop Memory and Storage	286
Proprietary memory modules	287
Laptop Screen, Graphics and Sound	287
Laptop Batteries	288
Whitebooks - ASSEMBLED LAPTOPS	289
ASSEMBLING COMPUTERS.....	292
INSTALLING THE CPU	293
INSTALLING THE RAM	294
MOUNTING THE MAINBOARD	294
INSTALLING HARD DISK	295
Physical Installation	296
INSTALLING FLOPPY DRIVE	297
INSTALLING AGP/PCI DEVICES	297
READY TO GO.....	298
Important Plug Types	298
BIOS SETUP	300
Configuring HARD DISK on the Machine	300
Formatting and Software Installation	301
DOS Formatting	309
EXERCISE	312
UPGRADING & TROUBLESHOOTING.....	314
NECESSARY TOOLS	314
BASIC WORKING TECHNIQUES	315
UPGRADING PC'S.....	315
DISPLAYS: THE CARDS	316

MONITORS	317
KEYBOARD	318
MEMORY	318
MOUSE	319
TROUBLESHOOTING YOUR SYSTEM	320
THE GROUND RULES	320
Power Supply Problems	321
Motherboard Problems	322
Keyboard Maintenance	323
Monitors Problem	324
Mouse Cleaning and Maintenance	325
Hard Disks Problems	326
TroubleShooting Printers	328
Serial Ports Problems	333
Modem Problems	334

Acknowledgments

We would like to thank all of the many people who helped bring this book to print. It's really fascinating for us to write a book, because what begins as a few thoughts ends up as a physical, tangible thing that goes into the hands of many readers. The trip from thought to thing is a long one and involves inputs from lots of people. Thanks goes to all of our editors, art personnel and staff at Silicon Media.

Munishwar Gulati

I would especially want to thank my wife and co-author Mini Gulati, who always cooperated and went through the whole book to rectify me at many places.

Mini Gulati

It has been a delight working with my husband and author Munishwar Gulati. He persuaded me to write, what I knew and rectified it, to give it a shape of chapter. My deep appreciation goes to him for his constant enthusiasm and hard working.

About this book

As you read the book, you will find the answers to many of the questions you have about Computer Hardware And Peripherals.

How the book is organised

Chapter 1 introduces you to Basics of Computers and various general concepts of Computers.

Chapter 2 tells you the various components and peripherals of computers.

Chapter 3 explains you the concept of Microprocessorm its classification on the basis of architecture and Historty of Microprocessor.

As you move on to **Chapter 4**, you are introduced to various types of memory used in Computers.

Chapter 5 covers all about motherboards, its layouts, BIOS, expansion slots, various types of buses and setting of BIOS.

Chapter 6 tells you about various types of floppy disks.

Chapter 7 tells you about Hard Disk Drives and its history from 20 MB to 2 TB.

Chapter 8 tells you to about various other secondary storage devices like tape drives, CD-ROM and wriers and DVD.

Chapter 9 tells you all about keyboard and Mouse.

Chapter 10 tells you about working with SMPS.

Chapter 11 discusses various display and video adapters.

Chaper 12 covers all about Laptops.

Chapter 13 tells you all about assembling Computers.

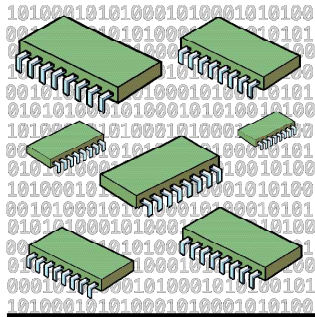
Chapter 14 tells you all about upgrading various components in Computers.

This book is a sincere effort for explaining the concepts of Computer ardware & Peripherals. We sincerely hope that you find this work to be informative and enjoyable.

As a reader, you are the most important critic and commentator of our books. We value your opinion and want to know what we are doing right, what we could do better, what areas you would like to see us publish in, and any other words of wisdom you are willing to pass.

As the publishing manager of the group that created this book, We welcome your comments. You can e-mail at info@siliconmedia.in.

Authors



CHAPTER 3

THE MICROPROCESSOR

CLASSIFICATION OF MICROPROCESSOR
TRENDS OF MICROPROCESSORS

The Microprocessor

A microprocessor is an integrated circuit built on a tiny piece of silicon. It contains thousands, or even millions, of transistors, which are interconnected via superfine traces of aluminum. The transistors work together to store and manipulate data so that the microprocessor can perform a wide variety of useful functions. The particular functions a microprocessor performs are dictated by software.

In general, a digital computer consists of an interconnected system of *processors*, *memories*, and *input/output* devices. These devices are connected by a collection of parallel wires called a *bus*. The lines in the bus can be classified into three groups: *address bus*, *data bus*, and *control bus*. In a simple microcomputer there is one microprocessor integrated circuit called the *central processing unit* (CPU) which acts as the “brain” of the computer. The CPU is composed of several distinct parts: the *Control Unit* fetches instructions from memory and decodes them, *Arithmetic and Logic Unit* (ALU), and *registers* which are special high-speed memory used to store temporary results and certain control information. Block diagram representation of the internal operation of the computer.

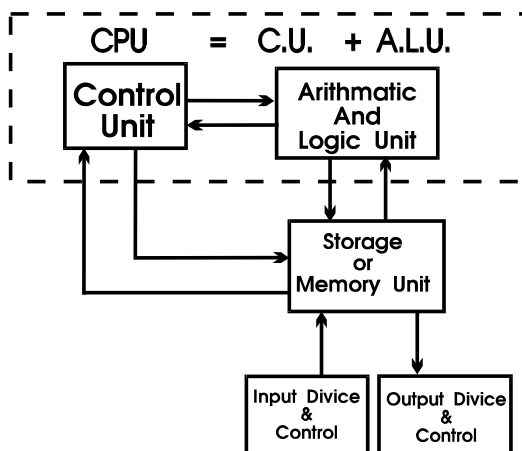


Fig 3.1 Block diagram of CPU

ARITHMETIC /LOGIC UNIT

In this area of the microprocessor, computing functions are performed on data. The ALU performs arithmetic operations such as addition and subtraction, and logic operations such as AND, OR and exclusive OR. Results are stored either in registers or in memory or sent to output devices.

CONTROL UNIT

The control unit provides the necessary timing and control signals to all the operations in the microcomputer. It controls the flow of data between the microprocessor and peripherals (including memory).

At present, microprocessors are available from many manufacturers. The earliest microprocessors include the Intel 4004, 8080A and 8085, Zilog Z80, Motorola 6800 and 6809, and MOS Technology 6500 series. The design of the IBM Personal Computer is based on the Intel 8088 microprocessor. This was followed by 8086, 80186, 80286 (10/12/16 MHz), 80386 and 80486. These chips are classified as first (80186), second (80286), third (80386) and fourth (80486) generation of X86 family. Similarly, the Pentium was the first fifth-generation X86. It was soon followed by the NexGen Nx586, Cyrix 6x86 and, eventually, the AMD K5. Intel's revised Pentium MMX and Centaur's odd little C6 followed on late in the fifth generation's market life. More advanced units like the Pentium Pro, Pentium II/III/IV, 6x86MX and AMD K6 are regarded as sixth generation designs.

COMPUTER REGISTERS

Computer instructions are normally stored in consecutive memory locations and are executed one by one. The control reads an instruction from a specific address in memory and executes it. Then continues by reading the next instruction in sequence and executes it, and so on.

This type of instruction sequencing needs a counter to calculate the address of the next instruction after execution of the current instruction is completed. It is also necessary to provide a register in the control unit for storing the instruction code after it is read from memory. The computer needs processor registers for manipulating data and a register for holding a memory address. There are various other registers required for different purpose. These registers are located in the processor, instead of in RAM, so data can be accessed and stored faster. The following are the basic registers, which a processor can have.

Data register (DR)

This register holds the operand read from memory.

Accumulator (AC)

The accumulator is an 8 bit register that is part of the arithmetic logic unit (ALU). This register is used to store 8 bit data and to perform arithmetic and logical operations. It stores a previously calculated value or a value loaded from the main memory. The accumulator is also identified as register A.

Instruction Register (IR)

Stores a copy of the instruction loaded from main memory.

Temporary Instruction Register (TIR)

As the CPU evaluates exactly what an instruction is supposed to do, it stores the edited instruction in the TIR.

Address Mask (AMASK)

When the CPU needs to know the address of a target word that an instruction is using, the AMASK is AND'ed with the instruction to eliminate the opcode, leaving only the desired address.

Memory Address Register (MAR)

This register contains the address of the place the CPU wants to work with in the main memory. It is directly connected to the RAM chips on the motherboard.

Memory Buffer Register

This register contains the word that was either loaded from main memory or that is going to be stored in main memory. It is also directly connected to the RAM chips on the motherboard.

Program Counter (PC)

This 16-bit register deals with the fourth operation, sequencing the execution of instructions. This register is a memory pointer. Memory locations have 16 bit address, and that is why this is a 16 bit register.

It stores the address of the macro-instruction currently being executed. The PC goes through a counting sequence and causes the computer to read sequential instructions previously stored in memory. Instruction words are read and executed in sequence unless a branch instruction is encountered. A branch instruction calls for a transfer to an instruction not in sequence in the program. The address part of a branch instruction is transferred to PC to become the address of the next instruction. To read an instruction, the content of PC is taken as the address for memory and a memory read cycle is initiated. PC is then incremented by one, so it holds the address of the next instruction in sequence.

Two registers are used for input and output. The **input register** (INPR) receives an 8-bit character from an input device. The **output register** (OUTR) holds an 8-bit character for an output device.

Flags

The ALU includes five flip flops that are set or reset according to data conditions in the accumulator and other registers. The microprocessor uses them to perform the third operation namely testing for data conditions.

For example, after an addition of two numbers, if the sum in the accumulator is larger than eight bits, the flip flop that is used to indicate a carry, called the carry flag (CY), is set to one. When an arithmetic operation results in zero, the flip flop called the Zero flag (Z) is set to one. The 8085/8080A has five flags to indicate five different types of data conditions. They are called Zero (Z), Carry (CY), Sign (S), Parity (P) and Auxiliary Carry (AC) flags. The most commonly used flags are zero and Carry.

These flags have critical importance in the decision making process of the microprocessor. The conditions (set or reset) of the flags are tested through software instructions. For example the instruction JC (Jump on Carry) is implemented to change the sequence of a program when the CY flag is set. The importance of the flags cannot be emphasized enough; they will be discussed again in applications of conditional jump instructions.

Stack pointer (SP)

The stack pointer is also a 16 bit register used as a memory pointer, initially, it will be called the stack pointer register to emphasize that it is a register. It points to a memory location in R/W memory, called the stack. The beginning of the stack is defined by loading a 16 bit address in the stack pointer (register).

80X86 REGISTERS

The 80x86 registers shown here are considered program visible since they are addressable by all applications. On the 80286 and above there are also program invisible registers used to control and operate the protected memory system. The earlier 8086, 8088, and 80286 contain 16-bit internal architectures. The 80386 and above contain full 32-bit internal architectures.

Multipurpose Registers

Many of these registers have a dedicated purpose for certain instructions, but are mostly considered general purpose. There are some restrictions on which registers can be used for certain instructions, but most of these restrictions are lifted when used as 32-bit registers.

EAX The *accumulator* register is typically used for arithmetic and logic instructions. EAX is 32-bits wide. AX is 16-bits wide and holds the least significant word of EAX. AH and AL are each 8-bits wide hold the most significant (high) and least significant (low) bytes of AX respectively.

EBX The *base index* register is sometimes used to hold an address offset. EBX is 32-bits wide, BX is 16-bits wide, and BH and BL are 8-bits wide.

ECX Sometimes holds the "*count*" for various instructions. ECX is 32-bits wide, CX is 16-bits wide, and CH and CL are 8-bits wide.

EDX

The *data* register holds part of the result of a multiplication or holds part of the dividend for a division. EDX is 32-bits wide, DX is 16-bits wide, and DH and DL are 8-bits wide.

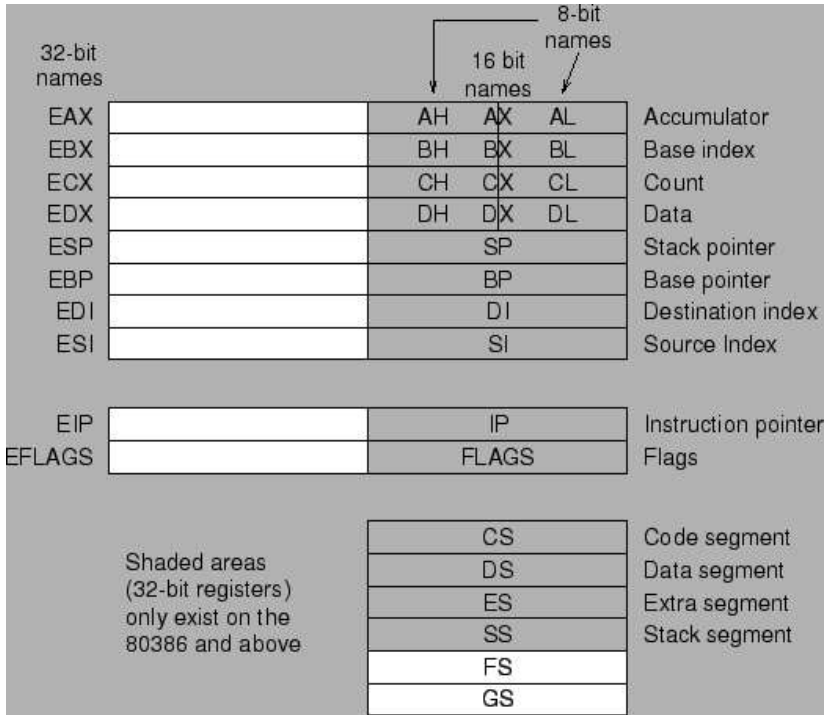


Fig 3.2 Registers in 32-bit architecture

EBP

The *base pointer* register typically holds the offset address for a stack frame used by a subroutine to reference local variables and procedure arguments. The 16-bit version is BP.

EDI

The *destination index* register often addresses destination string data. The 16-bit version is DI.

ESI

The *source index* register often addresses source string data. The 16-bit version is SI.

Special purpose registers

EIP

The *instruction pointer* (called the *program counter* on other machines) holds the offset address of the next instruction the CPU will execute. The 16-bit version is IP. This register is typically modified with jump and call instructions.

ESP

The *stack pointer* register holds the offset address for the processor stack. The 16-bit version is SP.

EFLAGS

The *flags* register indicates the current condition of the processor as well as control its operation. The

specific meaning of certain bits in this register are detailed below. The 16-bit version is FLAGS.

The FLAGS and EFLAGS register bit assignments:

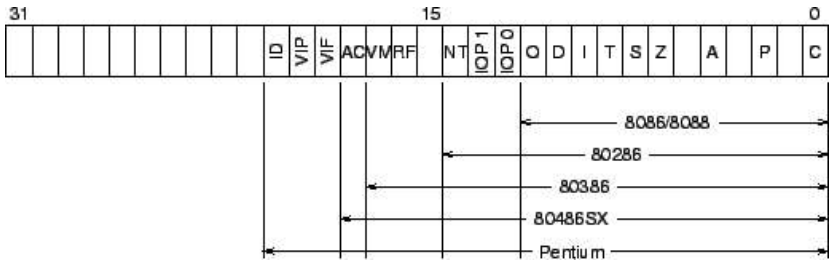


Fig 3.3 Various FLAGS & EFLAGS

C (carry)	Holds the carry (borrow) after addition (subtraction). Often used to indicate an error condition.
P (parity)	0 for odd parity, 1 for even parity. Not used much.
A (auxiliary carry)	Holds carry (borrow) between bits 3 and 4 (used mainly for BCD arithmetic).
Z (zero)	Set (Z = 1) if the result of an arithmetic operation is zero.
S (sign)	Set (S = 1) if the result of an arithmetic operation is negative.
T (trap)	Used for debugging.
I (interrupt)	Controls the operation of the INTR pin (if I = 1, the INTR pin is enabled).
D (direction)	For setting increment or decrement mode for string instructions.
O (overflow)	Set when an arithmetic result has exceeded the capacity of the machine.
IOPL (I/O privilege level)	Selects privilege level for I/O.
NT (nested task)	Current task is nested within another task.
RF (resume)	Used for debugging.
VM (virtual mode)	Selects 8086 virtual mode in a protected mode system.
AC (alignment check)	Set if a word (doubleword) is addressed on a non-word (non-doubleword) boundary.
ID (identification)	Indicates whether or not the CPUID instruction is supported.

Segment Registers

Additional registers, called segment registers, generate memory addresses when combined with another value called the offset. The details of how these registers are used differ greatly depending on whether the processor is in “real mode” or “protected mode”.

- CS** The *code segment* register references the the portion of memory that holds the currently executing program instructions (i.e. the code).
- DS** The *data segment* is for referencing a portion of memory that holds program data.
- ES** The *extra segment* is for referencing another portion of memory that holds other program data.
- SS** The *stack segment* is for referencing the processor stack memory.
- FS and GS** Other supplemental segment registers.

COMMON BUS SYSTEM

A personal computer may transfer data from disk to CPU, from CPU to memory, or from memory to the display adapter. It is not possible to have separate circuits between every pair of devices. The solution is a **Bus**. **The Bus is simply a common set of wires that connect all the computer devices and chips together.** Some of these wires are used to transmit data. Some send housekeeping signals, like the clock pulse. Some transmit a number (the “address”) that identifies a particular device or memory location. The computer chips watch the address wires and respond when their identifying number is transmitted. They then transfer data on the other wires.

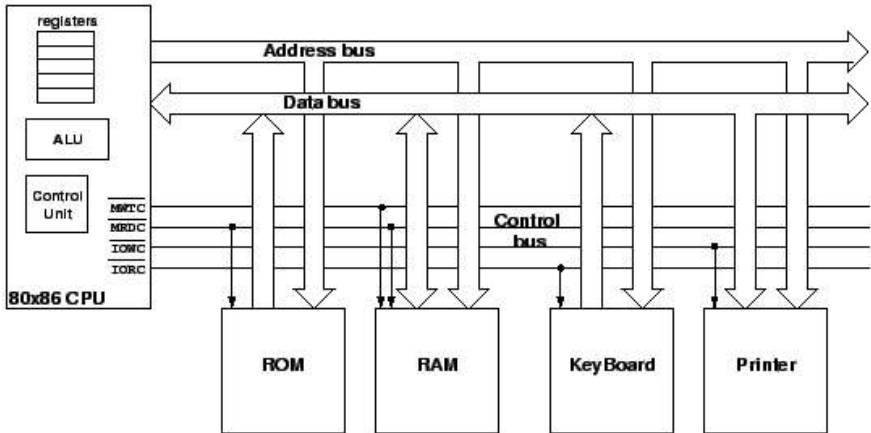


Fig 3.4 Common Bus System

Although earlier computers had one bus and all of its devices and chips ran at the same speed. But, starting with machines that used the 386 CPU, the memory and CPU of the system ran faster than the I/O devices. Today, the mainboard of the computer has separate bus connections to the CPU, memory, and I/O devices. The CPU bus may run at 100, 133, 200, or 266 MHz. Independently the memory will run at 100, 133, or 166 MHz and the I/O will run at 33 or 66 MHz or 100 Mhz.

Whether there is more than one Bus, or one Bus with different speeds, is a matter of perspective. A car drives down the local streets at 30 km per hour. Then it turns onto a highway and accelerates to 65. Is there one road system, or two? The important thing is that there is a connection that allows a flow of traffic between the two speed zones. Within the PC, data can flow from any chip to any other chip, but different parts of the path run at different speeds.

The buses could be categorized into three types :

Data Bus

The data bus carries the information to or from the memory. These lines are bidirectional data flow in both directions between the Processing Unit and peripheral devices. The four main types of data bus used in PCs are PCI, AGP, VESA Local, and ISA. Of these four, only the PCI bus and the AGP bus are widely used today. The VESA Local Bus and ISA bus are rarely found today in new computers, as both standards have been officially deprecated as viable hardware interfaces for high-speed PCs. The data bus influences the microprocessor architecture considerably. It determines the word length and the register size of a microprocessor. ISA and VESA Local are 8-bit bus, thus the largest number that can appear on these data bus is 11111111 (255_{10}). PCI is a 16 bit bus where as AGP is 32 bit bus.

Address Bus

The address bus carries the memory address. It's width is equal to the number of bits in memory address register. The address bus is unidirectional as bits flow in one direction only - from the Processing Unit to peripheral devices.

In a computer system each peripheral or memory location is identified with a binary number, called an address, and the address bus is used to carry a 16 bit address. It is a group of sixteen lines generally identified as A_0 to A_{15} . The number of address lines of the Processing Unit determines its capacity to identify different memory locations (or peripherals). For example, sixteen address lines are capable of addressing $2^{16} = 65536$ (generally known as 64K) memory locations.

Control Bus

It carries control signal among various units of a computers. It is comprised of various single lines that carry synchronization signals.

To communicate with a memory - for example, to read an instruction from a memory location- the Processing Unit places the 16 bit address on the address bus . The address on the bus is decoded by an external logic circuit, and the memory location is identified. The Processing Unit sends a pulse called Memory Read as the control

signal. The pulse activates the memory chip, and the contents of the memory location (8-bit data) are placed on the data bus and brought inside the microprocessor.

THE FETCH EXECUTE CYCLE

On most modern computers the memory system is *byte addressable*, meaning that memory is thought of as a linear array of bytes (i.e. each byte has a unique address). A program's data and instructions reside in memory. The CPU interprets a program's instructions via a *fetch-decode-execute* cycle which repeatedly carries out the following steps:

Fetch

Fetch the next instruction from memory (the address of the instruction is stored in special purpose register called the *instruction pointer* (IP) (or *program counter* (PC))). The IP is incremented to point to the next instruction.

Decode

Decode the instruction. If the instruction uses data in memory, it is fetched into internal CPU registers.

Execute

Execute the instruction. Store results in the proper place. Go to step 1 (Note that certain flow-control instructions may alter the IP).

Although the concentration is on CPU and memory, there are additional components in a computer such as the I/O modules which can interact with the processor. In an improved instruction execution cycle, we also have a third cycle known as the interrupt cycle.

MEMORY

Memory plays an important role in a microcomputer system. There are two types of memory - Read/Write Memory and Read Only Memory. The R/W Memory is made of registers, and each register is a group of flip-flops that stores bits of information. The microprocessor can read from or write into this memory. The second type of memory, the ROM, stores information permanently in the form of diodes; a group of diodes can be viewed as a register. The microprocessor can only read information from the ROM; it cannot write into this memory.

In a memory chip, all registers are arranged in a sequence and identified by binary numbers called memory addresses. To read from or write into a memory location, the microprocessor places the address on the address bus. The decoder decodes the address and identifies the register. The control signal R/W enables the I/O lines, and the data byte is either read from or stored in the memory location.

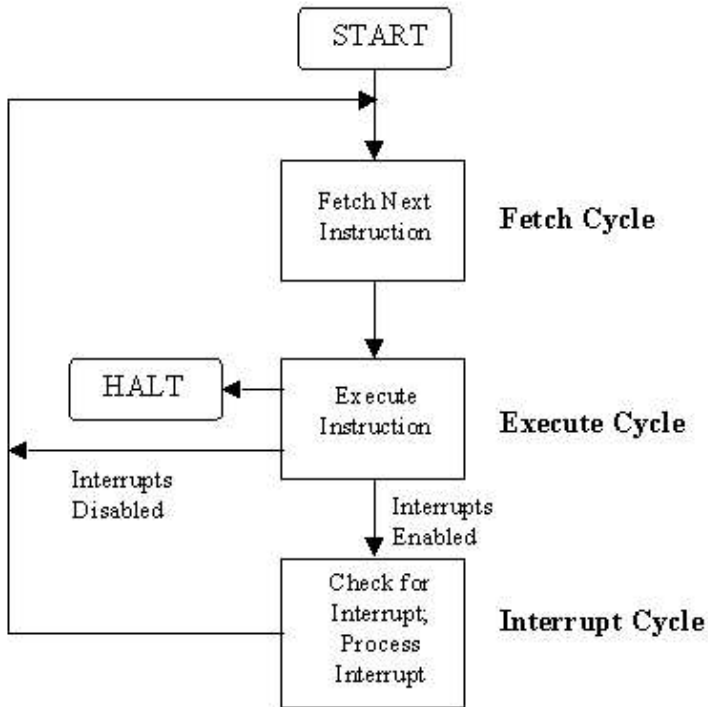


Fig 3.5 Various Cycles of Hypothetical Computer

INPUT/OUTPUT (I/O)

Microprocessor communicates with the Input/Output devices in the computer. The microprocessor accepts the input in binary form from input devices such as keyboards or floppy disks, and sends the output to Display unit or printer.

CLASSIFICATION OF MICROPROCESSOR

A microprocessor incorporates most or all of the functions of a central processing unit (CPU) on a single integrated circuit (IC). The first microprocessors emerged in the early 1970s and were used for electronic calculators, using binary-coded decimal (BCD) arithmetic on 4-bit words. Other embedded uses of 4- and 8-bit microprocessors, such as terminals, printers, various kinds of automation etc, followed rather quickly. Affordable 8-bit microprocessors with 16-bit addressing also led to the first general purpose microcomputers in the mid-1970s.

Classifications are characteristics that determine the type of microprocessor in a broad sense. Characteristics of a microprocessor are normally determined at the beginning of its design as a set of

design guide lines. For example, early in a chip's design it will be decided whether the chip will support variable length instructions or not. This initial design choice will have a significant impact on the final design. In many cases, you can classify microprocessor chips according to their design guide lines. On the basis of the fundamental characteristics a processor can be divided into the following categories:

- CISC
- RISC
- EPIC

CISC PROCESSORS

Complex Instruction Set Computing - this class of microprocessors have characteristics that fit most early microprocessor designs. As such their characteristics consisted of practical design choices for the time that they were invented (around the 1970's). For some applications, the rate at which data can be fetched from memory is a limiting factor to the performance of the application. These types of applications are still well suited to the design guide lines of CISC processors. Two common design choices that reduce High Memory Overhead are:

Variable Length Instructions

Variable length instructions were used in CISC processors to limit the amount of memory each instruction took up. It also made it so that any instruction length would fit in one cycle so it wouldn't have to make two cycles to get one piece of data.

Extensive Instruction Set

Frequently CISC supports SS (Storage to Storage) instructions. A SS instruction is the act of allowing data to be moved from one place in memory to another in one instruction. The good thing about SS instructions is CISC processors only have to use one instruction to perform what would take a RISC processor two instructions to do; one to load the data from memory into a register and another to store it in memory again at the new destination..

Another common attribute to CISC processors is support for some programming sequences that:

- Hinder a processor's performance
- Occur rarely
- Could have been initially defined as producing unpredictable results

Frequently this support occur because we didn't know the performance implications the sequence would have on future design choices. One of the significant examples of this is support for Self Modifying Code. Self modifying code was a great idea, back in the days when a fair amount of assembler code was written. Self modifying code is a process in which an upcoming instruction is modified before it is

reached. Basically, only the current instruction was being executed, so all future instructions had not been reached yet.

RISC PROCESSORS

Reduced Instruction Set Computing - RISC architecture is newer than the CISC architecture. After making CISC chips for some time, it was found out that some CISC design guide lines were not such a good idea. With these in mind, the RISC chip was created. Also memory has become cheaper by now and that had an significant impact on the chip's design. RISC concepts consists of:

Fixed Length Instructions

This makes it so that all the instructions are the same length. This can add speed in some cases because it is easier to predict what is coming. Although this is faster most of the time, it also makes it so that if you have an instruction that needs more information than the fixed length can hold, then you have to use two instructions instead of one.

Minimized Instruction Set

This makes the processor easier to design and predict where technology will be. The less circuitry on the die makes it so that there isn't much for an instruction to screw up on. The die is also smaller which requires less time to communicate from point to point. This allows the chip to run faster.

No Storage Instructions

This concept makes it so that the processor can not store the contents of one memory location to another using the same instruction. The source and destination can not both be stored on the same memory location.

No Automatic Detection of Self Modifying Code

This makes it so that the chip does not have to detect this rare condition. Instead it's up to the programmer/compiler to insure that this condition will never occur.

Aligned Operatos

An architecture can impose a programming restriction on the placement of operands. Typically, operands will be restricted to starting addresses that are evenly divisible by the size of the object. This can assure that any part of the operand never crosses a boundary of the size of the object. The main advantage of Aligned Operands is that each fundamental data type will be fully contained in a single line of memory. With this restriction the chip does not have

to concern itself with obtaining an operand from portions of two different lines. Although this restricts what operations the chip can perform, it eliminates the need for circuitry that impedes the overall performance of the chip. For example, without this restriction, all unaligned operands will have to spend two cycles in B stage of the pipe line: one to handle each line that the data is within.

EPIC PROCESSOR

Explicitly Parallel Instruction Computing - EPIC microprocessor architecture is a brand new classification. EPIC chips will start being produced in the year 2000, permitting new levels of processing speed. One of the special things about EPIC chips is the CPU not having to check for parallelism, or in other words, check which instructions it can execute in parallel. Instead the compiler is responsible for determining which instructions can be executed in parallel and encoding that information within the machine code. Also, if the CPU comes to a branch, it will use a technique called *predication*. In this process, the CPU explores all possible outcomes and when the final result comes, it takes the result of the actual path and throws away the others.

Bundled Instructions

Bundled Instructions are a very important design guide line to the EPIC architecture. New EPIC architecture packs multiple instructions into a very long instruction word for faster processing. This is called LIW encoding, even though Intel doesn't care for the name much. Each bundle also contains a flag that is set by the compiler to indicate which instructions can be executed in parallel. The CPU can compile these bundles in any order it chooses, and this causes the processor to stall less often. Unlike CISC and RISC processors, now the CPU doesn't have to scramble around, wasting time thinking about which instructions it can run in parallel.

Supports Predication

Predication is a new way EPIC processors handle branches. In past CISC and RISC chips, there were many times the CPU guessed the wrong branch, which causes the processor to stall. Now, when EPIC CPU's encounters a branch situation, it will begin executing the code along all paths of the branch, exploiting as much parallelism as possible. Frequently the right path to take on a branch depends on the result from instructions still in the pipeline. When this condition is encountered by traditional CISC and RISC based processors they will choose one of the paths of the branch to execute and hope that it is the one that is to be taken.

Speculative Loads

Since computer chips have been created, processors speeds have gotten much much faster. Processors have gotten so fast that many cycles are wasted while the CPU is waiting for data to come from

memory. Speculative loading is another key feature to the EPIC chip, and does a good job at keeping the number of cycles spent waiting for data to come from memory down to a minimum. Speculative loading allows the CPU to bring a load instruction up further in the instruction stream (even before a branch sometimes) and load data before it is needed. It then inserts a speculative load check in the spot where the data was originally supposed to be fetched. The idea is to keep the data fetching and the use of the data separate. When a speculative load instruction is placed in the instruction stream, the CPU doesn't have to sit around, doing nothing for lots of cycles waiting for the data to come from memory.

VLIW - Very Long Instruction Word

VLIW is another EPIC design guide line that has to do with bundled instructions. When VLIW was used in the past, it didn't work out the way they planned. EPIC is using VLIW, which is a process in which instructions are put into and compiled as bundles. Also, the CPU can process bundles at a faster rate than each instruction could be processed alone because each bundle can be processed in any order the CPU wishes. With EPIC, flags are set by the compiler to signal which bundles can be executed in parallel.

TRENDS OF MICROPROCESSORS

CPUs in personal computers have moved from 4-bit to 64-bit and show every sign of not slowing down. The amount of memory each can address has also gone up with each generation, from 1 megabyte for early micros to four gigabytes (remember, that's billions), and growing. You can expect new CPUs to continue to be smaller and faster. Also, as chip-making technology advances you can expect to see other parts of the computer integrated onto a single chip with the CPU.

Since the early 1970s, the increase in capacity of microprocessors has been known to generally follow Moore's Law, which suggests that the complexity of an integrated circuit, with respect to minimum component cost, doubles every two years. In the late 1990s, and in the high-performance microprocessor segment, heat generation (TDP), due to switching losses, static current leakage, and other factors, emerged as a leading developmental constraint.

Moore's law

Moore's law describes a long-term trend in the history of computing hardware. Since the invention of the integrated circuit in 1958, the number of transistors that can be placed inexpensively on an integrated circuit has increased exponentially, doubling approximately every two years. The trend was first observed by Intel co-founder Gordon E. Moore in a 1965 paper. It has continued for almost half a century and in 2005 was not expected to stop for another decade at least.

Almost every measure of the capabilities of digital electronic devices is strongly linked to Moore's law: processing speed, memory capacity, sensors and even the number and size of pixels in digital cameras. All of these are improving at (roughly) exponential rates as well. This has dramatically increased the usefulness of digital electronics in nearly every segment of the world economy. Moore's law describes this driving force of technological and social change in the late 20th and early 21st centuries.

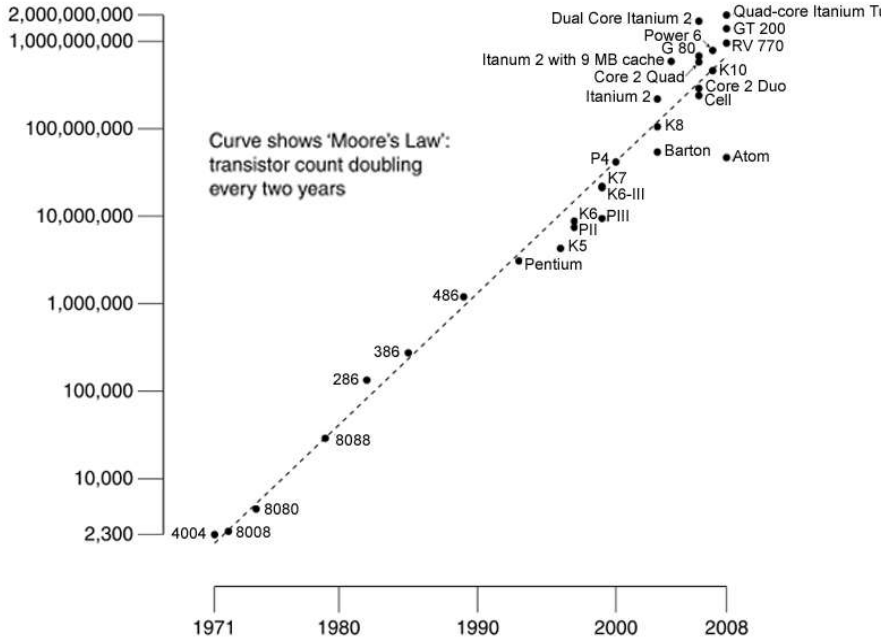


Fig 3.5 Moore's Law

There have only been five common brands of CPU in the market in recent times: Intel, AMD, Cyrix, Centaur and IBM. For microcomputers, the most common CPUs are built by Intel. Early models were the 8088 and 8086, an intermediate model is the 80286, and current models are the 80386 and i486 with the Pentium I, II, III and Pentium IV. New generations typically are introduced about every 12 months or so. (The most popular alternatives to Intel CPUs are built by Motorola and are the 68000, 68020, 68030, and 68040; used in Apple Macintosh and other computers.) The PowerPC is another generation CPU that brought even more computing power to the desktop. Following is the History and trends of CPU since its birth with Intel's 4004 chip, which may probably not completely from slowest-to-fastest ordering :

FIRST GENERATION MICROPROCESSORS

Intel 4004

Generally regarded as the very first single-chip microprocessor. It was designed almost by accident: Intel (Started by Robert Noyce and Gordon Moore, who left the Semiconductor company Fairchild in late sixties) were contracted by Busicom calculator, Japan - to create 12 custom chip for a desktop calculator - but Intel decided they could do the job more efficiently with one flexible, multipurpose chip, which could be reused again and again in different applications.



Fig 3.6 Intel 4004 CPU

The idea of using one chip was brought forward by Intel engineer Ted Hoff. Both Noyce and Moore embraced Hoff's solution and put forward the proposal with Busicom - which was accepted. Busicom funded the Project while keeping the right of the chip. Working on the concept for almost one year, the Intel brought the first tiny 0.74MHz (External Clock also 0.74 MHz) 4-bit 4004 chip containing 2250 transistors. Intel purchased the rights from Busicom at \$60,000. It had a short and unspectacular market life, and very little power even by the standards of the day, but it was the forerunner of the more successful 8-bit 8008 and then the first really successful CPU, the Intel 8080.

Intel 8088

The classic Intel CPU of 2 MHz containing 2500 transistors. This was the heart of most of the world's first microcomputers, including the Altair 8800. All modern PCs use descendants of the 8080. Indeed, an Athlon or a Pentium 4 can still run something very like 8080 code.

Despite its huge influence and massive sales, the 8080 was soon overshadowed in the marketplace by the more advanced Zilog Z-80, which built on the 8080's success and owed a lot to it. Introduced at 2MHz, the 8080 was pushed up to just over 3MHz as time went by.

MOS Technologies 6502

A development of the Motorola 6800 by ex-Motorola engineers, the 6502 and its numerous variants powered many of the high-profile early micros, including the Commodore Pet, Apple II, and Commodore 64, not to mention the hordes of marginally compatible Apple clones.

Zilog Z-80

The most successful microprocessor of all time. Zilog was a group of ex-Intel engineers who set out to improve on the 8080 but still maintain compatibility with it. The Z-80 sold in huge quantities, and was at the heart of most of the microcomputers of the CPM era. Apple and Commodore used the 6502, but nearly all the other successful manufacturers used a Z-80. Programmers loved it for its power and simplicity and Zilog pushed it to faster and faster clock speeds over the years. It was more than a match for the clumsy first-generation 16-bit CPUs like the Intel 8086 and can still do useful work. Incredibly, the 25-year-old 8-bit Z-80 is still in production!



Fig 3.7 Zilog Z-80 Chip

Intel 8085

The 8085, Intel's replacement for the 8080, had considerable success, but always played second-fiddle to the Z-80. Much improved over the 8080, it powered a fair number of industry-standard CPM machines, especially in the United States.

It was handicapped by an inability to run Z-80 extensions, and by Intel's haphazard programming tools—8080/8085 assembler mnemonics were very poorly thought out. Most programmers much preferred the dauntingly complex but very logical Z-80.

SECOND GENERATION MICROPROCESSORS

Intel 8086 and 8088

The chip that powered the original IBM PC, and thousands of other models too. All later X86 designs (286, 386, 486, Pentium, and so on) have built on this foundation.

There were four main 16-bit CPUs at this time. The Texas Instruments TI 9900 was an early leader but badly marketed. The Zilog Z-800 was powerful but late and buggy. The expensive Motorola 68000 was generally regarded as the best choice with its combination of simplicity and power. But the 8086 family ended up as easily the most successful of them all.

The main reasons were its compatibility with the 8080/8085 and Z-80 family, its relatively low price, and above all, the 8088 variant. The 8088 was a hybrid 8/16-bit design: 16-bit internally, with 8-bit I/O.

Motorola 68000 Family

The 68000 powered almost everything that didn't have an 8086 or 8088: Apple Lisa, Apple Mac, Commodore Amiga, Atari ST, and many others. The simple, practical design lasted well, and was updated with the 68020, 68030 and 68040. Programmers loved the 68000 (no-one ever loved an 8088) and 68000 code tends to be small, fast and efficient.

There have been numerous variations and developments within the 68000 family- much like the X86 family with its 8086, 286, 386 and so on. The last of them were roughly equivalent to the Pentium. By that time, Apple had switched to the Power PC, and most of the other 68000-based manufacturers had disappeared.



Fig 3.8 The 286-16 Chip



Fig 3.9 The 286-20 Chip

80286

The original 286 (more properly, 80286) was a vast power increase over the 8086/8088, particularly the 8/16-bit hybrid 8088. The 8MHz ones were very rare, as by the time most people could afford to buy 286s they were buying 12 and 16MHz ones.

The 286-16 were almost universal: the computer 286-16, 1MB RAM, a 256k VGA card and a 40MB IDE stepper hard drive from Seagate or Western Digital was an instant hit.

THIRD GENERATION MICROPROCESSORS

80386SX & 386DX

The original 386 was a full-blown 32-bit monster in its day, and very expensive. The 16-bit/32-bit hybrid 386SX came along three years later. Just like the 8/16-bit hybrid 8088 a few years earlier, the 386SX was designed to provide a way of using existing 16-bit motherboard components. The SX was much cheaper than the DX and more easily adapted to existing board technology.

The Intel 386 was easily the most significant X86 CPU of them all. The 386 was a huge advance over the 8086 and the '286. In fact, it was designed before the 286, but was too difficult to make and use at first.

The 386 had vastly better memory management than the 286, and built-in multi-tasking features to allow the development of powerful operating systems like OS/2 and Windows NT. The SX-33s and DX-40s that followed soon after were vastly faster.

Although it was substantially dearer than a 286, it was just as crippled by its slow 16-bit memory access. It did offer the benefit of the 386 instruction set - which was vastly better than that of the brain-dead 286 - but it didn't offer any particular performance advantage.



Fig 3.10 The 386-33 Chip

80486

A 486 chip was a fast 32-bit 386DX CPU, and a 387 NPU, and 8k of cache RAM, all on a single chip. By putting the cache and the NPU on the CPU chip, Intel were able to greatly speed communication between the three. Also, the 486 introduced mainframe techniques like pipelining into single chip microprocessors for the first time. With all these different components to fit onto one chip, it is no wonder that the 486 was physically much bigger than a 386.

Pipelining

Pipelining is a concept of performing a task after dividing it into various heads. To understand pipelining, imagine putting out a fire with buckets of water. If you are alone, you have to walk to the well, fill the bucket, walk to the fire, empty the bucket, walk back to the well, and so on.

This is how a 386 works, going to the RAM, getting the next instruction, decoding it, running it, going back to the RAM for the next one.

Now imagine have four or five people to help put the fire out. One fills the buckets, one empties them onto the fire, the others pass the buckets along. You are still only pouring one bucket at a time, but it's much faster. CPU pipelining works the same way. Instructions are loaded one after another into the RAM end of the pipeline and the CPU just takes them from the other end as needed. All modern CPUs are heavily pipelined.



Fig 3.11 The 486DX-33 Chip

There were many variants of 486 that were launched in the market. These included:

IBM 486SLC/2-50

A very odd little CPU which was essentially a clock-doubled 386SX.

Cyrix 486DLC-40

The DLC was Cyrix's first serious entry into the mainstream X86 market.

AMD 486DX-40

The 486DX-40 was just a short-term stepping-stone to the mighty 486DX/2-66.

Intel 486 DX/2-66

The classic 486. Intel were first to market with it and sold the largest number. AMD/Cyrix and IBM followed soon.

IBM 486BL/4-75

The very last IBM-designed X86 CPU, and the end of an era. Once again, it was based on licensed Intel 386 technology, and this time it was pushed out to the max. Unlike the SLC, the Blue Lightning was fully 32-bit and a serious performer.

T I 486DX/2-80

The processor chip was a good performer but it was that 40MHz system bus of the motherboard and I/O cards, VESA video cards in particular. There were quite a few IBM and Cyrix DX-2/80s on the tail end of the market after AMD's DX/4 arrived.

Intel 486DX/4-100

Another one of the great 486s. There were 1.6 million transistor and 5 models were launched - 2 from Intel and AMD each and one from Cyrix/IBM. The AMD and Cyrix/IBM DX/4s were all 3 Volt parts. Most of the Intel ones were 3 Volt chips with a built-in 5 Volt regulator, and could be used to upgrade any 486-33 or 66.

FOURTH GENERATION MICROPROCESSORS

Intel Pentium

Like the 486, the Pentium was a quantum leap in hardware, and it took several more techniques out of the mainframe and into the desktop. It was the first super-scalar X86 CPU, capable of doing more than one instruction per clock-tick. To do this, the Pentium had two separate integer maths units that could both run at the same time, as well as a much faster NPU and a 16k internal cache. It required all of Intel's formidable manufacturing expertise to produce it. Three times the size of a 486, the Pentium had 3.1 million transistors, and early examples consumed no less than 16 Watts and ran very hot. Such a big chip was expensive to produce too, and Pentium pricing was very, very high for a long time.



Fig 3.12 The Pentium 60 CPU

The Pentium was faster than any of its competitors in the early days, and a major achievement. It became the bestselling CPU in the world and, in various flavours, remained so for four years - a very long time in this industry. The Pentium NPU was nearly four times faster than the 486's NPU, and this let Intel start a push into the CAD workstation market which up until then had been dominated by RISC chips.

Following are the various variants of Pentium CPU of Intel :

- | | |
|------------|--|
| Pentium 60 | March 1993, 3.1 million transistor, Socket 4: 273-pin PGA. Intel never actually intended to manufacture a P-60. They got very hot, they had the famous floating-point bug, and motherboard manufacturers hadn't really got the hang of the new chip. |
| Pentium 66 | March 1993, 3.1 million transistor, Socket 4: 273-pin PGA, another other slow, hot, unreliable early Pentium. |
| Pentium-75 | September 1993, 3.3 million transistor. The 5 Volt Pentium 60s and Pentium 66s were big, hot, sluggish and fragile: these were cool, fairly quick, and very reliable. |

Pentium 90	March 1994, 3.1 million transistor. A good performer, and the first 'real' Pentium - in the sense that the original 60 and 66MHz versions were absurdly expensive and not particularly good performers.
Pentium 100	March 1994, 3.1 million transistor. The Pentium is one of the all-time classic CPUs, and the 100 was one of the all-time classic Pentiums. It ran a 66MHz main board, so the P-100 ran its RAM faster and had better input/output performance than the 60MHz-based Pentiums 90, 120 and 150.
Pentium 120	March 1995, 3.3 million transistors. It was little faster than the Pentium 100, and well behind the 133MHz version that came out just three months later.
Pentium 133	June 1995, 3.3 million transistors. Like the 100 and the 166 it ran a 66MHz bus to give it good I/O performance, and like those two it was the fastest X86 of all for quite some time till the 6x86 came out in Feb 1996.
Pentium 150	January 1996, 3.3 million transistors. It ticked faster than a P-133, but the bus was slower. In reality you were better off with any of the 66MHz bus chips: Pentium 133, Pentium 166, or 6x86-166.
Pentium 166	January 1996, 3.3 million transistors. The power-user's chip of choice for a long, long time, and still a very good performer indeed right up to the close of the century. Generally regarded as the classic Pentium.



Fig 3.13 The Pentium 166 CPU

Pentium Pro 200	November 1995, 5.5 million transistors. The Pro was a physically massive chip, or, more accurately, two chips bonded together - the CPU itself and a 256k or 512k secondary cache chip too. The advantage of this is that the processor could access the cache RAM at full clock speed. Unlike all previous Intel X86 CPUs, the Pro (and the very
-----------------	---

similar Pentium II which was developed from it) had a **RISC** core. The major strengths of the Pentium Pro were its very fast maths co-processor, the huge on-chip secondary cache, and the lack of a viable alternative for large server installations.

Pentium-200

June 1996, 3.3 million transistors. Although the Pentium 200s were much more expensive than the 166, they were only marginally faster because the clock-tripled 66MHz bus held them back. Other chips in this performance class did better: the 6x86-200 had a 75MHz bus, while the subsequent parts - Pentium MMX, K6 and 6x86MX - all had enlarged on-chip caches and better buffers to help them cope more gracefully with accessing RAM over a by-now barely adequate 66MHz bus.

Pentium MMX-166

January 1997, 4.4 million transistors. Although based on the Pentium Classic, the MMX had a smaller die to make it run cooler (despite having one-third more transistors), a host of detail changes and - most significantly - the primary on-chip cache was doubled to 32k. The larger cache in particular, made the MMX-166 about 10% faster than a 166 Classic, 5% faster than a Classic 200, or 1% faster than a 6x86 200 Classic.

Intel Pentium MMX-200

January 1997, 4.4 million transistors. The 166 was easily favourite Pentium MMX, but the 200 was a good one too, once Intel finally got real about the price-performance.



Fig 3.14 Pentium 200 CPU



Fig 3.15 Pentium MMX 200 CPU

Intel Pentium MMX-233

January 1997, 7.5 million transistor. Like all the higher-speed Pentiums, these tended to be I/O bound: the input and output from the CPU to the rest of the system was a bottleneck. It was the last

of the long-lived and enormously successful Pentium line, though not a particularly grand finale.

FIFTH GENERATION MICROPROCESSORS

The Pentium II refers to Intel's x86-compatible microprocessors introduced on May 7, 1997. Containing 7.5 million transistors, the Pentium II featured an improved version of the first P6-generation core of the Pentium Pro CPUs, which contained 5.5 million transistors.

Pentium II-233 January 1997, 7.5 million transistor. The Pentium-II was Intel's second effort with the P6 architecture. The P-II was a Pentium Pro core in a different shaped package, and with marginally less effective but much more cost efficient cache arrangements.

Pentium II-266 March 1997, 7.5 million transistor. The Pentium-II 266 will long be remembered as one of the great CPUs, but also one of the rarer ones.

Intel Celeron 266 April 1997, 7.5 million transistor. Quite possibly the worst-performing CPU ever made.

This was the "Pentium II SX", an extraordinary product. It came in the form of a massive and not inexpensive SEC cartridge, used a substantially more expensive motherboard, and had very poor performance by the standards of the day. Intel had felt forced to introduce it because the Pentium MMX was fading fast by then and the Pentium II was much too expensive for the mid-range.



Fig 3.16 The Pentium Celeron 266 CPU

Pentium II 300 September 1997, 7.5 million transistors. It offered a mere two percent better system performance than the 266, or six percent better than the K6-233.

Pentium II 333 January 1998, 7.5 million transistor. It was incredible speed at which Intel's production engineers were launching chips. P-II 333 is not

- underwhelming, it is just plain sluggish. With equal RAM and a much slower hard drive, K6-2/300 was easily superior to the 333.
- Pentium II-350 April 1998, 7.5 million transistor. The P-II 350 disappeared around the winter of 1999: it was a pointless product, as Intel's own Celeron 366 outperformed it.
- Pentium II 400 April 1998, 7.5 million transistor. Easily the fastest X86 chips on the planet when they were introduced, the 400 and 450MHz Pentium IIs were a showcase for Intel's production genius, and they remained faster than any non-Intel X86 for a long, long time: right up until early 1999 when the K6-III arrives and the Pentium III was just around the corner.
- Celeron 300 April 1998, 7.5 million transistor. Another slug. It would still be well remembered for its unacceptably low performance levels had it not been for the existence of the even slower 266. Like the 266, the Celeron 300 was an otherwise standard Pentium II with no secondary cache at all.
- Pentium II 450 and Pentium III 450 August 1998, 7.5 million transistor. The P-II 450 had one of the longer reigns as the world's fastest X86. Until the arrival of the K6-III/400 and then the Pentium-III/500 in March 1999, only an Alpha was faster.
- Celeron 300A August 1998, 7.5 million (plus 11.5m cache) transistors. On-chip secondary cache made this a vastly better performer than the woeful original Celeron, and a serious competitor to the fast-selling AMD K6-2.
- Celeron A 333 August 1998, 7.5 million (plus 11.5m cache) transistors. It was a serious performer with 128k internal cache.
- Celeron 366 February 1999, 7.5 million (plus 11.5m cache) transistors. It marked Intel's return to socket mounting instead of the huge and clumsy Slot 1 arrangement they used for the Pentium II/II and the earlier Celerons. It also gave them a chance to compete on more or less equal terms with AMD.
- Celeron 400 February 1999, 19 million transistors. Performance was very similar to the Pentium II 400. What the Celeron lost in cache size, it made up in cache access speed.
- Pentium III 500 May 1999, 9.5 million transistor. Intel's answer to AMD's 3DNow extensions was SSE; another variation on the single-instruction, multiple-data theme introduced by MMX.

Celeron 433 & 466

May 1999, 19 million transistors. With the release of the 433 and 466MHz Celerons, Intel reclaimed their lost market leadership from the K6-2 on the one hand, and improved the otherwise crippling sales.

Celeron 500

August 1999, 19 million transistors. Another goodie from Intel, though less common than the 466 Mhz because for most of its market life.

SIXTH GENERATION MICROPROCESSORS

Rather than just being a faster Pentium, the Pentium Pro, Pentium II/III, sixth-generation processors have many feature and architectural improvements. The core of the chip is very RISC-like, whereas the external instruction interface is classic Intel CISC. By breaking down the CISC instructions into several RISC instructions and running them down parallel execution pipelines, the overall performance is increased.

Pentium III 550 “Coppermine”

August 1999, 22 million transistors. While the original Pentium-III was a non-event, this second version of it was vastly improved. It was sufficiently different to belatedly justify the new name, and a worthy competitor to the Athlon. Like the Celeron and the K6-III, Coppermine had full-speed on-die cache of 256k.



Fig 3.17 The Pentium III 550 “Coppermine”

Pentium III 600 “Coppermine”

March 2000, 28.1 million transistors. Coppermine didn't just introduce better caching, several of the new Pentium IIIs ran on a 133MHz bus for faster communication between the CPU, the mainboard

and the RAM. Over time, the entire Pentium III family migrated to 133MHz bus.

Pentium III 650 & 666

May 2000, 28.1 million transistors. The Pentium III 650, with its 100MHz bus, was an underperformer. The 133MHz bus 666, on the other hand, was an excellent part once it became available after the by now customary six month delay.

Pentium III 700 & 733

June 2000, 28.1 million transistors. The 733 was significant in that it marked a brief return by Intel to their once-for granted status as the manufacturer of the fastest-clocked X86 in the world. As to whether it was actually faster than an Athlon 700.

Pentium III 800 & 850, 866

August-October 2000, 28.1 million transistors. There was an incredible variety of announced P-III parts. The only one of these three to sell well and deserve a mention was the 866.

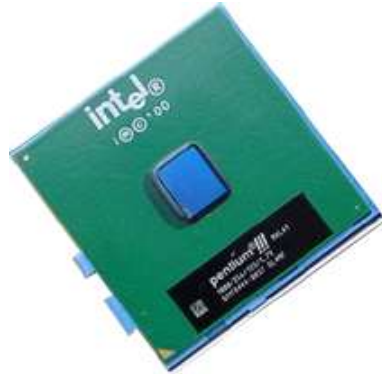
Celeron 566 and 600

March/August 2000, 19 million transistors. With these, the Celeron finally gained some multimedia instructions. Also, and perhaps more significant to most users, it got most of the advanced cache design from the "Coppermine" Pentium-IIIs—though only 128k, not the full 256k.

Celeron 633, 666, 700, 733 and 766

June-November 2000, 19 million transistors. The forgotten Celerons, and justly so. These grew up in the shadow of the AMD Duron, still crippled by their stone-age 66MHz bus speed, and utterly unable to compete.

Pentium III 1000 February 2000, 28.1 million transistors. The 1000MHz Pentium III was announced a bare two days after the Athlon 1000 and promised to be a faster, better part, but in reality 1000MHz P-IIIs didn't become readily available to the market at large until February 2001 - almost a full year from the original announcement. But despite arriving far to late to win any speed crowns, in a modest sort of way the Pentium III 1000 became one of favourite CPUs.

**Fig 3.18 Intel Celeron 566****Fig 3.19 Pentium III 1000**

- Pentium III 1133** An extraordinary debacle. In an attempt to maintain some credibility at the top of the performance tree, Intel announced the 1130Mhz Pentium III in July 2000. Several of the reviewers who were lucky enough to get early production samples sent to them had difficulties. Humiliatingly, Intel were obliged to withdraw the P-III 1133 from sale and issue refunds to those who had bought one.
- Pentium 4 1400** November 2000, 42 million transistors. The Pentium 4 did not improve on the old P6 design in either of the normal two key performance measures: integer processing speed or floating-point performance. Like most of the rest of the world, we were dumbfounded by this. At 1.5GHz, the Pentium-4 was not only inferior to the Athlons, but couldn't beat the old Pentium-III 1000! On a clock-for-clock basis, it was still worse than this.
- Pentium 4 1500** November 2000, 28.1 million transistors. To reach its very high design clockspeeds, the Pentium 4 has a very long 20-stage pipeline. For comparison, the P6 pipeline has twelve stages, the Athlon ten. The advantage of the longer pipeline is that it makes it possible to clock the chip faster, the disadvantage is that it makes branch mispredictions more of a problem than ever. When the processor correctly guesses what it is going to have to do next and pre-fetches the right instructions, it cooks right along; when it mispredicts, the penalty is severe. The 1500 was one of the two original Pentium 4s (along with the 1400). The other variants were 1600, 1700 and 1900 MHz.

Pentium 4 2000Mhz

November 2000, 42 million transistors. The Pentium 4 arrived. The 2000 was the first Intel chip to outpace AMD's best since the Pentium III 550's brief reign two and a half years before, and provided some justification for Intel's brave and very unorthodox design decision to sacrifice all else in the pursuit of high clock rates.

Intel Xeon Processor

The Intel Xeon processor is targeted for high-performance and mid-range, dual-processor workstations, dual and multi-processor server configurations coming in the future. The platform offers customers a choice of operating systems and applications, along with high performance at affordable prices. The processor is based on the Intel NetBurst architecture, which is designed to deliver the processing power needed for video and audio applications, advanced Internet technologies, and complex 3-D graphics.

SEVENTH GENERATION MICROPROCESSORS

Itanium™ Processor

The Itanium processor is the first in a family of 64-bit products from Intel. Designed for high-end, enterprise-class servers and workstations, the processor was built from the ground up with an entirely new architecture based on Intel's Explicitly Parallel Instruction Computing (EPIC) design technology. The processor delivers world-class performance for the most demanding enterprise and high-performance computing applications, including e-Commerce security transactions, large databases, mechanical computer-aided engineering, and sophisticated scientific and engineering computing.

Pentium M

Released in March, 2003, this processor was based on Pentium III core, with SSE2 SIMD instructions and deeper pipeline having 77 million transistors. It had variants from 900 MHz to 2.26 GHz. It was considered to be the Heart of the Intel mobile "Centrino" system. The Pentium M-branded processors had a maximum power consumption (TDP) of 3–25 W, and were intended for use in laptop personal computers (thus the "M" suffix standing for mobile). Pentium M had an improved instruction decoding and issuing front end, improved branch prediction, SSE2 support, and a much larger cache. The first Pentium M-branded CPU, code-named Banias, was followed by Dothan.

Pentium D

Released in May, 2005, Pentium D had Dual-core microprocessor with No Hyper-Threading and 800(4x200) MHz front side bus. Each CPU comprised two dies, each containing a single core residing next to each other on a multi-chip module package. The brand's first

processor, codenamed Smithfield, was followed by its successor nine months later, codenamed Presler. By 2005, the NetBurst processors reached a clock speed barrier at 3.8 GHz due to a thermal (and power) limit exemplified by the Presler's 130 W Thermal Design Power (a higher TDP requires additional cooling that can be prohibitively noisy or expensive). The final shipment date of the dual die Presler chips was August 8, 2008, which marked the end of the Pentium D brand and also the NetBurst microarchitecture.



