May, 2015

FEDERAL UNIVERSITY WUKARI

CSC 208 – HARDWARE SYSTEM AND MAINTENANCE

COMPUTER SCIENCE DEPARTMENT | AGU, Edward O.

Introduction:

A computer system is an electronic machine which accepts inputs through input devices, process the input using processing elements, store the processed information if need be using storage memory and gives out output through the output devices.

A computer system is composed of both hardware and software. The hardware is the equipment, or physical components, that make up the system while the software is the set of program instructions executed by the computer. Hardware consists of the machines that comprise a computer system, such as all the mechanical, electrical, and magnetic devices within the computer itself and all related peripheral devices. The peripheral devices are the input and output devices, which are outside the main part of the computer.

FUNCTIONAL UNITS OF A COMPUTER SYSTEM

Digital computer systems consist of three distinct units. These units are the input unit, processing unit and output unit. These units are interconnected by electrical cables to permit communication between them, allowing the computer to function as a system.

INPUT UNIT: Is the unit through which data and program instructions are read into the computer for processing. A computer must receive both data and program statements to function properly and solve problems. The input device is the means/mechanism of feeding data and programs into the computer. Computer input devices read data from a source, such as tapes and disks, and translate them into electronic impulses for transfer into the central processing unit. Examples of input devices are keyboard, mouse, scanner, joystick, light pen, and monitor. The input device helps the computer by sending data and programs into the system.

PROCESSING UNIT: The processing unit is the brain of a computer system, which is the central processing unit (CPU). The CPU processes data transferred to it from one of the input devices, processes it and then transfers either an intermediate or final result of the CPU to one or more output devices. The CPU is a chip containing millions of tiny transistors. It is the CPU's job to perform the calculations necessary to make the computer work and the transistors in the CPU manipulate the data. A CPU can be regarded as the decision maker and the computing center of the system.
The central processing unit (CPU) contains the electronic components that interpret and execute the functions needed to operate the computer. The CPU contains two major components: the arithmetic logic unit (ALU) and the control unit (CU). The arithmetic logic unit (ALU) is the component of the CPU that performs arithmetic (e.g. addition, multiplication) and logical operations (e.g. is the statement A and B true or false). The control unit (CU) is the component

of the CPU that directs and manages the execution of computer's instructions. Memory stores data and instructions for use by the CPU. There are two types of memory: random access memory (RAM) and read only memory (ROM). Random access memory (RAM) is the memory that can be read from and written to by the CPU. It is sometimes refers to as "working memory" or "main memory." Read-only memory (ROM) is the memory that can be read but not written to by the CPU. It is used to store permanent data and instructions. The CPU and memory are connected by a bus, a band of wires through which information is exchanged.
The CPU acts as the brain of the computer system with components: the control unit (CU), the arithmetic/logic unit (ALU), and the primary storage unit. Each unit performs its own unique functions. The control unit, an arithmetic/logic unit, and an internal storage (or main memory) is shown in Figure 2.3. Each unit within the CPU serves a specific function and has a particular relationship with the other units within the CPU. It is the CPU that executes stored programs and does all of the processing and manipulating of data.

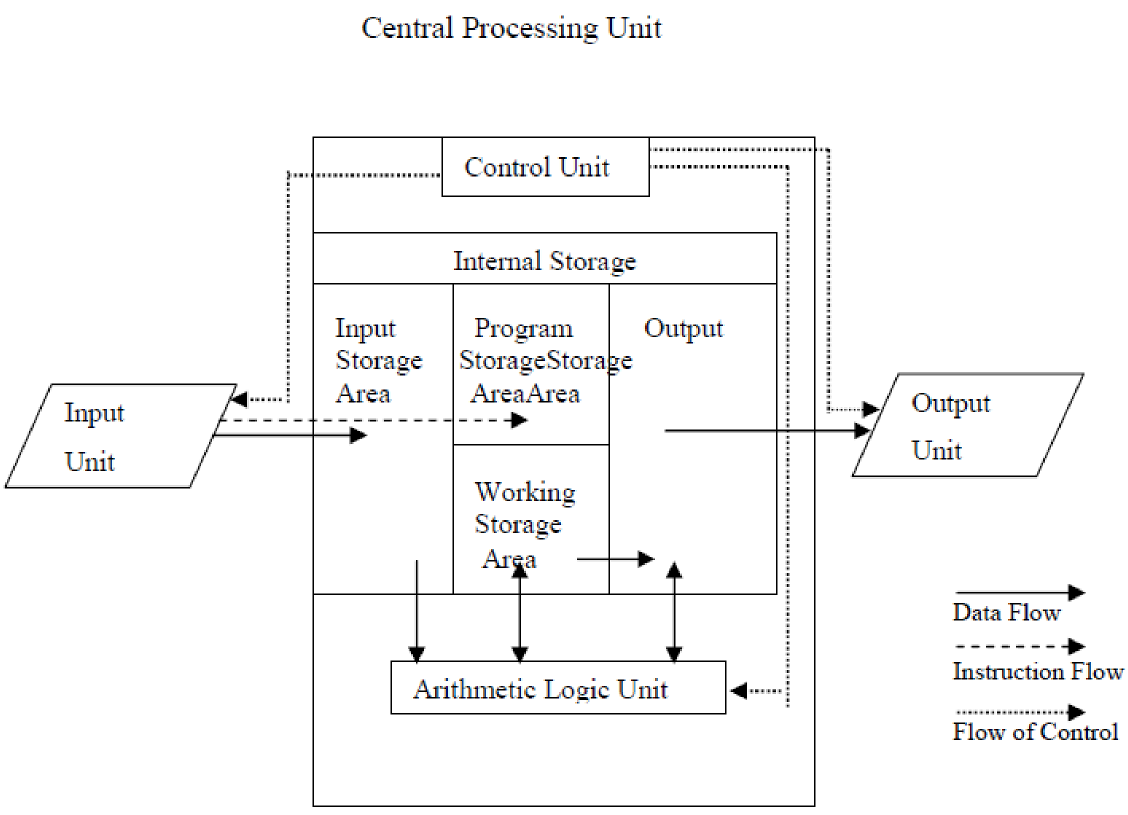


Figure 2.3: Functional units of the computer system

Control Unit: **-** The control unit is the circuitry that controls the flow of data through the processor, and coordinates the activities of the other units within it. It is considered to be the "brain within the brain", as it controls what happens inside the processor, which in turn controls the rest of the PC. The control unit directs the flow of operations and data; and maintains order within the computer. Dotted arrows in figure 2.3 indicate the flow of control.

The control unit selects one program statement at a time from the program storage area, interprets the statement, and sends the appropriate electronic impulses to the arithmetic/logic and storage units so they can carry out the instructions. The control unit does not perform actual processing operations on the data. The control unit instructs the input device on when to start and stop transferring data to the input storage area. It also tells the output device when to start and stop receiving data from the output storage area. The outputs of the control unit control the activity of the rest of the device.

A control unit can be thought of as a finite state machine. The control unit controls what is happening in the CPU. It does not process or store data; rather, it directs the sequence of operations. The control unit retrieves one instruction at a time from the storage unit. It interprets the instruction and sends the necessary signals to the ALU and storage units for the instruction to be carried out. This process is repeated until all the instructions have been executed. Another function of the control unit is to communicate with the input device in order to transfer program instruction and data into storage. It communicates with the output device to transfer results from storage to the output device.

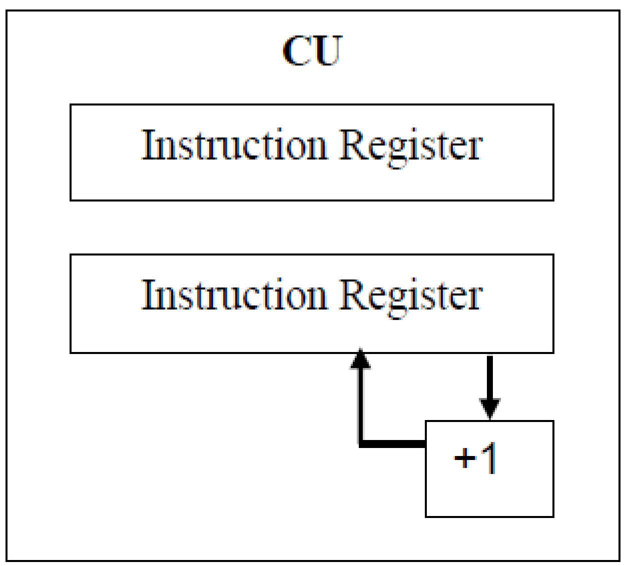


Figure 2.4(a): Control Unit Operation

Specialized electronic circuitry in the control unit is designed to decode program instructions held in the main memory. Each instruction is read from the memory into the instruction register as shown in figure 2.4(a). The process of reading an instruction is often referred to as the fetch-execute process. The CU directs the CPU through a sequence of different states. A device called the system clock governs the speed with which the CPU cycles from state to state. Clock speed is often measured in gigahertz (GHz) where a Giga hertz is one billion cycles per second. Thus, a 2.9 GHz processor could execute 2.9 billion cycles in one second.

***Arithmetic/logic Unit:* -** An arithmetic logic unit (ALU) is a digital circuit that performs arithmetic and logical operations. The ALU is a fundamental building block of the central processing unit (CPU) of a computer, and even the simplest microprocessors contain one for purposes such as maintaining timers. The processors found inside modern CPUs and graphics processing units (GPUs) accommodate very powerful and very complex ALUs. A single component may contain a number of ALUs. The accumulator is used to accumulate results. It is the place where the answers from many operations are stored temporarily before being put out to the computer's memory. The other general-purpose registers hold data on which operations are to be performed by the ALU as shown in figure 2.4(b).

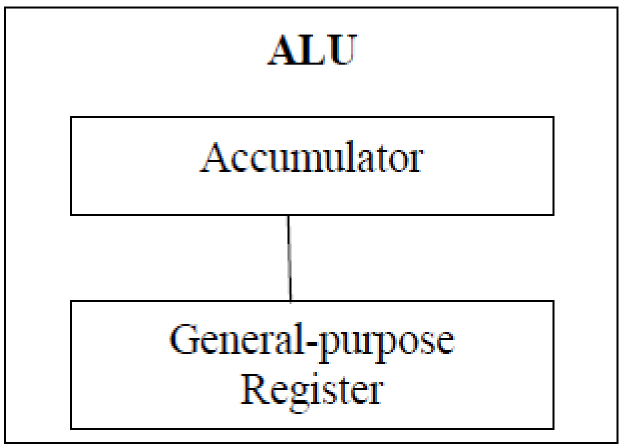


Figure 2.4(b): ALU Structure

The ALU executes arithmetic and logical operations. Arithmetic operations include addition, subtraction, multiplication and division. Logical operations compare numbers, letters and special characters. Comparison operations test for three conditions: equal-to condition in which two values are the same; less-than condition in which one value is smaller than the other; and greater-than condition in which one value is larger than the other. Relational operations (=, <, >) are used to describe the comparison operations used by the ALU. The ALU performs logic functions such as AND, OR and NOT. Other operations are bitwise logic operations (AND, NOT, OR, XOR); and bit-shifting operations (shifting/rotating a word by a specified number of bits to the left or right, with or without sign extension. These shifts can be interpreted as multiplications by 2 and divisions by 2). Through internal logic capability, it tests various conditions encountered during processing and takes action based on the result. As indicated by the solid arrows in figure 2.3, data flows between the arithmetic/logic unit and the internal storage unit during processing. Data is transferred as needed from the storage unit to the arithmetic/logic unit, processed, and returned to internal storage. Processing does not take place in the storage unit. Data may be transferred back and forth between these two units several times before processing is completed. The results are then transferred from internal storage to an output device, as indicated by the solid arrow in figure 2.3.

The arithmetic/logic unit (ALU) handles the execution of all arithmetic computations. It does not store data; it merely performs the necessary calculations. Functions performed by the ALU include arithmetic operations (addition, subtraction, multiplication, and division) and comparisons. Since the bulk of the computer processing involves calculations or comparisons, the capabilities of a computer often depend on the capabilities of the ALU. An ALU must process numbers using the same format as the rest of the digital circuit. The format of modern processors is usually the binary number representation of two's complement. Early computers used a wide variety of number systems, including one's complement, sign-magnitude format, and even true decimal systems, with ten tubes per digit.

ALUs for each one of these numeric systems had different designs, and that influenced the current preference for two's complement, as this is the representation that makes it easier for the ALUs to calculate additions and subtractions. The two's-complement number system allows for subtraction to be accomplished by adding the negative of a number in a very simple way, which negates the need for specialized circuits to do subtraction. The inputs to the ALU are the data to be operated on (called operands) and a code from the control unit indicating which operation to perform. Its output is the result of the computation. In many designs, the ALU also takes or generates as inputs or outputs a set of condition codes from or to a status register. These codes are used to indicate cases such as carry-in or carry-out, overflow, divide-by-zero, etc.

How the CPU works

The CPU is centrally located on the motherboard. Since the CPU carries out a large share of the work in the computer, data pass continually through it. The data come from the RAM and the units (keyboard, drives, etc.). After processing, the data is sent back to the RAM and the units. The CPU continually receives instructions to be executed. Each instruction is a data processing order. The work itself consists mostly of calculations and data transport.

OUTPUT UNIT: As program statements and data are received by the CPU from an input device, the results of the processed data are sent from the CPU to an output device. These results are transferred from the output storage area onto an output medium, such as a floppy disk, hard drive, monitor, printer, etc. The output devices help the computer to display information.

THE INSTRUCTION-EXECUTION CYCLE

Many types of personal computers can execute instructions in less than one-millionth of a second; supercomputers can execute instructions in less than one-billionth of a second.

The CPU performs four steps in executing an instruction:

1. The control unit gets the instruction from memory.
2. The control unit decides what the instruction means and directs the necessary data to be moved from the memory to the arithmetic logic unit.
3. The arithmetic logic unit performs the actual operation on the data.
4. The result of the operation is stored in memory or a register.

The first two instructions make up what is called the instruction time. The last two instructions make up what is called the execution time. The combination of these two is called a machine cycle. Each central processing unit has an internal clock or system clock, which produces pulses at a fixed rate to synchronize all computer operations. A single machine cycle instruction is made up of a number of sub-instructions, each of which must take at least one clock cycle. Each type of CPU is designed to understand a specific group of instruction called the instruction set.

How the CPU finds Instructions and Data

An address or a number that stands for a location in the computer memory identifies the location in memory for each instruction and each piece of data. An address may be compared to a mailbox except that the address can hold only one item, that is, a fixed amount of data, a number or a word, at any one time.

The following is an example of a simple case of adding two numbers together and placing the result in a location X. For the command execution such as: LET X = N1 + N2, is as shown in figure 2.6.

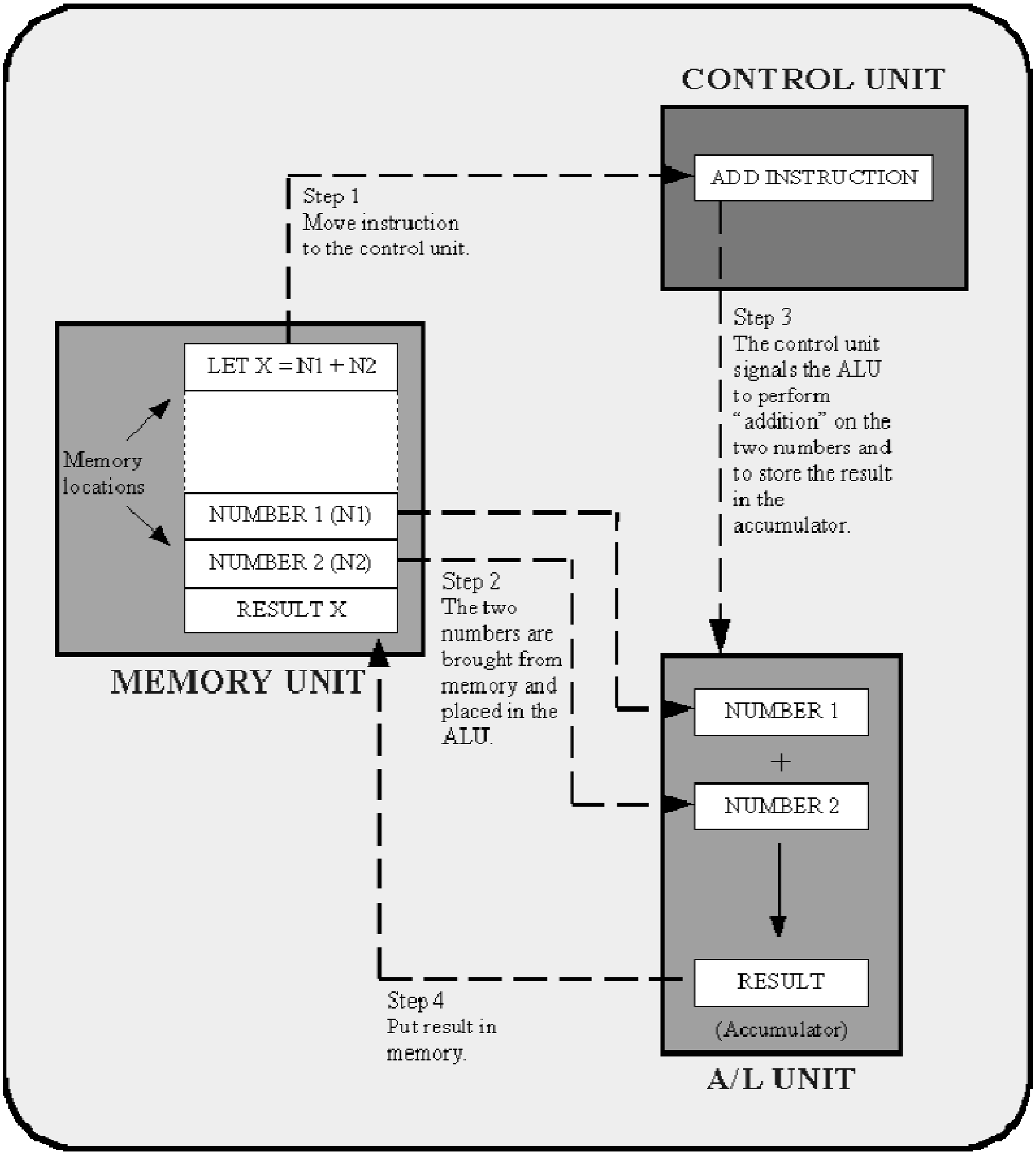


Figure 2.6: CPU process of execution

CPU OPERATION

The fundamental operation of most CPUs, regardless of the physical form they take, is to execute a sequence of stored instructions called a program. The program is represented by a series of numbers that are kept in computer memory. There are four steps the CPUs use in their operation: fetch, decode, execute, and writeback. The first step, fetch, involves retrieving an instruction (which is represented by a number or sequence of numbers) from program memory. The location in program memory is determined by a program counter (PC), which stores a number that identifies the current position in the program. In other words, the program counter keeps track of the CPU's place in the current program. After an instruction is fetched, the PC is incremented by the length of the instruction word in terms of memory units. Sometimes the instruction to be fetched may be retrieved from relatively slow memory, causing the CPU to stall while waiting for the instruction to be returned. This issue is largely addressed in modern processors by caches and pipeline architectures. The instruction that the CPU fetches from memory is used to determine what the CPU is to do.

In the decode step, the instruction is broken up into parts that have significance to other portions of the CPU. The way in which the numerical instruction value is interpreted is defined by the CPU's instruction set architecture (ISA). Most times, one group of numbers in the instruction, called the opcode, indicates which operation to perform. The remaining parts of the number usually provide information required for that instruction, such as operands for an addition operation. Such operands may be given as a constant value (called an immediate value), or as a place to locate a value; a register or a memory address, as determined by some addressing mode. In older designs, the portions of the CPU responsible for instruction decoding were unchangeable hardware devices but in more abstract and complicated CPUs and ISAs, a microprogram is often used to assist in translating instructions into various configuration signals for the CPU. This microprogram is sometimes rewritable so that it can be modified to change the way the CPU decodes instructions even after it has been manufactured.

After the fetch and decode steps, the execute step is performed. During this step, various portions of the CPU are connected so they can perform the desired operation. In this case of an addition operation, an arithmetic logic unit (ALU) will be connected to a set of inputs and a set of outputs. The inputs provide the numbers to be added, and the outputs will contain the final sum. The ALU contains the circuitry to perform simple arithmetic and logical operations on the inputs such as addition and bitwise operations. If the addition operation produces a result too large for the CPU to handle, an arithmetic overflow flag in a flags register may also be set. The final step, writeback, simply "writes back" the results of the execute step to memory. Sometimes the results are written to some internal CPU register for quick access by subsequent instructions and also results may be written to slower and larger main memory.

Some types of instructions manipulate the program counter rather than directly produce result data. These are generally called "jumps" and facilitate behaviour like loops, conditional program execution (through the use of a conditional jump), and functions in programs. Many instructions will also change the state of digits in a "flags" register. These flags can be used to influence how a program behaves, since they often indicate the outcome of various operations.

After the execution of the instruction and write-back of the resulting data, the entire process repeats, with the next instruction cycle normally fetching the next-in-sequence instruction because of the incremented value in the program counter. If the completed instruction was a jump, the program counter will be modified to contain the address of the instruction that was jumped to, and program execution continues normally.

MOTHERBOARD

The motherboard is the main printed circuit board in the computer, which the CPU plugs into; as do all of the "cards" (sound card, scanner cards, etc.), the bus, memory sockets, keyboard controller and supporting chips. This is the central nervous system of the computer, without it nothing runs. The motherboard is shown in figure 2.6.

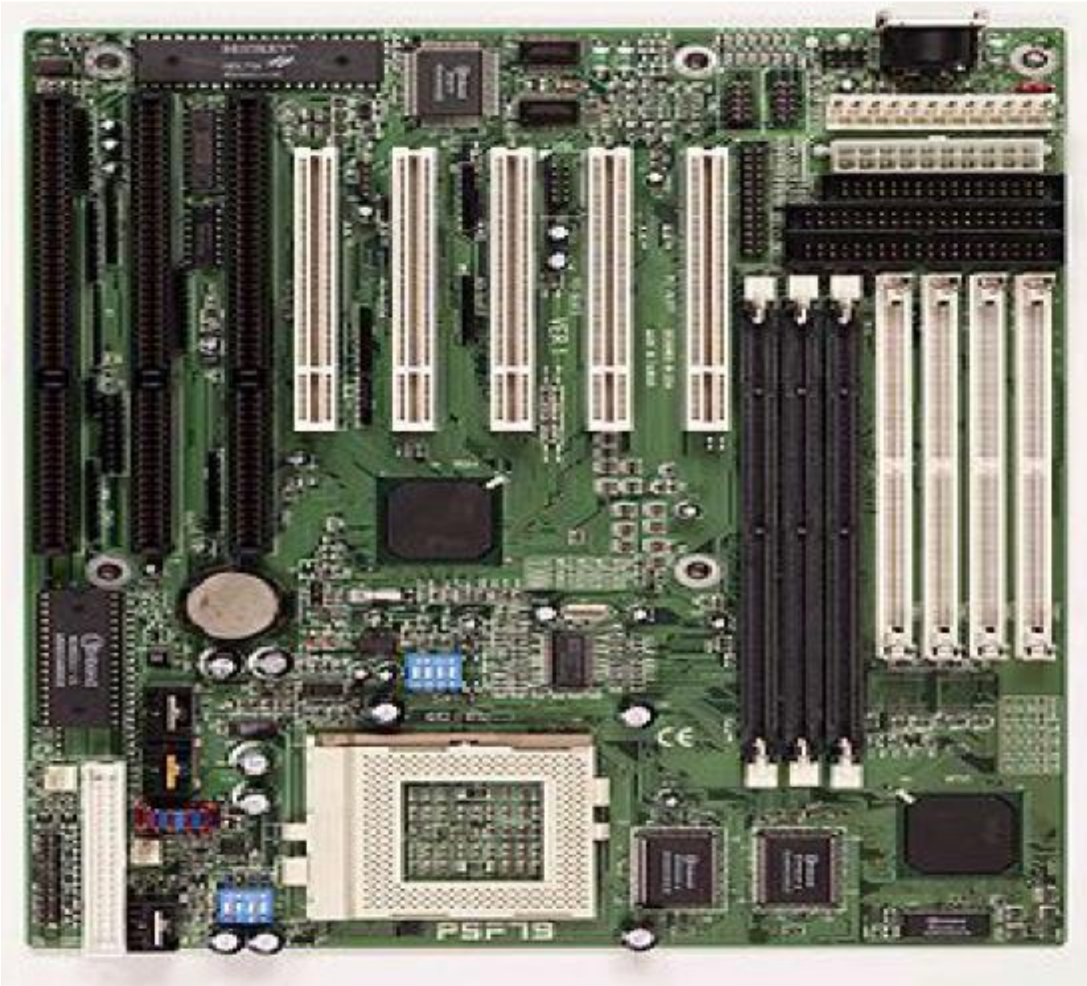


Figure 2.6: Motherboard

The microprocessor, (or CPU), is the brain of the computer. Figure 2.7 shows a slot 1 processor with heat sinks and a fan, which prevent it from overheating.

