EIILM UNIVERSITY
SIKKIM

COMPUTER PLATFORMS

Subject: COMPUTER PLATFORMS **Credits: 4**

SYLLABUS

Basic Computer Concepts, Architecture of Computer System, Basic Parts of computer system, Computers then and now, Classification of Computers, Performance of computer system, Bus architecture, Primary Storage, Secondary storage

Input devices, Output devices- Monitors, Printers etc., Computer selection, OS and its function, Command Interpretation, Process management –I, Process management II, Types of OS –I, Types of OS –II

Computer operations, Security & deadlock, Network administration, Upgrading Opportunities, Upgrading & Installation, Data communication system, Transmission Media, Computer networks

Network topologies, OSI Reference Model, Data Communication Hardware, Internet –I, Security & Electronic Infection, Hacking & Encryption, Back Up & Recovery

Suggested Readings:

1. P. K.Sinha; Computer Fundamentals, BPB Publications

2. Leon, Alexis; Leon, Mathews ; Fundamentals Of Computer Science And Communication Engineering, BPB Publications

3. Abraham Silberschatz, Peter Baer Galvin; Operating systems concepts, , John Wiley & Sons Publications

4. Operating System By Tann & Bomm

COMPUTER PLATFORMS

CONTENT

COMPUTER PLATFORMS

CONTENT

LESSON 1
BASIC COMPUTER CONCEPTS

Outline

- System Concepts
- Basic Operations of Computer System
- How computer works?
- The input-process-output model
- Components of a computer system

What is System ?

What are Basic Components Of A System

The System Concept

A system is a group of integrated parts that have the common purpose of achieving some objective(s). So, the following three characteristics are key to a system:

1. A system has more than one element.
2. All the elements of a system are logically related.
3. All the element of a system are controlled in such a way that the system goal is achieved.

Basic Three Elements Of Any System Is

- Input
- Output
- Processor

Processor is the actual part that converts input into output.

Lets say for music system

What will be the input?

What will be the output?

What will be the processor?

You might have observed by now that we have been referring to a computer as a system (computer system).

Since a computer is made up of integrated components (input and output devices, storage, CPU) that work together to perform the steps called fro in the program being executed, it is a system. The input or output units cannot function until they receive signals from the CPU. Similarly, the storage unit or the CPU alone is of no use. So the usefulness of each unit depends on other unit and can be realized only when all unit are put together (integrated) to from a system

How Will You Define a Computer

A computer is an electronic device capable of executing instructions, developed based on algorithms stored in its memory, to process data fed to it and produced the required results faster than human beings.

OR

An electronic instrument, which can solve problems by accepting data, performing various operations with the data, and supplying results. May be a personal computer or terminal, may be connected to a network, may provide access to the Internet.

You mean that you've got this far and you still don't know what a computer is? True, sometimes it can be hard to tell. A computer is the part of the system that contains a processing unit that does the thinking. Sometimes people refer to the computer as the entire package; the processing unit, keyboard, screen, disk drives the works. You may need to know if you have a computer on your desk or if you have a terminal connected to a computer somewhere else.

What are Basic Operations of Computer System?

The Four Operations of a Computer System

Input

Home computers are microcomputers. Input is supplied to the microcomputer with the use of a keyboard, a mouse, or another input device. These input devices may be called peripheral devices.

Processing

Processing is done inside the computer in an area called the central processing unit (CPU). Processing is the conversion of input to output.

Storage

Storage refers to holding information somewhere.
RAM, Random Access Memory, is short-term memory. It is volatile memory because the memory is automatically "erased" when the power is turned off or interrupted. The RAM memory is located inside the computer case on the motherboard. A motherboard is not the keyboard. The keyboard is what you type with. A motherboard holds RAM memory, electronic circuits and other computer parts including the central processing unit. ROM, Read-Only-Memory, is not volatile meaning the memory is still there when power is interrupted or turned off. When the computer is turned back on again, ROM memory is still in storage on the internal hard disk.

Output

Output is the result of a computer process. Output may be viewed on a monitor screen, heard through speakers, printed on

printers, and so forth. Output devices may be considered hardware and are also considered to be peripheral devices.

Controlling

Directing the manner and sequence in which all of the above operations are performed

How Computer Works?

A computer collects processes, stores and outputs information.

What Are Basic Operations Of Computer System?

Input

An input device lets you communicate with a computer. You can use input devices to enter information and issue commands. A keyboard, mouse and joystick are input devices.

Process

The Central Processing Unit (CPU) is the main chip in a computer. The CPU processes instructions, performs calculations and manages the flow of information through a computer system. The CPU communicates with input, output and storage devices to perform tasks.

Store

A storage device that holds information. The computer uses information stored on these devices to perform tasks. The hard drive, the tape drive, the floppy disk and the CD Rom drive are storage devices.

Output

An output device lets a computer communicate with you. These devices display information on a screen, create printed copies or generate sound. Monitor, printers and speakers are output devices.

A typical computer system consists of several parts.

Computer Case

A Computer case contains all the major components of a computer system.

Monitor

A monitor is a device that displays text and images generated by the computer.

Printer

A printer is a device that produces a paper copy of documents you create on the computer.

What is Input-Process-Output Model?

- Regardless of the type of work to be performed by a computer system, the work can be characterized by an input-process-output model.
- That is, a program receives input from an input device, performs some processing upon it, and produces some form of output.
- The operations performed during processing are limited to calculations, comparisons, merging, sorting, simple decision making. Many of these operations are carried out repetitively, many times per second.

What Are The Components of Computer System?

Note:

Components Of Computer System

There are two parts one that you can touch and another one that you cannot touch

Physical parts – hardware

Non-physical parts – software

A computer typically consists of

Hardware

- CPU (Central Processing Unit)
- Memory
- I/O (Input/Output) Devices

Software

- Data
- Computer hardware provides the physical mechanisms to store, manipulate, and input/output computer data. Examples of computer hardware include CPU, memory, and I/O devices.
- Computer software provides instructions to tell the hardware what tasks to perform. Examples of computer software include system software and application software.
- Computer data represents real-life data in a form that can be manipulated by computer hardware and software.

Following question should be answered: -

1. What are the five basic operations are performed by the computer system?

2. What is a system? Why is a computer often referred to as a computer system?

3. Differentiate between data and information.

LESSON 2
ARCHITECTURE OF COMPUTER SYSTEM

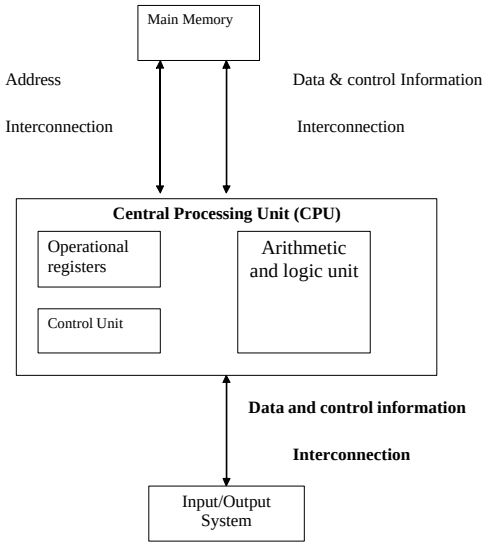
Outline

- VON Neumann Architecture
- Central processing unit
  - ALU (arithmetic and logical unit)
  - CU (Control Unit)
- Input unit
- Output unit
- Storage unit

Von Neumann Architecture

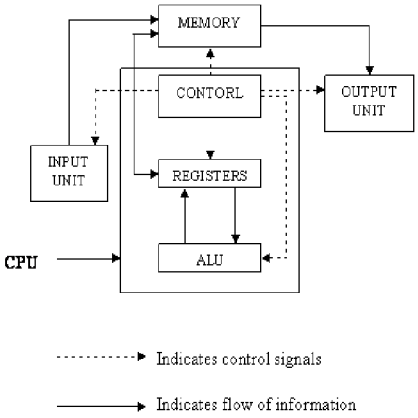
Key feature of von Neumann machine are

- The hardware of the von Neumann machine consist of
- A CPU which includes a ALU and CU
- A main memory system
- An input/output system
- The Von Neumann machine used stored program concept, i.e. the program and data are stored in the same memory unit. The computers prior to this idea used to store programs and data on separate memories. Entering and modifying these programs were very difficult as they were entered manually by setting switches and plugging and unplugging.
- Each location of the main memory of von Newmann machine can be addressed independently.
- Execution of instructions in Von Newmann machine is carried out in a sequential fashion (unless explicitly altered by the program itself) from one instruction to the next.



Basically we divide the computer system into four units one for inputting, storing, outputting and last for processing. here we have

a) Central processing unit
b) Input unit
c) Output unit
d) Storage unit



A) Central Processing Unit (CPU)

It is the main part of a computer system like the heart of a human being. It interprets the instructions in the program and executes one by one. It consists of three major units.

- Control Unit: It controls and directs the transfer of program instructions and data between various units.
- Arithmetic and Logic Unit (ALU): Arithmetic operations like (+,-,*,^,/), logical operations like (AND, OR, NOT) and relational operations like (<,>,<=,>=) are being carried out in this Unit.

Registers

They are used to store instructions and data for further use.

A processing unit in a computer interprets instructions in a program and carries them out. An instruction in general, consists of a part, which specifies the them out. An instruction in general, consist of a part, which specifies the operation to be performed and other parts, which specify the address of operand. In a processor, a strong of bits is used to code operations and another string of n operations in binary so that $2^x = n$. For example to code 16 operations we need 4 bits, since $2^4=16$.

- The CPU (or processor) is the heart of the computer. It consists of three main parts:
  - a Control Unit
  - an ALU (Arithmetic and Logic Unit)
  - and registers
- Instruction Register
- General-purpose Registers
- Program Counter

I/O (Input/Output) Interface

- The CPU performs the processing of data.
- Data and programs are stored in memory, and moved to and from the CPU as required.
- Signals representing data and instructions travel between system components along electronic pathways (sets of wires) called buses.

The control unit and the arithmetic logic unit of a computer system are jointly known as the Central Processing Unit (CPU). the CPU is the brain of any computer system. In a human body, the brain takes all major decisions and the other parts of the body functions as directed by the brain. Similarly, in a computer system, all major calculations and comparisons are made inside the CPU and the CPU is also responsible for activating and controlling the operation of other unit of a computer system.

Central processing unit is divided into two parts

i) Arithmetic logic unit-

The arithmetic logic unit (ALU) of a computer system is the place where the actual execution of the instructions takes place during the processing operation. To be more precise, all calculations are performed and all comparisons (decisions) are made in the ALU. The data instruction, stored in the primary storage prior to processing, are transferred as and when needed to the ALU where processing takes place. No processing is done in the primary storage unit. Intermediate results generated in the ALU are temporarily transferred back to the primary storage unit needed a later time. Data may thus move from primary storage to ALU and back again to storage many times before the processing is over. After the completion of processing, the final results, which are stored in the storage unit, are released to an output device.

The type and number of arithmetic and logic operations that a computer can perform is determined by the engineering design of the ALU. However, almost all ALU's are designed to perform the four basic arithmetic operations - add, subtract, multiply, divide and logic operations or comparisons such as less than, equal to, or greater than.

II) Control Unit-

How does the input device know that if is time for it to feed data into the storage unit? How does the ALU know what should be done with the data once they are received? And how is it that only the final results are sent to the output device and not the intermediate results? All this is possible because of the control unit of the computer system. By selecting, interpreting, and seeing to the execution of the program instructions, the control unit is able to maintain order and direct the operation of the entire system. Although, it odes not perform any actual processing on the data, the control unit acts as a central nervous system for the other components of the computer. It manages and coordinates the entire computer system. It obtains instructions from the program stored in main memory, interpret the instruction, and issue signals that cause other units of the system to execute them.

B) Central Processing Unit (CPU)

A flip-flop made of electronic semiconductor devices is used to fabricate a memory cell. These memory cells organized as a Random Access Memory (RAM). Each cell has a capability to store one bit of information. A main memory or store of a computer is organized using a large number of cells. Each cell stores a binary digit. A memory cell, which does not loose the bit stored in it when no power is supplied to the cell, is known as a non-volatile cell.

A word is a group of bits, which are stored and retrieved as a unit. A memory system is organized to store a number of words. A Byte consists of 8 bits. A word may store one or more bytes. The storage capacity of a memory is the number of bytes it can store. The address of the location from where a word is to be retrieved or to be stored is entered in a Memory Address Register (MAR). The data retrieved from memory or to be stored in memory are placed in a Memory Data Register (MDR). The time taken to write a word is known as the Write time. The time to retrieve information is called the Access time of the memory.

The time taken to access a word in a memory is independent of the address of the word and hence it is know as a Random Access Memory (RAM). The main memory used to store programs and data in a computer is a RAM. A RAM may be fabricated with permanently stored information, which cannot be erased. Such a memory is called a Read Only Memory (ROM). For more specialized uses, a user can store his won special functions or programs in a ROM. Such ROM's are called Programmable ROM (PROM). A serial access memory is organized by arranging memory cells in a linear sequence. Information is retrieved or stored in such a memory by using a read/write head. Data is presented serially for writing and is retrieved serially during read.

The purpose of memory is data storage.

There are two major types of memory:

· Primary Memory
  - Holds data and instructions during processing;
  - Has relatively limited capacity; and
  - is volatile-Secondary memory
  - Provides permanent long-term storage;
  - Has relatively high capacity; and
  - Is non-volatile.
- An example of primary memory is RAM (Random Access Memory) whereas an example of secondary memory is a HDD (Hard Disk Drive) or DVD-ROM (Digital Video Disc-Read Only Memory).
- Digital computers deal with data in binary form. Binary data is represented using just two digits: 0 and 1.
- Letters and other symbols are assigned unique binary codes.
- Primary memory consists of a set of locations defined by sequentially numbered addresses. Each location contains a binary number that can be interpreted as data or an instruction.

- Memory locations are called words.
- Words are eight bits (one byte) in size or a multiple of eight. A bit is a binary digit (0 or 1).

Common word sizes are 16, 32, and 64 bits.

- Memory is commonly measured in multiples of bits and bytes.
- 1 bit = 1 binary digit (o or 1).
- 1 byte = 8 bits.
- 1 KB (kilobyte) = 1024 bytes = 210
- 1 MB (megabyte) = 1024 KB = 220
- 1 GB (gigabyte) = 1024 MB = 230
- 1 TB (terabyte) = 1024 GB = 240

- Secondary storage stores binary data in several different forms.
- Hard and floppy disks, and tapes store data as magnetised spots.
- CDs (Compact Discs) and DVDs (Digital Video Discs) store data as pits detectable by laser light.

C) Input Unit

Computers need to receive data and instructions in order to solve any problem. Therefore we need to put the data and instructions into the computer. The input unit consists of one or more input device. The keyboard and mouse of a computer are the most commonly used input devices. Data and instructions must enter the computer system before any computation can be performed on the supplied data. This task is performed by the input init that links the external environment with the computer system. Data and instructions enter input units in forms that depend upon the particular device use. For example, data is entered from a keyboard in a manner similar to typing, and this differs from the way in which data is entered though a card reader which is another type of input device. However, regardless of the from in which is another type of input device. However, regardless of he from in which they receive their inputs, all input devices must provide a computer with data that are transformed into the binary codes that the primary memory of a computer is designed a accept. Units called input interfaces accomplish this transformation. Input interfaces are designed to match the unique physical or electrical characteristics of input device to the requirements of the computer system.

In short, the following functions are performed by an input unit :

1. It accept (or reads) the list of instructions and data from the outside world.
2. It converts these instructions and data in computer acceptable form.
3. It supplies the converted instructions and data to the computer system for further processing.

D) Output Unit

The job of an output unit is just the reverse of that of an input unit. It supplies information and results of computation to the outside world. Thus it links the compute with the external environment. As computers work with binary code, the results produced are also in the binary form. Hence, before supplying the result to the outside world, it must be converted to human acceptable (readable0 form. This task is accomplished by units called output interfaces are designed to match the unique physical or electrical characteristics of output devices (terminals, printers, etc.) to the requirements of the external environment.

In short, the following functions are performed by an output unit :

1. It accepts the results produced by the computer, which are in code from and hence cannot be easily understood by us.
2. It converts these coded results to human acceptable (readable) form.
3. It supplies the converted results to the outside world.

Following questions should be answered: -

1. What is an input interface? How it is differ from output interface?

2. List out the logical steps taken by a computer system along with the roles of main units of the computer system in each step while transforming an input data to useful information for presentation to the user.

LESSON 3
BASIC PARTS OF COMPUTER SYSTEM

Outline

- Software
  - System software
  - Application software
- Hardware
- Computer languages
  - Low level languages
  - High level languages

What are Two Basic Parts Of Computer System ?

Sofware and Hardware are Two Basic Parts of Computer System

- Software translates users needs into computer instructions. There are two basic types of software:
  - System software
- The most common type of system software is the operating system. System software also includes other specialised software such as compilers and debuggers.
  - Application software
- Is a set of programs written to solve a specific problem? Examples of application software include word processors, spreadsheets and Internet browsers.

A) Software

Your computer system cannot help you type a letter unless you have some type of software (program) to do this. A "word processing" program handles most typing. A software program for your computer can be purchased at stores like Office Depot, Staples, and computer stores. Check your local telephone directory.

Software is a set of electronic instructions that tell a computer what to do. You cannot see or touch software, but you can see and touch the packaging the software comes in.

i) Application Software

Application software lets you accomplish specific tasks. Popular application software includes Microsoft Office XP.

It is the set of programs necessary to carry out operations for a specified application.

Example

Programs

- To solve a set of equations
- To process examination results
- To prepare a Pay-Bill for an organization
- To prepare Electricity-Bill for each month.

ii) System Software

Operating system software controls the overall activity of a computer. Highly Advanced computers use Windows 2000 as an operating system software

These are general program written for the system, which provide the environment to facilitate writing of Application software. Some of the system programs are given below:

Compiler: It is a translator system program used to translate a High-level language program into a Machine language program.

Assembler: It is another translator program used to translate an Assembly language program into a Machine language program.

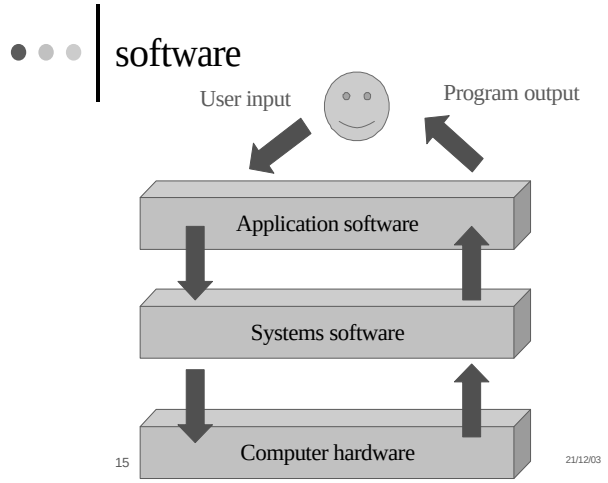
Interpreter: It is also a translator system program used to translate a High level language program into a Machine language program, but it translates and executes line by line.

Loader: It is a system program used to store the machine language program into the memory of the computer.

B) Hardware

Hardware is any part of a computer system you can see or touch.

i) Peripheral



A peripheral is any piece of hardware attached to a computer, such as a printer.

Computer Languages

What do you think whatever you write it is understood by your computer system? The answer is No. The computer understands only the language of 0's and 1's.

The computer languages are divided into three categories

a) Low level
b) Middle level
c) High level

A) Low level languages

i) Machine Language

The computers can execute a program written using binary digits only. This type of programs is called machine language programs. Since these programs use only '0's and '1's it will be very difficult for developing programs for complex problem solving. Also it will be very difficult for a person to understand a machine language program written by another person. At present, computer users do not write programs using machine language. Also these programs written for execution in one computer cannot be used on another type of computer. i.e., the programs were machine dependent.

ii) Assembly Language

In assembly language mnemonic codes are used to develop program for problem solving. The program given below shows assembly language program to add two numbers A & B.

Program code	Description
READ A	It reads the value of A.
ADD B	The value of B is added with A.
STORE C	The result is store in C.
PRINT C	The result in 'C' is printed.
HALT	Stop execution.

Assembly language is designed mainly to replace each machine code with and understandable mnemonic code. To execute an assembly language program it should first be translates into an equivalent machine language program. Writing and understanding programs in assembly language is easier than that of machine language. The programs written in assembly language are also machine dependent.

B) Middle level languages

Middle level languages are those languages, which have features of low-level languages and high-level languages. Example – C language

C) High Level Languages

High-level language is developed to allow application programs, which are machine independent. High-level language permits the user to use understandable codes using the language structure. In order to execute a high-level language program, it should be translated into a machine language either using a compiler or interpreter.

The high level languages commonly used are FORTRAN (Formula Translation), BASIC (Beginner's All-purpose Symbolic Instruction Code), COBOL (Common Business Oriented Language). Recently developed programming language such as Visual FoxPro, Visual Basic (VB), and Visual C++ (VC++) are more popular among the software developers. The following program written in BASIC language is to add two given numbers.

Program Code	Description
10 INPUT A,B	To read the value of A&B
20 LET C=A+B	A&B are added and result is stored in C
30 PRINT C	Print the value of C
40 END	Stop execution

Following Questions Should be Answered:

1. Hardware and software of a computer system are like two sides of a coin. Discuss

2. Define the terms hardware and software.

3. How many types of software are there? Give examples of each.

4. List out some of key functions performed by the system software of a computer system.

LESSON 4
COMPUTERS: THEN AND NOW

Outline

- History and evolution of computing devices
- Generations of computers
- Evolution of the Intel x86 processor
- Performance enhancements
- Classification of computers
- How computer works?

Which Was The First Calculating Device?

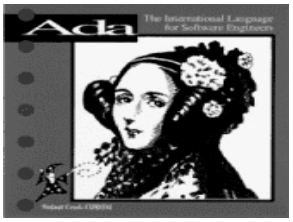
When the human race started doing some trade, it felt a need for a calculating device. The first calculating device, which was used 2000 years ago was called abacus and the improvements in the calculating device in that age were slow. The next change came after about 1600 years. Following this, the changes were frequent and the mechanical desk calculator was developed around 1800 A.D.

In 1833, Prof. Charles Babbage, the father of the computer, developed a machine called analytical engine, which was the vase for the modern digital computer.

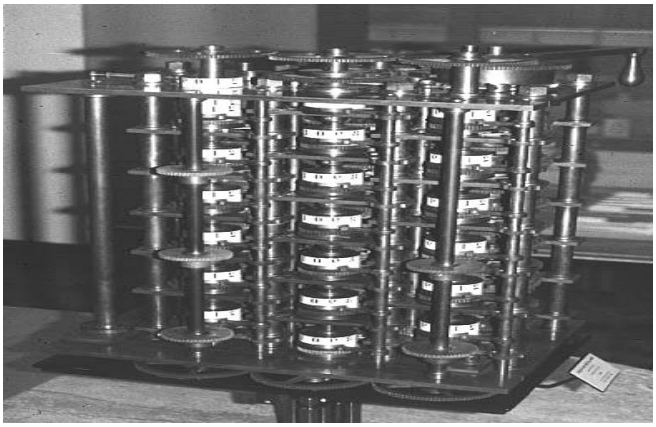
The electronic digital computers, which were introduced in 1950's, were using vacuum tubes. Following this, the development in the electronic components helped in the developed of digital computers also. The second-generation computers used transistors.

The introduction of integrated circuits (ICs), also know as chips opened the door for the developed of third generation computers. A very large number of circuit elements (transistors, diodes, resistors, etc.,) could be integrated into a very small (less than 5mm square) surface of silicon and hence the name IC. The third generation computers used small-scale integrated circuits (SSI), which contain about 10-20 components. When large-scale integrated circuits (LSI) (around 30,000 components) were developed, the fourth generation computers were produced.

- Mechanical and electromechanical devices preceded the electronic computer.
- Charles Babbage (1830's): the analytical engine, and Augusta Ada Byron, the first programmer.



- The analytical engine of Charles Babbage.
- This machine was actually built to Babb age's specifications in the 1990's and it worked.
- The technology at the time of Babbage could not produce mechanical parts that were sufficiently exact for it to work.



Electro Mechanical Devices

- Hollerith (1890's): electrical tabulating machine, and foundation of IBM (1924).
- Aiken's Mark 1. (1944) Based on Babbage's original design - built at IBM labs, electro-mechanical, weighed 5 tons. Admiral Grace Hopper worked as programmer on this computer, and coined the term 'bug' for a computer fault.

Electronic Devices

- To increase speed and reliability, the mechanical components of these early devices were replaced by electronic equivalents.
- These fully electronic devices became known as the first generation of computers.
- Successive generations of electronic computers have greater speed, smaller size, larger memory.

1st generation (1940-50's)

What Are Four Generations Of Computer System ?

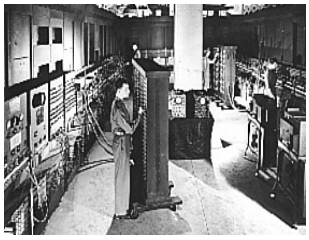
Generations of Computers

A) Ist Generation Computers

- Atanasoff & Berry (1939) first electronic digital computer.
- Turing (1943) Colossus machine built to decode encrypted messages. (*Turing Test)
- ENIAC (1946) first general purpose computer. (Decimal not binary machine)
- UNIVAC (1951) first commercial computer - commissioned for 1950 US census.

ENIC

- The world's first electronic digital computer was developed by Army Ordnance to compute World War II ballistic firing tables.
- When it was finished, the ENIAC filled an entire room, weighed thirty tons, and consumed two hundred kilowatts of power. It generated so much heat that it had to be placed in one of the few rooms at the University with a forced air-cooling system.



- Early computers (1st generation) had only one function. To change
- The function, you had to 'rewire'. In 1945 John von Neumann proposed the 'stored program concept', with memory to store both data and instructions.
- Loading a new program into memory then allows a new function for the computer.
- In Von Neumann's basic design, the program is executed in a sequential manner. Whilst still the model for most conventional computers, parallel processing has recently challenged this aspect of the traditional computing device.

B) 2nd generation (1959-1964)

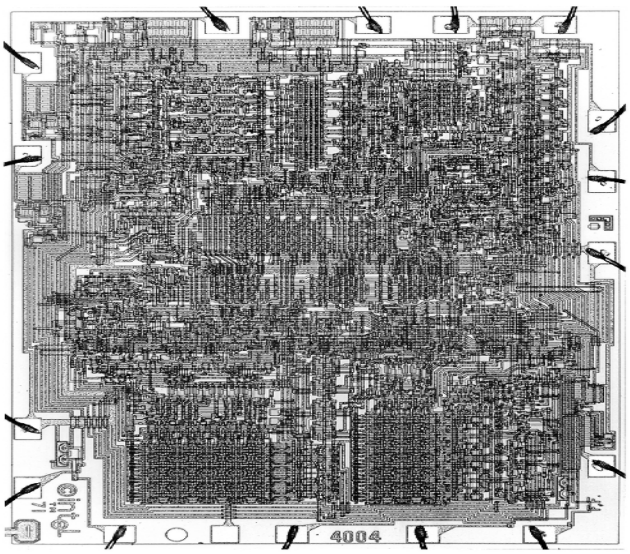
- Transistors replaced vacuum tubes. Far smaller, faster, more reliable, and required less energy.
- Development of symbolic languages to simplify programming – e.g. ADD instead of 00000101, and later high level languages such as Fortran and Cobol.
- Provision of system software (OS).
- Computers used by military, government and big business.
- Problems with transistor-based machines: 100's of 1000's of transistors required for powerful machines - expensive and cumbersome.
- This problem was solved by the invention of the integrated circuit.

C) 3rd Generation (1965-1970)

- Complete electronic circuit on single silicon chip, smaller than single transistor.
- 1000's of components on one chip (SSI to MSI).
- cheap, fast and reliable.
- used by small business.

D) 4th Generation (1971-now)

- Microprocessor contains ALL elements of CPU on one chip (Intel's4004, 1971).
- Progressively greater integration (from LSI to VLSI) - 100,000's to millions of components per chip.
- Inc. complexity of instruction sets, inc. no of bits handled at once, inc. amount of memory addressed.
- Advent of personal computers and other microprocessor controlled devices.
- Intel's 4004 microprocessor.



Evolution of Intel Microprocessors

Name:	8008	80386	80486	Pentium	Pentium4
Year:	1972	1985	1989	1993	2000
I. set:	66	154	235	>300	>400
Data bus:	8	32	32	64	64
Clock speed:	<1mh	33mh	50mh	133mh	1.3GH

Following Question Should be Answered:

1. Who is known as the father of modern digital computers and why?

2. Why modern digital computers often referred to as stored program digital computers?

3. What is meant by generation in computer terminology? How many generations are there till now?

4. Write short note on fifth generation computers (yet to come).

5. List out some various computers generations along with their basic characteristics.

6. Write short note on evolution of Intel computers.

LESSON 5
CLASSIFICATION OF COMPUTER SYSTEM

Outline

- Microcomputers
- Minicomputers
- Mainframes
- Supercomputers
- Macintosh
- Working of computers

Microcomputers (Pc's)

Classification Of Computers

- Desktop, laptop & palmtop usage
- Single user
- Many uses
- $2,000

Minicomputers

- Business & industry usage
- Multiple users
- Specified tasks
- $20,000 - $250,000

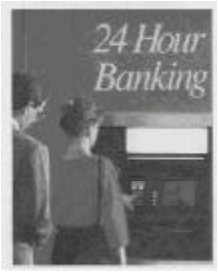
Mainframes

- Business & government usage
- Multiple user focus
- Centralized control
- Very powerful
- Multiple processors
- $400,000

Supercomputers:

- Scientific & business usage
- Multiple user focus
- Handles huge amounts of data
- Very powerful
- $500,000 - $35,000,000

One special Purpose Digital Computer, that you might use and not even think of as being a computer is the ATM machine!