Fig 3.20 Pentium M Chip



Fig. 3.21 Pentium D Chip

EIGHTH GENERATION MICROPROCESSORS

Intel Centrino Core Duo (Intel CD)

Released in January, 2006, Centrino took over Pentium brand which ruled the industry for almost 13 years. Centrino as a brand is not a mobile CPU - rather, the term covers a particular combination of mainboard chipset, mobile CPU and wireless network interface in the design of a laptop.

The Core brand was launched on January 5, 2006 by the release of the 32-bit Yonah CPU - Intel's first dual-core mobile (low-power) processor. Its dual-core layout closely resembled two interconnected Pentium M branded CPUs packaged as a single die (piece) silicon chip (IC). Hence, the 32-bit microarchitecture of Core branded CPUs - contrary to its name - had more in common with Pentium M branded CPUs than with the subsequent 64-bit Core microarchitecture of Core 2 branded CPUs.

Intel Core Duo (product code 80539) consists of two cores on one die, a 2 MB L2 cache shared by both cores, and an arbiter bus that controls both L2 cache and FSB access. Upcoming steppings of Core Duo processors will also include the ability to disable one core to conserve power.

Intel Centrino Core 2 Duo (Intel C2D)

Released in July, 2006, the Core 2 brand refers to a range of Intel's consumer 64-bit x86-64 single-, dual-, and quad-core CPUs based on the Intel Core microarchitecture. The single- and dual-core models

are single-die, whereas the quad-core models comprise two dies, each containing two cores, packaged in a multi-chip module. The introduction of Core 2 relegated the Pentium brand to the mid-range market, and reunified laptop and desktop CPU lines, which previously had been divided into the Pentium 4, Pentium D, and Pentium M brands.

The Core microarchitecture returned to lower clock rates and improved the usage of both available clock cycles and power when compared with the preceding NetBurst microarchitecture of the Pentium 4/D-branded CPUs.

Duo, Quad, and Extreme

The Core 2-branded CPUs include: "Conroe"/"Allendale" (dual-core for desktops), "Merom" (dual-core for laptops), "Merom-L" (single-core for laptops), "Kentsfield" (quad-core for desktops), and the updated variants named "Wolfdale" (dual-core for desktops), "Penryn" (dual-core for laptops), and "Yorkfield" (quad-core for desktops). (Note: For the server and workstation "Woodcrest", "Clovertown", "Tigerton", "Harpertown" and "Dunnington" CPUs.

Corei7

Officially launched in November 2008, the successor to the desktop version Core 2 brand is the Core i7, which is based on the Nehalem microarchitecture. Core i7 was officially launched on 17 November 2008 as a family of three quad-core processor desktop models. Nehalem has major changes from the Core microarchitecture. The FSB is replaced by a QuickPath interface, and the processor has an onboard memory controller. The Core i7 is used mainly for extreme gaming purposes whereas the Core 2 desktop series are used mainly for mainstream computing. The Core i7 is more of a high-end alternative to the desktop Core 2 series.



Fig 3.22 Corei7 Chip

The Core i7 identifier applies to the initial family of processors code named Bloomfield. Intel state that the Core i7 is meant to help

consumers decide which processor to purchase as the newer Nehalem-based products are released in the future. The name continues the use of the Core brand.

AMD AND OTHER CHIPS

- Cyrix 5x86-100 August 1995, 2 million transistors. A great little chip which only had about six months on the market before it got pushed off the production line to make room for the **6x86**. It started with a derivative of the 64-bit M1 core but trimmed it down to fit on a 486-type motherboard, in more or less the same way that the 386SX was a 32-bit chip with a 16-bit interface.
- AMD 5x86-133 November 1995, 1.6 million transistors. The AMD 5x86 was really a hotted-up 486 - a “DX/5-133” if you like, and much less advanced than the Cyrix one. The 5x86-133 was the last and the best of all the 486s, just as the DX-40 was the last and greatest of the 386s.
- Cyrix 5x86-120 October 1995, 1.4 million transistors. Like the AMD 486DX/2-80, these suffered from running a 40MHz bus. They were the fastest of the 5x86s and roughly equal to a Pentium-90, but at 40MHz bus speed they were prone to motherboard problems.



Fig 3.20 Cyrix 5x86

AMD K5 90

4.3 million transistors. In March 1996, the long-awaited fifth-generation AMD K5 arrived—but at 75 and 90MHz the K5 was a long, long way behind their competitor’s Pentium, 6x86 and Nx586 parts.

NexGen Nx586-133

The Nx586 was a very bold, innovative design and pioneered the use of a RISC core to execute X86 instructions. The Nx586 had two main disadvantages, the lack of an in built FPU, and its need for a special motherboard. The motherboard requirement was because the Nx586 had a



Fig 3.21 AMD 5x86

dedicated high-speed cache memory bus - an idea Intel have since adopted for the Pentium II/III family.

AMD K5-100

June 1996, 4.3 million transistors. The AMD K5 and the Cyrix M1 (or 6x86) marked a huge change in the CPU industry. The K5 and the M1 were both intended to out-Intel, to forget duplicating existing designs like the 386, the 486 and the Pentium, and lead the way with bold new engineering.



Fig 3.22 AMD K5-100

Cyrix 6x86-120

December 1995, 3.3 million transistors. The M1 and the AMD K5 weren't mere Pentium equivalents, they introduced whole new high-speed RISC design features which were 'impossible' given the arcane and over-complex X86 instruction set.

AMD K5-120

January 1996, 4.3 million transistors. When it eventually arrived, it was a bit cheaper than the Pentium-120, near enough to the same speed as made no difference, and equally eclipsed by the big-selling Cyrix 6x86 PR-150.

AMD K5-133

October 1996, 4.3 million transistors. The K5-133 arrived just as the mainstream was shifting to the 166MHz class (be it 6x86 or Pentium), and there was still just enough life in the 133MHz category to let AMD notch some sales up at long last. It was far from being a best-seller.

Cyrix 6x86-133

February 1996, 4.3 million transistors. Because the 6x86 did substantially more work per clock tick than a Pentium or first-generation K5, and a 6x86 at 120MHz (like the 150) was a little bit faster than a Pentium at 150MHz.

Cyrix 6x86-150

February 1996, 4.3 million transistors. The 6x86-150 remained very popular here right up until production ceased in June 1997, at which time it was still a major CPU being sold those days. IBM and Cyrix pushed it off the production line before its time was really up because their ever-increasing

yields meant that most of their parts could be sold at PR-166 and PR-200.

AMD K6-166 March 1997, 8.8 million transistors. The 166 was the least common of the K6 Classics and only ran as a mainstream product for about six months; it was soon overshadowed by the cheaper Cyrix 6x86 Classic 200 and 6x86MX-166, and the faster 200MHz versions of the Pentium MMX, 6x86MX and the K6 itself.

Cyrix IBM 6x86MX-166 May 1997, 6.6 million transistors. A very traditional Cyrix CPU with the traditional Cyrix virtues: generous mainstream performance at a knock-down drag-out price. It was considerably cheaper than the Pentium MMX-166 or the AMD K6-166, and noticeably faster than either.

IDT C6-200 December 1997, 5.4 million transistors. There is always a need for low-cost CPUs in non-performance critical roles, and the C6-200 filled it nicely. It was the second slowest new CPU you could buy in 1998, but it was cheap, simple to work with, and still easily faster than the Pentium-100s and 6x86-120s of the year before.

AMD K6-200 March 1997, 8.8 million transistors. Another one of the excellent K6 Classic family. The 200MHz variant was very popular for a short while before the Cyrix/IBM 6x86MX-200 and the long-delayed K6-233 hit the market in volume.



Fig 3.23(a) AMD K6-200



Fig 3.23(b) 6x86 200

Cyrix IBM 6x86MX-200

May 1997, 6.6 million transistors. The best of the first-generation 6x86MX chips, and indeed, the best 200MHz chip of them all. Almost forgotten now, the 6x86MX-200 nevertheless an all-time favourite CPUs, together with the Z-80, the evergreen 386DX-40, the 486 in both DX/2 and

DX/4 form, perhaps the 5x86-133, certainly the 6x86-200 Classic, the mighty K6-III/450, the early Durons, the Athlon C series Thunderbirds, and the K6-2/300.

AMD K6-233

March 1997, 8.8 million transistors. All things considered, the K6 was the best high-power CPU on the market for 1997 and into the first half of 1998. The K6-233 was roughly equal to the P II-233 for most tasks, cost far less, and used a standard Socket 7 motherboard.



Fig 3.24 AMD K6-233

AMD K6-266 and K6-2/266

April 1998, 9.3 million transistors. The K6-266 was our bestselling CPU for a short while in mid-winter 1998: it was about 5% faster than the Pentium II 233. The most significant thing about the K6-266 was the new 0.25 micron production process that made it possible.

AMD K6-2/300 AFR-66

April 1998, 9.3 million transistors. A cheaper, slower version of the mighty K6-2/300 that ran at 4.5 by 66MHz instead of 3.0 by 100MHz. Look for the “AFR-66” markings on the second line of text on the top of the chip to tell the difference.

IBM 6x86MX-266 April 1998, 6.6 million transistors. To run an MX-266 properly, you needed a motherboard that was stable at 83MHz - not all that common until some time after these came out - and very good quality RAM.

Cyrix M II-300

April 1998, 6.6 million transistors. Cyrix just loved changing names. The M1 became the 6x86, the M2 became the 6x86MX, and now the part that you had every right to expect would be called the 6x86MX 300 became the M-II 300.

AMD K6-300

April 1998, 8.8 million transistors. With the 266, the last of the original K6 series, and another good

one. From this time on all future K6 family releases would be K6-2s with a 100MHz Super 7 bus and 3DNow multi-media extensions.

AMD K6-2/300

April 1998, 9.3 million transistors. For 1998, this was the hot-selling CPU, and it rocketed AMD into market leadership almost overnight. The K6-2 family soon became the bestselling chips on the retail market, and brought a new world of affordable performance with them.



Fig 3.25 Cyrix M II-300

Cyrix MII 366

April 1999, 6.6 million transistors. New product from Cyrix was long overdue by the time these were announced. They were regarded as difficult and hot-running.

AMD K6-2/333

April 1998, 9.3 million transistors. Seriously fast in its day, but it never became a popular choice because it required a 95MHz bus and very few motherboards provided this early on.

AMD K6-2/400

April 1998, 9.3 million transistors. The 400, in other words, was not as big an improvement on the 350 or the 300 as the 300 was on the K6 Classic family.

Cyrix MII 466

September 1999, 6.6 million transistors. The last of the National Semiconductor/Cyrix “non-chips”. As with all the 6x86 family, these were intended to do more work per clock-tick than the AMD or Intel parts and run a lower multiplier, but these were an equivalent for business use to the Celeron 400 or 433.

AMD K6-2/450

February 1999, 9.3 million transistors. It was the fastest K6 available and ran at 2.4 Volts rather than the 2.2 Volts of the lesser K6-2 parts—higher voltages are good for faster clock speeds, but bad for heat generation and power consumption.



Fig 3.26 AMD K6-2/300

- AMD K6-2/475 April 1999, 9.3 million transistors. A marketing department part from AMD. Very few of these were shipped in real life, they were just a place-holder for the 500, which came out not much later.
- VIA Cyrix III June 2000, 11.2 million transistors. The design had nothing in common with the 6x86 family: it was the work of the IDT WinChip team. It slipped into a standard Socket 370 mainboard (the same board as a Celeron or Pentium-III) and it ran at 100MHz bus speed.



Fig 3.27 VIA Cyrix III



Fig 3.28 AMD K6-III 400

- AMD K6-2/500 August 1999, 9.3 million transistors. This was the first K6-2 not to be a performance leader: it came out at about the same time as the K6-III, which took over as AMD's fastest CPU until the Athlon was released. For much of its market life it was the same price as a K6-III/400.
- AMD K6-2/550 August 2000, 9.3 million transistors. The K6-2 design was running out of steam as multiplier increases had progressively less effect—something that happens to ever chip family in time. The K6-2/550 was as troublesome as the old and unlovely 486DX-2/80.
- AMD K6-III 400 Feb/May 1999, 21.4 million transistors. It seemed that AMD could do no wrong in '98 and '99: the K6-2 had been a great success, and this improved version of it was better yet. The K6-III had 64k of primary cache, 256k of full-speed secondary cache, and between 512k and 2MB of 100MHz tertiary cache too.
- AMD K6-2+ 500 April 1999, 15.3 million transistors. The K6-II+ ran at 2.0 Volts rather than the usual 2.4 or 2.2 Volts and, like the K6-III, had on-chip secondary cache as well as the mainboard-mounted tertiary cache. The difference was the size: 128k where the K6-III had 256k.
- AMD K6-III 450 Feb/May 1999, 21.4 million transistors. One of the all-time great CPUs. It was the long-time

unchallenged king of business CPUs, and a capable all-rounder too. It is astonishing that the K6-III was so little known: it comprehensively outperformed all other contenders in its MHz range.

AMD Athlon 500

August 1999, 22 million (plus cache) transistors. The Athlon was new from the ground up: the world's first seventh generation CPU. It featured very deep pipelining and multiple parallel units, including three pipelined floating-point units, plus 128k of on-chip primary cache - twice as much as any previous X86 and four times as much as the Pentium II and III family.



Fig 3.29 AMD K6-III 450



Fig 3.30 AMD Athlon

AMD Athlon 600

August 1999, 22 million (plus cache) transistors. The Athlon 600 was popular for a high-end part. In its first release form, the Athlon had four speed grades: 500, 550, 600 and 650. As is often the case, the second-top part was considerably less expensive and much more common than the very fastest one.

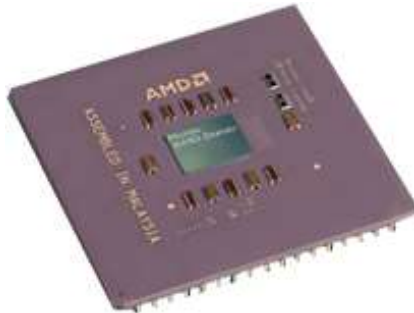


Fig 3.31 AMD Duran 600

AMD Duron 600

July 2000, 25 million transistors. The Duron was AMD's K6 replacement for the entry level, and stood to the Athlon Thunderbird as the Celeron did to the Pentium III - or almost, as the Celeron had three performance-robbing differences: smaller cache, less cache intelligence, and lower

bus speed, where the Duron had only one: the smaller cache. Otherwise it was identical to its bigger brother.

AMD Athlon 650 August 1999, 22 million transistors. The fastest Athlon of them all on first release, and the fastest X86 CPU on the planet for the three months until the 700 marked the start of AMD's relentless and spectacular push towards the 1000MHz milestone.

AMD Duron 650 July 2000, 25 million transistors. The 650 was a little faster than the 600 and it cost a little more, a little slower than the 700 and cost a little less.

AMD Athlon 700 October 1999, 22 million transistors. Perhaps the best of the Athlon Classics, and the last one to use the original 1/2 cache divider—the 750 and up ran their cache slower relative to the CPU itself and the performance gains from higher clockspeeds started to tail off.

AMD Duron 700 July 2000, 25 million transistors. The Duron 700 was right in the middle of the sweet spot in the last few months of 2000, then an attractive entry-level part up until AMD dropped it in favour of the 750 and 800 in April 2001.

AMD Athlon 750 November 1999, 22 million transistors. The 750, 800 and 850 ran with a 1/2.5 divider, the 900 and up with 1/3: the cache ran slower relative to the CPU speed.

AMD Athlon 800 January 2000, 22 million transistors. Like all the fastest Athlon Classics, these were not nearly as common as the parts in the 500 to 750MHz range.

AMD Athlon Thunderbird 700

June 2000, 37 million transistors. The second-generation Athlon Thunderbird did for the Athlon what the Coppermine did for the Pentium-III. Like the Coppermine, it had 256k of full-speed on-chip cache instead of 512k of cache at one-third or one half speed located off the chip.

AMD Athlon Classic 900, 950 and 1000

March 2000, 22 million transistors. Right from the start the Athlon was designed to be run at high clockspeeds but , the higher speeds provided less of a performance boost than might be imagined.

AMD Athlon Thunderbird 1000

June 2000, 37 million transistors. The fastest of the original first-release Thunderbirds was priced to match. But no-one much bought these either, as the particular appeal of that four-figure number was such that AMD felt able to command a hefty premium for it.

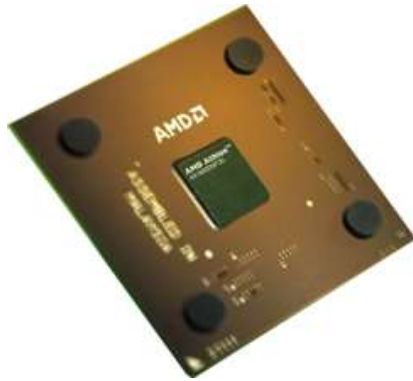


Fig 3.32 Thunderbird 1000 **Fig 3.33 AMD Athlon XP 1500+**

AMD Athlon Thunderbird 1100

August 2000, 37 million transistors. The last of the 200MHz bus Thunderbirds ("Athlon A") to sell well - the 1200, 1300 and 1400 were all overshadowed by the 266MHz bus Athlon Cs and were rather pointless.

AMD Athlon Thunderbird 1200 C

October 2000, 37 million transistors. Over their design life, most CPU families have a high point, a version that in later years will be regarded as the definitive one. Most commonly this is one of the last two or three speed grades before the changeover to a new design.

AMD Athlon XP 1500+

October 2001, 37.5 million transistors. The Athlon XP is essentially a revised Thunderbird. There is no massive change, rather a number of incremental improvements.

AMD Athlon XP 1700+

October 2001, 37.5 million transistors. The XP 1700+ was the clear price-performance leader when the XP arrived in late 2001, and it remained the one to have for a good six months.

AMD Athlon XP 1800+

October 2001, 37.5 million transistors. The fastest CPU on earth when it first arrived, well out in front of the original Pentium 4 2000, and very reasonably priced by comparison with that spectacular budget-buster.

AMD Athlon Thoroughbred

The fourth-generation Athlon, the Thoroughbred, was released 10 June 2002 at 1.8 GHz, or 2200+.

The Thoroughbred line received an increased front side bus clock, up to 333 MT/s from 266 MT/s. This improved the processor's memory and I/O access efficiency, and improved per-clock performance as a result.



Fig 3.33 AMD Opteron

AMD Opteron

The Opteron is AMD's x86 server and workstation processor line, and was the first processor to implement the AMD64 instruction set architecture (known generically as x86-64). It was released in April, 2003 with the SledgeHammer core (K8) and was intended to compete in the server and workstation markets, particularly in the same segment as the Intel Xeon processor.

AMD Athlon 64 (Single Core)

The Athlon 64 is an eighth-generation, AMD64-architecture microprocessor produced by AMD, released in September, 2003. It is the third processor to bear the name Athlon, and the immediate successor to the Athlon XP. It is AMD's first K8, eighth-generation processor core for desktop and mobile computers. Clock rate varied from 1800 to 2800 MHz.

AMD Sempron

The first Sempron CPUs, introduced in July 2004, were based on the Athlon XP architecture using the Thoroughbred or Thorton core. It has CPU clock ratings from 1.4 GHz to 2.7 GHz. Sempron has been the marketing name used by AMD for several different budget desktop CPUs, using several different technologies and CPU socket formats. The Sempron replaced the AMD Duron processor and competes against Intel's Celeron series of processors.

AMD Athlon 64 (Double Core)

The Athlon 64 X2 is the first dual-core desktop CPU manufactured by AMD. Introduced in year 2005, it has the CPU clock from 1.0 GHz to 3.2 GHz. In 2007, AMD released two final Athlon 64 X2 versions: the AMD Athlon 64 X2 6400+ and 5000+ Black Editions. Both processors feature an unlocked multiplier, which allows for a large range of overclocked settings.

Athlon 64 FX

The Athlon 64 FX is positioned as a hardware enthusiast product, marketed by AMD especially toward gamers. Unlike the standard Athlon 64, all of the Athlon 64 FX processors have their multipliers completely unlocked. In AMD Athlon 64 FX series, Clock rate varies from 2200 to 3200 MHz.

AMD Turion 64

Turion 64 is the brand name AMD applies to its 64-bit low-power consumption (mobile) processors. The Turion 64 and Turion 64 X2 processors compete with Intel's mobile processors, initially the Pentium M and currently the Intel Core and Intel Core 2 processors.

AMD Phenom

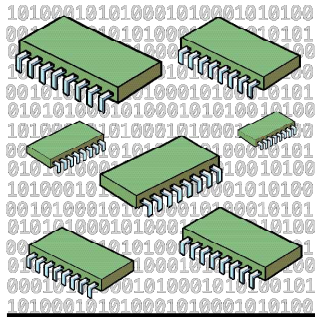
Phenom is the AMD desktop processor line based on the K10 microarchitecture, or Family 10h Processors, as AMD calls them. Triple-core versions (codenamed Toliman) belong to the Phenom 8000 series and quad cores in the AMD Phenom X4 9000 series. Introduced in 2007, it has CPU clock ratings from 1.8 GHz to 2.6 GHz.

EXERCISE

1. Fill in the blanks :
 - a) CPU acts as the _____ of the computer.
 - b) CPU consists of _____, _____ & _____.
 - c) The registers are located in _____.
 - d) Bus is set of wires that connect _____ and _____ together.
 - e) The address bus carries the _____, where as the control bus carries _____ among various units of a computers.
 - f) Microprocessor could be of three types _____, _____ & _____.
 - g) RISC stands for _____.
 - h) Segment registers, generate _____ when combined with another value called _____.

2. State True or False :
 - a) The ALU performs arithmetic operations as well as Logical operations.
 - b) Zilog 80 microprocessor is still being manufactured.
 - c) Accumulator is a part of RAM.
 - d) There are four main types of data buses used in PCs.
 - e) Motorola 68000 series is used in IBM computers.

3. Explain AMASK, PC, IR, MAR.
4. Explain Fetch Cycle.
5. Differentiate among CIS, RISC and EPIC. Which of them is latest technology and what are the plus point of the latest technology above two others?
6. What is variable length instruction? How it is different from fixed length instructions?
7. What is VLIW? In which technology it is used and why?
8. What is pipelining?
9. Who are the major players in the CPU manufacturing?
10. Discuss atleast 5 variants of Intel Pentium? Which other company has been able to produce better chips than Intel. Discuss 5 variants of that company also.



CHAPTER 4

MEMORY CONCEPTS

ROM-BIOS
RAM AND ROM
CONVENTIONAL MEMORY: THE 640 K LIMIT
EXTENDED MEMORY

Memory Concepts

All programs need computer memory to run. An operating system needs memory, pop-up calendar needs memory, a disc cache needs memory, a mouse driver needs memory, the more functional and feature laden the program, the more memory it consumes.

All computers have a basic, 640K memory allotment, this is referred to as conventional or base memory. Many applications can use this standard memory only. Such programs cannot utilize anything but the base memory even if your system has 16 MB of memory. The problem is that by the time you load start-up programs such as the operating system (COMMAND.COM), mouse network drivers, and a disc cache, conventional memory may be reduced to 550 KB or less.

It is a common misconception that more the memory that is installed, the more programs we can run. Having more memory is always desirable, but not all programs are capable of utilizing the entire memory in the computer. All memory is not alike, and memory can be configured to behave and operate in different ways. For example, on a 286 and later systems, one can usually configure memory as expanded or extended memory. Some programs can use only expanded memory, some can use only extended memory, and some can use either. Lotus 1-2-3 Release 2.4 uses expanded memory whereas Lotus 1-2-3 Release 3.4 uses extended memory.

Different techniques are available to make sure that every bit of free memory is put to good and efficient use. Before exploring these techniques in detail, understanding of some basic concepts is necessary.

The memory is organized into locations. Each memory location is known as one memory word. Each memory location is identified by a unique address. The first location's address is zero. This is specified by all zero's in the address. The number of bits necessary to specify the address depends on the capacity of the memory. If there are "n" address bits, it is possible to address the complete range of memory with 2^n locations or words. For example, if there are 10 address bits the maximum memory size is 1024 locations ($2^{10} = 1024$).

ROM-BIOS

ROMs which are used on the system board contains a piece of software called BIOS. ROM-BIOS, which stands for read-only memory -Basic Input\ Output System. DOS uses ROM-BIOS to perform low level tasks such as retrieving data from the disk or outputting data to a COM port, ROM-BIOS is therefore very hardware dependent.

BIOS contains a set of low level programs that directly manipulates hardware. There are two functions of the ROM -BIOS. First function is used when the computer is turned on to check and initialize programs that check to see that the computer is in working order. The

second and the most important part of the BIOS are its routines. These programs provide the detailed and intimate control of various parts of the computer particularly I/O.

If you add new hardware to an older system - for example, if you add new hard drive to a Compaq Deskpro 286 - you may find that you also need to replace the ROM-BIOS chips. These memory chips are not soldered into the circuit board; you may simply pull out the old chips and push in the new ones.

The hidden DOS file IO.SYS, located in the root directory of the boot disk, includes patches, updates, and extensions to ROM-BIOS. This file provides a convenient mechanism for extending PC's BIOS capabilities without replacing the ROM-BIOS chips.

When the IBM PC's were initially distributed, they included proprietary ROM-BIOS code developed by IBM. The emergence of non-IBM PC's, known as clones or PC compatibles, some vendors began developing their versions of ROM-BIOS. Some of the best known BIOS developers are American Megatrends, Award and Phoenix Technologies.

Every ROM-BIOS includes the manufacturers name, a version number, and the data when the BIOS was written. Users of DOS 6 can use the utility MSD. EXE to review the type of BIOS installed on your PC.

HOW DOES THE BIOS WORK?

The first part of the BIOS contains the start-up routines, which get the computer going when you turn on the power. There are two main parts to the start-up routines, which test to see that the computer is in good working order. They check the memory for defects and perform other tests to see that the computer is not malfunctioning. The second part of the start-up procedure is the initialization.

The initialization involves such things as creating the interrupt vectors, so that when interrupts occur the computer switches to proper interrupt handling routine. Initialization also involves setting up the computer's equipment. Many of the parts of the computer need to have registers set, parameters loaded, and other things done to get them ready to go.

HOW DO YOU ADD TO THE ROM-BIOS?

Any new experiment requiring special BIOS support, such as an optical disk, places its read-only BIOS memory some where in that block and includes in it a special marking, hex 55 AA, in the first two bytes. The BIOS cannot be located just anywhere. It has to be in a unique location that doesn't conflict with any other BIOS extensions and must begin on a 2 KB memory boundary. The standard BIOS, as part of its start-up routines, searches the BIOS extension area for identifying 55 AA signature. When its find one, it passes control over to whatever it needs to do to initialize its equipment and integrate itself into the rest of the BIOS. E.g. a BIOS extension for a new kind

of display adapter might change the interrupt vector for video services to direct them to the BIOS extensions rather than to the old BIOS video routines.

A BIOS extension performs whatever start-up and initialization work it has to do when the main BIOS passes control to it during the start-up procedure. When the BIOS extension is initialized, it passes control back to the main BIOS support for that equipment has been added.

Memory and Hardware

PC doesn't just have a single, homogenous kind of memory. It has a variety of hardware used to store information, each with different advantages and disadvantages. Knowledge of these is necessary to manage the memory effectively.

Chips and Disks

There are a number of physical ways to store a bit, to save it to memory for use or inspection later. You can use electricity and refer to an electric charge as a "1" and no charge as a "0". Or you could use magnetism and refer to a magnetized area as a "1" and an unmagnetized area as a "0". Computers happen to use both of those schemes, storing electric charges in tiny circuits on chips and magnetic fields on iron oxide coated disks and tapes. The information is the same, only the medium changes.

Disks are considered nonvolatile memory, i.e. they keep their stored information even when the computer's power is turned off. Many memory chips, on the other hand, are volatile, they lose their information when they lose power.

RAM AND ROM

The two basic classifications of memory are Read-only memory (ROM) and Random Access Memory (RAM).

When the PC is first turned on or the system is re-booted, the PC goes through a special initialization process. You may have noticed that the memory check being displayed or the disk drives spinning before DOS is loaded. The instructions that guide this process are stored in ROM. As its name implies, read only memory is read but never written. The instructions are fused into the memory chips during manufacturing and are permanent.

The random access memory, or RAM, is initially empty (or at least, it contains no meaningful data). The computer uses RAM to store programs and data temporarily. RAM is the life line of computer programs. ROM is similar to a book in which text is preprinted and cannot be changed. RAM on the other hand is like a chalkboard on which data can be written, erased and rewritten. ROMs are found on expansion boards like LAN, VIDEO, SCANNER etc. RAM stands for

Random Access Memory. Any bit of RAM can be read from and written to at any time. ROM's (Read only Memory) can only be read from. You can't write new information to a ROM. They are unchangeable to protect the important information, typically programs, stored inside them. All PC's use a mixture of RAMs and ROMs for their chip memory. The ROMs hold some fundamental programs that animate the computer, the larger amount of RAM is the working area to hold computer programs and data files.

ROM ,PROM, EPROMS, EEPROMS, AND FLASH EPROMS

There are certain types of ROM that can be written to in special circumstances. EPROMs can be removed from your computer, then erased and written to by a special machine with an ultraviolet light.

- ROM** The **ROM** is mask programmed at factory by the manufacture just before shipment, as per the contents ordered by the user. The contents are determined by metal masks used during chip fabrication. Once programmed in the factory the ROM becomes permanent memory and its contents are not changed.
- PROM** The **PROM** is a field programmable device. The user buys a blank **PROM** and enters the desired contents using **PROM (burner)**. Inside the **PROM** chip there are small fuses which are burnt open (cut) during programming. It can be programmed only once and is not erasable.
- EPROM** The **EPROM** can be erased by exposing it to ultra violet light for a duration of 6 to 40 minutes. Usually, an eprom eraser achieves this function. During programming, an electrical charge is trapped in an insulated gate region. The charge is retained for more than ten years because there is no leakage path. For erasing this charge UV light is passed through a quartz crystal window (lid). This exposure to UV light dissipates the charge. During normal use the quartz lid is sealed with a sticker.
- E²PROM** The **E²PROM** is programmed and erased electrically. It can be erased and reprogrammed about ten thousand times. Both erasing and programming take about 4 to 10 ms. In EPROM, any location can be selectively erased and programmed.

RAM

The PC has primary main memory called RAM. CPU can read from or write to " HIGH SPEED " memory. Memory is packaged in bank of 8/ 9 small chips or SIMM's . SIMM's are (single in line memory module).

There are different kinds of RAM's which are used in PC's.

Static RAM

STATIC RAM is the simplest kind of memory. It is called "static" because entered data stays there in power up. They need six transistors for each bit storage location. SRAM is quite faster than DRAM. SRAM's are quite expensive and are very rarely used as main memory.

Dynamic RAM

DYNAMIC RAM is a volatile RAM i.e. data is stored temporarily. Each bits of single transistor and capacitor. DRAM based system requires refresh circuitry. Each bit has row address and column address.

DRAM comes in the following different types of "packages":

SIMMs

Single In-line memory module (SIMM) - A slender circuit board dedicated to storing memory chips. Each chip is capable of holding 8 to 9 chips per board, the ninth chip usually an error checking chip (parity / non parity). The typical BUS from the chip to the motherboard is 32-bits wide.

The original SIMM had a 30-pin format but this has long-since been superseded by the 72-pin variety. A typical motherboard offers four SIMM sockets which will take either single-sided or double-sided SIMMs with module sizes of 4, 8, 16, 32 or even 64MB. As the Pentium processor has a 64-bit external data bus, two modules of the same size are required as a minimum configuration.

Speed

SIMM speeds are measured in nanoseconds (ns), 9 decimals, 1 billionth of a second. Speed is the time the computer takes to access one byte of information. The smaller the number the faster the RAM operates. 80ns is faster than 150ns; 1MB is faster at 80ns than 1MB at 100ns.

Banks

SIMMs are put into the computer in special slots called SIMM Slots, which are arranged in groups called Banks. Banks may consist of 1, 2, or 4 SIMM Slots.

Banks which have only one SIMM Slot would be referred to as Independent Slots. Computers with more than one Independent Slot can be upgraded with SIMM chips in any combination; the SIMMs which are added do not have to be the same size in megabytes.

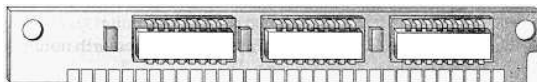


Fig 4.1 30 Pin SIMM

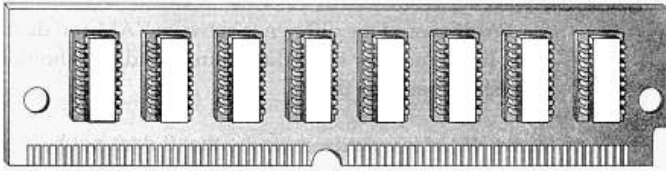


Fig 4.2 72 Pin SIMM

DIMMs

Dual In-line Memory Module (DIMM) - Type of Circuit board that holds memory chips. DIMMs have a 64-bit path because of the Pentium Processor requirements. DIMM's have been the prevalent technology since around 1997. Physically they look similar. One difference is that SIMM's are installed in pairs, whereas DIMM's are installed singly. Another is that they are "keyed" differently - their notch positions along the connector edge are in different locations. They also install differently. SIMM's are first inserted into the memory slot at a slight angle and then rotated into a straight-up position. DIMM's are inserted straight-down into the memory slot.

RIMMs

With the introduction of Direct RDRAM sometime in 1999 will come the RIMM module (the name is not an acronym, but a trademark of Rambus Inc.). RIMMs will use the same socket specification as a standard 100MHz DIMM. A PC's BIOS will be able to determine what type of RAM is fitted, so 100MHz SDRAM modules should work in a RIMM-compatible system. However, systems won't be able to use RIMMs unless both the chipset and BIOS support it - and Direct RDRAM support won't come until Intel's next generation of chipsets.

The following are common different types of DRAM:

Fast Page Mode DRAM

All types of memory are addressed as an array of rows and columns, and individual bits are stored in each cell of the array. With standard DRAM or FPM DRAM, which comes with access times of 70ns or 60ns, the memory management unit reads data by first activating the appropriate row of the array, activating the correct column, validating the data and transferring the data back to the system. The column is then deactivated, which introduces an unwanted wait state where the processor has to wait for the memory to finish the transfer. The output data buffer is then turned off, ready for the next memory access.

At best, with this scheme FPM can achieve a burst rate timing as fast as 5-3-3-3. This means that reading the first element of data takes five clock cycles, containing four wait-states, with the next three elements each taking three.

DRAM speed improvements have historically come from process and photolithography advances. More recent improvements in performance however have resulted from changes to the base DRAM architecture that require little or no increase in die size. Extended Data Out (EDO) memory is an example of this.

Extended Data Out DRAM

EDO memory comes in 70ns, 60ns and 50ns speeds. 60ns is the slowest that should be used in a 66MHz bus speed system (i.e. Pentium 100MHz and above) and the Triton HX and VX chipsets can also take advantage of the 50ns version. EDO DRAM doesn't demand that the column be deactivated and the output buffer turned off before the next data transfer starts. It therefore achieves a typical burst timing of 5-2-2-2 at a bus speed of 66MHz and can complete some memory reads a theoretical 27% faster than FPM DRAM

Burst Extended Data Out DRAM

Burst EDO DRAM is an evolutionary improvement in EDO DRAM that contains a pipeline stage and a 2-bit burst counter. With the conventional DRAMs such as FPM and EDO, the initiator accesses DRAM through a memory controller. The controller must wait for the data to become ready before sending it to the initiator. BEDO eliminates the wait-states thus improving system performance by up to 100% over FPM DRAM and up to 50% over standard EDO DRAM, achieving system timings of 5-1-1-1 when used with a supporting chipset.

Despite the fact that BEDO arguably provides more improvement over EDO than EDO does over FPM the standard has lacked chipset support and has consequently never really caught on, losing out to Synchronous DRAM (SDRAM).

SDRAM

The more recent Synchronous DRAM memory works quite differently from other memory types. It exploits the fact that most PC memory accesses are sequential and is designed to fetch all the bits in a burst as fast as possible. With SDRAM an on-chip burst counter allows the column part of the address to be incremented very rapidly which helps speed up retrieval of information in sequential reads considerably. The memory controller provides the location and size of the block of memory required and the SDRAM chip supplies the bits as fast as the CPU can take them, using a clock to synchronise the timing of the memory chip to the CPU's system clock.

This key feature of SDRAM gives it an important advantage over other, asynchronous memory types, enabling data to be delivered off-chip at burst rates of up to 100MHz. Once the burst has started all remaining bits of the burst length are delivered at a 10ns rate. At a bus speed of 66MHz SDRAMs can reduce burst rates to 5/1/1/1. The first figure is higher than the timings for FPM and EDO RAM because

more setting up is required for the initial data transfer. Even so, there's a theoretical improvement of 18% over EDO for the right type of data transfers.

PC100 (100 MHz speed) and PC133 (133 MHz speed) are two types of SDRAM. PC1600 and PC2100 are types of DDR. They are all DIMM's in physical construction. The type of memory, SDRAM or DDR, that can be used in a computer is a characteristic of the motherboard. Most motherboards support either SDRAM or DDR, but not both. However, there are some motherboards that allow either, but not both at the same time. SDRAM and DDR use different types of memory slots for installation. PC100 memory is designed for use with 100MHz SDRAM FSB. PC133 memory is designed for use with a 133MHz SDRAM FSB. PC1600 is designed for use with a 200MHz DDR FSB. PC2100 is designed for use with a 266MHz FSB. The newer technology is DDR.

Parity memory

Memory modules have traditionally been available in two basic flavours: non-parity and parity. Parity checking uses a ninth memory chip to hold checksum data on the contents of the other eight chips in that memory bank. If the predicted value of the checksum matches the actual value, then all is well. If it does not, then the contents of memory is corrupted and unreliable. In this event a non-maskable interrupt (NMI) is generated to instruct the system to shut down and thereby avoid any potential data corruption.

Parity checking is quite limited - only odd numbers of bit errors are detected (two parity errors in the same byte will cancel themselves out) and there's no way of identifying the offending bits or fixing them - and in recent years the more sophisticated and more costly Error Check Code (ECC) memory has gained in popularity.

ECC memory

Unlike parity memory, which uses a single bit to provide protection to eight bits, ECC uses larger groupings. Five ECC bits are needed to protect each eight-bit word, six for 16-bit words, seven for 32-bit words and eight for 64-bit words.

Additional code is needed for ECC protection, and the firmware that generates and checks the ECC can be in the motherboard itself or built into the motherboard chipsets (most Intel chips now include ECC code). The downside is that ECC memory is relatively slow - it requires more overhead than parity memory for storing data and causes around a 3% performance loss in the memory sub-system. Generally, use of ECC memory is limited to so-called mission-critical applications and is therefore more commonly found on servers than on desktop systems.

What the firmware does when it detects an error can differ considerably. Modern systems will automatically correct single-bit errors, which

account for most RAM errors, without halting the system. Many can also fix multi-bit errors on the fly or, where that's not possible, automatically reboot with the bad memory mapped out.

CACHE MEMORY

Cache memory is high speed memory, usually using SRAM technology, that interfaces between the processor and the system memory. Cache memory is special high speed memory, which sits on the processor's address and data buses in parallel with the much larger main memory. Its purpose is to supply the processor with the most frequently requested instructions and data. This allows the CPU to access the information many times faster than instructions and data located in main memory. Memory access has a significant impact on the overall speed of the computer. Cache memory is much more expensive than main memory. This means that the amount of cache on a system is usually a small percentage of the overall system memory. An algorithm is required to allow the CPU to make most of its accesses to the high speed cache rather than the slower RAM. There are three methods commonly used to organize cache memory: direct mapping, associative mapping and set associative mapping.

When it comes to access speed, processors are currently outstripping RAM memory chips by an ever-increasing margin. This means that processors increasingly having to wait for data going in and out of main memory. Using Cache Memory between the main memory and the processor, with the use of clever electronics and algorithms, helps ensure that the data the processor needs next is already in cache.

Level 1 Cache

L1 Cache is cache memory that is located within the CPU and is typically 32 kBytes. In modern CPU chips there is usually an L1 Data Cache and an L1 Instruction Cache.

Level 2 Cache

Level 2 Cache is a type of cache that is also located within the CPU chip in modern CPUs. (Previously, it was memory external to the CPU.) Typical Secondary Cache sizes are 512 kBytes or 1MByte (as usual, these kinds of numbers will be "dated" very soon).

HARDWARE & SOFTWARE LIMITS FOR MEMORY

There are limits. The number of memory slots is one constraint. Windows 95, Windows 98 and Windows 98SE users should stick to no more than 512MB. Windows ME and Windows XP support up to 4GB. There may be other limiting factors imposed by your computer or motherboard.

It's safe to say you will see a performance improvement in at least one way or another by having at least a total of 256MB, if your computer supports that much. For running today's computer software and

operating systems, it's entirely reasonable to equip your computer with up to 512MB of memory. Although there's no standard amount so different models of computers have different maximums.

CONVENTIONAL MEMORY: THE 640 K LIMIT

The regular RAM memory in your PC is called conventional or "BASE" memory. This was 256K in some early model PC's. Now most came with at least 512K. The hardware in a PC or XT allows a maximum of 1 MB. PC hardware that is 486-based can handle as much as 4096 MB (4 Gigabytes). But the software in both standard PC's and high powered 486 based PC's typically allows a maximum of 640 K. If there is 640K or less then the PC is limited solely by how many memory chips are present in the system. But if more than 640K of chips are plugged in, these are other limits as given below:

Hardware Limits for Conventional Memory

(i) Space - The physical place to plug in the memory chips and boards and the electrical connections to power them. Some Earlier PC's and small laptops - don't have any slots where you can plug in memory boards. At best they might offered a special socket for adding a few memory chips, perhaps as much as a megabyte. Big desktop PC's also with lots of slots used to have plenty of room for memory chips, but had a limited amount of electrical power for them.

(ii) Addresses - The processor's ability to find the chips for reading and writing information.

In earlier PC's, with six or more slots, it was theoretically possible to plug in boards with as much as 100 MB of RAM chips, but a typical PC processor couldn't address that much RAM.

Different generations of PC's use different processor chip that can address certain amount of memory, which depends on how many address bits at a time it works with. The direct address space is the memory address that a processor can reach directly, quickly and efficiently.

Hence the memory address space and the memory in the system are different. If the processor chip could reach a certain address, does not mean there is a memory chip waiting there to talk to. Similarly, the existence of a memory chip does not mean the processor can address it. The memory needs both.

The 8088 chip in the original PC, and in the XT, worked with 20 address bits at a time and could address 1 MB. (1024K). (The previous generation of computers worked with 16 address bits and could handle only 64K or 65536 bytes of memory). The IBM AT- and all the AT compatibles including those built around the 286 chip can address 16 MB.

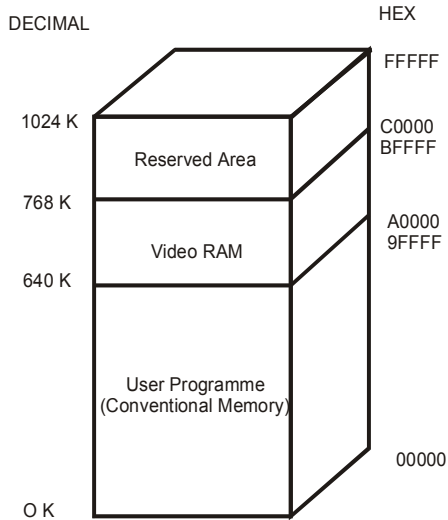


Fig 4.3 Distribution of 1024 K RAM

Software limit for conventional memory

None of the earlier systems, the PC's, XT's, AT's, or even the PS/2's can use more than 640 K of memory as conventional memory for their programs, no matter how many megabytes of chips of memory addresses they have. The software limits stops them from touching anything more.

DOS the operating system software that nearly every PC uses, only recognizes 640 K of conventional memory for the regular programs and data. That limit is the single largest factor in the memory management for PC's. Running into it was traumatic for computer owners who expected memory to keep growing. The limit is so fundamental to DOS, built right into the very care of the operating system, that it just cannot be eliminated. There are however schemes for slipping around the limit, for avoiding enforcement of the 640K law.

In fact even the 640 K is not available to the user. DOS uses a part of the 640K - from 40K to 100K or more for itself. (The amount depends on the version of DOS one uses (DOS 5.0 and above will use less).

High Memory

384 K above conventional memory is called the high memory or the reserved memory. The regular program and data can't use high memory the way they use conventional memory, in fact they are barred from using it at all. Any RAM chips that would have been at those addresses in high memory are preempted by the reserved uses. The messages to High Memory are automatically detoured to the

reserved use. That 384 K was reserved by the programmers who created DOS for BIOS and other ROMS.

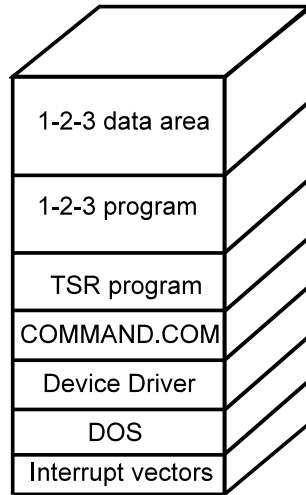


Fig 4.4 Memory Map using DOS applications

Base and Upper memory

Most DOS users have heard about the 640K barrier or ceiling. This ceiling represents the most memory a program can use, and it continues to be a source of frustration to this day. But it was described how the original PC engineers devised a memory addressing system that allowed the CPU to access 1 MB of memory. So why is there a 640K ceiling rather than a 1 MB ceiling?

Although there is 1 MB of addressable memory in a PC, not all of it is directly available to DOS programs. The area between 640K and 1 MB is reserved for use by special adapter boards and ROM-BIOS. Some memory, for example, is reserved for the video display.

It is safe to say that the amount of memory address space dedicated to upper memory is very generous, especially when you consider how precious base memory is. In many cases, one third or more of upper memory is not used at all. If you have a 386 or later system, you can steal this area of memory and use it for storing DOS programs.

Over the past few years, there have been many changes in memory technology. Unfortunately, there have been just as many changes in memory terminology. Some memory terms used in the late 1980's now have a different meaning.

For example, in old literature, you may see upper memory referred to as high memory. Worse still, the term high memory now has a different meaning. This can be a real source of confusion. The good news is that it appears that terminology has finally matured.

The sketch of the memory warehouse, shows the structure of base and upper memory. The memory warehouse illustrates a number of important details:

- The DOS command interpreter, COMMAND.COM, automatically uses the first portion of the 640K base memory. This leaves less than 640K for your programs.
- Device drivers (such as mouse and network drivers) and other TSR programs use up still more of your base memory.
- Although 8086 computers can address 1,024K of memory, most computers do not actually have that much general purpose RAM installed. The memory addresses in upper memory are reserved for memory located else where. For example, the ROM-BIOS memory is located on the ROM chips, and the video memory is located on the video card. In the sketch, they are shown as storage containers that occupy part of the warehouse space.

EXPANDED MEMORY

Arguably the most successful PC program and the one that first put PC's on the desktops of corporate America was Lotus 123. But corporate America soon demanded bigger spreadsheets, and PC's simply didn't have enough memory to do the job.

The History of EMS

The original expanded memory specification developed by Lotus and Intel are called LI-EMS. Believe it or not, this first specification was dubbed EMS version 3.0. Presumably, the authors believed we would give more credence to a third generation product. The original specification detailed how memory hardware should be engineered and how software should communicate with it.

MicroSoft joined forces with Lotus and Intel to produce a revised specification known as LIM-EMS version 3.2. The primary change in the specification between 3 and 3.2 was the removal of strict hardware controls, which in turn allowed more vendors to produce expanded memory products.

A number of other PC vendors believed that LIM-EMS was too conservative and didn't utilize the full potential of expanded memory. Ashton Tate (now part of Borland International), Quarterdeck Office Systems (the DESQ view folks), and board-maker AST Research banded together and developed a better mousetrap, called the Enhanced Expanded Memory Specification.

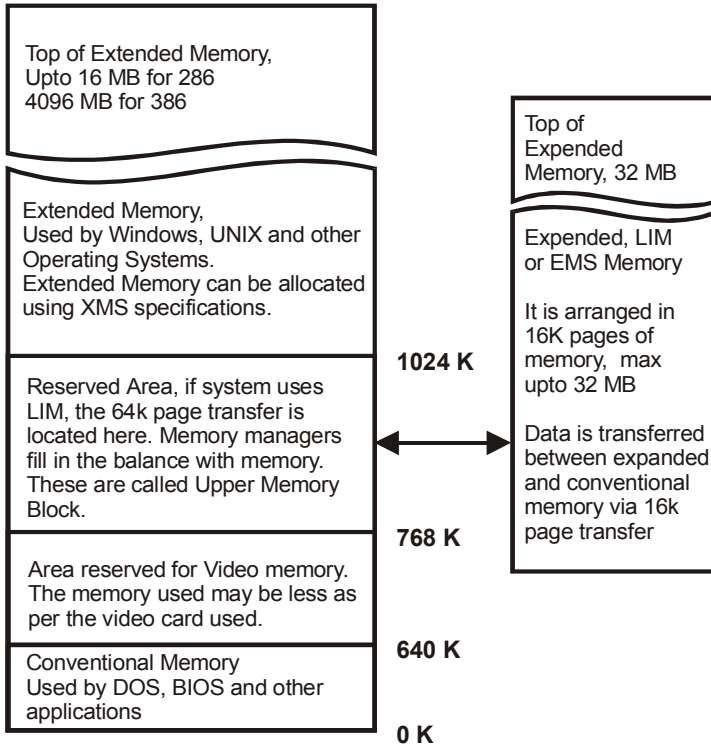


Fig 4.5 Memory Map with Expanded Memory

EEMS supported all the features of LIM-EMS 3.2. The most important new feature was the capability to load and run programs from the EMS page frame. The original specification was really oriented to managing data and not programs in expanded memory. By adding support for programs, EEMS provided an excellent vehicle for multitasking programs - that is, running more than one program at a time and switching between them.

While all this was taking place, other vendors, such as IBM, were developing their own EMS enhancements. The final chapter in the EMS story was written by the LIM team. In a cooperative effort to settle on one standard, the LIM team took the best features of all the hybrid EMSs and introduced LIM-EMS version 4.0. This specification incorporated most of the features of EEMS, EMS 3.2, and other vendors' variations. Today, LIM-EMS 4.0 stands as the dominant expanded memory specification for the PC industry.

Three major companies in the PC business got together to devise a way to make more memory available: Lotus Development Corporation was the spreadsheet company, MicroSoft Corporation was the operating system company, and Intel Corporation was the hardware manufacturer. They called their solution expanded memory.

As you have seen, there is a lot of wasted space in upper memory. The obvious solution to the memory shortage problem was to make better use of upper memory. But upper memory could only offer a few hundred thousand extra bytes at most, and it was becoming clear that users wanted spreadsheets that were two or three megabytes in size. The answer to this dilemma was memory paging.

Lotus, Intel, and MicroSoft designed the expanded memory specification, or EMS. This specification often is called the LIM-EMS, named after the three main contributors (Lotus, Intel, and MicroSoft). The theory behind LIMEMS is quite simple. A 64K area of upper memory is reserved for use as expanded memory. This area, called the EMS page frame, is subdivided into four 16K pieces. Those four pieces can be taken from a pool of 32 MB of memory.

A program cannot access all the expanded memory at one time. Rather, it must manage the data in 16K pieces. The process of assigning different blocks of memory into the page frame is known as bank switching.

In the memory warehouse sketch, the expanded memory pool is represented by 16K container trucks. This memory is separate from the main 1 MB warehouse and must be shunted in small pieces into the warehouse before DOS can access it. The truck analogy isn't quite accurate because the 16K pieces of expanded memory aren't actually moved into upper memory. A memory manager fools DOS into thinking the memory is located in upper memory. When DOS tries to manipulate the data in the expanded memory area, the memory manager intercepts the task and makes the real changes directly to the real memory, which is located somewhere in the 32 MB expanded memory area.

The expanded memory solution for 8086 computers is, at best, a work-around. Although it provides a lot more memory to applications, the bank switching process is not efficient. Dealing with memory in small 16K pieces is restrictive and time consuming. Ideally, software needs to access huge blocks of contiguous memory.

EXTENDED MEMORY

When Intel developed the 80286 microprocessor (286), the company decided to implement a new memory architecture that provided access to large volumes of memory and didn't suffer from the restrictions of bank switched EMS. The new memory architecture was called extended memory. The PC AT was the first IBM PC to use the 286 microprocessor.

Why the industry used the word extended, which sounds so much like expanded, is a mystery. The similarity of the names has been the source of confusion for years. It would have been more convenient if it were called something more distinguishable, like 286 memory or auxiliary memory but perhaps that would be too easy.

If you find yourself confusing the two memory types, remember that the p in expanded comes before the t in extended. Another, somewhat backward, memory jogger is to remember that extended memory was not available on the XT.

Extended memory blows the lid off the 1 MB addressable limit suffered by 8086 processors. Internally, the 286 uses 24 bits to address memory, compared with the 8086's 20 bits, yielding a maximum memory capacity of 16 MB. However, because it is backward compatible, the 286 microprocessor can function just like an 8086, thus ensuring that all software that runs on an 8086 also runs on the 286.

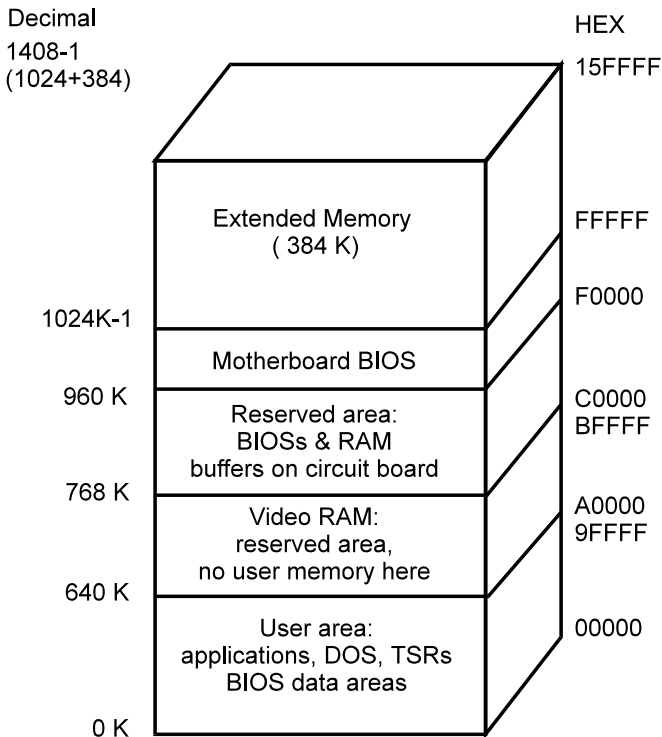


Fig 4.6 Memory Map with Extended memory

The 286 has two modes of operation. When the 286 is emulating the 8086, it is said to be in Real Mode. (They should have called it 8086 Mode). To gain full access to the 15 MB of memory beyond the traditional 1 MB limit, the microprocessor must shift into Protected Mode. Note that the memory above the 1 MB boundary is referred to as extended memory.

In Protected Mode, the system enjoys access to 16 MB of memory, if the PC has that much memory. That makes for a big warehouse. All the extended memory is contained within the warehouse and doesn't

need to be trucked in and out. Applications that run in Protected Mode can enjoy fast access to 16 MB of memory which is a lot more than 1 MB. Soon after the 286 was introduced, many industry figures expected all major software vendors to start writing software to take advantage of Protected Mode. But it just didn't happen.

The problem was simple. Any application written specifically to use Protected Mode will not run on 8086 based computers - in other words, original PC's and XT's. A Protected Mode program is useless to a PC or XT owner. For the first few years following the introduction of the 286, the majority of the PC market place was still dominated by the 8086 computers, and that was too large a market to ignore. Nonetheless, a few companies did introduce products written exclusively for Protected Mode. Lotus 1-2-3 Release 3 was the most popular. Even though Lotus Development introduced Release 3, the company made it clear that it would continue to enhance the Release 2 version of the product to keep 8086 customers happy.

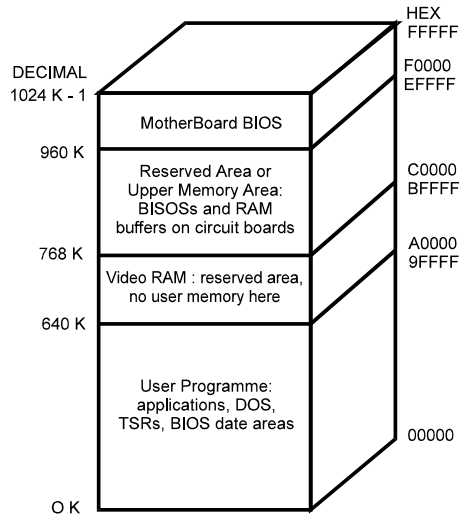


Fig 4.7 Memory Map with Reserved Area

DOS is Real Mode application and is still limited to accessing the first megabyte of memory. A Protected Mode program is launched from DOS in Real Mode. The program must then switch the microprocessor into Protected Mode to access extended memory. When the user quits the application, the microprocessor is switched back into Real Mode and control is passed back to DOS.

Why is DOS 6 a Real Mode program? If you are thinking that it would make sense to create a Protected Mode operating system, you're right. OS/2 is the best known Protected Mode operating system. Although originally developed by MicroSoft, OS/2 is now developed and supported by IBM. Meanwhile, MicroSoft provides Protected Mode

support in Windows and probably will do so in a future version of DOS.

As a real Mode application, DOS had no control or interest in Protected Mode, and so protected Mode was a free for all, with each program doing as it saw fit. A sort of extended memory anarchy prevailed. In 1988, some four years after manufacturers shipped the first 286 computers, the major PC vendors got together to develop some industry standards on the use of extended memory. The people that were most involved in expanded memory (Lotus, Intel, MicroSoft and AST) released a new specification to extended memory. Not surprisingly, this specification was dubbed the extended memory specification or XMS. The successor to the 286 microprocessor was the 80386, or 386 for short. The 386 could do everything the 286 could do, and it could do it faster. The clock speed of the 286 ranged from 6 to 12 MHz, whereas the 386 clock speed ranged from 16 MHz to 33 MHz and beyond.

As well as greater speed, the 386 offered some major new features - A 386 can access 4096 MB of RAM. Yes, four billion bytes. Even by today's standards, that's a big spreadsheet.