Figure 2.7: Processor with heat sink and fan

Figure 2.8 is the processor without the heat sinks and fan, being inserted into a slot 1 motherboard connection. Slot 1 processors have the microprocessor and level 2 cache memory mounted on a circuit board, (or card), which is enclosed inside a protective shell.

Figure 2.8: Processor without heat sinks and fan

The enclosed slot 1 processor card contains the central processing unit, (or CPU), with its level 1 cache memory. The central processing unit also contains the control unit and the arithmetic/logic unit, both working together as a team to process the computer's commands.
The control unit controls the flow of events inside the processor. It fetches instructions from memory and decodes them into commands that the computer can understand. The arithmetic/logic unit handles all of the math calculations and logical comparisons. It takes the commands from the control unit and executes them, storing the results back into memory.
These four (fetch, decode, execute, and store), are called the "machine cycle" of a computer.
These 4 basic steps are how the computer runs each and every program. The microprocessor's level 1 cache memory is memory that is contained within the CPU itself. It stores the most frequently used instructions and data. The CPU can access the cache memory much faster than having to access the RAM, (or Random Access Memory). Figure 2.9 is a picture of what is inside a Pentium 3 processor. The control unit, arithmetic/logic unit, and level 1 cache are contained within the center CPU chip. Level 2 cache memory is visible on the right-hand side of the processor card.

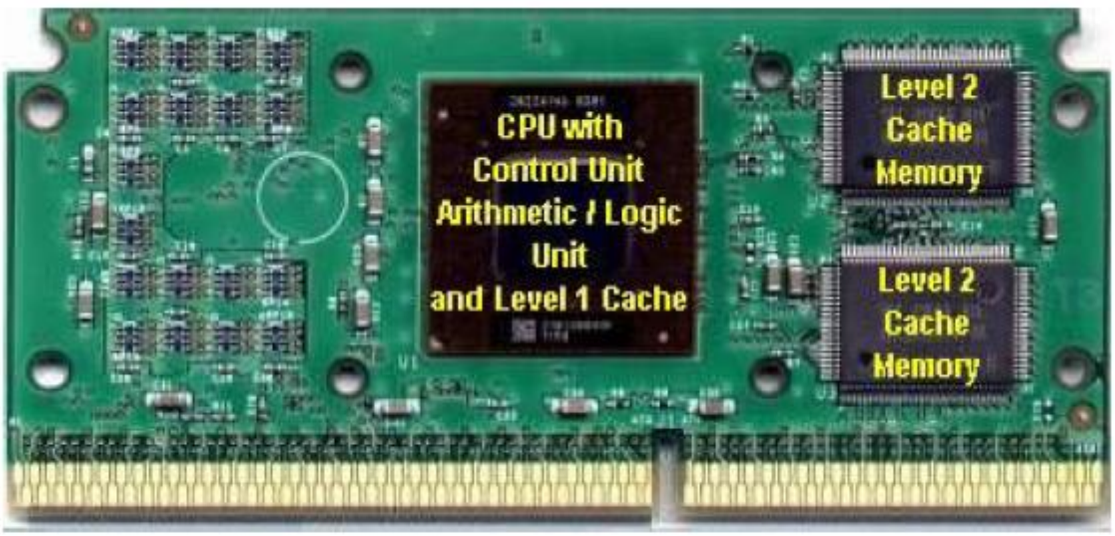


Figure 2.9: Memory card displaying the cache memory

Level 1 cache memory is memory that is included inside of the CPU itself. It is usually smaller and faster than level 2 cache memory. Level 2 cache memory is memory between the RAM and CPU. It is used when the level 1 cache memory is full or is too small to hold the intended data. Originally it was not directly on the CPU chip itself. Figure 2.9 shows level 2 cache memory on the processor card, beside the CPU.
Figure 2.10 is two photos of a CPU displaying the view of the CPU chip from the outside and the large map of the inside of the CPU showing the different areas and what their function is.

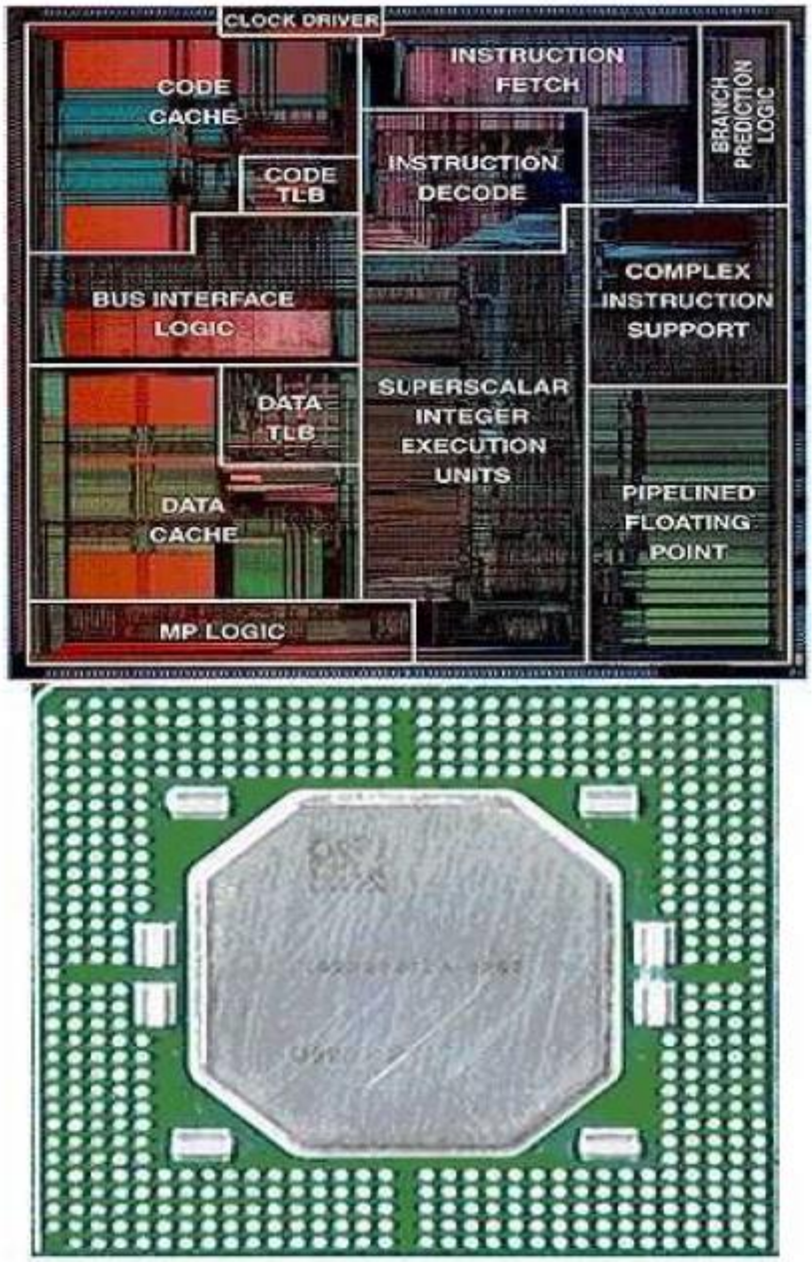


Figure 2.10: View of the CPU chip from the outside and inside

The clock driver is what times, or sets the pace, for the computer. The clock's speed is how CPUs are rated. A device called the system clock governs the speed with which the CPU cycles from state to state. Clock speed is often measured in gigahertz (GHz) where a gigahertz is one billion cycles per second. Thus, a 2.9 GHz processor could execute 2.9 billion cycles in one second. Each machine cycle consists of two beats. Each beat the control unit fetches and decodes data, which is called the "instruction cycle." At the same time the arithmetic/logic unit executes and stores data, which is called the "execution cycle." The speed of a clock is rated by how many beats per second it can accomplish. 1 billion beats per second is referred to as 1 GHz. For every

beat, (except the very first), a machine cycle is completed. A 3 GHz CPU can execute 3,000,000,000 instructions in a single second.
Different systems have different details, but in general all computers consist of components (processor, memory, controllers, video) connected together with a *bus*. Physically, a bus consists of many parallel wires, usually printed (in copper) on the main circuit board of the computer. Data signals, clock signals, and control signals are sent on the bus back and forth between components. A particular type of bus follows a carefully written standard that describes the signals that are carried on the wires and what the signals mean. For example, the PCI standard describes the PCI bus used on most current PCs. The processor continuously executes the machine cycle, executing machine instructions one by one. Most instructions are for an arithmetical, a logical, or a control operation. A machine operation often involves access to main storage or involves an i/o controller. If so, the machine operation puts data and control signals on the bus, and may wait for data and control signals to return. Some machine operations take place entirely inside the processor. These operations are very fast.
Information is moved between the CPU and memory across a connection called a bus. The amount of information that can be moved to or from memory at one time is referred to as the bus width.

Internal storage unit: - The internal storage unit is also known as primary storage or main memory. Primary storage holds all instructions and data needed for processing. Any final or intermediate results from calculations are also stored there. Primary storage is actually adjacent to the central processing unit. Instructions and data are rapidly transferred between the two units as needed. When all the data have been processed, the results stay in memory until the control unit causes them to be erased or the power is turned off. The storage unit serves four purposes; three of which are associated to retention (holding) of data during processing. First, data is transferred from an input device to the input storage area where it remains until the computer is ready to process it as indicated by the solid arrow figure 5.1. Secondly, a working storage area within the storage unit holds both the data being processed and the intermediate results of the arithmetic/logic operations. Thirdly, the storage unit retains the processing results in the output storage area, from where the processing results can be transferred to an output device. Fourthly, the program storage area contains the program statements transferred from an input device to process the data. The four areas, that is, input, working storage, output, and program storage are not fixed in size or location but are determined by individual program requirements.

MEMORY

Memory is another critical component in computers. The memory unit is the part of the computer that holds data and instructions for processing. Although it is closely associated with the CPU, it is actually separate from it. Memory associated with the CPU is also called primary

storage/primary memory/main storage/internal storage or main memory. When software is loaded from a floppy disk, hard disk or CD-ROM, it is stored in the main memory.
The two most important types of memory inside the computer are read-only memory (ROM) and random access memory (RAM). Computers can read data stored in ROM, but cannot write new data to it. With RAM, computers can read from and write to the memory. Without computer memory, every calculation on a computer would be stateless so there will be no way to preserve information from one moment to the next and every process would start on a clean slate. Many desktop PCs have the capacity for additional RAM. The user simply has to open the computer and plug RAM chips into the appropriate sockets on the motherboard.
Memory is very fast storage used to hold data. It has to be fast because it connects directly to the microprocessor. There are several specific types of memory in a computer:

Random-access memory (RAM): - This is really the main store and is the place where the programs and software loaded into the system are stored. When the CPU runs a program, it fetches the program instructions from the RAM and carries them out. If the CPU needs to store the results of calculations it can store them in RAM. RAM can have instructions READ from it by the CPU and also it can have numbers or other computer data WRITTEN to it by the CPU. The more RAM in the computer, the larger the programs it can run. When the computer is switched off, whatever is stored in the RAM gets erased.
RAM is used to temporarily store information with which the computer is currently working. When a program is run, it is loaded from the hard drive into RAM. Constant saving of work on the hard drive or diskette is essential, because whatever is in RAM is erased when the computer is turned off. RAM chips, the physical components that contain the memory are grouped in rows called SIMMs (or single in-line memory modules). These modules are small bars, usually containing eight, or nine memory chips. To add more memory to the computer, one or more SIMMs can be plugged. A memory chip is the integrated circuit that actually contains the RAM. The DIMMs and SIMMs are shown in figure 2.11.

DIMMs SIMMs

Figure 2.11: DIMMs and SIMMs

SIMM is an acronym for Single In-line Memory Module. It is a printed circuit board that holds several semiconductor memory chips and is used to add memory to a computer. It is plugged into

a SIMM socket on the motherboard and had 32-bit path to the memory chips. DIMM is an acronym for Dual In-line Memory Module. DIMM has 64-bit path. DIMM comprises of a series of dynamic random access memory integrated circuits. These modules are mounted on a printed circuit board and designed for use in personal computers, workstations and servers. DIMMs began to replace SIMMs as the predominant type of memory module as Intel P5-based Pentium processors began to gain market share. The main difference between SIMMs and DIMMs is that DIMMs have separate electrical contacts on each side of the module, while the contacts on SIMMs on both sides are redundant. Another difference is that standard SIMMs has a 32-bit data path, while standard DIMMs have a 64-bit data path. Since most processors have a 64-bit bus width, it requires SIMMs installed in matched pairs in order to complete the data bus. The processor would then access the two SIMMs simultaneously. DIMMs were introduced to eliminate this practice. Because the Pentium processor requires a 64-bit path to memory, SIMMs can be installed two at a time, but with DIMMs, one DIMM can be installed at a time.
The four types of RAM chips used in personal computers are dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), static RAM (SRAM), and Rambus Dynamic RAM (RDRAM). The CPU must constantly refresh DRAM or it will lose its contents. The type of DRAM used in most PCs is SDRAM, which is synchronized by the system clock and is much faster than DRAM. SRAM is faster than any DRAM and will retain its contents without having to be refreshed by the CPU. RDRAM is faster and more expensive than SDRAM.

Read-only memory (ROM): - Is a permanent type of memory storage used by the computer for important data that does not change. ROM contains the commands the computer needs to activate. The CPU can only fetch or read instructions from ROM.
ROM come with instructions permanently stored inside and these instructions cannot be over-written. ROM memory is used for storing special sets of instructions, which the computer needs when it starts up. When the computer is switched off, the instructions/contents of the ROM does not become erased but remains stored permanently as it is non-volatile.

Other types of non-volatile solid-state memory permit some degree of modification:
(a) *Programmable read-only memory (PROM)*, or one-time programmable ROM (OTP), can be written to or programmed via a special device called a PROM programmer. It uses high voltages to permanently destroy or create internal links within the chip and can only be programmed once.
(b) *Erasable programmable read-only memory (EPROM)* can be erased by exposure to strong ultraviolet light (for 10 minutes or longer), then rewritten with a process that again requires application of higher than usual voltage. Repeated exposure to UV light will eventually wear out an EPROM, but the endurance of most EPROM chips exceeds 1000 cycles of erasing and reprogramming. EPROM chip packages can often be identified by the prominent quartz "window" which allows UV light to enter. After programming, the window is typically covered with a label to prevent accidental erasure. Some EPROM chips are factory-erased before they are packaged, and include no window.

(c) *Electrically erasable programmable read-only memory* (EEPROM) is based on a similar semiconductor structure to EPROM, but allows its entire or selected contents to be electrically erased, then rewritten electrically, so that they need not be removed from the computer or camera or MP3 player. Writing or flashing an EEPROM is much slower (milliseconds per bit) than reading from a ROM or writing to a RAM (nanoseconds in both cases).
(d) *Electrically alterable read-only memory (EAROM)* is a type of EEPROM that can be modified one bit at a time. Writing is a very slow process and again requires higher voltage (around 12 V) than is used for read access. EAROMs are intended for applications that require infrequent and only partial rewriting. EAROM may be used as non-volatile storage for critical system setup information. In many applications, EAROM has been supplanted by CMOS RAM supplied by mains power and backed-up with a lithium battery.
(e) *Flash memory* (or simply flash) is a modern type of EEPROM invented in 1984. Flash memory can be erased and rewritten faster than ordinary EEPROM. Modern NAND flash makes efficient use of silicon chip area, resulting in individual ICs with a capacity as high as 32 GB as of 2007; this feature, along with its endurance and physical durability, has allowed NAND flash to replace magnetic in some applications such as USB flash drives. Flash memory is sometimes called flash ROM or flash EEPROM when used as a replacement for older ROM types, but not in applications that take advantage of its ability to be modified quickly and frequently. By applying write protection, some types of reprogrammable ROMs may temporarily become read-only memory.

Basic input/output system (BIOS): - Is a type of ROM that is used by the computer to establish basic communication when the computer is first powered on. The BIOS works closely with the CPU. BIOS is a specific kind of ROM. If the CPU is seen as the brain of the computer, then BIOS can be considered to be the spine. It is the job of BIOS to handle interactions between the software running on a computer and the machine's hardware components.

Caching: - Is the storing of frequently used data in extremely fast RAM that connects directly to the CPU.

Virtual memory: - Is space on a hard disk used to temporarily store data and swap it in and out of RAM as needed.

CMOS chip:- Complementary metal-oxide semiconductor chips are powered by a battery and thus do not lose their contents when the power is turned off. The chips contain flexible start-up instructions, such as time, date, and calendar, which must be kept current even when the computer is turned off. It can be reprogrammed unlike the ROM chips.

STORED-PROGRAM CONCEPT

A program can be defined as a series of instructions that direct the computer to perform a given task. In early computers, instructions had to be either wired on control panels and plugged into the computer at the beginning of a job or read into the computer from punched cards in distinct steps as the job progressed. This approach slowed down processing because the computer had to wait for instructions by a human operator. To speed up processing, the memory of the computer began to be used to store the instructions as well as the data. This development, the stored-program concept, was significant; since instructions were stored in computer memory in electronic form, no human intervention was required during processing.

The computer could proceed at its own speed-close to the speed of light. Modern computers can store programmes. Once instructions required for an application have been determined, they remain in memory until new ones are stored over them. This same process holds true for data as well. Data in the computer is held until new data is placed over it. Therefore, the same instructions or data can be used over and over again until they are changed. The process of accessing the same instructions or data over and over is called reading. Storing new instructions or data in computer memory is called writing. The basic characteristic, therefore, is known as non-destructive read/destructive write. When instructions or data are read, they do not replace or destroy previous instructions or data.

When new instructions or data are written into the computer memory, whatever was formerly there in the memory is destroyed. A series of instructions placed into memory is called a stored program.

MEMORY LOCATIONS

In order to direct processing operations, the control unit of the CPU must be able to locate each instruction and data in memory so each location in memory is assigned as address. Since each location in memory has a unique address, items can be located by use of stored-program instructions that provide their addresses. Sometimes data at a location must be changed, added to, or deleted during execution of the program. The programmer assigns a variable, or symbolic name, to the piece of data that is stored at that memory location. Regardless of the type of storage used, there will be a system of unique addresses for each type used.

TYPES OF PRIMARY STORAGE

MAGNETIC-CORE STORAGE: In computers manufactured during the 1960s and 1970s, the most common type of primary storage was magnetic-core storage. Magnetic cores are iron-alloy of doughnut-shaped rings. Millions of these cores are strung on sections of thin wire mesh to make up the storage unit. The cores, located at the intersections of two wires, represent data. The wires on the mesh were magnetized in various combinations to store data.

SEMICONDUCTOR MEMORY: In most computers manufactured since the mid-1970s, the primary storage unit relies on semiconductor memory components. Semiconductor memory consists of tiny transistor circuits on silicon chips. Several silicon chips are cut from a single wafer, a thin disk of silicon material. The tiny circuitry etched on the silicon chip contains storage cells that are electrically charged. Each cell is either on or off, representing the digits 1 or 0, respectively. The control unit of the CPU "reads" the memory by testing the electrical contents of each cell.
Semiconductor storage units have replaced magnetic core units for several reasons such as size, memory or storage capacity, etc. Semiconductors are less bulky than magnetic cores as one silicon chip, slightly bigger than one core, can hold as much data as thousands of core. Semiconductor units operate much faster than core units; and semiconductor storage costs less per bit than core storage. Semiconductor storage still has one serious drawback-since information is stored by using electrical charges, a constant power source is needed to ensure that no data are lost. If the power source to the semiconductor unit is interrupted, even for a fraction of a second, all data stored in the unit will be lost. Storage of this nature is said to be volatile but core storage operates with no magnetism and is non-volatile as it maintains its contents in case of power failure.

BUBBLE MEMORY: Another advance in memory technology was the introduction of bubble memory. This technology was used for both primary and secondary storage, as it incorporates the best features of both semiconductor and magnets core memory. Bubble memory is magnetic and therefore non-volatile. It is much faster and more compact than magnetic core memory. Bubble memory uses tiny charged spots or bubbles on a thin strip of magnetized film. Data are stored on the film by using the presence or absence of bubbles in specific locations. If a location contains a bubble, it represents the digit 1. If no bubbles are present, the location represents the digit 0. Difficulty of production, user acceptance of bubbles and high cost limited its wide spread and users in the industry.

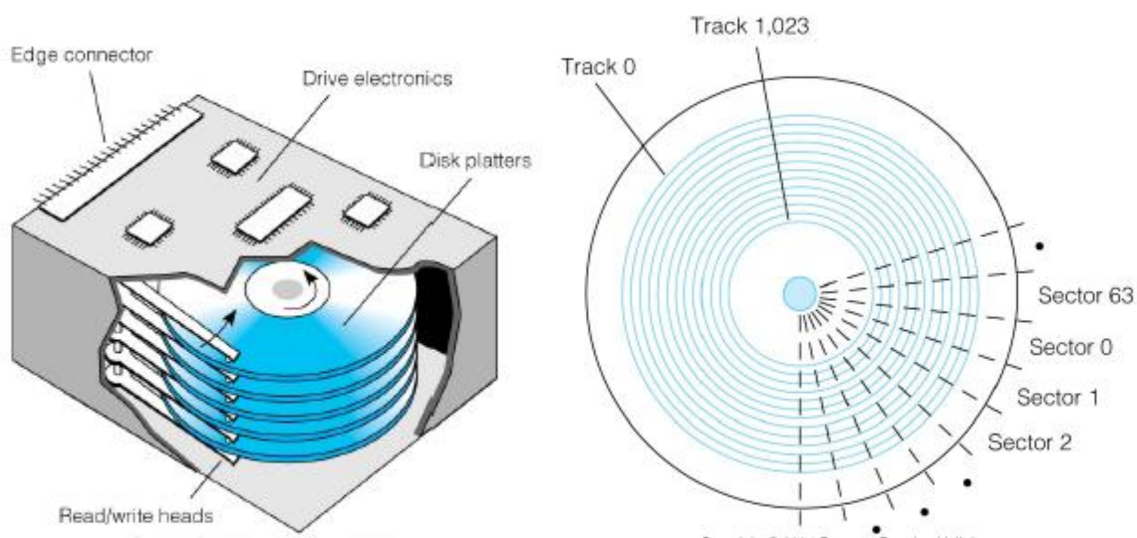
SECONDARY STORAGE

Sometimes, the amount of data required by a program exceeds the capacity of primary storage. Such data are stored in secondary storage (also called auxiliary storage or external storage) outside the main computer. Secondary storage hardware is device that permanently holds data and information as well as programs. The common types of secondary storage devices are magnetic tapes and magnetic disks, floppy disks, optical disks, smart cards, flash memory cards, and hard disks. These secondary storage media cost much less than primary storage so they allow larger volumes of data to be stored more economically. The secondary storage media are connected to the CPU. Once data have been placed in them, they can be accessed as required for processing. The retrieval of items from secondary storage is slower than from primary storage. After processing has been completed, the data or results can be written back onto the auxiliary medium.

MAGNETIC TAPE: A magnetic tape is a continuous plastic strip wound on a reel. It is a thin plastic tape coated with a substance that can be magnetized (figure 2.12). Data is represented by magnetized spots (representing 1s) or non-magnetized spots (representing 0s). The magnetic tape plastic base is treated with a magnetizable coating. Data are stored on a magnetic tape by magnetizing small spots on its surface. The computer can read these spots; though they are invisible to the human eye. Large volumes of data can be stored on a single tape since the data are packed so compactly. A magnetic tape is mounted on a tape drive when a program needs the information it contains. The drive has a read/write head that is an electromagnet. It detects the magnetic spots as the data are read. Pulses are sent from the head to the computer, which interprets the pulses as data. To write data on the tape, the head magnetizes the appropriate spots and erases any previously stored data.