The ATM is a special purpose computer, in that while it is good for getting or giving money to, it can't be used to register for classes!

The computers we will use in this course are general purpose computers. All we need to do to change what they can do is just load new software into them, such as software for doing one's income tax. These computers are also called desk top microcomputers because they fix on a desk or table top, and they have a single micro process chip that defines their capabilities.

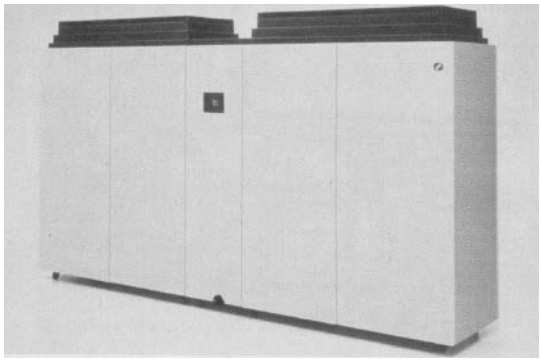
A large version of the computer is the minicomputer. While a terminal (keyboard and display), for a minicomputer could set on top of a desk, the works of a minicomputer would fill up all of the space in the desk or it could be a rack of equipment next to the desk.

Minicomputers usually support many uses.

The bigger commercial computers are mainframe computers. A small mainframe computer might be a rack of equipment about eight feet tall and ten to twenty feet long. It is keep in a special room, with a raised room, and heavy air conditioning.

Mainframe computers are used when you access your bank account from a automated teller. The bank's mainframes handle all the transactions.

A mainframe is a computer that can process and store large amounts of information and support many users at the same time. A terminal consisting of a keyboard and monitor, is used to input and output information on a mainframe.



One of the biggest mainframe computers are those at an IRS regional office. They take up many rooms and have hundreds of terminals, tape stands, and disk drives.

A special class of computers designed to do high speed processing are the super computers. They were the first computers to do a 24-hr weather forecast in less than 25 hours. They work best of weather forecasting and other types of computing where many calculations can be done in parallel. The early super computers were cooled with liquid air. Now days, they are air-cooled. Below are shown an early super computer, the Cray I, and a modern supercomputer, the CrayT3E

Now we also have portable personal computers, called lap top computers. Their display is a built in LCD color screen.

Now we also have palm top computers that fit in one's hand.

Work is also being done on a very small robot called a ninabot which will use a Nina computer for a brain!

Classification is becoming less useful because of rapid evolution of technology. Microcomputers of today are as powerful as minicomputers of few years ago.

A PDA (personal digital assistant) is a type of information appliance that is designed for communication and organizing personal information like addresses and dates. A PDA is typically small enough to carry around in a pocket.

There are several types of computer systems.

PC (Personal Computer)

A PC is a computer designed to meet the needs of a single person and usually refers to IBM-compatible computers. PCs are found in many businesses and are popular for home use.

Macintosh computers are found in many homes and are very popular in the graphics, publishing and multimedia industries. The Macintosh was the first widely used computer that offered a graphical display.

Macintosh

Macintosh computers, or Macs, were introduced by Apple Computer in 1984. Macintosh computers were the first home computers with a mouse, on-screen windows, menus and icons.

Macintosh Advantages

Desktop Publishing

The fast display of images on screen and true What You See Is What You Get (WYSIWYG) display have helped to establish Macintosh computers as the standard in the desktop publishing industry. Desktop publishing lets you create professional documents by integrating text and graphics on a page.

Easy to Use

The graphical interface of Macintosh makes this type of computer very easy to use.

Types of Macintosh Computers

There are many types of Macintosh computers available, including Power Macs, the iMac and notebook computers.

Power Mac

The latest Power Mac, called the G4, is a tower computer that comes with powerful processing capabilities. You can easily expand its capabilities as your needs increase. Apple also manufactures the Power Mac Cube, which is less than a quarter of the size of most tower computers.

iMac

The iMac is an all-in-one computer. An all-in-one computer contains many devices, such as a monitor, CD-ROM drive and speakers, in a single unit. The iMac is easy for beginners to set up and use.

Notebooks

The PowerBook and iBook are notebook computers. Like other notebook computers, the PowerBook and iBook are light-weight and come with a built-in keyboard and screen. You can buy a PowerBook with the same capabilities as a full-sized computer.

Operating System

An operating system is the software that controls the overall activity of a computer. Like the Windows operating systems. Macintosh operating systems use a Graphical User Interface (GUI, pronounced “gooey”). A GUI allows you to use pictures instead of text commands to perform tasks.

Versions

Mac OS 9

Many Macintosh computers currently use the Mac OS 9 operating system, which offers advanced Internet features. The Mac OS 9 is an improved version of previous Mac operating systems, such as Mac OS 8.

Mac OS X

The latest version of the Macintosh operating system is Mac OS X. This version offers many new features, including a new graphical user interface.

LESSON 6
SYSTEM PERFORMANCE

Outline

- Factors affect the performance of computer system.
- Components involved in determining the overall system performance?
- Performance measures
  - Clock speed
  - MIPS
  - Benchmarking Programs
  - CPI/IPC
  - FLOPS

Factors Affect The Performance Of Computer System

Performance Enhancements

Computer performance has been greatly enhanced by the following:

- Faster processing within CPU, & multiple CPU's.
- Wider, faster data and instruction paths (buses).
- More and faster memory, including advanced cache technology.
- Faster disk access.
- Faster, better quality display

Components Involved In Determining The Overall

How Will You Measure The Performance Of A Computer System?

Components Involved in Determining the Overall

Performance

a) CPU Development

- CPU Speed - determined by clock speed, instruction set and pipelining technologies, and super scalar architecture.
- Increased clock speeds have directly increased overall speed of CPU.
- Instruction sets have developed in two opposing directions - CISC and RISC. However, most recent CISC processors use a combination of pipelining, super scalar and RISC technologies to increase throughput.
- Pipelining overlaps the stages of the fetch-execute cycle, but completes execution of one instruction at a time. This is scalar processing.
- Most modern processors implement super scalar processing, using multiple execution units to complete the execution of several instructions at the same time.
- Super scalar processing is highly complex, and has problems such as instructions completing in wrong order, handling program branches, and register conflicts.

b) Bus Technology

- Wider and faster paths provided by improved BUS technology. Wide paths allow several words data/ instructions at once to move between components.
- Newer bus standards allow for substantially higher bus clock speed. E.g. System Bus speed for Pentium III was 133 mhz, cf 400 mhz for P4.
- Modern buses such as PCI use burst mode to transfer multiple data words once starting location is identified.

c) Memory Performance

Apart from having more memory available, several technologies have emerged to reduce performance gap between CPU and memory access:

- Synchronous RAM reads multiple sequential memory locations (burst transfer synchronized to the system clock)).
- Improved cache technology means that ~90% memory accesses satisfied through cache, not main memory.

d) Cache Technology

- Amount, speed, & levels of cache all increased.
- On-chip caching means less distance for data to travel and higher operating speed.
- Separation of level 1 cache into data and instruction cache improves search efficiency.
- Increased quantities of level 2 caches. E.g. Pentium 111 has 32kb L1 on CPU chip, 256 KB L2 in chip package.

e) Disc Access Performance

Several areas of hard drive technology contribute to increased

Performance.

- Increased data density means more data can be retrieved off a smaller physical area, reducing seek time.
- Increased rotation times reduce rotational delay, and data transfer time.
- Disk scheduling algorithms optimise data retrieval in multitasking environments.
- Recent disk controllers read several consecutive sectors and cache them for rapid access.
- High-end systems, e.g. servers may distribute data amongst multiple disks, allowing simultaneous retrieval of data from separate devices.

f) Display Technology

- The display is a collaborative effort between the display device, the graphics card and BUS. Speed, resolution and colour depth all improved.
- Both CRT and Flat screen technology has seen substantial improvements in screen refresh rates.

- Graphics accelerators, with large amounts of on-board RAM and sophisticated processors remove the display-processing load from the CPU.

Performance Of A Computer System

Performance Measures

There are five basic measures used to describe computer system performance:

A) Clock Speed

The clock speed of a CPU is defined as the frequency that a processor executes instructions or that data is processed. This clock speed is measured in millions of cycles per second or megahertz (MHz). The clock itself is actually a quartz crystal that vibrates at a certain frequency when electricity is passed through it. Each vibration sends out a pulse or beat, like a metronome, to each component that's synchronized with it.

By now, you've probably got a handle on the incomprehensible amount of data and information that is flowing around inside your computer. There are many different devices and components, each sending and receiving millions of bytes of information every second. Imagine the havoc if each of these components just sent their information where ever and whenever they wanted, or all at once. There has to be some sort of organization and cooperation between the different devices. That's where these little quartz crystals come in. Each device synchronized with the clock performs each instruction, transmits or receives data, transfers, reads, processes, etc. on each beat of the clock (clock cycle).

Early computers had a single system clock. The expansion bus and the CPU both ran at the same speed. It didn't take long before CPU technology had advanced to where they were faster than the expansion bus. This caused a real bottleneck. The CPU could process and send information faster than the expansion bus could possibly transfer it. The problem was to get each device to run at its own separate speed and still be in synch. What to do?

Well, first off, you give the expansion bus its own clock. Take a hint from music, and make that clock speed a sub-multiple of the system clock. In other words, if the expansion bus clock is running at 8 MHz, then the system clock would run at 16, or 32 etc. Of course, these are arbitrary numbers, the actual values can be pretty strange (like 14.31818 MHz) but you get the idea. Now, have the bus send a signal to the rest of the system that says, "don't send me anymore info until I'm through with this"! When the bus can handle more data, it sends a signal asking for more.

To begin with, the expansion bus was not really given a separate clock. The circuits and technology on the motherboard could split the clock signal or multiply it as needed by certain components. These circuits are called 'frequency dividers' or 'frequency multipliers'. Now, a 100 MHz clock can control the memory system in your computer. A frequency divider is used to reduce the speed by 1/3 to accommodate the 33 MHz PCI bus, and another divider to reduce the speed by 12 for the 8 MHz ISA bus. The CPU can have a frequency multiplier that will increase that 100 MHz speed to 500 MHz or 600 MHz for internal operations. So these devices are still maintaining synch even though they are running at different speeds.

Modern computer technology does allow for some asynchronous operation. Some motherboards now have separate clocks for the expansion bus, system bus, video, etc.. Although these devices run asynchronously, their individual clocks will still run at multiples or sub-multiples of each other which allows for the optimization of transfer rates and data exchange

B) MIPS

Millions of Instructions Per Second. The approximate number of commands that can be carried out in one second. A CPU's power is sometimes measured in MIPS for comparison to another CPU. This is calculated by dividing the number of instructions executed in running a program by the time required to run the program.(Millions of instructions per second) more accurate, but accuracy affected by RISC v/s CISC (200 MIPS cf 8008 .06 MIPS).

C) Benchmarking Programs

most accurate for a particular task, but are task dependent. A benchmark suite consists of set of programs that are believed to be a typical of the programs that will be run on the system. A systems score on the benchmark suit is based on how long it takes the system to execute all of the programs in the suite. Many different benchmark suites exist that generate estimates of a systems performance on different types of applications

Example D) SPEC suite

E) Advantages

i) Their performance results are based on total execution times , not rate of instruction execution.

ii) They average a systems performance across multiple programs to generate an estimate of its average speed.

iii) MIPS rating on only one program

It uses geometric mean rather than arithmetic mean to average the results of the programs contained in the benchmark suite.

Geometric mean of n values is calculated by multiplying the n values together and taking the nth root of the product. Arithmetic mean is calculated by adding all the values together and dividing by the number of values.

D) CPI/IPC

Another metric to describe computer performance clock cycles required executing each instruction known as cycles per instruction. The CPI of a given program on a given system is calculated by dividing the number of instruction executed in running the program IPC instructions executed per clock. High IPC values indicate that the reference program took fewer cycles to execute than low IPC values while high CPI values indicate that more cycles required than low CPI values. Thus a large IPC tends to indicate poor performance.

F) Megaflops and Gigaflops:

1) FLOPS - Floating point operations per second

2) Megaflops- millions floating-point operations/sec

3) Gigaflops- billions floating-point operations/sec.

Following questions should be answered :-

1. What factors affect the performance of computer system ?

2. What components are involved in determining the performance of computer system?

3. Give the full forms of the following terms

a) MIPS

b) CPI

c) IPC

d) FLOPS

4. Write a short note on clock speed.

Note:

LESSON 7
BUS ARCHITECTURE

Outline

- Bus
- Types of memory buses
  - Data bus
  - Address bus
  - Control bus
- Types of I/O buses
  - Data bus
  - Address bus
  - Control bus
- Interconnection architecture
  - Unibus
    - Daulbus
    - DMA

Bus

The interaction between the CPU and memory takes place very frequently. To enable this interaction, some type of connectivity is needed between the two units of a computer system. This connectivity channel is known as bus. Physically, a bus is set of wires, which carries a group of bits in parallel and has an associated control scheme. The bus width is defined as the number of parallel lines in the bus. Every computer has three types of buses for interconnecting the CPU and memory.

Memory Buses

Data Bus

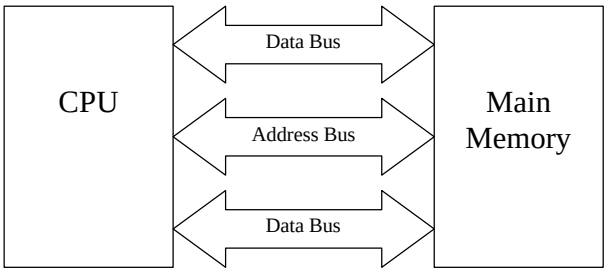
The data bus is used to transfer data between the CPU and memory. The one that you hear most about is the data bus. The bus width of the data bus is important parameter that affects the overall speed of a computer system. This is because each wire of a bus can transfer one bit at a time. Hence an 8-bit bus can move8 bits data time, a 16 bit bus can transfer two bytes, and 32 bit bus can transfer four bytes at a time. This is similar to a multi-lane highway. The wider the highway, the more traffic can flow simultaneously. Similar wider the bus enables more bits of data to travel, simultaneously resulting in faster exchange of data.

Address Bus

Every location in the memory has a unique address. The address of a location does not change, but the data stored in it can change. In order to retrieve some data from memory, it is necessary to specify the address of the location where the data is stored. The address bus is used to carry the address of a memory location whenever data is to transferred to or from memory.

Control Bus

In addition to spending address and exchanging data with the memory, the CPU also needs to send control signals to the memory to specify whether the data is to be read from or written to the specified address location. The control bus carries such signals, which are in the form of READ/WRITE commands.



I/O Buses

The flow of data between the CPU and memory, data also flows from between the CPU and I?O devices in a computer system. Hence just loke the buses between the CPU and memory every computer system also uses buses for interconnecting the CPU with I/O devices.

Data Bus

The data bus is used to transfer data between the CPU and I/O devices

- ISA bus

ISA stands for industry standard architecture. It is a 16-bit bus that can transmit data along either 8 or 16 data lines depending upon what kind of adapter card is used in the expansion slot.

- MCA Bus

MCA stands for micro channel architecture. It is 32-bit bus that transmits data along 32 data lines. Due to is wider bus width, it is significantly faster than the ISA bus. MCA expansion slots can not accept 8-bit or 16 bit adapter cards and required specifically designed 32-bit adapter cards.

- EISA

EISA stands for extended industry standard architecture. Since the expansion slots into an ISA bus could not work with the MCA bus, the MCA bus had the problem of upward compatibility. EISA bus was designed to solve this problem. Hence like MCA , EISA is also a 32-bit bus. However , unlike MCA , it was designed to accept and use the old ISA expansion slots. The EISA is fater than the ISA, but not as fast as the MCA – the price of its compatibility with the older 16-bit expansion slots.

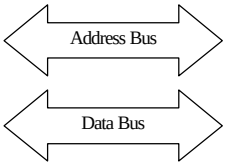
List the similarities between ISA, MCA and EISA buses

Address Bus

A computer system normally has multiple I/O devices like disk, tape, network etc simultaneously connected to it. Each I/O devices has a unique identifier associated with it. The address bus is used to carry the address bus is used to carry the address of the I/O device to be accessed by the CPU.

Control Bus

The control bus is used to carry commands such as Start, READ, WRITE, REWIND Tape, etc. from the CPU to I/O devices. It is also used to carry the status information of the I/O devices to the CPU.



INTERCONNECTION ARCHITECTURES

It defines how exactly these functional units of a Computer system and connected to each other.

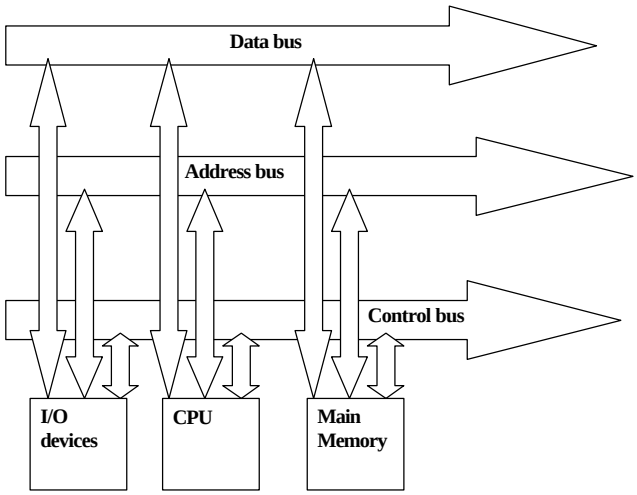
Three commonly used interconnection architecture:

- Unibus
- Dual bus
- DMA

Unibus Architecture

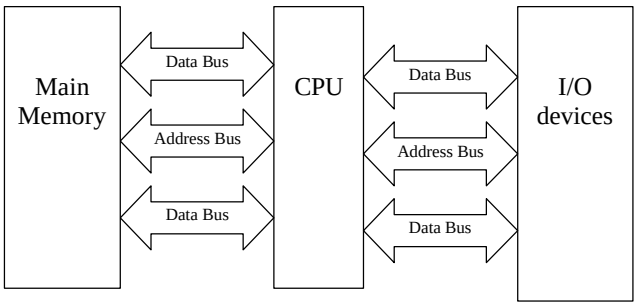
In this type of architecture there is a single data bus, a single address bus, and a single control bus used to interconnect the CPU and with both the memory and I/O devices. That is, the buses are shared for data transfer between the CPU and memory, as well as between the CPU and I/O devices.

What is The Advantage Of This Architecture?



Dual Bus Architecture

There is separate set of buses fro interconnecting the CPU with the memory and with I/O devices. Hence data transfer between the CPU and main memory is carried out by using the buses which interconnect the CPU and main memory, while data transfer between the CPU and I/O devices is carried out by using the buses which interconnect the CPU and I/O devices. In this case, one of the following two mechanisms is used for data transfer between the CPU and an I/O device.



DMA Architecture

DMA is a direct memory transfer. Architecture, which minimizes the CPU

Participation in I/O data transfer is often used.

How this DMA works?

LESSON 8
PRIMARY STORAGE

Outline

- Primary storage
- Ranking Criteria
- Storage location and addresses
- Storage Capacity
- RAM
  - Dynamic RAM
  - Static RAM
- ROM
  - EPROM
  - PROM
- Cache Memory
- Registers

Storage unit have two types of storage primary and secondary. Primary storage in which the data is stored temporarily and in secondary storage the data is stored permanently

Primary Storage

This chapter introduces the basic fundamentals related to the primary storage or the main memory of a computer system. In this chapter, you will learn about storage locations and addressed, how to determine the capacity of storage units, the difference between fixed and variable word-length storage organizations, and several other terms related to the main memory of a computer system.

Ranking Criteria

Any storage unit of a computer system is ranked according to the following criteria:

1. Access time. This is the time required to locate and retrieve stored data from the storage unit in response to a program instruction. A fast access time is preferred.
2. Storage capacity. It is the amount of data that can be stored in the storage unit. A large capacity is desired.
3. Cost per bit of storage. An obvious goal is to minimize this cost.

Based on the above-mentioned criteria, storage units are basically of two types – primary and secondary. As compared to secondary storage units, primary storage units have faster access time, smaller storage capacity, and highest cost per bit of storage. In this chapter, we will be concentrating only on the concepts of primary storage. We shall learn about different types of secondary storage devices in the next chapter.

Storage Locations And Addresses

A primary or internal storage section is basic to all computers. It is made up of several small storage areas called locations or cells. Each of these locations can store a fixed number of bits called word length of that particular primary storage. Thus, as shown in Figure 7.1, a given memory is dividing into N words, where N generally is some power of 2. Each word or location has a built-in and unique number assigned to it. This number is called the address of the locution and is used to identify the location. Each location can hold either a data item or an instruction, and its address remains the same regardless of its contents. The addressed normally start at 0 and the highest address equals the number of words that can be sorted in the memory minus 1.

For example, if a memory has 1024 locations, then the address ranges between 0 and 1023. Thus, at address 0 we find a word, at address 1 a second word, at address 2 a third word, and so on upto the final word at the largest address.

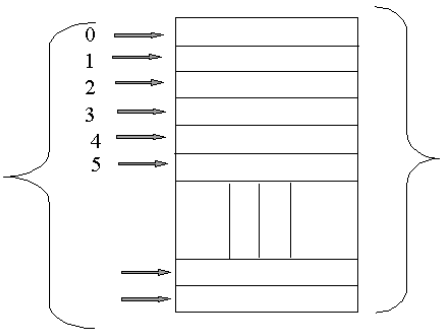


Figure 7 Organization of a Primary storage Unit having n Words.

There is an important difference between the address number and the contents of the address. A memory is like a large cabinet containing as many drawers as there are addresses in memory. Each drawer contains a word a word and the address of each word is written on the outside of the drawer. If we write or store a word say 10101010 at address 125, it is like placing the word 10101010 in the drawer labeled 125. Later, reading from address 125 is like looking in that drawer to see its contents which is now 10101010. We do not remove the word at an address when we read, but change the contents at an adds only when we store or write a new word. Thus, entering data into a storage location is destructive of previous contents, but retrieving data from a location is non-destructive.

Storage Capacity

The storage capacity of large computer systems is normally more than small systems. This capacity is defined in terms of bytes or words. A symbol commonly used to denote the storage capacity of a computer is the letter k (Kilo), which is equal to 2^{10}, or 1024. Thus a 32 Kilobyte memory is capable of storing 32 x 1024 = 32,768 bytes or characters. Memory sizes range from a few K bytes or words n small machines to several thousand K bytes or words in large machines. However, if the

memory capacity is stated in terms of words, it is necessary to know the word size in bits or bytes in order to determine the actual storage capacity of the computer. Thus, while specifying the memory capacity in terms of words, it is customary to specify the total number of bits per word along with the total number of words. So a memory with 4096 locations each with a different address and with each location storing 16 bits is called a "16-bit 4096-word memory, by or in the vernacular of the computer trade, a "4 K 16-bit memory". Similarly, a memory having 2^{15} words with each word of 16 bits is called a "32 K 16-bit memory". If the word size of a memory is 8 bits (equal to a byte) then it becomes immaterial whether the memory capacity is expressed in terms of bytes or words. Thus a memory having 216 words with each word of 8 bits is simply refereed to as 64 K memory(word size of 8-bits is implicit here).

Why More Bits

You might have heard about 8-bit computers, 16-bit computers, 32-bit computers, etc. This refers to the word size of a particular computer in terms of total number of bits. Word size is an important architectural factor. Small machines have word sizes of 8 or 16 bits; large machine word sizes are 32 bits or more. The obvious question that arises to ones mind is that what is the advantage of having more number of bits per word instead of having more words of smaller size?

For an answer to the above question, imagine a highway with eight lanes, and a heavy flow of traffic. If it is expanded to sixteen lanes, the flow of traffic speeds up considerably. 8 –bits refers to the number of lanes on a microchip. Greater bits means a more rapid flow of electronic signals. In other words, faster computer. Thus, what an 8-bit computer takes one minute to do; a 32-bit computer can do in one second.

In word-sized bunches. Therefore, even if the electronic circuits used are comparable in speed, small machines are slower than big machines. This difference is analogous to providing the user of a small machine with a small data shovel. Even though they both may be shoveling at comparable speeds, the user with the smaller shovel will be slower because more shovels will be slower because more shovelfuls are needed to move the same amount of data.

The primary storage section of some computers is designed to store a fixed number of characters (equal to its word-length in bytes) in each numbered address location. Such computers are said to be word addressable, and they employ a fixed word-length storage approach. In these computes, storage space is always allocated in multiples of word-length of 4 characters then this computer will require one word (4 bytes) to store the word "CAT" and two words (8 bytes) to store the word "BOMBAY".

In many computers, the primary storage section is also designed in such a way that each numbered address can only store a single character (A, B, 1,2, +, -, etc). Computers designed in this manner are said to be character addressable and they employ a variable word-length storage approach. Thus in these machines, only 6 bytes will be require to store the word "BOMBAY". Figure 7.2 summarizes the difference between the fixed-length and variable-length storage approach.

Both the fixed and the variable word-length storage systems have their own merits and demerits. The fixed word-length storage approach is normally used in large scientific computers for gaining speed of calculation. One the other hand, the variable word-length approach is use din small business computes for optimizing the use of storage space. For example, let us consider a fixed word-length machine with a word size of eight characters. If most of the data words to be stored are of les than five characters. If most of the data words to be stored are of less than five characters then more than half of the storage space will remain unused. This will not happen in case of machine with variable word-length primary storage because a character can be placed in every storage cell of this machine. However, word addressable computes possess faster calculating capability because they can add tow data words in a single operation. If the fixed – length word is eight characters, two eight digit numbers can be added in signal operation. On the other hand with a character addressable machine only one digit in each number is added during a single operation and eight steps would thus be needed to add two eight digit numbers.

Most of the today's business and scientific processing is handled by flexible computes, which can employ either a fixed word-length or a variable-length storage organization. The set of instructions available with these computers allow them to be operated as either variable or fixed – length computers.

RAM (Random access Memory)

Primary storage is usually referred to as random access memory (RAM) because it is possible to randomly select and use any location of this memory to directly store and retrieve data and instructions. Each separate location inside the memory is as easy to access as any other location, and takes the same amount of time, It is also referred to as red/write memory because information can be read from a RAM chip and can also be written into it.

02 a type of computer memory that can be accessed randomly; that is, any byte of memory can be accessed without touching the preceding bytes. RAM is the most common type of memory found in computers and other devices..

There are two basic types of RAM:

- dynamic RAM (DRAM)
- Static RAM (SRAM)

The two types differ in the technology they use to hold data, dynamic RAM being the more common type. Dynamic RAM needs to be refreshed thousands of times per second. Static RAM does not need to be refreshed, which makes it faster; but it is also more expensive than dynamic RAM. Both types of RAM are *volatile*, meaning that they lose their contents when the power is turned off.

In common usage, the term *RAM* is synonymous with *main memory*, the memory available to programs. For example, a computer with 8M RAM has approximately 8 million bytes of memory that programs can use. In contrast, *ROM (read-only memory)* refers to special memory used to store programs that boot the computer and perform diagnostics. Most personal computers have a small amount of ROM (a few thousand bytes). In fact, both types of memory (ROM and RAM) allow

random access. To be precise, therefore, RAM should be referred to as *read/write RAM* and ROM as *read-only RAM.*

ROM

A read only memory (ROM) is one in which information is permanently stored. The information from the memory can only be read and it is not possible to write fresh information into it. This the reason why it is called ROM. When the power supply is switched off, the information stored inside a ROM is not lost as it is in the case of a RAM chip. Such memories are also known as field stores, permanent sores or dead stores.

The most basic computer functions are carried out by wired electronic circuits. However, there are several higher-level operations that are very frequently used but will require very complicated electronic circuits for their implementations. Hence instead of building electronic circuits for these operations, special programs are written to perform these operations. These programs are called microprograms because hey deal with low-level machine functions and are essentially substitute for additional hardware. Microprograms are written to aid the control unit in directing all the

operation of ht computer system. ROMs are mainly use by computer manufacture's for storing these microprograms so that they cannot be modified by the users.

a) PROM

A variation of ROM chip is programmable read only memory (PROM). ROM chips are supplied by the computer system manufacture and it is not possible for a customer (user) to modify the programs stored inside the ROM chip. However, it is possible for a user to "customize" a system by converting his own programs to microprograms and storing them in a PROM chip. Once the users programs are stored n a PROM chip. Once the users programs are stored in a PROM chip. Once the users programs are stored in a PROM chip, they can usually be executed in a fraction of the time previously required. PROMs are programmed to record information using a special facility known as a prom-programmer. However, once the chip has been programmed, the recorded information cannot be changed, i.e. the PROM becomes a ROM and it is only possible to read the stored information. PROM is also nonvolatile storage, i.e. the stored information remains intact even if power is switched off.

b) EPROM

Once information is stored in a ROM chip or a PROM chip it cannot be changed or altered. However, there is another type of memory chip called erasable programmable read only memory chip called erasable programmable read only memory (EPROM) that overcomes this problems. As the name implies, it is possible to erase information stored in an EPROM chip and the chip can be reprogrammed to store new information suing a special prom-programmer facility. Information stored in an EPROM chip is erased by exposing the chip for some time to ultraviolet light. When an EPROM is in use, information can only be read and the information remains on the chip until it is erased. Eon can only be read and the information remains on the chip until it is erased.

EPROMs are mainly used by R & D personnel (experiments) because they frequently change the microprograms to test the efficiency of the computer system with new programs.