Another innovation introduced in the 386 was that extended memory, with some help from software drivers, could behave and act like expanded memory. This allowed 386 users with lots of memory to take advantage of this memory in applications written to use expanded memory.

Virtual Memory

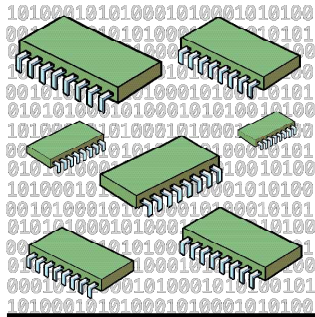
Probably the most significant new technology introduced in the 386 was the ability of the microprocessor to run in Virtual 8086 Mode. In this mode, the 386 supports multiple 8086 sessions - in other words, the system can be configured to use extended memory to emulate multiple 640K base memory areas. Each session appears to the software as though it is running in Real Mode. Products such as DESQ view 386 and Windows 3.x in Enhanced Mode take advantage of this capability, allowing you to run multiple DOS sessions concurrently.

The unused memory blocks in the upper memory area can be filled with extended memory. DOS can then access this memory to store itself, TSRs, and device drivers, freeing precious base memory for applications software.

After the 386, Intel introduced the 486 microprocessor. The 486 offers a faster clock speed than its predecessor, a built in math co-processor, and numerous other technological advances. In terms of memory management, the 486 and the further system are the same as the 386.

EXERCISE

1. Fill in the Blanks:
 - a) E²PROM is _____.
 - b) In EPROM, _____ light is passed for erasing.
 - c) SIMM, DIMM and RIMM are types of _____.
 - d) Conventional memory is limited to _____ in present PC.
 - e) Command.com uses _____ of base memory.
 - f) LIM EMS stands for _____.
 - g) A Protected Mode program is launched from DOS in _____.
2. What is ROM-BIOS and how it works.
3. What is memory and what is the difference between the types of memory RAM and SRAM?
4. What Is A SIMM? What is a 'parity' SIMM? How do you know whether or not my SIMMs are parity?
5. What does EDO mean? What is a DIMM?
6. What is the difference between FPM (Fast Page Mode) and EDO (Extended Data Output) DRAMS? Why is a SDRAM based memory faster than a FPM or EDO DRAM based memory?
7. What Are Nano-Seconds and how they are connected to memory?
8. What is a refresh Rate and Self-Refresh?
9. What is RDRAM and how does it differ from SDRAM?
10. Can you mix a 128MB PC100 SDRAM DIMM with a 128MB PC133 SDRAM DIMM? Can you mix and match parity and non-parity modules?
11. Several Memory module suppliers are offering PC133 modules with a performance setting of 3-2-2. What advantage does a 2-2-2 module provide?
12. Can you tell by looking at a module if it is SDRAM, FPM, EDO etc?
13. What is Expanded Memory? What is Extended Memory?
14. What is 286 protected mode and 386 enhanced mode?
15. What is virtual memory & how it works?
16. Write short notes on :
 - a) SIMM
 - b) DIMM
 - c) EDO RAM
 - d) ECC Memory
 - e) Cache Memory



CHAPTER 5

THE MOTHERBOARDS

SPECIFICATIONS
PHYSICAL LAYOUT OF MOTHERBOARDS
THE MICROPROCESSORS
EXPANSION SLOTS
TYPES OF I/O BUS
MOTHER BOARD SETTING

The Motherboards

In any computer organization mother board is the primary component. It incorporates all the fundamental units required for a computer. The Main components of a motherboard are

- The CPU logic
- Memory (RAM / ROM)
- INPUT (PORTS)
- OUTPUT (PORTS)

Thus we see that a motherboard holds most of the primary logic required for a computer and thus justifies it's name. To adapt a computer to use widely different types of secondary devices, like - the floppy drive, the hard disk, the display adapter, the printer.

In the motherboard, we need to connect the corresponding adapters. These adapters communicate with the CPU at the motherboard through address and data buses. Since different manufacturers are manufacturing different types of motherboard and adapters therefore it becomes necessary to have some standardized designs to make all the boards adaptable to each other. This is realized through the standardized designs of I/O slots called I/O buses. Here, we will discuss about various types and configurations of motherboard, the circuits located on them and features included on the latest versions.

SPECIFICATIONS

Technical specifications for any motherboard should include the following features :

- CPU Type
- Clock speed
- RAM capacity & installed
- Cache RAM if any
- Chip Set Details
- Bus Type

HISTORY OF MOTHERBOARDS

The IBM Personal Computer, commonly known as the IBM PC, is the original IBM PC compatible hardware platform. The main circuit board in an IBM PC is called the motherboard . It carries the CPU and RAM, and it has a bus with slots for expansion cards. On the motherboard are also the ROM subsystem, DMA and IRQ controllers, coprocessor socket, sound (PC speaker, tone generation) circuitry, and keyboard interface.

XT BOARD

The IBM Personal Computer XT, often shortened to the IBM XT, PC XT, or simply XT, was IBM's successor to the original IBM PC.

A very typical XT clone board, indistinguishable from a thousand others. This one is from the golden age of the XT clones, after they had finally pushed aside the CP/M and Apple DOS machines, and while the 286 was still more a high-end promise than a mass-market reality.

Beginning in 1985, the XT was offered in floppy-only models without a hard disk. XTs with the 256k-640k motherboard came standard with half-height floppy drives in place of the full-height drives, as well as the option for a 20MB half-height hard disk and 'enhanced' keyboard.

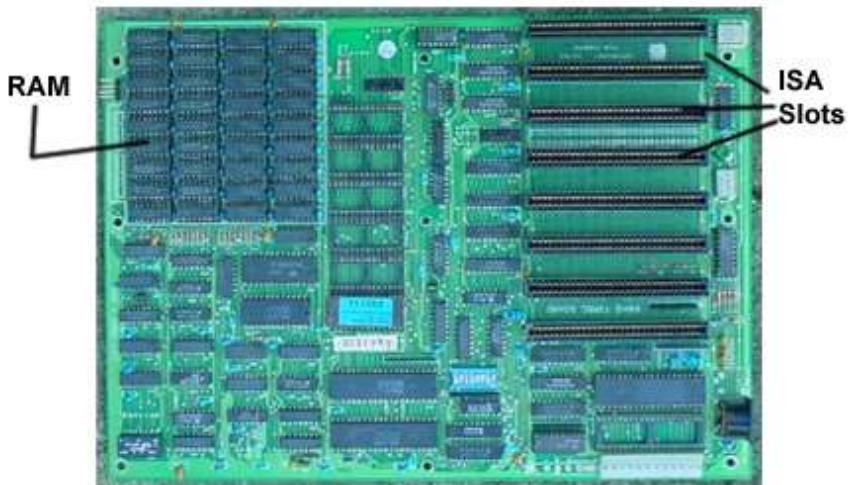


Fig 5.1 XT Board

Notice the expansion bus: eight 8-bit XT slots, which (in XT-class boards) always ran at the same speed as the system and the CPU. In the over-engineered and very conservative IBM PC, it was 4.77MHz—clone systems were often twice as fast as Big Blue's. Most of the later clone boards ran at either 8 or 10MHz. A very few even ran at 12MHz, but without the decoupled expansion bus that the 286 boards introduced, this was very difficult to make reliable.

Specification of XT Board

CPU	: Intel 8088-8 fitted, 8087 socket provided.
Speed	: 8 MHz.
Slots	: 8 8-bit
I/O	: None.
RAM	: 36 16-pin DIPP sockets, 640k fitted, 640k maximum.

Cache : None.

Chipset: None, Phoenix BIOS.

AT MOTHERBOARD

The Personal Computer/AT, also known as the IBM AT or PC AT or PC/AT, designed around the 6 MHz Intel 80286 microprocessor and released in 1984 as machine type 5170. Because the AT used various technologies that were new at the time in personal computers, the name AT stood for Advanced Technology; one such advancement was that the Intel 80286 processor used in the AT supported Protected mode. An 8 MHz version of the AT was the next to follow..

At this time, DIPP RAM and SIPP RAM was just giving way to SIMMs on the general market; The boards were having as many as four 256k SIMMs, or even an extravagant pair of 1MB SIMMs to make 2MB in total.

RAM arrangements were state of the art: sockets for 1MB of DIPP RAM at front left and centre.



Fig 5.2 16 MHz Headland Motherboard

Specifications of 16 MHz Headland Motherboard

CPU : AMD 286-16 fitted, 287 socket provided.

Speed : 16 MHz.

Slots : 5 ISA, 2 8-bit

I/O : None.

RAM : 20-pin DIPP sockets, or 4 30-pin SIMM sockets up to 16MB.

Cache : None.

Chipset: Headland H12A, AMI BIOS.

The Personal System/2 or PS/2 technology was next to invade the market. However, many of the PS/2's innovations, such as the 16550 UART, 1440 KB 3.5-inch floppy disk format, 72-pin SIMMs, the PS/2 keyboard and mouse ports, and the VGA video standard, went on to become standards in the broader PC market.

The PS/2 connection ports (later colored purple for keyboard and green for mouse —PC 97), are still used in some new computers today, although they are mostly replaced by USB connectors.



Fig 5.3 PS/2 Sockets

COMPUTER FORM FACTOR

A form factor in computing controls the physical dimensions of motherboard. Specifically, in the IBM compatible PC industry, standard form factors were introduced to ensure that parts are interchangeable across competing vendors and generations of technology, while in enterprise computing, form factors ensure that server modules fit into existing rackmount systems. Apart from controlling the dimensions of the motherboard, form factor also controls various other factors but they are not so prominent. Small form factors shrinks the size of motherboards, but further reduction in overall size is hampered by current power supply technology.

ATX MOTHERBOARDS

Designed by Intel in 1995, ATX overtook AT completely as the default form factor for new systems. ATX addressed many of the AT form factor's annoyances that had frustrated system builders. It was the first big change in computer case, motherboard, and power supply design in many years. The ATX form factor ensures that smaller boards (including microATX, FlexATX and mini-ITX) usually keep the basic rear layout but reduce the size of the board and the number of expansion slot positions.

In 2003, Intel announced the BTX standard, intended as a replacement for ATX. But still in 2009, the ATX form factor remains a standard for do-it-yourselfers; although BTX has made inroads into pre-made systems.

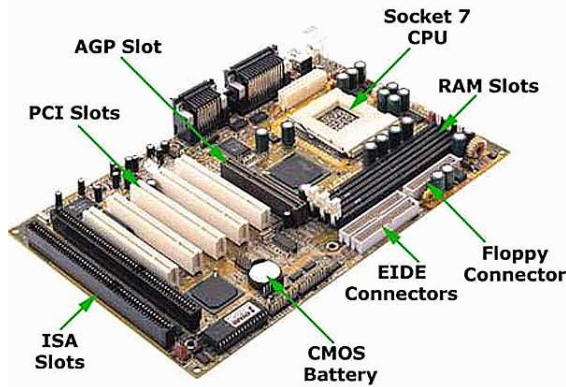


Fig 5.4 PS/2 Sockets

The official specifications were released by Intel in 1995, and have been revised numerous times since, the most recent being version 2.2, released in 2004.

A full size ATX board is 12 in × 9.6 in (305 mm × 244 mm). This allows many ATX form factor chassis to accept microATX boards as well.



Fig 5.5 PS/2 Sockets

ATX also made the PS/2-style mini-DIN keyboard and mouse connectors as standards. AT systems used a 5-pin DIN connector for the keyboard, and were generally used with serial port mice (although PS/2 mouse ports were also found on some systems). Many modern motherboards are phasing out the PS/2-style keyboard and mouse connectors in favor of the more modern Universal Serial Bus. Other legacy connectors that are slowly being phased out of modern ATX

motherboards include 25-pin parallel ports and 9-pin RS-232 serial ports.

The AT standard had only a keyboard connector and expansion slots for add-on card backplates. Any other onboard interfaces (such as serial and parallel ports) had to be connected via flying leads to connectors which were mounted either on spaces provided by the case or brackets placed in unused expansion slot positions.



Fig 5.6 A typical Black Plate of ATX Motherboard

ATX allowed each motherboard manufacturer to put these ports in a rectangular area on the back of the system, with an arrangement they could define themselves (though a number of general patterns depending on what ports the motherboard offers have been followed by most manufacturers). Generally the case comes with a snap out panel, also known as an I/O plate, reflecting one of the common arrangements. If necessary, I/O plates can be replaced to suit the arrangement on the motherboard that is being fitted and the I/O plates are usually included when purchasing a motherboard.

POWER SUPPLY CONNECTORS

The power supply connectors are two six-pin male header connectors. Plug the dual connectors from the power directly onto the board connectors.

There are several connectors on an ATX power supply. Several connectors with several different wires providing several different voltages.

Connector shown in Fig 5.7 is a 20-pin ATX power supply connector. As you can see from the color-code to the left of the photo, it supplies many different voltages (+3.3V, +5V, +12V, -12V, -5V) as well as a number of ground wires and a couple signal wires ("Pwr On" which is how the motherboard tells the power supply to turn on and "Pwr Good" which reports back that the voltage of 3.3V and 5V are within spec) There's also a +5V stand by (+5VSB) that constantly supplies power to the board even when the PC is off. This voltage is used for several different CMOS functions like Wake On LAN, Wake On Ring, etc.

The connector above is still an ATX main power connector, but is of the 24-pin variety. One each additional +3.3V, +5V, +12V and Ground have been added to the connector.

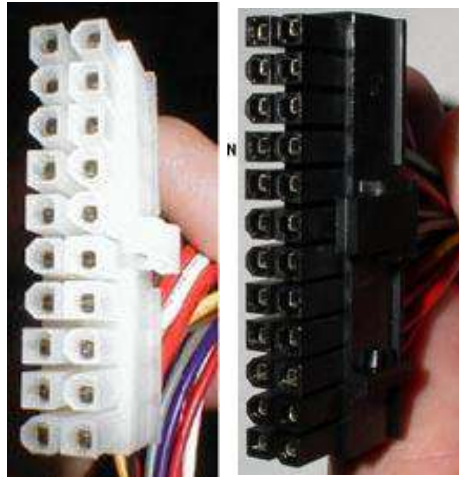


Fig 5.7 Power supply connector

The extra 4-pins simply provide more power to the motherboard over more wires. This reduces resistance, heat and can provide better voltages.

Color	Signal	Pin	Pin	Signal	Color
Orange	+3.3 V	1	13	+3.3 V	Orange
				+3.3 V sense	Brown
Orange	+3.3 V	2	14	-12 V	Blue
Black	Ground	3	15	Ground	Black
Red	+5 V	4	16	Power on	Green
Black	Ground	5	17	Ground	Black
Red	+5 V	6	18	Ground	Black
Black	Ground	7	19	Ground	Black
Grey	Power good	8	20	-5 V (obsolete)	White
Purple	+5 V standby	9	21	+5 V	Red
Yellow	+12 V	10	22	+5 V	Red
Yellow	+12 V	11	23	+5 V	Red
Orange	+3.3 V	12	24	Ground	Black

PHYSICAL LAYOUT OF MOTHERBOARDS

There may be wide variety of PC motherboards ranging from PC XT, PC AT(286), PC AT(386) etc. based on different types of chip sets. Thus, depending upon the manufacturers and the chip sets used along with other circuitries provided on board, there may be widely different types of physical layout of motherboards.

One main aspect which as a learner we should look into is the main circuits and the components which may be present on any class of motherboard.

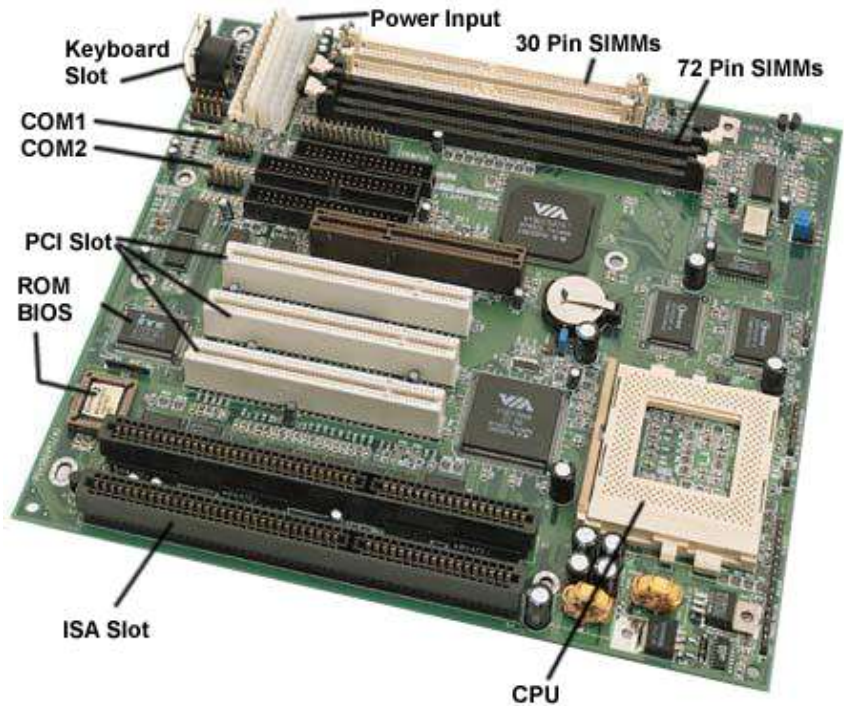


Fig : 5.8 A Typical Motherboard

Design variations in the existing motherboards are primarily because of the following reasons -

SIZE

Earlier versions of motherboards used to be very big and the circuits were discretely fabricated on them. There used to be separate controller chips for different operations like keyboard, interrupt control, DMA control, etc.

But the advent of VLSI technology has produced compact circuits and hence reduced the size of motherboards. Now we're having compact VLSI chip sets available from different manufacturers which incorporate the complete motherboard circuits following the ATX standards.

Various Varients of ATX Boards

There exist several ATx-derived form factors that use the same power supply, mountings and basic back panel arrangement, but set different standards for the size of the board. The two most popular are the Standard and Micro ATX sizes.

	Width	Length
FlexATX	9 in (229 mm)	7.5 in (191 mm)
microATX and EmbATX	9.6 in (244 mm)	9.6 in (244 mm)
Mini ATX	11.2 in (284 mm)	8.2 in (208 mm)
Standard ATX	12 in (305 mm)	9.6 in (244 mm)
EATX (extended ATX)	12 in (305 mm)	13 in (330 mm)
EEATX (enhanced extended ATX)	13.68 in (347 mm)	13 in (330 mm)
WTX (workstation ATX)	14 in (356 mm)	16.75 in (425 mm)

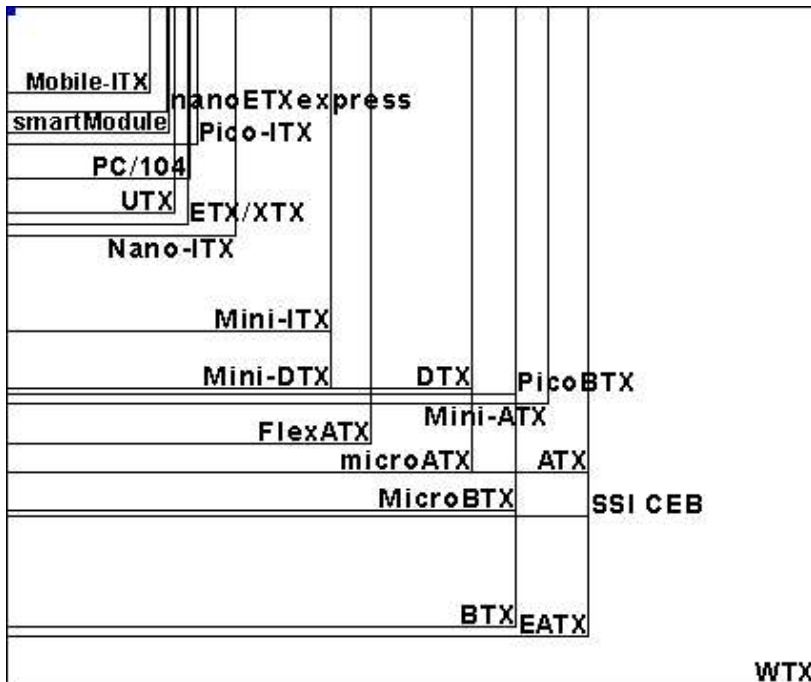


Fig 5.9 Varients of ATX Boards

CHIP SETS

A chipset or chip set refers to a group of integrated circuits, or chips, that are designed to work together. Various types of chip sets are available and they vary from manufacturer to manufacturer. Accordingly there are many types of chip sets which incorporate

different circuits of the motherboard. So depending upon the selection of the chip set the design of the motherboard varies.

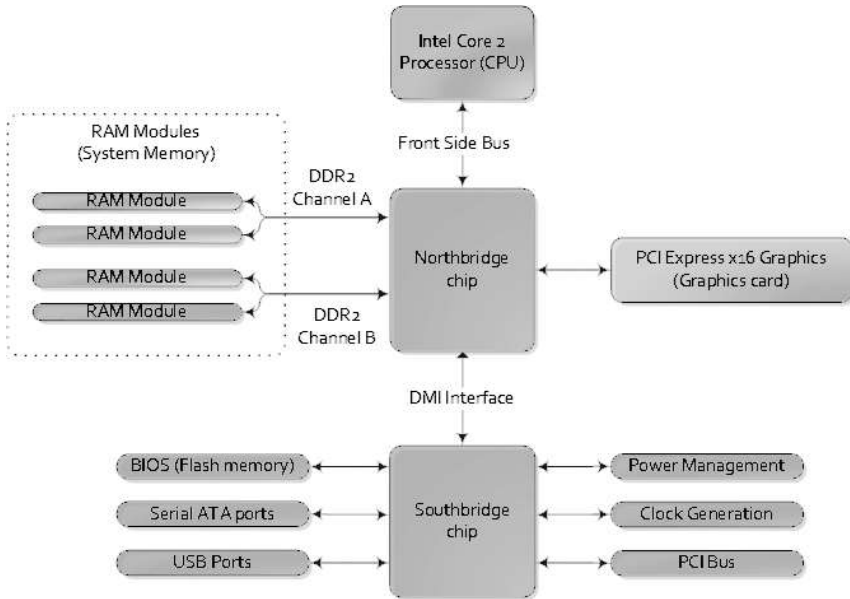


Fig. 5.10 Motherboard Chipset Diagram

Modern motherboard chipsets nearly always consist of two separate chips. These two chips on the motherboard are called the north bridge and the south bridge. Together, the north bridge and the south bridge handle all of the communication between the processor, RAM, video options, PCI slots, BIOS, ATA controller, USB ports, integrated modem, integrated LAN port and integrated sound. The chipset also determines the type of RAM that can be used.

The manufacturer of a chipset often is independent from the manufacturer of the motherboard. Current manufacturers of chipsets for PC-compatible motherboards include NVIDIA, AMD, VIA Technologies, SiS, Intel and Broadcom. Apple computers and Unix workstations from Sun, NeXT, SGI, and others have traditionally used custom-designed chipsets.

The North Bridge

The north bridge controls the communication and throughput speed between AGP video, the RAM and the CPU. The speed at which the memory can communicate with the CPU is referred to as the front-side bus speed or system bus speed. For those who are looking for ultimate speed, the north bridge's front-side bus speed is a key item to study. If you want a fast CPU and fast RAM, you need a fast front-side bus. Modern CPUs and RAM come in a variety of ratings for

front-side bus speeds and must be matched with an appropriate chipset.

For example, RAM can be purchased in a variety of speeds such as DDR200, DDR266, DDR333, or DDR400. Each number next to "DDR" represents a frequency speed rating. DDR266 should only be used with a front-side bus of at least 266MHz, and DDR400 requires at least a front-side bus speed of 400MHz. When the front-side bus speed is the same as the speed of the DDR RAM, this combination is described as being "synchronous." It is the motherboard chipset.

The South Bridge

The south bridge is often referred to as the I/O controller. The features of the south bridge change less often than the those of the north bridge. From chipset to chipset, nearly all south bridges have support for integrated sound, LAN and modem, ATA and USB.

However, just because a chipset can support these features does not mean that the motherboard manufacturer will make that particular feature available. There are also some I/O components that are not controlled by current south bridges, such as RAID and Serial ATA. These have their own chips, separate from current north and south bridges. This will change in the future, as Intel is incorporating Serial ATA into their upcoming Springdale chipset.

I/O BUSES

To provide different types of standardized buses like PC bus, ISA bus, EISA bus, or VESA bus, etc. different types of motherboards are designed which accommodate these particular types of I/O slots physically as all of them have different physical dimensions.

MEMORY MODULE

Depending upon the type of memory module used like DIPs, SIMMs, or SIPs, the design of RAM bank varies accordingly. Further to these variations design of motherboards also vary to provide various other features like floppy disk controller, IDE controller, Serial/Parallel ports, etc.

MOTHERBOARD COMPONENTS

Apart from of the main CPU, many other support chips called peripheral chips reside on the motherboard. The support chips are the programmable chips which support many important dedicated functions on the motherboard. The peripheral chips are designed to work for a particular CPU type.

CLOCK GENERATOR	8284(PC), 82284(PC/AT)
PIT (Programmable Interval Timer)	8253
INTERRUPT CONTROLLER	8259
DMA CONTROLLER	8237

BUS CONTROLLER 8288(PC), 82288(PC/AT)

PPI (Programmable Peripheral Interface) 8255

In the latest motherboards PC/AT, 386 or 486 no discrete chips are used. Instead specially designed CMOS VLSIs which incorporate one or more of the above mentioned circuits.

In addition to the CPU and peripheral chips other components available on the motherboards are

- Discrete registers, transceivers and buffers
- BIOS ROMs
- RAM banks
- Keyboard controllers
- CMOS chips and battery for storing SETUP
- Jumpers and connectors etc.

Many motherboards available today are also having some additional built-in controllers also i.e they have the controllers and circuitry mounted on them. To list a few, following are the different types of controllers integrated into the motherboards these days.

- Video display controller
- Serial/Parallel port controller
- IDE controller
- Floppy Drive Controller
- Game port controller

Further to these motherboards are having I/O slots, jumpers, connectors (power, keyboard, etc.), some resistances and capacitors combinations and crystal oscillators.

THE MICROPROCESSORS

It is the key working part of a PC, to understand a PC we should thoroughly understand a microprocessor operation. PC family is primarily based on the 8086 series of microprocessors. The basic computing power of a PC is decided by the microprocessor installed in it. Main specifications of a microprocessor are

- Data Bus Width
- Address Bus Width
- Speed of the Microprocessor

Various Intel microprocessors popular among the PC's these days are as follows :-

PART NO.	ADDRESS BUS	DATA BUS	MEMORY SIZE
8086	20	16	1M
8088	20	8	1M
80286	24	16	1M

80386(SX)	24	16	16M
80386(DX)	32	32	4G
80486(SX)	32	32	4G
80486(DX)	32	32	4G
PENTIUM	32	32/64	4G

As we have already discussed that the internal operations each microprocessor completes each instruction in three basic steps viz., fetch, decode, & execute.

THE CO-PROCESSOR

The instruction set of a normal processor from 8086 family is not optimized to do complex numerical calculations, CRT graphic manipulations, etc. Therefore specialized co-processors have been developed for these applications. The co-processors operate in parallel to the 8086 family processors with same buses and with same instruction stream. These specialized co-processors perform work that the main processor cannot do well. Various co-processors are available with corresponding processors, like 8087, 80287, 80387 and 80487 etc. The co-processors provide both speed and accuracy to our calculations. Besides 4 basic arithmetical operations (add, subtract, multiply, divide), co-processors also have specialized transcendental operations required for performing logarithmic and trigonometric operations.

One of the difficult things about trying to quantify the effect of adding a numeric coprocessor is that, not only do different applications use numeric coprocessors to different extents, but a single application will vary in its use of a coprocessor depending on the job it is tackling. For example, it is generally held that spreadsheet programs are heavy coprocessor users but it isn't difficult to find particular cases of spreadsheets that make no use of floating point arithmetic at all. For example, a spreadsheet that has been constructed to act as a database is very unlikely to perform any arithmetic calculation so it is not worth installing a coprocessor for such a task.

The sort of calculations that you have to do, to find out what impact a numeric coprocessor will have, is to estimate how much time your program spends in floating point calculation. This is much more difficult than you might expect. For example, suppose you have a spreadsheet consisting of numeric data and formulae and it takes 30 minutes to recalculate. A simple minded approach would suggest that if you add a numeric coprocessor that does floating point calculations five times faster than software the calculation would only take six minutes. If you try this out you will usually be disappointed. The reason is that the 30 minutes to do the calculation is not occupied solely by floating point calculations- data has to be found and results returned to the correct locations. In all probability no more than 75% of the 30 minutes is spent in performing floating point calculations and this is the only portion of the calculation time

that can be reduced by the addition of a coprocessor. In this case the calculation time would go down to only 12 minutes instead of the promised six.

Earlier, various Numeric co-processors were introduced in the market along with the normal CPU. A brief history is best summarised by the table below.

Co-processor	Manufacturer	CPU	Remarks
8087	Intel	8088/8088	designed before any standards
80287	Intel	80286/80386	not as power as 387
80387	Intel	80386	standard 386 coprocessor
80387SX	Intel	80386SX	standard 386SX coprocessor
FasMath83D87	Cyrix	80386	claims 10x faster than compatible 387
IIT-3C87	IIT Inc	80386	claims 2x faster than compatible 387
Abacus	Weitek	80386	Not compatible with 87

Later, 486 and onward, the co-processor was included in the main CPU and there was no need to add the co-processor on the motherboard.

MEMORY

It is the key portion of a PC. It is the area where all the programs and CPU executable codes are loaded. The microprocessor starts execution of instructions written in the memory once it is given the specific address. To standardize the memory organization in a PC so that all the software and routines developed at different platforms should be portable, a memory map has been accepted. This standardized memory organization is called the system memory map.

How Memory is Organised

System designers have reserved for internal usage a certain portion of the maximum amount of memory in each system. This reserved memory is used for things such as graphics and ROM programmers. The remaining memory is called the user memory and is the portion of memory in which a Pentium rating system loads itself and any operating programme.

In today's systems, video memory may be around 8-16 MB portion of the memory reserved for the storage of graphic and text, material of display. This memory is located from memory address A0000 to BFFFF. The lower resolution adapters may use only a portion of the video memory but the higher resolution VGA adapters use it all.

Adapter board ROM and RAM space is allocated for various adapter cards that need a place on their onboard control software. Some network boards use this space of RAM in order to transfer data to and from the system.

Motherboard ROM space is the space for motherboard control programmer that is also called BASIC INPUT-OUTPUT SYSTEM (BIOS). On IBM systems only the requisite BASIC INTERPRETER is also located in this space.

The onboard memory in a PC motherboard is organized in many different fashions using various types of memory modules. Depending upon the access time required for a particular portion of memory, type of the RAM or ROM is selected. Generally the user memory is organized using DRAM banks available in various different packages like DIP, SIMM, DIMM etc. To further enhance the speed of access to the programs, faster memories called CACHE RAMs are also provided which are organized using high speed static RAMs.

BIOS

The **BASIC INPUT-OUTPUT SYSTEM** resides in a set of EPROMs on the motherboard. They store the programs which are called the firmware that is the routines which need not be changed time and again. The ROM memory occupies a space in the system memory map. These chips store the programs that must be activated every time the system is turned on. Because these programs must be available immediately, they cannot be loaded from a secondary device such as disk drive. As a matter of fact, ROM programs instruct the system during boot-up.

ORGANIZATION OF ROM-BIOS

The ROM-BIOS is roughly divided into 3 functional parts. The first part of ROM-BIOS is the start-up routines, which get our computer going when we turn on the power. There are two main parts to the start-up routine -

- Power-on-self-test (POST) routines
- Boot routine

one is the power-on-self-test (POST) routines, which is used to test the computer for its working condition. They check the memory for defects and perform other tests to see that the computer is not malfunctioning.

The initialization involves things like creating the interrupt vectors, so that when interrupt occurs, the computer switches to the proper interrupt handling routine. Initialization also involves setting up the computer's equipment. Many of the parts of the computer need to have registers programmed and parameters configured in order to initialize them. Included in this initialization are steps that tell ROM-BIOS what equipment is present.

Some of that is learned by checking switch setting and jumper configuration inside the computer or reading a permanent memory (CMOS) that records the equipment configuration. This permanently recorded information is called the system setup and can be modified whenever the configuration is changed. The ROM-BIOS during initialization reads this configuration setup and confirms the presence of equipment by checking them for a particular response.

The very last part of the startup routines in the ROM-BIOS is the boot routine, which tries to initialize DOS or any other operating system that we have in our computer. The bootstrap process involves the ROM-BIOS attempting to read a boot record from the beginning of a disk. The BIOS first tries the drive A, and if that does not succeed, and the computer has a hard disk drive C, it tries the hard disk. If neither disk can be read, Then the ROM-BIOS goes into it's non-disk mode and displays a message saying the computer needs a boot disk. Once the ROM-BIOS is able to read the boot record from the disk, it hands over the control of the computer to the short program on the boot record. The boot process begins the process of loading DOS (or any other operating system) into the computer.

The other two parts of the ROM-BIOS form a key part of the running of the computer. These two parts are the hardware handling and the service handling. They function as two distinct but closely cooperating kinds of routines.

The service handling routines are there solely to perform work for our programs (and for DOS) by carrying out whatever services the programs need performed. They are things like a request to clear the display screen , or to switch the screen from the text mode to the graphics mode, or a request to read information from the disk, or write information on the disk. To carry out the service requests that our programs make, the ROM-BIOS has to work directly with the computer's I/O devices, like using ports to issue commands and pass and receive data with the various devices like printers, disks and so forth.

This shields our programs not only from the details of how the printer works, but more important, it shields our programs from the problems of error recovery.

While some of the hardware parts of our computer only require attention when we want them to do something, other parts call for attention that is completely separate from what our programs are doing.

BIOS FEATURES IN ADVANCED COMPUTERS

Apart of the regular services as discussed many advanced features are added into the latest computers these days like the 'System Set Up'.

Earlier setup (CMOS setup) which used to store the configuration of the computer used to be invoked from the floppy after booting. Now a days the system setup feature is available in the BIOS and it is invoked before booting an user command. Generally depending upon the type of the BIOS like AMI, EUROBIOS, etc., a message is prompted indicating the method to invoke the setup. Setup is also divided into the following types :-

Standard Setup

This setup program is used to record configuration of your system using the keyboard. Specifically you can set date, time, no. of floppy drives and their types, no. of hard disks and their types, video type, RAM, etc.

Advance Setup

This setup is also available in BIOS like AMI etc. Here you can set the following features :-

- CPU speed high/low
- Wait Status
- Password Facilities
- Booting Sequence A: C; / C; A:
- Video/ System Shadow ROM's etc.

A window at the bottom of the screen shows the instructions to modify, exit and save the setup configured. This complete information is picked up from the BIOS (ROM) and saved in CMOS RAM.

Other features available in the BIOS setup these days are the auto configuration to BIOS defaults, system self test and the built-in hard disk utilities. BIOS defaults set the configuration of setup parameters according to the factory settings. Self test provides the built-in testing routines for different hardware components inside the system. Hard disk utilities provide various hard disk management facilities like formatting, making bootable, etc.,

HOW DOES THE BIOS WORK

The first part of the BIOS contains the start-up routines, which get the computer going when you turn on the power. There are two main parts to the start-up routines, which test to see that the computer is in good working order. They check the memory for defects and perform other tests to see that the computer is not malfunctioning. The second part of the start-up procedure is the initialization.

The intiallization involves such things as creating the interrupt vectors, so that when interrupts occur the computer switches to proper interrupt handling routine.

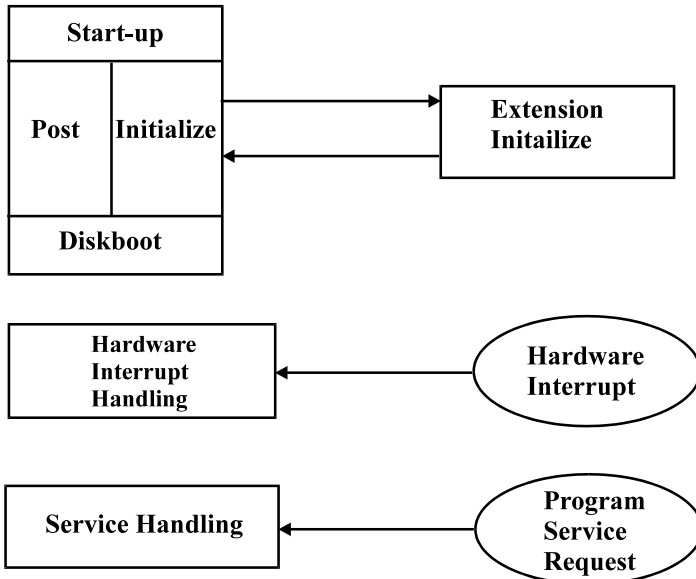


Fig 5.11 The Organisation of BIOS

Initialization also involves setting up the computer's equipment. Many of the parts of the computer need to have registers set, parameters loaded, and other things done to get them ready to go.

THE BASIC INPUT OUTPUT SYSTEM is a set of permanently recorded program routines in a chip which gives its system a fundamental operational characteristic.

CMOS SETUP

Following are the few screens of BIOS Setup of a typical Motherboard.

1. Turn on or reboot the system. The following message appears at the bottom of the screen:
2. Press the key to enter the CMOS setup program and the following screen appears :
3. Choose an option and press <Enter>. Modify the system parameters to reflect the options installed in the system. (See the following sections for more information.)
4. Press <ESC> at anytime to return to the Main Menu.
5. In the Main Menu, Choose "SAVE AND EXIT SETUP " or <F10> to save your changes and reboot the system. Choosing "EXIT WITHOUT SAVING " or <ESC> ignores your changes and exits the program.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.

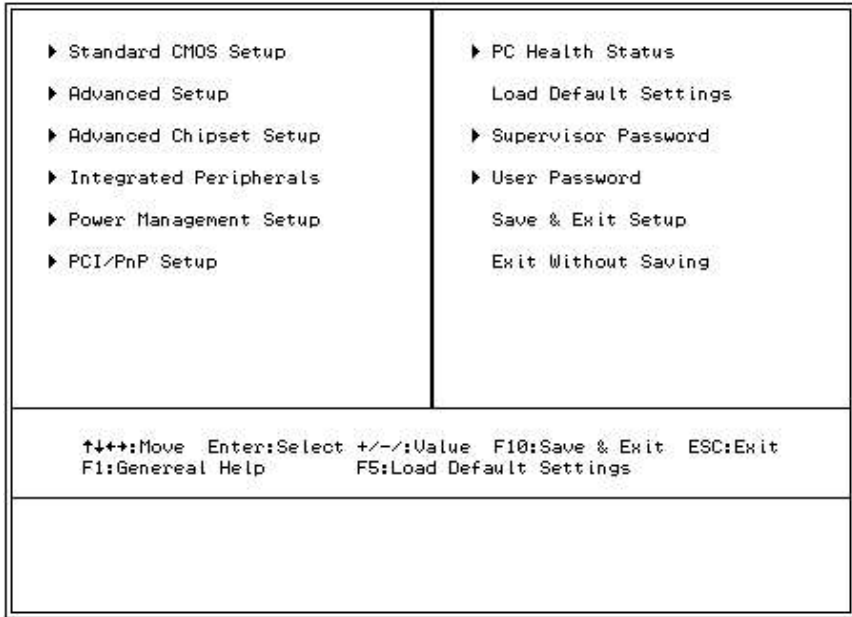


Fig 5.12 CMOS Setup Utility - Main Screen

Standard CMOS Setup

Run the Standard CMOS Setup as follows.

1. Choose "STANDARD CMOS SETUP" from the Main Menu and a screen with a list of items appears.
2. Use the arrow keys to move between items and to select values. Modify the selected fields using the PgUp/PgDn/+/- keys. Some fields let you enter numeric values directly.
 1. Date (mm/dd/yy) Type the current date
 2. Time (hh:mm:ss) Type the current time
3. Primary master and slave along with SATA drive options, which BIOS Auto detect. You can Press <Enter> to access submenu.
5. IDE BusMaster - Bus Mastering is when devices are allowed to talk, to each other directly without having to be controlled or use the processor at all, thus reducing processor occupancy. When this BIOS feature is enabled, the BIOS loads up the 16-bit busmastering driver for the onboard IDE controller. This allows the IDE controller to transfer data via DMA, resulting in greatly improved transfer rates and lower CPU utilization in real mode DOS and during the loading of other operating systems.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
Standard CMOS Setup

Date	Wed 01/09/2008	Help Item
Time	14:35:27	
▶ Primary IDE Master	Not Detected	
▶ Primary IDE Slave	ATAPI CDROM	
▶ S-ATA 1	Hard Disk	
▶ S-ATA 2	Not Detected	
▶ S-ATA 3	Not Detected	
▶ S-ATA 4	Not Detected	
IDE BusMaster	Enabled	
Processor Type	Intel(R) Pentium(R) D CPU 2.80GHz	
Processor Speed	2.80GHZ/800MHz	
CPU ID	0F44h	
Processor Cache	2048 KB	
System BIOS Version	Rio3U311	
Base Memory	640KB	
Extended Memory	767MB	
Memory Channel	Single Channel	
Memory Speed	533MHz	
On Chip UGA BIOS Version	008.0501.000.000	

↑↓←→:Move Enter:Select +/=:Value F10:Save & Exit ESC:Exit
F1:General Help F5:Load Default Settings

Fig 5.13 Standard CMOS Setup

When this BIOS feature is disabled, the BIOS will not load up the 16-bit busmastering driver for the onboard IDE controller. The IDE controller will then transfer data via PIO.

3. After you have finished with the Standard CMOS Setup program, press the <ESC> key to return to the Main Menu.

BIOS Advance Setup Features

This option is used for selecting boot options. Run the BIOS Features Setup as follows.

1. Choose “Advance Setup” from the Main Menu and a screen with a list of items appears.

Limit CPUID MaxVal It Allows you to determine whether to limit CPUID maximum value. Set this item to DISABLED for Windows XP operating system. Set this item to ENABLED for legacy operating systems such as Windows NT4.0.

Intel XD Bit Execute Disable Bit (XD-Bit) is a system feature that, if present and enabled, allows the notebook’s processor to distinguish between bits of code that should be executed and the ones that cannot be executed because they pose a threat to the system.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
Advanced Setup

		Help Item
Full Screen Logo show	Enabled	
Thermal Management	Enabled	
Limit CPUID MaxVal	Disabled	
Intel XD Bit	Disabled	
Quick Power on Self Test	Enabled	
Boot Up Numlock Status	On	
APIC Mode	Enabled	
1st Boot Device	LITE-ON COMBO SOHC	
2nd Boot Device	ST3200822AS	
▶ Hard Disk Drives	Press Enter	
▶ CD/DVD Drives	Press Enter	
Boot Other Device	Yes	
BIOS Write Protect	Disabled	

↑↓←→:Move Enter:Select +/-:Value F10:Save & Exit ESC:Exit
F1:General Help F5:Load Default Settings

Fig 5.14 Advance Setup Features

Quick Power On Self Test / Quick Boot

Enabling this setting will cause the BIOS power-on self test routine to skip some of its tests during bootup.

Boot Up Num Lock Status

This setting, when enabled, automatically turns on your NumLock key when the system is booted. Most systems default this to enabled.

APIC Mode

Enabling APIC mode will expand available IRQs resources for the system.

1st Boot Device

Select your required boot device to be the first boot device. For example if you want to boot from the floppy disk, select floppy drive to be the first boot device. Likewise, if you want to boot up from a Windows CD, select CD-ROM to be the first boot device instead.

Chipset Features Setup

The “Advance Chipset Setup” is used to control the values of the chipset registers. It deals with extremely low-level chipset functions. These functions control most of the system options in the computer.

Run the Advance Chipset Setup as follows :-

1. Choose “Advance Chipset Setup” from the Main Menu and a screen with a list of items appears.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
Advanced Chipset Setup

CAS Latency Time	By SPD	Help Item
Auto Share Memory Control	Enabled	
UGA Share Memory	256 MB	
Video Display Devices	Auto	
TV Standard	NTSC	
Expansion Mode	Disabled	
HPET in SB	Disable	

↑↓→←:Move Enter:Select +/-:Value F10:Save & Exit ESC:Exit
F1:General Help F5:Load Default Settings

Fig 5.15 Chipset Features Setup

CAS Latency Time	This controls the time delay (in clock cycles - CLKs) that passes before the SDRAM starts to carry out a read command after receiving it.
Auto Share Memory Control	Automatically shares memory for VGA device.
Video Display Devices Select	the Video Display devices from the available devices.
TV Standard	Select the mode for standard TV display from the listed options NTSC, PAL, PAL-M, PAL-60, NTSC-JAP, PAL-CN, PAL-N, SCART_RGB
Expansion Mode	Select the options from Enabled/ Disabled.

Integrated Peripherals

Allows you to configure, enable, or disable onboard ports. Select the various option from the screen shown below :

USB Controller:

This item enables the USB controller. Leave this at the default "Enabled" if you want to connect USB devices to your computer.

USB 2.0 Supports:

Enable this item if your system supports USB 2.0.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
Integrated Peripherals

		Help Item
USB Controller	Enabled	
USB 2.0 Controller	Enabled	
Legacy USB Support	Enabled	
Onboard AUDIO Function	Auto	
Onboard LAN Function	Enabled	
Onboard LAN Boot ROM	Disabled	

↑↓→←:Move Enter:Select +/-:Value F10:Save & Exit ESC:Exit
F1:General Help F5:Load Default Settings

Fig 5.16 Integrated Peripherals Setup

USB Legacy Support

This BIOS feature determines if the BIOS should provide legacy support for USB devices like the keyboard, mouse and floppy drive. This is a useful feature when using such USB devices with operating systems that do not natively support USB (e.g. Microsoft DOS or Windows NT).

When enabled, the BIOS will enable legacy support for USB keyboards, mice and floppy drives. You will be able to use these USB devices even with operating systems that do not support USB.

When disabled, the BIOS will not enable legacy support for USB keyboards, mice and floppy drives. You will not be able to use these USB devices with operating systems that do not support USB.

Onboard Audio function Enable or disable the onboard Audio options.

Onboard Lan Function Enables or disables the onboard LAN function.

Power Management Setup

The Power Management controls the mainboard “green” features that for the power saving Mode, Display turn off and HDD power down that together form the hardware power conservation scheme.

Run the Power Management Setup as follows :

1. Choose “Power Management Setup“ from the Main Menu and a screen with a list of items appears.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
Power Management Setup

ACPI Suspend Type	S3	Help Item
Run UGABIOS on S3 Resume	No	
Soft-off by PWR-BTTN	Delay 4 Sec	
PWRON After PWR-Fail	Power Off	
Resume On LAN	Disabled	
Wake-Up by PME	Enabled	
Power On by Ring	Disabled	
USB KB Mouse Wake Up from S3	Enabled	
PS2 Keyboard Wakeup	Enabled	
PS2 Mouse Wakeup	Enabled	
Resume on RTC Alarm	Enabled	
RTC Alarm Date (Days)	Every Day	
Time Alarm (hh:mm:ss)	12:30:30	

↑↓←→:Move Enter:Select +/-/:Value F10:Save & Exit ESC:Exit
F1:General Help F5:Load Default Settings

Fig 5.17 Power Management Setup

ACPI Suspend Type Here you can select the ACPI (Advanced Configuration and Power Management) suspend type that will be used. For S1 (Power On Suspend), power is still given to the system and the data is kept on the HDD and RAM but the processor is halted, whereas, in case of S3 (Suspend To RAM), the system/application state is saved to RAM (Random Access Memory) and power is cut to most components.

Soft-Off by PWR-BTN	This allows you to select the delay between when you press the power button to when the system turns off. If you set it to "3", then you will have to hold the button for three seconds. If you set it to "Instant-Off", then your system will power down as soon as you press the button.
Wake Up by PME	This allows you to select whether or not you would like your system to wake up from power saving mode when a PME event occurs.
Power On by Ring	If this setting is enabled, an incoming call will awaken the system from power saving mode.
USB wakeup from S3	If enabled, it will wake up the system in ACPI S3 using USB keyboard or Mouse.
Resume by Alarm	If this setting is enabled, you can select the exact time and date for when your system will power on.

PCI/PnP (Plug n Play) Setup

The "PCI/PnP Setup" is used to control the values of the chipset registers. It deals with extremely low-level chipset functions. These functions control most of the system options in the computer.

Choose "Advance Chipset Setup" from the Main Menu and a screen with a list of items appears.-

Resources Controlled By	If you find that you cannot get a particular expansion card to work properly, you might be able to solve the problem by changing this item to Manual and manually configuring the card. If you change this item to Manual, a series of items will appear that allow you to define the assignments of the system interrupt lines (IRQS) and Direct Memory Access (DMA) channels. By default, these items are set to PCIISA PnP.
Assign IRQ ForVGA	To assign an IRQ to VGA card if you enable this item.
IRQ Resources	Interrupts are in some ways the most "famous" system resources. An interrupt is basically what it sounds like, a message from one part of the computer to another (normally to the system processor) that tells it that it needs to stop what it is doing, and do something else instead.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
 PCI/PnP Setup

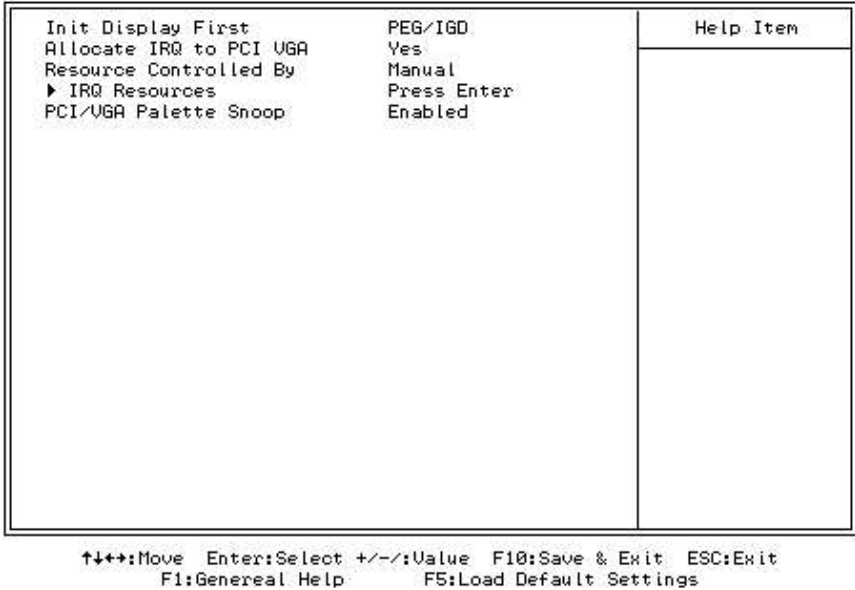


Fig 5.18 PCI/PnP Setup

Load Setup Defaults

This Main Menu item loads the default system values. These settings are recommended for optimum performance. If the CMOS is corrupted, when enter BIOS setup utility you must load setup default again. Choose this item and the following message appears:

“Load SETUP Defaults (Y/N) ? N”

To use the Setup defaults, change the prompt to “Y” and press <Enter>.

User Password Setting

This Main Menu item lets you configure the system so that a password is required every time the system boots or an attempt is made to enter the Setup program. The password cannot be longer than 8 characters.

Important : Keep a safe record of the new password. If you forget or lose the password, the only way to access the system is to discharge CMOS memory.

Save & Exit Setup

Select this item from the main menu and Type “Y” to save the values entered during the current session and then exit the BIOS Setup program. Type “N” to return to the Setup program.

CMOS Setup Utility - Copyright (C) 1985-2007, American Megatrends, Inc.
Supervisor Password

Supervisor Password : Not Installed	Help Item
Change Supervisor Password Press Enter	

↑↓←→:Move Enter:Select +/-:Value F10:Save & Exit ESC:Exit
F1:General Help F5:Load Default Settings

Fig 5.19 Password Setting

Exit Without Saving

Select this item from the main menu and type “Y” to exit the BIOS Setup program without saving the values entered during the current session. Type “N” to return to the Setup program.

EXPANSION SLOTS

As we know that in a basic computer the motherboard is having important components like CPU, RAM, ROM (BIOS), Keyboard Controllers, Support Circuitry etc. Now to connect any peripheral i.e., video display, printers, plotters etc. We need to connect corresponding adapters to the system i.e., the display adapters, I/O port cards, FDC's, and HDC's etc. These adapters are connected to the motherboard through the so called expansion slots or I/O slots. Now to maintain the compatibilities between various types of IBM compatible adapters manufactured by different manufacturers, these I/O slots are standardized in their designs. The first I/O slot standardized was PC BUS. An expansion slot consists of :-

- Address Bus
- Data Bus
- Control Bus

Then to utilize the upgraded versions of processors like 80286 80386 etc. New expansion slots were designed like ISA, EISA, VESA etc. As the newer versions of buses were developed many parameters and signals were included to accommodate the increasing need of

- Address Lines
- Data Lines
- Control Lines
- Bus Arbitration Signals
- Bus Transfer Rate

PROCESSOR BUS

The processor bus is the communication pathway between the CPU and immediate support chips. This bus is used to transfer data between the CPU and the main system bus.

The bus consists of electrical circuits for data, for addresses (the address bus, which is discussed in the following section), and for control purposes. In a 486-based system, for example, the processor bus consists of 32 address lines, 32 data lines, and a few lines for control purposes.

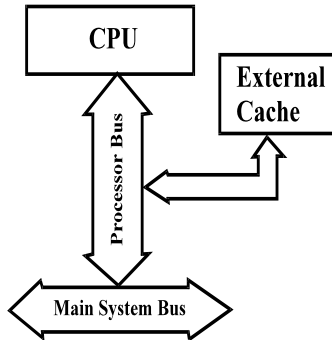


Fig 5.20 Processor Bus

Pentium's processor bus has 64 data lines, and associated control lines. The processor bus operates at the same base clock rate as the CPU and can transfer one bit of data per data line every one or two clock cycles. Thus, a 486- based system can transfer 32 bits of data at a time, whereas a Pentium can transfer 64 bits of data.

Common Pathways between CPU & immediate support chips transfers data between CPU & main system bus. e.g.. Data transfer between CPU and external cache. Data transfer rate is defined = Data width x clock speed of bus. e.g.. Pentium MB has 66MHz speed, Data transfer rate = 66MHz x 64 bits = 4,224 Mega bits/second = 4,224 / 8 = 528 Mega byte / second.

Transfer rate is same as bandwidth of bus. $66 \text{ MHz} \times 64 \text{ bits} = 4,224 \text{ megabits/second}$. $4,224 \text{ megabits/second} / 8 = 528 \text{ M/second}$ This transfer rate, often called the bandwidth of the bus.

THE MEMORY BUS

The memory bus is used to transfer information between the CPU and main memory - the RAM in your system. This bus is either of processor bus itself, or it is implemented by dedicated chipset that is responsible for transferring information between the processor bus and the memory bus.

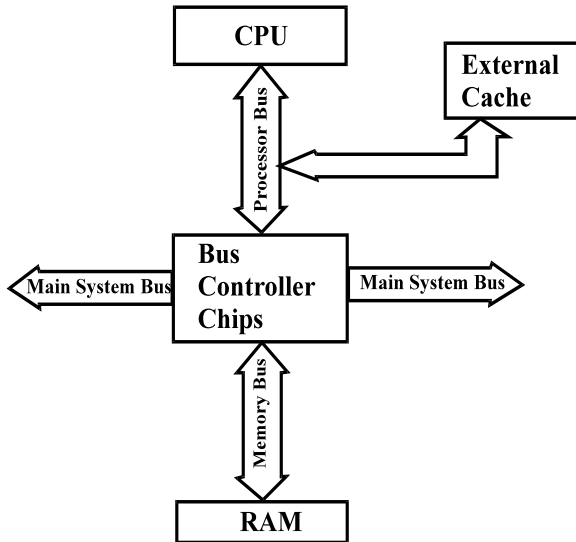


Fig 5.21 Memory Bus

Systems that run at motherboard clock speeds of 16 MHz or faster cycle at rates that exceed the capabilities of standard Dynamic RAM chips. Transfer information between CPU and main memory through bus controller chips, The information that travels over the memory bus is transferred as much slower rate than the information on the processor bus. The chip sockets or the slots for memory SIMMs are connected to the I/O bus.

DATA BUS

The data bus carries the information to or from the memory. The four main types of data bus used in PCs are PCI, AGP, VESA Local, and ISA. Of these four, only the PCI bus and the AGP bus are widely used today. The VESA Local Bus and ISA bus are rarely found today in new computers, as both standards have been officially deprecated as viable hardware interfaces for high-speed PCs.

ADDRESS BUS

The address bus actually is subset of the processor and memory buses. The address bus indicates what address in memory or what address on the system bus are used in a data transfer operation. The address bus indicates precisely where the next bus transfer or memory transfer will occur. The size of the memory bus also controls the amount of memory that the CPU can address directly. Subset of processor & memory bus indicates that address in memory is used for data transfer operation. Size of memory controls the amount of memory that the CPU can address.

CONTROL BUS

It carries control signal among various units of a computers.

TYPES OF I/O BUS

I/O buses are necessary for better system performance. This need for higher performance involves three main areas:

- Faster CPU's
- Increasing software demands
- Great video requirements

The main types of I/O architecture are:

- ISA
- Micro Channel Architecture (MCA)
- EISA
- VESA Local Bus (VL-Bus)
- PCI Local Bus
- PC-Card (formerly PCMCIA)

ISA BUS

ISA is acronym for Industry Standard Architecture, is bus architecture that was introduced with original IBM PC in 1981 and later expanded with the IBM PC/AT in 1984. ISA is a basis of the modern personal computer and the primary architecture is used in the vast majority of PC systems on the market today. It may seem amazing that such an antiquated architecture is used in today's high-performance systems, but ISA is a basis of the modern personal computer and the primary architecture is used in the vast majority of PC systems on the market today. It may seem amazing that such an antiquated architecture is used in today's high-performance systems, but this seems to be true for reasons of reliability, affordable and compatibility. Two versions of the ISA bus exist, based on the number of data bits that can be transferred on the bus at a time. The older version is an 8-bit bus, the newer version is a 16-bit bus. The original 8-bit version ran at

4.77 MHz in the PC and XT. The 16-bit version used in the AT ran at MHz.

The 8 Bit ISA Bus

The bus architecture is used in the original IBM PC computers. An adapter card with 62 gold contacts on its bottom edge plugs into a slot on the mother board that has 62 matching gold contacts. Electronically, this slot provides 8 DATA lines and 20 ADDRESS lines.