MAGNETIC DISK: A magnetic disk is a metal platter as shown in figure 2.12. The disk has concentric circles called tracks, which are enclosed within each other without touching. They are not visible but they are paths that the recorded data follow. The disk is coated on both sides with a magnetizable material, and magnetizing spots on the disk's surface store data. A center shaft connects the disks so that they can be rotated at the same time. A disk drive is used to rotate the magnetic disks. The data are read or written by a read/write head. These heads are connected to access arms, and all move together. By means of the access arms, the read/write heads move back and forth across the disk until the proper track is located. The disk is then rotated until the desired data are located on the track. In this way, magnetic disks allow the user to access the desired data directly. Thus, a magnetic disk is a form of direct access, or random storage. Accessing data on a disk does not destroy the stored data; it remains permanent.

Hard Disks



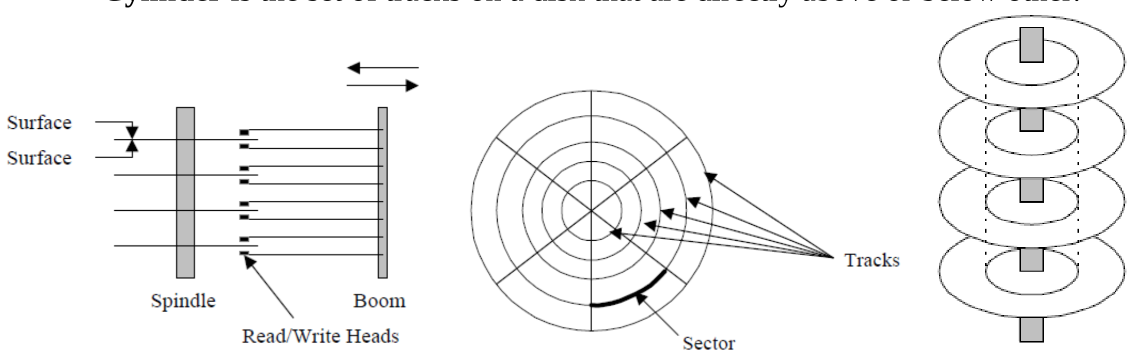
Computation of a hard disks capacity is identical to that for other disks but the numbers are larger. Breakdown of capacity for a 50 GB hard disk, assuming 11 platters, 264,528 tracks, 369 sectors per track:

12024 cylinders	x	22 heads (sides)	=	264528 tracks
264528 tracks	x	369 avg. sectors/track	=	97,610,823 sectors
97610823 sectors	x	512 bytes	=	49,976,754,984 bytes (Approx.)

If the actual numbers of sectors per track were taken into account instead of an average, the calculation would be more difficult and the resulting number of bytes would be higher.

A recap:

- A disk platter contains concentric **tracks**,
- Tracks are divided into **sectors**
- A sector is the smallest addressable unit in the disk.
- **Cylinder** is the set of tracks on a disk that are directly above or below other.



A Cylinder

FLOPPY DISKS: The flexible, or floppy, diskette was introduced in 1973. It is a removable flat piece of mylar plastic packaged in a 3.5-inch plastic case as shown in figure 5.12. These flexible diskettes are made of plastic instead of metal, which is why they have the nickname “floppy”. Data and programs are stored on the disk’s coating by means of magnetized spots, following standard on/off patterns of data representation. The plastic case protects the mylar disk from being touched by human hands. Each disk is enclosed in a protective jacket. They are miniature magnetic disks. Floppy diskettes are reusable and easy to store. They provide the sum random-access capabilities as magnetic disks and are very popular for use with microcomputers. A floppy diskette that stores data on both sides can store approximately 360k bytes of data.
Floppy disks are inserted into a floppy disk drive, a device that holds, spins, reads data from, and writes data to floppy disks. Read means that the data secondary storage is converted to electronic signals and a copy of that data are transmitted to the computer‘s memory. Write means that a copy of the electronic information processed by the computer is transferred to secondary storage.

Floppy disks have a write-protect notch, which allows the user to prevent a diskette from being written to.
There are other disks like the zip disks (special disks with a capacity of 100 or 250 megabytes); SuperDisks (disks with a capacity of 120 megabytes); and HiFD (capacity of 200 megabytes). Both SuperDisks and HiFD can read standard 1.44-megabyte floppies.

CARTRIDGE TAPES: Cartridge tapes are similar to cassette tapes, but they can store data more compactly. They require 90 percent less storage space than common magnetic tapes. They are used mainly for video games. Cartridges can contain random -access memory (RAM) and read-only memory (ROM). Game cartridges use ROM to protect the games from damage.

LASER DISKS: Laser or optical disks need concentrated beams of light to store data. The beam makes marks on the thin layer of metal or polymer (a synthetic substance) that coats the disks and the computer reads these marks. Laser disks hold much more data than magnetic disks hold. They also resist the fingerprints and dust that can harm magnetic disks. The disks are read in an optical disk drive by laser beams. When pictures are stored on laser disks, they are called videodisks. Images stored on video disks are high quality.

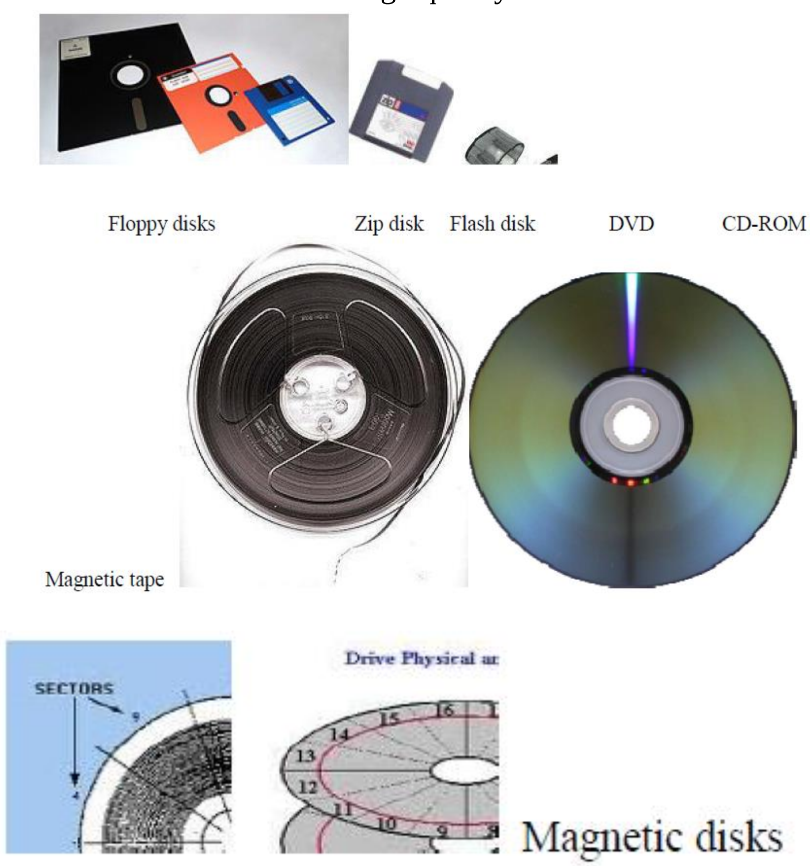


Figure 2.12: Storage devices

INPUT AND OUTPUT DEVICES

A computer system is much more than a central processing unit with different kinds of storage. Auxiliary/peripheral devices enter data into the system and receive output from the CPU. Input is data submitted to the computer for processing. Output is the information produced by the computer as a result of computer processing. The input and output operations are often referred to as I/O. I/O operations are important activities in any computer-based system because they are the communication links between people and the machine. If these interfaces are weak, the overall performance of the computer system suffers. To communicate with the computer, the user must use an input or output device.

These devices are located outside the main part of the computer; this made them to be called peripherals.

INPUT DEVICES

The past years have seen many advances in techniques used to enter data into computer systems. Compared to the early days of computing, when punched cards were used to enter programs and data, today's data-entry techniques have become quite sophisticated. One of the most significant advances in automated data entry has been the use of scanners. Even though many significant advances in the area of automated data entry have occurred, the majority of data entered into computer systems is still done manually by human data-entry personnel.

This is usually accomplished through the use of a computer keyboard. The data typed in is then stored on one of the computer system's storage devices. This process of manually entering data into a computer system is referred to as key-to-magnetic media, or keyboarding. Input hardware devices are categorized into three types- keyboards, pointing devices and source data entry devices.

Keyboards are traditional computer keyboards (104-105 keys for desktop computers and 85 keys for laptops); and specialty keyboards (ranging from touch-tone telephone keypads to keyboards featuring pictures) and terminals (dumb terminals, intelligent terminals, and Internet terminals). A dumb terminal, also called a video display terminal (VDT), has a display screen and a keyboard and can input and output but not process data. An intelligent terminal has its own memory and processor, as well as a display screen and keyboard. An Internet terminal provides access to the Internet.

Pointing devices control the position of the cursor or pointer on the screen. It includes the mouse and its variants (trackball, pointing stick and touch pads); the touch screen; and various forms of pen input (pen-based systems, light pens and digitizers). The trackball is a movable ball, mounted on top of a stationary device that can be rotated using the fingers or palm. A pointing stick looks like a pencil eraser protruding from the keyboard between the G, H, and B keys. The forerunner of the pointing stick is the joystick, which consists of a vertical handle like a gearshift lever mounted on a base with one or two buttons. Pointing sticks are used on laptop computers. A touchpad is a small, flat surface over which is slide using the fingers as in the way with the mouse. A touch screen is a video display screen that has been sensitized to receive input from the touch of the finger. Pen-based computer systems allow users to enter handwriting and marks onto a computer screen by means of a pen like stylus rather than by typing on a keyboard. The light pen is a light-sensitive pen like device connected by a wire to the computer terminal. A digitizer uses a mouse like copying device called puck, or an electronic pen, which can convert drawings and photos to digital data.

Source data-entry devices create machine-readable data on magnetic media or paper or feed directly into the computer's processor. The devices include scanning devices (such as imaging systems, bar-code readers, fax machines, mark-and character-recognition devices like MICR, OMR, OCR); audio-input devices; web cams and video-input cards; digital cameras; voice-recognition systems; sensors; radio-frequency identification devices; and human biology-input devices. Scanners use light-sensing equipment to translate images of text, drawings, photos, etc into digital form.

Imaging system or scanner or graphics scanner converts text, drawing, and photographs into digital form that can be stored in a computer system and then manipulate output through modem to another computer. Bar-code readers are photoelectric/optical scanners that translate the symbols in the bar code (vertical zebra striped marks seen the most manufactured retail products) into digital code.

Magnetic-ink character recognition (MICR) reads the strange-looking numbers printed at the bottom of checks. Optical-mark recognition (OCR) uses a device that reads pencil marks and converts them into computer-usable form. Optical character recognition (OCR) uses a device that reads pre-printed characters in a particular font and converts them to digital code. Fax machine or facsimile transmission machine, scans an image and sends it as electronic signal over telephone lines to a receiving fax machine, which prints out the image on paper.

KEYBOARDS: Key-to-magnetic media data entry comprises the most widely used method of input. Data is entered through the keyboard, which is similar to a typewriter. The keyboard has all the keys on typewriter keyboards, plus other keys unique to computers. The keyboard is shown in figure 2.13.

Figure 2.13: Keyboard

Keyboard Technology

The technology that makes up a typical PC keyboard is very interesting. This section focuses on all aspects of keyboard technology and design, including the key switches, the interface between the keyboard and the system, scan codes, and the keyboard connectors.

Key Switch Design

Several types of switches are used in keyboards today. Most keyboards use one of several variations on a mechanical key switch. A *mechanical key switch* relies on a mechanical momentary contact type switch to make electrical contact in a circuit. Some high-end keyboard designs use a totally different non-mechanical design that relies on capacitive switches. This section discusses these switches and the highlights of each design. The most common type of key switch is the mechanical type, available in the following variations:

- Pure mechanical
- Foam element
- Rubber dome
- Membrane

The *pure mechanical* type is just that--a simple mechanical switch that features metal contacts in a momentary contact arrangement. Often a *tactile feedback mechanism*-- consisting of a clip and spring arrangement to give a "clicky" feel to the keyboard and offer some resistance to pressing the key--is built in. Several companies, including Alps Electric, Lite-On, and NMB Technologies, manufacture this type of keyboard using switches primarily from Alps Electric. Mechanical switches are very durable, usually have self-cleaning contacts, and normally are rated for 20 million keystrokes, which is second only to the capacitive switch.

Microprocessor and controller circuitry of keyboard

They also offer excellent tactile feedback. Foam element mechanical switches were a very popular design in some older keyboards. Most of the older compatible keyboards, including those made by Keytronics and many others, use this technology. These switches are characterized by a foam element with an electrical contact on the bottom that is mounted on the bottom of a plunger attached to the key itself.

When the switch is pressed, a foil conductor on the bottom of the foam element closes a circuit on the printed circuit board below. A return spring pushes the key back up when the pressure is released. The foam dampens the contact, helping to prevent bounce, but unfortunately gives these keyboards a "mushy" feel. The big problem with this type of key switch design is that there is often little in the way of tactile feedback, and systems with these keyboards often resort to tricks such as clicking the PC's speaker to signify that contact has been made. Compaq has used keyboards of this type (made by Key-tronics) in many of their systems, but perhaps the most popular user today is Packard Bell.

Another problem with this type of design is that it is prone to corrosion on the foil conductor and the circuit board traces below. When this happens, the key strikes may become intermittent, which can be frustrating. Fortunately, these keyboards are among the easiest to clean. By disassembling this type of keyboard completely, you can usually remove the circuit board portion without removing each foam pad separately, and expose the bottoms of all the pads. Then you can easily wipe the corrosion and dirt off the bottom of the foam pads and the circuit board, thus restoring the keyboard to a "like-new" condition. Unfortunately, over time the corrosion problem will occur again. Stabilant 22a from D.W. can be used to solve this problem. Because of problems like this, the foam element design is not used much anymore and has been superseded in popularity by the rubber dome design.

Rubber dome switches are mechanical switches that are similar to the foam element-type but are improved

in many ways. Instead of a spring, these switches use a rubber dome that has a carbon button contact on the underside. As you press a key, the key plunger presses on the rubber dome, causing it to resist and then collapse all at once, much like the top of an oil can. As the rubber dome collapses, the user feels the tactile feedback, and the carbon button makes contact between the circuit board traces below. When the key is released, the rubber dome re-forms and pushes the key back up.

The rubber eliminates the need for a spring and provides a reasonable amount of tactile feedback without any special clips or other parts. A carbon button is used because it is resistant to corrosion and also has a self-cleaning action on the metal contacts below. The rubber domes are formed into a sheet that completely protects the contacts below from dirt, dust, and even minor spills. This type of design is the simplest, using the fewest parts. These things make this type of key switch very reliable and help make rubber dome-type keyboards the most popular in service today.

If rubber dome keyboards have a drawback at all, it is that the tactile feedback is not as good as many users would like. Although it is reasonable for most, some users prefer more tactile feedback than rubber dome keyboards normally provide.

The membrane keyboard is a variation on the rubber dome type in which the keys themselves are no longer separate, but are formed together in a sheet that sits on the rubber dome sheet. This severely limits key travel, and membrane keyboards are not considered usable for normal touch typing because of this. They are ideal in extremely harsh environments. Because the sheets can be bonded together and sealed from the elements, membrane keyboards can be used in situations in which no other type could survive. Many industrial applications use membrane keyboards especially for terminals that do not require extensive data entry but are used to operate equipment such as cash registers.

Capacitive switches are the only non-mechanical type of switch in use today. These are the Cadillac of key switches. They are much more expensive than the more common mechanical rubber dome, but they also are more resistant to dirt and corrosion and offer the highest-quality tactile feedback of any type of switch.

A capacitive switch does not work by making contact between conductors. Instead, two plates usually made of plastic are connected in a switch matrix designed to detect changes in the capacitance of the circuit.

When the key is pressed, the plunger moves the top plate relative to the fixed bottom plate. Usually a mechanism provides for a distinct over-center tactile feedback with a resounding "click." As the top plate moves, the capacitance between the two plates changes and is detected by the comparator circuitry in the keyboard.

Because this type of switch does not rely on metal contacts, it is nearly immune to corrosion and dirt. These switches are very resistant to key bounce problems that result in multiple characters appearing from a single strike. Capacitive switch keyboards are among the most expensive designs, but the quality of the feel and their durability are worth it.

Traditionally, the only vendors of capacitive key switch keyboards have been IBM and its keyboard division, Lexmark, which is why these keyboards have always seemed to stand out as superior from the rest.

Keyboards with Special Features

There are a number of keyboards on the market that have special features not found in the standard designs. These additional features can be simple things such as built-in calculators and clocks, to more complicated features such as integrated pointing devices, special character layouts, shapes, and even programmable keys.

Over the years, many have attempted to change the design of the standard keyboard in an attempt to improve typing speed and ergonomics. Around 1936, August Dvorak and William L. Dealy developed a modified character layout for the keyboard, which replaced the QWERTY layout we are all familiar with today.

The Dvorak-Dealy keyboard design is normally just called the Dvorak design for short. It featured different character positions on the keys designed to promote the alternation of hands during typing. The characters are arranged so that the vowels are in the home row under the left hand, while the consonants used most frequently are placed in the home row under the right hand. The theory was that this would dramatically improve typing speed; however, most tests show fairly modest improvements. The public being resistant to change, the Dvorak keyboard design has not achieved widespread popularity, and the familiar QWERTY layout is still by far the most common design.

A more recent trend is to change the shape of the keyboard instead of altering the character layout. This has resulted in a number of different so-called ergonomic designs. The goal is to shape the keyboard to better fit the human hand. The most common of these designs split the keyboard in the center, bending the sides back. Some allow the angle between the sides to be adjusted, such as with the Lexmark Select-Ease designs, while others are fixed, such as the Microsoft Natural keyboard. These split or bent designs more easily conform to the natural angle of the hands while typing. They can improve productivity and typing speed, as well as help prevent medical problems such as Carpal Tunnel Syndrome (tendon inflammation).

Virtually every keyboard company now has some form of similar ergonometric keyboard, and the same things apply with respect to quality and feel as with the standard keyboard designs. The Microsoft Natural Keyboard is manufactured for Microsoft by Keytronics, and uses the inexpensive light-touch key switches they are known for. For those who prefer a more rugged keyboard with higher quality switches, I recommend the Lexmark Select-Ease, Alps, NMB Technologies, or Lite-On keyboards. These keyboards are available with very high quality mechanical switches with a positive tactile feel to them. The Lexmark design, in particular, allows you to adjust the angle between the two sides of the keyboard from fully closed like a standard keyboard, to split at virtually any angle. You can even separate the two halves completely. It also features built-in palm rests, an oversized space bar, and cursor keys on both sides of the keyboard.

Although these ergonometric keyboards sound like a good idea, people are resistant to change, and none of these designs has yet to significantly displace the standard keyboard layout.

MOUSE:

The principal pointing tool used in computers is the mouse. A mouse is a hand-movable input device that controls the position of the cursor on the computer's screen. The mouse is rolled about on a desktop mouse pad and directs a pointer on the computer's display screen. The mouse pad is a rectangular rubber/foam pad that provides traction for the mouse ball. The cursor is a symbol on a computer display that shows where the next typed character will appear. The mouse pointer can be an arrow, a rectangle, or pointing finger. It is the symbol that indicates the position of the mouse on the display screen. On the bottom of a mouse is either a small ball similar to a ball bearing or a light sensor that detects that the mouse is being moved across a flat surface. On the top of the mouse are one to three pushbuttons that are used for activating commands. When the mouse is moved across a flat surface, the cursor on the computer's display moves accordingly. Using a mouse eliminates a considerable amount of typing. There are three main variations of the mouse- the trackball, the pointing stick, and the touch pad. The mouse is shown in figure 2.14.

Figure 2.14: Mouse

TOUCH SCREEN

A touch screen is an electronic visual display that can detect the presence and location of a touch within the display area with a finger or hand. Touchscreens are common in devices such as game consoles, tablet computers and smart phones. The touchscreen has two main attributes. First, it enables one to interact directly with what is displayed, rather than indirectly with a pointer controlled by a mouse or touchpad. Secondly, it lets one do so without requiring any intermediate device that would need to be held in the hand. They play a vital role in the design of digital appliances such as the personal digital assistant (PDA), video games, mobile phones, etc.

Figure 2.5: Touch screen

GRAPHICS TABLETS: Graphics tablet, also called digitizing tablet, graphics pad, or drawing tablet is a computer input device that allows one to hand-draw images and graphics. Graphics tables are flat board like surfaces directly connected to a computer display. The user draws on the tablet using a pencil-like device, and the image is transmitted to the display.

Graphics tablets enable the user to employ colours, textures, and patterns when creating images. These tablets are also used to capture data or handwritten signatures. It can also be used to trace an image from a piece of paper which is taped or secured to the surface.
Capturing data in this way, either by tracing or entering the corners of linear poly-lines or shapes is called digitizing. The graphics tablet is shown in figure 2.15.

Figure 2.15; Graphics tablet

LIGHT PENS: A light pen is an input device in the form of a light-sensitive wand used in conjunction with a monitor. It is a pen-shaped object with a light-sensitive cell at its end as shown in figure 2.16. Users can select from a list of choices displayed on the computer's screen by touching a light pen to the proper item.It allows the user to point to displayed objects, or draw on the screen, in a similar way to a touch screen but with greater positional accuracy. Light pens may also be used in highly technical fields for altering graphs and other drawings. A light pen can work with any cathode ray tube (CRT) -based display, but not with liquid crystal display (LCD) screens.

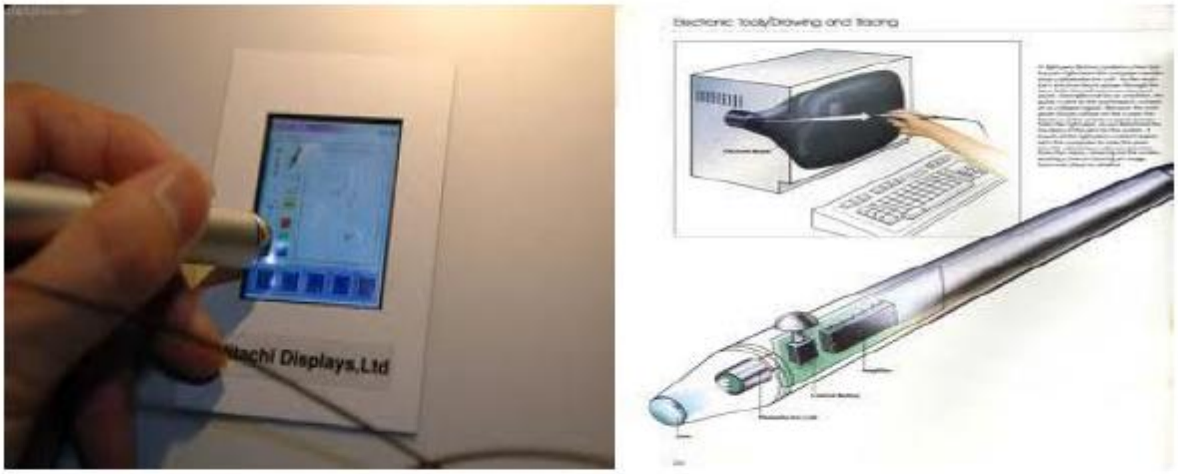


Figure 2.16: Light pen

JOYSTICK: A joystick is an input device used to control video games with one or more push buttons whose state are read by the computer. It is a stick that pivots on a base and reports its angle or direction to the device it is controlling. Joystick is used for controlling machines such as wheelchairs, surveillance camera, and trucks. It is also used in military jets as the principal flight control in the cockpit, where center stick or side-stick location is employed. A typical joystick is shown in figure 2.17.

Figure 2.17: Joystick

SCANNERS: Scanners are devices that read printed material so that it can be put in a computer readable form without having to retype, redraw, reprint, or rephotograph the material. Scanners have become very popular as desktop publishing software has become more and more popular. Scanners, in many different forms, also play an important role in source-data automation. Figure 2.18 is a diagram of a scanner.

Figure 2.18: Scanner

MICROPHONES: It is also possible for people to perform data entry by speaking into a microphone connected to a computer system. This capability is known as voice recognition.
At present, computers can identify and process a vocabulary of several dozen words or numbers and also to recognize an extensive vocabulary responding by creating transcripts of spoken text.