Regardless of the type of ROM chip use, they all serve to increase the efficiency of CPU by controlling the performance of a new specialized task. A generalized CPU can be made to meet the unique needs of different users merely by changing micro programs.

Cache Memory

A special very high-speed memory is sometimes used to increase the speed of processing by making current programs and data available to the CPU at a rapid rate. As you may be aware, CPU speeds are quite high compared to the access time of main memory. In many situations the performance of the processors are limited due to the slow speed of the main memory. A technique used to compensate for the mismatch in operating speeds is to employ an extremely fast, small memory between CPU and main memory whose access time is close to the processing speed of the CPU. This type of memory is called a high-speed buffer or cache memory. It is used to store segments of programs currently needed in the present calculations. By making active programs and data available at a rapid rate, it is possible to increase the performance rate of the CPU.

As the name implies, cache memory is a memory in hiding and is not addressable by the user of the computer system: Its purpose is to look ahead and to provide the CPU with currently needed information. Cache memory makes main memory appear to be faster and larger than it really is . It is very expensive as compared to the main memory and hence its size is normally very small. The most important effects of cache memory is in improving the memory transfer rates and thus raising the processor speed.

Registers

As the instruction are interpreted and executed by the CPU, there is a movement of information between the various units of the computer system. In order to handle this process satisfactorily and to speed up the rate of information transfer, the computer uses a number of special memory units called register. These register are not considered as a part of the main memory and are used to retain information on a temporary basis.

The number of registers varies among computers, as does the data flow pattern. Most computers use several types of registers, each designed to perform a specific functions. Each of these register posses the ability to receive information, to hold it temporarily, and to pass it on as directed by the control unit. The length of a register equals the number of bits it can store. Thus a register that can store 8 bits is normally referred to as 8-bit register. Although the number of registers that are common to all computers. The function of these registers is described below.

1. Memory Address Register (MAR): It holds the address of the active memory location. It is loaded from the program control register when an instruction is read from memory.

2. Memory Buffer Register (MBR) : It hold the contents of the memory word read from, or written in, memory. An

instruction word placed in this register is transferred to the instruction register. A data word placed in this register is accessible for operation with the accumulator register or for transfer to the I/O register. A word to be stored in a memory location must first be transferred to the MBR, from where it is written in memory.

3. Program Control Register (PC) : It holds the address of the next instruction to be executed. This register goes though a step-by-step counting sequence and causes the computer to read successive instructions previously stored in memory. It is assumed that instruction words are stored in consecutive memory locations and red and executed in sequence unless a branch instruction is encountered. A branch instruction is an operation that calls for a transfer to a nonconsecutive instruction. the address part of a branch instruction is transferred to the PC register to become the address of the next instruction. To read an instruction, the contents of the PC register are transferred to the MAR and a memory read cycle is initialed. The instruction placed in the MBR is then transferred to the instruction register

4. Accumulator Register (A) : This register holds the initial data to be operated upon, the intermediate results, and also the final results of processing operations. It is used during the execution of most instructions. The results of arithmetic operations are returned to the accumulator register for transfer to main storage through the memory buffer register. In many computers there are more than one accumulator registers.

5. Instruction Register (I) : It holds the current instruction that is being executed. As soon as the instruction is stored in this register, the operation part and the address part of the instruction are separated. The address part of the instruction is sent to the control section where it is decoded and interpreted and ultimately command signals are generated to carry out the task specified by the instruction.

6. Input/Output register (I/O) : This register is used to communicate with the input/output devices. All input information such as instructions and data is transferred to this register by an input device. Similarly, all output information to be transferred to an output device is found in this registers.

Table 1 summarizes the functions of each of these registers.

Table 1 Functions of Various Registers

SE NO.	NAME OF REGISER	FUNCTION
1.	Memory Address (MAR)	Holds the address of the active memory location.
2.	Memory Buffer (MBR)	Holds information on its way to and from memory.
3.	Program Control (PC)	Holds the address of the next instruction to be executed.
4	Accumulator (A)	Accumulates results and data to be operated upon.
5	Instruction (I)	Holds an instruction while it is being executed.
6.	Input/Output (I/O)	Communicates with the I/O devices

Following questions should be answered: -

1. Difference between a bit, a byte and a word.

2. What is a memory address? Describe with an example.

3. Give the full form of following abbreviations

a) RAM

b) ROM

c) EPROM

d) MBR

e) PROM

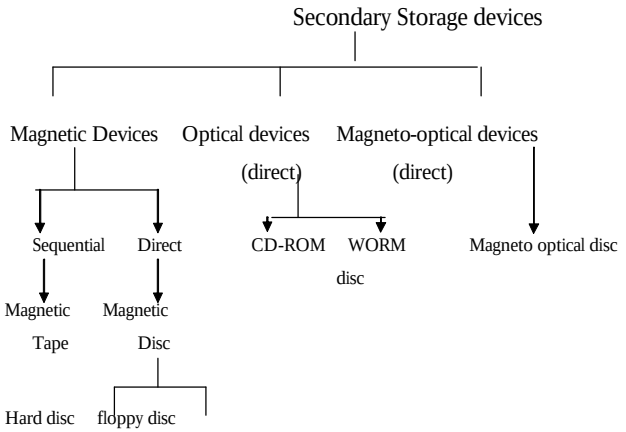
4. What are registers? Name some commonly used registers and briefly describe the function of each.

5. What is the impact of cache memory on the performance of computer system

LESSON 9
MAGNETIC DISC

Outline

- Classes of secondary storage media
- Magnetic Devices
  - Magnetic tape
  - Magnetic disks
    - Floppy disc
    - Hard disc
- Optical devices
  - CD-ROM
  - WORM DISC
- Magneto Optical devices



Secondary Storage

Secondary storage provides cheap, non-volatile high capacity storage.

Two classes of secondary storage media :

1. Direct Access Media (e.g. magnetic disk)- supports sequential or random access. i.e. can go straight to the required block.
2. Sequential Access media (e.g. magnetic tape)- supports only sequential type access. i.e. begin at one point and move past each block until required one is reached.

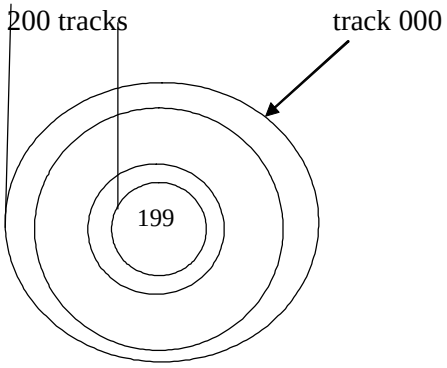
Magnetic Devices

Magnetic Disc

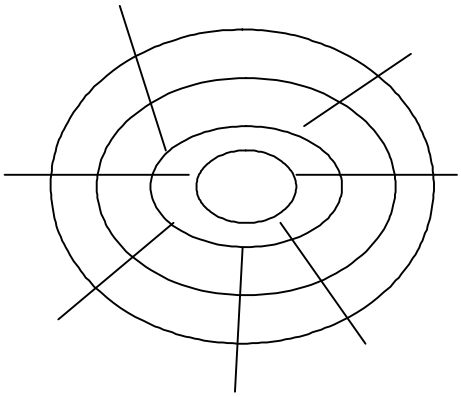
- Direct access storage devices (DASD's) of two basic types - floppy disks and hard disks. Some hard disks are portable (removable).
- Consist of platter(s) rotating under read/write head(s) with space organised into sectors, tracks, & cylinders.
- Data is recorded as magnetized spots on coating.

Storage Organization

For data recording, the surface of a disk is divided into a number of invisible concentric circles called tracks. The tracks are numbered consecutively from outmost to innermost starting from zero. The number of tracks varies greatly between disks, from as few as 40 on some small, low capacity disks to several thousand on large, high capacity disks



Each track is further subdivided into sectors. For this, in addition to the concentric circles, the disc surface is also divided into invisible pie-shapes segments. Thus if there are eight such pie shaped segments, each track will get divided into eight parts, and each of these eight portions of a track is called a sector.



Storage capacity depends on

- Number of recording surfaces
- Numbers of tracks per surface
- Number of sectors per track
- Number of byes per sector

Storage capacity of a disk system = number of recording surfaces X Number of tracks per surface X

What will be the storage capacity of a double-sided disc that has 400 tracks, 16 sectors per track and 512 bytes per sector?]

Hard Disc

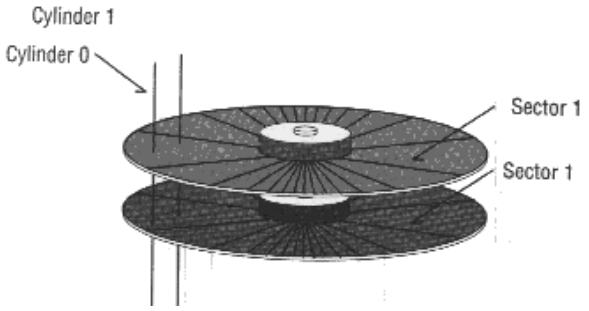
- Hard disks usually have multiple platters with R/W heads on each surface.
- Set of vertically aligned tracks is called a cylinder
- To access data, must specify cylinder and sector nos. Always read/write min of 1 sector.
- Winchester style sealed units allow high data density due to exclusion of contaminants. R/W head "floats" just above disk surface

A magnetic disk on which you can store computer data. The term *hard* is used to distinguish it from a soft, or *floppy*, disk. Hard disks hold more data and are faster than floppy disks. A hard disk, for example, can store anywhere from 10 to more than 100 gigabytes, whereas most floppies have a maximum storage capacity of 1.4 megabytes.

A single hard disk usually consists of several *platters.* Each platter requires two read/write heads, one for each side. All the read/write heads are attached to a single access arm so that they cannot move independently. Each platter has the same number of *tracks*, and a track location that cuts across all platters is called a cylinder. For example, a typical 84 megabyte hard disk for a PC might have two platters (four sides) and 1,053 cylinders.

In general, hard disks are less portable than floppies, although it is possible to buy removable hard disks.

Basically, tracks, sectors, and cylinders are the divisions of the hard drive platters where information is stored. A track is a concentric ring around the platter containing information. Since a hard drive typically has two or more platters, each storing data on both sides, these tracks line up on each platter. The identically positioned tracks on each platter are called cylinders. To better help you understand a track and cylinder, let's take a target used for target practice. You have a bunch of concentric circles, each bigger than the other, all sharing the same center, which is the bulleye. Now, each of the spaces between circles is similar to a track on a hard disk platter. Now, if you stack several of these targets on top of each other, each exactly the same, you can form a cylinder by simply taking a track and moving it down through all of the same tracks on the targets below.

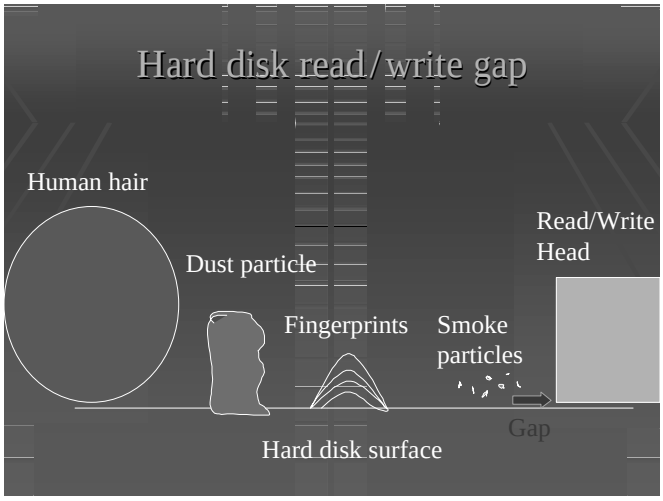


Since typical hard drives are too large to deal with by the track, each track is divided into sectors. Its not that a track could not be dealt with, but since a track can hold as much as 50K sometimes, this would not be practical for storing large files. So, sectors are basically slices of the track. Different drives have different numbers of sectors per track.

Each sector is given an identity during low-level formatting to aid the controller in finding what it needs in the appropriate sector. These sector numbers are written to the beginning and the end of each sector, called the prefix portion and the suffix portion respectively. These identities take actual space on the hard drive. This explains why there is a difference between the capacity of an unformatted disk and a formatted one. On a floppy, the disk itself can hold 2M or so of data. When formatted and the identities placed, the capacity reduces to 1.44M. The same holds true for a hard drive. Drive manufacturers know this and publish formatted capacities to indicate drive size.

Removable DISC

- SyQuest,
- Bernoulli,
- Zip



Magnetic Tape

A magnetically coated strip of plastic on which data can be encoded. Tapes for computers are similar to tapes used to store music.

Storing data on tapes is considerably cheaper than storing data on disks. Tapes also have large storage capacities, ranging from a few hundred kilobytes to several gigabytes. Accessing data on tapes, however, is much slower than accessing data on disks. Tapes are *sequential-access* media, which means that to get to a particular point on the tape, the tape must go through all the preceding points. In contrast, disks are *random-access* media because a disk drive can access any point at random without passing through intervening points.

Because tapes are so slow, they are generally used only for long-term storage and backup. Data to be used regularly is almost always kept on a disk. Tapes are also used for transporting large amounts of data.

Tapes come in a variety of sizes and formats.

Tapes are sometimes called *streamers* or *streaming tapes.*

Floppies

A soft magnetic disk. It is called *floppy* because it flops if you wave it (at least, the 5¼-inch variety does). Unlike most hard disks, floppy disks (often called *floppies* or *diskettes*) are portable, because you can remove them from a disk drive. Disk drives for floppy disks are called *floppy drives.* Floppy disks are slower to access than hard disks and have less storage capacity, but they are much less expensive. And most importantly, they are portable.

Floppies come in three Basic Sizes:

- 8-inch: The first floppy disk design, invented by IBM in the late 1960s and used in the early 1970s as first a read-only format and then as a read-write format. The typical desktop/laptop computer does not use the 8-inch floppy disk.
- 5¼-inch: The common size for PCs made before 1987 and the predecessor to the 8-inch floppy disk. This type of floppy is generally capable of storing between 100K and 1.2MB (megabytes) of data. The most common sizes are 360K and 1.2MB.
- 3½-inch: *Floppy* is something of a misnomer for these disks, as they are encased in a rigid envelope. Despite their small size, microfloppies have a larger storage capacity than their cousins — from 400K to 1.4MB of data. The most common sizes for PCs are 720K (double-density) and 1.44MB (high-density). Macintoshes support disks of 400K, 800K, and 1.2MB.

Optical Storage

a) Compact Disc CD-ROM

Wednesday, August 13, 2003

Known by its abbreviation, *CD*, a compact disc is a polycarbonate with one or more metal layers capable of storing digital information. The most prevalent types of compact discs are those used by the music industry to store digital recordings and CD-ROMs used to store computer data. Both of these types of compact disc are *read-only*, which means that once the data has been recorded onto them, they can only be read, or played.

- Small, removable, read only drives
- Same media as audio CDs
- Large capacity (an encyclopedia can be on a disk)
- CD-R drives write to them
- Write once, read many (WORM) drives
- Compatible with regular CD-ROM reading drives

a) Magneto Optical device

: Thursday, May 08, 2003

A type of disk drive that combines magnetic disk technologies with optical technologies, such as those used in CD-ROMs. An MO disk drive is designed so that an inserted disk will be exposed to a magnet on the label side and to the laser beam on the opposite side.

b) DVD

Last modified: Thursday, October 16, 2003

A type of optical disk technology similar to the CD-ROM. A DVD holds a minimum of 4.7GB of data, enough for a full-length movie. DVDs are commonly used as a medium for digital representation of movies and other multimedia presentations that combine sound with graphics.

The DVD specification supports disks with capacities of from 4.7GB to 17GB and access rates of 600KBps to 1.3 MBps. One

of the best features of DVD drives is that they are backward-compatible with CD-ROMs, meaning they can play old CD-ROMs, CD-I disks, and video CDs, as well as new DVD-ROMs. Newer DVD players can also read CD-R disks.

Performance Factors

- Hard drive speed is v. important because modern O/S use them as an extension of RAM (virtual memory).
- This can cause performance degradation as hard drives are ~ 106 times slower than RAM.
- Two measures of hard drive performance:
  - Access time; and
  - Data transfer rate.

Total Access Time

- Total access time = seek time + rotational delay + data transfer time
- Seek time = time to move head to correct cylinder, (max will be seek from inner to outer track).the time required to position the read/write head over the desired track is called the seek time. The seek time varies depending on the position of the access arms assembly when a read/write command is received. If the access arms assembly is positioned on the outer most track and the track to be reached is the inner most one then the seek time will be maximum, and it will be zero if the access arms assembly already happens to be on the desired track. The average seek time is thus specified for most systems which order of 10 to 100 milliseconds.
- Rotational delay (latency) = time for sector to move under head (max latency when head just past required sector). Once the heads are positioned on the desired track, the head on the specified surface is activated. Since the disk is continuously rotating, this head should wait for the desired data to come under it. This rotational waiting time i.e. the time required to spin the desired sector under the head is called latency.
- Data transfer rate- transfer rate refers to the rate at which data are read or written from or written to the disk. Once the read/write head is positioned ove the desired sector, the data read/written at a speed determined by the rotational speed of the disc.

If the rotational speed of a disk is 3600 rpm, the disk has 125 sectors/track and 512 bytes/sector then in one full revolution of the disk, what amount of data will be transferred.

Following questions should be answered: -

1. Distinguish between a sequential access, a direct access and a random access device. Give one example of each.

2. What factors determine the storage capacity of disks ?

LESSON 10
INPUT DEVICES

Outline

- Input output components
- Input devices
  - Keyboard
  - Mouse
  - Touch Screen
  - Voice reorganization system
  - Joystick
  - Track ball
  - Scanner
  - Audio input

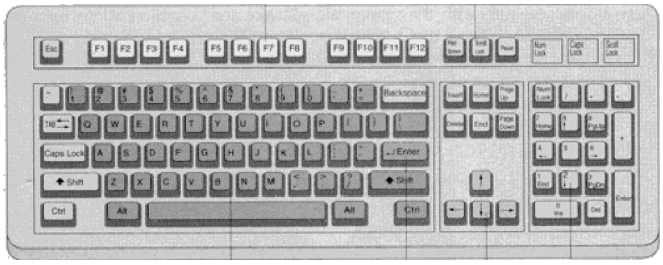
Input Devices

Input devices are the electronic or electromechanical equipment that provides a means of communicating with the computer system for the input of data. As previously stated, the processing unit requires this data in the appropriately formatted electronic signals. The input device that most users are familiar with is the keyboard. This converts a key press in a particular position on the board into an electronic signal sequence that can be interpreted by the processing unit as representing a particular character.

Input Output Components

- Parallel port
  - Parallel because it can move a whole byte at a time
  - Mainly used for connection to a printer
  - Nowadays it can handle bi-directional traffic
- Serial port
  - One bit at a time-Uses thin cable
  - Slow-inexpensive
- Universal serial bus (USB) replaces those
  - It is a new, rectangular port device
  - 4 wires
    - 2 for power
    - 2 for communication
  - USB has a hub/slave technology
  - One USB line can connect to a hub
  - The hub has several outgoing lines that connect to other devices
  - This is achieved by a complex set of software

Keyboard



This is what you use to type information into your computer. It is the most common input device. Keyboards usually have a Qwerty layout.

Keys on the Keyboard

- Typewriter keys
- Function keys
- Arrow keys
- Editing keys
- Numeric keypad
- Alt and Control keys
- Escape, windows key
- Obsolete keys

Different keys on the keyboard have different names. They are used to do different things.

Function keys perform a specific task. The program you are using sets this task. For example, pressing F1 usually displays the help screen.

Special purpose keys perform special preset tasks. For example, in Word 97 pressing the Print Scrn (Print Screen) key transfers a picture of your screen to the clipboard.

The typewriter keys are used for typing text. They can also be used to perform tasks that appear on the program's menus. For example, in Word 97, Ctrl+V (hold down the control and the V key at the same time) does the same action as Edit / Paste (select Paste from the Edit menu).

Pressing TAB moves you to the next tab in a word processing program and the next cell in a spreadsheet program.

Caps Lock sets to capitals all the text you write.

The Enter key causes data to be accepted by the computer. If you are typing, pressing the enter key begins a new paragraph.

The cursor control keys move your cursor (the flashing bar where text appears) around the screen.

Numeric keys allow you to enter numbers quickly.

Dvorak Keyboard

- An efficient keyboard for the English language, patented by August Dvorak, cousin of the composer.
- Result of efforts to study typewriting behavior.
- Studies in the US Navy suggest that the payback period would be 10 days from learning Dvorak.
- From my own experience, doubtful.

This is another common input device. A mouse is used to point to information on the screen and either Click on it or Click and drag the information.

- There are two-button and three button mice.
- Since MS uses two buttons, three button mice are a dying species.
- The third (middle) button can be emulated on two-button mice by pressing the left and the right button at the same time.
- Use of the mouse is best avoided! the two buttons
- Primary button — usually left
- Secondary button – usually right
- Actions in Microsoft windows
  - Left click
  - Right click
  - Left double click (time adjustable)
  - Left drag and drop
  - Right drag and drop

A Mouse can have one, two or three buttons.

On a two button mouse:

- To *Click* means to quickly press down and release the left mouse button.
- To *Double Click* means to give two clicks very quickly, one after the other.
- To *Right Click* means to click with the right mouse button. This will perform a different action to normal clicking (done with the left mouse button).
- To *Click and Drag* means to point the mouse at an object, hold down the left mouse button, move the mouse, and release the left mouse button. This moves the object you pointed to from where it was to where it is when you release the mouse button.
- A device that controls the movement of the cursor or pointer on a display screen. A mouse is a small object you can roll along a hard, flat surface. Its name is derived from its shape, which looks a bit like a mouse, its connecting wire that one can imagine to be the mouse's tail, and the fact that one must make it scurry along a surface. As you move the mouse, the pointer on the display screen moves in the same direction. Mice contain at least one button and sometimes as many as three, which have different functions depending on what program is running. Some newer mice also include a *scroll wheel* for scrolling through long documents.

Invented by Douglas Engelbart of Stanford Research Center in 1963, and pioneered by Xerox in the 1970s, the mouse is one of the great breakthroughs in computer ergonomics because it frees the user to a large extent from using the keyboard. In particular, the mouse is important for graphical user interfaces because you can simply point to options and objects and click a mouse button. Such applications are often called *point-and-click* programs. The mouse is also useful for graphics programs that allow you to draw pictures by using the mouse like a pen, pencil, or paintbrush.

There are three basic types of Mice:

1. Mechanical

Has a rubber or metal ball on its underside that can roll in all directions. Mechanical sensors within the mouse detect the direction the ball is rolling and move the screen pointer accordingly.

2.Optomechanical

Same as a mechanical mouse, but uses optical sensors to detect motion of the ball.

3.Optical

Uses a laser to detect the mouse's movement. You must move the mouse along a special mat with a grid so that the optical mechanism has a frame of reference. Optical mice have no mechanical moving parts. They respond more quickly and precisely than mechanical and opt mechanical mice, but they are also more expensive.

Cordless mice aren't physically connected at all. Instead they rely on infrared or radio waves to communicate with the computer. Cordless mice are more expensive than both serial and bus mice, but they do eliminate the cord, which can sometimes get in the way.

Touch Screen

Last modified: Friday, January 18, 2002

.INCLUDEPICTURE "http://63.236.18.118/RealMedia/ads/adstream_nx.ads/intm/sbc/www.webopedia.com@468x60-1,468x60-2,125x125-1,336 x280,336x280,125x800,cp1,cp2,cp3,cp4,cp5,cp6,cp7!336x280"

Mergeformatinet

A type of display screen that has a touch-sensitive transparent panel covering the screen. Instead of using a pointing device such as a mouse or light pen, you can use your finger to point directly to objects on the screen.

Although touch screens provide a natural interface for computer novices, they are unsatisfactory for most applications because the finger is such a relatively large object. It is impossible to point accurately to small areas of the screen. In addition, most users find touch screens tiring to the arms after long use.

Last modified: Monday, February 25, 2002

Voice Recognition

The field of computer science that deals with designing computer systems that can recognize spoken words. Note that voice recognition implies only that the computer can take dictation, not that it *understands* what is being said. Compre-

hending human languages falls under a different field of computer science called *natural language processing.*

A number of voice recognition systems are available on the market. The most powerful can recognize thousands of words. However, they generally require an extended training session during which the computer system becomes accustomed to a particular voice and accent. Such systems are said to be *speaker dependent*.

Many systems also require that the speaker speak slowly and distinctly and separate each word with a short pause. These systems are called *discrete speech* systems. Recently, great strides have been made in *continuous speech* systems - voice recognition systems that allow you to speak naturally. There are now several continuous-speech systems available for personal computers.

Because of their limitations and high cost, voice recognition systems have traditionally been used only in a few specialized situations. For example, such systems are useful in instances when the user is unable to use a keyboard to enter data because his or her hands are occupied or disabled. Instead of typing commands, the user can simply speak into a headset. Increasingly, however, as the cost decreases and performance improves, speech recognition systems are entering the mainstream and are being used as an alternative to keyboards.

Other Pointing Devices

- Replacements for a mouse are
  - Trackballs
  - touch pads
  - Joysticks
- Pen-based devices
  - Vary with respect to text recognition

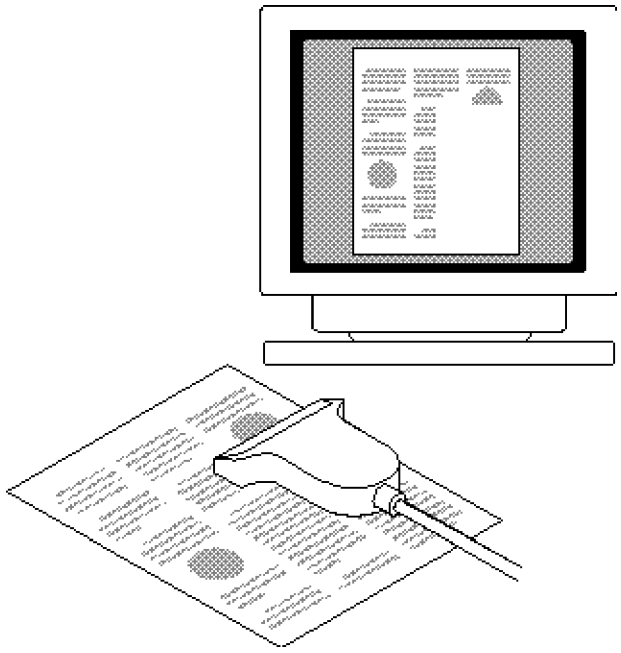
The Joystick

These are used mainly in computer games. Moving the joystick will move the car / person / object on the screen.

The Trackball

A trackball has been described as an upside down mouse. It is a small ball with buttons around it that is either part of the keyboard or attached to the keyboard. Moving the ball moves the pointer on the computer screen. The buttons are used to do the same things as the buttons on a mouse. A trackball is used instead of a mouse in notebook and laptop computers.

Scanner



A device that can read text or illustrations printed on paper and translate the information into a form the computer can use. A scanner works by digitizing an image — dividing it into a grid of boxes and representing each box with either a zero or a one, depending on whether the box is filled in. (For color and gray scaling, the same principle applies, but each box is then represented by up to 24 bits.) The resulting matrix of bits, called a bit map, can then be stored in a file, displayed on a screen, and manipulated by programs.

Optical scanners do not distinguish text from illustrations; they represent all images as bit maps. Therefore, you cannot directly edit text that has been scanned. To edit text read by an optical scanner, you need an *optical character recognition (OCR)* system to translate the image into ASCII characters. Most optical scanners sold today come with OCR packages.

Scanners differ from one another in the following respects:

- Scanning Technology: Most scanners use charge-coupled device (CCD) arrays, which consist of tightly packed rows of light receptors that can detect variations in light intensity and frequency. The quality of the CCD array is probably the single most important factor affecting the quality of the scanner. Industry-strength drum scanners use a different technology that relies on a photo multiplier tube (PMT), but this type of scanner is much more expensive than the more common CCD -based scanners.
- Resolution: The denser the bit map, the higher the resolution. Typically, scanners support resolutions of from 72 to 600 dpi.
- Bit Depth: The number of bits used to represent each pixel. The greater the bit depth, the more colors or grayscales can be represented. For example, a 24-bit color

scanner can represent 2 to the 24th power (16.7 million) colors. Note, however, that a large color range is useless if the CCD arrays are capable of detecting only a small number of distinct colors.

- Size and Shape: Some scanners are small hand-held devices that you move across the paper. These hand-held scanners are often called *half-page* scanners because they can only scan 2 to 5 inches at a time. Hand-held scanners are adequate for small pictures and photos, but they are difficult to use if you need to scan an entire page of text or graphics.

Larger scanners include machines into which you can feed sheets of paper. These are called *sheet-fed* scanners. Sheet-fed scanners are excellent for loose sheets of paper, but they are unable to handle bound documents.

A second type of large scanner, called a *flatbed scanner*, is like a photocopy machine. It consists of a board on which you lay books, magazines, and other documents that you want to scan.

Overhead scanners (also called *copy board* scanners) look somewhat like overhead projectors. You place documents face-up on a scanning bed, and a small overhead tower moves across the page.