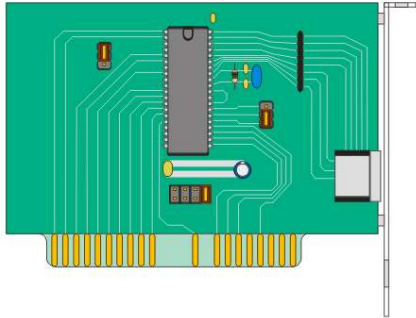


Fig 5.22 8 bit ISA Card

The 16 bit ISA Bus

The second generation x286 chip can handle 16 bits on the I/O bus at a time, compared with 8 bits in older CPU's. The second part of each expansion slot adds 36 connector pins to carry the extra signals necessary to implement the wider path. The extended 16 bit slots physically interfere with some adapter cards placing it in line with the older 62 line connector to provide some extra features.

- Eight more data lines
- Four more address lines
- Four more DMA channels
- Eight more IRQ levels

THE EISA BUS

The EISA bus provides 32 bit slots for use with 386DX or higher systems. The EISA bus adds 90 new connections with out increasing the physical length of the connector. It is just like the 8 bit ISA connector but with two rows of connectors. The EISA bus can handle 32 bit data @ 8. 33MHz cycle rate.

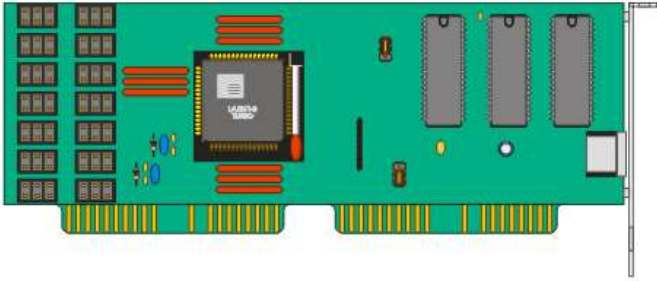


Fig 5.23 16 bit ISA Card

Other improvements include the introduction of shareable interrupt lines and an efficient bus mastering system. The shareable interrupt lines have been achieved by changing existing interrupt lines to be level rather than edge sensitive, as in the case of the ISA bus. A level sensitive interrupt allows the processor to scan the add-on boards to discover which of them needs attention.

The only disadvantage of this scheme is that you cannot mix ISA cards and EISA cards on one interrupt line. In other words only EISA cards can share. The bus mastering facilities are particularly powerful. A bus master can gain control of the bus very quickly and can then transfer data directly between other devices. The EISA hardware will take care of any problems of a 32 bit bus master trying to talk to a 16 bit device. In addition the EISA hardware will arbitrage between multiple bus masters to make sure that the most important get a fair share of the bus's time.

In theory the EISA bus can transfer data at 32 MByte/s (i. e. 32 bits or 4 Bytes x 8 MHz) in short bursts but 25 MByte/s is a more typical sustained transfer rate.

To make life easier for the user, each EISA card has a unique product identifier code that the processor can read to determine which card is where. This makes configuring EISA software and hardware easier than configuring ISA software and hardware.

Features of EISA Bus

- 32 bit data path
- Enough address lines for 4GB memory
- More I/O address up to 64 K
- 8. 33 MHz clock rate
- Auto Configuration
- Bus Mastering

THE MCA BUS

Perhaps the most important fact about the MCA bus is that it is completely incompatible with the ISA bus. This is both its advantage and its disadvantage. Not having to be compatible with any earlier design it was possible to start again and produce something simple and fast. The price for this is simply that the existing range of ISA boards cannot be used in an MCA based system.

MCA was introduced by IBM as a replacement for the ISA bus in its 286 and 386 based PS/2 machines. It is more complex than the EISA bus because it is implemented as a 16 bit and as a separate 32 bit bus. In other words there are two types of MCA bus. The 16 bit version is used in 286 and 386SX PS/2 models because there is no point in using a 32 bit expansion bus with a 16 bit processor! However even in 386 based PS/2 designs a mixture of 16 bit and 32 bit expansion connectors are used. The 16 bit MCA connectors are used for lower performance devices and the 32 bit MCA connectors are used primarily for memory and other high performance devices.

There is an additional, but minor, complication in that there is also the possibility of a video extension being added to any single MCA bus slot. The video extension allows any extra video card to override the internal VGA and make use of the machine's video output hardware. The video extension makes it possible to extend the display modes offered by the internal video hardware without having to replace it.

Only one of the MCA expansion slots in a machine has the video extension because the internal hardware can only be overridden by one expansion adaptor. None of the MCA buses can achieve the same result by adding an extra connector to the video card itself.

Thus inside an MCA machine you can find as many as three distinct types of expansion slot. The 32 bit MCA has many of the same specifications as the EISA bus. It is a full 32 bit data bus with 32 bit address lines. The 16 bit version is similar to the ISA bus. That is, it has a 16 bit data bus and a 24 bit address bus. The main technical difference between MCA and EISA is that MCA is an asynchronous bus whereas EISA is a synchronous bus. A synchronous bus works at a fixed clock rate no matter how fast a device is capable of transferring data but an asynchronous bus doesn't transfer data at a rate that depends on a fixed clock. Instead an asynchronous bus takes its timings from the devices involved in the data transfer. When the receiving device has the data it signals back to the sending device for more data. Clearly an asynchronous device has to be more complex than a synchronous one but it is potentially more efficient.

The original MCA specification resulted in a maximum transfer rate of 20 MByte/s which is of course slower than the EISA maximum of 32 MByte/s. However IBM have recently introduced a new method of data transfer using the existing MCA bus hardware that increases the data rate to 160 MByte/s. They have also introduced additional

modes to allow other devices to take over the MCA bus more effectively and efficiently.

These new developments emphasize the way IBM is making use of MCA to keep one step ahead of the competition without much regard to the actual needs of the user. Indeed there are many parts of the MCA bus that are 'undefined' or 'reserved for future use' and this stops many manufacturers from producing MCA designs for fear that IBM will simply change the rules and make the existing MCA machines look out of date. Other improvements include a change in the way that interrupts are handled. The ISA bus uses edge triggered interrupts but the MCA bus level triggered interrupts. This allows systems that use a large number of interrupts to be more efficient and more reliable.

MCA cards also have a unique identifying number which can be used by software - POS Programmable Option Select - to simplify installation.

Choosing the bus - ISA v EISA v MCA

At the end of this chapter it seems fitting to deliver judgement on the three buses. The issue facing any prospective 386/486 buyer is which type of bus to look for in their machine. Of course there is the small matter of cost to consider as well as performance. Obviously the aim is to find the lowest cost machine that will do the job well and continue to do that job for a reasonable number of years - typically five. The first thing to restate is that, although the bus issue is hotly contended by the participants and the technical press, the issue for most users is not critical. The reason, is twofold: only the most demanding applications depend on the bus for their success and many critical add-ons are now integrated into the main board and don't make use of the expansion bus. This situation may change if the nature of applications that personal computers are used for and characteristics of add-ons change.

The lowest cost bus to buy and run is currently the ISA bus. It has been around a long time and looks set to be around for a long while yet. There are more ISA add-on cards than either MCA or EISA cards. EISA is marginally easier to design and so in principle it should be cheaper than MCA but market forces have as much influence on actual costs as manufacturing costs. MCA adaptor cards are much smaller than either PC bus or ISA/EISA bus cards and this can cause designers a problem in crushing all the electronics needed into the space provided.

The MCA design was produced with the idea that users would use high density surface mounting components to reduce the space needed but this is an expensive constructional method for small runs or prototypes.

As far as performance goes MCA and EISA are more or less equal. The point is that they are in the same class when it comes to transfer rates and sophistication and any slight difference will have little impact on

observable performance. In other words once you decide that you need something more than ISA the choice between EISA and MCA will be based on something other than performance. The EISA, bus has one single powerful advantage over MCA - it is backward compatible with ISA. This would be enough for a clear cut recommendation in its favor if it wasn't for the fact that MCA is backed by IBM and at the time of writing there are more MCA machines and add-on boards than there are EISA machines and boards. However this situation is changing rapidly and manufacturers, other than IBM, are announcing cut backs or the abandonment of their MCA designs and are moving towards EISA. Also surveys seem to suggest that of the installed user base of MCA machines the majority were bought because they were IBM, rather than MCA, machines.

VESA LOCAL BUS

The I/O buses ISA, MCA, EISA have one thing in common that they are very slow i.e. @ 8.33 MHz. As the speed of the processor bus increased, the I/O bus realized only nominal speed improvements, primarily from an increase in the bandwidth of the bus. The I/O bus had to remain at the low speed because of the huge installed base of adapter cards. To run the BUS at a much faster speed it is very important to move the BUS closer to CPU.

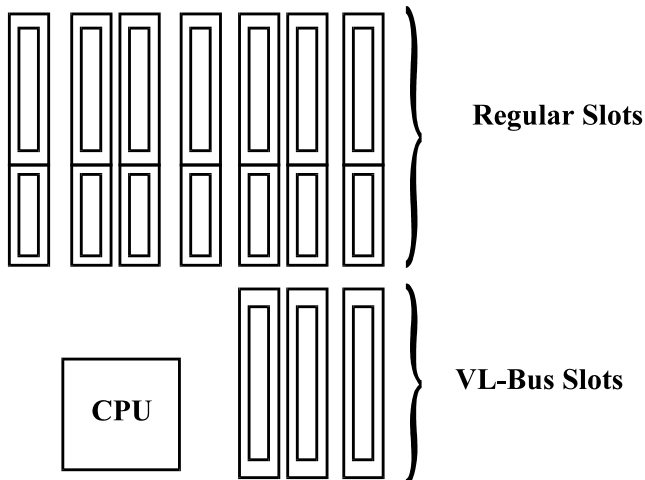


Fig. 5.24 VL BUS

The **VESA (VIDEO ELECTRONIC STANDARDS ASSOCIATION)** developed a standardize local bus or simply VL-BUS. The VL-BUS slot offers direct access to system memory at the speed of the processor itself. The VL-BUS can move 32 bit at a time, enabling the data to flow between the CPU and accompanied video subsystem or hard disk at a full data width. Regular VGA cards on EISA bus sets 600 K pixels/sec while local bus can set them at 50 -60 million pixels/sec. VL bus

is positioned as an extension of 16 bit ISA slot or EISA slot. VL bus extension has 112 contacts & uses MCA type of physical connector.

Drawback of VL Bus

- The VL BUS inherently is tied to the 486 processor bus.
- The VL BUS provides only up to 66 MHz of speed
- Electrical Limitations
- Card limitations

AGP BUS

AGP (Accelerated Graphics Port) was introduced to meet consumer demand for high-resolution 3D graphics in home computers. New software programs (especially games) require more and more video bandwidth for fancy textures, high frame rate animations, etc. While the AGP bus employs 66 MHz clocked PCI specifications, it also has the advantage of allowing large amounts of graphics data to be transferred directly between the computer's main memory and the AGP video card. This feature allows the video card to share system memory on demand. The AGP bus is designed strictly for video processing and does not have to share available bandwidth with other connected devices. Most high-performance video cards are now only available as an AGP version.

PCI BUS

PCI (Peripheral Component Interconnect) remains a strong industry standard for Pentium-class computers. acts as a bridge between MAIN BUS and CPU. It is called MEZZANINE BUS also because it adds another layer to the traditional bus configuration. PCI bypasses the standard I/O bus. It uses the system bus to increase the bus clock speed and take full advantage of the CPU's data path. Information is passed on the PCI bus @33 MHz, at the full data width of the CPU.

$$33\text{MHz} \times 32 \text{ bits} = 1056 \text{ megabits/sec}$$

PCI bus does not directly interface to the CPU. Rather it communicates with the CPU via a BRIDGE CIRCUIT that acts as a buffer between the specifics of a particular CPU and the BUS.

Advantages of PCI Bus

- Processor Independence
- Wider Data Path
- High Speed
- Backward Compatibility
- Bus Mastering
- Software Setup

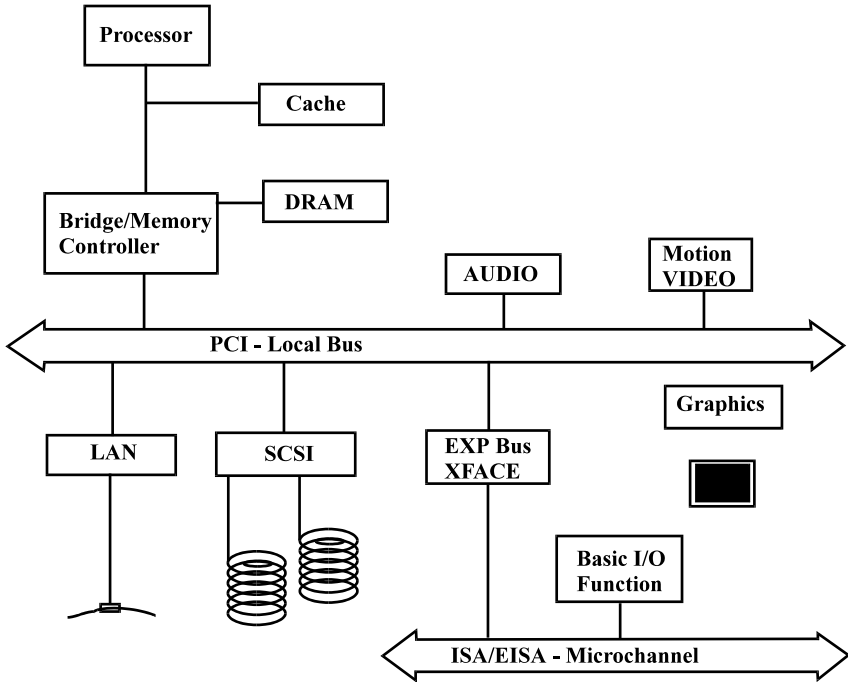


Fig. 5.25 PCI BUS

PCI EXPRESS

Since PCI Express is a serial based technology, data can be sent over the bus in two directions at once. Normal PCI is Parallel, and as such all data goes in one direction.

Each 1x lane in PCI Express can transmit in both directions at once. In the table the first number is the bandwidth in one direction and the second number is the combined bandwidth in both directions. Also please note that in PCI Express bandwidth is not shared the same way as in PCI, so there is less congestion on the bus.

The "x" in an "x16" connection stands for "by." PCIe connections are scalable by one, by two, by four, and so on.

When the computer starts up, PCIe determines which devices are plugged into the motherboard. It then identifies the links between the devices, creating a map of where traffic will go and negotiating the width of each link.

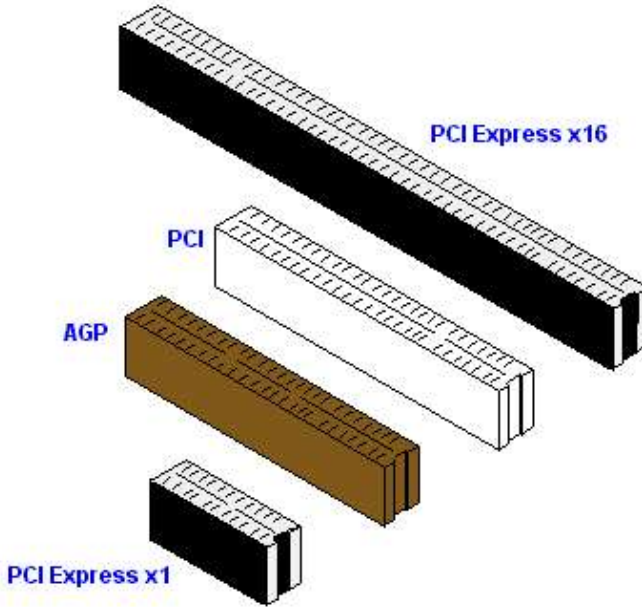


Fig 5.26 PCI Express Bus

The 32-bit PCI bus has a maximum speed of 33 MHz, which allows a maximum of 133 MB of data to pass through the bus per second. Different PCI Express specifications allow different rates of data transfer, anywhere from 400 MB to 6400 MB of data per second and beyond as shown in the table below :

PCI	133 MB/s
AGP 8X	2100 MB/s
PCI Express 1x	400 MB/s
PCI Express 4x	1600 MB/s
PCI Express 8x	3200 MB/s
PCI Express 16x	6400 MB/s
IDE (ATA100)	100 MB/s
IDE (ATA133)	133 Mb/s
SATA	150 MB/s
SATA II	300 MB/s
Firewire	400 Mb/s
USB 1	12 Mb/s
USB 2	480 Mb/s
Gigabit Ethernet	1000 Mb/s.

Table 5.1 Comparative of Expansions Slots

SUPPORT CIRCUITRY

By now we have seen that in a PC motherboard, a part of CPU and allied logic other important circuitries are-

- KEYBOARD INTERFACE : Through which the keyboard interacts with the CPU.
- Providing and interfacing the peripherals I/O, timer, PPI, interrupt controller etc.
- Providing interfacing for memory and I/O expansion boards.

Thus we see that only keyboard is the peripheral which is directly interfaced to the conventional PC motherboard. All the peripherals are interfaced through other adapters on I/O slots. We will list a few important circuitries on the motherboard and would make a detailed study in further sections.

RESET LOGIC

This circuit is responsible for providing the reset in the motherboard during

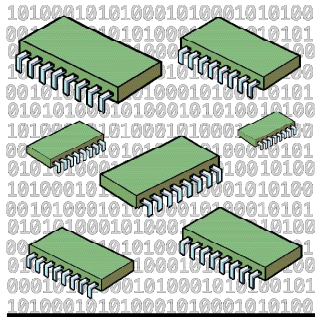
- Power-On (Power-Up reset)
- Pressing reset push button
- Pressing CTRL, ALT and DEL keys together

CPU NUCLEUS LOGIC

The nucleus logic is composed of CPU, clock generator, bus controller, address latch, data bus buffer, and interrupt controller. The CPU interacts with these chips and generates the address, data buses, control signals for memory, I/O operations.

EXERCISE

1. Fill in the Blanks :
 - a) A motherboard holds _____ required for a computer.
 - b) Co-processor was included with the main processor in _____ and later CPU.
 - c) Primary Master and Slave setting are done in _____ of BIOS setting.
 - d) POWER MANAGEMENT is the master control for the four power saving modes namely _____, _____, _____ and _____.
 - f) Hard disk are automatically detected by using the _____ option in BIOS setup.
 - g) The data bus carries the information to or from the _____.
 - h) Data bus could be of four major types namely _____, _____, _____ and _____.
 - i) MEZZANINE BUS is the name of _____ bus.
2. How you define the technical specification of the motherboards?
3. What are the reasons for design variations in Motherboards?
4. Discuss the various components of Motherboard.
5. What is co-processor and how it works?
6. What is a (Flash) BIOS ? How do you determine if your system has a Flash Bios chip ?
7. What is a BIOS? What is Setup?
8. What is CMOS? What is POST?
9. How do I determine which BIOS version I have?
10. Why do you get a “CMOS checksum invalid” error message during POST?
11. Why do you get the message, “Stack overflow, system halted, please adjust stacks in config.sys and try again” when you try to flash your BIOS?
12. How you activate the BIOS setup? What are the various settings used in the BIOS setup?
13. What is function of Swap Floppy drive option?
14. How can you use external Display and Sound adaptor?
15. Discuss various types of ISA bus. How MCA bus are better than ISA bus?
16. What is VESA Local Bus and what are its advantages?
17. What are the advantages of PCI bus?



CHAPTER 6

SECONDARY STORAGE - I

FLOPPY DISK DRIVES
DRIVE CONFIGURATION DEVICES
TYPES OF DISK DRIVE

Secondary Storage - I

FLOPPY DISK DRIVES

THEORY AND OPERATION

The best first step in understanding how a disk drive operates is to actually watch one in operation. The drive has the ability to record concentric circular tracks of information on one or both sides of diskette.

The circular plastic diskette is grasped by a cone shaped clamp and hub which also centres the disk as the drive door is closed. The main drive motor spins the disk inside the jacket at the very precise speed of 360/300 RPM. The record/ play back head is referred to as a read/ write head in a disk drive. It is mounted on a carriage capable of moving in and out on slides along a radius. Because the tracks are at discrete locations, this movement is conveniently driven by a type of motor called a stepper. Rotating the motor consists of pulsing it through one or more steps with the help of some controlling electronics. In a disk drive, the stepper motor drives the head in such a way that one step of the motor causes a head motion of one track.

On a 5.25 inch drive, 40 tracks per side are most often used with 48 Tracks per inch measured along a radius; 80 is an alternative with 90 tracks per inch, while 3.5 inch drives use 80 Tracks at 135 Tracks per inch, depending on the model. A new diskette generally starts life unrecorded. Formatting or initializing a diskette consists of having disk drive write some information at every tracks on the disk, just circles which have been recorded upon. The formatting operation writes markers for each track and section of track. The sections of tracks are called sectors. It is critically important that the head-positioning system be very accurate. In an economy drive, the heads are allowed to press (Load) against the disk surface as soon as the door is closed. In the luxury models, an extra solenoid operates a mechanism which normally holds back the head and only allows it to spring into operating position needed.

DRIVE COMPONENTS

Regardless of floppy drive types, a floppy drive consists of several basic common components which are :

Read Write Heads

Any disk drive today has two Read/Write heads making the floppy drive a double sided drive. The two magnetic heads are mounted across the disk on one carriage (Head carriage)

The surface is designed for minimum disk wear and maximum read out. Each head is constructed with a read/write gap used for data

reading and data writing and to erase gaps to erase the recorded track edge immediately after the recording.

Head Actuator

Head Actuator causes the head to move in and out over the surface of disk. This mechanism is with stepper motor which moves in both the directions. The stepping motor rotates 3.6 degree in 48 TPI and 1.8 degree for 96 TPI models 06 drives.

Stepper Motor

The stepper motor is a rotor type DC brushless motor which has the long life. The normal speed of rotation is 300 or 360 RPM. The speed is maintained at a stable condition against the load variations and environments changes.

Circuit Board

The electronics of the floppy drive is constructed which are read write circuit, control circuit and servo circuit.

Connectors

The two connectors are required for any disk drive -one for power to run the drive and the other to carry the control and data signals. A 4 in line connector is used for power, and 34 pin edge connector is used for data and control signals. Some of 3-1/2 inch drives use a tiny version of power connector.

DRIVE CONFIGURATION DEVICES

To configure a drive in the system, one must install it correctly considering the following items -

- Drive select jumper
- Terminating Resistor

Jumpers are easy to set, e.g. take care of are drive select jumpers, which determine which number the drive thinks it is. These are labelled DS0 to DS3. Install one jumper according to whether this drive will be 0, 1, 2 or 3 (Corresponding to A, B, C or D on most system)

The disk drive ribbon cable performs a peculiar twist between Drive 0 and Drive 1.

The object of this twist is to connect the two drives select signals from the controller to the same pin on each of the two drives connector. This is done so that both drive in this kind of PC can be jumpered as drive 1 at the factory. Next we have jumpers labelled HS and HM. HS stand for Head Solenoid, and installing the jumper causes the head to load when the drive is selected. HM stands for Head/Motor, and causes the head to load when Motor ON signal is active. Finally there is a jumper labelled MX. This stands for multiplex. Installing this

jumper causes that drive to always be active. Don't install this jumper.

TERMINATOR

Some of the signals on the drive cable are open collector signals. This means that controller only has the capability to either do nothing to this wire, or to pull it firmly to ground. It is expected that, at the drive, this line will be pulled up with a resistor, so that when controller is doing nothing, the line will be high.

The problem is that we need one pull up resistor for each signal, and we might have upto four drives on the cable. For this reason, all the lines which require a resistor are brought to an IC socket, where you can plug in a resistor pack. A resistor pack looks like an IC but, of course, it is full of resistors. This is the terminator. As with ICs, pin 1 of the resistor pack is marked with a 1, a dot or a notch, and should be oriented to agree with a similar marking on the socket, and to be in the same orientation as ICs on the same board.

It is necessary that there is one and only one terminator installed between all the drives, and it should be in the drive which is at the furthest point on the ribbon cable from the controller.

DOS REQUIREMENT

A brand new diskette can't be used until it has been formatted. Formatting places magnetic marker on the disk surface to define the sectors in which data is stored. When DOS encounters a disk that doesn't have these markers, it issues an error message like:

**General Failure error reading drive A
Abort, Retry, Fail?**

Formatting also creates a main directory and file allocation tables by which DOS keeps track of the placement of files. A newly formatted disk is ready to hold any kind of files including program files and data files. For the disk to make it bootable, it must have certain DOS system files transferred to it during formatting.

BASIC FORMATTING

To format a diskette in drive A simply enter :

FORMAT A: <enter>

The format command is carried out by a part of DOS that is kept in its own file FORMAT.COM when you want DOS system files to make the diskette bootable simply enter

FORMAT A:/ S <enter>

For 360 KB diskette the command will be

FORMAT A:/4/ S

DRIVE CABLING

Disk drives are typically connected to their adapter cards that carry data and control signals between the disk drive and the basic system. Floppy drives use a single ribbon cable to connect the system's floppy drives to the disk drive controller card. Generally, the cable has two edge connectors along its length that connect directly to PCB of drive. The other end of the cable has 34 Pin, two row BERG Header. The indicator stripe that marks the side of the cable which should be aligned with Pin 1 of the adapter's interface connector and the disk drive's signal connector. The floppy drive connected to the edge connector at the end of the cable will be drive A: and the floppy drive connected to the edge connector towards the middle of the cable will be designated as the drive B:. The indicator stripe on the signal cable must be oriented towards the edge connector pin no. 1. The location of this pin is marked on the drive's PCB.

TYPES OF DISK DRIVE

5-1/4 inch Drives

- 360 KB Double sided Double Density (**DSDD**)
- 1.2 MB Double sided High Density (**DSHD**)

3-1/2 inch Drives

- 1.44 MB Double sided Double Density (**DSDD**)
- 2.88 MB Double sided High Density (**DSHD**)

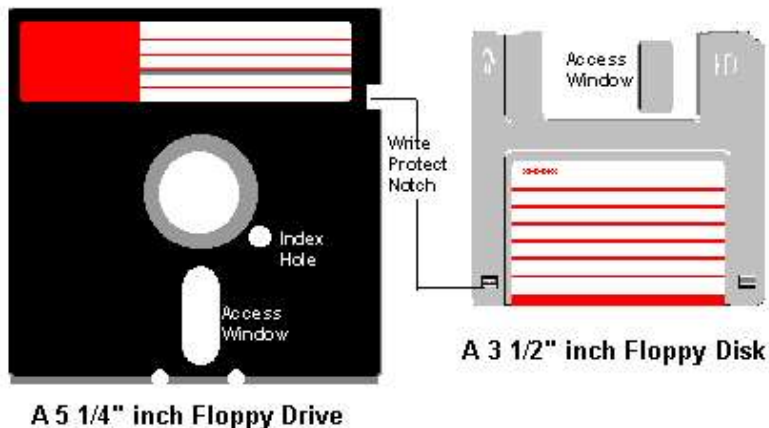


Fig 6.1 Floppy Disk

The following is the drive characteristics of different disk drive

Drive	Size	Tracks	Encoding	Rotation rate	Data Transfer Rate
360 KB	5 1/4"	40	MFM	300 RPM	250 K bits/sec

1.2 MB	5 ¹ / ₄ "	80	MFM	360 RPM	500 K bits/sec.
1.44MB	3 ¹ / ₂ "	80	MFM	300 RPM	250 K bits/sec.
2.88MB	3 ¹ / ₂ "	80	MFM	300 RPM	500 K bits/sec.

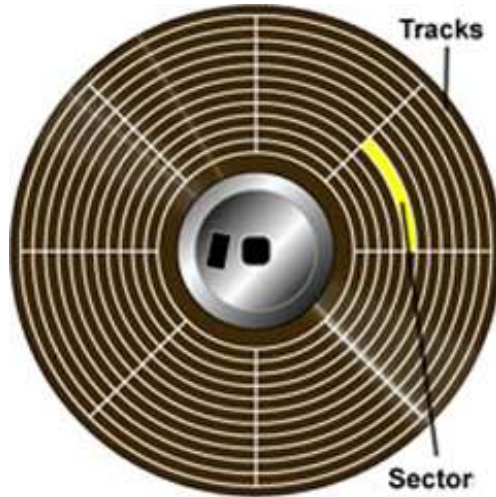


Fig 6.2 Tracks & Sectors

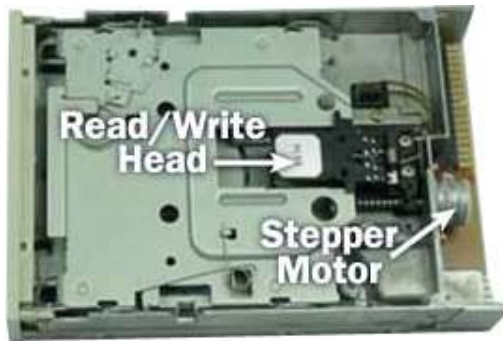


Fig 6.3 Floppy Drive

FLOPPY DISK CONTROLLER

The floppy disk controller in older PC used to be a separate PCB known as FDC. In the majority of PC's FDC was a part of multi I/O board or a separate board which is called SUPER IDE. For the last two years this super IDE card is integrated on the mother board itself thus saving a expansion slot.

The FDC is linked to the system bus of the PC and appears as a set of I/O ports to the CPU. In addition, it is connected to the Channel 2 of the DMA controller. The FDC interrupts the CPU at level 6. The FDC usually performs data transfer in DMA mode.

The FDC in the PC is designed around an intelligent programmable floppy disk controller IC. The original IBM design used NEC 765 IC. Intel's 8272, SMC's FDC 765A and Rockwell's R6765 are the other pin compatible ICs used by various clones. The term FDC IC used in this chapter represents any of these ICs. Most of the functions of the FDC are done by the FDC IC. Some additional functions are done by hardware circuits external to the FDC IC. The distribution of functions between the FDC IC and the external hardware are listed below:

FDC IC Functions

- (a) Generating MFM format for the data to be recorded
- (b) Decoding and executing commands, such as seek, read, format, etc.
- (c) Error detection by CRCC generation/checking
- (d) Data synchronizations

External Hardware Function

- (a) Selection of FDD
- (b) Switching-on motor in the FDD
- (c) Resetting the FDC IC
- (d) Enabling /disabling interrupt and DMA signals of FDC IC
- (e) Data separation logic
- (f) Write pre-composition logic
- (g) Line drivers for signals to the FDD
- (h) Line receivers for signals from the FDD

The FDC IC has provisions to perform device selection and motor-on control by appropriate commands. However the PC design does not make use of this feature. Instead, an external output port is used for performing these functions under software control.

INPUT/OUTPUT PORTS

The FDC has three I/O ports in the external hardware.

Data Port

This port is used by the software for three different purposes:

1. While issuing a command to the FDC IC, command and command parameter bytes are issued to the FDC IC through this port. The FDC IC stores the different parameters and the command in its internal registers.
2. After a command is executed, the FDC IC stores a set of status parameters in the internal registers. These are read by the CPU software through this port. The different status bytes are presented by the FDC IC in a specific sequence.
3. In the programmed mode and interrupt mode of data transfer, the data port is used for transferring data between the FDC IC and the CPU by IN or OUT instruction.

Main Status Register (MSR)

This port is used by the software to read the overall status information regarding the FDC IC and the FDDs. Before initiating a floppy disk operation, the software reads this port to confirm the readiness condition of the FDC and the disk drives and to verify the status of the previously initiated command/operation. The different bits of this register are given below:

Bit 0	--	FDD 0	:Busy in seek mode
Bit 1	--	FDD 1	:Busy in seek mode
Bit 2	--	FDD 2	:Busy in seek mode
Bit3	--	FDD 3	:Busy in seek mode
Bit4	--	FDC Busy	:Read/Write command in progress
Bit5	--	Non-DMA mode	
Bit6	--	DOI	: Indicates the direction data transfer between the FDC IC and the CPU
Bit7	--	RQM	: Indicates data register is ready for data transfer.

Digital Control Port

This port is used by the software to control certain FDD and FDC IC functions. The bit assignments of this port are listed below:

Bit 0 and 1:	Device number to selected
Bit 2	: RESET FDC IC
Bit 3	: Enable FDC interrupt and DMA request signals
Bits 4 to 7	: Turn ON the motor in disk drive 0, 1, 2, or 3 respectively.

DISK FORMAT

Each track is divided into a number of sectors. In a PC, the DOS organises 9 sectors (8 sectors in old DOS versions) on a track with 512 bytes of data on each sector. Each sector consists of an ID field and a data field. The field contains the address of the sector. The data field contains the actual data stored in the sector. The ID field is written in each sector of a track while the track is formatted. Formatting a diskette erases any previous data stored in it and writes the sector specific data pattern in the data fields of all the sectors. Thus a new diskette after formatting is, not a blank diskette.

The standard floppy diskette format used by IBM double density recording format consists of pre-index gap, address mark, post index gap, sectors and final gap.

Address marks are unique bit patterns used to identify the beginning of ID and Data fields. Certain bit cells in the address mark bytes do not contain a clock bit. There are four different types of address marks to identify the different types of fields. They are: Index address mark, ID address mark, Data address mark, and Deleted data address mark.

The ID field is seven bytes long consisting ID address mark, Cylinder number, Head number, Record number (Sector number), Sector length, and CRC bytes. The bad track is identified by writing hexa FF in the ID fields of the bad track. The gaps separating the different fields are created on the track when the diskette is formatted. These gaps serve multiple purposes, as listed below:

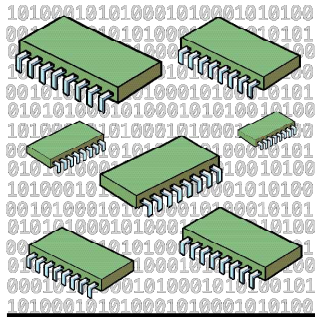
1. Providing Time interval for the FDD circuit to select the required mode: read/write
2. Compensation for variations in diskette rotational speed
3. Compensation for tolerances in diskette and FDD so as to maintain compatibility.

EXERCISE

1. Fill in the blanks
 - a) The formatting operation creates _____ & _____ on a floppy disk.
 - b) The floppy drive connected to the edge connector at the end of the data cable is drive _____.
 - c) 3¹/₂" floppy comes in two capacity _____ & _____.

2. State True or False:
 - a) Floppy drive contains different heads for reading & writing.
 - b) Two connectors are required for any disk drive to perform operation.
 - c) A 34 pin connector is used for power supply.

3. What is the significance of Stepper motor in the Floppy disk?
4. Discuss the various types of floppy drives? What are the major components that floppy drives are made of?
5. What are the function of Floppy disk controller?
6. Differentiate between high density and low density floppy disks.
7. Specify the significance of Drive select jumper & Terminating Resistor.



CHAPTER 7

SECONDARY STORAGE - II

HARD DISK DRIVE
HARD DISK OPERATION
HARD DISK CONTROLLERS AND INTERFACES
HISTORY OF HARD DISK

Secondly Storage - II

HARD DISK DRIVE

Hard disk is a Mass volume storage device. It is capable to store large volume of data in a single drive. A hard disk drive is a sealed unit that holds all data in a system. It is simply a means of storing information. When the hard disk fails, the results are almost never in consequential. To maintain the system properly, the working of hard disk unit should be perfectly understood and implemented.



Fig 7.1 Inside view of Hard Disk

The hard disk gets its name by the way these disks are constructed from rigid platters of metal, usually Aluminium. These platters are inflexible unlike in floppy drives. They cannot be removed from hard disk unit so the drives are also known as “Fixed disk drives”.

Sometimes the drives are also called “Winchester disk drives”. This term dates back to the 1960’s. During that time, IBM developed a 30 MB hard disk. This drive came to be known as 30- 30 drive and soon gained the nick name of “Winchester” after the famous Winchester 30-30 rifle. The term has no real technological or actual meaning and is now simply as a direct synonym for hard disk.

You may wonder how we are able to store this large volume of data in a single unit ? Why it is not possible in the case of Floppy disk drive. Answer for this question is you know very well that Floppy disk either consists of 80 tracks or 40 tracks. Here no. of tracks per disk (flexible disk) is fixed. Moreover, only one disk is present.

But in case of Hard disk no. of Platters (disk) and tracks Per Platter is also increased. By this way we are able to store large volume of data in a Hard disk drive.

Anyway we can’t simply increase number of Platter on the drive because if number of platters are increased obviously number of head

will also increase. So while manufacturing the Hard Disk drive we have to compensate by both the ways (i.e. by increase the Platter and by increase the no. of Tracks).

HARD DISK OPERATION

According to the operational point of view it performs Read/Write operation. So it belongs to the I/O peripheral. Most of you are aware of that how they are Recording Songs on your Audio cassettes. The same Magnetic recording is happening in hard disk. Now we will see the Read operation as well as write operation in the Hard disk drive. Where the data is stored in the Hard disk drive? It stores in the Platter.

WRITE OPERATION

Your Hard disk controller gives data in the form of digital. These digital signal is given to Hard disk drive PCB. These PCB consist of digital to Analog convertor. It converts the Analog signal. Then these signal is passed in to the write coil. Then it write operation.

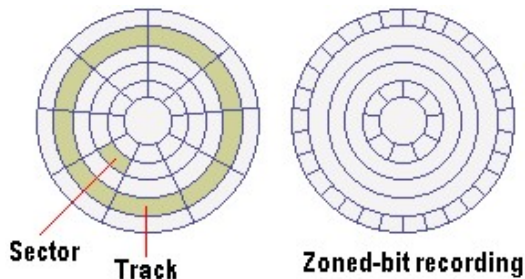


Fig 7.2 Tracks & Sectors

PHYSICAL OPERATION

The physical operation of a hard disk is similar to that of a floppy disk. A hard drive uses spinning disks with heads that move over the disks and store/read data in tracks and sectors. Hard disks usually have multiple platters, each with two sides on which to store data.

Most drives have two or three platters , meaning four to six sides. The same tracks on each side of each platter together makes up a cylinder. There is one head per platter side, and all the heads are mounted on a common carrier device or rack. The heads are moved in and out across the disk in unison and cannot move independently. Hard disks spin at a speed of 5,400 rpm to 7,200 rpm. And in future the speed may go up to 10,000 rpm..

The heads in a hard disk don't touch the platters during normal operation. When the system is powered off, the heads land on platters. While the drive is on, a cushion of air keeps the heads suspended in the air, a short distance above the platter. If this

cushion is disturbed by a particle of dust or shock of some kind the head may come into contact with platter spinning at 3600 rpm. This occurrence is called a head crash and may mean anything from a few lost bytes to a totally trashed disk. Most disks have special lubricants and hardened coatings to withstand much of this type of abuse.

CONSTRUCTION

A hard disk is made up of the following components:

- Disk Platter
- Read/ Write Head
- Head actuator
- Spindle motor
- Logic board
- Connectors

The platters, spindle motors, heads and head actuator mechanisms are usually contained in a sealed chamber called the Head disk assembly (HDA). The HDA is treated as a single component. The other parts are external to the drive, but built on the drive unit.

DISK PLATTERS

Hard disk stores data by magnetically saturated areas on the surface of disk platter. A typical hard disk has one or more platters that can be 5 1/4" (Golden Oldies) or 3 1/2" in diameter much like floppy disks. Each platter is made of a rigid metal usually aluminium for light weight (Golden Oldies). These platters are 1/8 inch in thickness.

The platters are coated with a magnetically retentive substance or media which is actually responsible for storing information. Two popular media are used for hard disk platters - oxide media and thin film media (Latest).

Oxide media is made of various compounds containing iron oxide as the active ingredient. A magnetic layer is created by a process where the aluminium platter is coated with a paste containing iron oxide particles. Spinning the platter at high speed evenly spreads this media across the disk. The surface is then cured and polished. After this a protective layer is added. Normally, this media is about 30 millionths of an inch thick.

As drive densities increase, the media needs to be thinner and more perfectly formed. Therefore most high capacity drives use plated media rather than the oxide coating.

Thin film media is aptly named because the thickness of the media is much less than can be achieved by the normal coating method. Thin film media also is known as plated or sputtered media. Plated media is manufactured by a process of placing the media material on the

disk with an electroplating material on the disk with an electroplating mechanism, much like chrome-plating.

The aluminium platters are immersed in a chemical bath that coat them with a layer of metallic films. The final layer is about 3 millionths of an inch. The advantage of the plated media is that since it is so hard, the likelihood of surviving contacts with the leads at high speed is much more. Because the thin film media can handle much greater densities with accuracy, they are mainly used these days. The number of platters, the drive has, is limited by drive's physical size.

The function of the read/write head is to create and later detect magnetic domains on the recording media by saturating small areas with a magnetic field. The read/write head doesn't access individual particles on the disk, but instead polarizes a segment of the particles as a group. The smallest group of magnetic particles that a specific storage head is capable of affecting is known as its magnetic domain. Heads in this technology, called MR (magnetoresistive), are capable of the a real densities required to sustain the disk drive industry's continued growth in capacity and performance. The technology is capable of supporting a real densities of at least 3 Gbits/in².

A hard disk drive usually has one read / write head for each platter side. All the heads are connected or "ganged" together on a single movement mechanism. All the heads, therefore move in unison across the platters. The basic design for a read/write head is composed of an iron core with a coil of wire wrapped around it. The iron core has a small chunk removed that serves as a gap for the magnetic field to pass through. It is at this gap that the magnetic field, induced into the core by current flowing through the coil, interacts with particles on the disk. This core/coil design can be used to both read and write magnetic fields on the disk.

Mechanically each head is on an arm that is spring loaded to force the heads into the platters. Each platter is "squeezed" by the two heads above and below it. At rest, the heads sit on the platter, but when the drive is spinning at full speed, the air pressure developed underneath the head lifts it up off of the platter's surface. The gap between the heads and the platter is 10 to 20 millionths of an inch. Since this gap is so small that if any particle of dust or dirt gets into this mechanism, it can cause the heads to oscillate and strike the platters. This occurrences is called a head crash and may damage the platter and possibly the head as well.

Two types of heads are used in the modern hard disk drive.

- Composite ferrite heads.
- Thin film heads.

The composite ferrite heads are the traditional type of magnetic head design. This type of head uses an iron oxide core with electromagnetic coils wrapping it. A magnetic field is produced when the coils are

energized, and a field is also induced when magnetic field is passed near the coils.

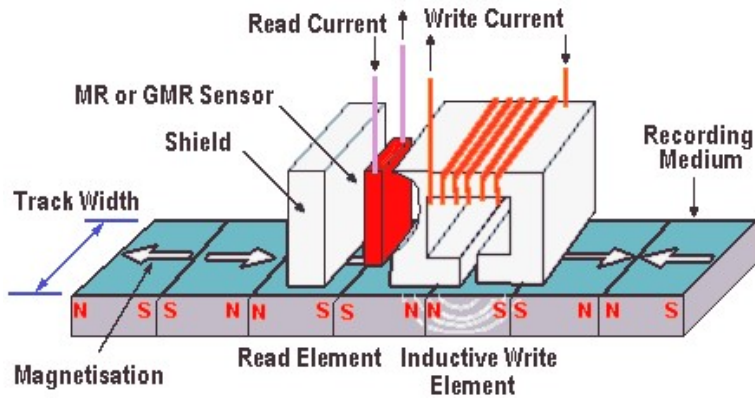


Fig 7.3 Recording on media

This gives the heads full read and write capability. These heads are larger and heavier than the thin film heads and must use a larger flying height to record on the disk. These heads are relatively inexpensive to produce. The film heads are actually a specially produced semiconductor chip. These heads are really a complex circuit.

They are produced by the same procedure as any semiconductor chip, with one exception - the shape here is critical. The heads must have a U shaped groove in the bottom to allow air pressure build up. These heads are light weight and can fly at a much lower than the composite types of heads. This reduced height allows for a much stronger signal to be picked up and this improves accuracy.

GMR technology

Giant Magneto-Resistive (GMR) head technology builds on existing read/write technology found in anisotropic MR, producing heads that exhibit a higher sensitivity to changing magnetisation on the disc and work on spin-dependent electron scattering. The technology is capable of providing the unprecedented data densities and transfer rates necessary to keep up with the advances in processor clock speeds, combining quantum mechanics and precision manufacturing to give areal densities that are expected to reach 10Gbits/in² and 40Gbits/in² by the years 2001 and 2004 respectively.

Current GMR hard disks have storage densities of 4.1Gbit/in², although experimental GMR heads are already working at densities of 10Gbit/in². These heads have a sensor thickness of 0.04 microns, and IBM claims that halving the sensor thickness to 0.02 microns - with new sensor designs - will allow possible densities of 40Gbit/in². The advantage of higher recording densities is that disks can be reduced in physical size and power consumption, which in turn

increases data transfer rates. With smaller disks for a given capacity, combined with lighter read/write heads, the spindle speed can be increased further and the mechanical delays caused by necessary head movement can be minimised.

Pixie dust

In the past decade, the data density for magnetic hard disk drives has increased at a phenomenal pace, doubling every 18 months and - since 1997 - doubling every year. The situation had left researchers fearing that hard drive capacities were getting close to their limit.

When magnetic regions on the disk become too small, they cannot retain their magnetic orientations and thus their data - over the typical lifetime of the product. This is called the “super paramagnetic effect”, and has long been predicted to appear when densities reached 20 to 40 billion bits (gigabits) per square inch, not a huge way away from the densities that had been reached by the start of the new millennium.

However, in the summer of 2001, IBM announced a breakthrough in storage technology that could prolong the life of the conventional hard disk drive for the foreseeable future. The key to the breakthrough is a three-atom-thick layer of the element ruthenium, a precious metal similar to platinum, sandwiched between two magnetic layers. That only a few atoms could have such a dramatic impact has resulted in scientists referring to the ruthenium layer informally as “pixie dust”.

Known technically as “antiferromagnetically-coupled (AFC) media”, the new multilayer coating is expected to permit hard disk drives to store 100 billion bits (gigabits) of data per square inch of disk area by 2003 and represents the first fundamental change in disk drive design made to avoid the high-density data decay due to the superparamagnetic effect.

AIR FILTERS

Hard disks have air filters but they are not changeable. Since hard disk is a sealed unit, no air circulations takes place and hence hard disks can run in extremely dirty atmosphere. The filter inside the hard disk unit is designed only to filter the smaller particles of media that are scraped off the platters during head take off and landing or any metal flakes or small particles that may dislodge inside the drive itself. As air circulates through the assembly, either through natural flow caused by the spinning platters or via help of a thin turbine located on the top and/or bottom spindle hub, it passes over the disks, out of the disk/head chamber, through the absolute filter and then back over the disks. Constant positive pressure (air being pushed through the filter as opposed to being vacuum drawn) must be maintained to insure that the contaminants remain lodged in the filter.

In addition to keeping debris off the disks, proper air flow through the head assembly also has the important function of regulating the temperature. A 5 1/4" platter spinning in air can generate up to 20 Watts of power. This becomes particularly limiting with multi-platter drives, which generate even higher power output. High temperatures increase the amount of stress placed on all drive parts, particularly integrated circuits and the read/write heads. Circulating the air not only allows the air molecules to make contact with the aluminum case (thereby radiating a portion of the heat outward), but more importantly unifies the temperature throughout the assembly and thereby prevents hot spots from developing.

HEAD ACTUATOR

The mechanical system that moves the heads is known as the head actuator. This mechanism moves the heads across the disk and positions them accurately over the desired cylinder. There are usually two types of head actuator mechanisms: Stepper Motor actuators and Voice coil actuators.

Stepper motor actuators

A stepper motor is an electrical motor that can step or move from position to position with mechanical detents. A stepper motor cannot position itself between step positions; instead it must stop at one of the predetermined positions. These motors are physically small and about 1 to 3 inches and can be square, round or cylindrical. These motors usually are outside the sealed HDA, although the spindle of the motor penetrates the HDA through a sealed hole. Stepper motors can be found in one of the four corners of a HDD.

The stepper motor is mechanically linked to the head rack by either a split steel band coiled around the motor spindle or a rack and pinion gear mechanism. As the motor steps, each detent or click stop position translates to the movement of one track through the mechanical linkage. Stepper motor mechanisms are affected by a variety of problems. The biggest problem is one of temperature. As the platters heat and cool, they expand and contract, which causes the tracks to move in relation to some predetermined track position.

The stepper mechanism cannot move in increments of less than a single track to correct for these temperature induced errors. The most widely used stepper motor actuator system transmits the rotatory stepping motion to the in and out motion of the head rack. The band is made of special alloys to limit thermal expansion and contraction as well as stretch.

One end of the band is coiled around the spindle of the stepper motor, and the other is connected directly to the head rack. A split metal transmits the rotatory stepping motion to the in and out motion of the head rack. The band is made of special alloys to limit thermal expansion and contraction as well as stretch. One end of the band is

coiled around the spindle of the stepper motor, and the other is connected directly to the head rack.

Voice coil actuators

A voice coil actuator works with pure electromagnetic force. The actual construction of this mechanism is much like that of a typical audio speaker, which uses a stationary magnet surrounded by a “voice coil “ that is connected to the speaker’s paper cone. Energizing the coil causes the coil to move, which produces sound from speaker cone. In a typical hard disk voice coil system, an electromagnetic coil moves on a track through a stationary magnet.

No contact occurs between the coil and the magnet other than magnetic interaction. The coil mechanism is connected directly to the head rack. As the electromagnetic coils are energized, they attract or repulse the magnet and cause the head rack to move. Such systems are quick and efficient, and most of these systems are much quieter than a stepper motor driven system. A special advantage of using a voice coil positions is automatic head parking.

Two main types of voice coil positions are available : linear voice coil actuator and rotatory voice coil actuator. The linear actuator move the heads in out over the platters in a straight line. The rotatory voice coil systems employ the same magnets and coils but now the coil is spun around spindle and is surrounding the stationary magnets. This rotator motion is transmitted directly to the heads.

SPINDLE MOTORS

The motor that actually spins the platters is called the spindle motor because it is connected to the spindle around which all the platters revolve. These motors are directly connected, no belts or gears are used. The motors must be noise free or they rumble to the platters disrupt reading and writing operations. The motors must be precisely controlled for speeds.

The platters on a hard disk revolve at 5400 to 10000 rpm and motor must have a control circuit with a feed back loop to monitor and control this speed precisely. The spindle motor is on the bottom of the drive just outside the sealed HDA. These motors consume about 12 volt power. The initial starting up requires at least 2 to 3 times of normal power for about 15 seconds until the drive platters have reached a stable rpm.

LOGIC BOARDS

Logic boards contain the electronics that control the drives spindle and head actuator systems, some hard disk drives even have the controller built directly on the drive. Logic boards are removable being simply mounted with screws to the hard disk drive. A couple of items must be set to configure a hard disk drive for installation in a particular system.

These configurations items are mounted on the disk drive logic board, which usually is on the bottom of the drive. The drive select jumper must be set or configured before installing the Hard Disk.

HARD DISK FEATURES

Speed & Performance

In a disk pack with 10 plates there are 20 surfaces. Out of 20 surfaces, the top most surface and the bottom most surface are not used for recording. Thus there are 18 surfaces on which information is recorded. A cylinder thus consists of 18 tracks in such a disk pack. A track is divided into sectors. Read and write operations on a disk start at sector boundaries. Typically 512 bytes are stored per sector. There are 50 sectors per track, 400 tracks per surface and 18 surfaces.

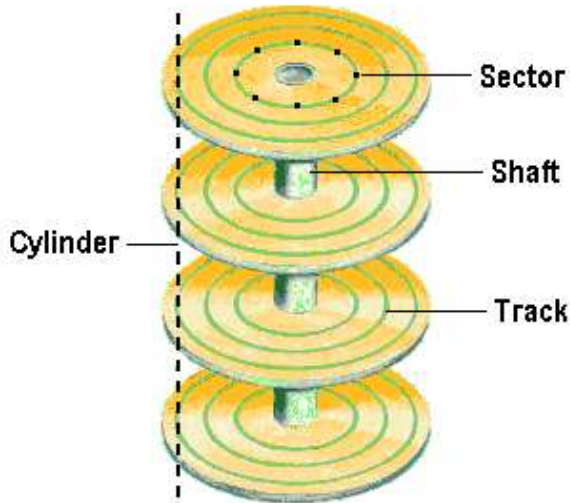


Fig 7.4 Tracks & Sectors

The total capacity of disk pack is -

$$\begin{aligned}
 & \mathbf{512 \times 50 \times 400 \times 18} && = \mathbf{18,43,20,000 \text{ bytes}} \\
 \text{or} & \mathbf{512 \times 50 \times 400 \times 18 \times 8} && = \mathbf{1,47,45,60,000 \text{ bits}} \\
 & && \text{(Because 1 byte = 8 bits)}
 \end{aligned}$$

The density of a recording on a tracks is of the order of

$$\mathbf{(50 \times 512 \times 8) / (12 \times \pi) = 5500 \text{ bits inch}}$$

for a disk pack of diameter 12 inches.

There's a read/write head for each side of each platter, mounted on arms which can move them towards the central spindle or towards the edge. The arms are moved by the head actuator, which contains a

voice-coil - an electromagnetic coil that can move a magnet very rapidly. Loudspeaker cones are vibrated using a similar mechanism.

The heads are designed to touch the platters when the disk stops spinning - that is, when the drive is powered off. During the spin-down period, the airflow diminishes until it stops completely, when the head lands gently on the platter surface - to a dedicated spot called the landing zone (LZ). The LZ is dedicated to providing a parking spot for the read/write heads, and never contains data. The time taken to reach the specified cylinder from the landing zone is known as the **seek time (T)**. The seek time varies depending upon the position of the arm assembly when the read-write command is received by the controller. It is normally measured in millisecond. The maximum seek time is the time taken by the head assembly to reach the innermost cylinder from the outermost cylinder or vice versa. The average seek time is usually specified on the disk and it is of the order of 30 milliseconds.

Once the head assembly is positioned on the specified cylinder, the head corresponding to the specified surface is switched. After selecting the head, there is a further delay because the specified sector has to reach the read-write head. This rotation delay is variable. The average rotational delay equals half the time taken by the disk to rotate once. This time is known as **latency time (T_L)**. E.g. for a disk rotating at 4800 r.p.m.

$$T_L = 0.5 * 60 / 4800 = 6.25 \text{ milliseconds}$$

The sum of average latency and seek time is known as the **average access time**.

Now after reaching the exact location of data, data is read by the head. The rate at which information is read from the disk is known as the **transfer rate**. In calculating the transfer rate, the speed of the rotation of disk plays an important role. For a disk rotating at 4800 rpm, one full revolution take 1/4800 minute. In one revolution, one track, which has say 50 sectors, is read. (One sector contained 512 bytes). Thus, the number of bytes read per revolution is thus

$$\begin{aligned} &= 50 \times 512 \\ &= 25600 \text{ bytes} \\ &= 25600 \times 8 \text{ bits.} \\ &= 25600 \times 4800 / 60 \text{ bytes per} \\ &\quad \text{second.} \\ &= 2048000 \text{ bytes per sec.} \\ &= 2000 \text{ MB per sec.} \end{aligned}$$

Thus **transfer rate**

Time to read one sector

$$\begin{aligned} &= 512 / 2000 * 1024 \text{ sec} \\ &= 0.25 \text{ millisecond} \end{aligned}$$

Thus **Average Access time**

$$\begin{aligned} &= 30 + 6.25 \text{ millisecond.} \\ &= 36.25 \text{ millisecond} \end{aligned}$$

Consider the following example:

A 7-disk pack has 400 track per surface. Out of the available surface, two (2) surface are not recordable. There are 10 sectors per track and each sector can store 512 bytes. If the disk rpm is 4800 and average seek time is 22 ms, calculate the following:

- a) Storage Capacity
- b) No. of cylinders in the disk pack
- c) No. of tracks per cylinder
- d) Transfer rate
- e) Average Access Time

Total no. of surface = No. of disk x 2 - unrecordable disk
 = 7 x 2 - 2 = 12

No. of track on the disk = 12 x 400
 = 4800

Total sectors on the disk = Total tracks x sectors/track
 = 4800 x 10 = 48000

Storage Capacity = Total sectors x sector size
 = 48000 x 512
 = 24576000 bytes = 24000 MB

No. of cylinders = No. of tracks per surface
 = 400 cylinders

Tracks/ Cylinder = No. of surface = 12

Transfer rate = No. of bytes per revolution x rpm/60
 = Bytes per sector x no. of sector x rpm/60

= 512 x 10 x 4800/60
 = 409600 bytes per sec

Average Access Time = Average Seek time + Average Latency Time

= 22 ms + Sector size/transfer rate
 = 22 ms + 512 x 1000 /409600 ms
 = 22 ms + 1.25 ms
 = 23.25 ms

Reliability

Reliability is measured in terms of Mean time Between Failures (MTBF). No measurable difference in actual reliability exists between drives based on the figures given by manufactures which can range from 20,000 hours to 50,000 hours.

HARD DISK CONTROLLERS AND INTERFACES

Selection of Hard disk controllers depends upon

- compatibility with the Hard disk drive or system you already have
- speed of the controller

The primary job of the Hard disk controller is to transmit and receive data and control signals to and from the drive. All hard disks can transmit and receive data just as fast as they spin, which means that as the drive is turning, raw data pulses are simply transferred to the controller, which then must send these signals to the motherboard of the system.

SCSI (SMALL COMPUTER SYSTEM INTERFACE)

SCSI is not a simple Interface. It is a stand alone, external bus capable of transmitting parallel data. The SCSI bus treats devices connected to the bus as logical devices. It communicates with these devices through a defined Protocol, designed specifically to eliminate hardware in-compatibilities.

It is a high end controller system, where the units are connected to a special (expensive) controller. The hard disks are generally of highest quality, fast and with a long life span. An ordinary SCSI controller can run 7 hard disks each of 45 GB - or more. The latest IBM Ultrastar 72ZX holds 72 GB and a cache of 16 MB.

Each device contains its own intelligent controller to control its mechanism that results in higher performance from each drive. In SCSI drive controller and disk controller are present in the drive itself. This way it improves the performance. This configuration allows all the control functions including motor control and data encoding, to be closely coupled with the drive's electronics, thereby decreasing transmission time and overhead.

The SCSI protocol allows devices to disconnect/ reconnect, which enables devices to queue up commands, disconnect from the bus (freeing it for other devices), and then reconnect later after completing the commands. If a single drive and host initiator are the only devices connected to the bus, this disconnect / reconnect sequence helps little. However, when multiple devices are connected to the bus it is very useful. The host can then utilize these disconnect wait times to issue separate commands to other devices connected to the same bus. The most powerful aspect of the SCSI bus for small computers is the expandability.

IDE & ATA

This is simple, primitive interface. IDE (Integrated Drive Electronics) and ATA (Advance Technology Attachment) are one or the same thing - basically a disk drive implementation designed to integrate the controller on the drive itself - much unlike SCSI. Data are delivered on the ISA bus, resulting in low transfer speed. Disks < 528 MB. Is not used on new PCs.