SOURCE - DATA AUTOMATION

Data entry has traditionally been the weakest link in the chain of data-processing operations. Although it can be processed electronically at extremely high speeds, significantly more time is required to prepare the data, enter it, and then check its accuracy before it can actually be processed. Another method of data entry, source-data automation, collects data about an event, in computer-readable form, when and where the event takes place. The most common forms of source-data automation are magnetic ink, character reader optical recognition devices etc.

Magnetic-ink was introduced in the late 1950s to sort, speed and check processing in the banking industry. Because magnetic-ink characters can be read by both human and machines, no special data-conversion step was needed. Magnetic-ink characters are formed with magnetized particles of iron oxide printed with magnetized ink, read by MICR equipment, producing a digitized signal. Each character is composed of certain sections of a seventy section matrix. The characters can be read and interrupted by a magnetic-ink character reader.

This process is called magnetic-ink character recognition (MICR). MICR reads the strange-looking numbers printed at the bottom of checks.

With MICR each character area is examined to determine the shape of the character represented. The presence of a magnetic field in a section of the area represents a 1 bit; the absence of a magnetic field represents a 0 bit. Each magnetic-ink character is composed of a unique combination of 0 bits and 1 bits. When all sections in a character area are combined and translated into binary notation, the character represented can be determined. MICR devices automatically check each character read to ensure accuracy.
Another form of source-data automation is optical recognition device. It read marks or symbols coded on paper documents and convert them into electrical pulses. The pulses can then be transmitted directly to the CPU or stored on magnetic tape for input at a later time.

The simplest approach to optical recognition is known as optical-mark recognition (OMR) or mark sensing. This approach is often used for machine scoring of multiple-choice examinations such as Joint Admission Matriculation Board (JAMB) now University Matriculation Examination (UME). A heavy lead pencil is used to mark the location of each desired answer. An optical-mark page reader senses the marks on the OMR document as the document passes under a light source. The presence of marks in specific locations is indicated by light reflected at those locations. As the document is read, the optical-mark data is automatically translated into machine language. When the optical-mark page reader is directly connected to the computer, thousands of forms of the same type can be read and processed in an extremely short period of time compared to entering the data by hand. Optical-mark recognition is also used in order writing, inventory control, surveys and questionnaire, and payroll applications.

Another type of optical reader, the bar-code reader, can read special line or bar codes. Bar codes are patterns of optical marks that represent information about the object on which the code appears. They are suitable for many applications, including point-of-sale (POS) systems, credit card verification, and freight identification to facilitate warehouse operations. Data is presented in bar code by the widths of the bars and the distances between them. The bar code readers are photoelectric (optical) scanners that translate the symbols in the bar code into digital code.

Optical character readers can read special types of characters known as optical characters. Some OCR devices can read the characters of several type fonts, including both uppercase and lowercase letters. OCR characters appear on utility bills and price tags on store/supermarket goods. A major difference between optical-character recognition and optical-mark recognition is that optical-character data is represented by the shapes of characters rather than by the positions of marks. However, both OCR and OMR devices rely on reflected light to translate written data into machine-readable form. Acceptable OCR input can be produced by computer printers, adding machines, cash registers, accounting machines, and typewriters. Data can be fed into the reader via a continuous form such as cash register tape or on cut forms such as phone or utility bills. When individual cut forms are used, the reader can sort the forms as well.

The most advanced optical-character readers can even read handwritten characters. Handwritten characters must be neat and clear, or they may not be read correctly. The system is not foolproof because handwriting can vary so much from person to person. The optical character readers reject any characters that cannot be interpreted. Devices that must read handwriting are often very slow. Machine-produced optical-character recognition has been used in credit card billing, utility billing, and inventory-control applications. Handwritten optical-character recognition has been used widely in mail sorting. Optical-character recognition systems are very reliable.

OUTPUT DEVICES

Output is data that has been processed into information by the computer. Output must be in a form that is convenient for the users. Printers, plotters, and visual display terminals can produce output. The principal kinds of output are soft copy and hard copy. Soft copy is data that is shown on a display screen or is in audio or voice form, while hard copy is printed output. There are several types of output devices such as printers, monitors and plotters.

Types of output are soft copy, hard copy, and other output (sound, voice, animation, and video). Printers and plotters produce hard copies of output, while monitor/display screen/visual display terminals (VDT) produce soft copy.

PRINTERS

A printer is an output device that prints characters, symbols, and graphics on paper or any other hardcopy medium. The resolution or quality of sharpness of the image is indicated by dpi, that is, dots per inch, which is a measure of the number of dots that are printed in a linear inch. Printers print processed data in form humans can read. To produce hard copy, the printer first receives electronic signals from the central processing unit. Printers can be grouped into two categories, according to whether or not the image produced is formed by physical contact of the print mechanism with the paper.

Impact printers do have contact with paper while nonimpact printers do not. In an impact printer, these signals activate print elements that are pressed against paper. Current nonimpact printers use heat, laser technology, or photographic techniques to print output.

Nonimpact printers are generally faster than impact printers because they involve less physical movement. They offer a wider choice of typefaces and better speed-to-price ratios than impact printers, and their technology implies a higher degree of reliability because they use fewer movable parts in printing. The disadvantages of nonimpact printers include the special paper requirements and the poor type image quality of some printers. Also, nonimpact printers cannot make carbon copies. Nonimpact printers can produce several copies of a page in less time than it takes an impact printer to produce one page with several carbon copies.

IMPACT PRINTERS*:*

Impact printers come in a variety of shapes and sizes. Some print one character at a time, while others print a line at a time. An impact printer forms characters or images by striking a mechanism such as a print wheel. Dot or wire-matrix printers and daisy-wheel printers are the principal types of impact printers. Figure 2.19 is the diagram of a printer-Dell inkjet.

Figure 2.19: Printer

Dot-matrix (also called wire-matrix) ***printers*** contain a print head of small pins, which strike an inked ribbon against paper to form characters or images. The matrix is a rectangle composed of pins. Certain combinations of pins are activated to represent characters. The dot combinations are used to represent various numbers, letters, and special characters. The quality of characters produced by a dot-matrix printer is determined by the resolution of the matrix used to produce the character. The more dots are printed in the matrix, the higher the quality of print. Near-letter quality (NLQ) characters produced by dot-matrix printers contain more dots placed closer together. Most dot-matrix printers have 9-, 18-, or 24- pin print heads. They print at speeds ranging from 80 characters per second in NLQ mode to 400 characters per second in draft mode.

Daisy-wheel printers use a daisy wheel, which is a flat disk with petal-like projections. Daisy wheels come in several type fonts that can be interchanged quickly to suit application needs. The daisy-wheel printer offers high-quality type and is often used in word-processing systems to give output a typewriter quality appearance.

Daisy-wheel printers can produce up to 3,000 characters per minute. Types of ***line-at-a-time printers*** include print wheel, chain, and drum printers. A print-wheel printer typical contains 120 – print wheels, one for each of 120 print positions on a line. Each print wheel contains forty-eight alphabetic, numeric, and special characters. The wheel rotates until the desired characters move into the corresponding print position on the current print line. When all wheels are in their correct positions a hammer drives the paper against the wheels and the entire line of output is printed. Print-wheel printers can produce about 150 lines per minute, which makes them rather slow compared to other line-at-a-time printers.

A ***chain printer*** has a character set assembled in a chain that revolves horizontally past all print positions. The printer has one print hammer for each column on the paper. Characters are printed when hammers press the paper against an inked ribbon, which in turn presses against appropriate characters on the print chain. The fonts can be changed easily on chain printers, allowing a variety of fonts, such as italic or boldface, to be used. Some chain printers can produce up to 2,000 lines per minute.

A ***drum printer*** uses a metal cylinder with rows of characters engraved across its surface. Each column on the drum contains a complete character set and corresponds to one print position on the line. As the drum rotates, all characters are rotated past the print position. A hammer presses the paper against the ink ribbon and drum when the appropriate character is in place. One line is printed for each revolution of the drum, since all characters eventually reach the print position during one revolution. Some drum can produce 2,000 lines per minute.

NON IMPACT PRINTERS:

Nonimpact printers form characters and images without direct physical contact between the printing mechanism and paper. It does not print characters by means of a mechanical printing element that strikes the paper. Instead, a variety of other methods are used. Nonimpact printers are faster and quieter than impact printers because they have fewer moving parts. Electrostatic, electrothermal, ink-jet, laser, and xerographic printers are examples of nonimpact printers.

An ***electrostatic printer*** forms an image of a character on special paper using a dot matrix of charged wires or pins. The paper is moved through a solution containing ink particles that have a charge opposite that of the pattern. The ink particles adhere to each charged pattern of the paper, forming a visible image of each character.

Electrothermal printers generate characters by using heat and heat-sensitive paper. Rods are heated in a matrix. As the ends of the selected rods touch the heat sensitive paper, an image is created. Both electrothermal and electrostatic printers operate quietly. They are used in applications where noise may be a problem. Some of these printers are capable of producing 5,000 lines per minute. The paper required for use in electrothermal printers is expensive and cost prohibitive.

Thermal printers use coloured waxes and heat to produce images by burning dots onto special paper. The coloured wax sheets are not required for black-and-white output. Thermal printers are expensive and they require expensive paper. It is the best for highest- quality colour printing.

In an ***ink-jet printer***, a nozzle is used to shoot a stream of charged ink towards the paper. Before reaching it, the ink passes through an electrical field that arranges that charged particles into characters. These printers can produce up to 12,000 characters per minute. Ink-jet printers spray small, electrically charged droplets of ink from four nozzles through holes in a matrix at high speed onto paper. Like laser and dot-matrix printers, ink-jet printers form images with little dots. It has a dpi of 300 to 1400. The advantages of ink-jet printers are- they can print in colour, quieter, and much less expensive than colour laser printers. The disadvantages are –they print in a lower resolution than laser printers and are slower.

Laser printers combine laser beams and electrophotographic technology to create output images. A beam of light focused through a rotating disk containing a full font of characters. The character image is projected onto a piece of film or photographic paper, and the print or negative is developed and fixed. The output consists of high quality, letter-perfect images. The process is often used in printing books. Like a dot-matrix printer, a laser printer creates images with dots. However, as in a photocopying machine, these images are produced on a drum treated with magnetically charged ink-like toner (powder), and then transferred from drum to paper. Laser printers run with software called page description language (PDL). This software tells the printer how to lay out the printed page, and it supports various fonts.

A laser printer comes with one of two types of PDL: PostScript (developed by Adobe) or PCL (Printer Control Language, developed by Hewlett-Packard). Printers have their own CPU, ROM, and RAM, usually 8MB. There are good reasons that laser printers are among the most common types of nonimpact printer. They produce sharp, crisp images of both text and graphics. They are quiet and fast, able to print 4 to 32 text-only pages per minute for individual microcomputers and up to 200pages per minute for mainframes. They can print in different font, that is, type styles and sizes.

The more expensive models can print in different colour. Laser printers have a dpi of 600 to 1200. Laser printers, which can operate at speeds of up to 21,000 lines per minute, are replacing the slower printers that have been used with word-processing systems in the past.

Xerographic printers use printing methods much like those used in common xerographic copying machines. For example, Xerox, the pioneer of this type of printing, has one model that prints on single S1/2-by-11-inch sheets of plain paper rather than on the continuous-form paper normally used. Xerographic printers operate at speeds of up to 4,000 lines per minute.

Multifunction printers – These are printers that do more than print. It combines several capabilities such as printing, scanning, copying and faxing. Xerox and Hewlett Packard make such machines. It takes up less space, cost less than the four separate office machines that they replace, but if one component malfunctions, nothing works.

PLOTTERS

Most times, the best way of computer to present information to a user is not in the form of text but in graphic form. A plotter is a specialized output device that prepares graphic, hard copy of information. It is designed to produce high-quality graphics in a variety of colours. It can produce lines, curves, and complex shapes. Plotters are often used to produce line, bar charts, graphs, organizational charts, engineering drawings, maps, trend lines, and supply and demand curves, etc, which are usually too large for regular printers. A typical plotter has a pen, a movable carriage, a drum, and a chart-paper holder. Shapes are produced as the pen moves back and forth across the paper along the x-axis while the drum moves the paper up and down along the x-axis. Both the paper movement and the pen movement are bidirectional. The pen is raised from and lowered to the paper surface automatically. Many plotters use more than one pen at a time. Changing the colours of the pens can change the colours of graphics. Plotter diagram is shown in figure 2.20.

Figure 2.20: Plotter

The three principal kinds of plotters are pen, electrostatic, and large-format. A pen plotter uses one or more coloured pens to draw on paper. It produces continuous lines, not patterns of dots. In an electrostatic plotter, paper lies partially flat on a table-like surface, and the toner is used in a photocopier-like manner. Large format plotters operate like an inkjet printer but on a much larger scale. It is used by graphic artists.

VISUAL DISPLAY DEVICES

Visual display devices are also called monitors, CRTs or screens. They are output devices that show programming instructions and data as they are being input and information after it is processed. The size of a screen is measured diagonally from corner to corner in inches. The most common sizes for desktop microcomputers are 13, 15, 17, 19, and 21; for laptop computers 12.1, 13.3, and 14.1inches. A typical screen can hold twenty-four lines, each containing eight characters. These terminals supply soft copy output, which means the screen image is not a permanent record of what is shown.
There are two types of monitors- CRT and flat panel. CRT (cathode ray tube) is a vacuum tube, used as a display screen in a computer or video display terminal. CRTs are well suited for applications involving inquiry and response, where no permanent (printed) records are required,

and can be used for capturing data transmitted to the screen for verification as it is keyed. Visual display devices have some advantages over printers. First, they can display output much faster than printers – some CRT terminals can display up to 10,000 characters in a second. Also, they are much quieter in operation than impact printers. It is usually possible to connect a printer or a copier to a CRT terminal, making it possible to obtain hard copy output of the screen contents. The flat panel display is made up of two plates of glass separated by a layer of a substance in which light is manipulated. Flat panel displays are used on portable microcomputers. They are much thinner and less bulky and require less power than the cathode ray tubes (CRTs) used in most display devices. They show the images less clearly. Looking at a flat panel display from an angle or in direct lighting makes the image faint or even invisible. Two common types of flat panel displays are liquid crystal display (LCD) and electroluminescent. The monitor is shown in figure 2.21.

Figure 2.21: Monitor

A type of CRT, known as graphic display device, is used to display drawings as well as characters on a screen. Graphic display devices are generally used to show graphs and charts, but they can also display complex curves and shapes. Graphic display devices are being used in highly technical fields such as the aerospace industry to aid in the design of new wing structures.

They are also used in computer-assisted design/computer-assisted manufacturing (CAD/CAM) areas, where objects can be designed and tested and the manufacturing process specified on the computer system in an interactive fashion Monitors are generally divided into three categories- monochrome, composite colour, and RGB (red-green-blue). Monochrome monitors display a single colour, such as white, green, or amber, against a black background. They display text clearly and are inexpensive. Most monochrome monitors are composite monitors so-called for their single video signal. They usually require no additional video circuitry in the computer. Composite colour monitors display a composite of colours received in a single video signal and are slightly more expensive than monochrome monitors. They have less clarity in displaying text than monochrome monitors. Images in a composite colour monitor are also less crisp than images on RGB monitors. RGB monitors receive three separate colours signals, one for each of three colours – red, green, and blue. Commonly used for high-quality graphics displays, they display sharper images than the composite colour monitors, but produce fuzzier text than monochrome monitors. They are more expensive than composite colour monitors. A colour graphics display adapter (circuit board) is necessary for using RGB monitors with most computers.

How CRT Works

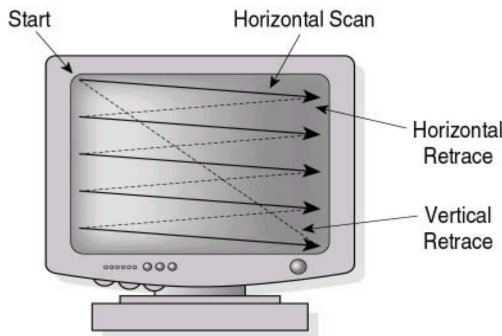
The wide end of the CRT is the display screen, which has a phosphor coating (a substance that can emit light when hit with radiation). When activated, the guns beam a stream of charged electrons onto the phosphorus coating. When the coating is hit with the right amount of energy (electrons), light is produced in a pattern of very small dots. This same technology is used in X-ray imaging, oscilloscopes, and other CRT devices. Similarly, monitors emit X-radiation. There is one dot for each primary color (RGB), and the dots are grouped in patterns close together. The name for a collection of all dots in a specific location is a pixel (which stands for picture element).

Image Formation and Refresh Rates

The human eye perceives the collection of pixels painted at the front of a CRT as a compound image, in much the same way as it interprets the pattern of ink dots in a newspaper as a photograph. The term *persistence* is used to define how long the phosphors on the screen remain excited and emit light.

The process of creating images is a sequential one. That is the image on the screen is not painted all at once. The stream is directed in rows, usually starting in an upper left corner. A series of raster lines are drawn down the face of the screen until the beam reaches the lower right, whereupon the process starts over. The persistence rate (how long a given line is visible) must hold for long enough to allow formation of a complete image, but not so long that it blurs the dots painted in the next pass.

These raster passes take place very quickly. The time required to complete a vertical pass is called the vertical refresh rate (VRR); the time required to pass once from left to right is known as the horizontal refresh rate (HRR). Generally speaking, faster is better. If the vertical rate is too slow, it can cause flicker, which is not only annoying, but can lead to eye strain. The larger the CRT, the faster the refresh rate must be to cover the entire area within the amount of time needed to avoid flicker. At 640 × 480 resolution, the minimum refresh rate is 60 Hz; at 1600 × 1200, the minimum rate is 85 Hz. Both the monitor and the display adapter produce the refresh rate.



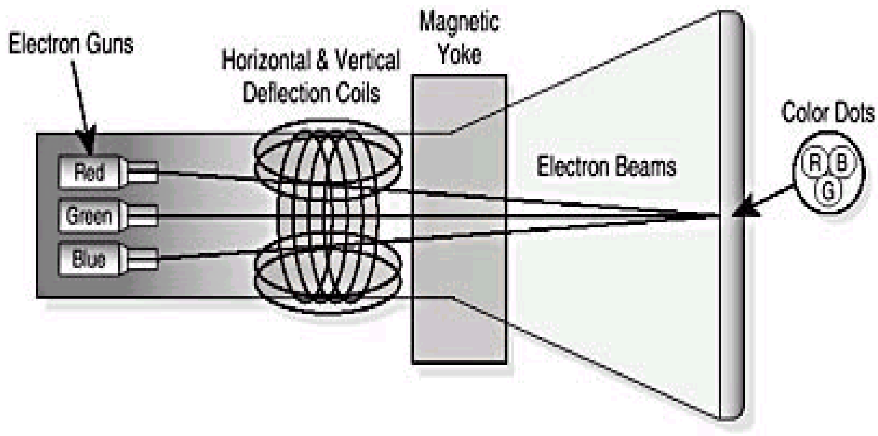
CRT Sweeping

Screen Resolution and Pitch

The term resolution refers to the degree of detail offered in the presentation of an image. The method of measurement varies, based on the medium. Photographic lenses, films, and paper are measured using lines per inch, whereas computer monitor manufacturers express resolution in pixels per inch. The greater the number of pixels per inch, the smaller the detail that can be displayed and, consequently, the sharper the picture.

Monitor resolution is usually expressed as a × b where a is the number of horizontal pixels, and b is the number of vertical pixels. For example, 640 × 480 means that the monitor resolution is 640 pixels horizontally by 480 pixels vertically. Modern monitors usually offer a variety of resolutions with different refresh rates. Price and quality should be compared at the maximum for both, along with two other factors, dot pitch and color depth.

Dot pitch is a term used to define the diagonal distance between the two closest dots of the same color, usually expressed in hundredths of millimeters. For example, you might see .25 dot pitch. Generally speaking, the smaller the pitch, the greater the number of dots, and the sharper the resulting image. The values for dot pitch are generally reflected in the monitor's price, and they are getting smaller as manufacturing technology improves. You should match the monitor's dot pitch and maximum resolution numbers to the needs of the customer, and install a graphics display card that will meet or exceed them.



Cathode Ray Tube

Do not confuse pixels with dots. A pixel is the smallest image unit the computer is capable of printing or displaying. It is usually the first number given in screen resolution: horizontal pixels × vertical raster lines. For example, 640 × 480 is the standard VGA resolution of 640 pixels per line, 480 lines deep.

Considerations in choosing a monitor

Cost and Picture Area

There is a direct link between the size of the picture tube and the cost of the monitor. The CRT is the most expensive part of the monitor. Graphical user interface (GUI) operating systems have increased the demand for bigger screens, to allow for more working area so that users can have more applications open at once or more working room for graphics.

Bandwidth

When referring to computer monitors, the term bandwidth is used to denote the greatest number of times an electron gun can be turned on and off in 1 second. Bandwidth is a key design factor because it determines the maximum vertical refresh rate of a monitor, measured in megahertz (MHz). Higher numbers are better. The lower the screens resolution, the faster the bandwidth. Therefore when comparing products, remember to measure bandwidth at the same resolution for each product.

Interlacing

Interlacing refreshes the monitor by painting alternate rows on the screen and then coming back and sweeping the sets of rows that were skipped the first time around. This increases the effective refresh rate but can lead to eye strain. Interlacing is found on less expensive monitors, and it should be avoided unless achieving the very lowest initial cost is the client's key concern.

Cost and Picture Area

There is a direct link between the size of the picture tube and the cost of the monitor. The CRT is the most expensive part of the monitor. Graphical user interface (GUI) operating systems have increased the demand for bigger screens, to allow for more working area so that users can have more applications open at once or more working room for graphics.

Power-Saving Features

Because they are the highest consumers of electrical current in the average PC, most new monitors provide some level of power-saving technology. Consequently, VESA (the Video Electronics Standards Association) has established a standard set of power economy controls to reduce power use when the monitor is idle. These are collectively referred to as DPMS **(Display Power Management Signaling) modes.**

DPMS technology uses monitors to gauge activity levels of the display. If there is no change in the data stream from the adapter, as set in either the BIOS (basic input/output system) or operating system controls, the monitor is switched to inactive status. The goal is to reduce power consumption while minimizing the amount of time required to restore the display to full intensity when needed. The following table lists DPMS stages, arranged in order from most to least power used.

Frequently turning a monitor ON and OFF places a lot of stress on its components. DPMS reduces the need to use the mechanical switch to turn the device on or off. You should advise clients without power-saving systems in place to turn on the display only when it is first needed and to turn it off at the end of each workday.

DPMS can be configured in one of three ways: using hardware, software, or a combination of both. When configuring a system for a new monitor, check the manufacturers' manuals for recommendations on appropriate settings and setup instructions.

Flat Panel Display

Flat-panel displays do the same job as CRT monitors but have some fundamental differences that make them more suitable for some environments than others.

Flat-panel displays (FPDs) are thin, bright display outputs that are gaining a foothold on desktops as a replacement for traditional CRT monitors. The most obvious benefit is the much smaller amount of desk space required, because there is no big case housing the electron gun, nor a heavy glass front. Because they don't rely on transitory phosphors to create an image, they are free from flicker (and produce no radiation). FPDs are also two to three times brighter than CRT screens. Since the screen is flat, this means that there is no distorted image at the edge of the viewing area, as there is with curved CRT monitors. FPDs are generally easier on the eyes and don't require a "warm-up" period to reach full color saturation.

With all those advantages, it would seem that only cost (they are much more expensive than CRT solutions) keeps them from sweeping the market. Not so. All technology has its pluses and minuses, and FPDs are no exception.

Take viewing angle, for example. If several people look at a CRT screen at angles up to 50 degrees off the center, they will all experience the same image—unless glare from a bright light source is reflected in someone's eyes. FPDs have limited optimal viewing angles, usually much narrower than CRTs.

Most FPDs also lack the range of resolutions of their CRT-based kin. To drive them, you will need a digital graphics adapter that is tuned to that FPD. The limited choice means an upgrade will usually require the purchase of a new card, and the cost will often be greater than a similar card for a comparable CRT monitor. The following table highlights the differences between FPD and CRT technology.

How Flat Panels Work

FPDs create an image made of pixels, just like their CRT counterparts, but they use different technology to accomplish that task. Several different types of FPDs are available today, varying in cost, image quality, and several other factors that affect both suitability to different computing applications and user acceptance.