This is used to put both text (writing) and graphics (pictures) into the computer. It digitizes the text and images i.e. it converts them into a form that can be used in the computer. This process is known as optical character recognition (O.C.R). Before the text can be used on the computer, it has to be converted into characters by optical recognition software.

a) Text and Code Scanners

- Bar-code readers
- Magnetic ink character recognition such as on bank checks.
- Optical mark recognition
- Optical character recognition
  - More general form of reading characters out of handwriting or print
    - Expensive and error-prone

b) Image Scanners

- Most widely used digitization device
- You can try to run OCR once you have the image
- Resolution measured in dots per inch. dpi
  - More dpi: crisper image
  - Fewer dpi: smaller file
- Audio Input
- Microphone with speech recognition software
  - This used to be mainly for those who cannot type
  - But the quality has improved, and maybe we will use it for more general input.
- Other analog to digital sound converters
  - Sound cards
  - MIDI board

Write some name of companies that provides cheap and best input devices like keyboard , mouse, scanner etc.

1. What are input output components?

LESSON 11
OUTPUT DEVICES- MONITORS

Outline

- Display Characters tics(performance characteristics)
  - Video modes
  - Colour depth
  - Display Speed
  - Displays
- Display Devices
  - CRT monitors
  - Flat panel monitors

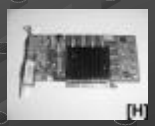
Video Monitor Screens

- Immediately see characters as typed
- Video screen provides a way for a computer user to receive messages from the computer
- Early versions character based
- Modern versions support pictures and graphics
- Connected to computer by special circuits that are often on cards
- Use special kind of memory, video memory, (VRAM) dedicated to video
- The more VRAM, the more detail

Monitor size measured like TV screen, diagonal line across screen

- Images on screen made of tiny dots called pixels, for picture elements
- Colors are additive
- Each pixel made up of red, green and blue dots in proximity
- Typical monitors are 72 dots per inch. Closer together dots (more dots per inch) means greater resolution
- When measured in columns and rows, the most common size is 640x480
- Color depth is measured in bits
- Monochrome 1 bit
- 256 color (common) is 8 bit
- True color (24 bit) has 16 million c
- Grey Scale, displays black, wh
- Multisynch monitors ca color systems, withi
- Two classes of
- Cathode R
- Liquid

DISPLAY system consists of video display adaptor (video card) and display device (monitor).



Adaptor and monitor must be compatible & work together to produce an image.

Display Characterstices (Performance Factors)

Display Characteristics

- Pixels: screen divided into units called pixels, which can represent text or graphics. More, smaller pixels, higher resolution. On 15" screen, 768 x 1024 resolution, each pixel = . 0116" .28 mm.

- Screen size : is measured diagonally across screen (e.g. 15", 17").
- Dot pitch: min distance between pixels. Determined when monitor manufactured, limits highest resolution of device.

Video Modes

Modern adaptors can operate in large no of video modes. Video mode determines 3 things:

1.Text Versus Graphics: all video cards (except really old ones) can display info in either text or graphical mode.

- In text mode, video information is stored as characters in a character set. Screen is divided into 80 columns by 25 rows.
- Text mode used under DOS, & sometimes during boot sequence. Usually systems run in graphics mode.

2. Resolution: Current standard resolutions are:

- 640x480 (VGA)
- 800x600
- 1024x768
- 1280x1024
- 1600x1200–Recent video cards all support SVGA, but variety of terminology & support variety of resolution/ colour depth combinations.

3. Interlacing : On an interlaced display, the electron beam first scans all the even numbered lines on the raster, then all the odd numbered lines, covering the whole frame in two scans. This doubles the vertical resolution without increasing the scan rate. The cost of this increased resolution is that any given pixel is refreshed only half as often. For an interlaced display to appear flicker-free, it may need a phosphor with a higher persistence.

5. Refresh rate: The image on your computer monitor doesn't just appear fully formed on the screen's phosphors: it's drawn line by line with beams fired from three electron guns at the back of the CRT. The frequency at which they redraw the image is called the refresh rate, and it's an important measure of how steady the image will appear. The amount of times a new picture is displayed per second, expressed in Hz (Hertz), the higher the refresh rate the better the picture will be. The higher the refresh rate the less flicker on the screen. This helps reduce eye strain.

4. Colour Depth: how many bits/pixel. Individual pixels represent either colour or grey scale. Many displays used in manufacturing Palmtops etc are monochrome.

- Each pixel represented by group of bits (2,4,8...). What info could two bits per pixel hold?
- A colour pixel is made up of RGB. If have 24 bits per pixel, gives 8 bits per colour, allowing 16 million diff colours.

Colour Depth

Color Depth	Colors displayed	Common Name
4-Bit	16	Standard VGA
8-Bit	256	256-Color Mode
16-Bit	65,536	High Color
24-Bit	16,777,216	True Color

What should be the minimum refresh rate?

Display Speed

- Speed: Screen images need to be redrawn to reflect changes in output. Video card and monitor together must be able to process image data and display it on screen fast enough to look right to human user.
- Complex moving images require far more processing and faster refresh than static text. Lowest satisfactory refresh rate for moving images is 60 times/second. Processing capacity depends on video adaptor.

Displays

- Graphics adaptors contain processor, memory, and DAC (digital to analog converter). They may be built-in to motherboard, or plugged-in.
- Graphic output produced by programs is stored in video memory (frame buffer) & read from by adaptor, which issues commands to create display.
- High resolution & colour, full screen video & 3D rendering all increase memory & processing requirements => High end cards have
- MBs fast RAM, fast processor, dedicated bus (AGP) to system memory & CPU.

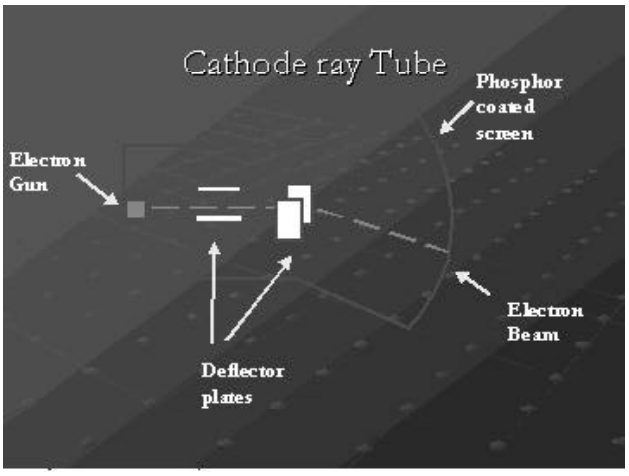
Display Devices

Two main types of display:

a) Video display unit using CRT technology

- Cathode Ray Tube (CRT) uses similar technology to television.
- Electron beam scans across phosphor-coated screen, phosphor emits light.
- Combination of RGB dots makes up each pixel. Intensity of each dot determines colour of the pixel.

Cathode Ray Tube



- A beam moves across screen in fixed path called raster scan.
- Screen refreshed 50-100 times/sec.

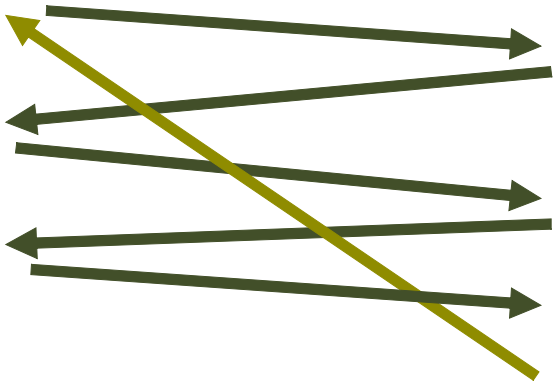
- Interlaced v/s non-interlaced.
- Dot pitch: distance between pixels. Determines precision of image.

Advantages of CRT:

- High contrast,
- Speed

Disadvantages

- size,
- power consumption,
- fragility.



B) Flat Panel Using LCD Technology

- Major advantages small footprint, low power, no EMR.
- Flat Panel displays most common in laptop and notebook computers, but also moving onto desktop.
- Most commonly use LCD technology in which light passes through cells polarized by electric current.
- Three cells (RBG) mix to create colour for each pixel
- Can be active or passive matrix.
- "Passive matrix" (STN) has one transistor per row and column and "scans" - less bright, simpler/cheaper to manufacture. Slow to update (problem for video/ animation). Dual scan version quicker and brighter.
- "Active matrix" (TFT) has one transistor per cell of matrix, is very bright, and more expensive.

There are many ways to classify monitors. The most basic is in terms of color capabilities, which separates monitors into three classes:

· Monochrome: Monochrome monitors actually display two colors, one for the background and one for the foreground. The colors can be black and white, green and black, or amber and black.

· Gray-scale: A gray-scale monitor is a special type of monochrome monitor capable of displaying different shades of gray.

- Color: Color monitors can display anywhere from 16 to over 1 million different colors. Color monitors are sometimes called *RGB monitors* because they accept three separate signals — red, green, and blue.

Following questions should be answered:-

1. Suppose if you want to buy a monitor for your new computer system, what should be the display characteristics you have to keep in mind?

2. What are three different things are determined by video modes ?

3. Write short note on
   - CRT monitor
   - Flat panel monitors

LESSON 12
OUTPUT DEVICES –PRINTERS

Outline

- Types of printers
- Characteristics of printers
- Impact printers
- Line printers
- Dot matrix
- Non Impact printers
- LaserJet
- Ink jet
- Plotters

A device that prints text or illustrations on paper. There are many different types of printers.

For Example

- Daisy Wheel: Similar to a ball-head typewriter, this type of printer has a plastic or metal wheel on which the shape of each character stands out in relief. A hammer presses the wheel against a ribbon, which in turn makes an ink stain in the shape of the character on the paper. Daisy-wheel printers produce letter-quality print but cannot print graphics.
- Dot-Matrix: Creates characters by striking pins against an ink ribbon. Each pin makes a dot, and combinations of dots form characters and illustrations.
- Ink-Jet: Sprays ink at a sheet of paper. Ink-jet printers produce high-quality text and graphics.
- Laser: Uses the same technology as copy machines. Laser printers produce very high quality text and graphics.
- Lcd & Led : Similar to a laser printer, but uses liquid crystals or light-emitting diodes rather than a laser to produce an image on the drum.
- Line Printer: Contains a chain of characters or pins that print an entire line at one time. Line printers are very fast, but produce low-quality print.
- Thermal Printer: An inexpensive printer that works by pushing heated pins against heat-sensitive paper. Thermal printers are widely used in calculators and fax machines.

Classification Of Printers According to Characterstics

Printers are also classified by the following characteristics:

- Quality of Type: The output produced by printers is said to be either letter quality (as good as a typewriter), near letter quality, or draft quality. Only daisy-wheel, ink-jet, and laser printers produce letter-quality type. Some dot-matrix printers claim letter-quality print, but if you look closely, you can see the difference.
- Speed: Measured in characters per second (cps) or pages per minute (ppm), the speed of printers varies widely. Daisy-wheel printers tend to be the slowest, printing about 30 cps. Line printers are fastest (up to 3,000 lines per minute). Dot-matrix printers can print up to 500 cps, and laser printers range from about 4 to 20 text pages per minute.
- Impact or Non-Impact: Impact printers include all printers that work by striking an ink ribbon. Daisy wheel, dot matrix, and line printers are impact printers. Non-impact printers include laser printers and ink-jet printers. The important difference between impact and non-impact printers is that impact printers are much noisier.
- Graphics: Some printers (daisy-wheel and line printers) can print only text. Other printers can print both text and graphics.
- Fonts: Some printers, notably dot-matrix printers, are limited to one or a few fonts. In contrast, laser and ink-jet printers are capable of printing an almost unlimited variety of fonts. Daisy-wheel printers can also print different fonts, but you need to change the daisy wheel, making it difficult to mix fonts in the same document.

What are the characteristics of Impact Printers ?

Impact Printers

- Produce image by impact of paper, ribbon and impact hammer
  - Line Printers (Typical, Daisy Wheel)
    - Fast
    - Character only!
  - Dot matrix
    - Impact hammer made up of tiny pins
    - Capable of crude graphics

Name any two types of Impact Printers.

a) Dot Matrix Printers

Dot Matrix printers form each character from a pattern of ink dots. The dots are printed by a print head containing a line of seven or nine hammers that can be struck against a ribbon. Inside the printer, a ROM chip stores the pattern of ink dots corresponding to each character. When the printer receives the character code for a letter, the ROM chip tells the print head which hammers to strike on the ribbon. A belt driven by a motor moves the print head back and forth as the paper is wound forward.

Dot matrix and high speed line impact printers are the only game in town for printing multi-part forms. The non-impact printers, such as laser printers and thermal printers, cannot make multiple copies in one pass like dot matrix printers and high speed line printers

Dot matrix printers remain an important part of back-office operations. Dot matrix printers are ideal for low volume impact printing. Like their high speed line counterparts, dot matrix printers offer the ability to print multi-part forms in a single pass.

b) Line printers

High-speed line matrix printers use inexpensive ribbons and are great for long, unattended print runs. These printers are highly robust and print comfortably at high speeds for long periods of time.

Commonly used in industrial applications, high-speed line printers last for many years and are ideally suited to factory and loading dock settings. High-speed line matrix printers are superior in handling complex forms, and they can print on a broad range of materials.

What are the characteristics of non Impact Printers ?

Non-impact Printers

- Can't do multi-part forms
- Laser Printers
  - Laser based, like copy machines
- Inkjet printers
  - Spray ink on paper
- Higher resolution
- Good for graphics
- Plotters
  - Automated drawing tool
  - Moves pen and/or per as per computer commands

Name Any Two Non-impact Printers

a) Ink Jet Printers

Ink jet printers work as their name suggests, by squirting jets of ink onto the paper through tiny nozzles. Ink jets provide good quality output at reasonable speed and at relatively low cost.

b) Laser Printers

A laser printer is like a photocopier attached to the computer as an output device. Laser printers are capable of high quality output, at a fast speed but at a relatively high cost.

Laser printers remain the fastest, most efficient way to reproduce text and images, as laser printers offer high resolution and quick speeds at a relatively low cost.

There are three main categories of laser printers:

- Color laser printers
- Continuous form laser printers
- Cut sheet monochrome laser printers.

Which laser printer will best suit your needs depends upon the requirements of your print job.

Continuous Form Laser Printers

Print Multi Part Forms with Dot Matrix Printers

While laser printers have become the office standard for letters and proposals, dot matrix printers remain an important part of back-office operations in many firms. Dot matrix printers are commonly used for printing invoices, purchase orders, shipping forms, labels, and other multi-part forms, as dot matrix printers are able to print through multi-part forms in a single pass.

They are best suited to production print runs that require laser quality output. Continuous form laser printers can also provide a straight paper path with pin-feed tractors that give highly accurate paper control. Continuous form laser printers are superior for applications that require tight forms registration. And the power stacking of continuous forms laser printers make them suitable for unattended production runs, where they are dedicated to specific types of print jobs.

What Do You Think Which One is Better Laser or Dot Matrix?

Cut Sheet Monochrome Laser Printers

Cut sheet laser printers use friction feed to advance the paper, suitable for desktop printers that run a variety of media, such as letterhead, envelopes and transparencies. Cut sheet laser printers are suitable for most print jobs shared in an office environment.

What Are Plotters?

Plotters

A plotter is a large-format printer capable of printing very large sheets for drawings, maps, posters and other applications.

Following questions should be answered: -

1. Give some characteristics according to which printers are classified.

2. Give the classification of printers.

3. Write short note on
   a) Impact printers

   b) Non impact printers

4. Discuss the features of laser printer and dot matrix printer

5. What are plotters ?

LESSON 13
COMPUTERS SELECTION

Outline

- Recommended and Minimum Requirements
- Which computer to buy
- Recommended Minimum Specification: Windows PC
- Buying Considerations

Whether you are buying a new computer to use on the campus network, or using a computer that you already own, you'll want a machine that suits your needs and works well in this environment.

In the past, Computing Services has conducted surveys on student computer ownership. If understanding past trends (PC vs. Mac, desktop vs. laptop)

Recommended And Minimum Requirements

A lower-powered computer may suffice for writing papers, but if you need to use more specialized software, that same machine will most likely be ineffective.

Depending on your major, the following is a sample of some common course software packages that you may be using. Remember that departmental academic software requirements may vary; contact your department for information specific to your major.

- Programming Software (Visual Basic, Visual C++, etc.)
- Graphics Software (PhotoShop, Fractal Painter, Illustrator, etc.)
- Standard Productivity Software (Word, Excel, PowerPoint, PageMaker, etc.)
- Communications Software (mail, telnet, ftp, Netscape, etc.)

Which Computer to Buy?

Before you purchase a system, we'd urge you to examine how you'd use that computer. Your work habits, possible fields of study, and interests should shape your decision.

- Decide between Apple Macintosh and Windows systems: Which platform matches your interests? In general, DOS/ Windows systems are less expensive for similar processing speed. Macintoshes can be easier to set up, to learn, and to use.
- Decide whether a portable computer or a desktop computer is more appropriate for your work habits. Desktop computers are generally less expensive but require that you do your work in one location. Laptop computers allow students to compute on the go, connecting to the campus network from a growing number of network jacks in the Morris Library, classrooms, study areas, computing sites, and other locations.

If your courses, activities, and research involve a lot of travel or fieldwork, you may also find a portable computer to be a good investment.

- Decide on a budgeting strategy, but plan for the long term. You could buy a powerful computer now or buy an inexpensive computer now and plan to upgrade it later on. Some vendors even have leasing programs that allow you to upgrade your computer every 2-3 years.
- Budget for software as well as hardware. Word processing and spreadsheet software can be purchased with most systems; however, you may also need other software for communications, graphics, and specialized projects. You should be aware that unauthorized copying of licensed software is illegal.

If you are planning on buying a computer to use for the next three or more years, we would recommend that the computer you select meet the following minimum specifications:

Recommended Minimum Specification:

Windows Pc

Desktop: Windows PC	Desktop: Macintosh
Intel Pentium 4 or work alike, 2.0 GHz Processor	PowerPC G4, 800 MHz Processor
Microsoft Windows XP Professional	Mac OS X 10.2.X w/Classic
Microsoft Office XP Standard *or* Professional	Microsoft Office v.X for Mac
256 MB of RAM	256 MB of RAM
30 GB Hard Disk	30 GB Hard Disk
17" SVGA CRT or 15" LCD Monitor	17" Studio Display
32X CD-R/W Drive	32X CD-R/W Drive
101-key keyboard	Apple USB Keyboard
Microsoft or Logitech Mouse	Apple USB Mouse
Sound card & powered speakers	Built-in 16-bit sound, speakers
Highly-Recommended Accessories: USB Inkjet Printer, USB A to B Cable, Surge Suppressor	

Buying Considerations

Purchasing a Computer

Do you want to buy a computer? If yes but before buying you have to keep following things in your mind

Hardware Considerations

It is easy to get confused when shopping for a computer. . . too many numbers and acronyms to make sense of. Here are a few general rules of thumb when shopping:

Larger numbers on the computer specifications will usually mean larger numbers that begin with $. Buy the largest that you can afford. Specifically, look for the following as minimums when purchasing a $1000 system:

15-inch or 17-inch monitor. These are becoming standard on computer systems

266 MHz or higher processor speed. These numbers are indicators of how fast the computer will be able to complete instructions. The higher the speed, the better. At the time this document was written, 500 MHz is the new industry benchmark. Processor speed is important if you want your computer to run really fast.

Pentium or Pentium II Processor. These are the standard CPU's installed on most Windows-based computer systems. If you are thinking of purchasing a Macintosh, take a serious look at the G3 family of products. They all use the latest, fasted CPU from the Motorola group. Also, if cost is a factor and you want to stay within a budget, you might consider buying a product with the Celeron CPU. This is a cheaper product manufactured by Intel to compete in the $1000-and-under computer retail market. Buy the best that you can afford.

What is this CELERON?

6.0 or 8.0 GB Hard Disk Drive

You can purchase disk drives that store gigabytes up in the teens. You might be able to purchase a computer with a smaller hard drive (3.2 to 4.1GB) and it may be enough for your needs. If the cost difference between the smaller drives and the larger ones is not too great, go for the bigger one.

32 MB RAM

Extra RAM is one of the best investments you can make for your computer. If you can afford to get one with 64MB or even 128MB, go ahead and get it. You'll never regret having more RAM in your computer, but you'll curse and stamp your foot if you have too little.

Color Inkjet Printer

Color inkjet printers are being marketed for unbelievably low prices. In fact, you can buy a really basic Lexmark for under $100. Be aware that the ink cartridges are very expensive ($25 to $40). A standard black cartridge will do about one ream (500 sheets) of paper if you do only text printing. Many color inkjets require a black cartridge and a color cartridge. This can get into your pocket quickly. Learning to use the Print Preview function in most applications can save you a bundle in ink costs!

Software Considerations

Most computers have an operating system installed when you purchase them. The most recent operating system installed on Intel-based machines is Windows 98. You can also buy computers that have Windows 95 or Windows NT installed instead (WinNT will cost you more). On the Macintosh side of the world, all current Macs come with MacOS8 or a variant. You don't have much latitude when you purchase their products. Some computer distributors include additional software as a "bundle" when you purchase their product. Here is a recommended list of generic software that you might want start with:

Word Processor

There are numerous word processors available. Windows 95/98/NT comes with a nice little package preinstalled called "WordPad". It is a scaled-down version of Microsoft's Word, a very powerful application. If you don't require spell checking, grammar checking, tables, columns, and the like, you might be able to get by with WordPad. If you need something better, check our Microsoft's Word or Corel's WordPerfect.

Internet Software

Many computers ship with Microsoft's Internet Explorer and/or Netscape's Navigator. Both will allow you to explore the World Wide Web. Both packages also include an Email program. These are two of the most useful Internet applications.

Depending on your computing needs, you might want to acquire a database/spreadsheet application, some games, financial packages, encyclopedias, and other useful applications. A good all-purpose program that you might give consideration to purchasing is Microsoft Works (or its competitor, Apple works). If you wish to explore the world of desktop publishing, consider Microsoft Publisher'98. Both Works and Publisher are under $100 at your friendly retail software store.

Buying Considerations	Additional Features
	Printer
Hardware	Scanner
Additional features	Network
Software	Digital Camera
Where to put the computer	Game controller
Problems and Support	
Computer Hardware Options	Computer Software Options
CPU	Operating System
Memory	Office Suite
Hard Disk	Anti-virus and Firewall
CD, DVD Drives	Desktop Publishing
Floppy Drives	Photo Editing
Video Card	Printshop/Creativity
Monitor	Basic Bookkeeping
Sound Card and Speakers	
Modem	Computer System Recommendations
Keyboard & Mouse	Desktop Systems
	Notebooks and Laptops

Following questions should be answered: -

1. Give the complete configuration of your computer system

2. What is the minimum specification to buy a new computer ?

3. What features should be kept in mind when you go to buy a new computer ?

LESSON 14
GROUP DISCUSSION
TOPIC – EVALUATE THE PERFORMANCE OF YOUR COMPUTER SYSTEM

Note:

LESSON -15
OPERATING SYSTEM& ITS FUNCTIONS

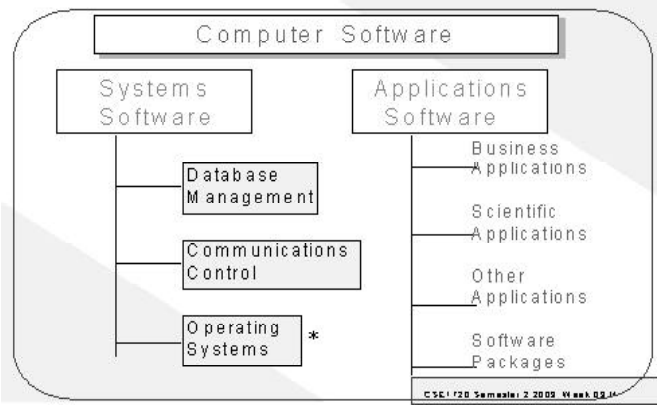
Outline

- Types of software
- Operating System-System software
- Define: Operating System
- OS-Resource Manager
- Functions of OS

What do you think operating system is a software or hardware. Operating System is a software.

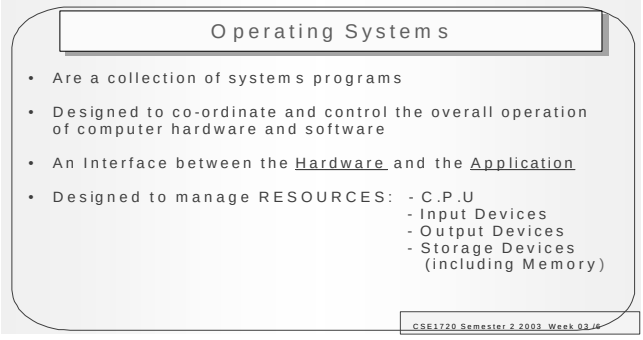
Software are of Two types

- System Software
- Application Software



An operating system is an important part of almost every computer system. The goal of this chapter is to introduce concepts related to operating systems and to show how this particular system software helps in the overall operating and usage of a computer system.

What Is An Operating System?



An operating system (often referred to as OS) is an integrated set of programs that controls the resources (the CPU, memory, I/O devices, etc.) of a computer system and provides its users with an interface or virtual machine that is more convenient to use than the bare machine. According to this definition, the two primary objectives of an operating system are as follows:

Control Hardware

An operating system controls the different parts of a computer system and enables all the parts to work together.

Run Application Software

An operating system runs application software, such as Office 2000 and Lotus Notes R5

Manage Information

An operating system provides ways to manage and organize information stored on a computer. You can use an operating system to sort, copy, move, delete or view files.

- Making a computer system convenient to use. A computer system consists of one or more processors, main memory and nay types of I/O devices such as disks, tapes, terminals, network interfaces, etc. Writing programs for using these hardware resources correctly and efficiently is an extremely difficult job, requiring in-depth knowledge of the functioning of the resources. Hence to make computer system usable by a large number of user's it became clear several years ago that some way is required to shield programmers from the complexity of the hardware resources. The gradually evolved solution to handle this problem is to put a layer of software on top of the bare hardware, to manage all parts of the system, and present the user with an interface or virtual machine that is easier to program and use. This layer of software is the operating system.

The logical architecture of a computer system is shown if Figure 1 As shown in the figure, the hardware resources are surrounded by the operating system layer, which in turn is surrounded by a layer of other system software (such as compilers, editors, command interpreter, utilities, etc.) and a set of application programs (such as commercial data processing applications, scientific and engineering applications, entertainment and education applications, etc.). Finally, the end users view the computer system in terms of he user interface provided by the application programs.

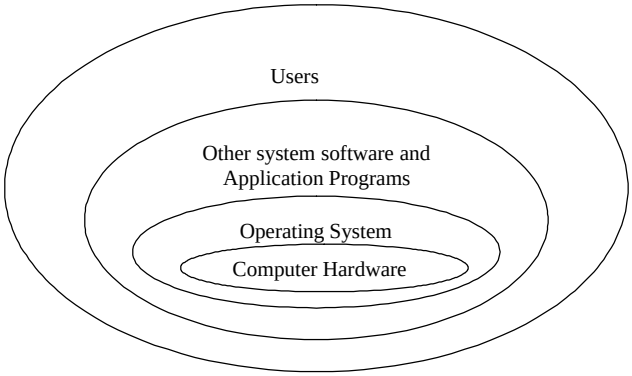
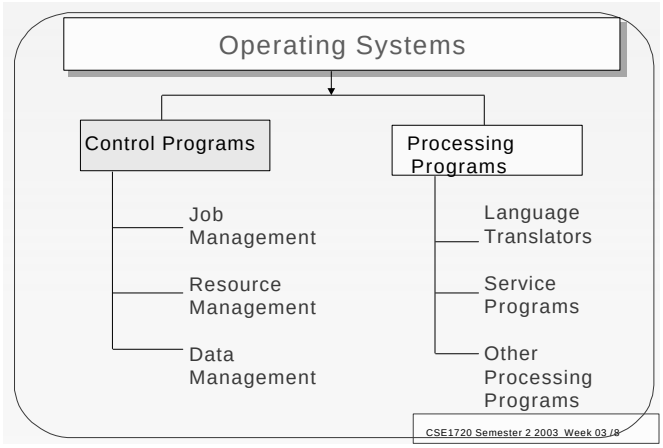


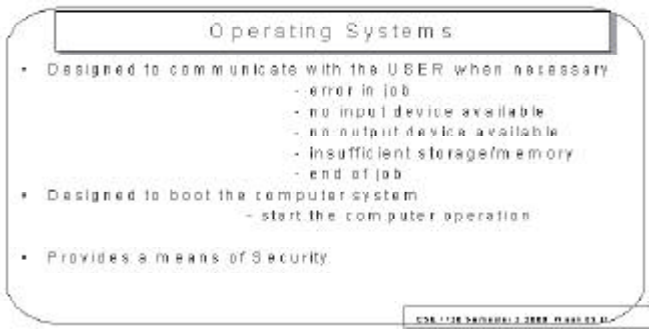
Figure-1 Logical architecture of a computer system.



The operating system layer hides the details of the hardware from the programmer and provides the programmer with a convenient interface for using the system.

The operating layer proved various facility and services that make the use of he hardware resources convenient, efficient, and safe. A programmer makes use of these facilities in developing an application, and the application while it is running, invokes the required services to perform certain functions. In effect, the operating system hide the details of the hardware form the programmer and provide the programmer with a convenient interface for using the system. In acts as an intermediary between the hardware and is user, providing a high-level interface to low-level hardware resources, and making it easier for the programmer and for application programs to access and use those resources.