IDE, which stands for Intelligent Drive Electronics or Integrated Drive Electronics is also known as the PC/AT interface, is a system level interface that connects the host computer to a hard drive via a 40-pin connector located on the host motherboard. The host CPU communicates with the intelligent drive using standard hard drive software routines embedded in the AT's ROM BIOS.

The IDE cable configuration is actually a direct bus interface that uses a 4-pin subset of the AT's 98 possible lines. DMA Request / Acknowledge lines, Interrupt Requests Lines, Memory Strobe lines, clocks and power lines have all been eliminated.

EIDE

This interface is an improvement of IDE. Data are delivered on the PCI bus with Bus Master control, which results in high transfer speed. Large disks (currently up to 40 GB) are used in such an interface. Such interface provide room for four units, which are typically connected directly to the system board. The disk of this interface are inexpensive and effective. Ultra DMA is the best interface among the EIDE varieties.

PIO

PIO(Program Input/Output) is a method of moving data between devices in a computer, in which all the information must pass through the processor. The ATA/IDE standard specifies three PIO transfer rates

Mode 0	3.3 Mbps
Mode 1	5.2 Mbps
Mode 2	8.3 Mbps

The new ATA-2 standard specifies two higher data transfer rates -

Mode 3	11.1 Mbps
Mode 4	16.6 Mbps

Direct Memory Access

DMA is the ability of an I/O subsystem to transfer data to and from a memory subsystem without processor intervention.

Ultra ATA/33

Ultra ATA/33 or Ultra DMA/33 is a protocol for transferring data between a hard disk drive through the system bus to the RAM. The Ultra DMA/33 protocol transfers data in burst mode at a rate of 33.3 Mbps, twice as fast as the previous DMA interface.

Ultra ATA/66

Most of the hard disk today are ATA compliant. These are also backward compatible with older PIO and Ultra/33 drives. The requirements for achieving Ultra ATA/66 performance from these new drives is an ATA/66 capable motherboard or IDE controller card and an 80-conductor IDE cable instead of usual 40 Pin. Ultra ATA/66 (Fast ATA-2) allows for a maximum burst transfer rate of 66.6 Mbps, double the maximum of Ultra ATA/33.

Fibre Channel

Fibre Channel is a high speed interface designed for multiple disk drive storage system. Fibre Channel drives can be installed or removed while the host system is operational. This interface can provide a maximum bandwidth of 100 Mbps and can support 126 devices.

Serial ATA

Serial ATA or Serial IDE is all set to replace the parallel IDE interface used currently. The new standard is a point to point protocol that connects each drive directly to IDE controller and will provide a bandwidth of 150 Mbps.

DRIVE CONFIGURATION

For configuring a Hard disk drive, the drive select jumper need to be set. Jumpers are small metallic pins that protrude from the end of a hard disc drive or from its circuit board, as shown below. Jumpers are used to configure the hard disc drive's options, such as defining master and slave or cable select.

Every pin set has a specific numbering scheme, which is determined by its location on the drive. As shown above left, if the pins are on the circuit board, the pin number assignments are printed on the circuit board around the pins. As shown in the Fig, if the jumper pins are on the end of the hard drive, Pin 1 will always be closest to the 4-pin power connector.

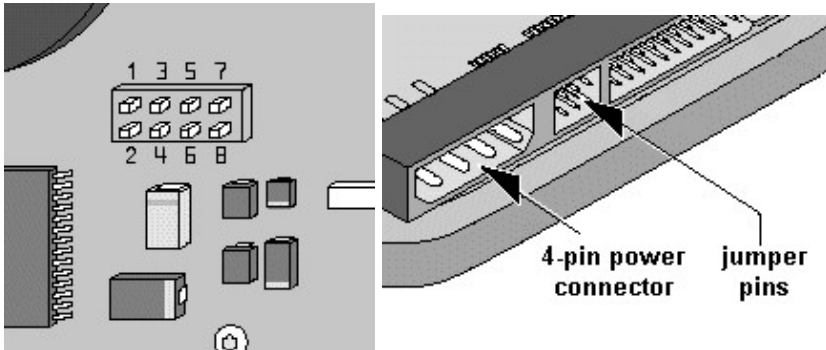


Fig 7.6 Jumpers on a Hard Disk

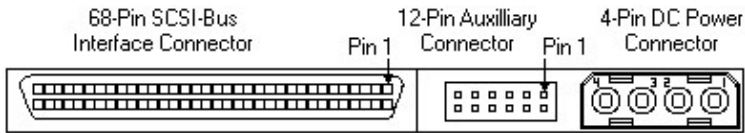


Fig 7.7 Pin 1 in Jumpers

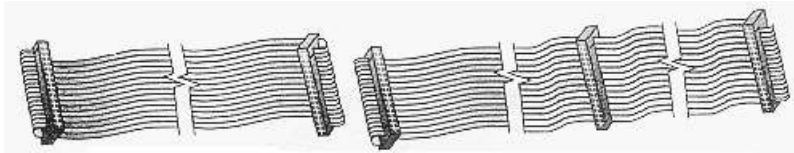


Fig 7.8 (a) IDE Drive Cable Fig 7.8 (b) Cable for two IDE drives

If you have only one drive in the system, you make the drive as master. If there are two drives, you can make your drive as master or slave. The jumper setting is supplied with the hard disk. The various options present are

- Drive is Slave
- Drive is master
- Drive is master with slave present
- Cable Select

HISTORY OF HARD DISK

IBM introduced the first hard disk in 1957, when data usually was stored on tapes. The first 305 RAMAC (Random Access Method of Accounting and Control) consisted of 50 platters, 24 inch diameter, with a total capacity of 5 MB, a huge storage medium for its time. It cost \$35,000 annually in leasing fees (IBM would not sell it outright) and was twice the size of a refrigerator.

The RAMAC was used for calculating the results at the Olympic games in 1960. In 1962 the disks were made removable introducing the

diskpack. In 1964 the CRC algorithm was in use giving a higher degree of data security on the disks.

The first model to use “float on air” technology for the read/write heads was named Winchester 3030 of 1973. Later, the disk platters shrunk to 14" and 8" diameter. They were installed in towers containing dozens of these magnetic platters.

In 1979 industry veteran Al Shugart - already well-known as the former CEO of floppy drive pioneer Shugart Associates - got together with Finis Conner to co-found a new company, Shugart Technology, which was soon to become Seagate. This was the beginning of storage as we know it. Shugart was responsible for the genesis of the two hard drive interfaces that still dominate the PC world today, SCSI and ST-506/412, and in 1980 produced the world’s first 5.25 inch hard drive: the 5MB ST-506. The 10MB ST-412 followed in 1981, and with its selection by IBM for the PC-XT, Seagate were on their way to becoming the world’s largest drive manufacturer.

Since then they have improved a lot. The modern hard disks are 3.5" diameter.

Till now, so many models have been launched that it is not possible to define all the hard drives of all the companies launched till today. For getting details of a particular model, you may visit the website of various hard disk manufacturer - www.westerndigital.com, www.seagate.com, etc. Following is the parameter of few hard drives, which will clearly show, how things moved in the hard disk industry over the period of time :

Seagate ST-412

1981, 4 heads, capacity of 10 MB, Data rate 5 Mbit/sec, Spin rate 3600 RPM, Seek time 85ms, Stepper Actuator, Platter capacity 5MB, Interface MFM, Form 5¼" full-height.

Seagate ST-225

Over the years, the ST-225 was undoubtedly the best-selling hard drive of all time. 1984, 4 thin film heads, capacity of 21.4 MB, Data rate 5 Mbit/sec, Spin rate 3600 RPM, Seek time 65ms, Stepper Actuator, Platter capacity 10.5MB, Interface MFM, Form 5¼" half-height.



Fig 7.9 Seagate ST 225

Western Digital

1986, 4 thin film heads, capacity of 21.4 MB, Data rate 5 Mbit/sec, Spin rate 3600 RPM, Seek time 85ms, Stepper Actuator, Platter capacity 10.7MB, Interface MFM, Form 3.5" half-height..

Seagate ST-251

1988, 6 thin film heads, capacity of 42.8 MB, Data rate 5 Mbit/sec, Spin rate 3600 RPM, Seek time 28ms, Stepper Actuator, Platter capacity 14.3MB, Interface MFM, Form 3.5" half-height..

Seagate ST-157A

6 thin film heads, capacity of 44.4 MB, Data rate 7.5 Mbit/sec, Spin rate 3600 RPM, Seek time 28ms, Stepper Actuator, Platter capacity 14.8MB, Interface IDE mode 0, Form 3.5" half-height.

The best of the IDE steppers, very reliable, and almost as fast as the Maxtor one.

Western Digital Caviar 140/280

4 thin film heads, capacity of 85.3 MB, Data rate 12 Mbit/sec, Spin rate 3395 RPM, Seek time 17ms, Stepper Actuator, Platter capacity 42.6 MB, Interface IDE mode 0, Form 3.5" half-height.

1992 Drive of the Year. These drives, and the ones that were to follow over the next few years, marked the time that Western Digital emerged from amongst the herd of other drive makers, began making a series of great drives, and became famous for their hard drives, rather than just for their old AT controller cards.

Maxtor 7213A

4 thin film heads, capacity of 212.8 MB, Data rate 22.6 Mbit/sec, Spin rate 3551 RPM, Seek time 15ms, Stepper Actuator, Platter capacity 106 MB, Interface IDE mode 2, Form 3.5" 1/3-height.

1993 Drive of the Year. These were the days when Windows 3.1 was new and DOS still reigned supreme, when a large space-hogging application came on no less than four floppy discs and everyone had twin drives: the days of the flip-top case and the 386DX-40 with 4MB of RAM in 30-pin SIMMs, of PC Tools and XTree Gold, of the Trident TVGA-9000 512k ISA video card. There were other 210MB drives; the Western Digital in particular was a worthy rival; the Seagate, the Quantum, and the Conner were all good performers- but in our view the Maxtor 7213 was the best of all.

Seagate Medalist 545xe

4 thin film heads, capacity of 545 MB, Data rate 35.8 Mbit/sec, Spin rate 3811 RPM, Seek time 14ms, Stepper Actuator, Platter capacity 545.5MB, Interface IDE mode 3, Form 3.5" 1/3 height.

1994 Drive of the Year. It was one of the first 500MB drives to sell in significant numbers, and with proven reliability and moderate cost it just kept on going and going. Seagate never needed to replace it. At 3800 RPM and 14ms it was never a speed-king, but always respectable. Cheap, reliable, simple to set up and very compatible, quick enough in its day.

Western Digital Caviar AC2850

4 thin film heads, capacity of 545 MB, Data rate 35.8 Mbit/sec, Spin rate 3811 RPM, Seek time 14ms, Stepper Actuator, Platter capacity 545.5MB, Interface IDE mode 3, Form 3.5" 1/3 height.

The Caviar 850 was an excellent substitute for the 850MB Decathlons when the Seagate drives were out of stock. The Western Digital 850 was not only quick, it proved to be very reliable and many of them were still in service for years to come.

Seagate Decathlon 850

4 thin film heads, capacity of 854.7 MB, Data rate 61.7 Mbit/sec, Spin rate 5376 RPM, Seek time 11ms, Stepper Actuator, Platter capacity 427MB, Interface IDE mode 4, Form 3.5" slimline.

1995 Drive of the Year. One of the all-time great hard drives. The Decathlon 850 was the fastest IDE hard drive in the world for a long, long time. They were very small, very cute, and very, very fast.



Fig 7.10 Caviar AC2850



Fig 7.11 Seagate Decathlon 850

IBM Deskstar 2 (31270/31700)

3/4 MR heads, capacity of 1.27/1.7GB, Data rate 62.1 Mbit/sec, Spin rate 4500 RPM, Seek time 10ms, Stepper Actuator, Platter capacity 850MB, Interface IDE mode 4, Form 3.5" slimline.

1996 Drive of the Year. These magnificent drives were an instructive example of IBM's great strength throughout the Nineties, their leading-edge technology. The four advanced Magneto-Resistive heads read and wrote more data per square inch than any other drive on the

market. In turn, this resulted in the highest data transfer rate around, and at a surprisingly low cost.

Seagate Medalist Pro

4 MR heads, capacity of 2.56 GB, Data rate 116 Mbit/sec, Spin rate 5397 RPM, Seek time 11ms, Stepper Actuator, Platter capacity 1.28 GB, Interface IDE mode 4, Form 3.5" slimline.

These tiny, beautifully finished drives were very, very fast - just as you'd expect from Seagate's slim-line Decathlon family.

IBM Deskstar 3 Family

6 MR heads, capacity of 3.28 GB, Data rate 88.4 Mbit/sec, Spin rate 5400 RPM, Seek time 9.5 ms, Stepper Actuator, Platter capacity 1.05 GB, Interface IDE mode 4, Form 3.5" slimline.

1997 Drive of the Year. These superb drives were a technician's delight and a graphic illustration of the way drive manufacturers in general and IBM Storage in particular were using expensive new technology to make cheaper, faster, more reliable drives. Magneto-Resistive heads were more expensive than the older thin-film heads but they allowed such high data density that only three platters were required to store over 3GB (an outstanding achievement at the time).

Seagate Cheetah Family

Cheetah 1(ST -34501/ST-19101)

8/16 MR heads, capacity of 4.55GB/9.1GB , Data rate 177 Mbit/sec, Spin rate 10,033 RPM, Seek time 7.7 ms, Stepper Actuator, Platter capacity 1.14 GB, Interface Fastwide SCSI, Form 3.5" one third height/ half height.

Cheetah 2 - 1998 (ST -34502/ST-39102/118202)

6/12/24 MR heads, capacity of 4.55GB/9.1/18.2GB , Data rate 231 Mbit/sec, Spin rate 10,025 RPM, Seek time 5.4 ms, Stepper Actuator, Platter capacity 1.52 GB, Interface Ultra SCSI, Form 3.5" one third height/one third height/half height.

Cheetah 3 - 1999 (ST -318203/ST-136403)

12/24 MR heads, capacity of 18.2 GB/36.4 GB , Data rate 308 Mbit/sec, Spin rate 10,016 RPM, Seek time 5.2 ms, Stepper Actuator, Platter capacity 3 GB, Interface Ultra 160, Form 3.5" one third height/half height.

Easily the world's fastest hard drive. For a long time there was nothing else spinning at 10,000 RPM, and nothing else even came close to the data transfer rate the Cheetah produced: up to 21.3 Megabytes a second. By way of comparison, a top-class 3 or 4GB IDE drive drive from the same era produced less than 7MB sec. Even the very best IDE drives could not match Cheetah 1 performance until nearly three years later when the third generation 7200s arrived.



Fig 7.12 Inside view of ST35530A

The series 2 and series 3 Cheetah were even faster. The data rate went up to an incredible 308M/bit sec, and the seek time was an astonishing 5.2ms-this was roughly half the seek time of most contemporary drives.

Seagate Medalist Pro 7200 Family (34520A/36530A/39140A)

4/6/8 MR heads, capacity of 4.56/ 6.51/9.15 GB, Data rate 194 Mbit/sec, Spin rate 7200 RPM, Seek time 9.5 ms, Stepper Actuator, Platter capacity 2.35 GB, Interface ATA 33, Form 3.5" slimline.

1998 Drive of the Year. With this drive Seagate finally reclaimed their traditional title as the high-performance leader. The Medalist Pro 7200 was the world's first 7200 RPM IDE drive, and its data transfer rate was almost 200Mbit/sec - roughly double the speed of most drives sold in 1997, and faster than all but a very few of the most expensive SCSI drives too. These were only a little more expensive than other IDE drives and were significantly faster than any competitor.

Western Digital Expert (AC313500/AC418000)

6/8 MR heads, capacity of 13.5 GB/ 18 GB, Data rate 224.3 Mbit/sec, Spin rate 7200 RPM, Seek time 9 ms, Stepper Actuator, Platter capacity 4.55 GB, Interface ATA 66, Form 3.5" slimline.

1999 Drive of the Year. Seagate had been the first with a 7200 RPM drive; they were quickly followed by IBM and, surprisingly, Maxtor - but none of the early 7200s gave the quantum performance leap that these did. They were the first second-generation 7200 RPM drive, easily faster than any of the previous three.

Samsung SpinPoint V6800 Family (SV0682D/SV1364D/SV2046D)

2/4/6 MR heads, capacity of 6.8 GB/13.6 GB/20.4 GB, Data rate 247 Mbit/sec, Spin rate 5400 RPM, Seek time 9.5 ms, Stepper Actuator, Platter capacity 6.8 GB, Interface ATA 66, Form 3.5" slimline.

Another quiet achiever from Samsung, with nothing much bar keen pricing to distinguish it from the crowd. This, of course, is a virtue in anything other than an out and out high-performance drive. Indeed, the more Samsungs we used, the more they reminded us of the best entry-level Seagate Medalists of old: plenty fast enough, surprisingly cheap, and ultra reliable.

Samsung SpinPoint V1020 (SV1022D/2044D)

2/4 MR heads, capacity of 10.2 GB/20.4 GB , Data rate 291 Mbit/sec, Spin rate 5400 RPM, Seek time 8.8 ms, Stepper Actuator, Platter capacity 10.2 GB, Interface ATA 66, Form 3.5" slimline.

2000 Drive of the Year. Samsung have come a long, long way since the horrid little black things they used to make in 120MB days. They were as cheap as or cheaper than competing drives, they remained readily available nearly all year round (itself no mean achievement in a year when major component shortages caused chaos throughout the storage industry).

Seagate Barracuda ATA III (ST-320414A/ST340824A)

2/4 MR heads, capacity of 20 GB/40 GB, Data rate 500 Mbit/sec, Spin rate 7200 RPM, Seek time 11.8 ms, Stepper Actuator, Platter capacity 20 GB, Interface ATA 100, Form slimline.

A rather surprising development from Seagate. Data transfer rates went up to an amazing 500Mbit/sec- a world record and a technological miracle in reality, but just a routine advance for one of the big modern hard drive manufactures - especially Seagate, for they have always been good at upping the internal data rate.



Fig 7.13 View of ST2042 with cover removed

Samsung SpinPoint V2040 (SV2042D/SV4084D)

2/4 MR heads, capacity of 20.4 GB/40.8 GB, Data rate 343 Mbit/sec, Spin rate 5400 RPM, Seek time 8.8 ms, Stepper Actuator, Platter capacity 20 GB, Interface ATA 66/100, Form slimline.

2001 Drive of the Year. Just like the V1020, the V2040 was quiet, cheap, fast, and best of all it has maintained Samsung's spotless reliability record.

Seagate Barracuda ATA IV (ST/3200011/ST-340016/ST-300021/ST-380021)

1/2/3/4 GMR heads, capacity of 20 GB/40 GB/60 GB/80 GB , Data rate 555 Mbit/sec, Spin rate 7200 RPM, Seek time 10.8 ms, Stepper Actuator, Platter capacity 40 GB, Interface ATA-100, Form slimline.

This fifth of the Seagate 7200 RPM IDE drives is astonishingly quiet. It has virtually no idle noise at all and ultra-quiet seeks too - even the Samsung 5400s seem loud by comparison with a Barracuda IV.

Samsung SpinPoint V40 (SV2001H/SV4002H/SV6003H/SV8004H)

1/2/3/4 GMR heads, capacity of 20 GB/40 GB/60 GB/80 GB , Data rate 443 Mbit/sec, Spin rate 5400 RPM, Seek time 8.9 ms, Stepper Actuator, Platter capacity 40 GB, Interface ATA-100, Form slimline.

Another quiet Samsung achiever, it seems. The jump to 40GB/platter drives has been a major hurdle for the industry.

Samsung SpinPoint P80

The P80 arrived mid-way through the year. It was available in a range of models: there were 3 basic flavours: 80, 120 and 160GB, all three available with standard ATA interfaces in 2MB and 8MB cache versions, and in SATA with 8MB. Data rate 741 Mbit/sec and Spin rate 7200 RPM. Samsung also listed a 60GB unit, but demand for that odd size was never strong and as far as we know it was not imported here. The most popular model by quite a margin was the SP0802N (2MB cache and 80GB).

The P80 has had an extraordinarily long market life: It continued as majority of Samsung's production as late as July 2005. It wasn't until April 2005 that a new 200GB model appeared to supplement it.

SATA TECHNOLOGY

The serial ATA (SATA) computer bus, is advancement of ATA (AT attachment). It is the latest storage-interface which connects host bus adapters on the motherboard to secondary storage devices such as hard disk drives and optical drives. All the motherboards and the laptop computer are having the SATA host adapter on the mainboard. A single cable with a minimum of four wires creates a point-to-point connection between devices.

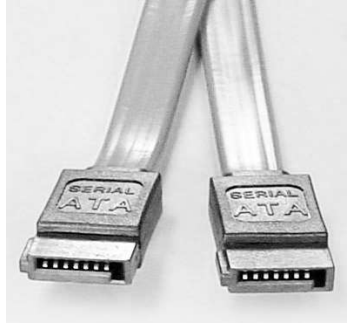


Fig 7.14 SATA connector

SATA use the same low level commands, but serial ATA host-adapters and devices communicate via a high-speed serial cable over two pairs of conductors. In contrast, the parallel ATA (the redesignation for the legacy ATA specifications) used 16 data conductors each operating at a much lower speed. This enables cables upto one meter as compare to IDE where the length is limited to 40cm.

SATA Road Map

	1st gen.	2nd gen.	3rd gen.
Approximate Data Rate	150MB/s	300MB/s	600MB/s
Approximate Bus Speed	1.5GB/s	3.0GB/s	6.0GB/s
Introduction	Mid 2002	Mid 2004	Mid 2007
2003	Serial ATA introduced		
2005	First 500 GB hard drive shipping (Hitachi GST)		
2005	Serial ATA 3G standardized		
2005	Seagate introduces Tunnel MagnetoResistive Read Sensor (TMR) and Thermal Spacing Control		
2005	Introduction of faster SAS (Serial Attached SCSI)		
2005	First Perpendicular recording HDD shipped: Toshiba 1.8-inch 40/80 GB		
2006	First 750 GB hard drive (Seagate)		
2006	First 200 GB 2.5" hard drive utilizing Perpendicular recording (Toshiba)		
2006	Fujitsu develops heat-assisted magnetic recording (HAMR) that could one day achieve one terabit per square inch densities.		
2007	First 1 terabyte hard drive (Hitachi GST)		
2008	First 1.5 terabyte hard drive (Seagate)		
2009	First 2.0 terabyte hard drive (Western Digital)		

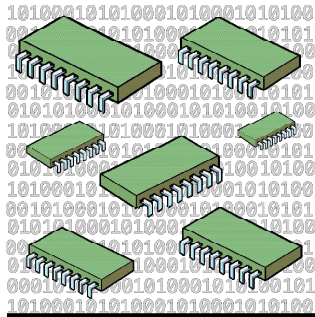
EXERCISE

1. Fill in the Blanks
 - a) The Hard disk speed normally varies from _____ to _____.
 - b) Hard disk head could of _____ and _____ type.
 - c) In a disk, information is stored on _____.
 - d) A hard disk drive usually has _____ read / write head for each platter side.
 - e) The size of head is _____ than the size of a dust particle.
 - f) Present GMR technology permits to store around _____ per square inch.
 - g) The Pixie dust technology would let you store _____ per square inch.
2. What is a megabyte?
3. What is secondary storage? How does it differ from a primary storage?
4. Who are the major players in Hard Disk manufacturing?
5. What are Winchester disk and why they are called so?
6. Discuss the various components used in construction of Hard disk?
7. What is MR technology. What advancements has been carried out in GMR technology?
8. What is Pixie dust?
9. How stepper motor actuators head mechanism is different from Voice coil actuators mechanism?
10. The following are different sorts of secondary storage media. Arrange them in ascending order according to their storage capacity and explain also.
 - (a) a standard reel of magnetic tape
 - (b) an 8-inch Winchester disk
 - (c) a cassette tape
 - (d) an 8-inch floppy disk
 - (e) a punched-card
11. Identify five advantages and five disadvantages of magnetic disk storage.
12. Write short notes on:
 - a) seek time
 - b) latency time
 - c) average access time
 - d) transfer rate
13. What is average seek time and average access time and how they are related?
14. A 3-disk pack having 600 tracks per surface has 10 sectors per track. If every sector could store an MB of data, calculate

the storage capacity of the disk. If disk transfers 1200 KB per second, calculate the disk rpm.

What are the main features of EIDE and Fast-ATA?

15. What are the main features of Ultra-ATA? How is Ultra ATA/100 different from Ultra ATA/33 or Ultra ATA/66?
16. What are ATA/66 cables for? How do they differ from the normal IDE cables?
17. What's the difference between IDE and SCSI?
18. What are the different kinds of IDE?
19. What are the different kinds of SCSI?
20. What benefits do higher RPM Hard Drives have?
21. What is the significance of the different jumper settings on hard drive?
22. What is FAT16? How is it different from FAT32?
23. What is defragmentation and how do you defrag your hard drive and why is it necessary?
24. What does a format do and when do you need to do it?
25. Does it matter how you connect the devices to the cable?
26. If you get an error saying "Non system disk or disk error". What does this mean?



CHAPTER 8

SECONDARY STORAGE - III

CARTRIDGE TAPE DRIVE
CD- ROM (COMPACT-DISK-READ-ONLY-MEMORY
MEDIA FOR CD-RECORDING
DVD
MAGNETO-OPTICAL TECHNOLOGY
SUPER-FLOPPIES (ZIP)

Secondary Storage - III

CARTRIDGE TAPE DRIVE

A tape backup unit makes backups much easier by avoiding the labour of swapping diskettes in and out of the machine. These units are expensive and they are useful only for backups and archiving files. Most are installed by adding an adapter card in one of the computer's internal slots. The tape drive itself may be housed in the computer, but many are independent units that may be shared between several machines.

Because IBM didn't offer its own tape backup unit at the time the technology first became popular, there is no single standard to which all units subscribe. They vary in the size of tape cartridges they use and the number of tracks written upon the tape. They also vary in the way data is formatted on the tape. It is all the same from the user's point of view, although some formats are considerably faster or more reliable than others, and some are better able to recover data from damaged media.

Under ideal circumstances, data is continuously transferred from hard disk to tape as quickly as the tape drive can absorb it. This phenomenon is called "streaming". It is more of an ideal than a reality, because most computers can't supply data quickly enough. The drive reacts by constantly starting and stopping, or by leaving blank gaps as the tape moves continuously.



Fig 8.1(a) Tape Drive

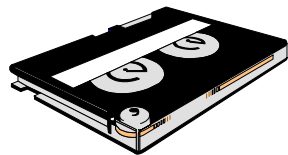


Fig 8.1 (b) Magnetic Tapes

TAPE OPERATION

Recording is done in a bit serial recording format, using a suspensive stream configuration illustrated in figure. The actual nine track recording continues this sequential pattern. Data is encoded to (0, 2) Run Length Limited (RLL) code and recorded at 10, 000 flux changes per inch (fci) on the nine tracks. The streaming format provides constant tape motion and very short interblock gaps with minimal format overhead. This format maximized data throughput efficiency.

Recording Method

Although magnetic tape is an extremely reliable, cost effective mass storage media, even the highest quality tapes are subject to errors caused by small imperfections in the tape oxide coating. For this

reason, Non-Return-to-zero (NRZI) recording method is used to reduce sensitivity to tape imperfection through lower density recording. In the NRZI method, however, the normal 8-bit binary sequence can result in a data pattern with a large number of successive zeros. To preclude the need for Read circuits that require complex data-bit window functions, the 8 bit bytes from the host are converted to a 0, 2 RLL code before being operations. The RLL code provides a 10 bit serial stream with no more than two consecutive zeros in any dot pattern. It also permits the data stream to be used to track the Read circuits to the data rate, so that speed variations between tape drives do not affect compatibility.

QIC tape uses a linear read/write head similar to those found in domestic cassette recorders. Typically this comprises a single write head flanked on either side by a read head. This allows the tape drive to verify data just written when the tape is running in either direction. If the data just written is verified by the read head, the buffer is flushed out and new data is acquired from the system memory. If errors are found, the segment is rewritten on the next length of tape.

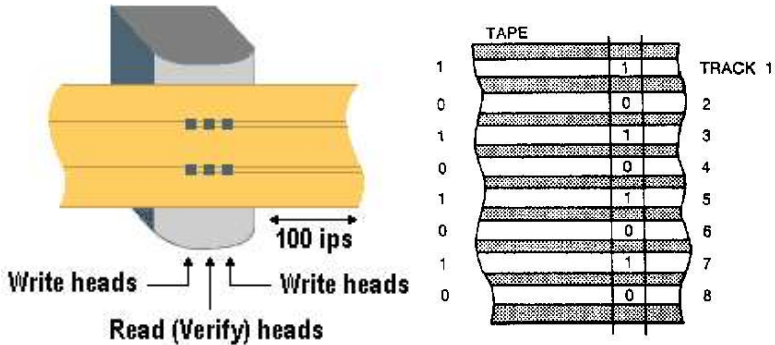


Fig 8.2 Reading & Writing on QIC tape and Parallel Format

Recording Format

The data format compatible to the quarter Inch Committee’s standard QIC (quarter-inch cartridge) - 24 Revision D. Even numbered tracks are recorded in the forward tape direction, odd numbered tracks in the reverse tape direction.

All data for interchanged is recorded in designated tape areas:-

- Even tracks from Load point (LP) marker to End-of-Tapes (EOT) marker.
- Tracks 3 and 5 from the early warning (EW) marker to Beginning of tape (BOT) marker.
- Tracks 1 and 7 from the EW marker to the LP marker. Tracks are recorded sequentially, from 0 through 8.

Data Block Formatting

The data block marker identifies the start of the data block. The data block contains 512 bytes of data for interchange. The data is encoded according to 0, 2 RLL transformation table.

The block address contains four bytes that provide a unique identification for the data block recorded on tape.

The cyclic Redundancy check (CRC) consists of two bytes calculated over the 512 bytes interchange data and 4 byte block address. CRC starts with an initial value which is all ones and uses this CRC generating polynomial

$$\mathbf{X^{16} + X^{12} + X^5 + 1}$$

The CRC is encoded, using the 0, 2 RLL transformation table. The postamble is recorded after the CRC and serves as on a guard band.

Interfacing

Data and commands are transferred to, or from the tape formats on an 8 bit bidirectional data bus, using asynchronous handshaking techniques to eliminate tight timing constraints upto four devices can be supported by the interface which is compatible with the standard QIC-02 Revision D.

Types of Media

Two primary standards are

- DC - 600 (3M Data Cartridge 600)
- DC - 2000

DC-600 is currently represents the best and most effective media type of all the standards DC-600 drives currently store 60 M, 125 M, 150 M, or 320 M, depending on the actual format and quality of tape used. The DC-2000 media only stores 40 M per tape and may be suitable for low end use.

DAT

DAT stands for Digital Audio Tape and, as its name implies, was originally conceived as a CD-quality audio format. In 1998, Sony and HP defined the DDS (Digital Data Storage) standard, transforming the format into one that could be used for computer data storage.

DAT technology is a 4mm tape which employs a technique called helical scan recording. This is the same type of recording as that used in video-tape recorders and is inherently slower than the linear type. For this reason, it is generally only used in environments where high capacity is the primary requirement.

The tape in a helical scan system is pulled from a two-reel cartridge and wrapped halfway around a cylindrical drum containing two read

heads and two write heads, arranged alternately. The read heads verify the data written by the write heads. The cylinder head is tilted slightly in relation to the tape, and spins at 2,000rpm. The tape moves in the opposite direction to the cylindrical spin, at less than one inch per second but because it is recording more than one line at a time, it has an effective speed of 150inch per second. Short diagonal tracks are written across the width of the tape, about eight times longer than the width. These contain about 128KB of data and an error correction code (ECC) each.

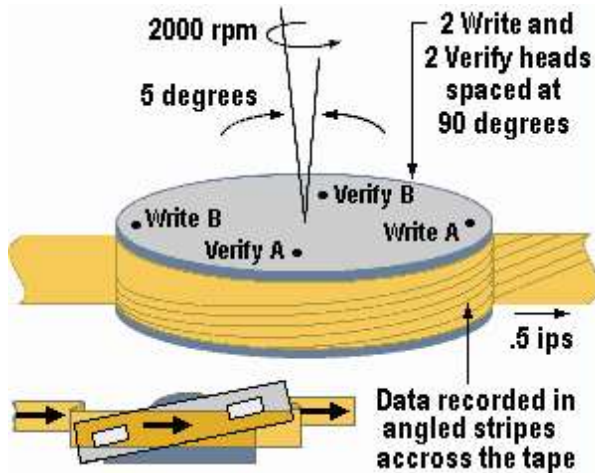


Fig 8.3 Recording in DAT

DAT tapes are available in two formats: DDS and DataDAT. The DDS protocol is the most common, and is represented by several standards, all of which are backward compatible:

Standard	Capacity	Max. DTR
DDS	2GB	55 KBps
DDS-1	2/4GB	0.55/1.1MBps
DDS-2	4/8GB	0.55/1.1MBps
DDS-3	12/24GB	1.1/2.2MBps
DDS-4	20/40GB	2.4/4.8MBps

8 MM

8mm tape technology was originally designed for the video industry. Its original purpose was to transfer high-quality colour images to tape for storage and retrieval. Now 8mm technology has been adopted by the computer industry as a reliable way to store large amounts of computer data. Similar to DAT but with generally greater capacities, 8mm also employs helical scan technology. A drawback to the helical scan system is the complicated tape path. Because the tape must be pulled from a cartridge and wrapped tightly round the spinning read/write cylinder, a great deal of stress is placed on the tape.

Standard	Capacity (n/c)	Interface	Maximum DTR
Standard 8mm	3.5/7GB	SCSI	32 MB/min
Standard 8mm	5/10GB	SCSI	60 MB/min
Standard 8mm	7/14GB	SCSI	60 MB/min
Standard 8mm	7/14GB	SCSI	120 MB/min
Mammoth	20/40GB	SCSI	360 MB/min
AIT-1	25/50GB	SCSI	360 MB/min

Mammoth

Exabyte has been a leader in the tape storage industry for more than a decade, pioneering the use of 8mm tape for backup, incorporating Sony's camcorder-based mechanisms into over 1.5 million tape drives. While camcorder-based mechanisms are adequate for low duty cycle applications, it is less appropriate for today's demanding server-based applications. Introduced in 1996, Mammoth is a more advanced and reliable technology and represents Exabyte's response to the requirements of this mid-range server market.

Advance Intelliget Tape (AIT)

Advance Intelliget Tape (AIT) Technology is the first multisourced tape standard targeted at the mid range server market. The technology is the result of several breakthrough including stronger and thinner media that is more stable and has better coating than previously available. In addition, new head technologies, higher level of intergration and a unique Memory in Cassate (MIC) feature have resulted in multi-Gbyte, high performance tape drive systems with low frequencies of error that are perfect for tape libraries and robotic applications associated with midrange systems backup.

AIT drive such as Segate Sidewinder drive ensure data integrity with read-while-write and Error Correction Code(ECC) technology. Read-while-write detects and corrects any write anomalies by rewriting the data after the dropout area of the tape has passed. ECC, which was developed by the aerospace industry, is a process of mathematically describing the data so that it can be reconstructed if lost.

Data Integrity is further ensured through hardware and media reliability. By starting with the best possible media, the ability to write data correctly is dramatically increased.

AIT's advance Metal Evaporated (AME) media technology is the cornerstone of the sidewinder 50's ability to reach an industry-leading bit error rate of less than 10⁻¹⁷.

The combination of head technology and AME tape creates a system that yields a head life of greater than 30,000 hrs (app. 10 yrs at 33% duty cycle - 8 hrs for seven days a week).

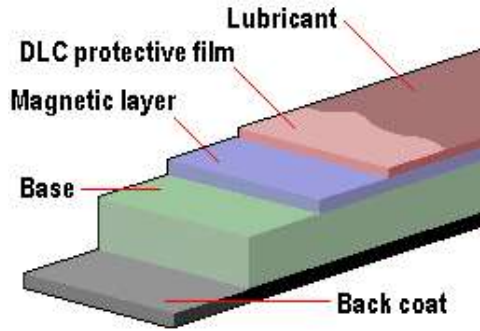


Fig 8.4 AME media Technology

There are other similar products such as Quantum's DLT-4000 with a head life of 10,000 hrs and Exabyte's Mammoth, which offers 16,000 hrs.

AIT drives such as Segate's Sidewinder 50 currently hold upto 50 Gbyte of data and execute transfer rate of upto 3 Mbps in native mode and upto 6Mbps in compressed mode.

Segate's Travan NS20

Desktop storage capacity are increasing day by day. Thus, the need for fast, accurate, easy to use, reliable and affordable backup is increasing.

To meet the need, Segate introduced the NS20 minicartridge tape drive with a native capacity of upto 10 GB and a compressed data capacity of 20 GB. The media of the Travan NS20 is coated with an alumunium oxide which effectively repel and removes any contamination that may be present.

INTERNAL AND EXTERNAL CTDS

Most of tape adapter cards provide an interface for external tape unit, connect the tape unit with the external cable and oftenly you have to use external power supply with an external unit, one can share the backup unit with many systems. All you have to do is equip each system unit with the required interface.

CTD INSTALLATION

1. The tools required to install tape adapter card and tape drives are:
 - Medium-sized flat blade screw driver.
 - Installation instructions for your tape drive.
2. Place the system unit switch to the OFF position.
3. Turn OFF all external equipment.
4. Remove the system unit cover.

5. Locate an unused expansion slot near the power supply.
6. Connect the 50 Pin ribbon cable to tape adapter taking care of red or blue stripe on the cable should be aligned with the side of the header labeled "1".
7. Install the tape drive according to manufacturers instructions.
8. Connect the power supply to tape drive using the connectors on the supply.
9. Connect 50 Pin cable to tape drive.

BACKUP USING SOFTWARE

Most manufacturers have written their own software, which is proprietary and used only by that particular manufacturer. In other words, even though you may be using a tape unit with the exact media and interface, if the software is different, you will not be able to interchange data between the units. The software called SYTOS (Sytron Tape Operating System) is becoming an industry standard.

Any software used for tape backup should be able to do certain basic operations like:-

- The software can backup an entire DOS partition (Image backup)
- The software can back up any or all files individually (File-by-File backup)
- The software allows a selective file-to-file restore from an image backup or file-by-file backup.
- The software combines several backup on a single tape.
- Backup can be completely verified.

Cassette Tape vs. Hard Disk

Let's look at the big differences between cassette tapes and hard disks:

- The magnetic recording material on a cassette tape is coated onto a thin plastic strip. In a hard disk, the magnetic recording material is layered onto a high-precision aluminum or glass disk. The hard-disk platter is then polished to mirror-type smoothness.
- With a tape, you have to fast-forward or reverse to get to any particular point on the tape. This can take several minutes with a long tape. On a hard disk, you can move to any point on the surface of the disk almost instantly.
- In a cassette-tape deck, the read/write head touches the tape directly. In a hard disk, the read/write head "flies" over the disk, never actually touching it.
- The tape in a cassette-tape deck moves over the head at about 2 inches (about 5.08 cm) per second. A hard-disk platter can spin underneath its head at speeds up to 3,000 inches per second (about 170 mph or 272 kph)!

- The information on a hard disk is stored in extremely small magnetic domains compared to a cassette tape's. The size of these domains is made possible by the precision of the platter and the speed of the medium.

CD- ROM (COMPACT-DISK-READ-ONLY-MEMORY)

Computers never seem to have enough memory. The more memory a computer has, the larger are the applications that are undertaken. Larger applications, in turn, create a demand for more memory capacity. The need to store and retrieve very large amount of data led to the development of optical-disk storage system. Because optical disk systems can store as much as 100 times the contents of current magnetic disk drives, they seem to afford a promising solution to the storage problem. An optical disk storage system is similar in principle to the laser-optical video disks sold for home use. As applied to computer-related information systems, optical disk store digitized data instead of the digitized elements of a picture.



Fig 8.5 CD-Player and CD-ROM

In optical-disk recording system an intense laser light source is focused on a rotating disk, which is coated with a reflective surface. Data recording is accomplished by turning the laser beam.

Tiny pits, visible only under a microscope, are burned or pressed into a thin coating of metal or other material deposited on disk. These pits may be on a 12-inch disk that records the sights and of movies, concerts, and sporting events. Or they may be on a compact (4.7-inch) disk (CD) that records high-quality music. In both cases, the pit patterns represent the streams of digital data that are used to encode images and sounds. When these optical disks are placed in disk playback devices, a beam of laser light is used to read the pit patterns and convert these patterns into the audio-visual signals.

The storage density of optical disks is enormous, the storage cost is very low, and the access time is relatively fast. Just one CD-ROM disk stores about 550 MB (a megabyte is 1.048 million bytes, so 550 megabytes is over 570 million alphanumeric characters). Thus, one small optical disk can store the contents of the of Encyclopedia Britannica. All 450 million characters of it, can be stored in a number of ways, from its original manuscript to an optical laser disk. This is what it would look like formatted in different media. The optical laser

disk is said to store information equivalent to 225,000 sheets or 33 published volumes, or 1,250 floppy disks of 360 KB each or 8 tape cartridges at 60 Mb each or 6 hard disks at 80 Mb each.

Advantages of CD-ROM

- High capacity storage. Optical disks have around 100 times higher capacity than magnetic disks of the same area. Single surface CD-ROMs have around 650 MB disk capacity
- Very high reliability. CD-ROMs are less prone to damage by dust and magnetic fields. Since there are no “reading heads”, the disks are not damaged by the heads
- Low storage costs.
Provide random access like magnetic tapes
- CD-ROMs are easily portable like floppy disks, inspite of a much higher capacity.
- CD drives are slower than magnetic disks.

DISK STRUCTURE

The CD-ROM disk is 120 mm (about 4. 72 inches) in diameter, 1. 2 mm thick and has a hole 15 mm across in the centre. The information is in the form of pits and lands. The pits are 0. 12 micrometer deep and about 0. 6 micrometer wide. The neighbouring turns of the spiral pattern of pits are 1.6 micrometer apart. In comparison, a human hair is around 50 micrometer in diameter.

The spacing corresponds to a track density of 16, 000 tracks per inch, much higher than the figures associated with a floppy (up to 96 tracks per inch) or hard disks (several hundred tracks per inch). Along these turns, pits and the spaces between them (called lands) run from 0. 9 micrometer long. The total length of the track on a CD-ROM disk is almost three miles. The total number of the pits on a CD-ROM is about two billion. The track width is about 600 nanometers wide and adjacent turns of the spiral track are about 1, 600 nanometers apart, making for a density of about 16, 000 turns per inch.

DATA ORGANIZATIONS

The lands and pits on a CD-ROM disk do not represent ones and zeros. The reason is complex to explain and understand. Briefly, each bit requires about 300 mm of length along the spiral track. Thus, if a sector has 2, 048 data bytes, comprising all zeros and ones, then it will have a pit of land measuring 4, 915, 200 nanometers or longer. The CD-ROM drive will have to maintain a precise clock that picks at an interval, representing 300 nanometers in distance, so that the next bit could be read at the right line. This becomes extremely difficult.

To overcome the difficulties, designers have limited the length of the pits and lands to a specified range. This ensures that a transition from one to the other occurs frequently enough for the CD-ROM drive

to drive a clock. The drive uses this clock to count the number of bits on pit or land, as determined by its length. The bits are stored on the disk in RLL (run length limited) format- the same self-clocking approach used for storing data on most hard disks. Each transition on a disk between a land and a pit represents a 'one' and a non-transition represents a 'zero'. The minimum practical length of a pit or land is 3 bits (900 nanometer); the maximum is 11 bits (3, 300 nanometers).

CD-ROM drives use an encoding scheme that converts data bytes into 14-bit patterns called channel bits. A 25 bit binary value provides 16, 384 bit patterns; more than 256 of these meet the requirements for the CD-ROM encoding format.

A one-to-one correlation is then made between each of the 256 binary patterns represented by an 8-bit byte and 256 of the channel bits that meet the CD-ROM formatting requirements. This method, called eight-to-fifty six modulation (EFM), requires a look-up table to correlate the 14-bit patterns on the disk to their 8-bit data values.

Standard CD-ROM sectors consist of 2, 352 bytes totally organized, including 12 bytes of synchronization data, followed by a 4-byte header, 2, 048 bytes of error corrected data and a 288-byte segment for ECCs/EDSS. As with audio CDs, CD-ROM disks are generally made using a 60 minute spiral, having 270, 000 sectors. These disks can, however, be made using up to a 74-minute spiral with 333, 000 sectors. The additional 14 minutes are located on the outer 5 mm of the disk. Since this area is the hardest to record well and keep clean, it is often left unused.

DRIVE MECHANISM

The data bits on a CD-ROM disk must pass under the read head at a constant rate known as Constant Linear Velocity (CLV). Because the sectors on the outer and inner edges of the disks are of the same length, achieving a CLV requires changing the rotational speed of the disk as the head position changes. As the head moves farther away from the centre, the rotational speed slows.

In contrast, conventional hard disks operate at a constant regular (rotational) velocity or CRV. As a result, the physical sector sizes grow longer as the tracks get farther from the centre. CD-ROM drive access time is very large (i.e. , slow) when compared to the sub 20 millionseconds access time of most hard disks. Here, we must remember that the sheer bulk of a CD-ROM optical head's assembly is the limiting factor. Hard disk read heads are small and light, while the laser, lenses and the assembly hardware make it difficult to move fast.

READING THE DATA

In order to read the data from the disk, a laser beam is focused on the spiral track of pits and the amount of light reflected back into the

objective lens is measured. Light striking one of the pits is diffracted (scattered) through such a wide angle that very little finds its way back into the lens. But when the spot focuses on the flat land between the pits, most of it is reflected back into the lens.

The modulated signal, produced by this combination of reflected and diffracted light, actually represents the information stored on the disk. The reflected light goes to a photo detector which produces a current proportional to the light intensity. The reflected light signal changes each time the laser beam moves from a pit to a land or vice versa. Through a decoding process, called Eight-to-fourteen-modulation, the information derived from this varying light signal is converted to digital data.

Until now, CD-ROM disks were mass produced by giant electronic companies. The disks are injection moulded in dies wherein all the format and datemarks appear as raised ridges or hills on one of the interior faces. The original (also called the master) disk is written with great precision by an Argonion laser.

After the master is developed, a negative father disk is used as a stamper to form a mother disk, whose negative sons become one face of the dyes, used in the injection moulding process that yields the final product. Polycarbonate is the material of choice because of its strength and durability. To safeguard against loss of data, the CDs must be protected from scratches on both the sides, i.e. the top and the bottom.

MEDIA FOR CD-RECORDING

There are two types of recordable media that are used for making your own CDs:

- CD-R discs (CD-Recordable)
- CD-RW discs (CD-ReWritable)

You can obtain blank CD-R or CD-RW discs where you purchased your CD-Recorder or at local computer retailers.

COMPACT DISK RECORDABLE (CD-R)

A more recent addition to the family of CDs, a Recordable Compact Disk (CD-R) enables one to record information on a CD-ROM disk. In other words, one can prepare a CD-ROM, putting in the information of his choice for reference. As just mentioned, prior to this, disks could be prepared only by giant electronic companies from master disks and pressed to manufacture final CD-ROM disks for distribution. The data etched was permanent and could only be retrieved and not erased or changed.

With CD-R, preparing a CD-ROM disk has finally become possible on a desk top computer system with a CD-ROM drive and the special CD-

R disks. Now, with the development of CD-Recordable disks, CD-ROMs are poised for a big take off.

CD-R has revolutionised the concept of portable data storage with a capacity of 700 MB. Since 1980, there were drives known as WORM (WRITE ONCE READ MANY). The user can create data on the WORM but cannot erase the data once created. Removable cartridges can be used with WORM drive. These WORM drives can store up to 63 0 MB.

The major advantages of WORM over HDD are

- Large storage capacity
- Unlimited life
- High reliability
- Lower cost

CD-R disks use the same technology as WORM, changing the reflectivity of the organic dye layer, which replaces the sheet of reflective aluminium in a normal Compact Disc. In the case of writable CDs, the substrate has a U-shaped groove (600 nanometers wide and 100 nanometers deep) covered by a sensitive layer of organic dye.

This is typically cyanine or phtalocyanine. These dyes are photosensitive organic compounds similar to those used in making photographs.

To protect the dye, a microscopic reflective layer of gold or silvery alloy is coated over dye. The colour of the CD-ROM you see in the market is determined by the colour of the reflective layer and the colour of the dye. For example, a silver/blue discs combine a silver reflective layer with a cyanine (Cyan colour) dye resulting in a gold appearance on the label side and a cyan appearance on the writing side.

Write and Read

During the Write process, the data is written on the disk by irreversibly altering its organic polymer dye layer. The altering process is accomplished using high-intensity laser, which heats the layer to a certain temperature, which, when reached, changes the layer from opaque to translucent.

The writing is done spirally on the CD-R, beginning from the centre and continuing to the outer edge of the database. The spiral track of conventional read-only CDs consists of a series of pits, etched into a polycarbonate substrate that is covered by a reflective layer.

When data is written, physical marks are made on the media surface by a low-powered laser and since, the marks are permanent, they can't be erased.

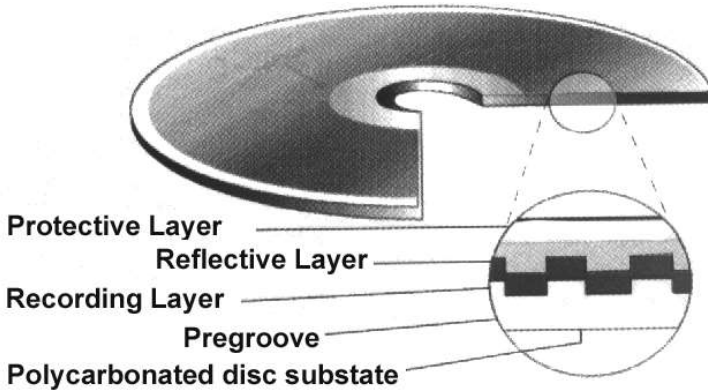


Fig 8.6 Different layers of CD-R

When the written CD is read, the laser from the read head of the driver reads the data by focusing on the reflective layer and the reflected light enters a photoelectric sensor. When laser hit the translucent area the reflected value is higher than the opaque area. The intensity of reflected laser light varies due to diffraction caused by the pits. The reflected light is modulated and modulation is transformed into electrical signals of 0s and 1s to represent a string of information.

CD-ROM disks have become the preferred medium for distributing information from operating systems to multimedia applications. Now, with the development of affordable CD Recordable disk, the use of CD-ROM with personal computers will grow substantially. As the price of CD-Recordable drivers becomes increasingly more affordable, more and more users will prefer to use CD-ROMs for storing volumes of information, with the ease of retrieval at their fingertips.

Virtually all CD-ROM and CD-Recorder drives can read CD-R discs. CD-R discs are excellent for permanent data storage and should be used when you do not need to erase the data. They are less expensive per CD than CD-RW discs. Here are some ideas for using CD-R discs:

- Use when you want to make a music CD that you can play on your home or car stereo CD player.
- Use when you want to copy a music CD for your own personal use.
- Use when you want to permanently save large project files.
- Use when you want to permanently archive or store files, such as your home records, photos, or videos.

CD-ReWriteable

CD-RW discs can only be used in CD-ReWriteable drives and MultiRead CD-ROM drives. You cannot play a music CD made on a CD-RW disc in most home or car stereo CD players. Here are some ideas for using CD-RW discs:

- Use when you need to write, erase, rewrite, and update all your large files on CD.
- Use when you want to make a practice CD or test the contents of a CD before making a permanent one.

The technology behind CD-RW is known as optical phase change, which in its own right is nothing radical.

The media of CD-RW is similar to that of CD-R having the similar structure but with significant detail difference.

A CD-RW disc's phase-change medium consist of a polycarbonate substrate, moulded with a spiral groove for guiding the servo meter, absolute time information and other data. The recording layer is sandwiched between dielectric layers that draw excess heat from the phase change layer during the writing process.

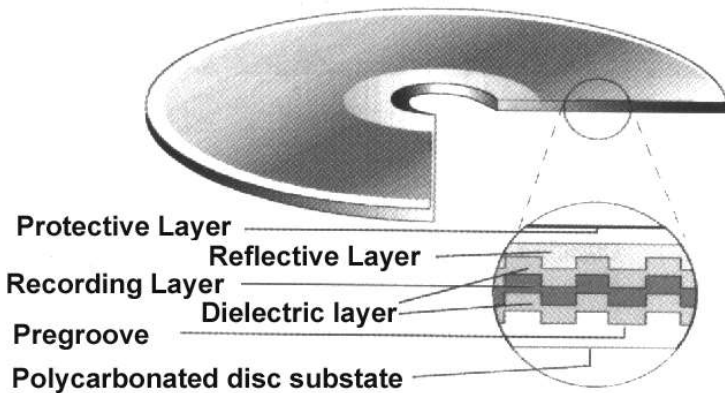


Fig 8.7 The dielectric layer allows the data to be re-written

In place of the CD-R disc's dye-based recording layer, CD-RW commonly uses a crystalline compound made up of a mix of silver, indium, antimony and tellurium. This rather exotic mix has a very special property: when it is heated to one temperature and cooled it becomes crystalline, but if it's heated to a higher temperature, when it cools down it becomes amorphous and can absorb laser beams. The crystalline areas allow the metallic layer to reflect the laser better while the non-crystalline portion absorbs the laser beam, so it is not reflected.

Speed of Reading

There are 2 kinds of CD-ROM/DVD Drives on the market, Fixed and Variable speed (not to be confused with Fixed and Variable Spin where the situations described below are reversed).

Thus, in a variable speed CD-ROM, your 32X CD-ROM drive wouldn't read data at 32X speed always. A 32X CD-ROM attains this speed only on a disk's outer tracks, whereas the reading speed on the inner

tracks (where data is stored in the beginning), is only about 12-16X. Data rarely fits the outer track so drive rarely reads 32X.

The CD-ROM in the market are described by the speed at which they can read, write or rewrite the data. You will often found 24X 10X 40X types of configuration in the CD-Writers. This indicates that the writing speed is 24X, rewriting speed is 10X, and reading Speed is 40X. Following table gives you fair idea that how much data will be read or wrote at any specific speed.

1x	150kb per Second
2x	300kb per Second
3x	450kb per Second
4x	600kb per Second
6x	900kb per Second
8x	1200kb per Second
10x	1500kb per Second
12x	1800kb per Second
16x	2400kb per Second
24x	3600kb per Second
32x	4800kb per Second
40 X	6400kb per second
52X	7200kb per second

DEVELOPING INDUSTRY STANDARDS

CD-ROM technology has been around for over ten years. As with all technological break through, however, the development of industry standards is an important part of the process of the technology becoming both popular and useful. Inventions such as the telephone, television, and the microprocessor did not become useful until standards for the technologies were developed. Without standards, you cannot communicate with everybody who has a particular technology. Television, for example, would not be wide spread today without the development of standards.

You would not be able to watch all televised shows if industry standards were not in place. The compact disc is a result of a joint venture by Sony and Philips. The CD audio specification, which describes the physical characteristics concerning the size of the pits and lands that represent data and their arrangement in a spiral, was announced in 1980.

This standard, known as the Red Book standard because of the color of the binder in which it was published, was introduced as a way to ensure that any audio CD would play in any audio CD player. This specification is relatively simple because an audio CD only contains music that is played from the disc in a stream.

The CD-ROM disc is an offshoot of the CD audio disc, and the CD-ROM specification, also known as the Yellow-Book, is an extension of the CD audio specification. This specification enabled the manufacturer to place different kinds of data on a CD-ROM.

INTERFACE TYPE

CD ROMS can be connected to the rest of your computer via a SCSI, IDE, EIDE, proprietary interface, or parallel port.

SCSI driver have the best overall performance. However, the cost is much higher than the IDE drivers. There are variety of proprietary CD-ROM drive interfaces out there in all shapes and sizes; on sound cards, as proprietary SCSI, and as a unique interface and adapter card bundled with CD- ROM drive. They are cheaper than standard interface. The mother boards of most ISA bus PC's include an IDE interface, but the drive houses the controller logic.

The 16 bit parallel interface can support two drives. When attaching both a HARD drive and a CD-ROM to the controller, the CD-ROM should be configured as the slave device.

SCSI is a good choice for those who need high performance and system flexibility. One SCSI host adapter card can provide a system interface for up to seven devices, and SCSI devices have a high data transfer rate.

EIDE is a rather new technology which is competing with SCSI-1, AN EIDE controller can support four devices, so it is easier to attach a CD-ROM. It is easier to install and cheaper to SCSI. It is downward compatible to IDE standards.

DVD

DVD once stood for **digital video disc** or **digital versatile disc**, but now it just stands for DVD - the next generation of optical disc storage technology. DVD is essentially a bigger, faster CD that can hold cinema-like video, better-than-CD audio, and computer data. DVD aims to encompass home entertainment, computers, and business information with a single digital format, eventually replacing audio CD, videotape, laserdisc, CD-ROM, and video game cartridges. DVD has widespread support from all major electronics companies, all major computer hardware companies, and all major movie and music studios. With this unprecedented support, DVD has become the most successful consumer electronics product of all time in less than three years of its introduction.

It's important to understand the difference between the physical formats (such as DVD-ROM or DVD-R) and the application formats (such as DVD-Video or DVD-Audio). DVD-ROM is the base format that holds data. DVD-Video (often simply called DVD) defines how video programs such as movies are stored on disc and played in a

DVD-Video player or a DVD computer. The difference is similar to that between CD-ROM and Audio CD. DVD-ROM includes recordable variations DVD-R/RW, DVD-RAM, and DVD+R/RW. The application formats include DVD-Video, DVD-Video Recording, DVD-Audio, DVD-Audio Recording, DVD Stream Recording, and SACD.

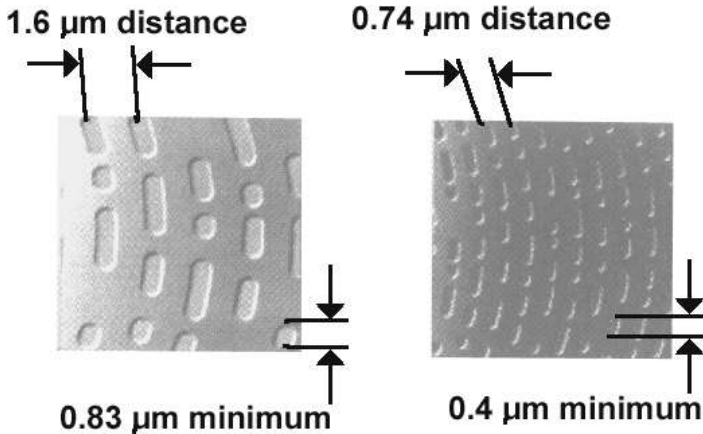


Fig 8.8 Data is more tightly packed in DVD-ROM

Size and capacity of DVD

There are many variations on the DVD theme. There are two physical sizes: 12 cm (4.7 inches) and 8 cm (3.1 inches), both 1.2 mm thick, made of two 0.6mm substrates glued together. These are the same form factors as CD. A DVD disc can be single-sided or double-sided. Each side can have one or two layers of data. The amount of video a disc can hold depends on how much audio accompanies it and how heavily the video and audio are compressed. The oft-quoted figure of 133 minutes is apocryphal: a DVD with only one audio track easily holds over 160 minutes, and a single layer can actually hold up to 9 hours of video and audio if it's compressed to VHS quality.