Liquid Crystal Displays

Even the most computer-illiterate individuals probably have experience with LCD technology. It is used for a wide variety of inexpensive applications, from digital watches to children's toys, from pagers and cell phones to ATMs. LCDs form an image by using transparent organic polymers sandwiched between a pair of polarizing filters, with some form of back-lighting.

The filters are set at a 90-degree angle to each other. In an uncharged state (no current applied), the crystals are aligned so that light can pass through the top filter. When a current is added, the crystals align to the electric field, blocking the transmission of light. Not all LCD panels are created equal. The greater the twist angle, the higher the contrast and the more responsive the display is to changes in current.

Color light-emitting diode (LED) displays have three adjoining cells, each equipped with a different color filter: one red, one blue, and one green. This allows a display that makes use of the RGB color system.

There are several different types of LCD displays, varying in quality of output and cost. Passive-matrix displays (PMDs) are the simplest, and they have been used in calculators and watches as far back as 1970. PMDs are too slow for today's demanding multimedia PCs.

Active-matrix displays use TFTs (thin film transistors; TFT also describes this type of display) at each pixel to control each pixel's on off state. TFT makes up the majority of both laptop and desktop FPDs today. The image is formed by an array of LCDs on a wire grid. The result is a faster response than the passive array.

Emerging Flat-Panel Technologies

Electroluminescent Displays

Electroluminescent displays (ELDs) actually emit light, rather than simply controlling the transmission of a back-light source. The light generation comes from phosphors layered between front and back electrodes. There are both passive- and active-matrix variations of ELDs, much like those in LED technology. Right now most ELD products are found in technical applications (medical and defense) as well as ATM machines. Vendors will have to overcome problems in the quality of color and the higher power usage of ELDs compared to existing PC screens for these displays to gain acceptance as an alternative to CRT or TFT products.

Plasma Display Panels

Plasma display panels (PDPs) work much like the fluorescent lights found in most offices by energizing an inert gas. Phosphor films are used to produce a color image. This technology is used to manufacture very large FPDs. Like fluorescent lights, PDPs are relatively inexpensive to produce, but lower contrast and brightness, as well as higher relative power consumption, have thus far limited their use for PC applications.

VOICE SYNTHESIZES

Audio output is becoming a more and more popular form of computer output. The computer to output information to users in sequences of sound that resembles the human voice uses a voice synthesizer. The computer generates sounds electronically from texts entered by human operators or stored within the system. Two common forms of audio output in use today are the voice that gives you a telephone number when you call information and the voice that tells you the price of grocery items at automated grocery store checkout counters.

Other output – sound, voice, animation, and video

Most PCs are multimedia computers capable of outputting not only text and graphics but also sound, voice, video and animation. Sound-output devices produce digitized sounds, ranging from beeps and chirps to music. Voice-output devices convert digital data into speech like sounds. Video output consists of photographic images, which are played at 15-29 frames per second to give the appearance of full motion. Another form of video output is videoconferencing, in which people in different geographical locations can have a meeting, can see and hear one another, using computers and communications.

INTEGRATED CIRCUITS – IC

INTRODUCTION

The separately manufactured components like resistor, capacitor, diode, and transistor are joined by wires or by printed circuit board (PCB) to form circuits. These circuits are called discrete circuits and they have following disadvantages:

1. In a large electronic circuit, there may be very large number of components and as a result the discrete assembly will occupy a very large space.
2. They are formed by soldering which causes a problem of reliability.

To overcome these problems of space conservation and reliability, the integrated circuits were developed.

Integrated Circuit (IC).

An IC comprises a number of circuit components like resistors, transistor etc. They are interconnected in a single small package to perform the desired electronic function. These components are formed and connected within a small chip of semiconductor material. In IC the following features are observed.

1. In IC, the various components are integral part of a small semiconductor chip and the individual components cannot be removed for repair and replacement as in discrete circuits.
2. It combines both active elements like diodes and transistors with passive components like resistors and capacitors in a monolithic structure, so the complete unit in a monolithic circuit. Their size is very small. To see connections between their various components, a microscope is needed.

All the components are formed within the chip and no component is seen projected above the surface of the chip.

Scale of Integration

The number of components fitted into a standard size IC represents its integration scale, in other words it's a density of components. It is classified as follows:

1. SSI – Small Scale Integration
   It has less than 100 components (about 10 gates).
2. MSI – Medium Scale Integration
   It contains less than 500 components or have more than 10 but less than 100 gates.
3. LSI – Large Scale Integration
   Here number of components is between 500 and 300000 or have more than 100 gates.
4. VLSI – Very Large Scale Integration
   It contains more than 300000 components per chip
5. VVLSI - Very Very Large Scale Integration
   It contains more than 1500000 components per chip.

Comparison of IC and Discrete Circuits

Integrated circuits and discrete circuits can be compared as follows:
Integrated Circuits (IC's)

1. I.C can work on low voltages.
2. They can handle limited amount of power.
3. They are very small in size
4. They are cheap
5. Complex circuitry on a chip may be used to obtain improved performance characteristics.

Discrete Circuits

1. Discrete circuits require comparatively more voltage.
2. Discrete circuits can handle much more power than IC.
3. Circuit with discrete components acquires large space.
4. Discrete circuits are costly than IC's.
5. The performance is not good.

Components Fabrication Process

Usually electronics components such as diodes, resistors and capacitors are made on monolithic integrated circuit (IC). In order to fabricate these IC, components impurities are added or diffused in specific place in the semiconductor wafer (i.e substrate) so PN junction can be made. Figure (a) shows the cross section area of the basic monolithic components.

All the four components are made inside the P-type substrate or wafer. N-type and P-type portion are made from N-type and P-type materials inside the P-type substrate. However this is done by diffusion process. In this process P-type and N-type materials (in the form of gas) are added in semiconductor wafer at high temperature. Wafer is placed in a high temperature furnace (of about 100^{o} C).

At first, a thin layer of silicon diode ($SiO_{2)}$ layer is made at particular areas of N-type layer which are subject to diffusion. The N-type material is diffused in to substrate. Now the first and large N-type potion is diffused inside the substrate.

Again a thin layer of SiO_2 that is grown over the other, new place is selected to diffuse P-type material inside the N-type material. This same process is repeated to diffuse last potion of N-type material.

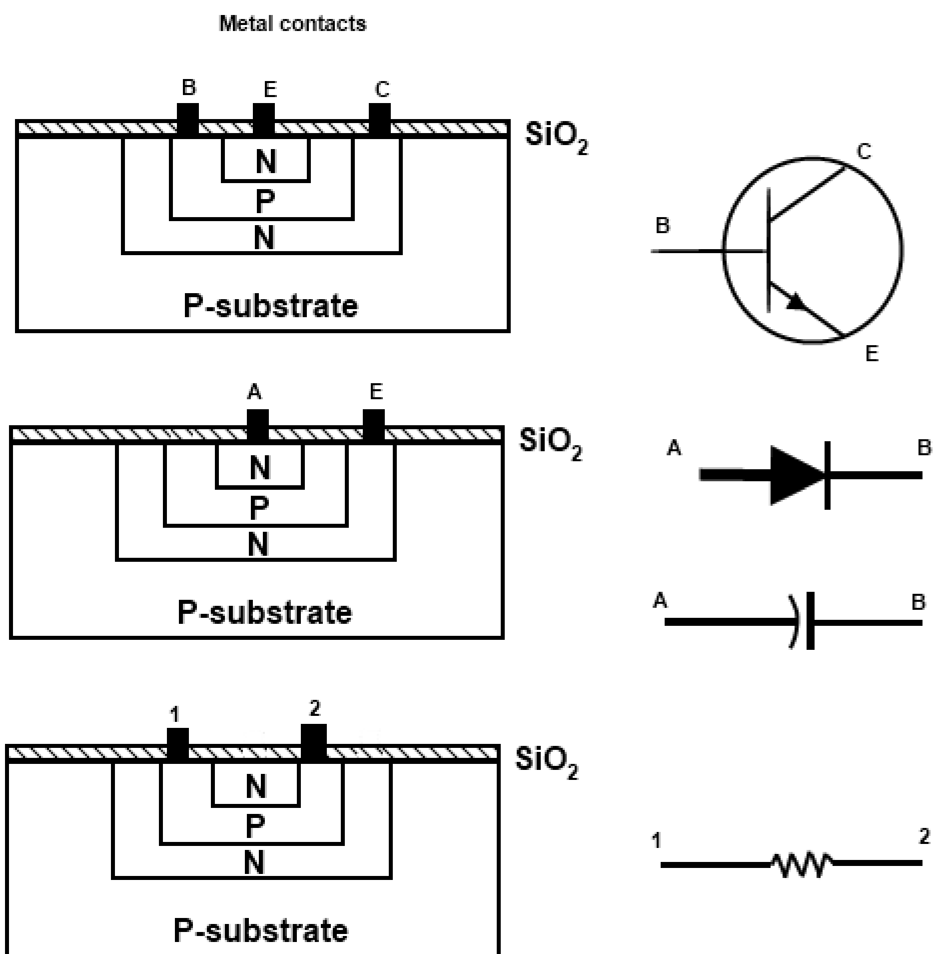


Figure (a)

An Overview of VLSI and IC Fabrication Process

The expansion of VLSI is 'Very-Large-Scale-Integration'. Here, the term 'Integration' refers to the complexity of the Integrated circuitry (IC). An IC is a well-packaged electronic circuit on a small piece of single crystal silicon measuring few mms by few mms, comprising active devices, passive devices and their interconnections.

The technology of making ICs is known as 'MICROELECTRONICS'. This is because the size of the devices will be in the range of micro, sub micrometers. The examples include basic gates to microprocessors, op-amps to consumer electronic ICs. There is so much evolution taken place in the field of Microelectronics, that the IC industry has the expertise of fabricating an IC successfully with more than 100 million MOS transistors as of today. ICs are classified keeping many parameters in mind. Based on the transistors count on the IC, ICs are classified as SSI, MSI, LSI and VLSI. The minimum number of transistors on a VLSI IC is in excess of 40,000.

The concept of IC was conceived and demonstrated by JACKKILBY of TEXAS INSTRUMENTS at Dallas of USA in the year 1958. The silicon IC industry has not looked back since then. A lot of evolution has taken place in the industry and VLSI is the result of this. This technology has become the backbone of all the other industries. We will see every other field of science and technology getting benefit out of this. In fact the advancements that we see in other fields like IT, AUTOMOBILE or MEDICAL, are because of VLSI. This being such important discipline of engineering, there is so much interest to know more about this.

What is VLSI?

VLSI is 'Very Large Scale Integration'. It is the process of designing, verifying, fabricating and testing of a VLSI IC or CHIP. A VLSI chip is an IC, which has transistors in excess of 40,000. MOS and MOS technology alone is used. The active devices used are CMOSFETs. The small piece of single crystal silicon that is used to build this IC is called a 'DIE'. The size of this die could be 1.5cms x 1.5cms. This die is a part of a bigger circular silicon disc of diameter 30cms. This is called a 'WAFER'.

Using batch process, where in 40 wafers are processed simultaneously, one can fabricate as many as 12,000 ICs in one fabrication cycle. Even if a low yield rate of 40% is considered you are liable to get as many as 5000 good ICs. These could be complex and versatile ICs. These could be PENTIUM Microprocessor ICs of INTEL, or DSP processors of TI, each costing around $10,000. Thus you are likely to make $50million out of one process flow. So there is lot of money in VLSI industry.

The initial investment to set up a silicon fabrication unit (called 'FAB' in short and also called some times as silicon foundry) runs into a few $Billion. In INDIA, we have only one silicon foundry-SCL at Punjab (Semiconductor Complex Ltd., in Chandigarh). Very stringent and critical requirements of power supply, cleanliness of the environment and purity of water are the reasons as to why there are not many FABS in India.

Complexity

Producing a VLSI chip is an extremely complex task. It has a number of design and verification steps. Then the fabrication step follows. The complexity could be best explained by what is known as 'VLSI design funnel' as shown in the Fig.1.1.

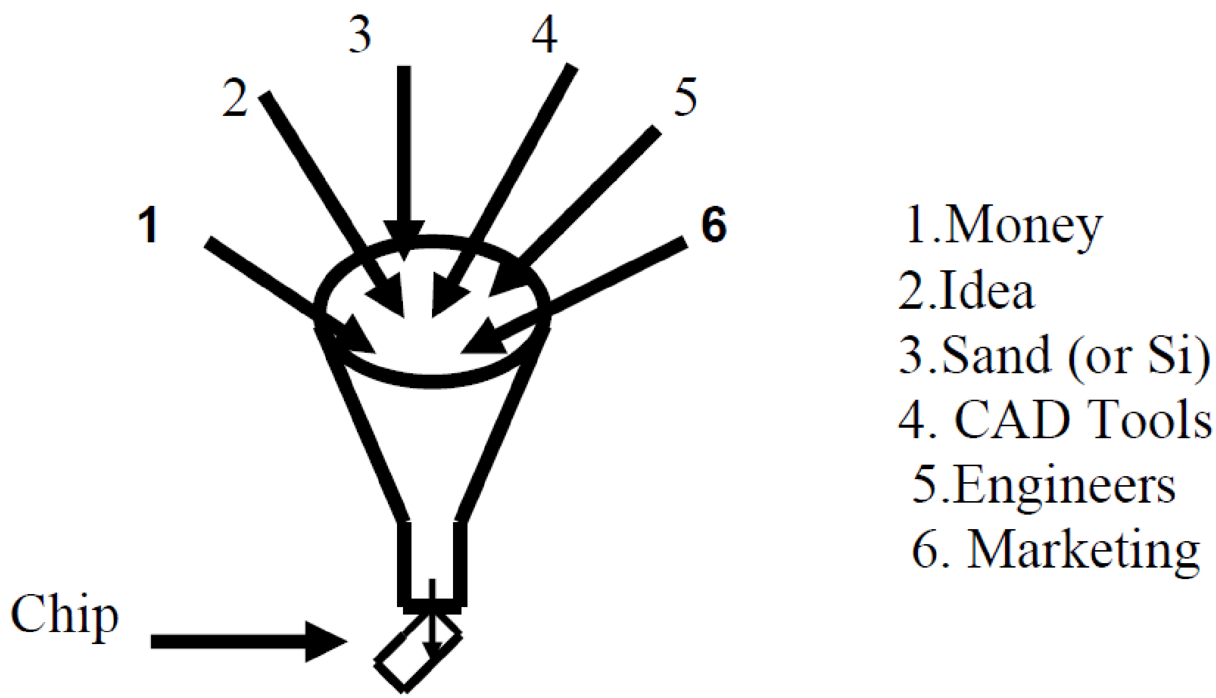


Figure1.1 The VLSI design funnel

To set up facilities for VLSI one needs a lot of money. Then the design starts at a highest abstraction in designer's mind as an initial idea. Engineers using CAD tools further expand this idea. One should have good marketing information also. Then all these are dumped inside the funnel along with a pile of sand as a raw material to get the wonderful item called "the VLSI chip'.

Design

A state of art of VLSI IC will have tens of millions of transistors. One human mind cannot assimilate all the information that is required to design and implement such complex chip. A design team comprising hundreds of engineers, scientists and technicians has to work on a modern VLSI project. It is important that each member of the team has clear understanding of his or her part of the contribution for the design.

This is accomplished by means of the design hierarchy. Any complex digital system may be broken down into gates and memory elements by successively subdividing the system in a hierarchical manner. Highly automated and sophisticated CAD tools are commercially available to achieve this decomposition. They take very high-level descriptions of system behavior and convert them into a form that ultimately be used to specify how a chip is manufactured.

A specific set of abstractions will help in describing the digital system, which is targeted for also chip. These are well depicted in the Fig.1.2 in a Y-chart. In this figure three distinct domains are marked in three directions in the form of letter Y. These domains are Behavioral, Structural and Physical. The behavioral domain specifies what a particular system does. The structural domain specifies how entities are connected together to effect the prescribed behavior (or function). The physical domain finally specifies how to actually build a structure that has the required connectivity to implement the required functionality.

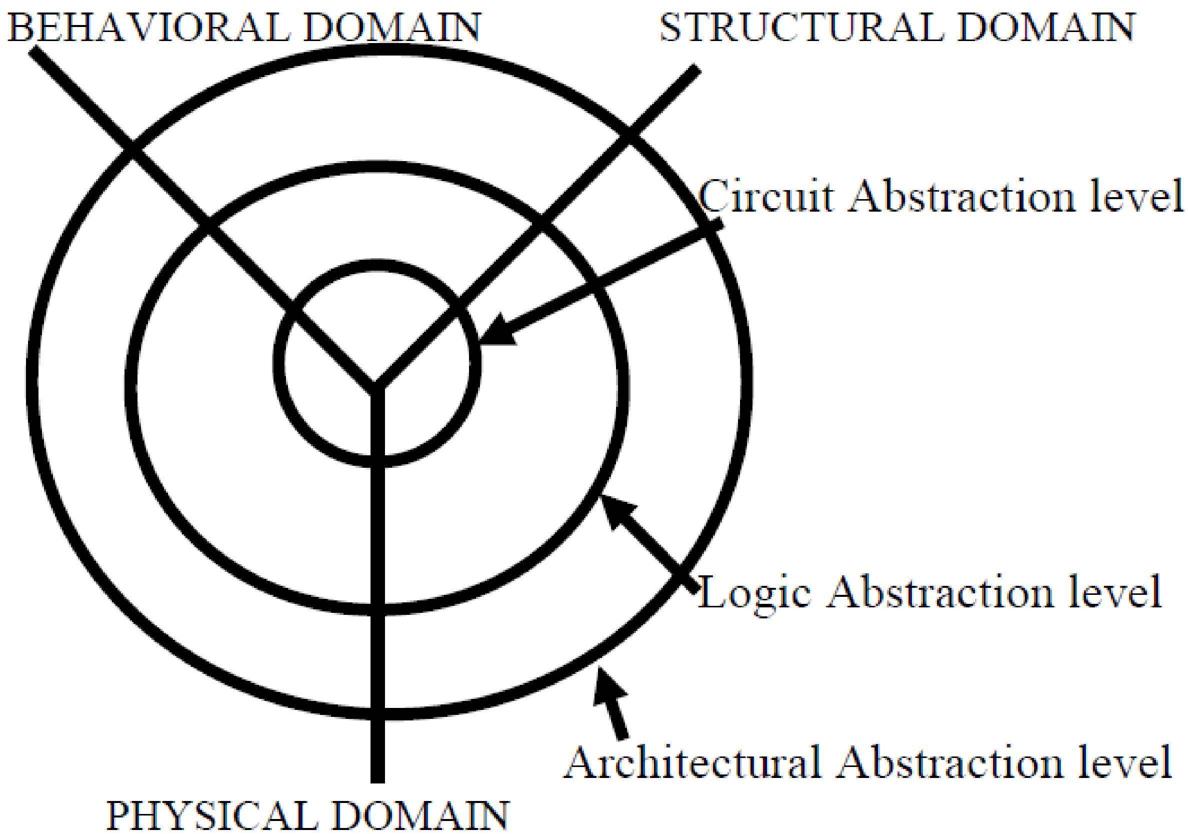


FIGURE 1.2 The Y-chart

Each design domain may be specified at a various levels of abstraction such as circuit, logic and architectural. Concentric circles around the center indicate these levels of abstraction. The design hierarchy is shown in the Fig.1.3 in the form of a flow.

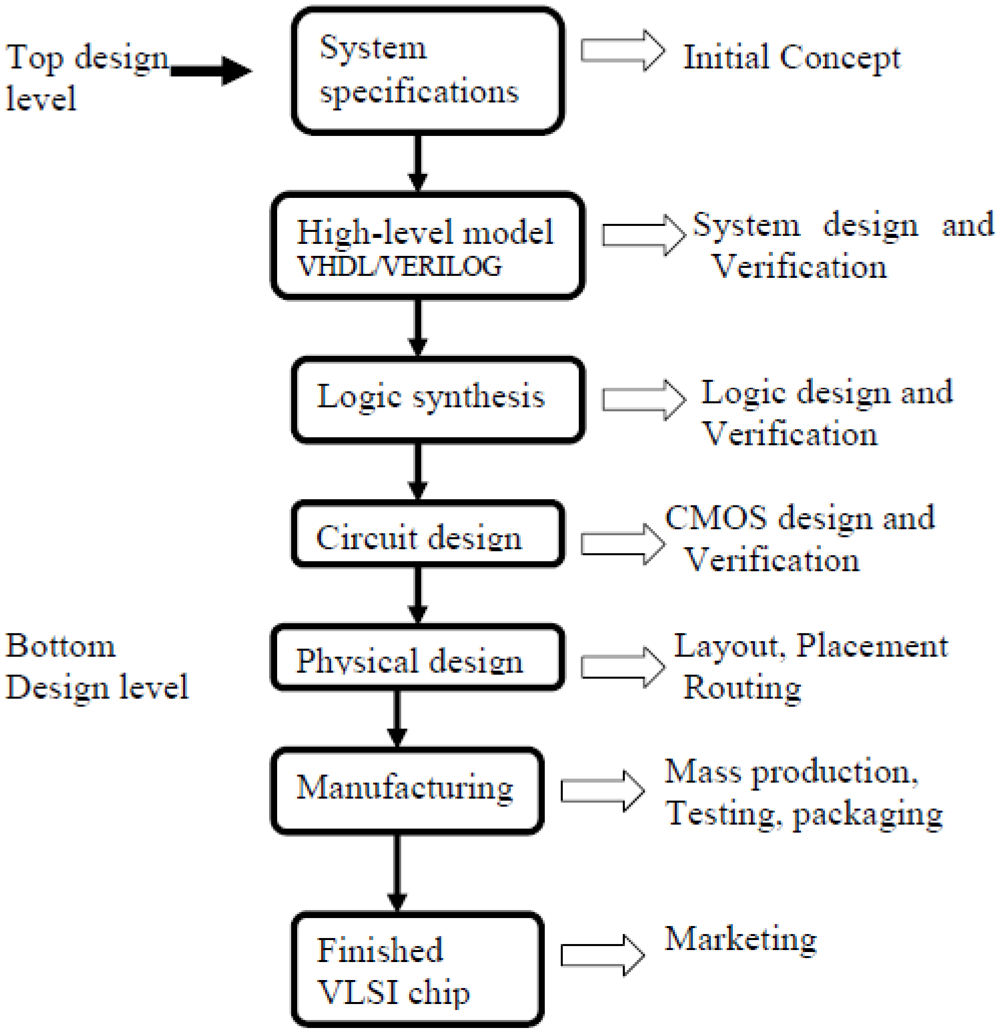


FIGURE 1.3 VLSI design flow

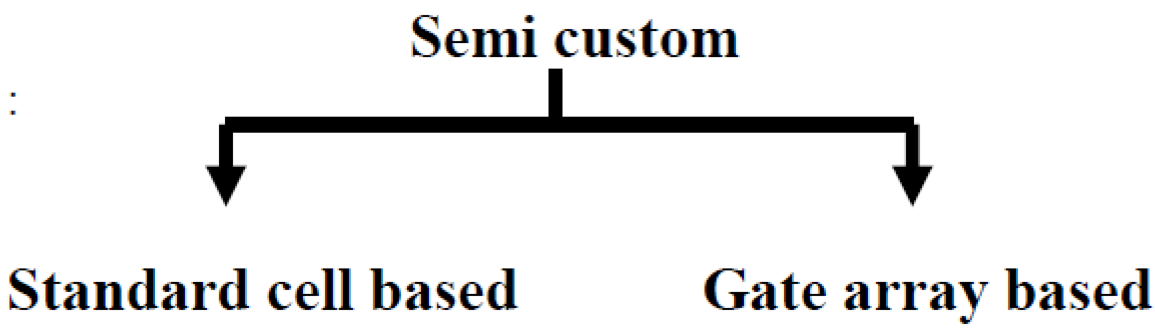
The starting point of a VLSI design (of a chip) is the system specifications. The specifications (specs) will provide design targets such as functions, speed, size, power dissipation, cost etc. This is the top level of the design hierarchy. These specifications are used to create an abstract, high-level model. The modeling is usually done using a high-level hardware description language (HDL), such as VHDL or VERILOG.

Using a HDL compiler the functional simulation is done. The next step in the flow is 'SYNTHESIS'. Synthesizer is a CAD tool (software), which generates a gate level net list (net list is the description of the logic gates and their interconnections) using the VHDL code as an input. This forms the basis for transferring the design to the electronic circuit level. In this level of design, the MOS transistors are used as switches, and Boolean variables are treated as varying voltage signals ('0' or '1' for digital) in the circuit.