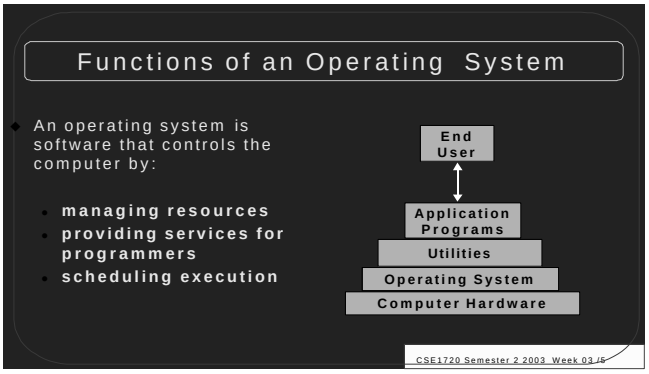
Why Operating System Is Known As A Resource Manager?



Managing the resources of a computer system. The second important objective of an operating system is to manage the various resources of the compute system. This involves performing such tasks as keeping track of who is using which resource, granting resource requests, accounting for resources usage, and mediating conflicting request from different programs and users.

Executing a hob on a computer system often requires several of its resources such as CPU time, memory space, file storage space I/O devices, and so on. The operating system acts as the manager of the various resources of a compute system and allocates them to specific programs and user to execute their jobs successfully. When a computer system is used to simultaneously handle several applications, there may be many, possibly confliction, requests for resources. In such a situation, the operating system must decide which request are allocated resources to operate the compute system efficiently and fairly (providing due attention to all users.). The efficient and fair sharing of resources among users and / or programs I s a key goal of most operating systems.

Main Function Of An Operating System



We saw that an operating system provides certain services to programs and to the use of those programs. The specific services provided will, of course, differ from one operating system to another, but there are some common types of function that we can identify. The main functions provided by most operating systems of today are as follows :

- Process Management. A process is a program in execution. During execution, a process needs certain resources such as CPU time, memory space, files, and I/O devices. A t a particular instance of time, a computer system normally consists of a collection of processes. The

process management module of an operating system takes care of the creation and deletion of processes, scheduling of various system resources to the different processes requesting them, and providing mechanism for synchronization and communication among processes.

- Memory Management. To execute a program, it must be loaded, together with the data it accesses, in the main memory (at lest partially). To improve CPU utilization and to provide better response time to its user, a compute system normally keeps several programs in main memory. The memory management module of an operating system takes care of the allocation and deal location of memory space to the various programs in need of this resource.
- File Management. All computer systems are used for storage, retrieval and sharing of information. A computer normally stores such information in units called files. Processes can read information from files and can create new files for storing newly generated information. Information stored in files is made persistent by storing them on a secondary storage media such as a magnetic disk. Files provide a natural and easy means of information sharing. That is, a file can be created by one application and then shared with different applications at a later time. The file management module of an operating system takes care of file-related activities such as organization, storing, retrieval, naming, sharing and protection of files. It allows programs to use a set of operations that allocation and layout of the secondary storage devices.
- Device Management. A computer system normally consists of several I/O devices such as terminal, printer, disk, and tape. The device management module of an operating system takes care of controlling all the computer's I/O devices. It keeps track of I/O request from process, issues commands to the I/O devices, and ensures correct data transmission to/from and I/O devices. It also provide an interface between the devices and the rest of the system that is simple and easy to use. Often, this interface is device independent, that is, the interface is same for all types of I/O devices.
- Security. Computer systems often store large amount of information, some of which is highly sensitive and valuable to their users. Users can trust the system and rely on it only if he various resources and information of a computer system are protected against destruction and unauthorized access. The security module of an operating system ensures this. This module also ensures that when several disjoint processes are being executed simultaneously, one process does not interfere with the others, or with the operating system itself.
- Command Interpretation. A user communicates with the operating system, for using the various systems resources, via a set of command provided by the operating system. The operating system also provides a simple language, known as command language (CL) or job control language (JCL), using which a user can put several commands together from the command set to describe the resource requirements of the job. The command interpretation module of an operating system take care of interpreting user command, supplied individually or in the form of command language, and directing interpreting user command, supplied individually or in the form of command language, and directing interpreting user commands, supplied individually or in the form of command language, and directing the system resources to handle the request. With this mode of interaction with system, the user is usually not too concerned with the hardware details of the system, or with how the operating system will direct the hardware to handle certain request.

In addition to the above listed major functions, operating systems performs few other functions such as keeping an account of which user (or processes) use how much and what kinds of computer resources, maintenance of log of system usage by all users, and maintenance of internal time clock.

Following Questions should be answered:-

1. Why OS is necessary for a computer system

2. Why OS is known as resource manger? What type of resources it handles?

3. Draw a logical architecture diagram of a computer system. Explain how an operating system helps in meeting these objectives.

4. "The operating system tends to isolate the hardware from the users ". Discuss this Statement.

5. List the various functions normally performed by an operating system.

LESSON 16
OPERATING SYSTEM STRUCTURE & COMMAND INTERPRETATION

Outline

- Structure
  - Layered Structure
  - Kernel
  - Monolithic & Micro kernel
- Resident & Non-Resident OS modules
- Command Interpretation
  - Command line interface
  - Graphical User interface
- Real Time Operating System
- Network and distributed system
- Measuring performance of computer system

Structure Of Os

Layered Structure

Most modern operating system organize their components into a number of layers (levels), each built on top of lower layers. The bottom layer (layer o) is the hardware, and the highest layer (layer n) is the user interface. The number of in-between layers and their contents vary from one operating system to another. How this is decided is explained below.

The main advantage of the layered approach is modularity. The layers are selected such that each layer uses the functions and services provided by its immediate lower layer. This approach greatly simplifies the design and implementation of the system because each layer is implemented using only those operations provided by its immediate lower level layer.

Kernel

The kernel of an operating system is its central controlling part that implements the most primitive of the system's functions. It is the only part of an operating system that a user cannot replace or modify. The precise interpretation of the system's function that are part of the kernel varies from one operating system to another. However, typical operating system kernels contain basic functions that are required for process management, memory management, device management, and low-level security features like access control. In some systems, the kernel is larger and provides for more than these functions, whereas in others, it is smaller.

Monolithic Kernel versus Micro kernel

The two commonly used models for kernel design is operating systems are the monolithic kernel and the micro kernel. In the monolithic kernel model, most operating system services such as process management, memory management, device management, the kernel provides file management, and security. As a result, the kernel has a large, monolithic structure.

On the other hand, in the micro kernel model, the main goal is to keep the kernel as small s possible. Therefore, in this model, the kernel is very small nucleus of software that provides only the minimal facilities necessary for implementing additional operating system services. The only services provided by the kernel in this model are low-level device management, a limited amount of low-level process management, and some memory management. All other operating system services, such as file management, additional process and memory management activities, and security are implemented as user-level server process. Each server process has its own address space and can be programmed separately.

Resident and Non-Resident Operating System Modules

With all the functionalities of an operating system implemented, it becomes large software. Obviously, all the functionalities of an operating system are not needed all the time. As the main memory capacity of a system is limited, it is customary to always keep in the system's memory only a very small part of the operating system and to keep its remaining part of an on-line storage device such as hard disk. Those modules of an operating system that are always kept in the system's main memory are called resident modules and those that are kept on hard disk care called non-resident modules. The non-resident modules are loaded into the memory on demand, that is, as and when they are needed for execution.

The system kernel should not be confused with the resident modules of the operating system. The two are not necessarily the same. In fact, for most operating system they are different. The following two criteria normally determine whether a particular operating system module should be resident:

Its frequency use, and Whether the system can operate at all without it.

For example, file directories can be maintained on disk and loaded into the memory when required. Status information for inactive processes can similarly be swapped out on disk. In fact, the resident part of an operating system is a subset of its kernel.

Command Interpretation

The command interpretation module (known as command interpreter) of an operating system serves as an interface for the user to communicate with the computer via its operating system. It provides a set of commands using which the user can give instructions to the computer for getting some job done by it. The commands supported by the command interpretation module are known as system calls. When a user gives instructions to the computer by using these system calls, the command interpreter takes care of interpreting these commands and directing the system resources to handle the requests. Thus the command interpreter hides the hardware details of the

system from the user. In this manner, it greatly contributes to the 'ease of use' objective of an operating system.

he two broad categories of user interfaces supported by various operating systems are command-line interface and graphical user interface. They are briefly described below.

Command-line Interface.

This is the textual user interface in which the user gives instruction to the computer by typing command. That is, to enter a command, the user uses the keyboard to type words and symbols. For example in Unix, the user has to type "del report.txt" to delete the file named report.txt. If the user types a command incorrectly, the command interpreter will respond with a message indicating that it did not understand the command. When this happens the user has to jut retype the command correctly.

There is no problem in typing simple commands as illustrated in the example above. However, users often need to give detail instructions to the computer about their jobs when they submit, them for execution. For example, while submitting a job for execution, a user may need to supply the following information.

- His/her identification for security and accounting purpose.
- The software and hardware resource requirements of the job.
- The I/O device requirements of the job.
- The directories in which the data files to be used by the job are stored.
- The action, which the system should take in exceptional conditions, such as when the input data is missing or incorrect, o r when an I/O device malfunctions.

To facilitate this, system, which support command-line interface, also support some type of command language (CL) or job-control language (JCL). Users can write codes in the JCL to give instructions to the systems. These coded statements tell the operating system such things as the name of the job, the user's name and account number, the I/O devices to use during processing, the compiler to use if language translation is needed, where to find the data files, and so on. The command interpreter is designed to interpret the codes written in JCL and invoke appropriate system actions.

Graphical User Interface

User interface plays an important role in ease of use. The more intuitive the user interface of a system is, the easier it is to use. An intuitive user interface allows a user to use the system effectively even if he/she has never seen it before. That is, how the user interface works is very obvious to any user. Such an user interface is called user friendly.

Graphical user Interface or (GUI) (pronounced "gooey") is a user-friendly interface that is much easier to learn and use than a command-line interface. Unlike command-line interface in which command are textual, GUI commands are graphical (pictorial). A GUI provides to the user a screen full of graphic icons (small images on the screen) or menus and allows the user to make a rapid selection form the displayed icons or menus to give instructions to the computer. A point-and-draw device is normally used to rapidly point to and select a particular graphic icon or menu item form the multiple options displayed on the screen. For example, we saw that in a system that uses command-line interface, to delete a file named report.txt we need to type a command like "del report.txt." However, in a system that uses GUI, simply using a mouse to drag the icon that represents the file until it is superimposed on an icon shaped like a trash can perform the same operation. Then releasing the mouse button that was used for dragging the file, causes the file to diaper into the bulging trash can.

Shell

The command interpreter of an operating system that serves as its user interface is often referred to as the shell because it forms the outer layer of an operating system covering the other modules of the operating system. The shell be a command-line interface or GUI. It is also possible to have many different shells for the same operating system. For example, an operating system can have a command-line shell as well as a GUI shell allowing the users to switch from one shell to another. Beginners often find it convenient to work with GUI shell, whereas advanced users find it more convenient to work with command-line shell.. Some operating systems also support multiple command-line shells. It is like supporting multiple JCLs and allowing the user to use a language that he/she is most comfortable with. For example, many Unix operating systems support three command-line shells known as C shell, Bourne shell and Korn shell.

Real-time Operating Systems

For many application, successful processing means that the application produces correct results when its processing completes. However, in real life, we often come across situation in which it is not only sufficient to obtain correct results, but the results must be produced within some specified time frame (deadline) to be useful.

A few examples of such application are:

- An aircraft must process accelerometer data within a certain period (say every 20 milliseconds) that depends on the specifications of the aircraft. Failure to do so could causes the aircraft to go away from its right course or may even cause it to crash.
- Failure to respond in time to an error condition in a nuclear reactor thermal power plant could result in a melt-down.
- Failure to respond in time to an error condition in the assembly line of an automated factory could result in several products units that will have to be ultimately discarded.
- A request for booking a ticket in a computerized railway reservation system must be processed within the passenger's perception of a reasonable time (say 3 minutes).

Systems that are designed for processing such application are called real-time system. That is, a real-time system is a system that must satisfy the requirement of producing the desired results before a certain deadline. If it delivers the correct result, but after the deadline, the system is regarded as having failed.

Thus timely production of the result of processing is as important as its correctness in real-time systems.

Operating system that are designed to handle the processing requirements of real-time systems are called real-time operating systems. The goals of CPU scheduling in traditional timesharing computer system are optimal throughout, optimal resource utilization, and fairness. In contrast, the main goal of CPU scheduling in real-time systems is to allow as many time-critical processes as possible to be processed in time to meet their deadlines.

Network And Distributed Operating Systems

Over the past few decades, advancements in microelectronic technology have resulted in the availability of fast, inexpensive computer, and advancements in communication technology have resulted in the availability of cost-effective and highly efficient computer networks. The net result of the advancements in these two technologies is that the price-performance ratio has now changed to favor the use of interconnected multiple computers in place of single, high-speed computer. Computer systems consisting of interconnected, multiple computers are called distributed computing systems. That is, a distributed computing system is a collection of computer systems interconnected by a communication network in which each computer has its own local memory and other peripheral devices, and the communication between any two computers of the system place by passing message over the communication network.

The operating system commonly used for distributed computing systems can be broadly classified into two types –network operating systems and distributed operating systems. The three most important features commonly used to differentiate between these two types of operating systems are system image, autonomy, and fault tolerance capability. These features are explained below.

1. System Image.The most important feature used to differentiate between the two types of operating system is the image of the distributed computing system from the point of view its users. In case of a network operating system, the users view the distributed computing system as a collection of distinct machines connected by a communication subsystem. That is the users are aware of the fact that multiple computers are being used. On the other hand, a distributed operating system hides the that multiple computers are being used. On the other hand, a distributed operating system hides the existence of multiple computers and provides a single-system image to its users. That is, it makes a collection of networked machines appear to its user as a virtual uniprocessor by providing similar type of user interface as provided by centralized operating systems.
2. Autonomy. In a network operating system, each computer of the distributed computing system has this own local operating system (the operating system of different computes may be the same or different), and there is essentially no coordination at all among the computers except for the rule that when two processes of different computers communicate with each other, they must use a mutually agreed on communication protocol. Each computer functions independently of other computers in the sense that each one makes independent decisions about the creation and termination of their own processes and management of local resources. Notice that due to the possibility of difference in local operating system, the system calls for different computers of the same distributed computing system may be different in this case.

   On the other hand, with a distributed operating system, there is a single system-wide operating system and each computer of the distributed computing system runs a part of this global operating system. The distributed operating system tightly interweaves all the computers of the distributed computing system in the sense that hey work in close cooperation with each other for the efficient and effective utilization of he various resources of the system. That is, processes and several resources are managed globally (some resources are managed locally). Moreover, there is a single set of globally valid system calls available on all computers of the distributed computing system.
3. Fault tolerance capability. A network operating system provided little or no fault tolerance capability in the sense that if 10% of the machines of the entire distributed computing system are down at any moment, at least 10% of the users are unable to continue with their work. On the other hand, with a distributed computing system. Therefore, the fault tolerance capability of a distributed operating system is usually very high as compared to that of a network operating system.

Difference Between Network Os And Distributed OS

The efficiency of an operating system and the overall performance of a computer system is usually measured in terms of the following :

Measuring Performance Of Computer System

1. Throughput. Throughput is the amount of work that the system is able to do per unit time. It is measured as the number of process that are completed by the system per unit time. For example, if n processes are completed in an interval of t seconds, the throughput is taken as n/t processes per second during that interval. Throughput is normally measured in processes/hour. Note that the value of throughput does not depend only on the capability of a system, but also on the nature of jobs being processed by the system. For long processes, throughput may be one

processor/hour; and for short processed, throughput may be 100 processes/hour.

2. Turnaround Time. From the point of view of an individual user, and important criterion is how long it takes the system to complete a bob submitted by him/her. Turnaround time is the interval from the time of submission of a job to the system for processing to the time of completion of the job. Although higher throughput is desirable from the point of view of overall system performance, individual users are more interested in better turnaround time for their jobs.

   Response time. Turnaround time is usually not a suitable measure for interactive system, because in an interactive system a process can produce some output fairly early during its execution and can continue executing while previous results are being output to the user. Thus another measure used in case of interactive systems is response time, which is the interval from the time of submission of a job to the

3. System for processing to the time the first response for the job is produced by the system.

In any computer system, it is desirable to maximize throughput and tom minimize turnaround time and response.

LESSON 17
PROCESS MANAGEMENT-I

Outline

- Process Management
- Process management early system
- Multiprogramming
- Multitasking
- Multithreading
- Multiprocessing
- Time sharing

Process Management

A process (also called job) is a program in execution. The main objective of the process management module of an operating system is to manage the processes submitted to the system in such a manner so as to minimize the idle time of the various processors (CPU, I/O processors, etc.) of the computer system. In this section, we will learn about some of he mechanisms commonly used in modern operating systems to achieve this objective. We will also see how these mechanisms have gradually evolved from the early days of computers.

Process Management Early Sysems

In early computer systems, a job was typically executed in the following manner:

- A programmer would first write the program on paper.
- It was then punched on cards or paper tape containing the program and data was then submitted at the reception counter of the computer center.
- The deck of cards or the paper tape containing the program and data was then submitted at the reception counter of the computer center.
- An operator would then take the card deck or paper tape and manually load it into the system form card reader to paper tape reader. The operator was also responsible for loading any other software resource (such as a language compiler) or setting hardware devices required for the execution of the job. Before loading of the job, the operator had to use the front panel switches of he computer system to clear the main memory to remove any data remaining from the previous job.
- The operator would then set the appropriate switches in the form panel to run the job.
- The result of execution of the job was then printed on the printer, which was brought by the operator to the reception counter, so that the programmer could collect it later.

The same process had to be repeated for each and every job to be executed by the computer. This method of job execution was known as the manual loading manual loading mechanism because the jobs had to be manually loaded one after another by the computer operator in the computer system. Notice that in this method, job-to-job transition was not automatic. The manual transition from one job to another caused lot o computer time o be wasted sine the computer remained idle while the operator loaded and unloaded jobs and prepared the system for a new job. In order to reduce this idle time to the computer, a method of automatic job-to-job transition was devised. In this method, known as batch processing, when one job is finished, the system control is automatically transferred back to the operating system which automatically performs the housekeeping jobs (such as clearing the memory to remove any data remaining from the previous job) needed to load and run the next job. In case of batch processing systems, jobs were typically executed in the following manner.

- Programmers would prepare their program and data card decks or paper tapes and submitted them at the reception counter of he computer center.
- The operator could periodically collect all the submitted programs and would batch them together and then load them all into the input device of the system at one time.
- The operator would then give a command to the system to start executing the jobs.
- The jobs were then automatically loaded from the input device and executed by the system one-by-one without any operator intervention. That is, the system would read the first job from the input device, execute it, print out its result on the printer, and then repeat these steps fro each subsequent job till all the jobs in the submitted batch of jobs were over.
- When all the jobs in the submitted batch were processed, the operator would separate the printed output for each job and keep them at the reception counter so that the programmers could collect them later.

The batch processing mechanism helped in reducing the idle time of a computer system to a great extent because transition form on job to another did not require any operator intervention. Another major time saving made possible by batch processing was reduction is set-up time by batching of similar jobs together by the operator.

For example, if all FORTRAN compilation jobs are batched together, the system needs to load the FORTRAN compiler only once for processing all these jobs.

The obvious question is how the computer separates one job from another from a batch of jobs for automatic job-to-job transition. Moreover, how does the system know which compiler or what hardware devices are to be used by a particular job when there is no operator intervention? To facilitate all these, control statements and job control languages (JCLs) were introduced along with the concept of batch processing. The control statement are used by the operating system to identify a

new job and to determine what resources are needed by the job during its execution. These control statements are written in a language known as the job control language. Usually every program and data sets are preceded and followed by JCL statements.

When a program is prepared for a computer run, it is necessary to prepare the job control statements and place them in proper order along with the program, before the program is fed to the computer system. Thus, each program has, besides the program itself; a set of instructions called JCL instructions, which instruct the operating system on the identity and requirements of the job. JCL statements tell the operating system things such as the name of the job, the user's name and account number, the I/O devices to be used during processing, the assembler or compiler to be used if language translation is required, and so on. Figure 10.2 shows a sample deck of cards prepared for the compilation and execution of a COBOL program in a batch processing system.

Notice from the figure that in order to distinguish a control card form data or program cards, a special character of pattern on the card identifies them. Several systems used the dollar-sign character ($) in the first column to identify a control card, while IBM's JCL used slash marks (//) in the first two columns. Other systems used some other character or codes. The basic idea in selection of these special characters or code was that no program or data card should have these characters or code.

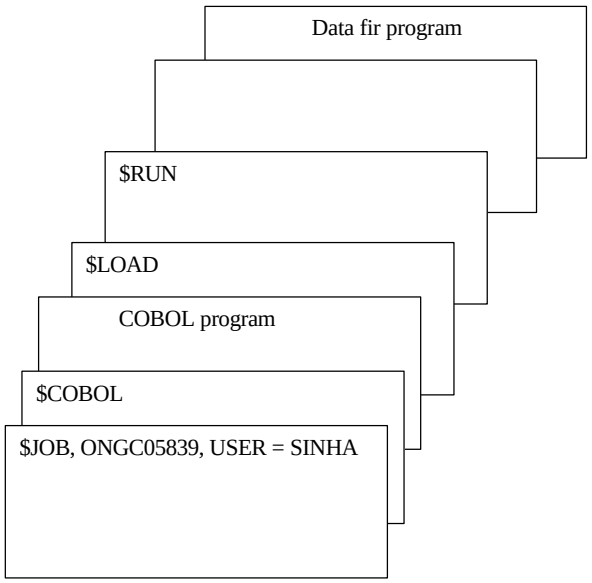


Figure 2. Illustrating the use of job control statements and the structure of a sample deck of cards submitted for processing in a batch processing system.

Multiprogramming

In case of both manual loading and batch processing of jobs, the jobs are loaded into the system and processed one at a time. That is, once loaded only after the completion of the previous job. As shown in Figure 3, in such a situation the job which is currently loaded and is being executed will be the sole occupant of the user's area of the main memory (the operating system always resides in a part of the main memory) and it will have the CPU exclusive available for itself. The situation shown in Figure 3 is that of a uniprogramming system in which only one job is processed by the system at a time and all the system resource are exclusively available for the job until it completes.

It was observed that a job does not need the CPU for the entire duration of its processing. This is because in addition to doing computation (for which CPU is needed), a job often needs to perform I/O operations (such as reading or writing some data to a tape or disk, waiting for some data to be input from the keyboard, and printing some results) during the course of its processing. In fact, depending on the CPU utilization during the course of processing, jobs are broadly classified into the following two types.

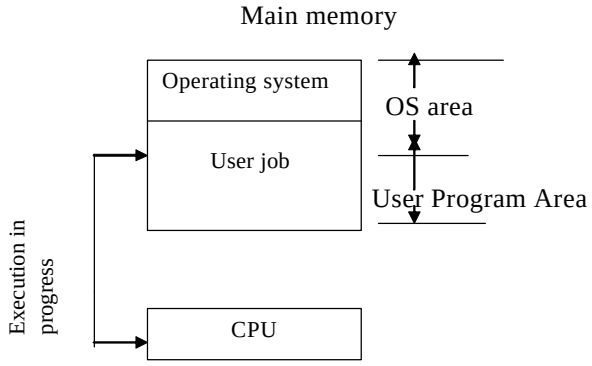


Figure 3. A uniprogramming system model

- CPU-bound jobs. These jobs mostly perform numerical calculations, with little I/O operations. They are so called because they heavily utilize the CPU during the course of their processing programs used for scientific and engineering computations usually fall in this category of jobs.
- I/O-bound jobs. These jobs normally input vast amount of data, perform very little computation, and output large amount of information. They are so called because during the course of their processing, their CPU utilization is very low and most of the time they perform I/O operations. Programs used for commercial data processing application usually fall in this category of jobs.

In a uniprogramming system, the CPU will be idle whenever the job that is currently being processed by the system performs I/O operations. With CPU-bound jobs, the CPU idle time may not be significant, but for I/O bound jobs the CPU may be idle 80 to 90% of the time. Moreover, since I/O devices are slower than the CPU by 20 to 100 times, the CPU idle time is significant even for CPU-bound jobs that perform little I/O. The concept of multiprogramming was introduced to minimize the idle time of the CPU by organizing multiple jobs in the system so that the CPU always has something to execute. How is done is explained below.

Multiprogramming is the name given to the interleaved execution of two or more different and independent programs by the same computer. In Figure 3 we have been introduced to the notion of having two programs in the main memory at the same time – the operating system for overall system control and the user program for performing user's job. In multiprogramming, this concept is carried one step further by placing tow or more user programs in the main memory and executing them concurrently. With multiple user programs simultaneously residing in the main memory, when a user program that was executing (using the CPU) starts performing I/O operations, the CPU is allocated to another user program in the main memory that is ready to use the CPU instead of allowing the CPU to be idle. The CPU switches from one program to another almost instantaneously. Thus in multiprogramming several user program share the time of the CPU to kept it busy.

It is important to note here that multiprogramming is not defined to be the execution of instructions from several programs simultaneously. Rather, it does mean that there are a number of programs available to the CPU (stored in the main memory) and that is portion of one is executed, then a segment of another , and so on. Although tow or more user programs reside in the main memory simultaneously, the CPU is capable of executing only one instruction at a time. Hence at any given time, only one of the programs has control of the CPU is impossible. In some multiprogramming systems, only a fixed number of jobs can be processed concurrently (multiprogramming with fixed tasks) (MFT), while in others the number of jobs can vary (multiprogramming with variable tasks) (MVT).

At the particular time instance shown in the figure, job A is not utilizing the CPU since it is busy writing output data on to the disk (I/O operations). Hence the CPU is being utilized to execute job B, which is also present in the main memory. Job C, also residing in the main memory, is waiting for the CPU to become free. Actually, as shown in Figure 5, in case of multiprogramming all the jobs residing in the main memory will be on one of the following three states-running (it is using the CPU), blocked (it is performing I/O operations) and ready (it is waiting for CPU to be assigned to it). In our example, jobs A, B and C are in blocked, running and ready states respectively. Since job C is in the ready state, as soon s the execution of job B is completed or job B requires doing I/O operation, the CPU will start executing job C. In the meanwhile, if job A completes its output operation, it will be in the ready state waiting of the CPU. Thus in multiprogramming, the CPU will never be idle as long as there is always some job to execute. Note that although many jobs may be in ready and blocked states, only one job can be running at any instant.

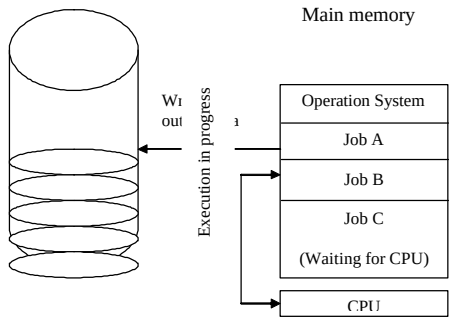


Figure 4

The area occupied by each job residing simultaneously in the main memory is known as a memory partition. The actual number of partitions, and hence jobs, allowed in the main memory at any given time varies depending upon the operating system in use. Moreover, those jobs awaiting entry into main memory are queued on a fast secondary storage device such as a magnetic disk. The first job from this queue will be loaded into the main memory as soon as nay one of the jobs already occupying the main memory is completed and the corresponding memory partition becomes free.

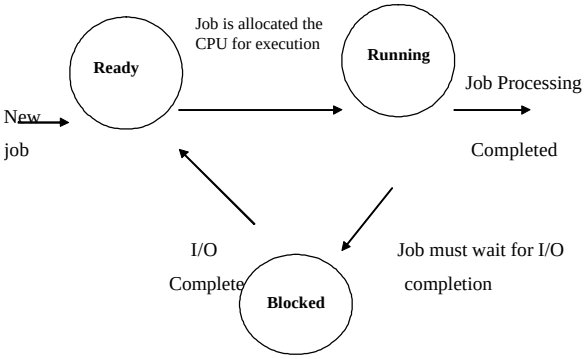


Figure-5 The three different states in which jobs may be after getting loaded in the main memory in a multiprogramming system.

Give some examples of blocked process

Explain how CPU is switched from one job to another in a multiprogramming system and how is it ensured that a temporarily suspended job will execute correctly when the CPU is given back to it after some time

Requirement of Multiprogramming Systems

Multiprogramming systems have better throughput than uniprogramming systems because CPU idle time is drastically reduced. However, multiprogramming systems are fairly sophisticated because they require the following additional hardware and software features :

1. Large Memory. For multiprogramming to work satisfactorily, large main memory is required to accommodate a good number of user programs along with the operating system.

2. Memory Protection. Computers designed for multiprogramming must provide some type of memory protection mechanism to prevent a job in one memory partition from changing information or instruction of a job in one memory partition. For example, in Figure 10.4 we would not want job A to inadvertently destroy something in the completely independent job B, or job C. In a multiprogramming system this is achieved by the memory protection feature, a combination of hardware and software, which prevents one job from addressing beyond the limits to its own allocated memory area.
3. Job Status Preservation. In multiprogramming, when a running job gets blocked for I/O processing, the CPU is taken away from this job and given to another job that is ready for execution. At a later time the former job will be allocated the CPU to continue its execution. Notice that this requires preserving of the job's complete status information when the CPU is taken away from it and restoring this information back before the CPU is given back to it again. To enable this, the operating system maintains a process control block (PCB) for each loaded process. A typical process control block is shown in Figure 10.6. With this arrangement, before taking away the CPU from a running process, its status is preserved in its PCB. Thus the process can continue execution without any problem.

Process identifier
Process state
Program counter

Values of various CPU
Register

Accounting and scheduling information

I/O status information

•
•

Figure 10.6. A typical process control block (PCB).