At a rough average rate of 4.7 Mbps (3.5 Mbps for video, 1.2 Mbps for three 5.1-channel soundtracks), a single-layer DVD can hold a little over two hours. A two-hour movie with two soundtracks can average 5.2 Mbps (with 4 Mbps for video). A dual-layer disc can hold a two-hour movie at an average of 9.5 Mbps (close to the 10.08 Mbps limit).

A DVD-Video disc containing mostly audio can play for 13 hours (24 hours with dual layers) using 48/16 PCM (slightly better than CD quality). It can play 160 hours of audio (or a whopping 295 hours with dual layers) using Dolby Digital 64 kbps compression of monophonic audio, which is perfect for audio books.

How DVD works?

In DVD, the tracks are placed closer together, thereby allowing more tracks per disc. The DVD track pitch (the distance between each) is

less than the half of CD's 1.6 microns. The pits, in which the data is stored, are also a lot smaller, thus allowing more pits per track. The packing of as many pits as possible onto a disc is however, the simple part and DVD real technological breakthrough was with its laser. Smaller pits means smaller laser to read the data. This was achieved by reducing the wave length of the laser from 780nm infrared light in CD to 635nm red light.

Also, the DVD specification also allows information to be scanned from more than one layer of a DVD simply by changing the focus of the read laser. Instead of using an opaque reflective layer, it's possible to use a translucent layer with an opaque reflective layer behind carrying more data. Although, this doesn't double the capacity of the disc as the second layer can't be as dense as first, but it does enable a single disk to deliver 8.5 GB of data without having to be removed from the drive and turn over. The interesting part is that second layer of the DVD may be read from the both side.

MAGNETO-OPTICAL TECHNOLOGY

As implied by the name, these drives use a hybrid of magnetic and optical technologies, employing laser to read data on the disk, while additionally needing magnetic field to write data. An MO disk drive is so designed that an inserted disk will be exposed to a magnet on the label side and to the light (laser beam) on the opposite side. The disks, which come in 3.5in and 5.25in formats, have a special alloy layer that has the property of reflecting laser light at slightly different angles depending on which way it's magnetised, and data can be stored on it as north and south magnetic spots, just like on a hard disk.

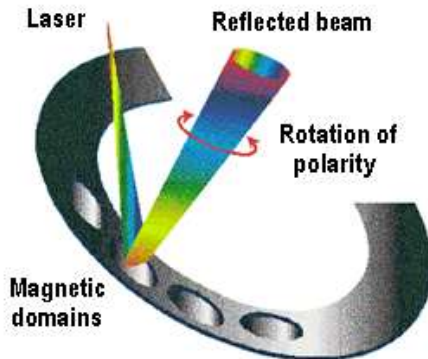


Fig 8.9 The Kerr Effect - The polarity of the magnetic particle determines the rotation of the beam

While a hard disk can be magnetised at any temperature, the magnetic coating used on MO media is designed to be extremely stable at room temperature, making the data unchangeable unless

the disc is heated to above a temperature level called the Curie point, usually around 200 degrees centigrade. Instead of heating the whole disc, MO drives use a laser to target and heat specific regions of magnetic particles. This accurate technique enables MO media to pack in a lot more information than other magnetic devices. Once heated the magnetic particles can easily have their direction changed by a magnetic field generated by the read/write head.

MO media

An MO disk is constructed by “spattering” a number of films onto a high-strength polycarbonate resin substrate base - the same material as used in “bullet-proof” glass - and coating the entire disk with an ultra-violet hardened protective resin. The recording film itself is made of alloy including a number of different metal elements, such as Tb (terbium), Fe (iron) and Co(cobalt). This is sandwiched between protective dielectric films which, as well as providing thermal insulation, also enhance the rotation of polarisation angle so that the sensor is better able to detect the “Kerr Effect”.

SUPER-FLOPPIES (ZIP)

Without doubt, the most popular device in this category is Iomega’s Zip drive, launched in 1995. The secret of the Zip’s good performance (apart from its high 3,000rpm spin rate) is a technology pioneered by Iomega (based on the Bernoulli aerodynamic principle) which actually sucks the flexible disk up towards the read/write head rather than vice-versa. The disks are soft and flexible like floppy disks, which makes them cheap to make and less susceptible to shock.

The Zip has a capacity of 94MB and is available in both internal and external versions. The internal units fit a 3.5in bay, come with a choice of SCSI or ATAPI interface and are fast, with an average 29ms seek time and a data transfer rate of 1.4 KBps. External units originally came in SCSI or parallel port versions only. However, the Zip 100 Plus version, launched in early 1998, offered additional versatility, being capable of automatically detecting which of these interfaces applied and operating accordingly.

The 200MB to 300MB range is best understood as super-floppy territory. This is about double the capacity of the would-be floppy replacements with performance more akin to a hard disk than a floppy disk drive. Drives in this group use either magnetic storage or magneto-optical technology. The magnetic media drives offer better performance, but even MO drive performance, at least for the SCSI versions, is good enough to allow video clips to be played directly from the drives.

In the summer of 1999 Iomega altered the landscape of the super-floppy market with the launch of the 250MB version of its Zip drive. Like its predecessor, this is available in SCSI and parallel port versions - the latter offering sustained read performance around

twice the speed of the 100MB device and sustained write speed about 50% faster. The actual usable capacity of a formatted 250MB disk is 237MB - explained by the fact that, like most hard drive and removable media manufacturers, Iomega's capacity ratings assume that 1MB equals 1,000,000 bytes rather than the strictly correct 1,048,576 bytes. The Zip 250 media is backwards compatible with the 100MB disks, the only downside being the poor performance of the drive when writing to the older disks.

FLASH DRIVE

Pen drive, Thumb drive or Flash drive is a small, portable flash memory card that plugs into a computer's USB port and functions as a portable hard drive.



Fig 8.10 A typical USB drive

USB flash drives are compact file storage devices to save your information externally. Unlike most removable drives, a USB drive does not require rebooting after it's attached, does not require batteries or an external power supply, and is not platform dependent. USB drives are available in capacities ranging up to about 65 gigabytes (GB), depending on manufacturer, in a corresponding range of prices.

USB flash drives have less storage capacity than an external hard drive, but they are smaller and more durable because they do not contain any internal moving parts.

HISTORY OF USB

IBM started selling the USB flash drives in December 2000, using the DiskOnKey product (Product of an Israeli company M-Systems). This had a storage capacity of 8 MB, making it five times the capacity of a floppy disk. At the same time Trek Technology out of Singapore was selling a model called "ThumbDrive". Later in 2004, a court in United State revoked one of the Trek's patent. Since then, the development and technology of the USB storage devices has continued to improve. Flash drives are removable and rewritable, and are a relatively dense form of storage. They are still less expensive and will hold more data than a CD/DVD.

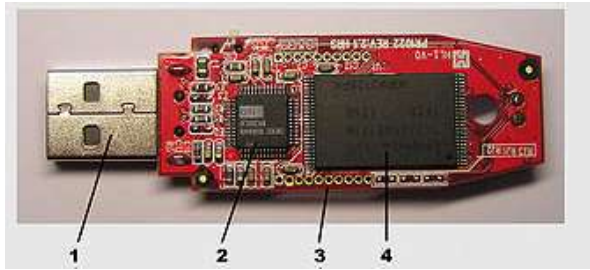


Fig 8.11 Internal Circuit Board of USB

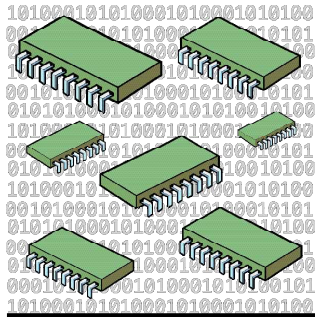
- 1. USB connector 2. USB mass storage controller device
3. Test points 4. Flash memory chip**

Modern flash drives have USB 2.0 (Second generation USB) connectivity. However, they do not currently use the full 480 Mbit/s (60MB/s) the USB 2.0 Hi-Speed specification supports due to technical limitations inherent in NAND flash. The fastest drives currently available use a dual channel controller, although they still fall considerably short of the transfer rate possible from a current generation hard disk, or the maximum high speed USB throughput.

The amazing technology of a USB storage device has shown that it can make it through a washing machine and retain its memory in most cases after drying thoroughly. Some companies have done testing in which the device was run over by a Jeep, frozen in dry ice, submerged in acid, and fired with a mortar and the little wonder has passed under all the extreme conditions.

EXERCISE

1. Fill in the blanks
 - a) MIC feature is available in _____ Technology.
 - b) The number of the pits on a CD-ROM is about _____.
 - c) The track width on CD-ROM is _____.
 - d) The tpi of CD is _____ whereas that of Floppy is _____.
 - e) The two type of recordable media in CD-ROM are _____ and _____.
 - f) Data can be rewritten on _____ type of media.
2.
 - a. How are data stored on magnetic tape?
 - b. How are these data retrieved?
 - c. What are some advantages and limitations of magnetic tape?
3. Which medium, tapes or disks, allows faster access to data, and why?
4. State two advantage of magnetic tape storage over punched-card storage.
5. Briefly describe and discuss the difference between -
 - (a) a sequential-storage medium, and
 - (b) a random-access storage medium.
6. Explain the following two terms and give an example of a medium suitable for each :
 - (a) random access storage
 - (b) serial access storage
7. What are the various type of Tape media that is available?
8. Discuss how and how much data is stored on CD-ROM.
9. Discuss the various advantages of CD-ROM over other media.
10. Discuss the construction of CD-RW media and how data is stored on the media?
11. Explain the meaning of 24X 24X 52X configuration.
12. What are the various interfaces available for CD-ROM?
13. What is DVD and how its capacity increases in the same area as of CD-ROM?
14. Write short notes on MO technology and ZIP drive technology?



CHAPTER 9

KEYBOARD & MOUSE

KEYBOARD
KEYBOARD TECHNOLOGIES
KEYBOARD INTERFACE
MOUSE

Keyboard & Mouse

KEYBOARD

The Input devices are the means by which you feed information into your computer, the primary means by which you interact with your personal computer. The various available devices span an entire range of technologies, although they work in different ways, all accomplish the same task, allowing you to communicate with your computer. While most people use a keyboard as their principal means of communication to their computer, we can also use a mouse, tablet or trackball to communicate our commands and to draw directly on the screen. All input devices are similar in one respect : they issue a signal that the computer can interpret as either a command or as a usable data. The keyboard is the most friendly input peripheral. Both program and data can be keyed in through it. In addition, certain commands to software can be given from keyboard.

HOW KEYBOARDS WORK

In principle, all keyboards work in the same way. They are a collection of switches. When you press a switch, an electrical signal is sent to the computer. That signal is interpreted as a binary value. The specific value will be looked up in a special table of characters, and the corresponding character or action in that table is then represented on the screen and stored in the computer's RAM.

The keyboard controller in your PC receives serial data via the cable from your keyboard, checks the parity, translates the keyboard scan codes and places the result (as a byte of data) in its data port (I/O address 60H).

Upon receiving the information from the keyboard that a key has been pressed, the keyboard controller sends an interrupt signal on the IRQ1 line to the interrupt controller to start the process that tell the microprocessor that a valid keystroke has occurred.

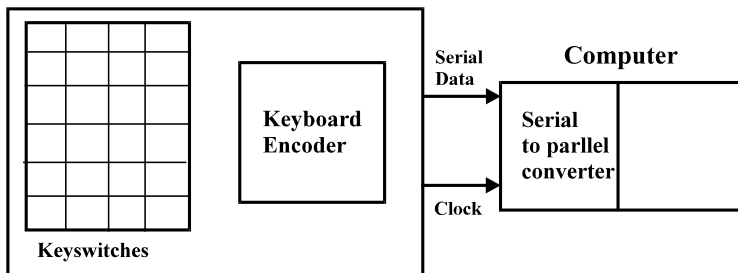


Fig 9.1 Serial Keyboard

There are two types of Keyboards. A serial keyboard sends the data bit by bit, in a serial fashion as shown in figure. The computer converts

the data into a parallel byte. A parallel keyboard sends the data as a byte in parallel form, all the bits are sent simultaneously on different lines (wires). The cable between the keyboard and the computer should have more wires in a parallel keyboard, as shown in figure.

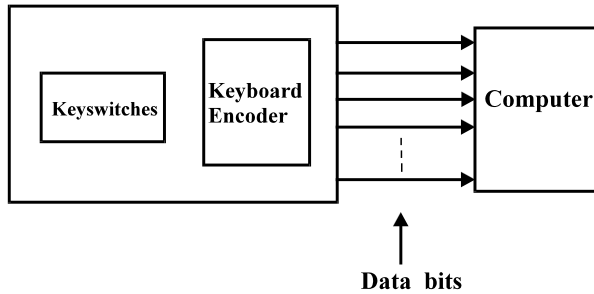


Fig 9.2 Parallel Keyboard

The PC XT Keyboard

IBM has offered nearly a dozen different keyboards with its range of Personal Computers. The first keyboard supplied with PC / PC XT were of the 83 key variety. It put two vertical rows of function keys at the left of the main alphanumeric keypad and forced cursor controls to share the same keypad with a calculator style array of numbers for direct data entry. The Enter key was small and ambiguously identified with a bent arrow legend, and no indicators were provided for the three locking shift keys: Caps Lock, Num Lock, and Scroll Lock.



Fig 9.3 XT KEYBOARD

The complaints about the original design concerned mainly the layout of peripheral keys. The left hand function keys did not correspond to the bottom-of-the-screen listings of function key assignments used by most programs.

The lack of indicators led to a lot of mistyping of numbers for cursor movements and capital letters for lower case. Spreadsheets needed both a numeric keypad and cursor-control keys. Also, the Enter key was too small.

The PC AT Keyboard

Later IBM introduced a new keyboard layout with the AT. Compared to the PC keyboard, the PC AT keyboard had a new key called SYS REQ, pressing this key and releasing this key causes generation of a software interrupt by the keyboard BIOS. In addition, three status lights Caps Lock, Scroll Lock and Num Lock keys can be controlled by software. Keyboard sends data only when both KB CLK and Keyboard data KBDATA lines are high.

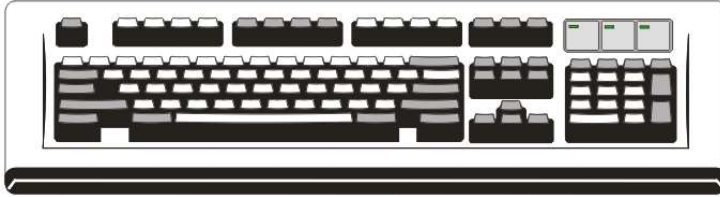


Fig 9.4 AT Keyboard

Keyboard data format in an AT is different from that of an XT. The data format is 11 bits wide and has one start bit, scan code of the depressed/released key (odd) parity bit and a stop bit. In fact, the difference was much deeper. Unlike the PC keyboard, the AT keyboard was made programmable and was given its own set of commands, which could be relayed to it through the system unit. This fact alone made the AT keyboard incompatible with PC and XT system units. Although the connectors are the same, a PC/XT keyboard won't work when plugged into an AT and an AT keyboard won't work when plugged into a PC or XT.

ENHANCED KEYBOARDS

With the introduction of upgraded AT, IBM offered another new keyboard, termed the Advanced keyboard by IBM, commonly also called the Enhanced keyboard. The advance it made was a greater endowment of keys, to a total of 101 in the standard international models. The keys additions were several. A new, dedicated cursor-control was provided, separate from the combined numeric and cursor pad, and several other control keys were duplicated in another small pad. Two new function keys F11 and F12 were added, and the whole dozen were moved to a top row, above and slightly separated from the alphanumeric area. Duplicate Ctrl and Alt keys were provided at either side of the space bar, and Caps Lock was moved to the former location of the Ctrl key.

COMPATIBLE KEYBOARDS

Compatible computer makers have striven to keep pace with IBM and have adapted their keyboards to the prevailing standard. Thus they have followed suit in adopting the advanced design, its drawbacks notwithstanding. Many third-party vendors introduced IBM compatible keyboards are designs that offer a redundant set of

function keys across the top of the board as well as in a pair of columns along the left side. Various designs have sought to isolate a set of cursor arrow keys from the numeric keypad. Several keyboards also include extra key caps to allow you to swap the position of the left hand Ctrl and Caps Lock Keys so that Ctrl falls in its more familiar position. The electrical relocation is handled either through another keyboard switch or through software that you must run on the host computer.

Several manufacturers have introduced devices that incorporate a trackball or some altered form of a mouse to the keyboard itself. These keyboards include a special cable that incorporates both the keyboard wiring and a serial circuit for the mouse or trackball; the cables end with two plugs that must be attached to different ports on your computer.

One improvement made by many compatibles manufacturers is the inclusion of a compatibility switch, usually located on the bottom of the keyboard. The two positions of this switch allow you to select the electrical compatibility of the keyboard between the PC/XT and AT standards. One keyboard can thus be used with either type of system unit, provided you've put the switch in the correct position.

KEYBOARD TECHNOLOGIES

No matter how their keys are arranged, all keyboards have the same function, detecting the keys pressed down by your fingers and relaying this information to your computer. Even though two keyboards may look identical, they may differ considerably in the manner in which they detect the motion of your fingers. The technology used for this process- how the keyboard works electrically- can effect the sturdiness and longevity of the keyboards.

Several technologies have been used for keyboards, the two most common designs in PCs are the capacitive and hard contact keyboards.

CAPACITIVE KEYBOARDS

Capacitive keyboards are generally built around an etched circuit board. Two large areas of tin-and-nickel-plated copper form pads under each switch key. The pads of each pair are neither physically nor electrically connected to one another.

Pressing any key on the keyboard forces a circle of metalized plastic down, separating a pair of pads that lie just below the key plunger. Although the plastic backing of the circle prevents making a connection that would allow electricity to flow between the pads, the initial proximity of the pads results in a capacity change. Separating them causes a decreasing to 2 to 6 picofarads. The reduction of capacitance causes a small but detectable current flow in the circuitry leading to the pads. When the key cap is released, the spring expands, returning the key cap to its original height and moving the plunger away from

the metallic pads, causing the current- flowing through the circuits leading to the pads-to return to its original level.

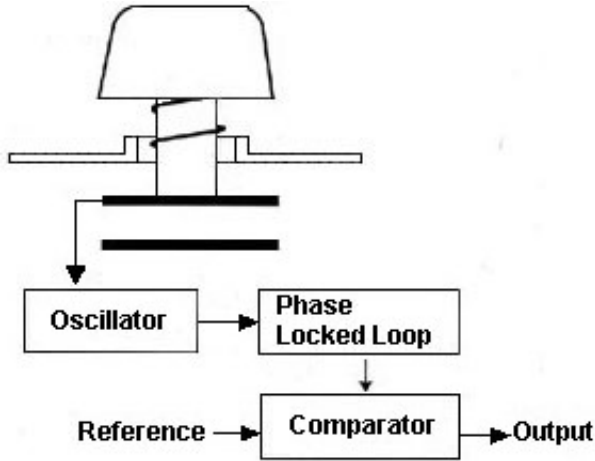


Fig 9.5 Capacitive Switch

HARD-CONTACT KEYBOARDS

Many of the keyboards used by today's compatible computers are of a very similar hard contact design. In a hard-contact keyboard, pressing the key cap collapses a foam rubber dome. The collapsed dome presses against a sheet of plastic on the bottom of which is a metallic area connected to the rest of the keyboard's circuit board. The metallic surface contacts a similar surface on another plastic sheet, permitting current to flow through the printed circuits connected to each of the pads.

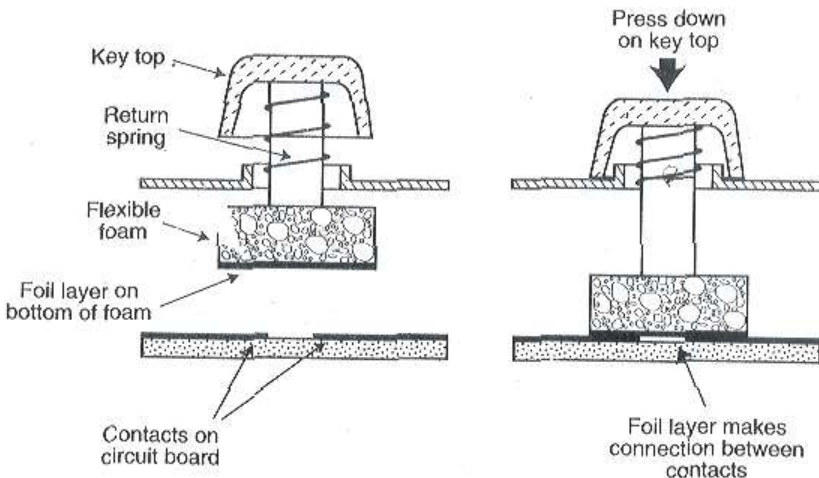


Fig 9.6 Mechanical Switch

When the key cap is released, the rubber dome springs back to its original shape, releasing pressure on the plastic sheet. The plastic also returns to its original position, breaking the electrical circuit and cutting off the flow of current.

KEYBOARD CABLING

The scan codes are sent from the keyboard to computer serially so only one wire conductor is needed to convey the keyboard data information. A second conductor is required to serve as a return path for the data signal, and, as a ground, it serves as common return for all the other circuits in the keyboard cable. To synchronize the logic in the keyboard with that in your computer, a separate wire is used for a keyboard clock signal. A fourth and final wire is used to supply the keyboard with the five-volt direct current power that it needs to operate. These four conductors are all that is necessary to link keyboard to computer.

Most PCs use standard five-pin DIN connectors for their keyboard connection.

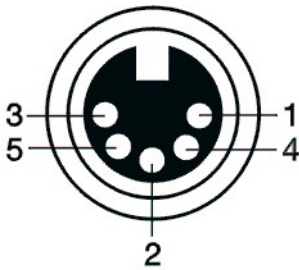


Fig 9.7(a) 5 Pin DIN



Fig 9.8 (b) Mini DIN 6 socket

- | | |
|-------|--|
| Pin 1 | for keyboard clock |
| Pin 2 | keyboard data |
| Pin 3 | assigned to carry a signal to reset the keyboard |
| Pin 4 | ground |
| Pin 5 | the five volt electrical supply |

Advanced keyboards generally use a modular AMP connector on the rear of the keyboard, allowing the cable itself to be easily replaced.

- | | |
|-------|----------------|
| Pin 1 | reserved |
| Pin 2 | keyboard data |
| Pin 3 | ground |
| Pin 4 | keyboard clock |
| Pin 5 | five volts |
| Pin 6 | reserved |

When looking at the gold contacts of the connector, the contacts are labeled in reverse alphabetical order from left to right.

The removable cable also allows one keyboard to serve both the old AT line and the new PS/2 line, which use different system-board input connectors. It has the following assignments.

KEYBOARD ORGANIZATION

There are several ways of organizing the key switches inside a keyboard. The matrix organization is the popular technique followed in large keyboards used in computers.

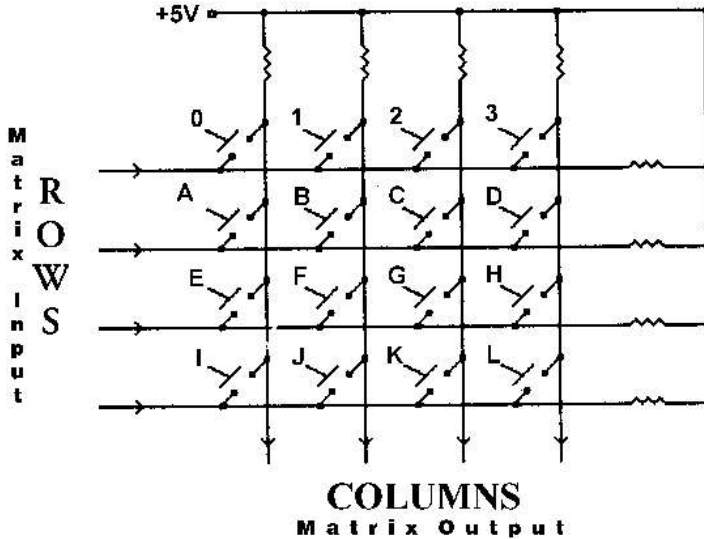


Fig 9.9 Keyboard Matrix

A matrix keyboard in which the keys are organized in rows and columns. Each key has a unique set of coordinates. A row number and a column number. When a key is pressed, the keyboard electronics has to determine the pressed key's row number and the column number. This encoding is performed in matrix keyboards by scanning technique.

The scanning method uses rows as inputs to the matrix and the columns as the outputs from the matrix. It scans the matrix row by row to determine the key press. A scan counter is used to scan the matrix. An eight bit counter is necessary for scanning a 16x16 matrix. The four MSBs D-4 D-7 are decoded to activate one of the rows at a given time. The column outputs are connected as inputs to the multiplexer. The four LSBs of the scan counter D0-D3 are the select inputs to the multiplexer.

According to this bit pattern, one of the column outputs is selected by the multiplexer and made available as key pressed signal. Initially the counter is cleared. Now a low input is applied to R0 and high input is applied to all other row inputs. The C0 is selected by the multiplexer.

If the key R0 in row 0 and column is not pressed, C0 will be high. If it is pressed, C0 will be low. Accordingly, the key pressed signal is high or low. When key pressed signal is high, it advances the scan counter by 1. Now the counter value will be 000000001. As a result the low input at R0 will be received at C1 through the closed switch 01.

The key pressed signal becomes low indicating that a key press has been detected. Now the counter is frozen. Its four MSBs D4-D7 give the row number and the four LSBs give the column number. There eight bits are transmitted to the computer by the keyboard electronics. Suppose no key in row 0 ks pressed, then the counter continues to advance. After reaching 00001111, it has completely scanned all the keys in row 0 and has not found a key press. Next it increments to 00010000. Now row 1 scanning commences. The R1 input is made low and all the other row inputs are made high. This method is known as rotating 0 since 0 is applied to the rows one by one in sequence.

In modern keyboards, a microprocessor serves as keyboard encoder. The microprocessor generates the row inputs as well as reads the column outputs. In this method, an output port is used to generate row inputs and an input port is used to sense the column outputs. The firmware for the microprocessor does all the other required function.

The Keyboard and Scan Codes

The internal microprocessor in all IBM keyboards identifies the keys that are pressed and converts the information so derived into scan codes, which are then sent serially to the host computer. Regardless of which type of key cap is used, pressing it causes a change in the current flowing through the circuits associated with that key cap.

A microprocessor such as the Intel 8048, built into the keyboard constantly scans circuits leading to the key caps. It detects the increase or decrease in current, the processor can tell both when a key has been pressed and when it's been released. Each key has a unique set of codes, even if to the users, the keys seem identical. The processor can, for example distinguish between the left and right shift keys. To distinguish between a real signal and an aberrant current fluctuation, the scan is repeated hundreds of times each second. Only signals detected for two or more scans are acted upon by the processor.

Depending on which key's circuit carries a signal to the microprocessor, the processor generates a number, called a scan code. There are two scan codes for each key, one for when the key is pressed and the other for when it's released. The processor stores the number in the keyboard's own memory buffer, and it loads the number in a port connection where it can be read by the computer's BIOS. Then the processor sends an interrupt signal over the keyboard cable to tell the processor that a scan code is waiting for it.

An interrupt tells the processor to drop whatever else it is doing and to divert its attention to the service requested by the interrupt. The BIOS reads the scan code from the keyboard port, and sends a signal to the keyboard that tells the keyboard it can delete the scan code from its buffer. If the scan code is for one of the ordinary shift keys or for one of the keys that are considered to be special shift keys and toggle keys- Ctrl, Alt, Num Lock, Caps Lock, Scroll Lock, or Insert- the BIOS changes two bytes in a special area of memory to maintain a record of which of these keys has been pressed.

KEYBOARD INTERFACE

The keyboard interface receives the scan code in a serial format from the keyboard, assembles the serial data into parallel eight bit scan code, and generates interrupt request you the interrupt logic. The system software follows specific protocol with the keyboard microprocessor for the data transfer and control sequences. The keyboard interface hardware consists of following functional sections.

1. Serial to parallel converter
2. Interrupt generation
3. Scan code port (port A PPI)

Key board interface logic figure shows the block diagram of keyboard interface. The serial to parallel converter is enabled by software. Once the scan code is assembled, interrupt request is generated and this interrupt request freezes the shift register. The interrupt service routine enables the shift register output which is connected to the input port A of PPI. The service routine also clears the shift register so as to prepare it for next scan code.

Bidirectional Communication

The keyboard microprocessor transmits the scan code to the PC system over a serial interface. There are two bidirectional lines between the keyboard and the mother board. One line is for serial DATA and other line for the clock signal. The four signals are +5v, Gnd, DATA and CLOCK. These signals are carried by a shielded four wire cable which is flexible coiled wire. In some keyboards fifth wire also is there which gets the reset signal from mother board. There is no standard connector on the keyboard side, But in the PC side, all the keyboards use a five pin DIN connector (male). The keyboard sends the scan code by transmitting a clock signal and an eight bit serial data. Before transmitting the scan code, the keyboard microprocessor checks the status of both CLOCK LINE DATA LINE. Only if both the lines are high the keyboard can send DATA.

Circuit Description

The LS232 is a special purpose IC functioning as 8 bit bidirectional shift register. It is configured as a serial to parallel shift register with tri state output. The D0, SE, DS, S/P are pulled high. The KB CLK signal enters at the D input of the LS175. Two D FLIP FLOPS in this

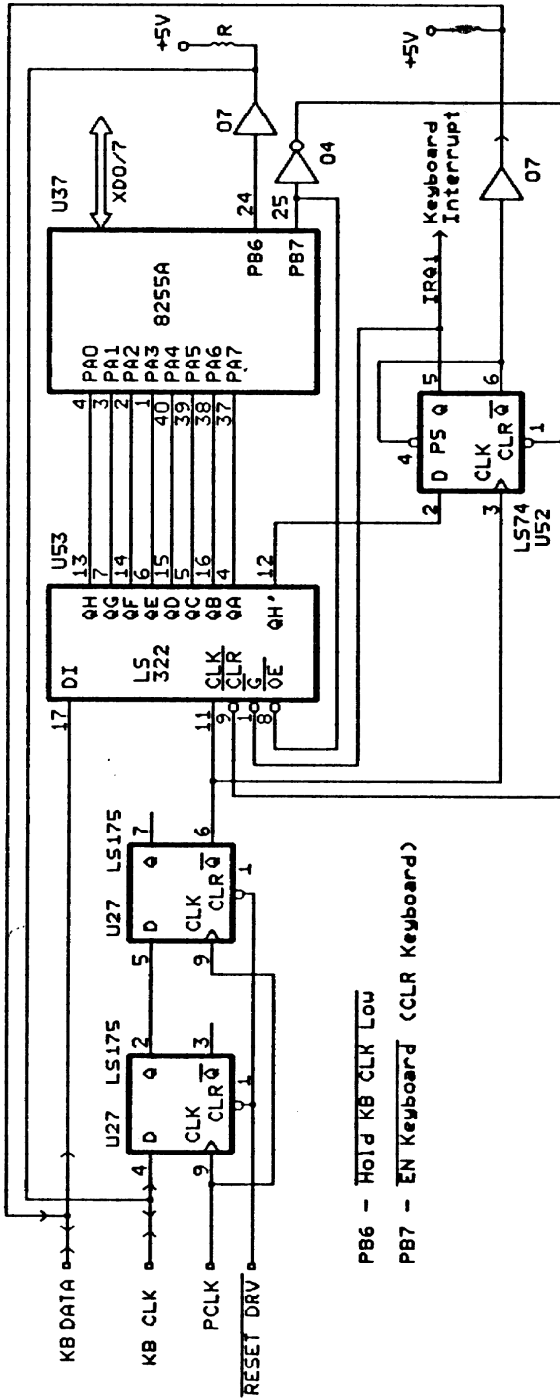


Fig 9.10 KEYBOARD LOGIC

chip are used for keyboard interface. The PCLK signal is a clock input to the LS175. The Q output of the first flip flop is connected to the D input of the second flip flop. The Q output of the second flip flop goes to the CLK input of the shift register and pin 3 of the flip flop LS74. The KB DATA signal is connected to the DI input of the shift register. During reset sequence, the RESET DRV goes low and clears both the D flip flops in the LS175. Initially, POST firmware makes PB7 high. The shift register output is tri stated since PB7 output is connected to OE input of the shift register. The PB7 is inverted by the LS74) and the inverted output is at pin 12 being low clears both the shift register and the IRQ1 flip flop LS74. Thus initially IRQ1 is low. The IRQ1 is connected to G of the shift register. Hence shift register operation is now enabled. Subsequently the POST firmware makes PB7 low so that the reset signal to the shift register and flip flop is removed. Since the shift register is enabled, it scans the DI input for each clock received at CLK input. The bit 0 or 1 at this time is put inside the shift register in the QA position. When the next clock comes, the new data bit is put in QA position and the older bit is shifted to QB position.

As each new data bit is received, the already received bits are shifted right by one position. When eight bits are received the already received bits are shifted right by one position.. When eight data bits are received in this way, the QH output goes high. The next shift clock sets the IRQ1 flip flop, LS74. The shift operation is immediately disabled since the G becomes high. The IRQ1 is connected to the interrupt controller as the level 1 interrupt. When 8088 branches to ISR, the following sequences are done by the keyboard service routine.,

1. Reads the scan code through PORT A of PPI. The scan code now can be converted into ASCII code by software.
2. PB7 is made high and performs three operations.
 - a) Tri stating the outputs of LS322.
 - b) Clearing the IRQ1 flip flop.
 - c) Clearing the shift register contents.
3. PB7 is made low. Now the shift register outputs are enabled. The reset signal to the IRQ1 flip flop and the shift register are removed. The above sequence is repeated for every code. When IRQ1 is set, the 8088 may not honour it immediately. Depending on the mask status of the IRQ1 in 8259 and the interrupt enable flag of 8088, the IRQ1 will be ignored. The 8088 software can force the KBCLK signal low by making PB low. This disables the keyboard interface operation. The keyboard performs an internal rest and self test.

MOUSE

On most computer systems there are two standard ways to input information - the keyboard and the mouse. A mouse is a pointing device used to point at and select items on-screen. There are other

kinds of pointing devices, including joysticks, track balls, light pens, touch screen and graphic tablets etc.

The mouse and its cousin, the trackball, have survived those other, more awkward methods of navigating with the keyboard. Mice can never replace the keyboard, but they can supplement the keyboard by doing tasks such as moving and pointing to on-screen objects, tasks for which the cursor keys are ill-suited. Until we reach the point where we simply talk to our PCs, mice will be an integral part of our systems.



Fig 9.11 Mouse has either two-button or three buttons

A track ball is simply a mouse mounted so that the ball is rotated with your fingers instead of on the surface of your desk. Most mice consist of nothing more than a plastic shell with two or three buttons built into it. Out of one end of the mouse is a cord that connects with your PC. This cord is used to send information to the PC about movement and button clicks. If you turn the mouse over, on the bottom is a small ball that rolls as you move the mouse across a surface. This provides the movement information that the mouse needs in order to operate. The best roller balls are made of a textured rubber that provides a bit of friction as the mouse moves.

Mouse come in two basic versions - Opto-electronic Mouse and Optical Mouse.

OPTO-ELECTRONIC MOUSE

The Opto-electronic mouse has become the most popular pointing device for the newest breed of operating environments-graphic interfaces represented by Windows, etc. With the mouse, you control your PC by pointing to images instead of typing in commands. Here's how the mouse translates the movements of your hand into the actions on screen.

Mouse has a ball of steel for weight and rubber-coated for grip. When you move mouse, it rotates and it drives two rollers, which touch it internally, one each for x and y displacement. A third spring-loaded roller holds the ball in place against the other two.

These rollers then turn two disks with radial slots cut in them. Each disk rotates between a photo-detector cell, and each cell contains two offset light emitting diodes (LEDs) and light sensors. As the disk turns, the sensors see the light appear to flash, showing movement, while the offset between the two light sensors shows the direction of movement.

Also inside the mouse are a switch for each button, and a microcontroller which interpret the signals from the sensors and the switches, using its firmware program to translate them into packets of data which are sent to the PC. Serial mouse use voltages of 12V and an asynchronous protocol from Microsoft comprised of three bytes per packet to report x and y movement plus button presses. PS/2 mice use 5V and an IBM-developed communications protocol and interface.

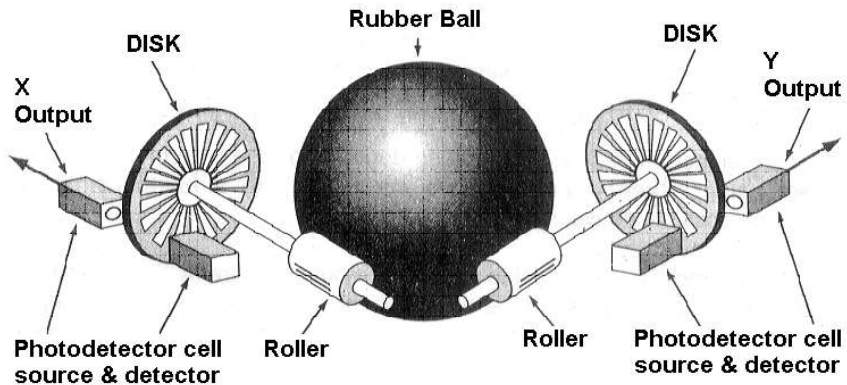


Fig 9.12 Opto Mechanical Mouse Mechanism

OPTICAL MOUSE

The Optical mouse uses a special pad with a reflective surface and grid lines to detect motion and transform it into cursor movements. Lines in one direction are blue, those in other black. On the bottom of the mouse are two LED's that emit light and two photo transistors that receive the light bounced back. Each break in the light signal represents a defined distance. The color and number of breaks in the light determine direction and distance.

Developed by Agilent Technologies and introduced to the world in late 1999, the optical mouse actually uses a tiny camera to take 1,500 pictures every second. Able to work on almost any surface, the mouse has a small, red light-emitting diode (LED) that bounces light off that surface onto a complimentary metal-oxide semiconductor (CMOS) sensor.

The CMOS sensor sends each image to a digital signal processor (DSP) for analysis. The DSP, operating at 18 MIPS (million instructions per second), is able to detect patterns in the images and see how those

patterns have moved since the previous image. Based on the change in patterns over a sequence of images, the DSP determines how far the mouse has moved and sends the corresponding coordinates to the computer. The computer moves the cursor on the screen based on the coordinates received from the mouse. This happens hundreds of times each second, making the cursor appear to move very smoothly.

The absence of moving parts means the mouse's traditional enemies - such as food crumbs, dust and grime - are all but completely avoided. The Optical Mouse works on nearly any surface, such as wood, paper, and cloth - although it does have some difficulty with reflective surfaces, such as CD jewel cases, mirrors, and glass.

Optical mice have several benefits over wheeled mice:

- No moving parts means less wear and a lower chance of failure.
- There's no way for dirt to get inside the mouse and interfere with the tracking sensors.
- Increased tracking resolution means smoother response.
- They don't require a special surface, such as a mouse pad.

WIRELESS MOUSE

The mouse is an extremely popular technology to interact with your computer. While the majority of mice are connected directly to the computers peripheral input via a cord and are powered by the computers main power, wireless mice are becoming extremely popular to give computer users cordless accessibility to their mice and to interact with their computer.

Wireless mice usually work via radio frequencies commonly referred to as RF. RF wireless mice require two components to work properly a radio transmitter and a radio receiver.

The radio frequency (RF) receiver usually connects to the computer's peripheral mouse input. It receives these RF signals, decodes them and then sends these signals directly to the computer as normal. RF receivers usually come in a few styles. The majority come as built in components that connect to the mouse input, others can be a separate card that is installed in one of the many expansion slots of computers and the third type of receiver is a separate unit that is then connected to a cable going directly to the computer's peripheral input. Since the technology has been mastered, most wireless mice have integrated receivers that plug into a computer's peripheral input and are very small in size.

802.11b or 802.11g

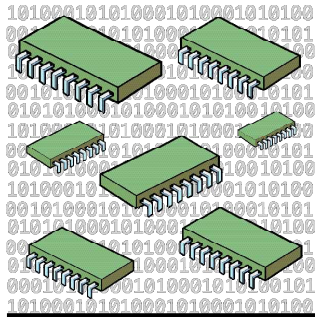
Wireless mice mainly use Radio frequencies to transmit data from the mouse to the computer. The most common type of RF used is the 802.11b or 802.11g. These frequencies operate at 2.4 gigahertz and at these frequencies a mouse can transfer data at pretty quick speeds, usually either 11 Mbps or 56 Mbps.

Bluetooth RF

Another technology that is emerging and now beginning to be used in large numbers for wireless mice is Bluetooth RF technology. Bluetooth is commonly known to connect computers to peripherals such as keyboards, printers, PDA's, and headsets. Bluetooth is similar to 802.11b and 802.11g in that it uses 2.4 gigahertz frequencies, however, it also uses software called adaptive frequency hopping to choose frequencies that have no or little interference.

EXERCISE

1. State True or False:
 - a) All input devices issue a signal that the computer can interpret as either a command or as a usable data.
 - b) Mouse controller in the computer receives parallel data.
 - c) In capacitive keyboards, pressing the key cap collapses a foam rubber dome.
 - d) When you move mouse, two rollers touches the surface on which you have kept the mouse and rotates the rollers.,
 - e) In Intellimouse, ball and rollers also exist.
2. What is keyboard?
3. What was the method used for inputting data when keyboards were not there?
4. Differentiate between XT and AT keyboard.
5. Differentiate between PS/2 and 5 pin DIN socket.
6. Explain the working of Capacitive keyboard. How its working is different from Hard contact keyboard.
7. What is keyboard matrix?
8. What do you understand by Bi-directional communication?
9. What is mouse? How it works?
10. What is Opto-electrical mouse? How it's different from Optical Mouse?



CHAPTER 10

SMPS

SWITCHED MODE POWER SUPPLY
POWER MANAGEMENT
POWER SUPPLY CONNECTOR
POWER SUPPLY TECHNOLOGIES

SMPS

SWITCHED MODE POWER SUPPLY

One of the most important components in PC is the power supply. Power supplies convert potentially lethal 110-115 or 220-230 volt alternating current (AC) into a steady low-voltage direct current (DC) usable by the computer.

In India, the power supply in a PC is designed to convert the 220 V AC into something (DC) that the computer can use, which is specifically both 5V and 12V DC. Usually, the digital electronics components and circuits of the system (MOTHER BOARDS, ADAPTER CARDS, DRIVES LOGIC CARDS) use 5V while the motors (DRIVES FAN) use 12V supply. A power supply is rated by the number of watts it generates.

One thing that the power supply does to ensure that the system does not run without proper power levels is to prevent the computer from starting up until all the correct power levels are present. Each power supply completes internal checks and tests before allowing the PC to start-up. The power supply sends to the motherboard a special signal called " POWER GOOD " signal. If the signal is not there the computer will not run. The effect of this setup is that when the AC voltage dips, the power supply becomes overstressed or overheated, the power good signal goes down, forcing a system to reset or complete shut down.

The power good line is wired to the 8284 or 82284 clock generator chips, which control the clock and reset lines to the microprocessor.

Advantages of SMPS

- Low weight and smaller size come about because operation is at a significantly higher frequency range and the inductive elements are vastly smaller and much cheaper.
- Higher efficiency occurs because the power transistor is switched very rapidly between saturation (FULL ON) and cut off (FULL OFF) and there is very little power dissipation and it follows there are reduced heat sinking requirements.
- Wider AC input voltage range results from the flexibility in selecting the switching frequency and the transistor duty cycle which makes voltage adoption unnecessary.
- Reduced costs occur owing to the absence of large bulky power transformers, a huge reduction in volume and power dissipation, smaller material requirements and also smaller semiconductor devices.

POWER MANAGEMENT

Power management was designed for convenience as it easier to have your computer go into power standby and be able to press a key on your keyboard or move your mouse and instantaneously be back where you were and for saving power. This is especially important on portable computers when using the battery as your main power source.

Power Supply Form Factors

Currently in the industry there are eight power supply form factors. Each of these form factors can have various amounts of configurations and power output levels.

PC / XT	AT/Desk
AT/Tower	Baby AT
ATX	NLX
SFX	LPX

Out of these eight, the two different types of computer power supplies are widely used in present day computers. Your computer has one or the other: AT (left) ATX (right)



Fig 10.1 (a) AT SMPS



Fig 10.1 (b) ATX SMPS

The AT and ATX power supply can be largely identified by the motherboard power socket. Generally, AT socket is 12 Pin socket as shown in Fig. ATX socket is 20 Pin connector. ATX power supply supports the function of “**Soft Power On Momentary Switch**” which connects the front panel switch to the 2-Pin SOFT-PWR on the system board. While the power switch on the ATX power is turned on, the full power will not go into the system board until the front panel switch is momentarily pressed. Also the ATX power supply switches off the power as soon as you shut down the Windows.



Fig 10.2 (a) AT Socket



Fig 10.2 (b) ATX Socket

In AT SMPS, there is a cable for ON/OFF switch, where as in ATX SMPS, there is no ON/OFF, there os only an ON switch is provided for momentarily passing the full power to motherboard.



Fig 10.3 Inside view of an ATX SMPS

The AT power supply connectors are two six-pin male header connectors. Plug the dual connectors from the power directly onto the board connectors.

Most of power supply have two leads. Each lead has six wires. Two of which are black, orient the connectors, so the black wires are in the middle.

The black wires should be in the middle



Fig 10.4 AT SMPS connector

Pin	Wire Colour	Function	Connector
1	Orange	“Power Good”	P8
2	Red (XT No Wire)	+5V DC	P8
3	Yellow	+12V DC	P8
4	Blue	-12V DC	P8
5	Black	Ground	P8
6	Black	Ground	P8

7	Black	Ground	P9
8	Black	Ground	P9
9	Black	Ground	P9
10	Yellow	-5V DC	P9
11	Red	+5V DC	P9
12	Red	+5V DC	P9

POWER SUPPLY CONNECTOR

The following illustration is the typical female connector which would be used to connect to a device such as a CD-ROM or Hard Drive. This connector is referred to as a large Molex connector. Additionally to these type of connectors you may also find a small Molex which is generally used for the floppy disk drive.

Pin	Wire Color	Signal
1	Yellow	+12v
2	Black	Ground
3	Black	Ground
4	Red	+5v

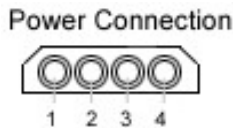


Fig 10.5 Power connection for Hard drives & CD-ROM

POWER SUPPLY TECHNOLOGIES

- Linear
- Switched Mode

Linear power supply and switched mode power supply differ mainly in the method of regulation of the output voltage. Linear power supply uses series or shunt stabilizer and SMPS uses switching transistor as main component for regulation. The way to regulate is to derive from the source more power than needed by the load even under worst conditions and then dissipate excess energy through heat sink. In the linear stabilizer a lot of current and voltage is present at the same time, excessive power is dissipated through heat to achieve stabilization. In SMPS, switch is substituted for linear stabilizer and it can be done in two ways.

1. When switch is open there is lot of voltage present but no current is flowing through it. So power dissipation through it is zero.
2. When the switch is closed, there is lot of current flowing through it but no voltage across it, so power dissipated is again zero.

In switching regulator, stable voltage is obtained by opening and closing the switch regularly. Average value of the output can be obtained by altering the ON/OFF time of switch and adding filter.

There are three basic types of converters :

- Flyback
- Forward
- Boost

Flyback Convertor

In flyback circuit magnetic field builds up in the inductor as long as switch remains closed. When the switch is opened, the inductor function as an energy source, the voltage across the inductor is reversed and the conducting diode passes the energy to the reservoir capacitor. The output voltage is reversed with respect to the input voltage.

Forward Convertor

Forward convertor does not reverse the placing of the input voltage. The capacitor is charged via the inductor when the switch is closed. The difference between the input and the output voltage is available on the inductor. Switch remains closed when the capacitor is charged and when opened, the magnetic field of the inductor is weakened through flyback diode.

Boost Convertor

In this, energy is stored in the inductor when the switch is closed. When switch is opened, this energy is supplied to the load at the output diode.

WORKING PRINCIPLE

In SMPS mainly supply is filtered and then rectified bridge rectifier to produce DC supply for the switching section. The input DC supply is switched (chopped) with the help of power convertor and the energy transferred to the output through a high frequency ferrite transformer. The switching section can be implemented in many different ways using switching transistor. Control section is implemented using special dedicated integrated circuit like TL494, SG3524, LM 3524 further additional circuits might be employed to provide current limit and over voltage protection.

CALCULATING PC POWER REQUIREMENTS

The computer system works on Direct Current of different voltages. Modern computer logic circuits operate by switching voltages with the two different logic states coded as two voltage levels, termed high and low. every family of logic circuits has its own voltage standards. Mostly PCs are built around the requirements of Transistor-Transistor Logic, which is most commonly referred to by its abbreviation TTL.

TTL circuits normally operate from a 5-volt supply. PC soften require other voltages as well. The motors of most disk drives, hard disks and floppy disks require 12 volts to make them spin. Other specialized circuits in a PC some times require bipolar electrical supplies. A serial port, for example, signals logic state by varying voltages between positive and negative in relation to ground. Consequently, the - 5 and -12 volts must be available inside every PC, at least if it hopes to use any possible expansion boards.

POWER SUPPLY FAN

The Power supply fan could be of two types ; 230-volts AC or 12-volts DC. The label on the fan will have the rating printed on it. In case the label is not present, the rating of the fan can be found out by measuring the voltage at the point inside the SMPS where the power for the fan is drawn i.e. the point where the wires from the fan are connected.

POWER SUPPLY CONNECTIONS

All Computer Power supplies have two kinds of connectors, two of them go to the system board, and the rest are designed to mate with tape and disk drives.

Mass Storage Power - The tape and disk drive connectors supply 5 and 12 volts to operate those devices. The connectors come in two sizes, which are polarized so that you cannot install them improperly.

System Board Power - The Two system board power connectors on standard power supplies are not identical. Each has its own repertoire of voltages. On most PC power supplies, these connectors are labeled P8 and P9. The lower number attaches to the matching connector on the PC system board, typically the one nearer the rear of the chassis.

POWER LINE IRREGULARITIES

Over voltages - The deadliest power line pollution is overvoltage, lightning like high potential spikes that sneak into PC and actually melt down its silicon circuitry. Often the damage is invisible, except for the very visible lack of image on Monitor. Other times, you can actually see charred remains inside computer as a result of overvoltages.

Under voltages - An undervoltage occurs when your equipment gets less voltage than it expects. An undervoltage can range from sags, which are dips of but a few volts, to complete outages or blackouts. Duration can vary from nearly instantaneous to hours.

Noise - Noise is a nagging problem in the power supplies of most electronic devices. It comprises all the spurious signals that wires pick up as they run through electromagnetic fields. In many cases, these signals can sneak through the filtering circuitry of the power supply and interfere with the signals inside the electrical device.

CIRCUIT EXPLANATION

The power supply circuit is using Pulse width modulation (PWM) technique. In this circuit main power supply is applied to main filter through fuse and a thermistor to protect power supply from any high current, and transients. Further it is given to rectifier section in which 4 nos. IN5402 diodes are used as bridge rectifier. The DC thus obtained is given to switching transistors C3039 (Q2 and Q3) through DC filter to smother out the ripples. In this PWM TL494 IC controls output voltage, maximum output power and the pulse width. The switching transistors thus derives high frequency transformer T1.

The secondary of this transformer provides different step down voltages which are rectified, filtered. This provides +5 V/15 Amp, +12V/4A, -5V/0.5A and -12V/0.5A in which +5V is the main controlled output.

A power good signal is generated about 100 ms after the source voltages have reached their minimum sense level to tell the computer that all voltages are OK and computer can operate normally. The failure of power good output of SMPS will stop your PC working.

What to look for in case of problem

1. Fuse blown : Check input bridge rectifier (D3) diodes IN5402 and switching transistors Q4 and Q5 and diodes across Q2, Q3 collector emitter.
2. Fuse Okay but no output - Check the IC pin 12 (17U), check all output diodes (At the output of T5 transformer) check transistor Q1 in PG circuit.
Check the wave forms at switching transistor base-emitter, collector-emitter.
3. The SMPS Fan not functioning - Open the SMPS and check for dry solder at the points where the fan is connected. Or Fan has to be replaced. Make note of the type (230V AC or 12V DC) and replace with an equivalent.

POWER SUPPLY

The power supply converts the AC voltage to dc voltage. The essential components of a basic power supply are shown in figure. The transformer does the functions listed below :

1. Isolation from the main supply
2. Reducing the voltage level from 220V or 110 V to 5V or 12V as required (step down). The rectifier converts the AC voltage at the transformer output into dc voltage. The filter removes the ripples in the rectifier output. The regulator maintains a constant output voltage in spite of the variations in the input voltage, load currents and component characteristics due to change in temperature.

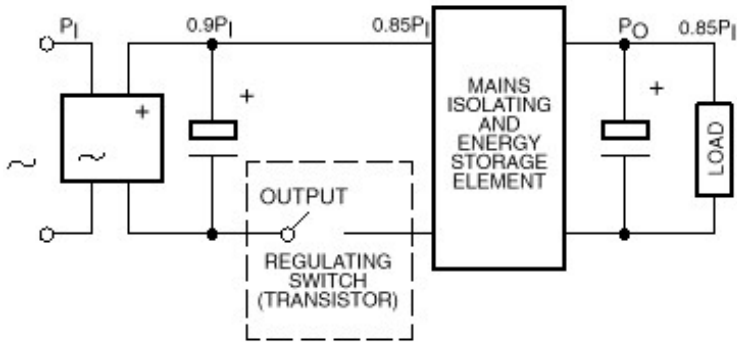


Fig 10.6 Basic Power Supply Unit

The linear power supply and the switch mode power supply differ mainly in the method of regulation of the output voltages. In a linear power supply more power than required by the load is derived from the source. The surplus power is dissipated across the regulator. By controlling the amount of surplus power being dissipated, the output voltage is maintained constant even when the load current varies. Thus there is a considerable amount of power loss resulting in poor efficiency. In the SMPS, a switching transistor is used in the place of a linear stabilizers.

The transistor is opened and closed periodically (i.e., switched ON and OFF) at a high frequency. In this principle, there is no power loss due to stabilisation, because there is no power dissipation across the transistor, whether it is open or closed. In short, there is no continuous output but a pulse waveform of high frequency. The switching rate is of the order of 20kHz. The output filter smoothens the output voltage. The dc to AC conversion by the switching logic is present in any SMPS.

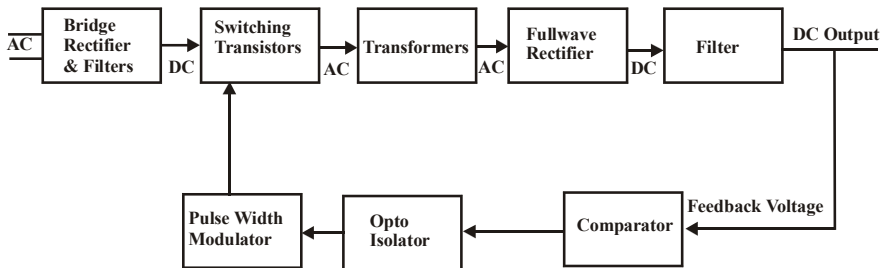


Fig 10.7 SMPS with Primary Side Control

Two different types of switching logic control are followed in SMPS designs. One approach is Pulse Width Modulation (PWM) technique and the second approach is resonant mode approach. The PWM switch mode technique is commonly followed in SMPS used in PCs. In the PWM principle, a dc voltage is chopped at a high frequency. The

resultant square wave voltage is stepped up or stepped down by using a transformer. The transformer's secondary output is rectified and filtered. The regulation of the output voltage is achieved by varying the duty cycle of the square wave. Thus in an SMPS the output voltage is compared to the reference and by controlling the duty factor of the drive signal applied to the switching circuit, the output is maintained constant. The voltage control circuit can be incorporated on either side of the transformer.

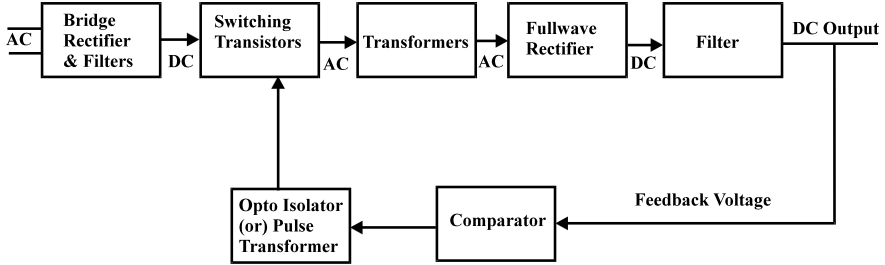
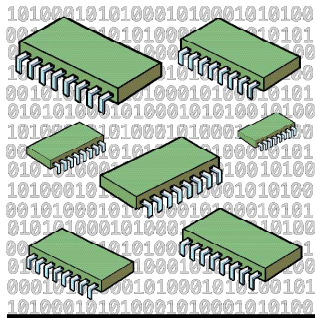


Fig 10.8 SMPS with secondary side control

The bridge rectifier converts the AC voltage into DC voltage. The filter removes the ripples at the output voltage. The switching transistor converts the DC voltage into square wave AC. The transformer steps down the voltage. The AC square wave is rectified by the full wave rectifier. The DC output from the rectifier is filtered by the filters. The DC voltage is compared with a reference voltage by the comparator. The error voltage at the comparator output is fed to the pulse width modulator. The opto-isolator provides isolation for the control loop against any pick up from main path. The output of the optoisolator is amplified at by the driver stage. The driver output controls the switching transistor.

EXERCISE

1. State True or False :
 - a) A power supply is rated by the number of watts it generates.
 - b) Cabinets with ATX power supply must have an ON/OFF switch.
2. What is SMPS? Discuss its advantages.
3. Differentiate between AT and ATX power supply. How can you recognize whether a power supply is AT or ATX by simply looking at it?
4. What is a converter? How many type of converter are there?
5. What is Pulse width modulation?