Transistors cannot be decomposed further and these are the components, which have to be siliconized. To make a transistor on silicon, one should know the layout details of the same at various integration layer levels. This corresponds to 'physical design'. The physical design level constitutes the bottom of the design hierarchy. After the design process, the VLSI design project moves on to the manufacturing line. The final result is a finished electronic VLSI chip.

Design styles

The above said design process is called ‘top-down’ approach. That is because design moves from the highest abstraction level to the lowest level of abstraction namely transistor. The other design approach starts at the transistor (switch) level. Using transistors logic gates are built. Using gates, smaller functional units such as adders, registers and multipliers are built. These are used to build bigger blocks and the system. This approach is known as ‘bottom up’. This is also called the ‘Full – custom’ approach. This approach suits well for small projects. For most of the projects ‘semi-custom’ approach is used. Semi-custom is classified as below:



In standard cell based design, design details up to synthesis will be same as top-down approach. After synthesis, standard cells such as flip-flops, Registers and Multipliers are borrowed from the library to fit (mapping) into the design. These cells would have been custom designed and fully characterized for a given technology. The design cycle time is shorter compared to full custom and the IC is cheaper.

Gate array based design makes use of the transistors’ array or the array of standard gates. Wafers are pre-diffused having M x N array of transistors or Gates in silicon. It is up to the designer to use whatever the number of transistors and connect them in whatever the manner he needs. This is to say that the only design, which is done in this style, is the design of the interconnections, VDD, VSS buses and the communication paths. Masks have to be prepared only for the above said metal layers. Therefore this approach has got a very short design cycle time.

Full custom design as already discussed, stats at the transistor (switch) level. Here the logic style (CMOS, Bi CMOS, CCMOS,DYNAMIC CMOS etc.) will be decided and the transistor layout details (W/L ratio) will be designed. So this is sometimes called as ‘hand crafted design’. In this approach, silicon area is optimized and the IC comes out to be of high density. The performance is also very high, and the design is very flexible (can easily change).This approach is suitable for large and complex designs. Circuit simulator such as SPICE (Simulation Package with IC Emphasis) and the layout editor such as Magic or Lassi are used.

The design approach using FPGAs (Field Programmable Gate Array) could also be shown under semi-custom approach. FPGAs are off the shelf readily available CMOS VLSI ICs. These ICs have to be programmed (configured) to suit the designer’s requirement. There are 2 levels of programmability; one at the intra CLB level and the second at the inter CLB level. The front end design can be done using third part tools but one has to use the vendor specific tools for the back end to generate BIT file (analogous to the GDS-IITAPE step of full custom design approach).

The bit file is downloaded to a chosen FPGA. This programmed chip has to be interfaced with other components to finish the PCB card. Many such cards will give you a complete system. This is highly suited for proto typing of systems. As there is no fabrication step (a major step), this design cycle time is shortest among all the approaches.

Only drawback is that the designer may not be able to use the resources on the chip to the full extent.

Concept

VLSI should be thought of as a single discipline that deals with the conception, design and manufacture of complex ICs. Carver Mead is the gentle man who did pioneering work towards VLSI in 1970s. He came out with the standard definition with regard to the formation of a MOS transistor on silicon, which states that 'when polysilicon cuts across the diffusion, a transistor is formed at the intersection'. Thus he observed that the digital IC could be viewed as a set of geometrical patterns (polygons) on every layer that is going to be integrated on the silicon surface. Thus an IC will comprise of innumerable polygons of conducting (metal and polysilicon), semi conducting (silicon) and non-conducting (insulator such as SiO2) layers at various levels of integration.

Groups of patterns represent different logic functions and these are replicated throughout the IC. Thus the complexity is broken down using the concept of repeated patterns that were fitted together in a structured manner.

In VLSI two units are quite often used. They are micron (1**µ**m $=10^{-4}$ cm $= 10^{-6}$ meter) and Angstrom (1A0 = 10^{-8}cm $=10^{-10}$meter). If the channel length of a transistor = 2**π** = 0.18 **µ**m, then we refer that technology to a 0.18 **µ** technology. The size of the transistor has been reducing ever since the concept of IC was conceived since 1958. In 1970 Gordon Moore predicted the growth of microelectronics in terms of number of transistors that could be fabricated on a chip. He projected that the number of transistors would get doubled every 18 to 24 months. This has been established as 'MOORE'S LAW'. The silicon industry is facing a tough challenge to keep the pace with the law. On the other hand it is not possible to manufacture a functional design because of defects in the silicon crystal structure that cannot be avoided. The larger the area of the circuit, the higher the probability that a defect will occur. Even a single bad transistor or connection (because of the defect) would make the chip unusable. Therefore the philosophy is to keep the overall size of the chip small.

Peripheral Interface Adapter

Motorola MC6820 and MC6821 Peripheral Interface Adapters

The Peripheral Interface Adapter (PIA) is a peripheral integrated circuit providing parallel I/O interfacing capability for microprocessor systems. Common PIAs include the Motorola MC6820 and MC6821, and the MOS Technology MCS6520, all of which are functionally identical but have slightly different electrical characteristics. The PIA is most commonly packaged in a 40 pin DIP package.

The PIA is designed for glueless connection to the Motorola 6800 style bus, and provides 20 I/O lines, which are organized into 2 8-bit bidirectional ports (or 16 general-purpose I/O lines) and 4 control lines (for handshaking and interrupt generation). The directions for all 16 general lines (PA0-7, PB0-7) can be programmed independently. The control lines can be programmed to generate interrupts, automatically generate handshaking signals for devices on the I/O ports, or output a plain high or low signal.
It was used in Atari 800 family of computers (to provide four joystick ports to the machine). It was also used in the Apple I to interface the ASCII keyboard and the display.

Hardware Design

Microcontroller Concept

One way of looking at a computer system is to consider the successive "translations" that occur from the high level code (a programming language such as C++) to the electrical signals that "communicate" with the hardware. A computer system can be broken down into multiple levels or layers to show the translation of a specific instruction into a form that can be directly processed by the computer hardware. Such hierarchical levels are discussed in detail in *Structured Computer Organization* by A.S. Tanenbaum. This hierarchy is shown in Figure 2-1

High Level	Sum := Sum + 1
Assembly	MOV BX,SUM INC (BX)
Machine	1101010100001100 0010001101110101 1111100011001101
Register Transfer	Fetch Instruction, Increment PC, Load ALU with SUM ...
Gate	
Circuit	

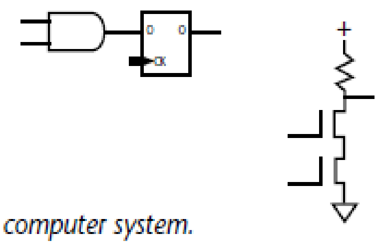


Figure 2-1: "Layers" of a computer system.

Language translators such as compilers and assemblers translate high-level code into machine code that can be executed by the processor.

Organization: von Neumann vs. Harvard

The von Neumann machine, with only one memory, requires all instruction and data transfers to occur on the same interface. This is sometimes referred to as the "von Neumann bottleneck." In common computer architectures, this is the primary upper limit to processor throughput. The Harvard architecture has the potential advantage of a separate interface allowing twice the memory transfer rate by allowing instruction fetches to occur in parallel with data transfers. Unfortunately, in most Harvard architecture machines, the memory is connected to the CPU using a bus that limits the parallelism to a single bus. The memory separation is still used as advantage in microcontrollers, as the program is usually stored in *non-volatile memory* (program is not lost when power is removed), and the temporary data storage is in *volatile memory.*

Non-volatile memories, such as *read-only memory* (ROM) are used in both types of systems to store permanent programs. In a desktop PC, ROMs are used to store just the start-up or bootstrap programs and hardware specific programs.

Volatile *random access memory* (RAM**)** can be read and written easily, but it loses its contents when power is removed. RAM is used to store both application programs and data in PCs that need to be able to run many different programs. In a dedicated embedded computer, however, the programs are stored permanently in ROM where they will always be available. Microcontroller chips that are used in dedicated applications generally use ROM for program storage and RAM for data storage. Memory technology is crucial to the design and understanding of embedded computers.

Microprocessor/Microcontroller Basics

There are three groups of signals, or buses, that connect the CPU to the other major components. The buses are:

• Data bus
• Address bus
• Control bus

The *data bus width* is defined as the number of bits that can be transferred on the bus at one time. This defines the processor's "word size." Many chip vendors define the word size based on the width of an internal data bus. For the purposes of this class, however, a processor with eight data bus pins is an 8-bit CPU. Both instructions and data are transferred on the *data bus* one "word" at a time. This allows the re-use of the same connections for many different types of information. Due to packaging limitations, the number of connections or pins on a chip is limited.

By sharing the pins in this way, the number of pins required is reduced at the expense of increased complexity in the external circuits. Many processors also take this a step further and share some or all of the data bus pins to carry address information as well. This is referred to as a *multiplexed address/data bus*. Processors that have multiplexed address/data buses require an external address latch to separate and hold the address information stable for the duration of a data transfer. The processor controls the direction of data transfer on the data bus.

The *address bus* is a set of wires that are used to point to the memory or I/O location that is to be read from or written to. The address signals must generally be held at a constant value for some period of time before, during, and after the data is transferred. In most cases, the processor actively drives the address bus with either instruction or data addresses.

The *control bus* is a collection of signals that determine what kind of information is on the data bus and determines where the data will go, in conjunction with the address bus. Most of the design process is concerned with the logic and timing of the control signals. The timing analysis is primarily involved with the relative timing between these control signals and the appearance and disappearance of data and addresses on their respective buses.

Microcontroller CPU, Memory, and I/O

The interconnection between the CPU, memory, and I/O of the address and data buses is generally a one-to-one connection. The hard part is designing the appropriate circuitry to adapt the control signals present on each device to be compatible with that of the other devices. The most basic control signals are generated by the CPU to control the data transfers between the CPU and memory, and between the CPU and I/O devices. The four most common types of CPU controlled data transfers are:

1. CPU reads data/instructions from memory *(memory read)*
2. CPU writes data to memory *(memory write)*
3. CPU reads data from an input device *(I/O read)*
4. CPU writes data to an output device *(I/O write)*

Design Methodology

The address decode and control logic shown in Figure 2-2 is the key part of the design, which requires attention to timing analysis to guarantee signal logic and timing compatibility between the other blocks. The simplified timing diagram for such a system is shown in Figure 2-3. Figure 2-3is a generic diagram and represents a typical example of a bus cycle for a typical CPU.

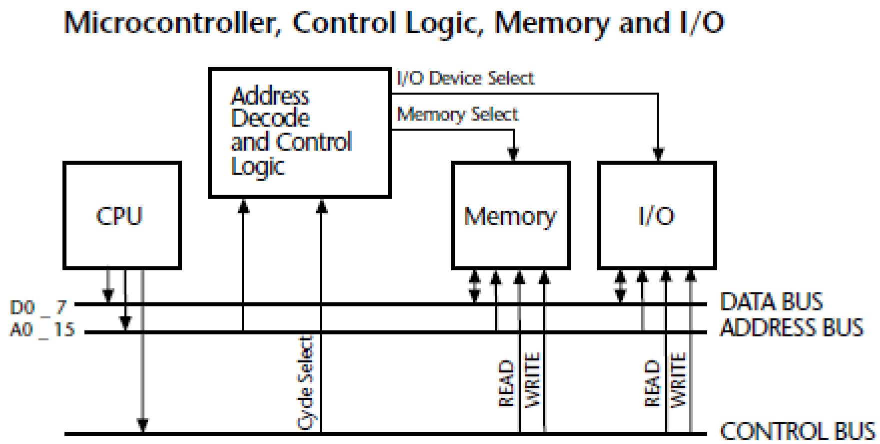


Figure 2-2: Microcomputer busses.

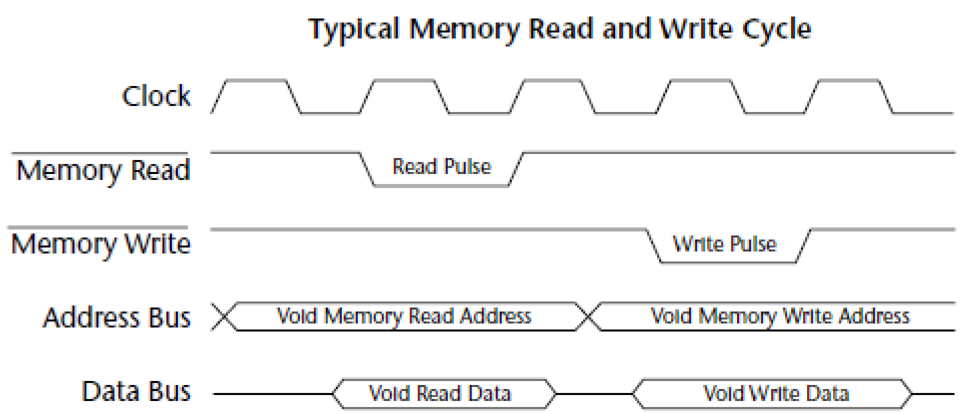


Figure 2-3: Generic bus timing example.

We see that there are two cycles:

• **Memory Read.** The processor places an address on the address bus, and activates the memory read signal by pulling it low, which causes the selected memory location to be placed on the data bus.

• **Memory Write.** The processor places an address on the address bus, data to be written on the data bus, and activates the memory write signal by pulling it low, which causes the selected memory location to be loaded with the data the CPU placed on the data bus.

Understanding Operational Amplifier (Op Amp)

The term operational amplifier, abbreviated op amp, was coined in the 1940s to refer to a special kind of amplifier that, by proper selection of external components, can be configured to perform a variety of mathematical operations. Early op amps were made from vacuum tubes consuming lots of space and energy. Later op amps were made smaller by implementing them with discrete transistors. Today, op amps are monolithic integrated circuits, highly efficient and cost effective.

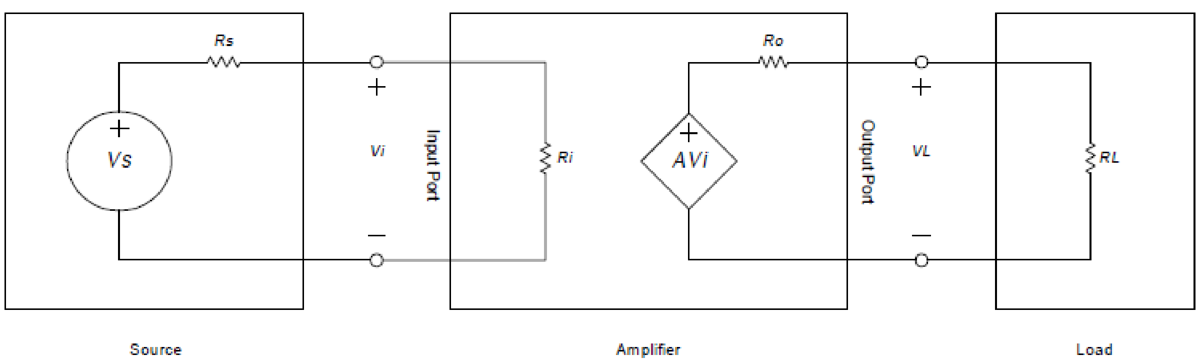
Amplifier Basics

Before jumping into op amps, let's take a minute to review some amplifier fundamentals. An amplifier has an input port and an output port. In a linear amplifier, output signal = *A* + input signal, where *A* is the amplification factor or gain.
Depending on the nature of input and output signals, we can have four types of amplifier gain:

- Voltage (voltage out/voltage in)
- Current (current out/current in)
- Transresistance (voltage out/current in)
- Transconductance (current out/voltage in)

Since most op amps are voltage amplifiers, we will limit our discussion to voltage amplifiers.
Thevenin's theorem can be used to derive a model of an amplifier, reducing it to the appropriate voltage sources and series resistances. The input port plays a passive role, producing no voltage of its own, and its Thevenin equivalent is a resistive element, *Ri*. The output port can be modeled by a dependent voltage source, *AVi*, with output resistance, *RO*. To complete a simple amplifier circuit, we will include an input source and impedance, *VS* and *RS*, and output load, *RL*. Figure 1 shows the Thevenin equivalent of a simple amplifier circuit.

Figure 1.Thevenin Model of Amplifier with Source and Load

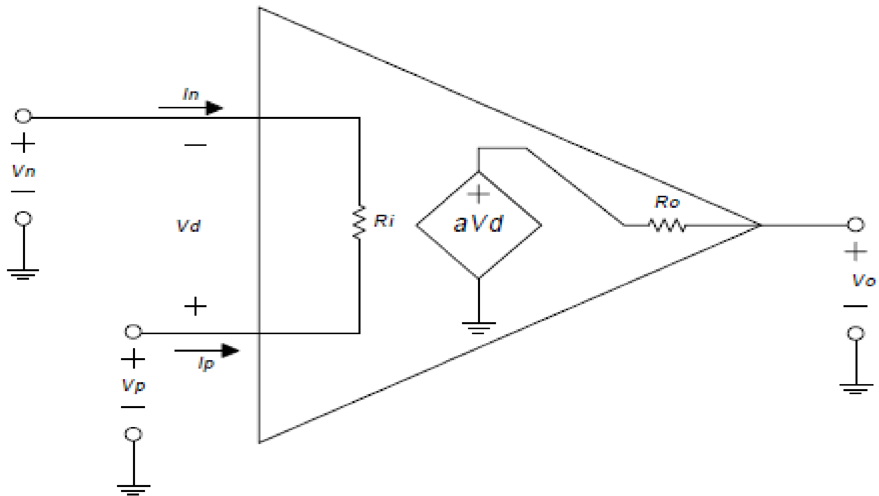


It can be seen that we have voltage divider circuits at both the input port and the output port of the amplifier. This requires us to re-calculate whenever a different source and/or load is used and it complicates circuit calculations.

Ideal Op Amp Model

The Thevenin amplifier model shown in Figure 1 is redrawn in Figure 2 showing standard op amp notation. An op amp is a differential to single-ended amplifier. It amplifies the voltage difference, $Vd = Vp - Vn$, on the input port and produces a voltage, VO, on the output port that is referenced to ground.

Figure 2. Standard Op Amp Notation



We still have the loading effects at the input and output ports as noted above. The ideal op amp model was derived to simplify circuit calculations and is commonly used by engineers in first order approximation calculations. The ideal model makes three simplifying assumptions:

1. Gain is infinite $\mathbf{a} = \infty$ (1)
2. Input resistance is infinite $\boldsymbol{Ri} = \infty$ (2)
3. Output resistance is zero $\boldsymbol{RO} = \mathbf{0}$ (3)

Applying these assumptions to Figure 2 results in the ideal op amp model shown in Figure 3.

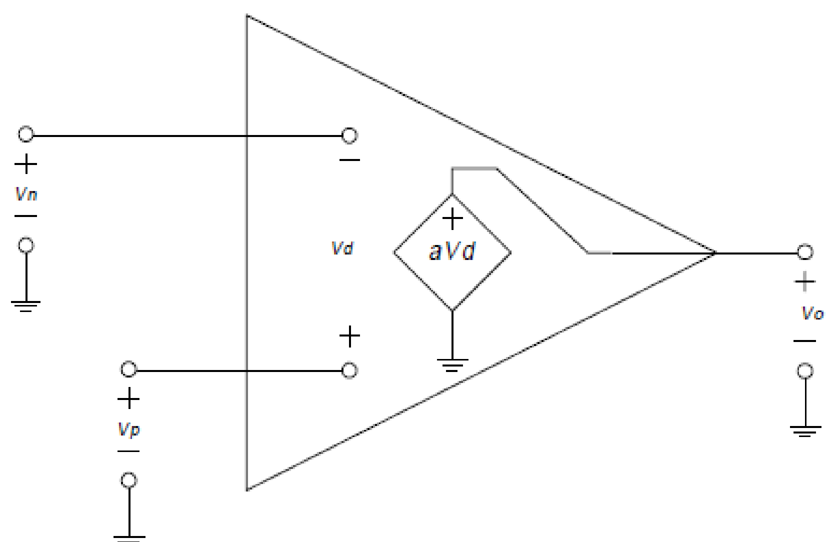


Figure 3. Ideal Op Amp Model

Other simplifications can be derived using the ideal op amp model:

- **In = Ip = 0** (4)

Because Ri = ∞, we assume *In* = *Ip* = 0. There is no loading effect at the input.

- **Vo =** a **Vd** (5)

Because *RO* = 0 there is no loading effect at the output.

- **Vd = 0** (6)

If the op amp is in linear operation, *VO* must be a finite voltage. By definition *VO* = *Vd´* a. Rearranging, *Vd* = *VO* / a . Since a = ∞,*Vd* = Vo / ∞ = 0. This is the basis of the virtual short concept.

- **Common mode gain = 0** (7)

The ideal voltage source driving the output port depends only on the voltage difference across its input port. It rejects any voltage common to *Vn* and *Vp*.

- **Bandwidth** = ∞ (8)
- **Slew Rate** = ∞ (9)

No frequency dependencies are assumed.

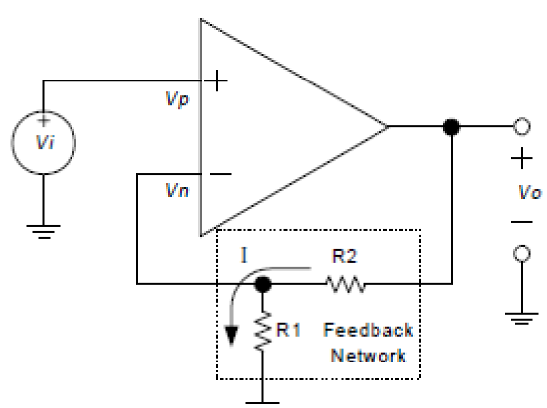
- **Drift = 0** (10)

There are no changes in performance over time, temperature, humidity, power supply variations, etc.

Non-Inverting Amplifier

An ideal op amp by itself is not a very useful device since any finite input signal would result in infinite output. By connecting external components around the ideal op amp, we can construct useful amplifier circuits. Figure 4 shows a basic op amp circuit, the non-inverting amplifier. The triangular gain block symbol is used to represent an ideal op amp. The input terminal marked with a + (Vp) is called the non-inverting input; – (Vn) marks the inverting input.

Figure 4. Non-Inverting Amplifier



To understand this circuit we must derive a relationship between the input voltage, *Vi*, and the output voltage, *VO*. Remembering that there is no loading at the input,

$$V_p = V_i \qquad (11)$$

The voltage at *Vn* is derived from *VO* via the resistor network, *R1*and *R2*, so that,

$$V_n = V_o \frac{R_1}{R_1+R_2} = V_o b \qquad (12)$$

Where,

$$b = \frac{R_1}{R_1+R_2} \qquad (13)$$

The parameter b is called the feedback factor because it represents the portion of the output that is fed back to the input. Recalling the ideal model,

$$V_O = aV_d = a(V_p - V_n) \qquad (14)$$

Substituting,

$$V_O = a(V_i - bV_O) \qquad (15)$$

And collecting terms yield,

$$A = \frac{V_o}{V_i} = \left(\frac{1}{b}\right)\left(\frac{1}{1+\frac{1}{ab}}\right) \qquad (16)$$

This result shows that the op amp circuit of Figure 4 is itself an amplifier with gain A. **Since the polarity of Vi and VO are the same, it is referred to as a non-inverting amplifier**. *A* is called the close loop gain of the op amp circuit, whereas *a* is the open loop gain. The product *ab* is called the loop gain. This is the gain a signal would see starting at the inverting input and traveling in a clockwise loop through the op amp and the feedback network.