4. Proper Job Mix. A proper mix of I/O-bound jobs is required to effectively overlap the operations of the CPU and I/O devices. If all the loaded jobs need I/O at the same time, the CPU will again be idle. Hence the main memory should contain some CPU-bound and some I/O-bound jobs so that at least on job is always ready to utilize the CPU.
5. CPU Scheduling. In a multiprogramming system, often there will be situations in which two or more jobs will be in the ready state waiting for CPU becomes free, the operating system must decide which of the ready jobs should be allocated at the CPU for execution. The part of the operating system concerned with this decision is called the CPU scheduler, and the algorithm it uses is called the CPU scheduling algorithm.

Following questions should be answered :-

1. "The process management schedule is an important part of an OS". Discuss.

2. Distinguish between uniprogramming and multiprogramming.

3. What are the requirements of multiprogramming?

4. Distinguish between CPU bound and IO bound Jobs.

5. What are process states ?

LESSON 18
PROCESS MANAGEMENT-II

Outline

- Multitasking
- Multithreading
- Multiprocessing
- Time sharing

Multitasking

Technically speaking, multitasking is the same as multiprogramming. That is, multitasking is the system's capability to concurrently work on more than one task (job or process). This means that whenever a task (job or process) needs to perform I/O operations, the CPU can be used for executing some other task (job process) that is also residing in the system and is ready to use the CPU.

Many authors do not distinguish between multiprogramming and multitasking because both the terms refer to the same concept. However, some authors prefer to use the term multiprogramming for multi-user system (systems that are simultaneously use by many users such as mainframe and server class system), and multitasking for single-user systems(systems that are used by only one user at a time such as a personal computer or a notebook computer). Note that even in a single-user system, it is not necessary that the system work only one job at a time. In fact, a user of a single-user system often has multiple tasks concurrently processed by the system. For example, while editing a file in the foreground, a sorting job can be given in the background. Similarly, while compilation of a program is in progress in the background, the user may be reading his/her electronic mails in the foreground. In this manner, a user may concurrently work on many tasks. In such a situation, the status of each of the task is normally viewed on the computer's screen by partitioning the screen into a number of windows. The progress of different tasks can be viewed on different windows in a multitasking system.

Hence for those who like to differentiate between multiprogramming and multitasking, multiprogramming is the concurrent execution of multiple jobs (of same or different users) in a multi-user system, while multitasking is the concurrent execution of multiple jobs (often referred to as tasks of same user) in a single-user system.

What are the Differences Between Multitasking and Multiprogramming ?

Multithreading

Threads are a popular way to improve application performance. In traditional operating systems the basic unit of CPU utilization is a process. Each process has its won program counter, its own register states, its own stack, and its own address space (memory area allocated to it). On the other hand, in operating systems, with threads facility, the basic unit of CPU utilization is a thread. In these operating systems a process consists of an address space and one or more threads of control .Each thread of a process has its own program counter, its own register states, and its own stack. But all the threads of process share the same address space. Hence they also share the same global variables. In additions, all threads of a process also share the same set of operating system resources, such as open files, signals, accounting information, and so on, Due to the sharing of address space, there is no protection between the threads of a process. However, this is not a problem. Protection between processes is needed because different processes may belong to different user. But a process (and hence all its threads) is always owned by a single user. Therefore, protection between multiple threads of a process is not necessary. If protection is required between two threads of a process, it is preferable to put them in different processes, instead of putting them in a single process.

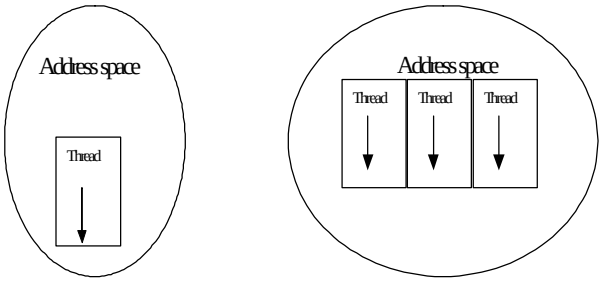


Figure-1(a) Single-threaded and (b) multithreaded processes. A single-threaded. Process corresponds to a process of a traditional operating system.

Threads share a CPU in the same way as processes do. At a particular instance of time, a thread can be in any one of several states – running, blocked, ready, or terminated. Due to these similarities, threads are often viewed as miniprocesses. In fact, in operating system with threads facility, a process having a single thread corresponds to a process of a traditional operating system [see Figure1 (a)]. Threads are often referred to as lightweight processes and traditional processes are preferred to as heavyweight processes.

Motivation for Using Threads

The main motivations for using a multithreaded process instead of multiple single-threaded processes for performing some computation activities are as follows :

1. The overheads involved in creating a new process are, in general, considerably greater than those of creating a new thread within a process. This is mainly because when a new process is created, its address space has to be created from scratch, although a part of it might be inherited from the process's parent process. However, when a new thread is created, it uses the address pace of its process that need not be created from scratch.
2. Due to the sharing of address space and other operating system resources among the threads of a process, the overhead involved in CPU switching among peer threads is very small as compared to CPU switching among processes having their own address spaces. This is the reason why threads are called lightweight process.
3. Resource sharing can be achieved more efficiently and naturally between threads of a process than between processes because all threads of a process share the same address space.
4. Finally, the use of threads is also motivated by the fact that a set of threads using a shared address space is the most natural way to program many applications. For example, in an application that uses the producer-consumer model, the producer an the consumer processes must share a common buffer. Therefore, programming the application in such a way that the producer and consumer are two threads of the same process makes the software design simpler.

Multiprocessing

Up to this point we have considered systems with a single CPU. However, we have already seen that the use of I/O processors improves the efficiency of a computer system by making possible concurrent input, processing, and output operations. The CPU can perform arithmetic and logical operations on parts of one or I/O processors on other parts of program concurrently carry out more programs while I/O operations. Figure 2 shows the architecture of a computer CPU, memory and I/O processors.

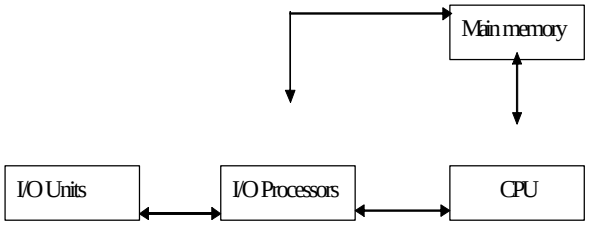


Figure 2Architecture of a computer system showing its CPU, memory and I/O processors.

The idea of use of I/O processors to improve the performance of a computer system was carried one step further by designing systems that make use of more than one CPU. Such systems are called multiprocessing system. The term multiprocessing is used to describe interconnected computer configurations or computers with two or more CPUs that have the ability to simultaneously execute several program. In such a system, instructions form different and independent programs can be processed simultaneously execute several programs. In such a system, instructions from different and independent programs can be processed simultaneously by different CPUs or the CPUs may simultaneously execute different instructions from the same program. The basic organization of a typical multiprocessing systems is shown in Figure.

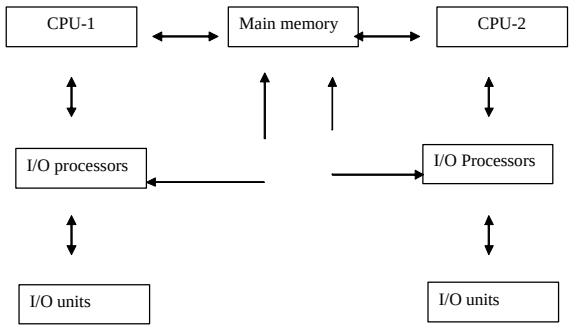


Figure 3 Basic organization of a typical multiprocessing system.

Tightly and Loosely Coupled Multiprocessing Systems

Multiprocessing systems are basically of two types – tightly-coupled systems and loosely-coupled systems. In tightly-coupled systems, there is a single system-wide primary memory that is shared by all the processors. On the other hand, in loosely-coupled systems, the processors do not share memory, and each processor has its own local memory. In contrast to the tightly-coupled systems, the processors of loosely-coupled systems can be located far from each other to cover a wider geographical area.

Multiprocessing

Multiprogramming is the interleaved execution of two or more processes by a single-CPU computer system. On the other hand, multiprocessing is the simultaneous execution of two or more processes by a computer system having more than one CPU. To be more specific, multiprogramming involves executing a portion of one program, then a segment of another, etc., in brief consecutive time periods. Multiprocessing, however, makes it possible for the system to simultaneously work on several program segments on one or more programs.

Advantage and Limitations of Multiprocessing

Multiprocessing systems typically have the following advantages:

1. Better Performance. Due to multiplicity of processors, multiprocessor systems have better performance than single-processor systems. That is, the multiple processors of such a system can be utilized properly for providing

shorter response times and higher throughput than a single-processor system. For example, if there are two different programs to be run, two processors are evidently more powerful than one because the programs can be simultaneously run on different processors. Furthermore, if a particular computation can be partitioned into a number of sub computations that can run concurrently, in a multiprocessor system all the sub computations can be simultaneously run with each one on a different processor. However, the speed-up ratio with n processors or subprograms, a certain amount of overhead is incurred in keeping everything working correctly. This overhead, plus contention of shard resources, lowers the expected gain from additional processors.

2. Better Reliability. Due to multiplicity of processors, multiprocessors system also have better reliability than single-processor systems. In a properly designed multiprocessors systems, if one of the processors breaks down, the other processor(s) automatically takes over the system workload until repairs are made. Thus a complete breakdown of such systems can be avoided. For example, if s system has 4 processors and one fails, then the remaining 3 processors can be utilized to process the jobs submitted to the system runs only 25% slower, rather than failing altogether. This ability of a system to continue providing service proportional to the level of non-failed hardware is called graceful degradation feature.

Multiprocessing system, however, require a very sophisticated operating system to schedule, balance, and coordinate he input, output, and processing activities of multiple processors. The design of such an operating system is a complex and time taking job. Moreover, multiprocessing systems are expensive to procure and maintain. In addition to the high charge paid initially, the regular operation and maintenance of these systems is also a costly affair.

Difference between Process and Program

Timesharing

Time-sharing is a mechanism to provide simultaneously interactive use of a computer system by many users in such a way that each user is given the impression that he/she has his/her own computer. It uses multiprogramming with a special CPU scheduling algorithm to achieve this.

A time-sharing system has many (even hundreds of) user terminals simultaneously connected to the same computer. Using these terminals, multiple users can simultaneously work on the system. The multiprogramming feature allows multiple user programs to simultaneously reside in the main memory. The special CPU scheduling algorithm used in a time-sharing system allocates a very short period of CPU time one-by-one to each user process, beginning from the first user process and proceeding though the last one, and then again beginning from the first one. This short period of time during which a user process get the attention of the CPU is known as a time slice, time slot, or quantum and is typically of the order of 10 to 100 milliseconds. That is, when the CPU is allocated to a user process, the user process will use the CPU until the allotted time slice expires 9the system's clock sends an interrupt signal to the CPU after every time slice), or until the process needs to perform some I/O operation, or if the execution of the process is over during this time period. Notice here that in a time-sharing system, the CPU is taken away from a running process when the allotted time slice expires, even though the process can continue to run. Hence the process state diagram of a time-sharing system is as shown in

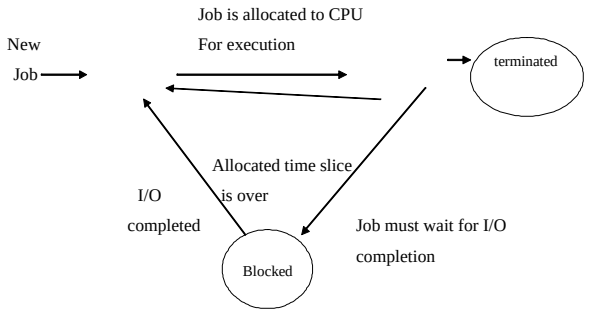


Figure 4. The process state diagram for a time-sharing system

Now let us see how the CPU scheduling algorithm mentioned above gives an impression to each user that he/she has his/her own computer. Let us assume that the time slice of a time-sharing system is 10 milliseconds. That is, the CPU scheduling algorithm of this system allocates 10 milliseconds to each user process one-by-one in a circular fashion (when the last process is over, it comes back to the first process). Suppose the processing speed of the system's CPU is of the order of 500 million instructions per second. That is, it can execute 500 x 106 x 10-3 x 10 = 5 x 106 = 5 million instructions in 10 milliseconds. This is large enough for substantial progress of a single user process. Now suppose there are 100 milliseconds is allocated to each user process one-by-one, a particular user will get the CPU's attention once in every 10 x 100 milliseconds = 1 second. As human reaction time is normally of the order of a few seconds, a particular user will not notice any delay in executing his/her commands and will normally feel that he/she is the sole user of the system. In this manner, each user is given the impression that he/she has his/her own computer, whereas actually a single computer is shared among many users.

To Achieve Timesharing - What Are The Requirements?

Time-sharing systems typically require the following additional hardware and software features:

1. A number of terminal simultaneously connected to the system so tht multiple users can simultaneously use the system in interactive mode.

2. A relatively large memory to support multiprogramming.
3. Memory protection mechanism to prevent one job's instruction and data from other jobs in a multiprogramming environment.
4. Job status preservation mechanism to prevent one job's instructions and data from other jobs in a multiprogramming environment.
5. A special CPU scheduling algorithm that allocates a very short period of CPU time one-by-one to each user process in a circular fashion.
6. An alarm clock mechanism to send an interrupt signal to the CPU after every time slice.

Advantage of Time-Sharing Systems

Although time-sharing systems are complex to design, they provide several advantage to their user. The main advantages of time-sharing systems are as follows:

- Reduces CPU Idle Time. It will be waste of precious CPU time to allow a single user to use a large computer interactivity because his/her speed of thinking and typing is much slower than the input, a time-sharing systems can service many other user. In this manner, time-sharing systems help in reducing the CPU idle time to a great extent, increasing the system throughput.
- Provides Advantages Of Quick Response Time. The special CPU scheduling algorithm used in time-sharing systems ensures quick response time to all users. This feature allows users to interact with the system more rapidly while working on their problem. For example, a time-sharing system can be effectively used for interactive programming and debugging to improve programmers efficiency. Multiple programmers can simultaneously proceed step-by-step, writing , testing and debugging portions of their programs or tying out various approaches to a problem solution. The greatest benefit of such a system is that errors can be encountered, corrected, and work can continue immediately for all the simultaneous users of the systems. This is in contrast to a batch system in which errors are corrected offline and the job is resubmitted for another run. The time delay between job submission and return of the output in a batch system is often measured in hours.
- Offers Good Computing Facility To Small Users. Small users can gain direct accesses to much more sophisticated hardware and software than they could otherwise justify or afford. In time-sharing systems, they merely pay a fee for resources used and are relieved of the hardware, software, and personnel problems associated with acquiring and maintaining their own installation.

Following questions should be Answered :

1. Define multiprogramming . Explain how multiprogramming ensures effective utilization of main memory and CPU

2. What are threads? Differentiate between a single threaded process and a multithreaded process.

LESSON 19
TYPES OF OPERATING SYSTEM-I

Outline

- MS-DOS
- Windows
- Windows 3.1
- Windows ME

MS-DOS
MS-DOS stands for Microsoft Disk Operating System. MS-DOS displays lines of text on the screen. You perform tasks by typing text commands.

Windows
Windows is a Graphical User Interface (GUI, pronounced "gooey"). A GUI allows you to use pictures instead of text commands to perform tasks. This make Windows easier to use than MS-DOS.

UNIX
UNIX is a powerful operating system used by many computers on the Internet. There are many different versions of the UNIX operating system available.

Mac OS
Mac OS is a Graphical User Interface (GUI) for Macintosh computers.

Platform
A platform refers to the type of operating system used by a computer, such as Windows or UNIX. Programs used on one platform will not usually work

on another platform. For example, you cannot use Word for Windows on a computer running UNIX.

MS-DOS
MS-DOS is an operating system that performs tasks using text commands you enter. And MS-DOS stands for Microsoft Disk Operating System.

Enter a Command

Command Prompt
The command prompt (C:\>) tells you that MS-DOS is ready to accept a command.

Command
You enter a command to perform a task or start a program. A single computer what you want to accomplish.

Cursor
The cursor is the flashing line on the screen. The cursor indicates where the text you type will appear.

IS MS-DOS Multi-user OS ?

File Organization
Like folders in a filing cabinet, MS-DOS uses directories to organize the data stored on a computer. The root directory (C:\) is the main directory. All other directories are located within this directory. A path describes the location of a file.

File Name
When you store a file on a computer, you must give the file a name. An MS-DOS file name cannot contain any spaces. A file name consists of a name and an extension, separated by a period. The name describes the contents of a file and can have up to eight characters. The extension identifies the type of file and consists of three characters.

Utilities
MS-DOS 6.0 and later versions include special programs, called utilities, to protect files and optimize a computer.

Can you have two operating systems on your PC?

Windows 3.1

Windows 3.1 works with MS-DOS to control the overall activity of a computer.

Is Windows 3.1 A True Operating System ?

Program Manager

The Program Manager is the control center where you start programs.

Program Icon

A program icon lets you start a program, such as a word processor. An icon is a small picture that represents an item, such as a program.

Group Icon

A group icon contains program icons. For example the Games group icon contains several games.

Window

A window is a rectangle that display information on the screen. Each window has title bar that display the name of the window.

File Manager

The File Manager lets you view and organize all the files stored on your computer. Windows 3.1 uses directories to organize information, just as you would use folders to organize papers in a filing cabinet.

Desktop

The desktop is the background area of the screen.

Windows 95

Windows 95 is the successor to Windows 3.1. This operating system is more graphical and easier to use than Windows 3.1.

My Computer

My Computer lets you browse through all the folders and documents stored on your computer.

Recycle Bin

Recycle Bin stores documents you delete and allows you to recover then later.

Network Neighborhood

Network Neighborhood lets you view the folders and files available on your network.

Start Button

The Start button lets you quickly access programs and documents.

Taskbar

The taskbar contains the Start button and displays the name of each open window on the screen.

Windows Explorer

Windows Explorer shows you the location of each folder and document on your computer. You can use Windows Explorer to move, open, print or delete documents.

Shortcut

A shortcut provides a quick way to open a document or program you use regularly.

Windows 98

Windows 98 is the successor to Windows 95. This operating system is very similar to Windows 95 but includes many new and improved features. There are currently two version of Windows 98. Windows 98 Second Edition includes many enhancements and updates to the original version of Windows 98.

My Documents

My Documents provides a convenient place to store your documents.

Internet Explorer

Internet Explorer lets you browse through information on the World Wide Web.

Quick Launch Toolbar
The Quick Launch Toolbar lets you quickly access commonly used features, including Internet Explorer and Outlook Express.

Windows 98 Features

Computer Maintenance
Windows 98 is more reliable than Windows 95 and includes many tools you can use to find and fix problems with your computer. For example, you can check your hard drive for errors, remove unneccessary files and defragment your hard drive to improve its performance.

FAT32
FAT32 is a file system that better manages data large hard drives to reduce wasted space. Windows 98 can convert your hard drive to FAT32 without disrupting your current programs and documents.

Internet
Windows 98 includes several programs that allow you to view and exchange information on the Internet Explorer, Outlook Express and FrontPage Express. Outlook Express lets you exchange electronic mail with people around the world. FrontPage Express allows you to create your own Web pages.

Multiple Monitor Capability
Windows 98 has the ability to display the Windows desktop on multiple monitors. This makes working with several open documents or programs easier.

Windows ME
Windows ME is the successor to Windows 98. Windows ME stands for Windows Millennium Edition. And if you have used Windows 98, you will already be familiar with the way Windows ME looks and works.

Windows ME Features

Record Videos
Windows ME allows you to record, edit and save videos on your computer using Windows Movie Maker. Once you have finished working with a video, you can e-mail the video to your friends and colleagues or place the video on a Web page.

Manage Multimedia Files
Windows ME contains an enhanced version of Windows Media Player that helps you manage your multimedia files. You can use Windows Media Player to play a variety of multimedia files and listen to radio stations over the Internet.

Restore Your Computer
If you are experiencing problems with your computer, you can use the System Restore feature to return your occurred. For example, if you have accidentally deleted program files, you can restore your computer to a time before you deleted the files.

Set Up a Home Network
If you have more than one computer at home, you can use the Home Networking Wizard to help you set up a network so you can exchange information between the computers. A home network is also useful or sharing equipment and playing multiplayer games.

Update Windows
The Windows Update feature allows you to automatically update your computer with the latest Windows features available on the Internet. When you are connected to the Internet, Windows will automatically check for updates, determine which updates apply to your computer and notify you when the updates are available.

Windows NT is a powerful version of the Windows operating system that provides advanced networking and security features.

LESSON 20
TYPES OF OPERATING SYSTEM-II

Outline

- Windows NT
- Windows 2000
- Unix
- Linux
- Mac OS 9

Windows NT

Windows NT is available in two main versions.

Windows NT Workstation

Windows NT Workstation is a version of the Windows NT operating system that is used on client/server and some peer-to-peer networks. Many powerful applications are designed specifically to run on Windows NT Workstation. Many programs designed for the Windows 95 and Windows 98 operating systems will also perform better on Windows NT Workstation.

Windows NT Server

Windows NT Server is a version of the Windows NT operating system that is used on client/server networks. Windows NT Server is designed to support the heavy processing demands of a network server. The client computers on a network running Windows NT Server can use a variety of operating systems, such as Linux, Windows 98 and Mac OS 9.

Windows 2000

Windows 2000 is the successor to Windows NT. This operating system offers improved networking and security features.

There are several versions of Windows 2000 available.

Windows 2000 Server and Windows 2000 Advanced Server

Windows 2000 Server and Windows 2000 Advanced Server are found on large client/server networks. These operating systems are both designed to support heavy network processing demands. Windows 2000 Advanced Server can support a larger network than Windows 2000 Server, which makes the Advanced Server version suitable for large businesses amd Internet service providers.

Windows 2000 Professional

Windows 2000 Professional is commonly used on client/server networks but can also be used on peer-to-peer networks. This operating system offers increased stability and provides tools that can help you maintain your computer. Windows 2000 Professional is intended for business use.

Write the features of windows XP

UNIX

UNIX is an older powerful operating system that can be used to run a single computer or an entire network. It is the oldest operating system still in widespread use today.

Versions

Many companies have owned UNIX since its development in the late 1960s. Today, there are several versions of the UNIX operating system available. Popular UNIX operating systems for personal computers include UnixWare by SCO Solaris by Sun Microsystems and Linux, which is available for free on the World Wide Web. Other versions of UNIX that are more popular as network operating systems include HP-UX by Hewlett-Packard and AIX by IBM.

Internet

Many of the first computers used to establish the Internet ran the UNIX operating system. Even today, UNIX is the most widely used operating system for servers on the Internet.

UNIX Features

Power

The UNIX operating system is very powerful. UNIX is harder to install and set up than most other operating systems, but provides greater control over a computer's resource and power. A computer's performance may be significantly improved when running UNIX.

Security

UNIX has many built-in security features to protect information from being accidentally deleted or accessed by unauthorized users. UNIX's strong security features are one of the reasons UNIX is such a popular operating system on the Internet.

Multitasking

UNIX was originally developed as the operating system for a single large computer, called a mainframe computer. Since multiple users can access a mainframe computer at the same time, UNIX was developed to run many programs and perform numerous tasks at once, called multitasking. UNIX's multitasking capabilities make it an efficient operating system for networks.

Linux

Linux is a UNIX-based operating system that is available for free on the World Wide Web. Many companies, such as Red Hat, Corel and Mandrake, create easy-to-use versions of Linux that you can purchase. Red Hat Linux is a popular version of Linux that comes with the GNOME desktop environment. GNOME displays pictures on the screen to help you perform tasks.

Home Directory

The Home directory allows you to browse through the folder that stores your personal files.

Help Web Pages

The Help Web pages let you quickly display Web pages that offer help on using Linux.

GNOME Panel

The GNOME Panel contains the Main Menu button and Application Launchers and displays the name of each open window on the screen.

Main Menu Button

The Main Menu button gives you quick access to programs.

Application Launchers

Application Launchers allow you to quickly access commonly used programs, such as the help system and Netscape Communicator.

Linux Features

Software

Red Hat Linux includes a wide variety of software, such as a drawing program, spreadsheet program, calendar program and a simple text editor. You will also find an on-screen calculator, an address book and many games that you can buy.

Command Line

You can work in Linux using a Graphical User Interface (GUI) such as GNOME or using the command line. The command line displays text without any graphics. Working in the command line allows you to perform many tasks more quickly than in a GUI, though you must know the proper commands.

Accounts

When you install Linux, the root account and one or more user accounts are created. You can work in Linux using either the root account or a personal account. The root account is useful for performing administrative and maintenance tasks. When

performing daily tasks, you should work in a user account. You can set up a different user account for each person who uses your computer.

MAC OS 9

Mac OS 9 is a popular version of the Macintosh operating system. This operating system offers many advanced multimedia and Internet features.

MAC OS 9 Features

Multimedia

Mac OS 9 offers advanced sound, video and graphics capabilities. For example, the QuickTime player allows you to manage and play multimedia files, such as songs and movie clips. Mac OS 9 also offers a speech recognition program that allows your computer to recognize voice commands.

Internet

Mac OS 9 offers several features that helps you find and exchange information on the Internet. For example, the Sherlock 2 feature allows you to efficiently search for information in categories such as News, Reference, People and Shopping.

Mac OS X (ten) is the latest version of the Macintosh operating system.

MAC OS X Features

Graphical User Interface

Mac OS X offers a new Graphical User Interface (GUI) that is designed to make sure operating system easier to use than previous versions. The new (GUI) features photo-quality icons and an area at the bottom of the screen, called the Dock, where you can store items you frequently access.

Graphics

Mac OS X combines several advanced graphics technology, including Portable Document Format (PDF), QuickTime and OpenGL, to provide enhanced graphics capabilities. This improves the display of graphics in programs such as desktop publishing programs and games.

Operating System Core

Mac OS X offers an improved operating system core, or kernel. This improved kernel helps ensure that your computer hardware and software work together efficiently and provides more stability than previous versions of the Macintosh operating system.

Note:

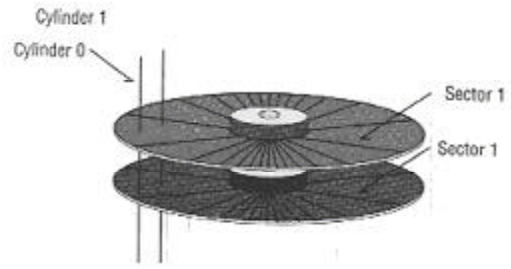
LESSON 21
COMPUTER OPERATIONS- FILE/DIRECTORY STRUCTURE

Outline

- How the information is stored on the hard disc?
- Types of file system
  - FAT
    - FAT 16
    - FAT 32
  - NTFS
  - HPFS
- Unix – file structure

How The Information is Stored

Basically, tracks, sectors, and cylinders are the divisions of the hard drive platters where information is stored. A track is a concentric ring around the platter containing information. Since a hard drive typically has two or more platters, each storing data on both sides, these tracks line up on each platter. The identically positioned tracks on each platter are called cylinders. To better help you understand a track and cylinder, let's take a target used for target practice. You have a bunch of concentric circles, each bigger than the other, all sharing the same center, which is the bulleye. Now, each of the spaces between circles is similar to a track on a hard disk platter. Now, if you stack several of these targets on top of each other, each exactly the same, you can form a cylinder by simply taking a track and moving it down through all of the same tracks on the targets below.



Since typical hard drives are too large to deal with by the track, each track is divided into sectors. Its not that a track could not be dealt with, but since a track can hold as much as 50K sometimes, this would not be practical for storing large files. So, sectors are basically slices of the track. Different drives have different numbers of sectors per track.

Each sector is given an identity during low-level formatting to aid the controller in finding what it needs in the appropriate sector. These sector numbers are written to the beginning and the end of each sector, called the prefix portion and the suffix portion respectively. These identities take actual space on the hard drive. This explains why there is a difference between the capacity of an unformatted disk and a formatted one. On a floppy, the disk itself can hold 2M or so of data. When formatted and the identities placed, the capacity reduces to 1.44M. The same holds true for a hard drive. Drive manufacturers know this and publish formatted capacities to indicate drive size.

Types Of File System

There are two types of disk formatting: low-level and high-level. These both are done in the preparation of a hard drive for use. First, one low-level formats, and then partitions, then high-level formats. A low-level format turns the platter from a blank slate to a divided slate. It defines the data areas: creates tracks, separates into sectors, and writes the ID numbers to each sector. Low-level formatting typically has to be done using a specific utility released by the manufacturer of the drive. Almost all drives sold are already low-level formatted at the time of sale. In most cases, you do not need to ever do it again.

Partitioning segments the drive into separate areas, each capable of running its own operating system. At this point, the hard drive needs to be high-level formatted. High level formatting is the type of formatting most people think of when thinking of formatting. During high-level formatting, the file allocation tables (FATs) are dropped in. This is a "table of contents" for the drive, allowing the drive to later find files on the drive wherever they sit on the platter. The drive cannot be used until it is high-level formatted. This can be performed with the DOS FORMAT command, through Windows (if the drive is not master), or through a third party utility.

There are four types of file systems.

- FAT16. The is the file system used by DOS, Win95 and Win98, unless you convert. It supports an eight letter file name max, with a three letter extension under DOS. With Windows it supports up to 255 characters. With this system, a partition can be no larger than 2G.
- FAT32. The 32 means 32-bit. This is an optional file system introduced with Windows 95 OSR-2. The file allocation units are stored as 32-bit numbers. The main advantage is that it allows for partitions of up to 2048G with smaller clusters. There is an option to convert FAT16 to FAT32 under Windows 98, and Windows NT5 will support it as an option.
- High Performance File System (HPFS). An unpopular file system only used with OS/2 or early versions of NT. File names can be 256 characters and partitions can be 8G.
- Windows NT File System (NTFS). Just like the HPFS, but only for NT.