CHAPTER 11

DISPLAYS & VIDEO ADAPTER

COMPUTER DISPLAYS - MONITORS
LCD (LIQUID CRYSTAL DISPLAY) MONITOR
GRAPHICS CARD
VESA ADAPTER

Displays & Video Adapter

A video display module is an output device which processes data from the computer to be displayed on a Cathode Ray Tube (CRT) such as that used in a Television set. In fact, personal computer systems sometimes uses a TV set to display output, linking the computer to it with a small RF (radio frequency) modulator tuned to an unused TV channel. More often, however, the display device is a video monitor, a unit containing a CRT tube, video circuits and power supplies but no RF or sound circuitry. The video display module is the interface between the computer and display. It converts digital data into video signals.

COMPUTER DISPLAYS - MONITORS

The Video monitor or display has rapidly become one of the most popular methods of displaying computer data. As an output device, the monitor can be used to display alphanumeric characters and non-text graphics as well. At the heart of the monitor is the Cathode Ray Tube (CRT) familiar to us from Television receivers we have in our homes. The basic difference between the television and a monitor is that no RF demodulation electronics are necessary in the video monitor.

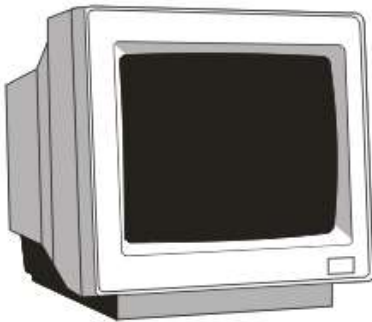


Fig 11.1 (a) Monitor



Fig 11.1 (b) Inside view of Monitor

CATHODE RAY TUBES (CRTS)

Whilst confusion surrounds the precise origins of the cathode-ray tube, or CRT, it's generally agreed that German scientist Karl Ferdinand Braun developed the first controllable CRT in 1897, when he added alternating voltages to the device to enable it to send controlled streams of electrons from one end of the tube to the other. However, it wasn't until the late 1940s that CRTs were used in the first television sets.

The cathode ray tube is the most popular electronic display device in existence. It is also one of the oldest, being invented in 1878 by an

English man named Crookes. Over 80% of the electronics display used world wide are CRTs. The CRT has been in use for so long that a large number of different methods have been developed to display information on it. A thorough understanding of the basic physical operation of the basic CRT will enable you to better understand the different encoding techniques used.

How CRTs Work

A CRT is an evacuated glass tube with an electron gun in the neck of the tube and a rectangular, fluorescent - coated surface opposite the electron gun.

All vacuum tubes are essentially electron accelerators. They use electrostatic charges to manipulate the motion of electrons through a vacuum. The beam of accelerated electrons is directed at a phosphor covered screen. When the electron beam impacts the phosphor coating, it stimulates the phosphor into emitting the light which displays the information on the CRT's screen.

The position of the beam along the face of the screen can be manipulated through horizontal and vertical deflection coils, attached to the tube, which cause the beam to deflect, due to electromagnetic attraction and repulsion principles. By applying different signals to the different coils, the beam can be positioned anywhere along the face of the screen.

The electron beam begins at the upper left hand corner of the screen, and sweeps across the screen to the upper right hand corner (as you are facing the screen), leaving a line across the screen, one line down from the first trace (horizontal retrace). At this point, the horizontal sweep begins producing the second display line on the screen. The scanning continues until the horizontal sweeps reach the bottom of the screen. At this point, the electron beam is blanked again and returned to the upper left corner of the screen (vertical retrace), completing one field.

The number of characters disyllable in a row depends upon the bandwidth of the circuit and the granularity of the monitor. Some constraint is also imposed by the access time of the character-generator read-only memory (CGROM). Better displays are hence more expensive.

Inside CRTs

Illustrates schematically and physical construction of a typical CRT. The first element of the CRT is the source of electrons and is called the cathode. Most CRTs use an indirectly heated cathode. This means that the actual electron emitter is heated by a separate filament called a heater. This heater is constructed of thoriated tungsten which is heated by the passage of low voltage electrons through it. This entire cathode assembly of cathode and heater is the source of

electrons that is manipulated to form the eventual display on the CRT's screen.

The next element of the CRT is control grid. The control grid has the function of controlling the number of electrons per unit time entering the accelerator section of the tube's neck. The control grid controls the level of luminosity displayed on the screen. The video luminosity information is coupled to the control grid by the video amplifiers of the display.

The third element of the CRT is the 'focus anode'. In the focus section the electron beam is accelerated and focused into a tight a beam as possible. The high voltage electrostatic type of focusing is used in virtually all CRTs.

The fourth active element of the CRT is usually simply called the 'anode'. Its function is to rapidly accelerate the electron beam on its way to collision with the screen's phosphors. The acceleration of the electron beam because of accelerating anode which is held at the high voltage, impacts kinetic energy to the electrons that make up beam. The kinetic energy is translated into light by collision with the phosphors on the CRT's screen.

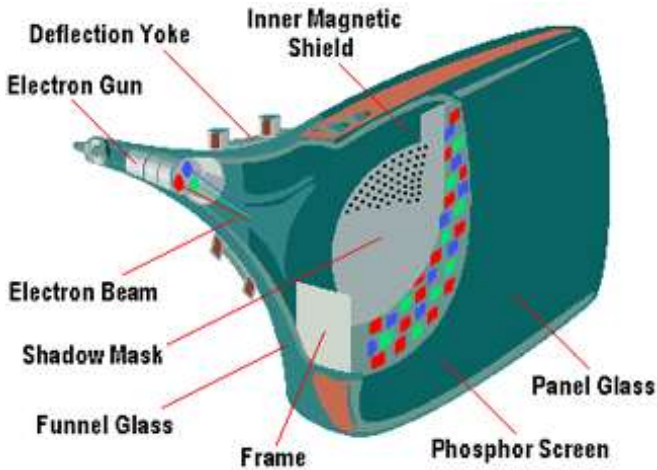


Fig 11.2 Inside view of CRT

The beam must be directed to the location on the screen that is desired to be illuminated. The process of directing the electron beam to the proper location is called 'deflection'. Deflection is accomplished in two manners, electrostatic and electromagnetic.

Electromagnetic deflection is the preferred method. In this technique, a "Yoke" is placed around the neck of the CRT. This yoke is actually two coils of wire, one coil producing magnetic field vertically, the other horizontally.

This yoke controls the position of the electron beam by establishing a variable magnetic field through which the beam must pass. The electron beam is deflected by magnetic interaction with the yoke's field. Electromagnetic deflections also offers the maximum deflection angle possible, making it suitable for large screen CRTs. After all stages of processing, the electron beam meets its intended destiny and collides with the phosphor coating on the screen. The result is light illustrates the different stages of processing that the electron beam undergoes on its journey to the screen.

COLOUR CRT

The most basic type of CRT contains only single colour phosphors and a single electron gun assembly. The full color CRT contains not one but usually three guns. Each of these three guns is aimed only at a specific colored phosphor. Phosphors are chemicals which emit light when excited by a stream of electrons: different phosphors emit different coloured light. One gun for each of the primary colors - Red, Blue and Green. Each gun produces a distinct electron beam which can range over the entire screen. Each gun is, however, directed so that it can only illuminate a single colored phosphor, regardless of its location on the screen.

The base is the monitor's "screen" and is coated on the inside with a matrix of thousands of tiny phosphor dots. Each dot consists of three blobs of coloured phosphor: one red, one green, one blue. These groups of three phosphors make up what is known as a single pixel.

If a pixel is required to display red, then the red phosphor of the pixel is impacted exclusively by the Red gun's electron beam. Colors other than the primary are produced by mixing the primary colors.

For example, if a pixel is required to be yellow then both the red and green phosphors at that pixel's location are activated by simultaneous operation of both the Red and Green guns. The resulting mixture of red and green creates a yellow spot on the screen. Each of the three guns has its own cathode, control grid and focusing anode, but all three guns share the same accelerating anode and deflection yokes.

Before the electron beam strikes the phosphor dots, it travels thorough a perforated sheet located directly in front of the phosphor. Originally known as a "shadow mask", these sheets are now available in a number of forms, designed to suit the various CRT tube technologies that have emerged over the years. They perform a number of important functions:

- they "mask" the electron beam, forming a smaller, more rounded point that can strike individual phosphor dots cleanly
- they filter out stray electrons, thereby minimizing "overspill" and ensuring that only the intended phosphors are hit

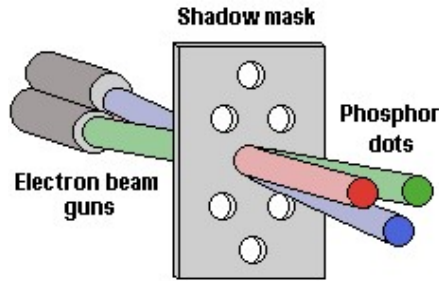


Fig 11.3 Colour CRT

- by guiding the electrons to the correct phosphor colours, they permit independent control of brightness of the monitor's three primary colours.

Disadvantage of CRT's

The CRT is still a formidable technology. It's based on universally understood principles and employs commonly available materials. The result is cheap-to-make monitors capable of excellent performance, producing stable images in true colour at high display resolutions.

However, no matter how good it is, the CRT's most obvious shortcomings are well known:

- it consumes too much electricity.
- its single electron beam design is prone to misfocus.
- misconvergence and colour variations across the screen.
- its clunky high-voltage electric circuits and strong magnetic fields create harmful electromagnetic radiation.
- it's simply too big.

Newer display systems are now available in the form of TFT and LCD monitors that use the same principles as that of your digital watch.

LCD (LIQUID CRYSTAL DISPLAY) MONITOR

A monitor that uses LCD (Liquid Crystal Display) technologies rather than the conventional CRT technologies used by most desktop monitors.

Until recently, LCD panels were used exclusively in notebook computers (laptops) and other portable devices. In 1997, however, several manufacturers began offering full-size Liquid Crystal Display Monitors as alternatives to CRT monitors.

The main advantage of LCD displays is that they take up less desk space and are lighter. Currently, however, they are also much more expensive. LCD (Liquid Crystal Display) is the technology used for displays in notebook, smaller computers, portable devices including

Mobile Telecoms (Telecommunication) pagers, phones, PDAs, EPOS and other instrumentation monitors.

LCD TECHNOLOGY

A liquid crystal display is a thin, lightweight display device with no moving parts. It consists of an electrically-controlled liquid-polarizing liquid trapped in cells between two transparent polarizing sheets. The polarizing axes of the two sheets are aligned in the same direction. Each cell is supplied with electrical contacts that allow an electric field to be applied to the liquid inside.

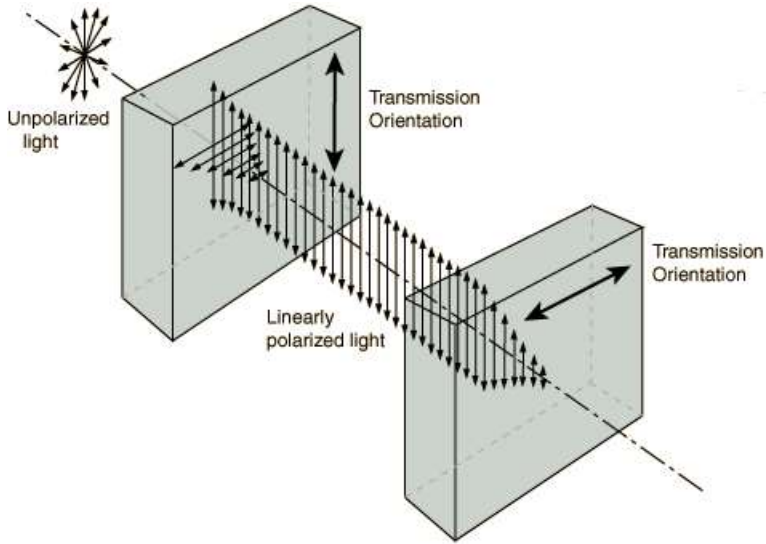


Fig 11.4 Two polarised sheets aligned in vertical direction

Before an electric field is applied, the molecules in the liquid are in a relaxed state. Light can pass through the liquid without any effect on its direction of polarization, so the entire assembly appears transparent. When an electric field is applied, the molecules in the liquid twist into an excited state, causing the axis of polarization of the light to rotate. Thus the light that has been polarized by passing through the first sheet, then rotated by passing through the liquid, now has the wrong polarization to penetrate the second sheet. The result is that the activated part of the display appears dark.

The crystals in an LCD could be alternatively arranged so that light passed when there was a voltage, and not passed when there was no voltage. However, since computer screens with graphical interfaces are almost always lit up, power is saved by arranging the crystals in the no-voltage-equals-light-passing configuration.

Transmissive and Reflective Displays

LCDs can be used in transmissive or reflective modes. A transmissive LCD is illuminated from one side and viewed from the opposite side. Activated cells therefore appear dark while inactive cells appear bright. This type is used in high-brightness applications such as pocket television receivers. The lamp used to illuminate the LCD in such a product usually consumes more battery power than the LCD itself.

A reflective LCD, as used in pocket calculators and digital watches, is viewed by ambient light reflected in a mirror behind the display. This type has lower contrast than the transmissive type, because the ambient light passes twice through the display before reaching the viewer. The advantage of this type is that there is no lamp to consume power, so the battery life is long. A small LCD consumes so little power that it can run from a photovoltaic cell.

Colour Displays

The liquid crystal used in LCDs rotates all visible wavelengths equally, but additional refinements have been added to the basic LCD to produce a colour display.

In a colour LCD each pixel is divided into three sections, one with a red filter, one with a green filter and the other with a blue filter. The pixel can be made to appear an arbitrary colour by varying the relative brightness of its three coloured sections.

Active & Passive Displays

An LCD is made with either a passive matrix or an active matrix display grid. The active matrix LCD is also known as a thin film transistor (TFT) display. The passive matrix LCD has a grid of conductors with pixels located at each intersection in the grid. A current is sent across two conductors on the grid to control the light for any pixel. An active matrix has a transistor located at each pixel intersection, requiring less current to control the luminance of a pixel. For this reason, the current in an active matrix display can be switched on and off more frequently, improving the screen refresh time.

Medium-sized displays, such as those in monochrome personal organisers and pocket television sets, have a passive matrix structure. This type has one set of contacts for each row and column of the display, rather than one for each pixel. However, the disadvantage is that only one pixel can be addressed at any instant. The other pixels have to remember their last state until the control circuit has time to revisit them. This results in reduced contrast and a poor response to fast-moving images. As the number of pixels increases, this type of display becomes less and less attractive.

For high-resolution colour displays such as large LCD monitors, an active-matrix system is used. The LCD panel contains, besides the

polarising sheets and cells of liquid crystal, a matrix of transistors. These devices store the electrical state of each pixel on the display while all the other pixels are being updated. This method provides a much brighter, sharper display than a passive matrix of the same size.

GRAPHICS CARD

Since the early days of personal computers, most graphics cards have been translators, taking the fully developed image created by the computer's CPU and translating it into the electrical impulses required to drive the computer's monitor. All cards today also feature some form of 3D acceleration, which is a specialised functionality that speeds up the display of 3D images and objects-the kind that are found in games and other applications. With this card, special effects like fog, mist and other environmental and lighting effects can also be simulated. The basic components of a graphics card are the processor, the video RAM and the interface. Graphics cards are identified by the graphics processors they use, their interface and the amount of video RAM they feature.

REFRESH MEMORY

Alphanumeric characters sent by the computer to the display module are loaded into successive locations of a block memory. This memory is usually 8 bits wide to accommodate - 8 bit ASCII code. Memory length depends upon how many characters are to be displayed on the screen.

The portion of the memory containing the characters currently displayed, called refresh memory, is accessed once per frame to refresh the display. It is generally accepted - including by standards bodies such as VESA - that a monitor requires a refresh rate of 75Hz or above for a flicker-free display. A computer's graphics circuitry creates a signal based on the Windows desktop resolution and refresh rate. This signal is known as the horizontal scanning frequency, (HSF) and is measured in KHz. A multi-scanning or "autoscan" monitor is capable of locking on to any signal which lies between a minimum and maximum HSF. If the signal falls out of the monitor's range, it will not be displayed.

Adapter	Horizontal scan rate	Verticalscan rate
CGA	15.750 kHz	60 Hz
EGA	22.000 kHz	60 Hz
VGA	31.500 kHz	70 Hz

The display module is within the host computer, memory is physically located on the display module, characters are stored there by the CPU via the address and data buses, using standard memory reference instructions from the CPU's instruction set. However, refresh memory is accessible to the CPU only during horizontal and vertical retrace periods.

RESOLUTION

Resolution is a term often used interchangeably with addressability, but it more properly refers to the sharpness, or detail, of the visual image. It is primarily a function of the monitor and is determined by the beam size and dot pitch (sometimes referred to as “line pitch”). An image is created when a beam of electrons strikes phosphors which coat the base of the monitor’s “screen.

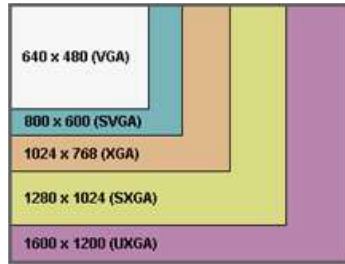


Fig 11.5 Difference among various Resolution

A pixel represents the smallest piece of the screen that can be controlled individually, and each pixel can be set to a different colour and intensity. A complete screen image is composed of thousands of pixels and the screen’s resolution - specified in terms of a row by column figure - is the maximum number of displayable pixels. The higher the resolution, the more pixels that can be displayed and therefore the more information the screen can display at any given time.

1981	CGA	Colour Graphics Adapter	640x200 160x200	None 16
1984	EGA	Enhanced Graphics Adapter	640x350	16 from 64
1987	VGA	Video Graphics Array	640x480 320x200	16 from 262144 256
1990	XGA	Extended Graphics Array	1024x768	16.7 million
	SXGA	Super Extended Graphics Array	1280x1024	16.7 million
	UXGA	Ultra XGA	1600x1200	16.7 million

VESA’s SXGA standard is used to describe the next screen size up - 1280x1024. SXGA is notable in that its standard ratio is 5:4, while VGA, SVGA, XGA and UXGA are all the traditional 4:3 aspect ratio found on the majority of computer monitors.

All SVGA standards support the display of 16 million colours, but the number of colours that can be displayed simultaneously is limited by

the amount of video memory installed in a system. The greater number of colours, or the higher the resolution or, the more video memory will be required. However, since it is a shared resource reducing one will allow an increase in the other.

COLOUR DEPTH

Each pixel of a screen image is displayed using a combination of three different colour signals: red, green and blue. The precise appearance of each pixel is controlled by the intensity of these three beams of light and the amount of information that is stored about a pixel determines its colour depth.

DOT PITCH

The maximum resolution of a monitor is dependent on more than just its highest scanning frequencies. Another factor is dot pitch, the physical distance between adjacent phosphor dots of the same colour on the inner surface of the CRT. Typically, this is between 0.22mm and 0.3mm. The smaller the number, the finer and better resolved the detail.

SPEED

There are many methods of classifying the speed of various functions that CRT must perform in order to display an image. In some cases, these functions are described in terms of time, others are expressed in terms of frequency. The electron beam scans the screen horizontally beginning at the upper left hand corner of the screen. It scans one horizontal line to the right edge of the screen and then it returns to the left side to begin scanning the next line in exactly the same fashion.

While the beam returns from right edge to the left, its intensity is suppressed and no image is displayed during this time. This is called 'horizontal blanking interval' and is abbreviated HBI. The time it takes the beam to scan a horizontal line is called the 'horizontal scanning frequency' (fH) and is expressed in kilohertz. The horizontal scanning frequency of CRTs varies from 15.75 kHz to 64 kHz with horizontal blanking times from 12 μ s to less than 3.5 μ s.

After the electron beam completes a single screen full of horizontal lines, it must return from the bottom of screen to the top and begin the next series of lines. During this bottom to top return process, the beam's intensity is suppressed and no image is displayed. The time it takes is called 'vertical blanking interval' and is abbreviated tVB. The time it takes for the beam to scan an entire series of horizontal lines and return to the top to begin again is called "vertical scanning frequency" (fv) and is expressed in Hz. Today's CRT have vertical scanning frequencies from 40Hz to over 70 Hz and vertical blanking times from 1ms to 0.5 ms.

Another CRT function that is time quantified is the response of electron beam to being turned on or off and is abbreviated TEB. This period determines how fast CRT can go from an illuminated dot on the screen to one which is dark and its time is usually less than 20ns.

VIDEO MEMORY

The memory that holds the video image is also referred to as the frame buffer and is usually implemented on the graphics card itself. Early systems implemented video memory in standard DRAM. However, this requires continual refreshing of the data to prevent it from being lost and cannot be modified during this refresh process. The consequence, particularly at the very fast clock speeds demanded by modern graphics cards, is that performance is badly degraded.

An advantage of implementing video memory on the graphics board itself is that it can be customised for its specific task and, indeed, this has resulted in a proliferation of new memory technologies:

Video RAM (VRAM)

A special type of dual-ported DRAM, which can be written to and read from at the same time. It also requires far less frequent refreshing than ordinary DRAM and consequently performs much better

Windows RAM WRAM

As used by the hugely successful Matrox Millennium card, is also dual-ported and can run slightly faster than conventional VRAM

EDO DRAM

It provides a higher bandwidth than DRAM, can be clocked higher than normal DRAM and manages the read/write cycles more efficiently

SDRAM

Similar to EDO RAM except the memory and graphics chips run on a common clock used to latch data, allowing SDRAM to run faster than regular EDO RAM

SGRAM

Same as SDRAM but also supports block writes and write-per-bit, which yield better performance on graphics chips that support these enhanced features

DRDRAM

Direct RDRAM is a totally new, general-purpose memory architecture which promises a 20-fold performance improvement over conventional DRAM.

VIDEO GRAPHICS ARRAY (VGA)

VGA (for Video Graphics Array) technology dates to 1987. VGA provides a basic graphics resolution of 640 pixels (picture elements or dots) across by 480 pixels down with 16 colors.

All the previous graphics adapter CGA and EGA had used a digital monitor. A digital monitor has three colour control lines, one each for Red, Green and Blue. Each of these lines can either be ON or OFF so the maximum number of colors a digital monitor can give 16 if you use an additional Intensity control line.

VGA uses an analog colour signal which allows an almost infinite variety of colors. Basic VGA requires 4 bits of memory per pixel to display 16 colors $2^4=16$, for 256 colors you need 8 bits ($2^8=256$) or 1 byte per pixel. 16.78 millions colors require 24 bits. ($2^{24} = 16.78$ million) or 3 bytes per pixel. For 16 colors VGA and super VGA, you need only 256 K byte of video memory. For 256 colors, 1024 x 768 video, you need 1 Megabyte. Other resolutions and color depths need 512 K byte memory.

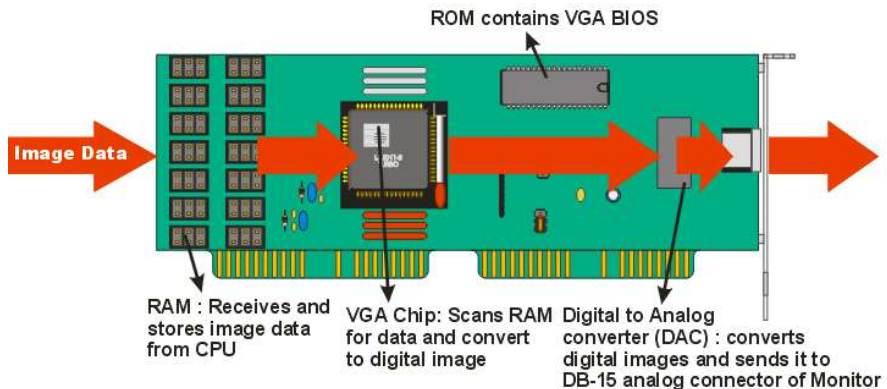


Fig 11.5 Video Board

RAMDAC

Many times per second, the RAMDAC reads the contents of the video memory, converts it into an analogue RGB signal and sends it over the video cable to the monitor. It does this by using a look-up table to convert the digital signal to a voltage level for each colour. There is one Digital-to-Analogue Converter (DAC) for each of the three primary colours the CRT uses to create a complete spectrum of colours. The intended result is the right mix needed to create the colour of a single pixel. The rate at which the RAMDAC can convert the information, and the design of the graphics processor itself, dictates the range of refresh rates that the graphics card can support. The RAMDAC also dictates the number of colours available in a given resolution, depending on its internal architecture.

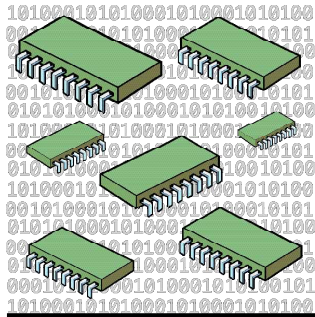
VESA ADAPTER

To speed the transfer of video information from system's microprocessor to the screen is to put a quicker connection between them - side step the slow expansion bus with a more direct link. Putting display circuitry on a PC's motherboard is not enough since such systems electrically connected their system board video circuitry to the PC through the same Input/ output interface used by the expansion bus. The only difference is the elimination of a connector. This design doesn't necessarily improve performance, however. Local bus video design take the video circuitry of the expansion bus and add a new link direct to the microprocessor. The local connection which can be either hard-wired to video circuitry on the motherboard or plugged into a new expansion - bus design operates at memory speed and with a full 32 bit bus width, giving it an appreciably higher speed potential than an expansion bus connection.

Moreover, while the high transfer rates promised by advanced expansion buses like micro-channel and EISA requires special software commands, local bus designs work at full speed with all traditional applications- no modifications or special drivers required. VESA (Video Electronics Standards Association) adapters uses the 32 bit local bus of microprocessor and hence transfer of video information from system's microprocessor to the screen is increased considerably. These adapters are used typically graphical environments like windows and CAD programs such as AUTOCAD.

EXERCISE

1. State True or False :
 - a) Video display module converts analog data into video signals.
 - b) RF demodulation is required in television and not in monitor.
 - c) Yoke in CRT is used to deflect the beam.
 - d) LCD is normally used in mobile communication.
 - e) More resolution indicates more number of pixels.
2. What is CRT?
3. Discuss various elements of CRT?
4. What is shadow mask?
5. What is a pixel ?
6. How coloured CRT are different from mono CRT?
7. What is Liquid Crystal display? How it works?
8. Differentiate between Active and Passive display.
9. Explain the following terms:
 - a) Colour Depth
 - b) Resolution
 - c) Dot Pitch
10. What is VGA technology?



CHAPTER 12

LAPTOPS

LAPTOP HISTORY
INSIDE LAPTOP

Laptops

A laptop computer, or a notebook, is a battery or AC-powered personal computer generally smaller than a briefcase that can easily be transported and conveniently used in temporary spaces such as on airplanes, in libraries, temporary offices, and at meetings. A laptop typically weighs less than 2.5 Kg and is 3 inches or less in thickness. It includes a screen, keyboard, and a mousepad or trackball, which serves as the mouse. As laptops are designed to be used anywhere, they have a battery which allows them to operate without being plugged into a power outlet and a power adapter that allows them to use power from an outlet and recharges the battery.

LAPTOP HISTORY

Alan Kay of the Xerox Palo Alto Research Center originated the idea of a portable computer in the 1970s. Kay envisioned a notebook-sized, portable computer called the Dynabook that everyone could own, and that could handle all of the user's informational needs. Kay also envisioned the Dynabook with wireless network capabilities. Arguably, the first laptop computer was designed in 1979 by William Moggridge of Grid Systems Corp. It had 340 kilobytes of bubble memory, a die-cast magnesium case and a folding electroluminescent graphics display screen.



Fig 12.1 World's First laptop

In 1983, Gavilan Computer produced a laptop computer with the following features:

- 64 kilobytes (expandable to 128 kilobytes) of random access memory (RAM)
- Gavilan operating system (also ran MS-DOS)
- 8088 microprocessor
- touchpad mouse

- portable printer
- weighed 9 lb (4 kg) alone or 14 lb (6.4 kg) with printer

The Gavilan computer had a floppy drive that was not compatible with other computers, and it primarily used its own operating system. But this Gavilan computer was a flop.

Other companies, like Apple, introduced other portable computers in following years. The first commercially viable machine, however, was the IBM PC Convertible, introduced in 1986. The PC Convertible featured:

- An 8088 microprocessor
- 256 kilobytes of memory
- Two 3.5-inch (8.9-cm) floppy drives
- An LCD display
- Parallel and serial printer ports
- Space for an internal modem

A software suite including basic word processing, appointment calendar, telephone/address book, and calculator software

Weighing in at a hefty 12 lbs (5.4 kg), the PC Convertible was sold for \$3,500. It was the first portable computer with the clamshell design used in today's laptops. The success of the PC Convertible was the catalyst for competitors like Compaq and Toshiba to switch to the clamshell design in their portable computers. And so began the era of the laptop computer.

LAPTOP TODAY

Among the best-known makers of laptop computers are IBM, Apple, Compaq, Dell, Lenovo, Sony, LG, Samsung, Asus, BanQ and Toshiba and dozen other brands. Laptop computers generally cost more than desktop computers with the same capabilities because they are more difficult to design and manufacture. A laptop can effectively be turned into a desktop computer with a docking station, a hardware frame that supplies connections for peripheral input/output devices such as a printer or larger monitor. The less capable port replicator allows you to connect a laptop to a number of peripherals through a single plug.

Laptops come with displays that use thin-screen technology. The thin film transistor or active matrix screen is brighter and views better at different angles than the STN or dual-scan screen. Laptops use several different approaches for integrating a mouse into the keyboard, including the touch pad, the trackball, and the pointing stick. A serial port also allows a regular mouse to be attached. The PC Card is insertable hardware for adding a modem or network interface card to a laptop. CD-ROM and digital versatile disc drives may be built-in or attachable.



Fig 12.2 Inside Laptop

All these fundamental design difference and requirement of inherent portability, all the components of the Laptops are defined in such a way that they:

- Fit into a compact space
- Conserve power
- Produce less heat than desktop components

Often, these differences make the components more expensive, which can contribute to higher laptop prices.

ADVANTAGES OF LAPTOP

Portability is usually the first feature mentioned in any comparison of laptops versus desktop PCs. Portability means that a laptop can be used at any place—not only at home and at the office, but also during commuting and flights, in coffee shops, in lecture halls and libraries, at clients' location or at a meeting room, etc. The portability feature offers several distinct advantages:

More Output

Using a laptop in places where a desktop PC can't be used, and at times that would otherwise be wasted. For example, an office worker

managing his e-mails during an hour-long commute by train, or a student doing his/her homework at the university coffee shop during a break between lectures.

Immediacy

Carrying a laptop means having instant access to various information, personal and work files. Immediacy allows better collaboration between coworkers or students, as a laptop can be flipped open to present a problem or a solution anytime, anywhere.

Up-to-date information

If a person has more than one desktop PC, a problem of synchronization arises: changes made on one computer are not automatically propagated to the others. There are ways to resolve this problem, including physical transfer of updated files (using a USB stick or CDs) or using synchronization software over the Internet. However, using a single laptop at both locations avoids the problem entirely, as the files exist in a single location and are always up-to-date.

Connectivity

A proliferation of Wi-Fi wireless networks and cellular broadband data services (HSDPA, EVDO and others) combined with a near-ubiquitous support by laptops[39] means that a laptop can have easy Internet and local network connectivity while remaining mobile. Wi-Fi networks and laptop programs are especially widespread at university campuses.

Size

Laptops are smaller than standard PCs. This is beneficial when space is at a premium, for example in small apartments and student dorms. When not in use, a laptop can be closed and put away.

Ease of Access

Most laptops have doors on the underside that allow the user to access the memory, hard drive and other components, by simply flipping the laptop to access the doors. For desktops the user must usually access the backside of the computer, which is harder if it's in an area with little space.

Low power consumption

Laptops are several times more power-efficient than desktops. A typical laptop uses 20-90 W, compared to 100-800 W for desktops. This could be particularly beneficial for businesses (which run hundreds of personal computers, multiplying the potential savings) and homes where there is a computer running 24/7 (such as a home media server, print server, etc.)

Quiet

Laptops are often quieter than desktops, due both to the components (quieter, slower 2.5-inch hard drives) and to less heat production leading to use of fewer and slower cooling fans.

Battery

A charged laptop can run several hours in case of a power outage and is not affected by short power interruptions and brownouts. A desktop PC needs a UPS to handle short interruptions, brownouts and spikes; achieving on-battery time of more than 20–30 minutes for a desktop PC requires a large and expensive UPS.[41]

All-in-One

Designed to be portable, laptops have everything integrated in to the chassis. For desktops (excluding all-in-ones) this is divided into the desktop, keyboard, mouse, display, and optional peripherals such as speakers, and a webcam. This leads to lots of wiring. It can also lead to massive power consumption.

Extras

In comparison to low-end desktops, even low-end laptops include features such as a Wi-Fi card, and Express Card slot, and a memory card reader.

INSIDE LAPTOP

MOTHERBOARD

Laptop motherboards are highly make and model specific, and do not conform to a desktop form factor. Unlike a desktop board that usually has several slots for expansion cards (3 to 7 are common), a board for a small, highly integrated laptop may have no expansion slots at all, with all the functionality implemented on the motherboard itself; the only expansion possible in this case is via an external port such as USB. Other boards may have one or more standard, such as ExpressCard, or proprietary expansion slots. Several other functions (storage controllers, networking, sound card and external ports) are implemented on the motherboard.[28]

LAPTOP PROCESSORS

The microprocessor works with the operating system to control the computer. It essentially acts as the computer's brain. The CPU produces a lot of heat, so a desktop computer uses circulating air, a fan and a heat sink - a system of plates, channels and radiator fins used to draw heat off of the processor - to cool off. Since a laptop has far less room for each of these cooling methods, its CPU usually:



Fig 12.3 Laptop Motherboard

Runs at a lower voltage and clock speed

This reduces heat output and power consumption but slows the processor down. Most laptops also run at a higher voltage and clock speed when plugged in, and at lower settings when using the battery.

Mounts to the motherboard without using pins

Pins and sockets take up a lot of room in desktop PCs. Some motherboard processors mount directly to the motherboard without the use of a socket. Others use a Micro-FCBGA (Flip Chip Ball Grid Array), which uses balls instead of pins. These designs save space, but in some cases mean that the processor cannot be removed from the motherboard for replacement or upgrading.

Has a sleep or slow-down mode

The computer and the operating system work together to reduce the CPU speed when the computer is not in use or when the processor does not need to run as quickly. The Apple G4 processor also prioritizes data to minimize battery drain.



Fig 12.3 Heat Sink of Laptop

Some laptops use desktop CPUs that are set to run at lower clock speeds. Although this can improve performance, these laptops typically run much hotter and have a significantly reduced battery life.

Laptops usually have small fans, heat sinks, heat spreaders or heat pipes to help dissipate the heat from the CPU. Some higher end laptop models reduce heat even further with liquid coolant kept in channels alongside the heat pipe. Also, most laptop CPUs are near the edge of the unit. This allows the fan to move the heat directly to the outside instead of across other components.

LAPTOP MEMORY AND STORAGE

A laptop's memory can make up for some of the reduced performance that comes from a slower processor. Some laptops have cache memory on or very near the CPU, allowing it to access data more quickly. Some also have larger busses, allowing data to move between the processor, motherboard and memory more quickly.



Fig 12.4 Laptop Memory

Laptops often use smaller memory modules to save space. Memory types used in laptops include:

- Small Outline Dual Inline Memory Module (SODIMM)
- Dual Data Rate Synchronous RAM (DDR SDRAM)
- Single data rate Synchronous RAM (SDRAM)

PROPRIETARY MEMORY MODULES

Some laptops have upgradeable memory and feature removable panels for easy access to the memory modules.

Like a desktop, a laptop has an internal hard disk drive, which stores the operating system, applications and data files. However, laptops generally have less disk space than desktops. A laptop hard drive is also physically smaller than that of a desktop. In addition, most laptop hard drives spin more slowly than desktop hard drives, reducing both heat and power consumption.

Desktop computers have multiple bays for installing additional drives, such as CD and DVD ROM drives. However, space in a laptop is in much shorter supply. Many laptops use a modular design, allowing a variety of drives to fit in the same bay. These drives come in three different designations:

Hot swappable The computer can stay on while changing the drive.

Warm swappable The computer can stay on while changing the drive, but the corresponding bus (the path the drive uses to send data to the CPU) must be inactive.

Cold swappable The computer must be off during the swap.

In some cases, these drive bays are not just limited to drives but will also accept extra batteries.

LAPTOP SCREEN, GRAPHICS AND SOUND

A graphics processing unit (GPU) is a microprocessor that handles the calculations necessary for 3-D graphics rendering. Like a CPU, a GPU produces a lot of heat. Most laptops have graphics capability built into the motherboard or have smaller graphics cards with a GPU designed specifically for laptop use. GPU manufacturers ATI and nVidia both make GPUs specifically for laptops. Laptops frequently share memory between the CPU and the GPU, saving space and reducing power consumption.



Fig 12.5 Laptop display

Many people don't notice a laptop's reduced graphics performance. Laptops have plenty of processing power for Web surfing and productivity applications. However, they may struggle with the latest 3-D games. A few specialty laptops, designed for gaming enthusiasts, include more powerful GPUs and additional video memory.

A laptop displays its graphics on a liquid crystal display (LCD) screen. Most screens measure between 12 and 17 inches, and the size of the screen affects the overall size of the laptop.

In addition, laptop screens can be:

- Black-and-white (16 grayscale) or color (65,536 colors)
- Active or passive matrix
- Reflective or backlit

Active matrix displays have sharper images and are easier to read, and backlit screens are better for low-level lighting conditions.

Most laptops also have sound cards or integrated sound processing on the motherboard as well as small, built-in speakers. However, there is generally not enough space inside a laptop for a top-of-the-line sound card or a high-quality speaker. Gaming enthusiasts and audiophiles can supplement their laptops' sound capabilities with external sound controllers, which use USB or FireWire ports to connect to the laptop.

LAPTOP BATTERIES

Laptops and desktops both run on electricity. Both have small batteries to maintain the real-time clock and, in some cases, CMOS RAM. However, unlike a desktop computer, a laptop is portable and can run on batteries alone.



Fig 12.7 Laptop Battery

Nickel-Cadmium (NiCad) batteries were the first type of battery commonly used in laptop computers, and older laptops sometimes still use them. They have a life of roughly two hours between charges, but this life decreases with each charge based on the memory effect. Gas bubbles form in the cell plates, reducing the total amount of available cell space for recharge. The only way around this is to discharge the battery completely before recharging it. The other

drawback of NiCad is that if the battery charges too long, it can explode.

Nickel-Metal Hydride (NiMH) batteries are the bridge between NiCad and the newer Lithium-Ion (LiIon) batteries. They last longer between charges than NiCad but overall have a shorter total lifespan. They suffer from the memory effect, but to a lesser extent than NiCad batteries.

Many laptops with LiIon batteries claim to have a 5-hour battery life, but this measurement can vary greatly depending on how the computer is used. The hard drive, other disk drives and LCD display all use substantial battery power. Even maintaining wireless Internet connectivity requires some battery power. Many laptop computer models have power management software to extend the battery life or conserve battery power when the battery is low.

LiIon batteries are the current standard for laptop computers. They are light and have long life spans. They do not suffer from the memory effect, can be charged randomly, and won't overheat if overcharged. They are also thinner than any other battery available for laptops, making them ideal for the new ultra-thin notebooks. LiIon batteries can last for anything from about 950 up to 1200 charges.

WHITEBOOKS - ASSEMBLED LAPTOPS

Many people build custom PCs for themselves or customers. These home-made computers, called whiteboxes, represent a significant portion of the computer market. Some people also modify their computers for cosmetics or performance. This is called modding. But what about building or modding a laptop?

A homemade laptop is called a whitebook. Whitebooks represent about 5 percent of the notebook market right now, and this number is slowly rising. The industry has done a pretty good job of keeping end users out of laptops. They've made it difficult to open, modify and get parts for a laptop. In addition, opening the laptop chassis voids the manufacturer's warranty in most cases.

It is still difficult to find parts to build a laptop from the ground up, but vendors like ASUS and ECS allow some customers to order blank laptop shells. They are especially open to resellers who build whitebooks and sell them to customers. In addition, people can modify or upgrade what came with the shell.

A laptop shell consists of:

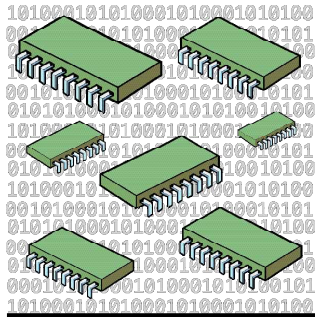
- Chassis
- Motherboard
- Optical drive
- Display
- Keyboard

- Graphics Card

This means that anyone wanting to build a whitebook must find:

- Microprocessor
- RAM
- Wireless card

At this point, there are no real standards for the form factor (shape and design) of laptop parts. Processors designed for laptops are available for sale, but finding a motherboard for those chips is a different story. Hard drives are pretty standard, and SODIMM system memory is easy to come by, but other parts may take some digging.



CHAPTER 13

ASSEMBLING COMPUTERS

INSTALLING THE CPU
INSTALLING THE RAM
MOUNTING THE MAINBOARD
INSTALLING HARD DISK
INSTALLING FLOPPY DRIVE
INSTALLING AGP/PCI DEVICES
READY TO GO....
BIOS SETUP

Assembling Computers

Probably one of the most fun aspects of assembling computer is buying the right components. Needs differ from one person to the next and price plays a critical role in the equipment you choose. The general rule of thumb for buying computer equipment is to buy the best components you can afford.

At present date the following configuration is one of the best configuration available:

Processor:	Intel 2.0 GHz
RAM:	256 MB stick of Hundai PC133 RAM
Mainboard:	VIA/Intel Chipset Motherboard
Case:	ATX Minitower Cabinet
Sound Card:	In built on Motherboard
Video Card:	Inbuilt on Motherboard
Phone Modem:	GVC/ Motorola 56k Modem
Hard Drive:	Samsung/Segate 80 GB HD (IDE Device)
CD/DVD-ROM Drive:	52X CDROM(IDE Device)
3.5" Drive(s):	1.44 MB 3.5" Floppy Drive
Speaker :	Creative/Samsung speakers
Monitor :	15" Samsung/L/Microtek monitor
UPS :	Any UPS of 650 VA with 1/2 batteries

Other Equipments Required

First and foremost make sure you have some good tools. You should not be running around for something mid-way, so its best to have all the tools within your reach. All you really need is a Phillis head screwdriver and a plane screwdriver with a magnetic head. If the head is not magnetic, you may keep a pair of needle nose pliers (just incase something falls in the case).

Screw Drivers

Screw drivers are classified by their over all length and the width of the working edge in mm. The size and type of screwdriver must be chosen to match the screw that is being driven.

Screw drivers are used to drive the screws and keep them in proper position.

i) Common screw drivers.



Fig 13.1 Replacable tip Screwdriver



Fig 13.2 Flat tip Screwdriver

ii) Philips type Screwdriver

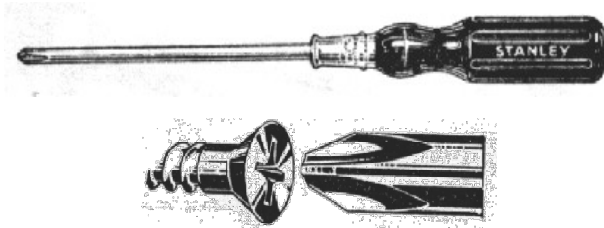


Fig 13.3 Philips Screwdriver

Tweezer

Tweezer is used for holding small job while soldering made of steel. Identified by their total length.



Fig 13.4 Tweezer

It is used to grip the wires or components and take or keep them at required places.

Before you start, make sure you have all potentially electro-statically charged items neutralized including you, static electricity has the potential to ruin your components. To neutralize any static electricity which you may be carrying, simply plant your feet firmly on the ground and touch the case without power supply plugged in. But this is just an arrangement. For neutralize any static electricity, you can also pick up a static wristband.

INSTALLING THE CPU

It's advisable to install your processor while the motherboard is outside of the case, as it gives you much more free space in which to work. First, make sure that your CPU has either the supplied heatsink and or fan attached. Next, simply plug in your Processor to the CPU connector, usually located in the upper left hand corner of your Mainboard. Depending on what type of CPU you are installing, be it Slot 1, Slot A, or Socket 370, there is only one correct way to install the CPU, gently slide the CPU into position, and check to make sure you have it firmly in place. Never exert even a simple pressure to plug in your processor into the slot.

If the fan & heatsink is not an integral part of CPU, put the heatsink over the CPU and then fix the fan over the heatsink in such a way that the heatsink can't move.

INSTALLING THE RAM

Like the CPU, this should also be done while the Mainboard is out of the case. You should be able to find your RAM slots located in the upper right hand corner of your Mainboard. Most of today's mainboards contain 2-4 DIMM connectors, while some newer (and more expensive) Intel i820 based mainboards contain RIMM slots, RIMM memory modules are a faster (and MUCH more expensive) type of RAM compared to today's SDRAM DIMMs.

Installing the RAM sticks is a snap - literally. By using DIMM modules it does not matter what size (in Megabytes) they are, or where/what order you position them in.

MOUNTING THE MAINBOARD

Make sure any jumpers are set properly and that your CPU and RAM have already been installed. Most of today's newer mainboards don't have hardware jumpers, the CPU's Speed/Voltage/FSB settings are all automatically set via the Mainboard's own 'soft jumpers' (located in the BIOS).

If you do happen to buy a Mainboard that requires jumpers to be set, refer to your Mainboard's manual, it will guide you through proper jumper setup procedures.

Your cabinet may have have a detachable Mainboard installation plate (If not, you may have to insert the motherboard into the case, which may be a bit tricky). Second, you'll want to check and see what actually connects your Mainboard with the installation plate. The most popular (and probably best) method is the use of metal risers. These screw in and out of both the Mainboard and the case, providing a good safe hold on the board. The second method is that of using plastic spaces. These snap in and out of your Mainboard, but are often easy to break, and can be hard to use.

After fixing the motherboard on the case plate using the screws or plastic spaces, line up the Mainboard with the case so that the I/O connectors on the motherboard line up nicely with I/O connectors on the case's I/O shield plate.

Finally, connect the case's power supply with the Mainboard and do the same with any Mainboard fans, etc. This should be very easy to do.

Now, you are ready to plug in the wires that connect the case's LED lights (power, and HD lights), power switch, restart switch and, if applicable, the internal speaker. Every Mainboard needs its case wires arranged differently so consult your manual for correct

arrangement of these wires. If you don't get any of these steps done right the first time, don't panic, it's easy to reposition your Mainboard, or reconnect a wire differently if needed.

INSTALLING HARD DISK

When planning a hard disk installation, consider the following items to assure a successful and reliable result.

Drive Bay - Obviously, a suitable drive bay is required to physically hold the drive. The size requirement depends on the drive itself. The rear access space requirements for signal and power cables is often entirely overlooked in early planning or deferred until attempted installation. Consider measuring the actual dimensions of the drive bay and comparing them against the various drive's mechanical specifications.

Power Connector - A spare drive power connector is usually present on most power supplies, but a 'splitter' or an adapter may be required to connect to the drive.

Cable length - Both to reduce clutter in the computer case and to insure a minimum of noise and delay, select the shortest cables possible. Keep in mind, though, that cables should not be under tension, nor should they rest on sharp metal corners.

Front Clearance - Drives installed in "hidden" bays may protrude into case hardware, causing misfits or even electrical shorts. Insure sufficient clearance is available.

Vertical v/s Horizontal - Most drives are not designed for front or rear oriented as top or bottom, but may otherwise be mounted as desired. Avoid diagonal orientation.

Hole Pattern - One would assume that hole patterns in computer cases are in the correct locations, but aberrations do occur, and if holes need to be moved or installed, all fragile components- system boards, other drives, etc.- should be removed prior to any drilling or filing.

Power Supply - The reduced power requirements per Megabyte of newer drives seldom require power supply upgrades, but older supplies may not be sufficient and should be replaced.

Controller Types - The controller type selected must be compatible with the drive or vice versa, but insure that any required I/O slots, interrupts, DMA channels, etc. are not in conflict with other installed devices. In many cases, these parameters are user setable via jumper pins on the controller card.

BIOS Compatibility - If dealing with an AT class machine, inspect the BIOS Drive Table to determine if the drive is currently supported or if a modification might be required.

Logical Installation - The installation procedure requires several steps, and each one has a particular order of progression to the next. The drive has to be configured correctly, physically installed, formatted, configured to receive an operating system and have the systems software installed. Hard disk can be divided to accommodate several different operating systems.

An outline of the major steps to follow in the installation of a hard disk for any systems is as following :

- (a) Drive configuration
 - Drive select jumpers
- (b) Physical Installation
- (c) Machine to disk configuration
- (d) Formatting and software installation
 - Low level format
 - Defect mapping
 - Interleave selection
 - Partitioning
 - High level format

For knowing more about setting the drive jumpers, refer to Chapter “Secondary Storage II”.

PHYSICAL INSTALLATION

After configuring the jumper setting, attach cables and mount the drive. Align pin 1 on the 40/68/80 pin interface along with the pin 1 of the interface cable. Align the other end of the cable to the pin 1 on the system host adapter on the Motherboard. Secure the drive in any orientation using the two screws per side. Do not overtighten the screws.



Fig 13.5 Fixing Harddisk in Cabinet

The hard drive can be laid horizontally with the top and bottom parallel to the bottom of the computer case, or on it's right or left side. Never install the hard drive vertically, with either the front or back plates facing upward. All these things are mandatory for proper physical installation of the drive.

When attaching the IDE devices with their corresponding cables make sure that the red line that runs along the side of the cable matches up correctly with the first pin (one closest to power connector) in the drive's pin array. It's also a good idea to place the master device on the end of the cable, and any slave devices on the inner most connector of the same cable. You should now be able to attach the power connectors to the devices; this should be easy to do as the 4-pin connectors were designed to go in only one way.

INSTALLING FLOPPY DRIVE

Installing the 3.5" floppy drive is much the same as installing an IDE device. One notable difference is that they use a different type of power connector (which is smaller) and a lesser 34-pin cable. Make sure to align the red line on the ribbon cable to the first pin on the back of the 3.5" floppy.

INSTALLING AGP/PCI DEVICES

By this point you should give yourself a pat on the back, you're nearly finished. It's a good idea to first install your AGP card and give your machine a little test boot. Installing an AGP video card is easy, you simply slide the card into the AGP port (usually the brownish connector closest to the CPU), make sure it's in firmly, and screw the edge connector into the case. Now you should try booting the computer.

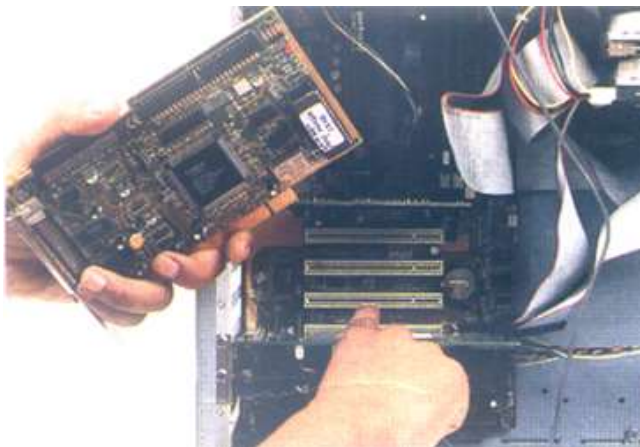


Fig 13.6 Installing PCI Card

READY TO GO....

Once all the things are in place, you are ready to go. What you are supposed to do now is to connect the bunch of cables to the Cabinet. For every cable that will be attached to the cabinet, there is a port.

IMPORTANT PLUG TYPES

Below are the most common connectors on your PC. You should be familiar with the functions of the individual plugs, since you'll have to detach them all and then reattach them in the upcoming installation.

- Standard connector of a graphic card. This is where you will connect the 15-pin plug of the monitor later on. Be careful when connecting this plug, since the pins bend very easily.

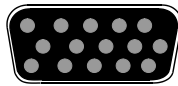


Fig 13.7 15 pin plug

- Connector for a parallel interface. This is where you usually attach your printer to the PC identifier LPT1.

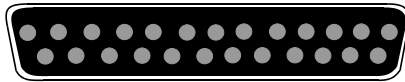


Fig 13.8 25 pin Printer plug

- 9-pin connector for a serial interface cable. A mouse or a modem can be connected here if you have the right kind of cable.

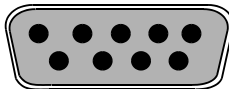


Fig 13.9 9 Pin serial plug

- 25-pin connector for a serial interface cable. You can also connect a mouse or a modem here, but you need a cable with a corresponding 25-pin socket.

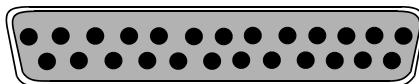


Fig 13.10 25 ping serial plug

- This connector is where you attach a joystick if you have one. This connector isn't standard PC equipment. There is often duplicate of this plug on the sound card, which can either connect a joystick or serve as a MIDI interface.

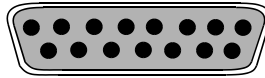


Fig 13.11 15 Pin Game port

- The keyboard plug has to be attached to this connection correctly, or you won't be able to communicate with your PC. Don't ever connect the plug while the computer is running!

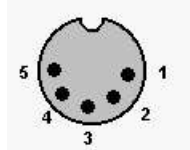


Fig 13.12 Keyboard plug

- This connector is where the PC is attached to the power socket. This plug should be tightly in the socket so that your computer doesn't lose power.

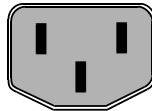


Fig 13.13 Power Plug

After Browsing the various plug types, do the following to start the system.

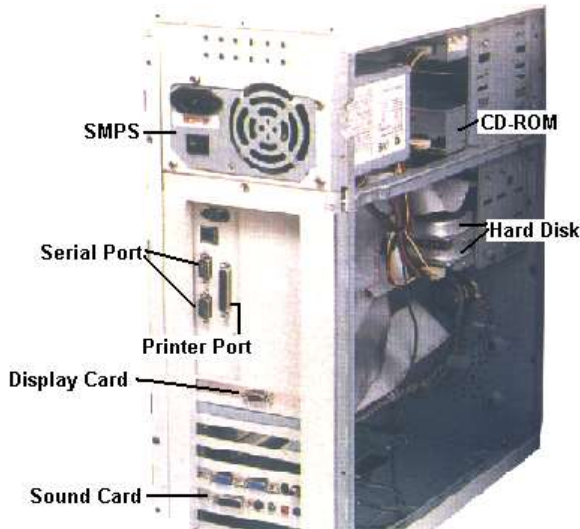


Fig 13.14 Back View of System Unit

- Attach the Monitor. Monitor normally has two cables coming out of it - a power cable and a data cable. The power cable is connected to the SMPS (or to the power switch) and data cable to the display port (A 15 pin port).
- Now attach the Keyboard to the keyboard port and Mouse to the COM1 port of PS/2 port.
- Connect the printer data cable to the parallel port.
- If you have an external modem, connect it to the COM2 port.
- Attach the Speakers and MIC to the sound card.

Now give power to the Cabinet - attach a power cord into SMPS and plug it into power socket.

BIOS SETUP

Now, it's time to setup the BIOS. Pressing either the Delete or F2 key after the computer posts lets you access the computer's BIOS. Browse through your Mainboard's manual for configuring the BIOS.

If your Mainboard is jumper-less, the first thing that you need to do is to check to make sure the CPU's clock speed is correct. The BIOS usually will automatically detect your CPU and set BIOS accordingly, but sometimes it will miss-detect your CPU forcing you to manually set the correct parameters.

Next important step is to configure your Hard disk on the BIOS. BIOS do have an auto detection facility, but in case it doesn't have, you have you to manually set the correct parameters.

CONFIGURING HARD DISK ON THE MACHINE

In order to begin the setup procedure, several specific items about the hard disk drive, controllers and the system ROM BIOS needs to be known before hand. Before installing the drive, write down and store all salient data from the label, pasted on the drive, in a separate location for future reference so you won't have to remove the drive to find it later.

- Number of cylinders
- Number of heads
- Starting cylinder for write pre-compensation
- Range of acceptable head step pulse timing
- Location of all defects by cylinder and head

After fixing the drive on the system, start the computer and press the configuration key (Del or F2 in most of the computers) to start BIOS setup. Enter the drive details in the standard setup. If entering the data in the standard setup, enable the LBA mode if your drive is LBA.

You can also select the Auto-detect option to recall the device parameters, if your BIOS has one. If auto-detect option is not available, and LBA mode is disabled, select user-defined drive type

and enter the LBA drive parameters defined on the hard drive label. If setup program doesn't accept the LBA parameters, use the CHS parameters.

In many BIOS setup, Auto option is available in the drive type. You may also select this option so as to setup to recall the parameters when ever, you start the computer.

Once you define the required parameter, save the BIOS setting and exit. Now boot from a floppy drive or a CD-ROM (To boot from a CD-ROM, you need a bootable CD-ROM and you must set the First boot device to CD-ROM in the Advance Setup of BIOS).

FORMATTING AND SOFTWARE INSTALLATION

Three basic steps are to be followed to complete this process, they are:

- Low level format
- Defect mapping
- Interleave selection
- Partitioning
- High level format

Low level format

The low level format includes drive defect mapping and interleave selection. A low level format is real format, which means that the actual tracks and sectors of the disk are outlined and written. Any newly purchased or acquired drive has to go through this process. An improper low level format can cause the drive to loose data and have many read/write failures so care must be taken while low level formatting.

When to Low-Level Format a Drive

There are four basic reasons to low-level format a hard disk. These reasons apply to ST506 and ESDI systems, where the controller and drive are separate components.

- If you've purchased a new controller, or new drive, it's a good idea to reformat. The low level format "marries" the drive and controller, calibrating them for each other.
- Just to refresh the sector IDs on your disk.
- You wish to experiment with the interleave factor to increase your data transfer rate from your hard disk. We covered interleave factors previously. Recall that an interleave factor can only be changed with a low level format.
- You suspect a faulty format, as in a case where a disk is accidentally formatted as a 20 MB drive when it is actually a 10MB drive.

How to Low-Level Formatting

Assuming that you want to low-level format a drive how do you do it? There are several things to consider.

- Low level format programs are either generic or they are specific to a particular controller. (Notice that low level format programs are specific to controllers, not drives.)
- A controller's low level format program may be stored on a ROM on the controller or it may be a separate DOS program.
- If your controller comes with a low level format program use it over a generic program.
- You usually activate a built in low level format program with the DOS program DEBUG.