Closed Loop Concepts and Simplifications

Substituting a = ∞ (1) into (16) results in,

$$A = \frac{1}{b} = 1 + \frac{R_2}{R_1} \qquad (17)$$

Recall that in equation (6) we state that *Vd*, the voltage difference between *Vn* and *Vp*, is equal to zero and therefore, *Vn* = *Vp*. Still they are not shorted together. Rather there is said to be a virtual short between *Vn* and *Vp*. The concept of the virtual short further simplifies analysis of the non-inverting op amp circuit in Figure 4. Using the virtual short concept, we can say that,

$$V_n = V_p = V_i \qquad (18)$$

Realizing that finding Vn is now the same resistor divider problem solved in (12) and substituting (18) into it, we get,

$$V_i = V_o \frac{R_1}{R_1+R_2} = V_o b \qquad (19)$$

Rearranging and solving for A, we get,

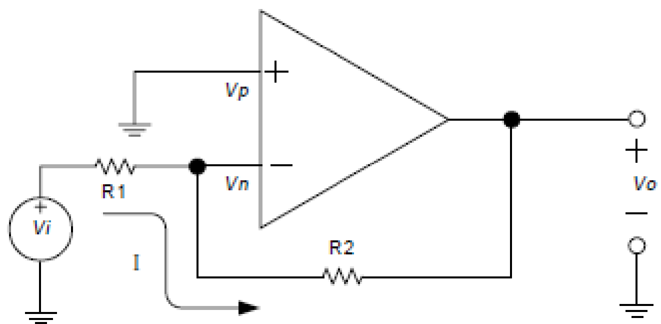
$$A = \left(\frac{1}{b}\right) = 1 + \left(\frac{R2}{R1}\right)$$

The same result is derived in (17). Using the virtual short concept reduced solving the non-inverting amplifier, shown in Figure 4, to solving a resistor divider network.

Inverting Amplifier

Figure 5 shows another useful basic op amp circuit, the inverting amplifier. The triangular gain block symbol is again used to represent an ideal op amp. The input terminal, + (Vp), is called the non-inverting input, whereas – (Vn) marks the inverting input. It is similar to the non-inverting circuit shown in Figure 4 except that now the signal is applied to the inverting terminal via R1 and the non-inverting terminal is grounded.

Figure 5. Inverting Amplifier



To understand this circuit, we must derive a relationship between the input voltage, *Vi* and the output voltage, *VO*.
Since *Vp* is tied to ground,

$$V_p = 0 \qquad (20)$$

Remembering that there is no current into the input, the voltage at *Vn* can be found using superposition. First let *VO* = 0,

$$V_n = V_i \left(\frac{R_2}{R_1 + R_2}\right) \qquad (21)$$

Next let *Vi* = 0,

$$V_n = V_o \left(\frac{R_1}{R_1 + R_2}\right) \qquad (22)$$

Combining,

$$V_n = V_o\left(\frac{R_1}{R_1+R_2}\right) + V_i\left(\frac{R_2}{R_1+R_2}\right) \quad (23)$$

Remembering equation (14), $VO = aVd = a(Vp- Vn)$, substituting and rearranging,

$$A = \frac{V_o}{V_i} = 1 - \left(\frac{1}{b}\right)\left(\frac{1}{1+\frac{1}{ab}}\right) \quad (24)$$

Where,

$$b = \frac{R_1}{R_1+R_2}$$

Again we have an amplifier circuit. **Because $b \leq 1$, the closed loop gain, *A*, is negative, and the polarity of *VO* will be opposite to *Vi*. Therefore, this is an inverting amplifier**.

Closed Loop Concepts and Simplifications

Substituting a = ∞ (1) into (24) results in

$$A = 1 - \frac{1}{b} = -\frac{R_2}{R_1} \quad (25)$$

Recall that in equation (6) we stated that *Vd*, the voltage difference between *Vn* and *Vp*, was equal to zero so that *Vn* = *Vp*. Still they are not shorted together. Rather there is said to be a virtual short between *Vn* and *Vp*. The concept of the virtual short further simplifies analysis of the inverting op amp circuit in Figure 5 Using the virtual short concept, we can say that

$$V_n = V_p = 0 \quad (26)$$

In this configuration, the inverting input is a virtual ground. We can write the node equation at the inverting input as

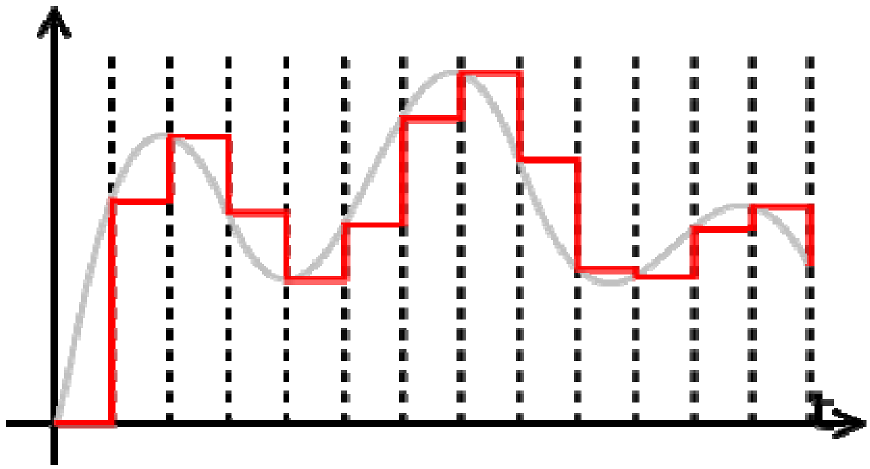
$$\frac{V_n - V_i}{R_1} + \frac{V_n - V_o}{R_2} = 0 \quad (27)$$

Since *Vn* = 0, rearranging, and solving for *A* we get

$$A = 1 - \frac{1}{b} = -\frac{R_2}{R_1} \quad (28)$$

The same result is derived more easily than in (24). Using the virtual short (or virtual ground) concept reduced solving the inverting amplifier, shown in Figure 5, to solving a single node equation.

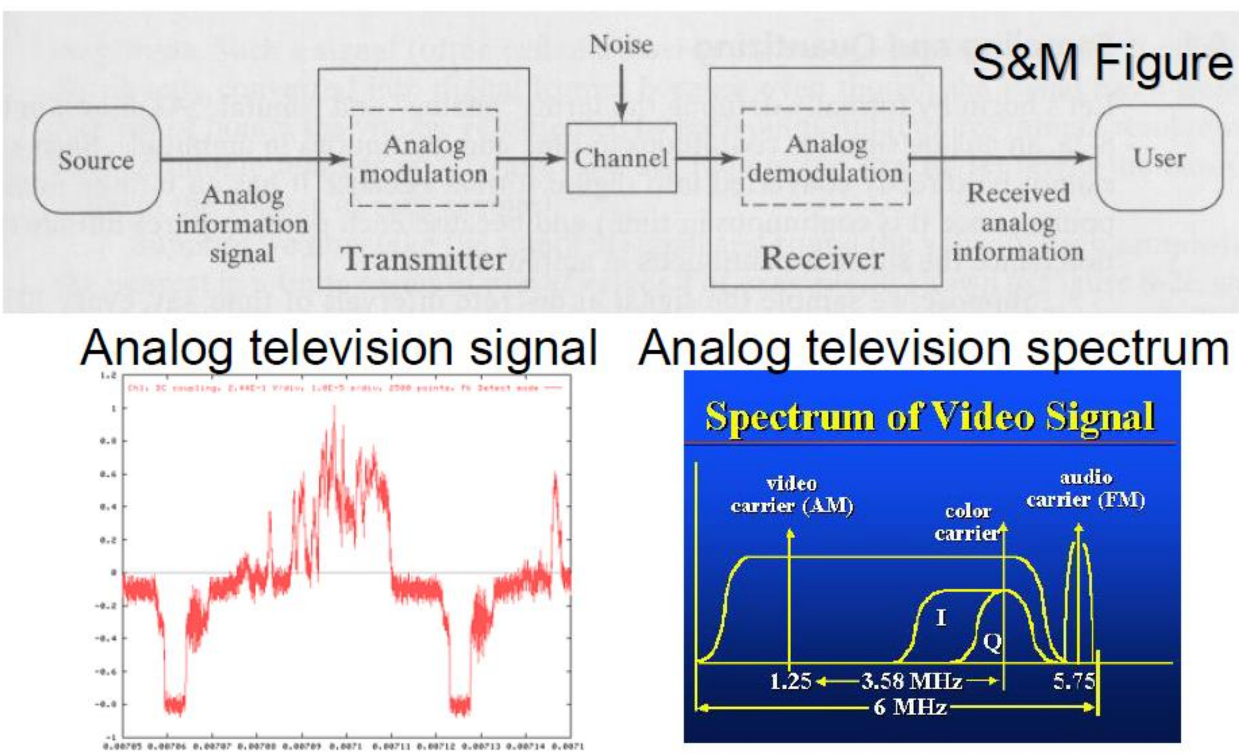
Analog-to-Digital and Digital to Analog Conversion



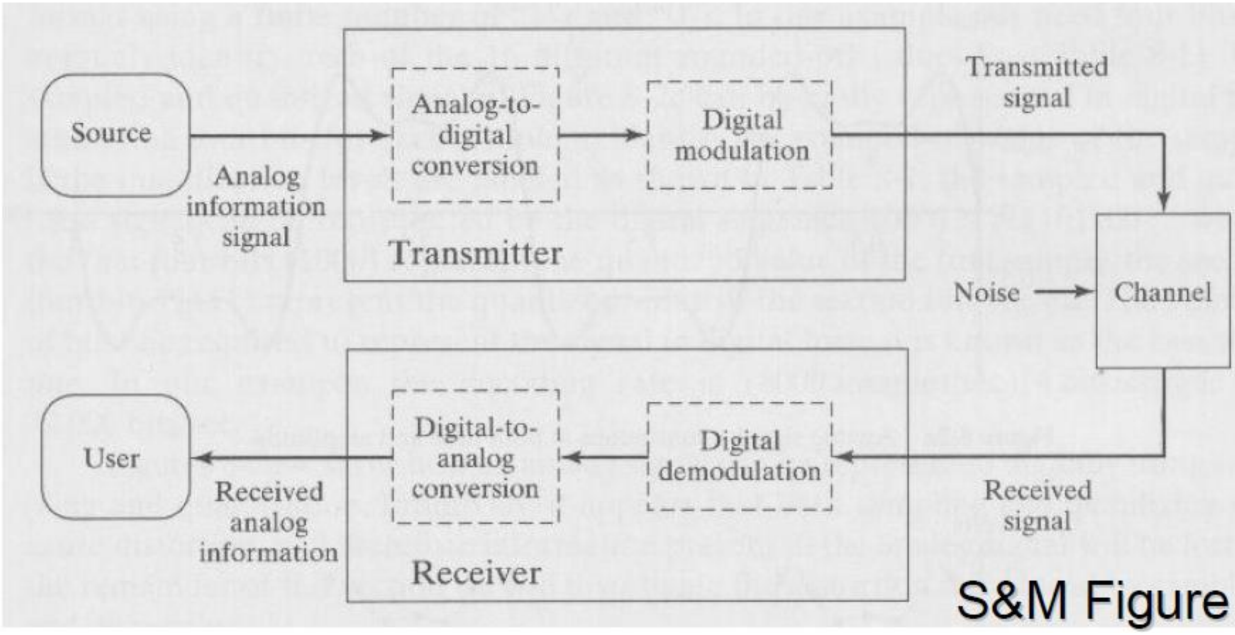
Analog and Digital Signal Samples

Sampling and Quantization

Traditional analog transmission (AM, FM and PM) are less complex than digital data transmission that have been the basis of broadcasting and communication for over 100 years.

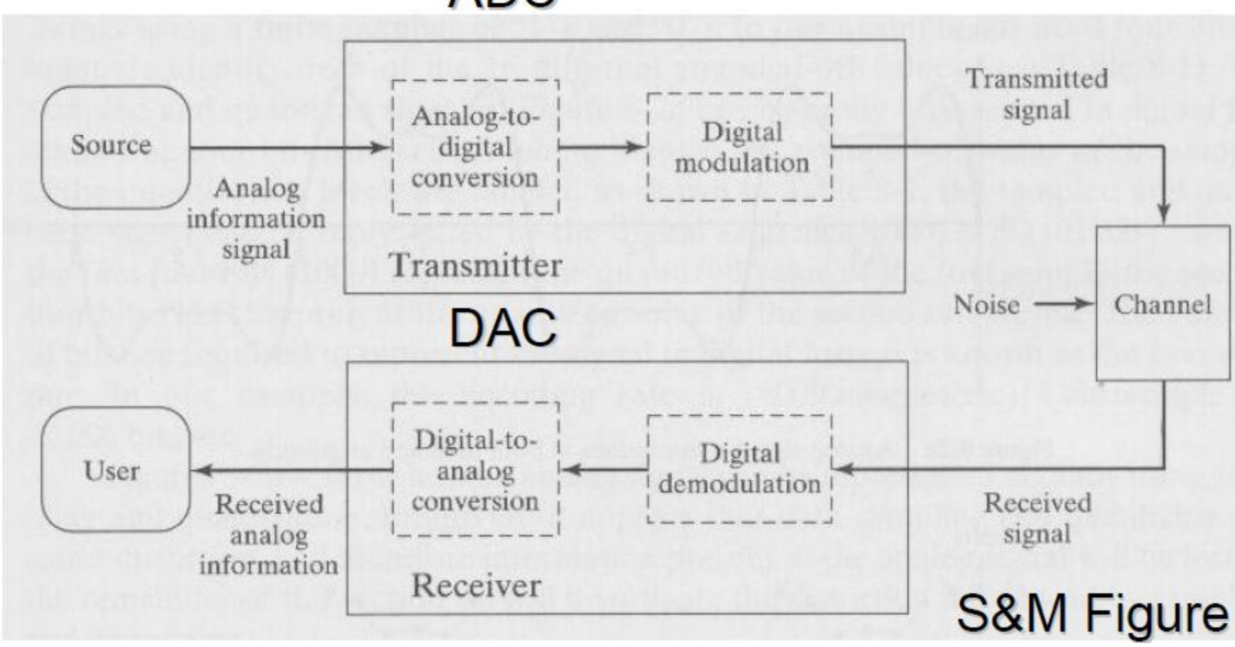


Digital data transmission (PAM, ASK, PSK, FSK and QAM) is more complex but (perhaps) offers higher performance with control of accuracy and easier storage, simpler signal processing for noise reduction, error detection and correction and encryption.

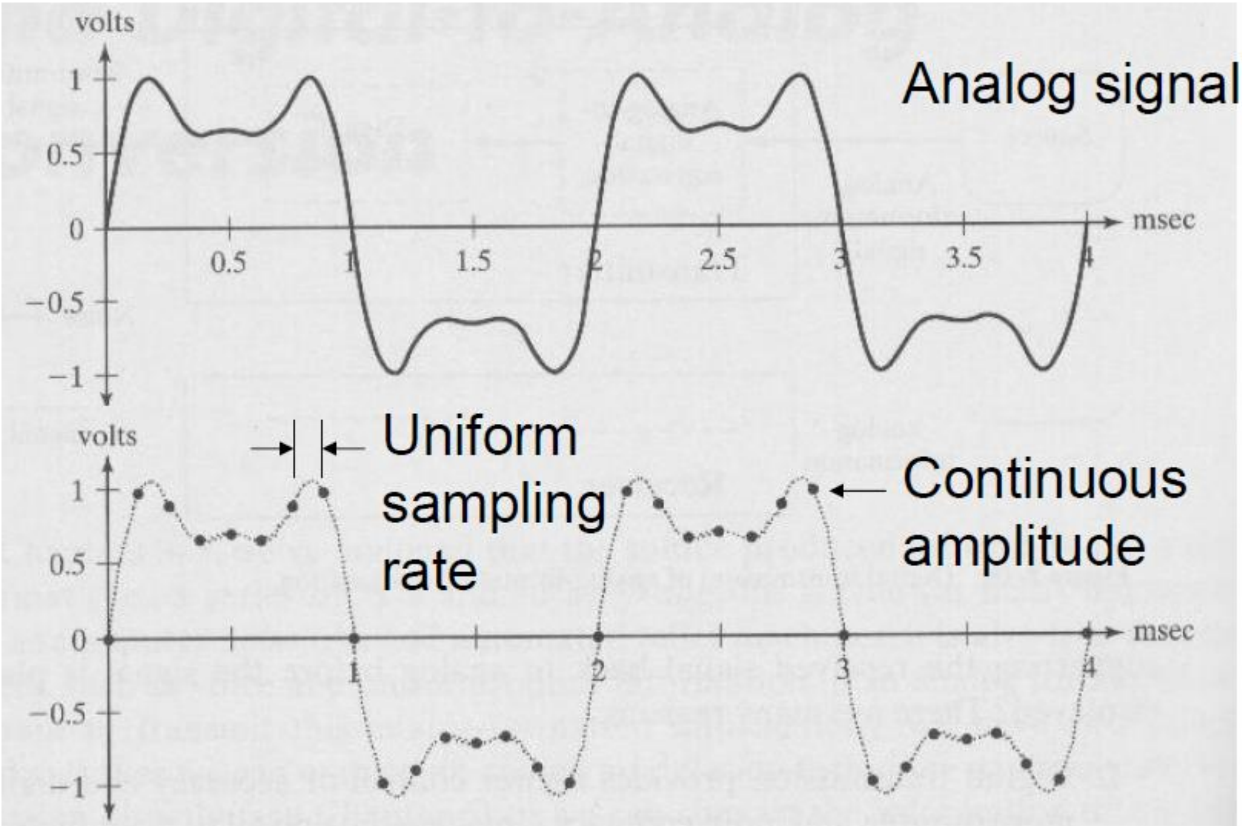


Digital data transmission requires analog-to-digital (ADC) and digital-to-analog (DAC) converters. The ADC process utilizes sampling and quantization of the continuous analog signal.

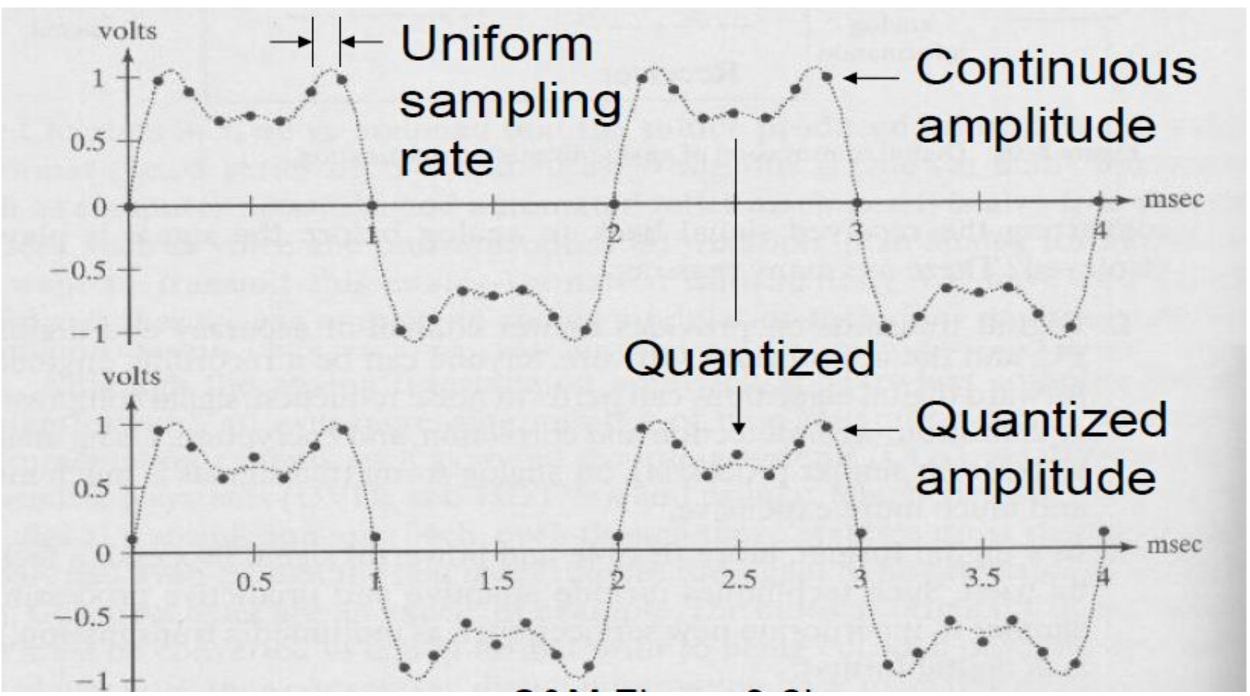
ADC



ADC sampling occurs at a uniform rate (the sampling rate) and has a continuous amplitude.



The continuous amplitude sample is then quantized to n bits or resolution for the full scale input or 2^n levels.



Here n = 4 and there are $2^4 = 16$ levels for a full scale input of 2 V (± 1 V). The step size = 2 V / 16 = 0.125 V and the quantized value is the midpoint of the voltage range.

Table 8-1 Quantization Values and Binary Assignments for the Sampled Signal

Voltage Range of Signal $x(t)$	Rounded-off (Quantized) Value	Binary Assignment for Quantized Value
$-1.000 \le x(t) < -0.875$ v	-0.9375 v	0000
$-0.875 \le x(t) < -0.750$ v	-0.8125 v	0001
$-0.750 \le x(t) < -0.625$ v	-0.6875 v	0010
$-0.625 \le x(t) < -0.500$ v	-0.5625 v	0011
$-0.500 \le x(t) < -0.375$ v	-0.4375 v	0100
$-0.375 \le x(t) < -0.250$ v	-0.3125 v	0101
$-0.250 \le x(t) < -0.125$ v	-0.1875 v	0110
$-0.125 \le x(t) < 0.000$ v	-0.0625 v	0111
$0.000 \le x(t) < 0.125$ v	0.0625 v	1000
$0.125 \le x(t) < 0.250$ v	0.1875 v	1001
$0.250 \le x(t) < 0.375$ v	0.3125 v	1010
$0.375 \le x(t) < 0.500$ v	0.4375 v	1011
$0.500 \le x(t) < 0.625$ v	0.5625 v	1100
$0.625 \le x(t) < 0.750$ v	0.6875 v	1101
$0.750 \le x(t) < 0.875$ v	0.8125 v	1110
$0.875 \le x(t) \le 1.000$ v	0.9375 v	1111

SYSTEM MAINTENANCE

MAINTENANCE TECHNIQUES

There is preventive maintenance, reparative and corrective maintenance. Reparative maintenance techniques are applied to a system that has totally broken down, a nonfunctioning system. Corrective maintenance techniques are applied whenever it is discovered that a system output does not conform to its specifications. The most common maintenance techniques attached to system hardware is preventive maintenance in order to prevent the system from untimely breakdown.

It is good practice to prevent problems before they become difficult to handle. The procedures used for preventing hazards to computer systems and associated peripherals and subsequently rules and safety precautions to adhere to when undertaking repair or upgrade jobs so as to prevent injuries or even death are discussed in this unit.

Preventive maintenance is something that customers often forget or do not believe is necessary. However, nothing could be further from the truth. For example, suppose you have a customer who frequently reschedules their preventive maintenance time. Because you are unable to perform the maintenance, neither of you would be aware that the fan in the power supply had stopped working. As a result, instead of just having to replace the power supply, the processor and some of the memory failed due to the excessive heat. This unscheduled downtime will result in a loss of productivity and a larger bill.

By performing regular maintenance on computer equipment, the life of the components themselves as well as locating potential problems will be extended to save the customer time and money.

BASIC MAINTENANCE

Computers are vital component of both the workplace and the home. What is less usual is the knowledge required in order for the systems to continue to perform at the speed they are capable of. To preserve the performance of a computer, basic computer maintenance needs to be applied. Windows is generally the operating system of choice when it comes to the home PC or laptop and is globally recognized and used. When it comes to using a PC nearly everyone involved will learn what they need to know on the Microsoft Windows operating system.

Unfortunately basic steps to ensure that the computer continues to operate as expected are things which are not taught as commonly as they should be. Simple computer maintenance techniques can be applied so that the PC will continue to run as efficiently as possible. As Windows runs, it leaves behind a trail recording every single activity whilst it is switched on. Each application that is loaded, website that is visited and file created will all generate vast amounts of information

that is stored on the computer's hard drive. Log files, temporary files and cookies from the internet are all stored each second your computer is running and performing duties. If these files are not regularly cleaned the machine's performance will gradually deteriorate.
It is with sensible computer management that the constant use of Windows does not have to be detrimental to performance. End users do not have to be technical experts to take basic steps to ensure their computer or laptop continues to run as quickly as it did on the first day of use. Windows provides its own tools to the user that can be used to keep itself clean. Under the 'Start' > 'all programs' > 'accessories' > 'system tools' menu is a variety of easy to use applications which can assist the end user with basic computer maintenance.