Which File System You Have In Your Computer System?

The FAT file system (specifically FAT32) is most used in PC's today. The main problem with the original FAT16 was the

inefficient use of disk space when defining clusters, or groups of sectors. The clusters were rather large, causing wasted space because a small file would still take up the entire cluster even though that cluster could hold more. With FAT32, 4 billion clusters are allowed, ultimately allowing 4K clusters. This significantly reduces disk waste. The concept of FAT explains why one can run out of disk space even when you are not storing the disk's full capacity in files. For example, a 1G hard drive can run out of space with 160MB to spare. This is due simply to the FAT structure. With the original FAT, each cluster could hold 32K. But, if you are storing an 8K file, it still takes up a complete cluster, leaving the other 24K to waste. This wasted space is called slack. The only way around this is to re-partition the hard drive to two or more partitions. As the partition gets smaller the wasted space gets less. This, then is a tradeoff. The convenience of one partition, or the wasted disk space. With FAT32, the wasted space is much less because of the smaller clusters.

No matter what file system is used, a boot sector is written to the beginning of each disk in the first sector. This sector contains the boot program which tells the system what to do when you first turn it on. This is called the boot sector.

That completes the basic view of how files are stored on a hard drive.

Every computer we use today almost always has a hard drive. I'm not sure how else you could run one, quite frankly. Hard Drives store information on little platters, which are shaped like Frisbees, but made of metal and coated with a magnetic substance. But how exactly is the data organized on those little platters? They do it with a thing called a file system. File systems are mostly operating system dependent. This means that almost every operating system has a different type of file system. Windows is the most widely used operating system in the United States and Europe; so Windows-compatible file systems are pretty much the standard that most other operating systems can at least read from.

FAT

The Dos/Windows file system is called File Allocation Table, or just FAT for short. There are 3 types of this FAT file system:

- FAT12. An old 12-bit file system which is mainly used on floppy disks and REALLY small hard drives.
- FAT16. A 16-bit file system used by DOS-based machines as well as PCs using older versions of Windows such as 3.x. Windows 9x is FAT16 compatible, but by this time FAT32 was becoming the new norm.
- FAT32. A 32-bit file system most common today.

Whether a partition is going to use FAT12 or FAT16 is based mainly on it's size and version of DOS you are using. If the partition or disk is 16MB or less in size, it is going to use FAT12. If the partition is between the size of 17MB and 2048MB, it will use the FAT16 file system.

FAT uses clusters to store files in. Each cluster is a group of sectors. The computer gives each cluster it's own address, just as each house in a neighborhood has it's own address. The operating system then keeps track of which files are stored in which clusters. The partition size as well as the type of FAT system being used determines the cluster size. The cluster size is important because only one file can be stored in a cluster at a time. If you have a cluster that is 32KB, and you are storing a file in it that is only 1KB, you are wasting 31KB of space on your hard drive. That's not very big when you look at it in terms of a 2GB hard drive, which has 2,097,152KB on the drive, but when you take a look and realize that there are no files that are perfectly going to fill up 32KB, you are wasting space with every file you have on your hard drive. How can this be avoided? The smaller the partition you have, the smaller the clusters you will have, and the less wasted space you will have. Below is a table showing the exact cluster sizes you will get with partition sizes.

Partition size	FAT Type	Cluster size
< 16MB	FAT12	4KB
17MB-32MB	FAT16	2KB
33MB-256MB	FAT16	4KB
257MB-512MB	FAT16	8KB
512MB-1GB (1024MB)	FAT16	16KB
1GB-2GB (2048MB)	FAT16	32KB

There is a slight problem with the FAT16 file system. What would you do if you had a hard drive larger than 2GB and wanted only 1 partition? FAT32 solved that. It now supports drives up to 2048GB. 2048GB = 2 terabytes. FAT32 also solved the problem of large cluster sizes. Below is another table, but this compares the size of a FAT32 with its cluster sizes.

FAT32 Partition Size	Cluster Size
< 260MB	512bytes (½KB)
260MB-8GB	4KB
8GB-16GB	8KB
16GB-32GB	16KB
32GB-2048GB (2TB)	32KB

Unfortunately, the ability to make a 260MB FAT32 partition is limited due to the fact that the major program that makes partitions limits FAT32 drives to 512MB. The improved nature of FAT32 has caused this file system to be the most common type used. All later versions of Windows either work primarily with this file system or at least support it (like Windows 2000

and XP, which both allow FAT32 as an option over NTFS). On version of Windows supporting both FAT16 and FAT32, Microsoft usually bundles a conversion utility with the operating system to allow you to convert a FAT16 partition to a FAT32 partition without data loss.

NTFS

NTFS (NT File system) is the major file system used by Microsoft's Windows NT, 2000 and XP. (Each of these also support the FAT file system). NTFS has features to improve reliability, such as transaction logs to help recover from disk failures. To control access to files, you can set permissions for directories and/or individual files. NTFS files are not accessible from other OSs such as DOS

For large applications, NTFS supports spanning partitions or volumes, which mean files and directories, can be spread out across several physical disks. The corporate and power users mainly use this file system.

HPFS

High Performance File system is basically a mix between NTFS and FAT. While FAT offers the 8.3 file naming system (8 characters, then a dot, then 3 more characters) HPFS will support up to 256 characters in a file name. This file system is mainly used by OS/2, which was IBM's answer to Microsoft Windows.

Ext2

Ext2 is a file system used by Linux. The main people that use it are those that run the many versions of the Linux Operating system.

Unix - File Structure And Management

Overview

Directories and their structure

Unix is made up of a hierarchical file structure. The fundament of this structure is the directory. It contains the records and the locations of all existing files. Unix has many such directories. For example, as you enter your username and password, you are automatically located in a current directory, which is your own home directory. Within your home directory, you can create sub-directories for various purposes. There are also such distinctions as parent and child directories. For example, you've created a new directory and called it mhcnews. MHCNEWS is a child directory of the home directory, and your home directory is a parent of mhcnews.

You can also visualize the Unix structure as a "tree with branches." The root of this tree is a directory called root. It has many branches coming out of it (sub-directories), and those have others as well (files, etc). Here's a picture.

A Word on File Management

The good thing to remember about directories is to keep them well organized. Just as you save lots of information on disk, if there is too much, it is hard to keep track of it. Same goes for Unix directories. In naming your files and directories, make sure that when you come back half a year later, you can still remember the contents of the file by the name you gave it. Also, the good part about keeping your directories small is that if you accidentally erase a directory, it doesn't change the structure of the tree as a whole. Those removed files are only one branch of the tree, and the tree itself remains undamaged.

The Unix Language of Pathnames

Let's take Jane's directory and put it into Unix language. It would be /user/jane/collegepapers. The "/" at the beginning is an abbreviation for root. If you wanted to see John's mhcnews (which is at /user/john/mhcnews), you would have to go back to user directory (/user), and from there go to John's dir, and see what he has to say about MHCnews. This is the most misunderstood technique by many new users. In order to go back a directory, Unix uses this command../ . In our example this would mean that from "college papers" we would back up to "Jane" directory and in order to get to "user", we have to back up yet another directory. Here's the Unix for it:

Our current directory /user/jane/collegepapers

Backing up to "Jane" ../

New current directory /user/jane

Backing up to "user" ../

We're now in "user" /user

Moving to John's mhcnews /user/john/mhcnews

Just like in the last example, there are two slashes in one line, to get straight into mhcnews; we could do the same while backing up directories. Instead of going../ For each backup, we could use../../ to go directly back to "user."

Another detail worth mentioning is the slash itself. Unix has several similarities with DOS. It is easy to confuse the DOS' \ with Unix's / . Remember to use backslash with Unix - / !

The Unix File System

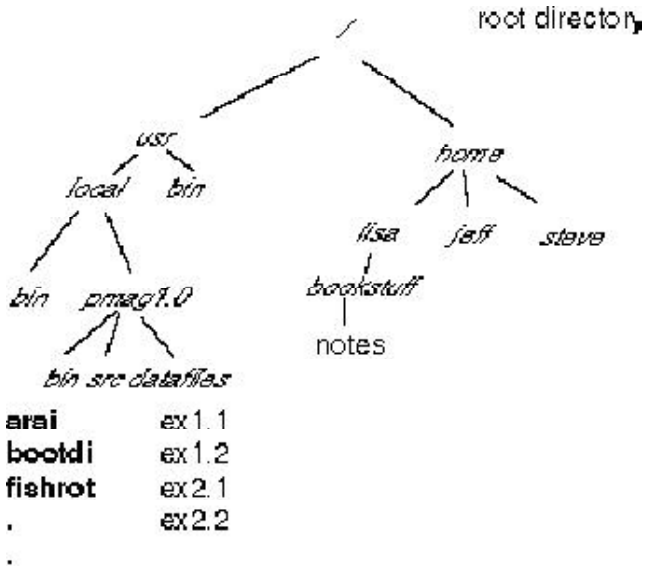
Fundamental to the UNIX operating system is the concept of directories and files. On windows based operating systems, directories are depicted as ``folders" and moving about is accomplished by clicking on the different icons. In UNIX, the directories have names and are arranged in a hierarchical sequence with the top directory being the ``root" directory, known as ``/" (see Figure A.1. Within the ``/" directory, there are subdirectories (e.g. usr and home). In any directory, there can also be ``files" (e.g. ex1.1, ex1.2) in the figure. Files can be ``readable", ``writable" and/or ``executable".

2#2

When you log in, you enter the computer in your ``home" directory. To refer to directories, UNIX relies on what is called a ``pathname". Every object has an ``absolute" pathname, which is valid from anywhere on the computer. The absolute pathname always begins from the root directory /. So the absolute pathname to the home directory Lisa in Figure A.1 is /home/Lisa. Similarly, the absolute pathname to the directory containing pmag1.0 executables is /usr/local/pmag1.0/bin. There is also a ``relative" pathname, which is in reference to the current directory. If user ``Lisa" is sitting in her home directory, the relative pathname for the file notes in the directory book stuff would be book stuff/notes. When using relative pathnames, it is useful to remember that. / Refers to the current directory and../. Refers to the one above.

A program called the shell handles commands typed at the command line prompt. There are many different sorts of shells (e.g. sh, csh, jsh, ksh, bash, tsh) that have a different look and feel but they all perform the task of interpreting between the user and the ``kernel'' which is the actual UNIX operating system. In order to execute a command, the shell needs to know where the command is. There are several ``built-in'' commands, but most are programs that are either part of the operating system, or something someone wrote (like the ones referred to here). There are any numbers of places where programs are kept, so the shell looks in particular places determined by your ``path'' environment variable. To instruct the shell to look in directories other than the default directories (for example in /usr/local/pmag1.0/bin), ask your system administrator to add this directory to your ``path''. Otherwise, you can always type the absolute pathname for any program (e.g. /usr/local/pmag1.0/bin/fishrot to execute fish rot.

Figure A.1: UNIX file system



Note:

LESSON 22
SECURITY & DEADLOCK

Outline

- Security
  - User Authentication
  - Access Control
  - Cryptography
    - Symmetric
    - Asymmetric cryptosystems
- Deadlock
- Spooling

Security

Security in computer systems deals with the protecting the various resources and information of a computer system against destruction and unauthorized access. A total approaches to computer security involves both external and internal security. External security deals with securing the computer system against external factors such as fires, floods, earthquakes, stolen disks/tapes, leading out of stored information by a person who has access to the information, and so on. For external security, the commonly used methods include maintaining adequate backup copies of stored information at place far away from the original information, using security guards to allow the entry of only authorized persons into the computer centre, allowing the access to sensitive information to only trusted employee/users, and so on. Internal security, on the other hand, mainly deals with the following aspects.

1. User authentication.

Once a user is allowed physical access to the computer facility, the user's identification must be checked by the system before the user can actually use the facility. This requirement is taken care of by user authentication mechanisms.

2. Access control.

A computer system contains many resources and several types of information. Obviously, not all resources and information are meant for all user. Therefore, even when a user passes the authentication phase and is allowed to use the computer facility, a way is needed to prohibit the user from accessing those resources/information that he/she is not authorized to access. This requirement is taken care of by access control mechanisms.

3. Cryptography.

A third type of internal security is often used to protect unauthorized access of information that is highly critical in nature. That is, even if a user somehow manages to gain access to some information that he/she is not authorized to access, a way is needed to ensure that the user cannot make use of that information. This requirement is taken care of by cryptography mechanisms.

Below we will discuss about the commonly used mechanisms for providing these three different types of internal security in computer systems.

User Authentication

User authentication deals with the problem of verifying the identify of a user (person or program) before permitting access to the requested resource. That is, an authentication mechanism prohibits the use of the system (or some resource of the system) by unauthorized users by verifying the identity of a user making a request.

Authentication basically involves identification and verification. Identification is the process of claiming a certain identify by a user, while verification is the process of verifying the user's claimed identity. Thus, the correctness of an authentication process relies heavily on the verification procedure employed.

The three basic approaches to user authentication are as follows:

1. Proof by Knowledge. In this approach, authentication involves verifying something that can only be known by an authorized user. Authentication of a user based on the password supplied by him/her is an example of proof by knowledge. Authentication methods based on the concept of proof by knowledge are again of two types – direct demonstration method and challenge – response method. In he direct and challenge – response methods. In the direct demonstration method, a user claims his/her identity by supplying information (like typing in a password) that the verifier checks against pre-stored information. On the other hand in the challenge-response method, a user proves his or her identity by responding correctly to the challenge questions asked by the verifier. For instance, at the time of initially registering in a system as a user, the user picks a function, for example, x + 18. When the user logs in, the system randomly selects and displays a number, say 105, in which case the user must type 123 for authentication to be successful.
2. Proof by Possession. In this approach, a user proves his/her identity by producing some item that can only be possessed by an authorized user. The system is designed to verify the produced item to confirm the claimed identity. For example, a plastic card with a magnetic strip on it that has a user identifier number written on it in invisible, electronic form may be used as the item to be produced by the user. The user inserts the card in a slot meant for this purpose in the system's terminal, which then extract the user identifier number from the card and checks to see if the card produced belong to an authorized user. Obviously can be ensured only if the item to be produced is unforgeable a safely guarded.

3. Proof by Property. In this approach, the system is designed to verify the identity of user by measuring some physical characteristics of the use that are hard to forge. The measured property must be distinguishing, that is, unique among all possible users. For example, a special device (known as a biometric device) may be attached to each terminal of the system that verifies some physical characteristic of the user, such as the person's appearance, fingerprints, hand geometry, voice, and signature. In deciding the physical characteristic to be measured, an important factor to be considered is that the scheme must be phycologically acceptable to the user community. Biometric systems offer the greatest degree of confidence that a user actually is who he/she claims to be, but they are also generally the most expensive to implement. Moreover, they often have user acceptance problems because user see biometric devices a unduly intrusive.

In practice, a system may use a combination of two or more of these authentication methods. For example, the authentication mechanism used by automated cash-dispensing machines in banks usually employs a combination of the first two approaches. That is, a user is allowed to withdraw money only if he or she produces a valid identification card and specifies the correct password corresponding to the identification number on the card.

To provide good security with password-based authentication, it is important that passwords are kept secret and passwords are chosen in such a manner that they are difficult to guess.

Access Control

Once a user or a process has been authenticated, the next step in security is to devise ways to prohibit the use or the process from accessing those resources/information that he/she/it is not authorized to access. This issue is called authorization and is dealt with by using access control mechanisms.

When talking about access control in computer system, it is customary to use the following terms:

1. Objects. An object is an entity to which access must be controlled. An object may be an abstract entity, such as a process, a file, a database, a tree data structure, or a physical entity, such as a CPU, a memory segment, a printer, and a tape drive.

   Each object has a unique name that differentiates it from all other objects in the system. An object is referenced by its unique name. In addition, associated with each object is a "type" that determines the set of operations that may be performed on it. For example, the set of operations possible on objects belonging to the type "data file" may be Open, Close, Create, Delete, Read and Write, whereas for objects belonging to the type "program file," the set of possible operations may be Read, Write, and Execute.

2. Subjects. A subject is an active entity whose access to objects must be controlled.

   That is, entities wishing to access and perform operations on objects and to which access authorizations are granted are called subjects. Example of subjects is processes and users.

3. Protection Rules. Protection rules define the possible ways in which subjects and objects are allowed to interact. That is, protection rules govern the subjects' access to objects. Therefore, associated with each (subject, object) pair is an access right that defines the subset of the set of possible operations for the object type that he subject may perform on the object. The complete set of access rights of a system defines which subjects can perform what operations on which objects. At any particular instance of time, this set defines the protection state of the system at that time.

   With the above-mentioned concepts, access to objects by subjects by subjects is controlled in the following manner. Suppose subject S requests for an access on object O, where r belongs to the set of operations that may be performed on O. To validate this access request, the access controls modules for the security system check whether access r is permitted for subject S on object O in the current protection state of the system. If yes, the access is permitted; otherwise the access is denied.

Cryptography

Cryptography is a means of protecting private information against unauthorized access in that situation where it is difficult to ensure access control. The basic idea behind this security technique is that if it is not possible to ensure access control, it is better to prevent comprehension of information.

Basic Concepts and Terminologies

Two primitive operations employed by cryptography are encryption and decryption. Encryption (also called enciphering) is the process of transforming an intelligible information (called plaintext or cleartext) into an unintelligible form (called ciphertext). Decryption (also called deciphering) is the process of transforming the information back from ciphertext to plaintext.

Encryption is basically a mathematical function (encryption logarithm) having the following form:

C = E (P, Ke)

Where P is the plaintext to be encrypted, Ke is an encryption key, and C is the resulting ciphertext. Decryption of C is performed by a matching function (decryption algorithm) that has the following form.

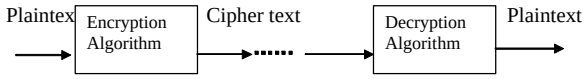
P = D (C, Kd)

Where Kd is the decryption key. Note that the decryption function D is the inverse of the encryption function E. Therefore we have

D (E, (P, Ke), Kd) = P

To prevent the plaintext from being easily revealed, it must be possible to transform a given plaintext into a large variety of possible ciphertext selected by a specific parameter. The keys Ke and Kd serve as this parameter. That is, the function parts remain the same but he keys are changed as often as necessary.

The above-described general structure of a cryptosystem is illustrated in figure 1



Symmetric and Asymmetric Cryptosystems

There are two broad classes of cryptosystem –

- Symmetric
- Asymmetric.

Symmetric

In a symmetric cryptosystem, either both the encryption key (Ke) or decryption key (Kd) are the same or one is easily derivable form the other. Usually, a common key (K) is used for both enciphering and deciphering. For security, it is important that the key of a symmetric cryptosystem is also known as shared-key or private-key cryptosystems.

Asymmetric

In an asymmetric cryptosystem, on the other hand, the decryption key (Kd) is not equal to the encryption key (Ke). Furthermore, it is computationally impractical to derive Kd from Ke. Because of this property, only Kd needs to be kept secret and Ke is made publicly know. Asymmetric cryptosystems are also known as public – key cryptosystems.

The Data Encryption Standard (DES) cryptosystem is the best known and the most widely used symmetric cryptosystem today. On the other hand, the Rivest-Shamir-Adleman (RSA) cryptosystem is the first published and practically the most satisfactory asymmetric cryptosystem today.

Deadlock

There are several resources in a system for which the resource allocation policy must ensure exclusive access by a process. For example, a printer needs to be exclusively allocated to a process till the process finishes using it because its simultaneous use by two or more processes will led to garbed printed output. Since a system consists of a finite number of units of each resource type (for example, there printers, six tape drives, four disk drives, two CPUs, etc.), multiple concurrent processes normally have to compete to use a resource. In this situation, the sequence of events required to use a resource by a process is as follows:

1. Request. The process first makes a request for the resource. If the requested resource is not available, possibly because it is being used by another process, the requesting process must wait until the requested resource is allocated to it by the system.
2. Allocate. The system allocates the resource to the requesting process as soon as possible. It maintains a table in which it records whether each resource in free or allocated and, if it is allocated, to which process. If the requested resource is currently allocated to another process, the requesting process is added to a queue of process waiting for this resource. Once the system allocates the resource to the requesting process, that process can exclusively use the resource by operating on it.
3. **Release.** After the process has finished using the allocated resource, it resources, it release the resource to the system. The system table records are updated at the time of allocation and release to reflect the current status of availability of resources.

The request and release of resources are system calls, such as request and release for devices, open and close for files, and allocate and free for memory space. Notice that of the three operations, allocate is the only operation that the system can control. A process initiates the other two operations.

With the above mentioned pattern of request, allocation, and release of resources, if the total request made by multiple concurrent processes for resources in time. Care must be taken that the strategy applied cannot cause a deadlock, that is, a situation in which competing process prevent their mutual process even though no single one request more resources than are available. It may happen tat some of the processes that entered the waiting state (because the requested resources were not available in the time of request) will never again change state, because the resources they have requested were not available at the time of request) will never again change state, because the resources involved are said to be deadlocked. Hence, deadlock is the state of permanent blocking of a set of processes each of which is waiting for an event that only another process in the set can cause. All the processes in the set block permanently because all the processes are waiting and hence none of them will ever cause any of the events that could wake up any of the other members of the set.

A deadlock situation can be best explained with the help of an example (see Figure 2). Suppose that a system has two tape drives T1 and T2 and the resource allocation strategy is such that a requested resource is immediately allocated to the requesting process if the resource is free. Also suppose that two concurrent process P1 and P2 make requests for the tape drives in the following order.

1. P1 requests for one tape drive and the system allocates T1 to it.
2. P2 requests for one tape drive and the system allocates T2 and it.
3. P1 requests for one more tape drive and enter a waiting state because no tape drive is presently available.
4. P2 request for one more tape drive and it also enter a waiting state because no tape drive is presently available.

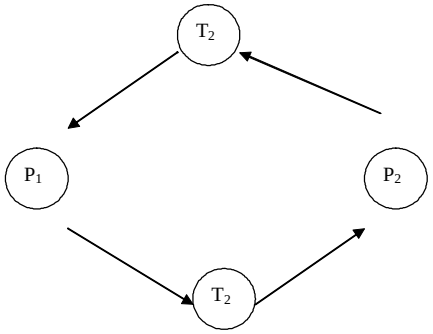


Figure 2 An example of a deadlock situation involving processes P1 and P2 and resources T1 and T2

From now on, P1 and P2 will wait for each other indefinitely, since P1 will not release T1 until it gets T2 to carry out its designated task, that is, not until P2 has released T2 , whereas P2 will not release T2 until it gets T1. Therefore, the two processes are in a state of deadlock. Note that the request made by the two processes are totally legal because each is requesting for only two tape drives, which is the total number of tape drives available in the system. However, the deadlock problem occurs because the total requests of both processes exceed the total number of units for the tape drive and the resource allocation policy is such that it immediately allocates a resource on request if the resource is free.

Give Real life example of Deadlock

Spooling

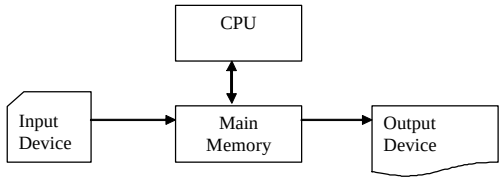
Spooling stands for simultaneous Peripheral Operation On-Line. It is a mechanism to deal with the following issues:

- To reduce the speed mismatch between slow I/O devices and CPU for better utilization of the CPU.
- To convert mutually exclusively I/O devices into non-mutually exclusive I/O devices so that they can be simultaneously used by multiple processes in a multiprogramming system.

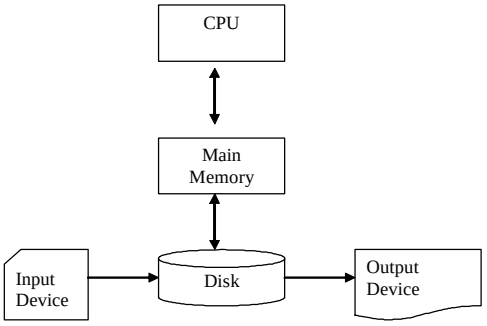
Dealing with Speed Mismatch

The speed mismatch between a slow I/O device such as a printer and the CPU is normally of the order of few thousand. Thus, while a slow I/O device is inputting/outputting a few bytes to/from the main memory, the CPU is idle for a large percentage of I/O time. Spooling reduces the idle time of the CPU by placing all data that comes form an input device or goes to an output device on a magnetic disk (see Figure 3). The primary reason for doing this is to keep the program and data readily available to the fast and expensive CPU on a high speed I/O medium such as a disk. Reading/Writing of data from/to a disk is much faster than reading/writing from/to a slow I/O device resulting in considerable reduction in CPU idle time during I/O operations involving slow devices.

In most computer systems, special low-cost I/O processors are used for spooling the input data from a slow input device on to the disk or for outputting the spooled output data from the disk on the slow output device. These I/O processors function independent of the main processor (CPU). This enables the main high-speed, expensive CPU to be fully



(a) Mode of data I/O without spooling facility.



(b) Mode of data I/O without spooling facility.

Figure 3 Illustrating spooling mechanism

Devoted to main computing jobs. The process of spooling is transparent to the user programs. In general, spooling makes better use of both the main memory and the CPU.

Dealing With Mutually Exclusive I/O Devices

There are several I/O devices that have to be used in dedicated mode for correct system operation. Such devices are called mutually exclusive devices. Printer is a typical example of such a device. For effective utilization of such devices, an operating system often converts them into non-mutually exclusive virtual devices by using spooling. For this, the operating system creates a special process, called a daemon, and a special directory, called a spooling directory. Now when a process makes a request to use the printer, instead of allocating the printer to the process, the operating system opens a file for the process's output data in the spooling directory, which is stored on the disk. All output data to be printed on the printer by the process is written in this file. The file is closed when the process completes printing. The daemon then actually prints the file form the disk on to the printer. By protecting the printer against direct use by the users, the problem of having someone keeping it opens unnecessarily long is eliminated.

LESSON 23
NETWORK ADMINISTRATION

Outline

- File security
- Workgroups(User Management)
  - Workgroups and files
  - Workgroups and file permissions

File Security

The Unix workgroups system provides flexible file security. Workgroups allow you to specify the level of access different types of users have to your files and directories. In other words, who can use your files and how.

Introduction to Workgroups

The basic idea of the workgroups system is to organize users who work together into groups. For example, at a university users might be divided into faculty, staff and student groups. At a business users might be grouped by job categories such as management, sales and secretarial. Users in the same workgroup can give each other special privileges to access files. For example, all the users in the managers group can edit the Employee Val file while all the users in the sales group can edit the NovSales file. Every user is a member of at least one group and users can be members of multiple groups. Execute the groups command to display the groups that you belong to.

For example,

```
$ groups
```

```
Managers Sales
```

shows that this user is a member of both the managers and sales groups. The first group shown is considered the primary group. In the above example, the primary group is managers.

Workgroups and Files

Every Unix file or directory is associated with exactly one user (also called the file's owner) and one group. A newly created file or directory is associated with the username and primary group of the person who created it. For example, suppose Kathy whose username is kathyr and primary group is sales starts a text editor and creates the file NovSales. The NovSales file will be associated with the user kathyr and group sales. To see what user and group a file belongs to use the long file list option, -l, with the ls command.[1]

```
$ ls -l
-rw-r--r--  1 kathyr  sales  0 Oct 08 OctSales
-rw-r--r--  1 kathyr  sales  0 Nov 08 NovSales
-rw-r--r--  1 kathyr  sales  0 Dec 17 EmployeeEvals
```

You can associate any file that you own with any group you are a member of using the chgrp command. For example, kathyr could associate the file EmployeeEvals with the managers group instead of the sales group.

```
$ chgrp managers EmployeeEval
$ ls -l
-rw-r--r--  1 kathyr  sales     0 Oct 08 OctSales
-rw-r--r--  1 kathyr  sales     0 Nov 08 NovSales
-rw-r--r--  1 kathyr  managers  0 Dec 17 EmployeeEvals
```

The -R option to the chgrp command recursively descends through a directory, including all subdirectories, and changes the group associations. For example, if EmployeeEvalDir is a directory with two subdirectories, SalariedDir and HourlyDir, each with several hundred employee evaluations

```
$ chgrp -R managers EmployeeEvalDir
```

will change the group associated with the EmployeeEvalDir, SalariedDir and HourlyDir directories as well as each evaluation within the directories.

Workgroups and File Permissions

A file's owner can control access to the file via the Unix file permission system. The file permission system allows you to define the level of access to individual files and directories for three different types of users.

Abb.	Person	Description
u	user	The user associated with the file (i.e. the file's owner).
g	group	Members of the group associated with the file.
o	other	Everyone else.

There are three kinds of access or permissions each different type of user can be given.

Abb.	Permission
r	read permission
w	write permission
x	execute permission

These permissions have different meanings when applied to a directory versus a regular file. Permissions allow the following access to a regular file.

Permission	Required To
read	• Copy and view a file. • Access a file with commands such as *cat* and *grep*.
write	• Edit, delete and overwrite a file.
execute	• Run a program or shell script.

Permission	Required To
read	• List directory contents with *ls*. • Access directory with commands such as *find*.
write	• Create, edit, rename and delete files and subdirectories within a directory.
execute	• *Cd* into a directory. • List directory contents with *ls*. • Create, edit, rename, access and delete files or subdirectories within the directory. • Execute a program or shell script within the directory.

A permission is said to be turned on if it is available. If not, it is said to be turned off. See the read, write and execute permission glossary entries for more detailed descriptions of these permissions.