Using DEBUG to Low Level Format

If your controller has a BIOS on it, then that BIOS may contain a low level formatting program. Here's how to use it.

First, it's probably a good idea to reboot your system without any memory resident (TSR) programs. Format programs are timing dependent, and may not work correctly with some TSRs. Definitely remove any automatic head parking programs or disk cache programs.

Next, try DEBUG, a program that comes with DOS. If your controller has a ROM on it, you activate its low level format program with DEBUG's G command, the Go and Execute command. You've got to know where the built in format program is. On XT controllers, the most common address is C800:5 or C800:6 although.

Adapter controllers use C800;CCC.AT controllers that have ROM commonly store the low-level format program at CC00; 5, such as the Western Digital 1006V-SR2 or 1003-SR2. As always, check your controller documentation to find the program.

Using Generic Low-Level Format Programs

Many 16-bit ISA-type ST506 and ESDI controllers format in pretty much the same way. That led to the creation of a number of generic low level format programs. You'll find one example on the CD-ROM disc, a program called HDTEST.

Good generic low level formatted is included with Disk Manager. You want only to low-level format the drive not partition and high level format it, so activate Disk Manager with DM/M. At various times in PC history. DM has been the most general purpose low level formatted around. The On Track folks apparently work hard to keep up with the variety of controllers out there, so make sure you've got a copy around.

Using Specific Low-Level Format Programs

Not all controllers have BIOSs on board, so they can't low level format from DEBUG. The manufacturers of those controllers offer low level format programs either via a disk shipped with the controller or

through some communications service like CompuServe, the Internet or a bulletin board.

Even some SCSI and IDE vendors offer these low level format programs, although of course in the vast majority of cases the only effect of these "low level format programs" is to erase all the user data on the disk, not the sector 1Ds, but that's not all bad-after all, you might want to zap the data for security purposes.

Partitioning

After the Hard disk has been physically formatted but before it is logically formatted, it must be partitioned. Partitioning means preparing the boot sector of the disk so that the DOS FORMAT program can operate correctly. Partitioning is also necessary to allow several operating systems to coexist on a single hard disk. The DOS FDISK program is the accepted standard for partitioning hard disks.

Clusters & Disks

The name of the file system used in all versions of DOS and Windows is called the File Allocation Table, or FAT system. One of the cornerstones of FAT organization is the fact that a FAT based system allocates space to files in units called clusters. A cluster is an area on a hard disk ranging from 2K to 64K.

When a FAT-based operating system (again, like Windows or DOS) allocates space, it does it cluster by cluster. Each cluster is allocated to one, and only one file. The size of a cluster on a logical drive is determined by the size of the logical drive, as you can see in the following Table:

SIZE OF LOGICAL DRIVE (IN MB)	CLUSTER SIZE (IN K)
0-16	4
16-127	2
128-255	4
256-511	8
512-1023	16
1024-2047	32
2048-4095	64

Suppose, then, that you save a 400 byte file on an 80 MB logical drive. Clusters on an 80MB drive are 2K in size, 2048 bytes, and so those 400 bytes are placed in this 2048-byte cluster, wasting 1648 bytes in the cluster. That's wasteful, but consider the greater amount of wasted space if you stored that 400 byte file on a logical drive 3GB in size, where clusters are 64k-65,536 bytes long. The cluster would have 400 useful bytes, and 65,136 bytes thrown away. Since cluster sizes are controlled by logical drive sizes, then keeping logical drive

sizes down keeps cluster sizes down, which leads to more efficient disk space usage.

Partitioning a Drive

Suppose We got that 1.2 GB drive and we want to chop it up into some smaller drives. It might look like Figure.

We use FDISK to do this. The disk gets divided into two partitions, a “primary DOS partition, “ and an “extended DOS partition.” The first step is to divide the example 1.2 GB drive into a primary DOS partition of 900 MB and an extended DOS partition of 300MB.

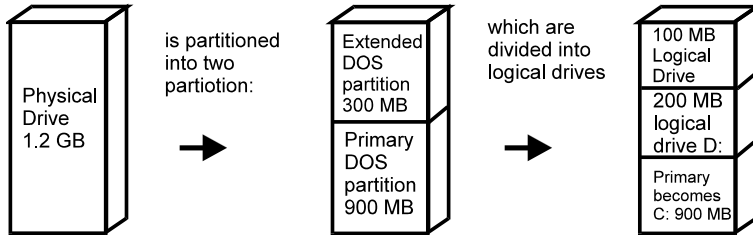


Fig 13.15 Partitioning Drives

Next, the extended DOS partition gets divided up into “logical” drives. You can have as many as you like each of any size. If you wanted, you could create drives D:, E:, F:,G:, H:, I:, and so on, each 1 MB in size.

First, start up FDISK. It shows a screen like the one in Figure:

```

Microsoft Windows Milleneium
Fixed Disk setup Program
(C) Copyright Microsoft Corp. 1983-1993

FDISK Options
Current fixed disk drive 1

Choose one of the following:

1. Create DOS partition or Logical DOS Drive
2. Set active partition
3. Delete partition or Logical DOS Drive
4. Display partition Information
5. Change current fixed disk drive

Enter choice: [1]

Press Esc to exit FDISK

```

Fig 13.16 Using Fdisk

The FDISK screen may look a little different from above because here we got five options on the screen. The fifth option, change current fixed disk drive, only appears if you've got two physical drives on the system. To create a DOS partition, press 1. The screen then looks like Figure:

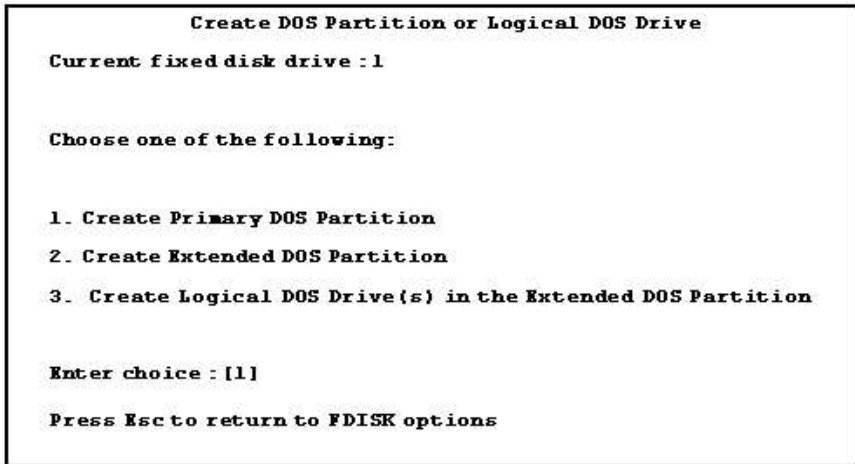


Fig 13.17 Creating DOS Partition of Logical Drive

The disk being used here is the master of an EIDE master/slave combination. The master becomes the first physical hard disk, and DOS requires that the first physical hard disk must have a primary DOS partition on it. It's perfectly acceptable for the second physical hard disk to only have an extended DOS partition without a primary, but the first disk must have a primary DOS partition. After I press 1 to create a primary DOS partition, the system thinks for a minute or two, and then a screen like the one in Figure appears:

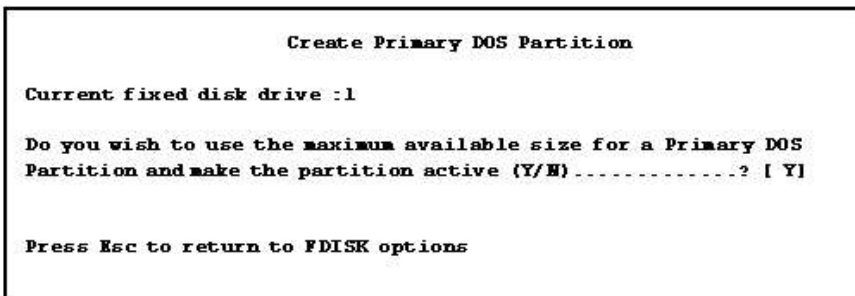


Fig 13.18 Creating Primary DOS Partition

This question asks whether or not to just make the C: drive one big partition. We don't want to do that, so press N. Since we don't want to do this all as one partition, the screen shown in Figure appears:

```

                Create Primary DOS Partition

Current fixed disk drive :1

Total disk space is 1203 Mbytes (1 Mbyte= 1848576 bytes)
Maximum space available for partition is 1203 Mbytes (100%)

Enter partition size in Mbytes or percent of disk space (%) to
create ia Primary DOS Partition..... (1203)

No partitions defined

Press Esc to return to FDISK options

```

Fig 13.19 Creating Primary DOS Partition

We can then either punch in the target megabytes of the primary partition or a percentage so we'll just enter 900 and then press. The screen in Figure then appears.

```

                Create Primary DOS Partition

Current fixed disk drive :1

Partition  Status  Type  volume Label  Mbytes System      Usage
C:         PRI  DOS          901  UNKNOWN
75%

Primary DOS partition created, drive letters changed or added

Press Esc to continue

```

Fig 13.20 Creating Primary DOS Partition

Notice that it's already assigned the drive letter C: to the partition because C: by definition is the primary DOS partition in the first physical hard disk. The system value is UNKNOWN because we haven't formatted it yet. As prompted, press Esc to continue. That returns the control to the original screen, but now it bears a warning message WARNING ! No partitions are set active disk 1 is not startable unless a partition is set active. This means that while you're allowed to partition your drive into up to four partitions, you've got to tell the system which partition to boot from. That partition is called the active one.

It can be remedy by setting the newly created partition to be active; just press 2, and a screen appears like the one in Figure.

```

                Set Active Partition
Current fixed disk drive:1

Partition  status  Type  volume Label  Mbytes System      Usage
C: 1      PRI  DOS           901  UNKNOWN
75%

Total disk space is 1203 Mbytes (1 Mbyte= 1848576 bytes)

Enter the number of the partition you want to make active..... [ 1 ]

Press Esc to return to FDISK options

```

Fig 13.21 Setting Active Partition

Now let's create the extended DOS partition, and chop it up into two logical drives. Press 1 to get back to the create partition menu, and this time press 2 to Create Extended DOS Partition. You 'll see the screen shown in Figure

```

                Create Extended DOS Partition
Current fixed disk drive :1

Partition  status  Type  volume Label  Mbytes System      Usage
C: 1      A    PRI  DOS           901  UNKNOWN
75%

Total disk space is 1203 Mbytes (1 Mbyte= 1848576 bytes)
Maximum space available for partition is 302 Mbytes (25%)

Enter partition size in Mbytes or percent of disk space (%) to
create an Extended DOS partition..... [302]

Press Esc to return to FDISK options

```

Fig 13.22 Creating Extended DOS Partition

Take all of the space that it offers, pressing Enter. It confirms Extended DOS Partition created. Press Esc and you'll see a screen like Figure:

```

Create Logical DOS Drive(s) in the Extended DOS Partition

No logical drives defined

Total Extended DOS Partition size is 302 Mbytes (1 Mbyte= 1848576 bytes)
Maximum space available for logical drive is 302 Mbytes (100%)

Enter logical drive size in Mbytes or percent of disk space (%) [302]

Press Esc to return to FDISK options

```

Fig 13.23 Creating Logical DOS drive

FDISK wants to create a single logical drive of 302 MB, but we don't want to, so we replace the 302 with 200 , and press Enter. The screen then looks like Figure

```

Create Logical DOS Drive (s) in the Extended DOS Partition

Drv   Volume Label      Mbytes   System      Usage
E:                201      UNKNOWN     66%

Total Extended DOS Partition size is 302 Mbytes (1 Mbyte= 1848576 bytes)
Maximum space available for logical drive is 101 Mbytes (20%)

Enter logical drive size in Mbytes or percent of disk space (%).....(101)

Logical DOS Drive created drive letters changed or added

Press Esc to return to FDISK options.

```

Fig 13.24 Creating Logical DOS drive

Now you can see that our first logical drive has been created at 201 MB, and that it's got the drive letter E:

FDISK shows that there's only 101 MB left, so press 1- to create the second logical drive. The screen looks like the one shown in Figure below:

FDISK indicates that there's no space left. Press Esc three times, and the system will reboot, the partitioning is done.

Create Logical DOS Drive(s) in the Extended DOS Partition					
Drv	Volume	Label	Mbytes	System	Usage
E:			201	UNKNOWN	66%
F:			101	UNKNOWN	34%

All available space in the Extended DOS Partition is assigned to logical drives.

Press Esc to continue

Fig 13.25 Creating Logical DOS drive

How DOS Names Partitions

If you have one physical drive with say, three logical drives defined on it, how are the drives named? Simple: the logical drive in the primary DOS partition is named C:, and the two in the extended partition are named D: and E:.

But suppose you have two physical drives that are divided each into three logical drives? How are they named?

Answer it depends on whether or not there's a primary DOS partition on the second drive. It is acceptable to DOS for the second physical drive to only contain an extended partition, unlike the first physical drive, which must have a primary DOS partition before it can have an extended partition. It is also acceptable to DOS if the second physical drive has both a primary and an extended DOS partition.

- If the second drive has both a primary and extended partition then the primary partition on the first drive is C:, the primary partition on the second drive is D: the two (in our example) logical in the extended partition on the first drive are named E and F, and the two logicals in the extended partition on the second drive are named G: and H:.
- If the second drive has only an extended partition then the primary partition on the first physical drive is C: again and this time the logicals in the extended partition on that first physical drive are named D: and E:

The three logicals in the extended partition on the second drive are named F:, G:, and H:, In above 1.2 GB drive example, we ended up with three drives on the hard disk with names C:, E: and F: because the second physical hard disk had a primary DOS partition and so laid claim to the D: drive name.

DOS FORMATTING

Finally you run the DOS FORMAT program. It does not actually overwrite sectors and physically format hard disks.

- **The partition record or Master Boot Record (MBR)** This contains the disk partition information to divide the physical drive into logical drives. On auto-configure controllers, it also contains a few bytes that describe the disk. This resides on cylinder 0, head 0, sector 1. The remaining sectors on the first cylinder/head are unused.
- **The DOS boot record.** Originally this only contained a snippet of code used to start booting up the system. In later versions of DOS, other disk ID information was added. Among other things the DOS boot record contains a pointer to the FAT, so if the boot record is damaged the FAT will look strange to DOS. The boot record resides on DOS sector 0, which is cylinder 0, head 1, sector 1.

DOS numbers sectors sequentially starting at cylinder 0, head 1, sector 1, Sectors 2 to 17 of that cylinder/head are the next 16 DOS sectors. The FAT starts there; its length depends of the size of the logical disk. The next group are the 17 sectors on cylinder 0, head 2. Next is head 3, and so on. First increment sectors until you're out of them, then increment the head numbers until exhausted then cylinders.

- **The FAT (File Allocation Table.)** The FAT is a map of what clusters are associated with what files. DOS keeps two copies of the FAT, the primary and the secondary. There are two kinds of FAT's a 12 bit FAT, with up to 4096 12 bit entries, requiring up to 6K total on the disk for each copy, and a 16-bit FAT, with up to 64K 16 bit entries, requiring up to 128 K for each copy. Under DOS 3X and later, floppies and 10MB hard disks use the 12 bit FAT, and 20+MB hard disks are the 16-bit FAT. DOS2.x uses the 12-bit FAT for all disk types.
- **The Root Directory** The root directory is the basis of the tree structured file system. There are 128 entries for a disk with a 12 bit FAT and 512 entries for a disk with a 16 bit FAT. Entries means room for directory information for a file .Twelve bit FAT disks can only have 128 files in their root directories. Try to create a 129th, and you get an file creation error message.

The root directory immediately follows the second copy of the FAT on the disk.

- **The Data Area(where the files go)** Actual user data goes here. It follows the root directory.
 - **IBMBIO.COM or IO.SYS.** If the disk is bootable, the first directory entry and the first cluster refer to the first "hidden file", IBMBIO.COM (for PC-DOS) or IO.SYS (or MS-DOS).
 - **IBMDOS.COM or MSDOS.SYS.** If the disk is bootable, the second directory entry will refer to the second "hidden file IBMDOS.COM (for PC DOS or MSDOS.SYS(for MS-DOS)

The FORMAT command creates the DOS boot record, FATs, and root directory- it does not touch the MBR or the user data area. Since the FORMAT program does not destroy or overwrite data in the data area, this implies that formatted hard disks can be recovered.

Hard Disk Software Installation

There's more to setting up a hard disk than FORMAT C:/S.

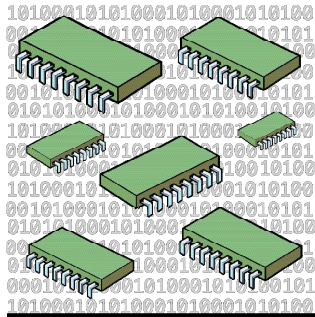
Support for multiple operating systems has caused hard disk formats to become more complex than floppy formats. Where a floppy only requires FORMAT A:a hard disk requires three steps. You may not have done all three as the first one (or two) is often done by the dealer. The three steps are :

- Low-Level or physical format
- Partition creation
- Operating system (high-level) format (DOS, OS/2, NT)

In general, you're not going to low level format a modern IDE, EIDE, or SCSI hard disk, so you'll probably skip that step altogether- but if your hard disk truly lets you low level format it, then low level formatting is a very important step.

EXERCISE

1. Write the configuration of any system of your lab.
2. Write down the sequence of installing various components while assembling the computer.
3. Once all the components are installed, what you have to do before installing various softwares into it.
4. Locate various ports available in your computer?
5. What is low level formatting?
6. What is the difference between DOS Partition and Logical drives.



CHAPTER 14

UPGRADING & TROUBLESHOOTING

NECESSARY TOOLS
BASIC WORKING TECHNIQUES
UPGRADING PC'S
TROUBLESHOOTING YOUR SYSTEM

Upgrading & Troubleshooting

Before you start working on your PC, you should take some precautions. Many computer components are very sensitive, and can be damaged by static electricity. Most of the time, you produce the static electricity yourself. Even the friction of your shoes on a synthetic carpet can generate enough charge to destroy a sensitive component.

In addition, you must never install PC components in a computer that is turned on. You shouldn't have any electrical problems if you observe the following rules:

- Do not wear shoes with a rubber or synthetic soles. If you have a carpet, these shoes can cause a voltage that is "lethal" to your components.
- Before you start working with PC components, you should ground your charge. To ground means that you get rid of any existing voltage by touching a grounding conductor. Touching a radiator, a faucet, or any similar metal object that is connected to the ground should do the trick.
- Avoid touching the chips on the components. Touch the cards only on the edges.
- Do not use any magnetized tools (such as magnetized screwdrivers) around the components. Their polarizing effect can damage certain chips.
- Make sure that the PC is disconnected from the power supply. Simply switching it off may be enough, but it is safest to actually unplug the power supply.
- Unplug the PC from the wall socket before you install anything. If you leave it plugged in, you endanger not only the component you are installing, but also the PC as a whole.

If you observe these suggestions from the start, you shouldn't have problems with your installations.

NECESSARY TOOLS

You'll need several tools for this installation. In this section, as well as in our video, we have listed the most important ones.

- A screw-setting screwdriver makes it easier to loosen and fasten screws, but a regular Phillips screwdriver will also do. You must have a screwdriver to do the installation.
- Flat-nosed pliers and tweezers help to set jumpers.
- A flashlight will help you see what you are doing inside the computer case. A small spotlight that sits on the floor or on your work table would be even better, freeing both hands for the installation.

BASIC WORKING TECHNIQUES

- Always turn off the monitor and the computer itself before starting to work on your PC.
- Unplug the power cord at the back to disconnect the power supply.
- Carefully loosen the other cables (the keyboard, the mouse, the monitor, and the printer cable). Some cables are simply plugged in, but others screwed into place. Before removing the latter, you'll have to loosen the screws.
- When all of the cables are detached, start removing the screws of the case, which are usually located at the back. Desktop cases have additional screws on the underside.
- When all the screws are removed, remove the case by pulling it to the back. This works for the tower as shown here, as well as for a desktop case.
- If you are installing a plug-in card, you need to remove the old card or, in the case of an empty slot, remove the cover on the computer case. Loosen the screw and pull out the cover or remove the old card.
- To install a drive in the computer, you'll need to remove a cover from the case. Simply push the cover from the inside out to release the drive duct.

UPGRADING PC'S

Upgrading a PC can be simple. You can do much of it yourself - even if you've never flipped open the cabinet cover. But you have to take care of the detail : you can't miss a step.

To carry out the upgrade.

Run CMOS setup and any other software installation. If it does not work, get into troubleshooting.

Of all the steps, the actual upgrade- plugging in the memory or drive, or whatever- is a small part. The main part is careful planning.

PC Workshop covers a range of PC upgrades and construction projects, both simple and advanced, in clear, practical steps. The idea is to let you carry out the upgrade, without a consultant or whiz kid friend around.

This month we cover simple upgrades : monitors keyboards, the mouse, coprocessors and memory. These are quick and clean, but can give you lots of benefits for the investment. And they're real easy, but they need planning.

In all cases, watch out for the planning process : what you require, what you need to buy. These processes are difficult to reverse if you make a mistake. After all this the installation is usually a breeze.

DISPLAYS: THE CARDS

If you need to upgrade your display, you usually need to change the display adapter card in the PC. You may also need to change your monitor. Keep in mind that some integrated, singleboard PCs from vendors such as Olivetti have the video circuitry on the motherboard, and don't let you carry out such upgrades easily.

Whether you're buying a new PC, or upgrading here's a straight, unqualified recommendation go for VGA.

The monitor and card will cost you upward of Rs 7,500. If you find some one who will swap you M/CGA display and card, you may get something for it in the range of Rs 1,000...3,000.

The easiest "upgrade" is a simple software setting that could have a miraculous effect.

For VGA, decide on your driver card and how much memory it should have. A good bet is to go for a Trident or similar card with 1 MB of RAM - so that you can run Windows at 1,024 x 768 (Super VGA resolution) and 256 colors. But actually, you should stick to 800 x 600 resolution with 14" monitors, because anything higher gives you real tiny text.

Should I buy a card for 256 colors even with a black and white monitor? Sure. You'll get smoother gray shades. With VGA, you must buy and use 256 colors. Most Windows users in India tend to use the basic 16 color VGA driver for Windows, even with a 256 color capable card. To change this, get hold of the window 3.1 drivers for your VGA card (Say Trident) and choose 800 x 600, 256 colors. Or, if you're using windows for Workgroups 3.11, use one of its own SVGA drivers.

Are there more sophisticated cards you can upgrade to? Sure, especially if you use Windows and intensive graphics. Accelerator cards can really speed up graphics on your 386 or 486 more than a processor upgrade could do More details next month.

Upgrading the Display Card

Keep yourself grounded say, barefoot on concrete, or touch a water pipe near to your workbench before you work. Never work on carpeted floors you can destroy your PC with static electricity.

- Remove the mains power cord from the PC. Remove the cover.
- Locate the old card. It's where your monitor is plugged in.
- Remove the monitor cord, unscrewing the mounting screws if necessary.
- Remove the screw that holds the adapter card in place, keep this screw you will need it for the new card.
- Lift the old card, gently rocking it out.

- Plug in the new card. If it doesn't appear to fit (possible because the old card was in a 8 bit socket) try any of the other sockets, after removing the covering rear bracket strip.
- Tighten the mounting bracket screw.
- Carefully check all the connections. Go through points 1 to 8 to ensure that they were followed properly.
- On a PC, you may need to run the SETUP program in case there are any error on bootup start up your system after connecting the monitor and keyboard.
- Close the cover, connect your monitor and restart.

MONITORS

If you're buying a new monitor, don't go for anything but VGA for SVGA).

Standard 14" mono monitors (black and white) capable of supporting SVGA cost Rs 3,0004,000. You can also get a 9" VGA monitor. This looks very cute and very sharp and is a great space saver (and conversation piece). But it's not very useful for Windows work; you have to squint a bit to read small text on it.

When you buy any monitor, check:

- Is the picture square, stable and sharp under windows?
- Is the picture nearly full screen? If there's a gap around it, it should be no more than a half inch gap, so that a 14" display gives you at least a 13" picture. Measure this diagonally bottom left corner to top left corner with a scale.
- Any gray shades are okay? Try to see what an illustration, or better still a photograph, looks like on it.
- Are the rotary controls for brightness, contrast etc., smooth and effective?
- When switching modes, say from windows to DOS (graphics to text), does the monitor behave itself, so there's no excessive picture bounce or jitter or snow?
- When you change a Windows resolution (driver) from say 640 x 480 to 800 x 600, does the picture stay full screen, and not vertically compressed (like a Cinemascope "film on Doordarshan)? If it does become narrower, there should be an accessible control knob for "vertical size" with which you can adjust the picture back to full screen.
- Is the tilt and swivel stand smooth and effective? Cheaper monitor vendors save here, and the stand doesn't work properly or becomes loose.
- Is there a power cable that plugs into your PC, rather than demanding a separate mains socket? This is better.

Color is great thing if you can afford it, but avoid cheap bargains here. A good SVGA color monitor starts at Rs 7,000. It should have a

rectangular picture, without bulges or other distortion. All other checkpoints are the same as for a mono monitor.

Two other, finer points need to be looked at, however, especially for color. One is dot pitch which dictates how fine the picture will be. Good color monitors have a 29 mm dot pitch, or smaller. Cheaper ones have a .39 mm dot pitch. You can always tell the difference on a Windows screen, though you can't measure dot pitch when buying. Anyway, watch out for this specifications on the vendor's sheet.

The other is the ability to handle a non interlaced display with a high refresh rate. A non interlaced display is smoother and has less flicker similarly, a 72 Hz or higher refresh gives a more stable picture. A monitor may be able to support up to 800 x 600 non interlaced but beyond that only interlaced: that's okay, but it should be specified in the data sheets.

KEYBOARD

If you're buying a new keyboard, check that the keys are smooth and can be operated with a gentle pressure, with no more than a faint click for "tactile feedback". Stiff keys with a loud "clack" get very troublesome later. Some keyboards use keys with smaller "travel" and a more rubbery feel. These don't feel as great initially but may last a lot longer.

Buy a 104 key keyboard specially designed for Windows 95 and upward.

Upgrading the Keyboard

- Power down your PC. Make this a habit even when you remove a keyboard or plug in.
- Remove the old keyboard connector. Align the new one a notch on it will correspondent with one on the socket, so it goes in smoothly only one way. Don't force it.
- When you position your keyboard, make sure you don't stretch the connector or connecting spiral cord. It's easily damaged.
- Power up the system and then test if the key board is functioning properly. Test if the Num Lock Caps Lock and Scroll Lock keys are okay.

MEMORY

This is the commonest upgrade, especially when people are considering running systems like Windows. In fact, as far as popularity goes, it matches a CPU upgrade.

Static RAM, which is fast and expensive, is used only for processor cache. If you have zero cache or only 64 kB of it, consider moving to 128 kB. Your system speed will improve.

Dynamic RAM is used for main memory, mostly in SIMM modules in modem PCs. Installing these is easy if your motherboard accommodates SIMMs. You need to know, however, that there are usually eight sockets in two banks of four each and each bank needs to be filled with the same SIMM type (size and speed); some board designs also demand that both banks use the same type. You has to fill at least one full bank. Using 1 MB SIMMs you get 4 or MB, 4 MB SIMMs give you 16 or 32 MB.

The latest PCI boards are having 4 banks for SIMM installation and may be 2 banks for DIMM installation (72 pins). The latest SIMMs and DIMMs are also coming in 16MB, 32 MB, 64 MB, 128MB and so on....

Removing or plugging in SIMMs or DIMMs is not difficult, though it takes a bit of getting used to. A few are screwed in most use small, spring apart, tilt the SIMM forward, and slowly pull it out. Always handle SIMMs by the edges, and never touch the chips or any metal connectors. To plug in the new SIMMs, reverse the process; insert the SIMM at an angle (it will go in only one way), and gently straighten up until it clicks Don't force it.

After ward, you need to get into CMOS Setup. Enter Standard CMOS Setup; the increased memory should be listed as extended memory there. When you exit and save, the new memory configuration will be stored.

MOUSE

A mouse is a must for Windows, even though in theory you can run windows without one. It comes as a standard mouse, typified by the Microsoft Mouse variety or as a trackball for the desktop. The standard mouse is easier to use, but you can get used to the trackball.

When you buy, make sure you use a serial mouse (unless you have a PS/2), and that the mouse has a good feel and buttons which give a soft but positive feel, and a faint click. A Microsoft mouse compatible mode will help ensure more compatibility, Also ensure that the connector will fit your PC. If the mouse has a DB -9 serial port connector, and your PC has only larger DB-25 types, you'll need a DB9-to-25 port adapter. Various mouse units to day, such as the Microsoft, ship with this.

The Microsoft mouse is the most comfortable of them, but it's also the most expensive - around Rs 1500. The Logitech is a fine mouse, too, at around Rs 600. You can also buy the more common True Mouse or Witty Mouse (around Rs 400) but these don't last you very long.

Installation is easy just plug it into the serial port. If there's are drivers for Windows and DOS software, install them, and set Windows up correctly. Specify COM 1: or COM 2:, and use this in setup.

In case of Windows 95 or higher version, the switch off the system and restart, it'll automatically detect the new hardware.

TROUBLESHOOTING YOUR SYSTEM

Simple operations that can help you diagnose and eliminate minor problems on your PC. Troubleshooting is something that most users believe is strictly for the maintenance guys. But do you know that most frequently occurring problems with PCs are minor and well taken care of by any user of average intelligence?

And no, you don't need to know about ISA buses and CPUs, any more than you need to know engine thermodynamics to change a spark plug.

Why should you do your own basic troubleshooting?

- Because you can, without being a techie.
- It saves you money.
- It helps you understand the system better.
- It lets you take better care of the PC after all, who would care more about your PC, you or an engineer who sees it twice a year?
- It can bail you out in a crisis where you can't wait for a repair person-just as knowing how to change a flat tyre can save you on the highway.
- But just as you and I don't open up our car engines, there are some things in PCs that are left strictly to qualified personnel.

This section will contain tips to troubleshoot common problems associated with the power supply motherboard, keyboards, monitors, mouse, hard disks, floppy disks printers, serial ports and modems. Rather than make this domain excessively garrulous, we will concentrate at the symptoms and the remedy to the same.

THE GROUND RULES

Whenever you are attempting troubleshooting of any sort, always follow the rules given below:

- Never test any equipment until you power it down and unplug the mains cord.
- If you have to work with or near live circuitry, ensure that you are properly insulated, sit on a wooden stool, or wear rubber slippers.
- For all other work especially with sensitive electronic circuitry you have to ensure just the reverse you need to be grounded (a wire between your metal watch strap and a water tap is okay), or stay barefoot on the floor not on carpeting. This way, you avoid having nasty static discharges damaging electronics. At least touch a grounded object water taps, window frames before working.
- Maintain a log of symptoms system not booting, monitor not working- along with the action you took to remedy it and what

happened after you fixed it. This can help you easily arrive at a solution.

POWER SUPPLY PROBLEMS

The SMPS, or the switched mode power supply, provides power to your PC's circuitry. Common problems include burnt out windings and blown power components. But the commonest problem is a blown fuse.

System dead:

- Is the cord firmly in the socket?
- Is the power cord okay- is there wear and tear, especially at either end? Replace with a working cord, just to be sure.
- Check if the socket can supply power. Run another electrical instrument like a fan or a lamp on the same socket and see if it works.
- Check if there are any blown fuses.

System comes on momentarily, then dies:

- Check the rating of the SMPS and the load on it. This sort of a problem commonly occurs on overload. Replace the SMPS with one of a higher rating in such a case. It could also be due to an inadequate CVT or servo or UPS, especially with a color monitor. If you find the PC running okay without the monitor, but resetting when you switch the monitor on this, is your problem.

System fails after working for a while;

- Check the load on the SMPS. replace the SMPS with a higher-rated unit. If required.
- Check the operating temperature. If it is too high, you may need more ventilation - perhaps an extra cooling fan, mounted on the case and radiating inward. Try keeping the CPU in a cooler place. 486DX2 systems without a heatsink or CPU cooling fan will fail in this manner after a while.

System locks up :

- Probably a software problem. But if it occurs while running simple DOS commands or in tested software, the symptom may have hardware faults. If you find that your mains supply is okay, run the PC through the full memory test on startup (don't bypass it.) Sometimes a BIOS setting lets you select this. You may have memory error-a bad RAM chip.

Loud groaning noise while working:

- Usually, a defective SMPS fan. Have it serviced or replaced. Some old hard disks and floppy disks are also noisy, especially if not mounted properly.

System reboots:

- Probably errant software or viruses. If it happens with all software and you haven't found a virus on scanning, check out the mains line- there could be other heavy equipment on that line. An air conditioner coming on can trigger such a reset in an improperly rated line when there's a PC on it.

MOTHERBOARD PROBLEMS

Motherboard problems are more difficult to tackle. The problem could also be because of some errant software, especially TSRs. If the problem occurs randomly, it could be a hardware fault.

Software problems:

- Clear all TSRs from AUTO EXEC. BAT and all device drivers from CONFIG.SYS. Reboot the system and use it normally for a while. If the problem does not reoccur, it means that one of the TSRs or device drivers is either corrupted, or at least fails to coexist properly with other software. Selectively try to load them one by one and see if your system hangs after loading a particular driver or TSR, don't load it.

Hardware problems:

- If the problem cannot be cleared by the above method, then it is probably a hardware fault, like memory chip fault, or a clash between IRQs or bad I/O chips.
- Run the BIOS SETUP utility and see if there are any conflicts, or any CMOS problems. The BIOS SETUP can be run by pressing the DEL key on bootup, in some systems, especially EISA ones. you press ctrl alt esc.
- Run utilities like QAPLus, which has a comprehensive set of tests for memory chips. If such a utility reports an error in the memory chip, it doesn't necessarily mean that the chip is bad there could be a loose connection. Power down, open the case and gently settle the memory chips down with your thumb (or straighten them, if they are SIMMs). If the chips are still trouble some, you may have to get them replaced.
- Ensure that all add on cards are firmly mounted in their sockets. Gently remove them and see if the edge connectors are dirty; clean them if required with cloth lightly moistened with petrol or spirit.
- Parity errors are most likely due to bad memory chips. Replace them if necessary. SIMMs are easy enough to replace- separate the metal clips on the edge lightly and swivel out the SIMM. Older RAM chips are tougher to extract you may need a technician to do this and the chips aren't easily available either.

KEYBOARD MAINTENANCE

The major aspect of maintenance for keyboards is abstinence. Abstinence, that is, from spilling things into the keyboard. The SafeSkin is one protection approach. Another is just to be careful.

Periodically disconnect the keyboard and remove it from the general area of the PC. Pull the keytops off of it. Then hold it upside down and blow it clean with compressed “air”. The keytops aren’t that tough to remove -We use one of those “chip puller” tools that come with PC toolkits (recall that they re dangerous to use to actually pull chips with), and find it to be just the tool for the job. If things are particularly messy, you can use a spray cleaner. In extremes, you can soak a sticky keyboard in water overnight - but make sure it’s completely dry before you try using it.

Keyboards Problem

Curiously the commonest reason for a keyboards not working is because it is either locked, or not fully in the unlocked position, or an XT /AT switch (behind the keyboard) that has gone over to the wrong position. Other reasons include a worn out cable, damaged connectors, or stuck keys.

- Ensure that the keyboard is not locked. Most BIOS will issue a message like “Keyboard Error. Press <F1> to resume” in case the keyboard is locked. Unlock the keyboard, check that the XT/AT switch under the keyboard is okay, and then press F1.
- Ensure that the cable is not worn out. Physically inspect the cable for wear and tear.
- Make sure that the keyboard is firmly seated in its socket. Unplug the keyboard and physically inspect the socket as well as the plug for wear and tear, Replacing the socket means using a soldering iron on the motherboard a qualified technician’s job.
- Use a utility like QAPLus to check the keyboard keys individually.
- Sometimes you can pop out individual keys and clean them in petrol or spirit. More often, you’d be able to open the keyboard and use an old (but clear) toothbrush to clear the dirt, though this doesn’t clean the key mechanisms themselves.
- Here’s a simple maintenance tip. Keyboards get damaged by little else than dirt. Keep them covered when not in use. And right now - reach for a clean sheet of newspaper, take the keyboard out, hold it upside down, and shake it hard over the paper

Troubleshooting

As with other items, you may want to view the keyboard as disposable. Replacement keyboards cost Rs. 850 or less from discount houses. There are some simple things which you can do before throwing out a keyboard, however.

Is It Plugged In?

On the back of many computers there are two identical ports, a mouse port and the keyboard port. Plugging the keyboard into the cassette port won't work. You could even destroy the keyboard.

Is It One key or All keys?

If only one key is malfunctioning, check that key's spring. Remove the key by grabbing it with your fingers and pulling up. For the tough keys, fashion a hook from a paper clip or, again use a chip puller. Under the key you will see a spring.

Replace the key cap and see if the problem goes away. If not, try pulling the spring out just a little. Then replace the key cap.

Checking the Cable Continuity

Next test continuity of the cable. Turn the keyboard upside down so that the cable is coming out of the back of the keyboard, to the right. Remove the two screws. The bottom plate will swing back and up to remove.

You will now see the cable splits to a single wire, which is grounded to the bottom plate, and a cable with a flat-jaw connector. Push apart the jaws of the connector to release. You can then test each of the five wires for continuity with your ohmmeter.

MONITORS PROBLEM

Many problems with a monitor are related to improper setup, or a corrupted setup. A monitor is solid state, with few mechanical parts to wear out (except for switches and controls, which do occasionally give trouble).

Monitor does not work, repetitive beeps on bootup:

- A monitor has two cords -one for the power supply, and one for receiving data from the computer. Check if both of them are properly connected. Most monitors receive power through the Pc unit, hence do not need to be separately switched off. Sometimes you might forget to check this.
- Run the BIOS setup utility on bootup, and check if the monitor type is correctly entered. You may need to use a working monitor from elsewhere, because if the display is dead you can't run setup very easily!
- Check the display adapter card in the PC and ensure that it is firmly placed in its slot.

- Check the motherboard jumper to see if the monochrome \CGA settings are proper- these jumpers are usually well labelled.

Monitor LED works, but no display:

- Check the display's brightness and contrast controls.
- Check the cable attached to the display adapter.

Monitor works, but produces jumbled display:

- Adjust the vertical/horizontal hold knobs at the rear of the monitor.
- The PC may be using a display mode not supported by the monitor.
- Software driver being used is either corrupted, or not suited for the monitor (usually in Windows).

Monitor makes a hissing sound:

- The monitor is being set to a mode that it does not support often a software problem. Run the setup utility of the software and choose a different type of monitor.

MOUSE CLEANING AND MAINTENANCE

On a mechanical mouse, the little wheels get gunk in them, so clean them. Just remove the mouse ball and examine the wheels. Clean the mouse about twice a year, unless your desk is covered with dirt or dust, as dirt and dust are just gunk in its fetal stages.

On an optical mouse, just clean the sensors, you can often just polish them with a soft cloth. Keep the mouse pad clean as well so that the grid lines are visible to the sensors.

Mouse Problem

Mice are prone to frequent problems due to their mechanical nature, and their ability to pick up dirt.

Mouse does not work:

- Check the connections. Check that the mouse software driver is loading and the correct port- COM 1 or COM2 is indicated. Most modern mice, including the Microsoft, auto detect this. In case of a bus mouse, check that the mouse plugs into the port, and the mouse card plugs properly into the expansion bus.
- Make sure that the mouse driver is properly loaded in the AUTOEXEC.BAT or CONFIG. SYS.

Erratic cursor movement:

- In trackball mice, this usually happens because the mouse trackball is dirty. In optical mice this happens because the

mousepad is dirty and the light sensor cannot track the movement precisely.

- Remove the ball and wipe it with a clean piece of linen, and clean the movement sensors which are visible when the ball is removed. A “Johnson’s ear bud” dipped in petrol or spirit is ideal. Then replace the ball and turn the retaining latch to lock it. Do not dip the ball in alcohol. If it is very dirty, a cloth lightly moistened in petrol or spirit will clean it adequately.
- In case of an optical mouse, clean the pad, the sensors below the mice, and try again. Make sure the pad is properly oriented some pads do not work when rotated 90 degrees.

Mouse pointer moves only horizontally or vertically:

- This problem occurs only in trackball mice. There are two movement sensors which detect horizontal and vertical motion. If the vertical one of them is dirty, the mouse can only sense horizontal motion and vice versa. Clean that sensor. If it’s defective a maintenance shop may be able to change it or change the mouse.

The buttons do not work:

- This problem occurs more often in cheaper mice the buttons give way first. Unfortunately, this sort of problem cannot usually be rectified, though some vendors may give you a replacement. If he doesn’t and you are forced to buy a new mouse, go in for a better brand this time. The leftmost button usually gives way first. If you’re good with a soldering iron, you can swap the two buttons, or better still, remove the middle button and solder it in the leftmost button’s place. If you’re not you can swap left and right buttons in windows to use the mouse for a while more.....

HARD DISKS PROBLEMS

Hard disks are complex, and trouble shooting them is complicated by the fact that problems can arise because of the hard disk itself, the controller or the software.

System does not boot from the hard disk:

- On power on, If you don’t get any errors during the POST, and instead get the message “Non system disk or disk error...”, the system is not booting from the disk.
- Check that there is no floppy in drive A. If there is just remove the floppy door and press any key. You can avoid this problem by setting the “Drive bootup sequence” to C, A” and the Floppy drive seek at boot to disabled” in the BIOS setup.
- The harddisk is probably not formatted. Boot from a DOS bootable disk and format the hard disk.
- The CMOS information is probably corrupted. Use the BIOS setup utility to ensure that the hard disk type is proper.

- The MBR (Master Boot Record) on the hard disk may be corrupted. Boot from a floppy and use the DOS utility SYS to make the hard disk bootable.
- There could be something wrong with either the controller or the drive itself- a job for a qualified technician.

Files on disk corrupted:

- This could either be due to logical errors related to the boot sector or FATs, or to physical errors on the disk.
- Run a utility like CHKDSK to check for logical disk errors. Use the "/f" switch of CHKDSK to clear errors, replying N when CHKDSK asks whether you want to convert lost chains to files. If you have DOS 6. use SCANDISK.
- Use a utility like NDD or SpinRite to check for bad sectors. If you have DOS6, you don't need to use them-SCANDISK will do the job for you.

Drive too slow:

- Check the BUFFERS directive in the CONFIG.SYS- it could be too low.
- Use a caching software like SmartDrive.
- Fragmentation could be the cause- use Compress or SpeedDisk to defragment your drive.
- There may be controller or hard disk media errors-contact your vendor.
- The interleave could be wrong use SpinRite to correct the situation. But not with modern IDE drives- they don't worry about interleave.

Critical errors on the hard disk:

- If your hard disk gives critical errors (Ones that end with "Abort, Retry, Fail?"), then you may be in trouble.
- If possible, make a complete backup of your hard disk. If it's an older disk, do a low level format with either the BIOS SETUP or a third party software or type G= C800:5 from within DEBUG. Do not low level format IDE drives. Run FDISK to partition the drive, and format your logical drives using FORMAT. This will remove many errors.
- If it is not possible to backup your hard disk, use a utility like NDD or Spin Rite to salvage as much data as possible. Then, backup your disk and follow the steps given above. Just because some utility salvages your data, it doesn't necessarily mean that the problem is solved you still need to run NDD or reformat the drive.
- If problems persist, it would be advisable to change the drive or the controller which ever is malfunctioning. Contact your vendor for advice and replacement.

Floppy disks problems

Floppy drives are rugged, but none the less have their problems- from dust, heat and misuse.

Read/Write errors on floppy:

- The disk could be damaged-try formatting it to mark bad sectors. If you get errors like "Track 0 bad", get rid of the diskette it is useless after checking it out on a known, working drive.
- The disk could be of the wrong density- check the density of the drive and the diskette. A 360 KB disk can be read in both 1.2 MB as well as 360 KB drives. But a 1.2 MB disk will only work properly on a 1.2 MB drive.
- Drive heads may be dirty use a head cleaning disk with the accompanying solution to clean the drive heads.
- The drive heads may be damaged or misaligned this is a job for your vendor. Get them fixed.

System refuses to read from drive:

- The drive may not be getting any power. At the DOS prompt, type DIR A: (or B:) The drive light must glow. If it doesn't then the drive is probably not getting any power. Open the case and test the power supply to the drive. If it glows, but still doesn't work, there is probably something wrong with the data cable- check the connection of the data cable on both sides to ensure that it is proper.
- The controller may have failed, or the drive may be having problems related to head alignment call your vendor.

Drive does not read floppies formatted on other drives:

- The disk and the drive may be incompatible for example the drive could be a 360 KB one, and the disk 1.2 MB. Or this could be a 360 KB disk you've formatted on a 1.2 MB drive - even with a correct disk and format, a 360 kB will often not read that disk!
- The drive heads may be misaligned-consult your vendor.
- The controller may have failed- consult your vendor.

TROUBLESHOOTING PRINTERS**Components of Printer**

Printers can be a real maintenance headache. As they produce tangible results (pieces of paper), malfunctions with printers can be more upsetting than their wholly electronic brethren. For example, try telling Windows that you've got a Hewlett-Packard laser printer configured as a PCL 5 device when you actually have a PostScript HP printer, then print a graph. The printer will start spewing out pages, as if angry for being misrepresented.

As with other peripheral troubleshooting, start your diagnosis from the three main pieces: a controller (the parallel or serial interface board), a cable, and a printer. The parallel port is simple enough that it is commonly included on other expansion boards (although it's undergone some changes over the years).

Parallel Ports

The parallel port was originally devised as a high speed, low-cost alternative to a serial port for printer interfaces. Until 1976, you'd connect a cheap (you know, under \$10,000) printer to a minicomputer or a microcomputer via an RS-232 port.

The problem with doing that was that serial port hardware was expensive. Adding a serial port to a computer could raise the computer's price by \$250 to \$1000. That got in the way of selling printers, and so a printer company, Centronics, decided to do something about it. RS-232 allows cables of up to 50 feet in length; the Centronics interface is only reliable for six feet, actually less for modern ports. RS-232 is bi-directional and quite flexible as to the kind of things it can support; Centronics was designed to be unidirectional and really was only aimed at printers. RS-232 is serial, employing only two wires for data; Centronics is an 8 bit parallel interface, with eight wires for data. The result was a Centronics interface that was easy to cobble together for just a few dollars and offered terrific throughput, as much as 500,000 bits per second compared to RS-232's 20,000 bits per second.

Since then, the parallel port has continued to evolve. First, some laptops offered bi-directional parallel ports. Then those bi-directional parallel ports were further modified to become Enhanced Parallel Ports, EPP ports for short. EPP ports include higher speed transfer capabilities. Some printers have bi-directional parallel ports as well and so those printers can talk back to their computers, informing them of things like what exact kind of printer they are, or sending error messages like out of paper rather than the generic printer error that's the greatest level of detail available on unidirectional parallel port interfaces.

Maintenance

A few things can be done to maintain printers of all types. Vacuum out the paper chaff periodically from the inside of the printer. Determine if there is a belt tightening mechanism for the printer- usually a motor moves the print head via a belt. Find the correct tension values. Keep a replacement belt on hand.

If you're using an impact printer like a dot matrix or daisy wheel printer (and if you're using a daisy wheel, it's time to reevaluate your priorities in life), use a dry, soft cloth to clean both the paper path and the ribbon path. Most manufacturers suggest cleaning every six

months, as the ribbon path can build up a film of inky glop which causes the ribbon to jam. To do this, go to a drug store and buy a dispenser box of 100 clear plastic gloves. Use them when working on the printer (but not chips and boards- that plastic can build up some mean static) so that you don't have to wash your hands for hours to remove the ink.

Most printers do not need to be lubricated in every day use. In fact, oil can do considerable damage if applied to the wrong places. If you thoroughly disassemble the printer, then you will probably have to lubricate various points as you reassemble it.

Here's a tip the will extend the life of both your ribbon and your head, put some WD40 lubricant on a used ink ribbon. Let it soak overnight. It'll produce good output the next day, and you won't damage the print head- WD40 is a good lubricant for print heads. Let me stress however that this only applies to ink ribbons. If you have a thermal transfer printer, like an IBM Quietwriter or a Okidata 20, this will not work.

Daisy Wheel Printers

Daisy wheel printers require that the daisy wheels be replaced about annually. Some variations, like Spinwriter thimbles, last longer but must be replaced eventually. (The Spinwriter also needs a striker shield replaced every few months. It is a thin flat piece of plastic which sits near where the thimble strikes the paper.) Their belts must be replaced now and then.

Be sure to replace the printing elements-daisy wheels and the like-before they wear out, as a worn daisy can damage a printer. You know when they're wearing because you see fading on the more common characters first, the "e" and the "t"- between them these two letters constitute about 20 percent of the characters in typical English text.

Dot Matrix Printers

The expensive part of a DM printer which dies is the print head. Print for too long a time and the head just burns up. This is not as much of a problem for the newer printers as it was for the old Epsoms and Okis. The models out today have a thermistor which shuts down the printer temporarily if the print head overheats. If the thermistor becomes ill, the printer shuts down regularly. Generally thermistors are pretty robust, but if such a thing happens try changing the thermistor first, it's a lot cheaper than a print head.

Replacing print heads is not economical on many printers because of the high price that manufacturers charge for replacements, often it is a significant percentage of the entire cost of the printer.

Dot matrix printers are in general, very reliable, but keeping the print head cool is vital. Don't stack things on or around the printer.

Ink Jet Printers

A variation on the typical dot matrix printer is the ink jet printer. Rather than hammering at a ribbon, the ink jet squirts a narrow jet of ink at the paper. This is very quiet, but the jets tend to clog, leaving partial letters on the page. The answer here is simple remove the cartridge and push on the ink sack with a long thin tool, like a straightened out paper clip. The ink will push out the small holes, unclogging them. With the HP Thinkjet, it's a pretty regular procedure. The Seiko color ink jet has the same problem.

Thermal -Transfer Printers

When you were a kid, did you ever make pictures by coloring with crayons onto a piece of heavy paper and then using an iron to transfer the colored wax to another piece of paper. Thermal transfer printing works something like that. Very hot pins are pressed onto a wax or wax/resin coated ribbon, and the wax or resin melts and is transferred to the paper beneath it. The difference is that the paper has to go through the process four times, once for each color (normally cyan, yellow magenta, and black).

The issues involved with thermal-transfer printers are much like those with other printers: keep them cool and keep them plugged in. If you're having problems with print quality, such as smearing, you might try using another manufacturer's paper or ribbons.

Laser Printers

The laser printer is very similar to a copy machine. Having said that, it's amazing that they are as reliable as they are.

The most common laser engine is made by Canon. The HP LaserJet, the AppleWriter, the Canon A1/A2 the QMS Kiss, and others are all built around the Canon engine. These need no maintenance except for a new cartridge every 3000 copies or so. The cartridges cost in the neighborhood of \$70, and according to HP, contain all that is needed for routine maintenance. So, you are performing routine maintenance every time that you change your cartridge.

It's okay to recharge cartridges .Make sure your refill company completely rebuilds the insides, including replacing the photoelectric drum. Avoid the "drill and fill" vendors, as they don't replace the insides, and using that kind of refill will lead to a lower quality print image and may damage the laser printer.

Lasers require proper ventilation and a fair amount of power. Other than that, don't pour any Cokes in them and they last a long time. Never ship a laser with a toner cartridge in place. It can open up and cover the inside of the laser with toner.

Infrared Printers

For members of the "virtual office" who have no desk but only a laptop to carry around and plug in as needed, there's a way to eliminate at

least one of the cables you need to plug in printers with infrared ports are available. Rather than requiring a parallel cable to connect the printer to your computer, a ray of infrared light shines between a transceiver on the computer and one on the printer.

Data travels between the computer and printer via an infrared connection. Light, sound, and electric impulses all pulse at a certain rate per second: this rate is called the frequency-the more pulses, the higher the frequency. Signals with higher frequencies can transmit data more quickly (each pulse can carry a bit of data), but have a shorter range and are more prone to interference than lower frequency signals, because anything that interferes with the signal will affect more data than it would if there were fewer pulses per second. Infrared light has a high frequency. Thus, the computer sends a beam of infrared light to the receiver on the printer, The devices have a pretty good range.

At this point, computers and printers with infrared ports exist but aren't terrifically common. IBM's Thinkpad is one computer than has an infrared port, and HP's LaserJet 5P has the infrared receiver, and can also work as a regular printer.

If you're having trouble with a wireless printer, check the following:

- Are the infrared ports on the printer and computer both clean and unblocked?
- Does it help to move the printer/computer to one side?
- Did someone stand between your printer and computer during the print job?

Other than that, troubleshooting an infrared printer is much like troubleshooting any other (probably laser) printer. The tricky thing about wireless printers is the connections.

Printers Problems

Most problems related to printers spring from two factors- careless handling and forgetting basic tasks like connecting the data cable or the power cable. If your printer fails to print, do not panic-just check that it is properly connected to the right port (usually parallel; a few use the serial ports), that its power cord is properly connected and the power switch is turned on.

Printer does not power on:

- Check if the power cord is properly connected.
- Ensure that the printer is switched on.
- Inspect the power cord, physically and then with a multimeter (if available) for connectivity. Try a different power cord, one that is known to work. If the printer turns on this time, the problem was probably with the cord- replace it.

- Check for blown fuses, and ensure that the power socket can supply electricity. Run other electrical equipment like a fan or a lamp to make sure.

Printer works but prints garbage:

- This could possibly be a software problem. Check software settings, and load a different software driver if necessary. An Epson driver should work with all printers, except for Post Script printers. Turn the printer off, wait momentarily and then turn it on again. This will clear the buffer.
- Most printers also include a self test. Run this test to ensure that all characters are properly formed.
- Many printers use DIP switches- check with your manual, and set the switches are set properly. This often affects their “emulation mode”- and therefore how they print.

Printer works, but gives very dull output:

- On a DMP, check the ribbon it is probably old and worn out. Replace if necessary. Alternately, you can also have it reinked, but avoid reinking a ribbon that is very old. If the nylon appears to be coming apart, throw the ribbon- the nylon pieces can damage the head.
- If this occurs on a new ribbon reduce the distance between the ribbon and the paper on the other hand, if it starts smudging, increase the distance. This is done by a lever whose exact position varies from printer to printer- the only way to find out where this lever is located on your system is to see the manual.
- In very old printers with a new ribbon, this usually indicates that the head is defective replace it.

SERIAL PORTS PROBLEMS

A number of external devices like modems and mouse connect to the computer through a serial port. For this reason, troubleshooting is hampered by the fact that the problem could be associated with either the external device, or the serial port itself. But more often the problem occurs because of improper serial port settings- check them before you proceed with further troubleshooting. In most cases, it may be all that you need to do.

External device appears dead:

- Check power to device. If the external device has any sort of self test, perform it.
- Ensure that the cables are plugged into the correct sockets.
- Ensure that software settings like speed, parity, etc, match hardware settings.
- Check cable wiring, try another cable if possible.

External device works, but produces junk:

- Ensure that software settings like speed, parity, etc. match hardware settings.
- Perform device self test if possible.

Loss of data during transmission:

- This occurs when one device transmits while another is not ready. Check cable for wear and tear and ensure that handshaking is enabled.

Characters are lost randomly:

- Reduce the communication speed.
- Use a serial card with a 16550 UART chip.

External device hangs after some time:

- Check for possible hardware errors like worn out cables. or loose connections.
- Deactivate all TSRs and try again-if the error doesn't occur again, don't use the particular TSR which caused the hangup.

MODEM PROBLEMS

Modems are pretty hard to trouble shoot, mainly the internal variety. The external ones are easier to handle, largely because of the different LEDs which indicate the status of the modem and as such are a boon to anybody troubleshooting a modem.

Modem unresponsive:

- Modem not powered-check power cord, cable and fuse.
- Data cable faulty- try another one.
- Ensure that the data cable is connected to the right serial port.
- Comm software using wrong parameters-check and use the correct parameters.
- Modem command echo could be off- try sending ATQOE1 to your modem via your comm software.

Modem doesn't dial:

- Modem probably not receiving command- check for faulty connections. Wrong dialling method-pulse instead of tone or vice versa specified.
- Modem dials, but doesn't connect: Check modem configuration it could be wrong.
- Remote modem could be using a different mode determine remote modem setup and configure your modem accordingly.

Modem connects, but outputs garbage:

- Faulty modem at either end.
- Bad line conditions- try a lower speed.
- PC to modem link could be using a wrong baud/parity setting. Ensure that the settings are the same.

File transfer errors:

- Bad line conditions and no error correction ensure that your modem uses MNP 4 error correction protocol. Disk full errors- this happens only for down loads.
- Too less time for download- some BBSs will refuse to download if time for download is insufficient. Use a ZModem protocol- it allows you to start a file transfer in one session and continue it in another.

File transfer takes too long:

- Check transmission speed it could be too low.
- Ensure that your modem uses MNP5 data compression protocol. If it doesn't, try software emulation MNP- even though it is slower than hardware compression, it is better than no compression at all. Not all modems at the other end will recognize this, though.

Computer Hardware & Peripherals

Everything you wanted to know about.....

About the Author

Munishwar Gulati is a Graduate from World famous IIT Roorkee, and has about 15 yrs of experience in computer industry. He has written many computer based titles, from DOS to Windows 2000, to COBOL, to desktop publishing, to Visual Basic, to web designing and programming etc., for various training centers all over India including DOEACC, CEDTI, ET&T, HARTRON, HILTRON etc.

Co-Author Mini Gulati is a graduate in Computer Science and has rich experience in analyzing systems and developing ERP solutions for small and medium scale industries. She is also co-authoring many of the upcoming titles.

ISBN 81-87870-21-4



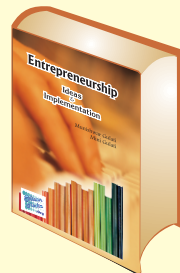
Other titles in Everything..... Series

- Introduction to Computers
- Windows 98 and Information Technology
- Word/Excel/Powerpoint 2000
- Office XP/2000
- Access 2000
- Foxpro for Windows
- Business System and Tally
- Programming in C
- Programming in C++
- Unix/Linux
- System Analysis and Design
- Adobe Photoshop 6.0
- HTML/DHTML 4.0
- VBScript
- Oracle
- Core Java
- Information Technology
- Fundamentals
- Basic Electronics
- Networking and Windows NT

Entrepreneurship Ideas & Implementation

Number of Pages	525
Language	English
Price	Rs. 300/-

Whether you are presently running your own business or just planning to start one, this book can be the best business investment you have ever made. Entrepreneurship - Ideas & Implementation, gives you all the guidelines and tools you need to start, operate and succeed in a business of your own.



SILICON MEDIA PRESS

www.siliconmedia.org