Disk Cleanup and Disk Defragmenter would be the two easiest applications to use, and two of the most effective in home computer maintenance. Old files that are built up and useless to the average home user can be removed with a simple scan and delete operation using Disk Cleanup. Downloaded program files, temporary internet files, recycle bin and general windows temporary files can all be safely removed to free up system resources. The Disk Defragmenter is also another simple but effective computer management tool.
When Windows installs new applications and writes new user files to the disk it generally does this in a linear fashion. However, when a PC is used extensively files are relocated over time, applications are uninstalled and new ones are installed. This change on the hard disk means the linear pattern that windows prefers is broken, with large gaps between the files that it is looking for. This is called fragmentation. By running the application, it aligns all of the files on the hard disk that are movable to reduce the gaps. This means that when a file is clicked to be opened the loading time is no longer than expected.

MAINTENANCE PROCEDURES AND APPROACHES

Manufacturer Suggested Guidelines

Before beginning any preventative maintenance procedures, it is critical that you consult the manufacturer's documentation. Vendors include the information on the proper cleaning materials to use when cleaning or maintaining their components. Failure to follow these guidelines could result in component degradation, complete failure, or avoided warranty.
Never assume that you already know what you can use with what device; instead, take the time to review the documentation. Remember that customer service is not only solving a problem, but also ensuring that you do not generate them.

Liquid Cleaning Compounds

There are several liquid cleaning compounds that are used when you perform preventive maintenance. However, it is important to keep in mind that you should always refer to the manufacturer's documentation prior to using any cleaning compound on a component. This is

because every vendor may use different materials in the component itself. In addition, some vendors require specialized cleaning compounds that can be purchased from them.
Various forms of alcohol are frequently used in cleaning computer components, such as isopropyl alcohol and denatured alcohol. These items are generally used to clean contacts and are applied to special disks used for cleaning floppy drive read/write heads. Mild detergent can be used on the outside of the monitor, the computer case, and on keyboards.

Cleaning Contacts and Connectors

Few people realize, or remember, that contacts and connectors require cleaning. The reason people forget is because connectors are seldom removed from their respective sockets.
Nevertheless, these items do get dirty over time and do require cleaning. Most components can be cleaned with a cotton swab that has been coated with isopropyl alcohol. However, many manufacturers recommend that you use a pencil eraser to clean the contacts on expansion cards. As with most computer components, it is important that you consult the vendor’s documentation to ensure that their product will not have an adverse effect from the materials that you use.

Cleaning Tools

There are several cleaning tools that you should keep in your maintenance kit. One nice item components, you can use a rubber knife to dislodge particles that the vacuum or dust-free cloth could not remove. Never use a metal knife or other metallic object when cleaning the computer or its components, as you can either damage them or cause injury to yourself through electrostatic discharge (ESD). ESD occurs when two charged objects come into contact with each other. The charge is transferred between the two objects until both objects have an equal charge.
Compressed air comes in handy when you are cleaning components, such as a keyboard, or areas that a vacuum cannot get to. Compressed air is distributed in an aerosol-type can and is available at any computer or electronics store. The air is compressed in the can, hence its name, by extreme pressure and dispelled through a nozzle that is similar to the old aerosol cans. You would use compressed air when you need to blow dust, dirt, or other unwanted debris out and away from a component.
Dust-free, lint-free disposable cloths or wipes are also cheaply available and should be used whenever you need to wipe the surface of a component. Normal cloths naturally attract lint and dust, while dust-free or lint-free cloths do not. This helps to ensure that you are not contributing harmful materials to sensitive components.

Floppy Drive and Tape Head Cleaning

If you think the read/write heads on a floppy drive need cleaning, you need to obtain a floppy drive cleaning kit. The kits contain a special cleaning disk that looks like a normal floppy and a small bottle of isopropyl alcohol. The disk has an access hole, similar to a regular disk, on which you place a few drops of the isopropyl alcohol. Then you insert the disk into the drive and the drive will spin up. When the read/write heads attempt to read it, the disk will spin over the heads and clean them. These kits are relatively inexpensive and can be found at most computer and electronics stores.

When you clean a tape drive, you are performing the same function as with a floppy drive. The only difference is that you would use a special tape-cleaning cartridge instead of a cleaning disk with isopropyl alcohol. The cleaning cartridge looks exactly like a normal tape cartridge, although some manufacturers use a different color (usually white or beige) casing around a cleaning cartridge to help differentiate it from a normal cartridge.

Vacuuming

It is important to vacuum the inside of a computer case whenever you open it. Dust and dirt particles get sucked into the case through the air ducts and deposit themselves anywhere they can. These particles can conduct an electrical charge, resulting in possible damage to the delicate electronic components inside. Most offices do not have a vacuum that you will be able to use, as there are usually cleaning crews that carry their own equipment with them. Whenever you have to open a printer for repair, it is important that you vacuum out the interior. Bits of paper and dust have a tendency to accumulate on the inside of a printer at a rate that exceeds the interior of a computer. The reason that this happens is due to the nature of the printer; that is, producing output on paper medium. Small bits of paper can tear off and become lodged inside the printer from clearing a paper jam. The dust produced during the printing process itself also contributes to the mess.

The role of a computer service professional

Matching the rapid pace of change in the industry, the role of the computer professional is constantly changing, too. Not too many years ago, the only tools needed to repair a computer were a screwdriver, needle-nose pliers, the documentation for the computer, a boot disk with a few utilities, and a good MS-DOS reference manual. The screwdriver is still the standard repair tool, but the technician is confronted with a wider array of case types, motherboard designs, processor types, and operating systems, and a wider array of customer needs. Today's computer professional needs to be a technician, scholar, and diplomat rolled into one, as you can see by the table that follows.

Title	Skills
Technician	You are able to troubleshoot and repair hardware and software efficiently and quickly.
Scholar	You have the wisdom and perseverance to seek answers to what you don't

	know and build your base of knowledge. Learning never stops.
Diplomat	You are able to instill in the user (your customer) the confidence that you are in control and can fix things, even when you are encountering problems for the first time. You are able to resolve the problem, even if your customer's (lack of) understanding of the computer might be part of that problem.

WORKSHOP TOOLS, SAFETY RULES AND REGULATIONS

Using the proper tools

To troubleshoot and repair PC systems properly, you need a few basic tools. If you intend to troubleshoot and repair professionally, there are many more specialized tools you will want to purchase. These advance tools allows you to more accurately diagnose problems and make the jobs easier and faster. The basic tools that should be in every troubleshooter's toolbox are:

- Simple hand tools for basic disassembly and reassembly procedures, including a flat blade and Phillips screwdrivers (both medium and small sizes), tweezers, an IC extraction tool, and a part grabber or hemostat.
- Diagnostic software for hardware and testing components in a system.
- A multimeter that allows accurate measurement of voltage and resistance.
- Chemicals such as contact cleaners, components freeze sprays, and compressed air for cleaning the system.
- Foams swabs, or cotton swabs if foam is not available.
- Small wire ties for "dressing" or organizing wire.

The following tools can also aid in effective system repairs:

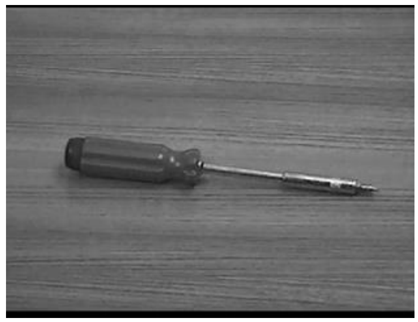
- Memory testing machine, which are used to evaluate the operation of SIMMS (Single Inline Memory Modules), DIP (Dual Inline Pin) chips, and other memory modules.
- Serial and parallel wrap plugs to test serial and parallel ports.
- A network cable scanner, if a network is used.
- A serial breakout box if a lot of the system operates over serial cables, such as UNIX dumb terminals.

In addition, an experienced troubleshooter will probably want to have soldering and desoldering tools to fix bad serial cables. These tools are discussed in more detail in the following section.

Hand Tools

When you work with the PC system, it immediately becomes apparent that the tools require nearly all service operations are very simple and inexpensive. You can carry most of the required tools in a small pouch. Even a- top- of- the – line “master mechanics” sets fit inside a briefcase-size container.
In this section, you learn about the tools required to make a kit that is capable of performing basic, board- level service on the PC systems. One of the best ways to start such a set of tools is a small kit sold especially for servicing PCs.
The following list shows the basic tools that you can find in one of the small PC tool kits:

- 3/16-inch nut driver
- Chip extractor
- Chip inserter
- Tweezers
- Claw- type parts grabber
- T10 and T15 torx drivers
- 1/4- inch nut driver
- Small Philips screwdriver
- Small flat- blade screwdriver
- Medium Philips screwdriver
- Medium flat- blade screwdriver



A Phillips screwdriver

You use nut driver to remove the hexagonal- headed screws that secure the system unit covers, adapter boards, disk drives, power supplies, and speakers in most systems. The nut drivers work much better than conventional screwdrivers.

Some manufacturers have substituted slotted or Phillips-head screws for the more standard hexagonal-head screw; allowing standard screwdrivers to be used for those systems.

Chip- extraction and insertion tool are used to install or remove memory chips (or other smaller chips) without bending any pins on the chip. Usually, you pry out larger chips, such as some microprocessors or ROMs, with the small screwdriver. Larger processors such as the 486, Pentium, or Pentium pro chips require a chip extractor if they are in substandard LIF (Low Insertion Force) socket. These chips have so many pins on them that a large amount of force is required to remove them, despite the fact they call the socket “low insertion force”. If you use a screwdriver on a large physical- size chip like a 486 or Pentium, you risk cracking the case ship and permanently damaging it. The chip extractor tool for removing these chips has a very wide end with times that fit between the pins on the chip to distribute the force evenly along the chips underside. This will minimize the likelihood of breakage. Most of these types of extraction tools must be purchased specially for the chip you are trying to remove.

Fortunately, motherboard designers have seen fit to use mostly ZIF (Zero Insertion Force) sockets on the systems with 486 and larger processors. The ZIF socket has a lever which when rose releases the grip on the pins of the chip, allowing it to be easily lifted out with your fingers.
The tweezers and parts grabber can be used to hold any small screws to jumper blocks that are difficult to hold in your hand. The parts grabber is especially useful when you drop a small part

into the interior of a system; usually, when you can remove the part without completing disassembling the system.

Finally, the Torx driver is a special, star-shaped driver that matches the special screws found in most Compaq system and in many other systems as well.

Although this basic set is useful, you should supplement it with some other hand tools, such as:

- Needle nose pliers
- Vise or clamp
- Small flashlight
- File
- Hemostats
- Wire stripper or wire cutter

Pliers are useful for straightening pins on applying or removing jumpers, crimping cables, or grabbing small component, such as jumpers. The wire cutter or stripper, obviously, is useful for making or repairing cables or wiring. The metric nut drivers can be used in many clone or compatible system, as well as in the IBM PS/2 system, all of which use metric hardware. The tampers–proof Torx drivers can be used to remove Torx screw with the tamper-resistant pin in the center of the screw. A tamperproof Torx driver has a hole drilled in it to allow clearance for the pin.

You can use a vise to install connectors on cables and to crimp cables to the shape you want, as well as to hold parts during delicate operations. In addition to the vise, radio shark sells a nifty “extra hands” device, which has two movable arms with alligator clips at the end. This type of device is very useful for making cables or other delicate operations where an extra set of hands to hold something might be useful.

You can use the file to smooth rough metal edges on cases and chassis, as well as to trim the faceplates on disk drives for a perfect fit. The flashlight can be used to illuminate system interiors, especially when the system is cramped and the room lightning is not good. I consider this tool very essential.

Another consideration for your tool kit is an ESD (electrostatic discharge) protection kit. This kit consists of a wrist strap with a ground wire and a specially conducted mat, also with its own ground wire. Using a kit like this when working on a system will help to ensure that you never accidentally zap any of the components with a static discharge.

The ESD kits, as well as all other tools and much more, are available from variety of tool vendors. Specialized Products Company and Jensen tools are two of the most popular vendors of computer and electronic tool and of service equipment. Their catalogues show an extensive selection of very high- quality tools. With a simple set of hand tools, you will be equipped for nearly every PC repair or institution situation.

PRECAUTIONS

A short circuit can cause physical damage to equipment and personnel. It can cause fire outbreak, component damage, permanent disability, or even death. The ground plug provides a direct connection to ground, giving the electricity an alternate path away from equipment and people.

AC is used for transporting low-cost power to end users. However, a computer's electronic components will not run on AC power, they need a steady stream of DC. The PC's power supply performs several tasks, but the main function is to convert AC into DC. A computer's power supply combines two components to handle this job: a step-down transformer and an AC/DC converter. DC is electrical energy that travels in a single direction within a circuit. (The electrical energy in a thunderstorm is another example, but not very practical to electronic applications.) DC current flows from one pole to another, hence it is said to have polarity. The polarity indicates the direction of the flow of the current and is signified by the + and – signs.

Hardware Vocabulary Terminologies and its equivalency definitions

Vocabulary Term	Definition
555 Timer	A semiconductor device that controls various modes of on/off states in electrical systems. The 555 timer is one of the most widely used types of integrated circuits.
Amplification	The process of increasing the size or strength of a signal.
Amplifier Gain	The ratio of an output signal's amplitude divided by the input signal's amplitude. Amplifier gain measures the degree to which a signal has been strengthened.
Amplify	To increase the size or strength of a signal.
Amplitude	The maximum height or depth of a sine wave.
AND	A logic function where both A and B must be true to trigger C. AND employs a logic that works similar to a series circuit.
Anode	The positive terminal of an electric current flow.
Antimony	An element used to dope silicon to create an N-type semiconductor.
Astable	A mode in which a 555 timer produces a

Vocabulary Term	Definition
	continuous string of on/off pulses at a preset frequency.
Avalanche Breakdown	A process that occurs in a diode when high voltage causes free electrons to travel at high speeds, colliding with other electrons and knocking them out of their orbits. The result is a rapidly increasing amount of free electrons.
Avalanche Diode	A specialized diode that acts as a relief valve for excess voltage. Avalanche diodes are sometimes used in combination with Zener diodes.
Base	The input control terminal of a semiconductor device. Also called a gate.
Bipolar Transistor	A three-layer semiconductor device that can conduct current in either direction. Bipolar transistors provide current gain and voltage amplification in a circuit.
Bistable	A mode in which a 555 timer changes from 1 to 0 or from 0 to 1 when current is applied. The timer will remain in the transitioned state until it receives a new input.
Boron	An element used to dope silicon to create a P-type semiconductor.
Breakdown	A condition that occurs when maximum reverse voltage in a diode is exceeded. Breakdown will cause a diode to fail and pass a large amount of current in the reverse direction.
Breakdown Voltage	A term used to describe the level of AC or DC voltage that results in the failure of a semiconductor device.
Capacitance	The ability to store an electrical charge.
Cathode	The negative terminal of an electric current flow.
CMOS	Complementary metal-oxide semiconductor. The CMOS IC is the most popular type of digital IC because of its low power usage and high immunity to noise.
Collector	The output terminal of a semiconductor device. Also called a drain.

Vocabulary Term	Definition
Common Base	A transistor amplifier in which the circuit shares the base terminal for both inputs and outputs.
Common Collector	A transistor amplifier in which the circuit shares the collector terminal for both inputs and outputs.
Common Emitter	A transistor amplifier in which the circuit shares the emitter terminal for both inputs and outputs.
Complementary Metal-Oxide Semiconductor	CMOS. The most popular type of digital IC because of its low power usage and high immunity to noise.
Demodulate	To change the amplitude or frequency of an analog signal.
Depleted Region	The part of a PN junction in which there are no electrons or holes. The depleted region prevents current from flowing.
DIAC	Diode alternating current switch. A DIAC can conduct current in either direction, but not until breakdown voltage has been exceeded.
Digital IC	Digital ICs process only on/off signals. These devices can be found in microprocessors, memory chips, and microcomputers.
Diode	A semiconductor device that acts as a one-way valve for electrical current.
Diode Alternating Current Switch	DIAC. A semiconductor device that can conduct current in either direction, but not until breakdown voltage has been exceeded.
Diode Array	A group of diodes arranged in a highly organized manner. Diodes are combined into arrays for greater efficiency and reliability.
DIP	Dual-inline package. A popular type of integrated circuit packaging that has two rows of external connecting terminals.
Doping	The process of changing the conductive properties of silicon by adding trace amounts of other elements.
Drain	The output terminal of a semiconductor device. Also called a collector.

Vocabulary Term	Definition
Dual-Inline Package	DIP. A popular type of integrated circuit packaging that has two rows of external connecting terminals.
Electrostatic Discharge	Static electricity that can damage integrated circuits.
Emitter	The terminal in a semiconductor device that is connected to the source supply of voltage. Also called the source.
ESD	Electrostatic discharge. Static electricity that can damage integrated circuits.
FET	Field effect transistor. A FET is a semiconductor device that outputs current in proportion to its input voltage. FETs use a small amount of control current to regulate a larger output current.
Fiber Optic	A filament made of thin, flexible glass or plastic through which light is transmitted. Optical fibers are bundled into groups to form fiber optic cable.
Field Effect Transistor	A semiconductor device that outputs current in proportion to its input voltage. FETs use a small amount of control current to regulate a larger output current.
Flat Pack	A type of integrated circuit packaging that is extremely thin and flat. A quad flat pack or QFP has leads projecting from four sides.
Flip-Flop Circuit	A circuit that changes from 1 to 0 or from 0 to 1 when current is applied.
Forward Bias	A condition in which a PN junction allows current to flow in one direction only.
Full-Wave Rectification	A type of current conversion that uses both parts of the AC sine wave, both positive and negative, to produce a DC output with a single polarity.
Gate	The input control terminal of a semiconductor device. Also called the base.
Gate Current	A small amount of forward-biased current in the middle PN junction of an SCR that controls a larger amount current flowing through the SCR.
Half-Wave Rectification	A type of current conversion that uses only one

Vocabulary Term	Definition
	half of an AC waveform to convert into intermittent DC. This can be the positive half or negative half of an AC wave, depending on how the diode is connected to the circuit.
Holding Current	The minimum current which must pass through a semiconductor device in order for it to remain in the ON state.
Hole	An extra opening in the outer orbital shell of an atom into which an electron can move.
IC	Integrated circuit. A miniaturized electronic circuit that combines a variety of components like transistors, resistors, capacitors, and diodes all into one incredibly small piece.
IEC	The International Electrotechnical Commission. An international organization that prepares and publishes all standards for electrical, electronic, and related technologies.
Impedance	Resistance to current flow in an AC circuit.
Integrated Circuit	A miniaturized electronic circuit that combines a variety of components like transistors, resistors, capacitors, and diodes all into one incredibly small piece.
Intrinsic Semiconductor	A type of semiconductor that has a low level of electron movement at any temperature above 0°C (32°F). Silicon is an intrinsic semiconductor.
JFET	Junction field effect transistor. A solid state device that can be used as an electronically controlled switch.
Leakage Current	The amount of current required to keep a device active when it is not operating. Also known as load current and residual current.
LED	Light emitting diode. A semiconductor device that emits a narrow spectrum of light in a forward direction.
Logic Gate	A programming function that processes true and false signals.
Low Current Drop	A condition that occurs when the current falls

Vocabulary Term	Definition
	below the minimum value required to operate a semiconductor device.
Monostable	A mode in which a 555 timer produces a single pulse for a preset amount of time in response to an input.
MOSFET	Metal oxide semiconductor field effect transistor. A type of transistor that is controlled by voltage rather than current.
Multivibrator	A device that controls various modes of on/off states in electrical systems. A multivibrator behaves like a pendulum moving back and forth, or a ball bouncing up and down.
NAND	A logic function where both A and B must be false to trigger the output. If one input is present, the output will not be turned on.
NOR	A logic function where the output will be triggered if neither A nor B are present.
NOT	A logic function where the output will not be triggered if a specified input is present.
N-Type Semiconductor	A type of semiconductor that moves current by creating extra electrons that are easily excited into movement.
One-Shot	A series of timer instructions that creates a delay for only one scan.
OR	A logic function where the output is triggered if either A or B are true. OR works similar to a parallel circuit.
Oscillator	A device that produces a continuous string of on/off pulses at a preset frequency.
PCB	Printed circuit board. Semiconductor devices are often mounted on PCBs because the electrical paths on a PCB are perfect for the needs of most semiconductors.
PIN diode	A specialized diode that has a layer of intrinsic semiconductor material between the P and N materials. PIN diodes are also used as photodetectors.

Vocabulary Term	Definition
PN junction	The area in a semiconductor where P-type and N-type materials are located next to each other. A PN junction allows current to flow in one direction only.
Polarity	Having a positive or negative charge.
Printed Circuit Board	PCB. Semiconductor devices are often mounted on PCBs because the electrical paths on a PCB are perfect for the needs of most semiconductors.
P-Type Semiconductor	A type of semiconductor that moves current by creating holes for the movement of valence electrons.
Pulsating DC	A type of DC current in which half of the original AC signal wave is blocked, so that the resulting DC signal rises from zero to a maximum, and then returns to zero.
QFP	Quad flat pack. A type of integrated circuit packaging that is extremely thin and flat. A QFP has leads projecting from four sides.
Quad Flat Pack	A type of integrated circuit packaging that is extremely thin and flat. A quad flat pack or QFP has leads projecting from four sides.
Rectifier	A device that converts AC into DC.
Rectify	To convert into direct current.
Reverse Bias	A condition in which a PN junction does not allow current to flow.
Reverse Breakdown Voltage	The amount of reverse bias that will cause a diode to break down and conduct in reverse.
Reverse Firing	A method of turning off an SCR by applying negative voltage to the gate. Reverse firing requires a high amount of gate current.
Sandwich	A three-layer NPN or PNP arrangement of semiconductor material.
Schottky diode	A specialized diode used in the electronics industry and in radio frequency applications because of its fast switching speed and high frequency capability.

Vocabulary Term	Definition
SCR	A solid state switching device that turns current on and off. SCRs use a small amount of current to switch hundreds of amps without being damaged.
Semiconductor	A solid state device that is less conductive than a conductor, but more conductive than an insulator. The most common semiconductor material is silicon.
Signaling	The transmission of electrical signals.
Silicon	An element from which almost all semiconductors are made.
Silicon Controlled Rectifier	A solid state switching device that turns current on and off. SCRs use a small amount of current to switch hundreds of amps without being damaged.
Silicon Dioxide	A glass-like material that is naturally an excellent insulator. Silicon dioxide is used as the gate insulating material in a MOSFET.
Single-Phase Rectifier	A semiconductor device that converts single-phase AC into DC. In a single-phase rectifier, the sine waves produced by the AC power supply reach their peak at 90° simultaneously.
Solid State	Purely electronic with no moving parts.
Source	The terminal in a semiconductor device that is connected to the source supply of voltage. Also called the emitter.
Substrate	The surface or medium that serves as a base for other materials or components.
Switching	The process of making or breaking an electric circuit, or selecting between multiple circuits.
Terminal	A connecting point in a circuit to which a wire can be attached to connect a component.
Three-Phase Electric Power	A type of AC power that is generated with three waveforms that are not in phase with each other.
Three-Phase Rectifier	A type of rectifier that uses a three-phase AC electric power source to produce a relatively smooth DC. Three-phase rectifiers provide much smoother DC output voltage than single-phase rectifiers.

Vocabulary Term	Definition
Thyristor	A three-terminal semiconductor switching device. The two main types of thyristors are SCRs and TRIACs.
TO package	Transistor-outline package. A type of integrated circuit packaging that is cylindrical in shape and looks like a little can.
TO-5 can	A commonly used TO package.
Transistor	A solid state device that can be used either as an amplifier or as a switch. There are two basic types of transistors: bipolar and field effect.
Transistor-Outline	A type of integrated circuit packaging that is cylindrical in shape and looks like a little can.
TRIAC	Triode alternating current switch. A gated switching device that will conduct current in either direction.
Triode Alternating Current Switch	TRIAC. A gated switching device that will conduct current in either direction.
UJT	Unijunction transistor. A voltage-controlled switch that controls current.
Unijunction Transistor	UJT. A voltage-controlled switch that controls current.
Valence Electron	An electron found in the outermost orbit of an atom.
Varactor Diode	A specialized diode that changes its level of capacitance depending on the level of reverse bias applied to the diode. Also known as a varicap diode.
Varicap Diode	A specialized diode that changes its level of capacitance depending on the level of reverse bias applied to the diode. Also known as a varactor diode.
Zener Diode	A specialized diode that is used as a voltage regulator. Zener diodes maintain voltage at a constant, predefined value when they are subjected to reverse voltage.