File Permission List

The -l (long file listing) option with the ls command shows the permissions associated with files and directories.

```
$ ls -l
-rwxr--r--   1 kathyr   staff  1024 Oct 15 prog1
-rwxr-xr-x   1 kathyr   staff  1024 Nov 05 prog2
-rw-r--r--   1 kathyr   staff     0 Nov 08 file1
-rw-------   1 kathyr   staff    16 Nov 07 file2
drwx------   2 kathyr   staff   512 Nov 08 sdir/
```

The file permissions list is in the first section of ls output. It consists of 10 single character columns. The first column of the file permissions list is d if the listing is for a directory, - for a regular file as well as identifying some more unusual file types.

The next three columns define the permissions available to the user. They are in the order rwx (read, write, execute) and will have a - if that permission is turned off. If a file is not a program or script, it will not have the execute permission turned on, even for the user. The next three columns define the permissions available to members of the group associated with the file. The final three columns of the file permissions list define the permissions available to others.

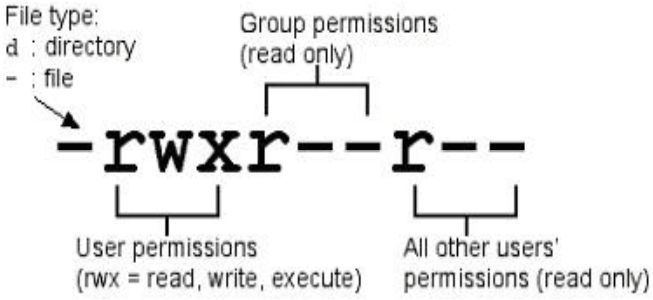


Figure 1: File Permissions List

In Figure 1, the permissions list indicates a regular file because the first column is a dash, -. The user has read, write and execute permissions. The group has read permission but not write or execute permission. Likewise others have only read access to the file.

Changing File Permissions

You can change a file or directory's permissions list using the chmod command.

chmod [who op permission] filename

who can be any combination of

u	(user)
g	(group)
o	(other)
a	(all) (i.e. ugo)

op sets, adds or removes permissions. It is specified as

=	(set permission exactly), or
+	(add permission), or
-	(remove permission) permission can be any combination of
r	(read)
w	(write)
x	(execute)

Examples

1. chmod u=rw,g=,o= file1

Set the permissions on file1 so that only the user has read and write access. The members of the file's group and others have no access to the file. Because the = operator is used, the permissions are set exactly.

2. chmod u=rwx, g=rx,o= dir1

This sets the permissions on the directory dir1 so that the user has read, write and execute access, the group has read and execute access and others have no access.

3. chmod ug=rw,o= file1

Set the permissions on file1 so that both the user and members of the file's group have read and write access. Others have no access to the

file.

4. chmod u=,g=rw,o= f1

Set the permissions on f1 so that the group has read and write access while the user and others have no access. .

5. chmod g+w file1

For file1, add write permission for the group. There is no affect on any other permissions.

6. chmod a-w,a+r read only

Remove write and add read permission to the file read-only for everyone (user, group and other).

7. chmod go-wr *

Remove read and write permission for the group and others for all All files and current directories in a file

8. chmod go-xwr dir1

Remove all permission for the group and others for the directory dir1. They can no longer list the contents of the directory with the ls command, cd into the directory or use files within the directory.

Note:

LESSON 24
GROUP DISCUSSION
CONTRAST THE FUNCTIONS OF DIFFERENT TYPES OF OPERATING SYSTEMS

Note:

LESSON 25
GROUP DISCUSSION
SECURITY IS AN IMPORTANT FACTOR

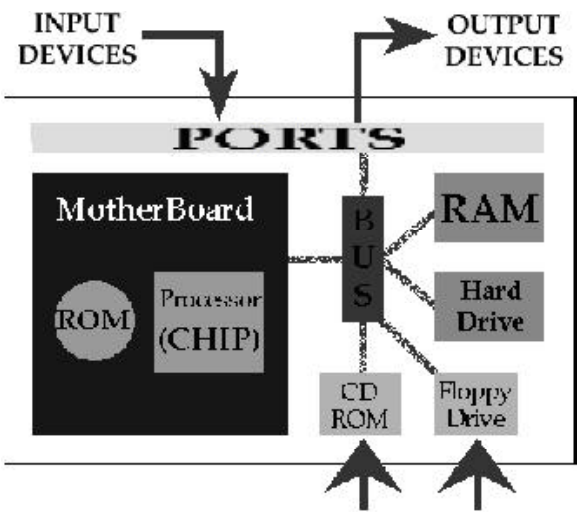
Note

LESSON 26
UPGRADING OPPORUNITIES- ADDING OF CARDS

Outline

- What ports can you add?
- Buying a port expansion card
- Adding a new port card
- Adding an external USB hub
- Adding ports to a portable PC

Diagram of the CPU



Here is an indisputable reason to upgrade your computer system: you buy a new peripheral, and have no free ports to connect it to. It happens, and it happens a lot. Let's say you want to add a second printer to your system (like a color photo printer; they're quite popular these days). Most printers only come with a single parallel port, to which you probably have your existing printer connected. So when you go to plug in your new printer, you're out of luck.

The answer, of course, is to add a second parallel port to your system. (Another solution is to buy a USB printer and connect to a free USB port; let's not ignore easy fixes!)

Or maybe you've become enamored of USB devices, and completely filled all of your system's USB connections. Buy one more USB-compatible peripheral, and then what do you do? Here you have a choice; you can increase the number of connections available by adding an external USB hub, or you can add an extra USB card to your system.

Here's another one. You want to connect your new digital video camcorder to your PC to edit your home movies. But your camcorder connects via Fire Wire, and your PC doesn't have a Fire Wire connector. The solution? Add a Fire Wire port to your system unit.

This chapter shows you how to add extra ports to your computer system. Some of these upgrades are external (like the USB hub) and some require opening up your PC's case. Fortunately, they're all relatively easy procedures.

What Ports Can You Add?

Let's do a quick rundown on the types of ports you can add to your computer system:

- Parallel ports are used primarily for connecting printers to your system
- Serial ports are used to connect some types of mice, keyboards, and printers, as well as PDAs, PC cameras, and other assorted peripherals
- USB ports are used to connect just about any type of peripheral; peripherals conforming to the new USB 2.0 standard run much faster than the existing 1.1 devices
- FireWire (IEEE 1394) ports are used to connect fast digital devices, such as digital video recorders and digital still cameras
- SCSI ports are used to connect fast external hard drives

Caution

To install a USB 1.1 port on your system, your computer must be running Windows 98 or later. To install a USB 2.0 port, your computer must be running Windows XP with Service Pack 1 installed. To install a Fire Wire port, your computer must be running Windows 98 or later.

Most computers today come with a single parallel port, one or two serial ports, a Fire Wire port, and anywhere from a couple to a half-dozen USB ports. As you can see, if you add enough new devices to your system, you easily run out of available connectors.

Buying A Port Expansion Card

Fortunately, adding a new port-of any kind-is as easy as installing a new card (like the one in Figure 1) in your system unit.

Figure 1 A typical port expansion card-this one adds two USB ports.

In addition, many port expansion cards contain more than one type of port. So you can install one card to gain extra serial and parallel ports, or get a combination of USB and Fire Wire ports on a single card.

Note

If you have the choice, you're better off buying a peripheral that plugs into an existing USB port than one that requires a new

parallel or serial connection. It's easier to connect a device to an open USB connector than it is to install a new parallel or serial card.

When you're shopping for a port expansion card, check out the product models from these manufacturers—most of these companies also make external port expansion devices:

- Adaptec (http://www.adaptec.com)
- Belkin (http://www.belkin.com)
- IOGEAR (http://www.iogear.com)
- SIIG (http://www.siig.com)
- StarTech (http://www.startech.com)

Upgrade #1: Adding A New Port Card

To install a port expansion card, grab your Phillips head screwdriver and follow these steps:

1. Close Windows and power off your PC.
2. Disconnect your PC from its power source.
3. Open the system unit case.
4. Locate an unused PCI card slot (see Figure 2).

Figure 2 Locate an empty PCI slot; if you have several empty slots, leave an empty slot between an old card and the new one you're installing.

5. Remove the cover to the selected slot (on the back of the system unit).
6. Insert the card into the open slot (see Figure 3).

Figure 3 Use firm pressures to push the card into the slot—and make sure that the card is firmly seated.

7. Screw the card bracket into place, using the screw from the slot cover (see Figure4).

Figure 4 Be careful when screwing the card into place; it's easy to slip and damage nearby components with the tip of your screwdriver.

8. Reinstall the system unit cover and reconnect the PC to its power source.
9. Restart your computer.
10. When prompted, install the drivers for your new card.

That's it. As soon as your computer is configured for the new ports, they're ready to use!

Upgrade #2: Adding an External Usb Hub

Here's an even easier upgrade. If your system already has one or more USB connectors, you can add extra ports by using an external device called a USB hub, shown in Figure 5. The hub connects to an open USB port on your PC; you can then connect multiple USB devices to the hub.

Figure 5 Add a USB hub to increase the number of USB ports on your system.

Follow these steps to connect a USB hub to your system; there's no need to turn off your computer before proceeding.

1. If you don't have a free USB port on your system, unplug one of your current USB peripherals.
2. Plug the USB hub into the open USB port on your system unit.
3. If you purchased a self-powered USB hub, connect it to its power source.
4. Connect any new or existing USB peripherals to the USB hub.

Upgrade #3: Adding Ports to a Portable Pc

Adding a new port to a portable PC is particularly problematic. (My apologies for the alliteration!) You can't add a new port via internal card, so you're limited to going the USB route, or adding an expansion port on a PC Card. Either of these options is equally viable.

If you go the PC Card route, you need to be aware of which type of PC Card slot you have on your portable. Older portable PCs used a 16-bit PC Card slot, which isn't very versatile; you can't use it for USB, FireWire, or fast SCSI adapters. Most newer portables have 32-bit PC Card slots, which can handle just about anything you throw at them. (Check the specifications page of your PC's instruction manual to determine what type of PC Card slots you have.)

Caution

Some older PCs have trouble working with some external USB hubs; they might require constant rebooting to recognize the hub. If this happens to you, try plugging the hub into a different internal port, or try using a different brand of hub.

When it comes time to expand the ports on your PC, remember these important points:

- You can add just about any type of port to your system by installing a port expansion card.
- To install a port expansion card, all you need is a free PCI expansion slot-and a screwdriver.
- An easier way to expand the number of USB ports in your system is to use an external USB hub.
- You can expand the ports on a portable PC via either USB or PC Card.

What should be the health issues while upgrading ?

LESSON 27
GROUP DISCUSSION
UPGRADING OPPORTUNITIES OF YOUR COMPUTER SYSTEM

Note:

LESSON 28
UPGRADING AND INSTALLATION

Outline

- Upgrading
- Upgrading Opportunities
- Guidelines
- Installing And Uninstalling Software

Upgrading

Upgrading to a new computer is like moving into a new house

PC Upgrade Commander analyzes the applications, data and settings, as if they were rooms in a house. Then it safely packs the items up, loads them into the moving van, ships them to the new home, where it safely unpacks the items and places them in the proper rooms of the new computer!

It is commonly assumed by many struggling to keep up with technology that a computer has reached obsolescence the day it is purchased. However, while year-old or even five-year-old computers may not be able to run the latest feature-ridden applications, they can be perfectly adequate for working on a network, browsing the Internet, producing documents, and doing other nonprofit work. Sometimes, those older computers can perform even better with the right adjustments.

Upgrading Opportunities

You can Upgrade

- RAM
- Hard drive
- Modem etc

Name other things that you can use to upgrade your system.

Upgrading your computer's RAM, hard drive, modem, or other hardware can be a cost-effective way to improve the technology available to you and increase productivity when your budget is tight. While it is always nicer to be able to purchase the fastest computers on the market, sometimes it may not be necessary. Here are some guidelines for deciding when to upgrade:

Is An Upgrade the Answer?

Many times, you choose to upgrade computers for the wrong reasons. You can't turn a slow computer into a fast one simply by increasing the size of a hard drive. Likewise, changing the computer's processor may not make a crashing computer more reliable.

Guidelines

Upgrading a computer's RAM, hard drive, or processor almost always involves getting under the hood of a computer. This means that you need to know something about what you are doing to avoid causing damage. While replacing an older modem with a faster model is not rocket science, it can be a daunting (but rewarding) experience for the novice user. . Outsourcing upgrades to a technician is another option, but that cost additional time and money.

Type of Upgrade	Guidelines	Difficulty
RAM	A RAM upgrade can be a quick and easy way to speed up your computer. This is the best route to making any computer faster because more RAM means less time that your computer must pause to access the hard drive. One way to tell if you need more RAM is to see and hear if your hard drive is constantly being accessed. A more precise way is to look at your system's Task Manager which will provide information on your system's performance.	Moderate
Modem	Upgrading a computer's modem will effectively speed up download time from the Internet; but most computers already equipped with a 56k modem, which is about as fast as they get. In order to get faster internet access you need to upgrade your Internet service to DSL or cable. New modems are usually provided with your DSL or cable service.	Moderate
Hard drive	Replacing a hard drive can be a momentous task because of the need to transfer files from the old drive to the new. You would also need to reinstall the operating system and all the software. It may be best to consider adding an additional drive instead of entirely removing the older one. Most motherboards can handle up to four hard drives.	Advanced
Processor	This is not an easy upgrade, and should be approached with caution. A faster chip will help make the computer process applications faster, but the additional speed may be negligible without sufficient RAM. The difficult part about installing a new processor is finding one that will work with your motherboard.	Advanced

In conclusion, the easiest way to speed up an older computer is too add more RAM. If your computers seem kind of sluggish and the programs your running seem to take up a lot of memory, RAM will make your computer perform a lot faster. If your computer is low on storage, it can also be quite easy to add an additional hard drive.

Ask yourself whether a component upgrade is worth it before embarking on that course of action. Many times it may be better to make do with the equipment available until it is time to make another planned computer purchase. Upgrading computer components can help in certain situations when you know what you are doing, but it is important to balance the time and cost of making the upgrade versus the time and productivity you may be losing from not making it.

The best advice I can give is to not run off and upgrade your computers unless you really need to. Most work can be done quite effectively without a high performance machine, but your shouldn't force to work on machines that constantly crash or slow down when running your basic productivity software.

Installing And Uninstalling Software

Installing

Most software that you purchase will be distributed on CD-ROM or Diskette Sets. If the software is distributed on CD-ROM, inserting the disk into the CD-ROM drive will probably launch the setup process. All you will have to do is watch the screen and agree to a few questions. The program will be automatically installed and additions will be made to the

system.

Software that is sold on diskette sets will not activate the self-installation procedures. You will have to do this manually. Follow these easy steps:

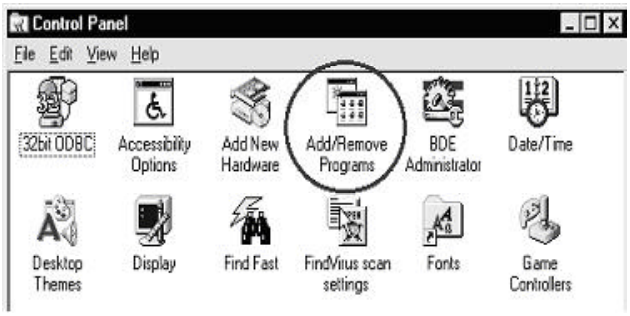
1. Insert the first diskette (Disk 1) in the diskette drive slot
2. Click the 

button and choose "Run" from the menu
3. Type "a: setup" (without quotes) and click on the OK button

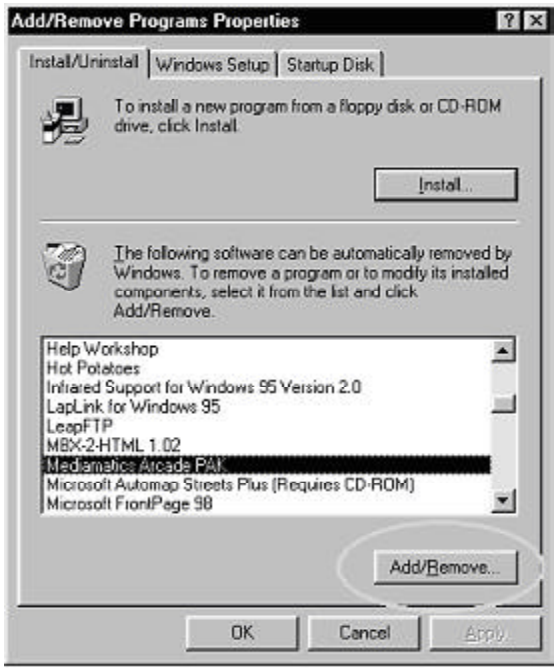
The installation program will begin. If the program is distributed on more than one disk, the setup program will ask you to exchange disks periodically. Be sure that you insert the disk than the program specifies.

Uninstalling

In the event that you wish to remove software from the computer, most commercial products build in an Uninstall routine when they are installed. To best method of accessing this routine is to open the Control Panel. Click the Start Button, move the mouse pointer upward until it is on top of the word "Settings".A submenu will appear that contains the Control Panel icon. Move your mouse to the right until the arrow is over the Control Panel Folder and then click the left mouse button once. Doing so will bring up the following screen:



Clicking the "Add/Remove Programs" icon will then allow you to uninstall software from a list. Carefully choose the software package that you want to remove and then click the "Add/Remove" button. The computer will take care of the rest. It is a really good idea to make sure that you don't have the program running on the computer when you try to uninstall it.



LESSON 29
GROUP DISCUSSION
TOPIC – MARKET RESEARCH

Note:

LESSON 30
DATA COMMUNICATION SYSTEM

Outline

- Definition
- Why have data communication?
- General Communication System
- Signal Types
- Bits Bytes Packets
- Transmission Impairments
- Communication Modes

Definition

- Two or more computers connected via a communications medium
- A host computer together with communications equipment and terminals
- The hardware and software that enable the user to access a remote device.

Why Have Data Communication?

Communication between Computers Allows:

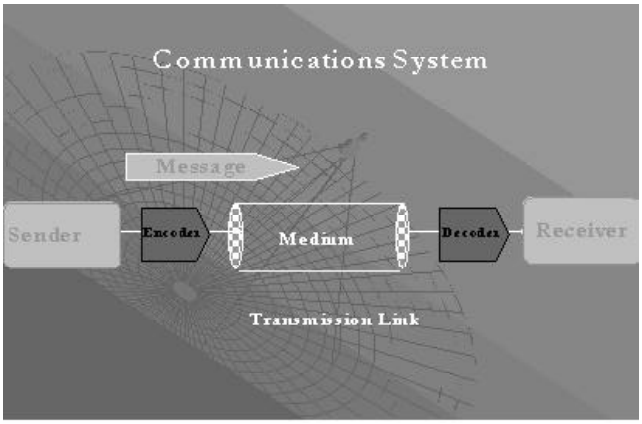
- Sharing of resources
- Distribution of processing functions
- Opportunity for centralized management control
- Compatibility of dissimilar equipment and software
- Increased performance with reduced cost.

List some resources that can be shared communication networks

General Communication System

Essential Features:

- Sender
- Receiver
- Medium
- Message



Signal Types

- Message may take many forms in data communications, - e.g. a user file containing graphics, text, or sound data, or machine generated request, or status message.
- Senders and receivers may be humans using a computer or a variety of programmable devices.
- Medium used can be space, telephone wire, optic fiber etc.
- Encoders translate message into signals suitable for medium and back again for receiver (e.g. modem).
- Data is usually transmitted using one of three types of signals:
  - An electrical voltage (co-axial cable, wire);
  - An electromagnetic wave (satellite, wireless LAN); or
  - Light pulses (optic fiber).
- E.g. phone line carries analog electrical voltage signals to represent sound.
- Data is represented by CHANGES in the signal as a function of time.

Protocols/Standards

Sets of rules or standards that let computers communicate over the Internet. A formal description of message formats and the rules two computers must follow to exchange those messages. Protocols can describe low-level details of machine-to-machine interfaces (e.g., the order in which bits and bytes are sent across a wire) or high-level exchanges between allocation programs (e.g., the way in which two programs transfer a file across the Internet).

- HTTP (Hypertext Transfer Protocol) allows transfer of Web pages via a browser.
- FTP (File Transfer Protocol) – Allows the transfer of one or more files from one machine to another across the Internet

- TCP/IP Transmission Control Protocol / Internet Protocol) - The protocols, or conventions, that computers use to communicate over the Internet.
- SMPTP Simple Mail Transfer Protocol – A protocol used to send e-mail on the Internet. SMTP is a set of rules regarding the interaction between a program sending e-mail and a program receiving e-mail.

Name Five More Protocols

BITS, BYTES AND PACKETS

- Another important role of protocols is to determine format of data being sent, so sender & receiver agree on meaning of stream of bits.
- Message may be sent as single bytes, complete message, or (most commonly) message broken up into pieces called packets, with exact format determined by protocol.
- Each packet contains header information such as source & destination address, amount of data, error-checking code.

Bandwidth

- Bandwidth determines amount of data that can be transmitted per unit time. Greater bandwidth = greater possible data transmission rate.
- Available bandwidth varies between different media, e.g. co-axial cable, optic fiber, satellite etc all have different bandwidth characteristics.
- Each media can only transmit signals with certain ranges of frequencies. Higher the central frequency, greater the possible frequency range, & greater potential bandwidth of the media.

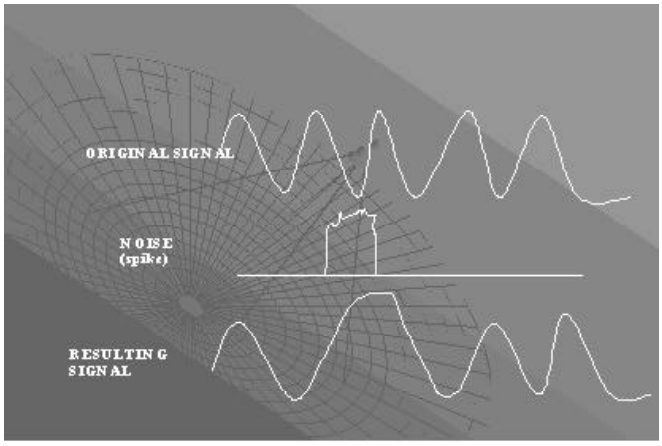
Transmission Impairments

Final signal received differs from the one sent because of impairments to transmission, possibly resulting in errors. Impairments are:

a) Noise - caused by various factors: thermal, crosstalk, and impulse.

b) Attenuation - signal strength decreases over distance. Corrected by use of repeaters (digital) and amplifiers (analog).

c) Delay Distortion - Higher frequency parts of signal travel faster (affects especially digital data).

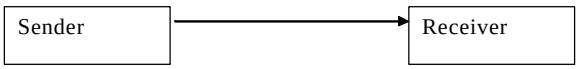


Which device we use to deal with transmission Impairments?

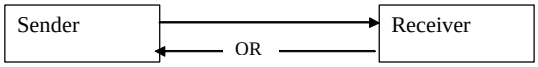
Communication Modes

Elementary types of data flow over a transmission path.

a) Simplex - only allows information flow in one direction. Devices connected to such a circuit are either a send only or receive only device. For example, a data collection terminal on a factory floor or a line printer. - e.g. host to dumb terminal, radio & TV broadcast

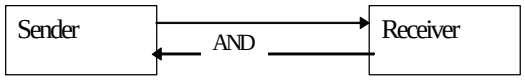


b) Half Duplex - allows information flow in both directions, but only one at a time. It requires two wires this is the most common type of transmission for voice communication because only one person is supposed to speak at a time. It is also used to connect a terminal with a computer. The terminal might transmit data and then the computer responds with an acknowledgement. E.g. press-to-talk radio systems, older modems.



c) Full Duplex - allows simultaneous information flow in both directions. This involves a special switching circuit and requires a small amount of time is approximately 150 milliseconds. With high-speed capabilities of computer, this turnaround time is unacceptable in many instances. Also some applications require simultaneous transmission

in both the directions. In such cases, a full duplex system is used that allows information to flow simultaneously in both directions on the transmission path. It requires four wires. E.g. (common in computer communications), telephone.



Note:

LESSON 31
TRANSMISSION MEDIA

Outline

- Conducted Media (guided)
  - Twisted Pair (transmit electrical signals)
  - Coaxial Cable (transmit electrical signals)
  - Optical Fiber (transmit light)
- Radiated Media (unguided. transmit electromagnetic waves)
  - Microwave
  - Radio wave
  - Infrared
- Media selection criteria

Conducted Media

Wires

- Wires are most commonly used medium due to availability & low cost.
- Transmit electrical signals.
- Disadvantage
  - Low transmission rates over long distances due to relatively limited bandwidth.
  - Max data rate depends on shielding, encoding scheme etc.

New technology allows incr. Rates

- 3 Basic types
- Single wire
- Twisted pair
- Bundled wires

i) Twisted Pair

- Two insulated copper wires twisted together in a spiral pattern to reduce external interference. Used in LANs & existing telephone system.
- May be shielded or unshielded, & max data rates vary depending on degree of twisting and shielding.
- Can carry both analog and digital signals.
- Signal susceptible to interference and attenuation. Repeaters (digital) and amplifiers (analog) are required at least every few km (depending on data rate etc).

ii) Coaxial Cable

- Can operate at higher frequency (higher data rate) than twisted pair for both analog and digital data.
- Typical data rate is 10 Mbps (> 500Mbps possible)
- Single cable may be divided into several channels, using "frequency guard bands" to prevent signals from interfering with each other. This is called BROADBAND transmission. It allows the cable to be used for a variety of transmission needs - voice, data, video.

Advantages

- High transmission capacity
- Less susceptible to cross talk and other external electromagnetic signals (which could corrupt data transmission) due to extensive shielding

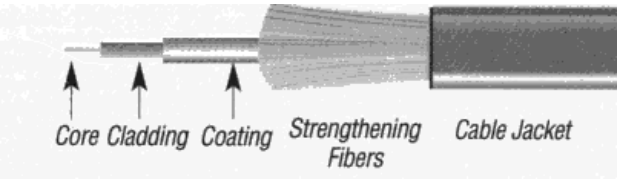
Disadvantages

- Cost higher than twisted pair

Give some practical uses of co-axial cables

iii) Optical Fiber

- Thin, flexible fiber, made of glass or plastic woven together to form core. Core surrounded by glass/plastic cladding, then plastic protective layer
- Transmits light signals
- Uses two types of Transmitters
  - LED (Light Emitting Diode)
  - ILD (Injection Laser Diode) for long distance transmission
- Uses a photo-detector for receiver



Advantages

- Very large bandwidth (very high data rates)
- Attenuation is lower than wire/cable
- Electromagnetic isolation
- Good security

Disadvantages

- Cost

- Difficult to extend/add to

Radiated Media

Microwave

- Common media for long distance transmission (around 50 Km)
- Works in direct line of sight due to high frequencies used (1011 Hz) ? Attenuation can be problem, and worsened by rain. Interference from other microwave signals is high; hence assignment of frequencies is controlled.
- Potential data rate is around 1 Gbps.
- Commercial transmission speed approx 100Mbps

Satellite

- Conceptually a microwave relay station in space.
- Point-to-point broadcast transmission.
- Older satellites stationed in equatorial, geosynchronous orbit at altitude of 35,880 kms
    - Geosynch. Orbit means satell. Remains stationary relative to a given position on earth
- –At this altitude only 3 satell. Required to have all points on earth within range
- Newer systems use multiple low cost satellites in low (700 - 2000 kms) and medium (~10,000kms) earth orbits.
- Uplink and downlink frequencies are different
- Transmission speed approx 50 Mbps
- Propagation delay - amount of time it takes for a signal to travel from source to destination
- Signal travels at speed of light approx. 300k km/sec. This is insignificant on terrestrial link: (5 km link 17ms), but may be significant on satellite link and noticeable during voice communication.

Infrared Transmission

- Uses light signals at frequency between visible light & radio waves.
- Line-of-sight technology
- High bandwidth, but over small distance only
- Uses – LANs, device connection e.g. printer to computer

Media Selection Criteria

Many different criteria. Most important one is often:

1. Cost - made up of many components.
2. Medium e.g.: Coaxial cable vs. fiber optic
3. Leased link vs. switched
4. Hardware e.g. Modems vs. fiber optic transceiver, repeaters, multiplexes
5. Software
6. Installation
7. Maintenance e.g. parts, personnel, test equipment
8. Security - ease of interception/tapping
9. Error Rate - immunity to noise
10. Environment - government regulations
11. Capacity - data rate
12. Response Time - speed and propagation delay
13. Availability and Expandability

Note:

LESSON 32
COMPUTER NETWORKS

Outline

- Network components
- Network classification
- Switching Techniques
- Broadcast networks
- Switched networks
  - Circuit
  - Message
  - Packet
  - virtual

Network Components

A computer network made up of components such as:

- Computers (servers and workstations)
- Network interface cards (or similar connection device)
- Network cabling, hubs & other specialist hardware
- Network operating system &/or other specialized software

Network Classification

A network is classified according to some of the following characteristics:

- Size - LAN, MAN, WAN
- Ownership - Public [Internet, Telstra], Private
- Transmission Mechanism - Switched (Packet and Circuit)
- Broadcast
- Topology

Network Categories

- WANS are used when computers are located in distant sites, generally use a mesh topology. Often involve public switched data/telephone network (PSTN, PSDN).
- MAN is a special WAN covering a metropolitan area.
- LANS generally confined to a limited geographical area - building, industrial complex, university campus, involving short distances (0.1 - 50 Km).
- LANS involves generally one private owner - public carrier not involved, therefore deregulated. Able to support high data rates between attached stations

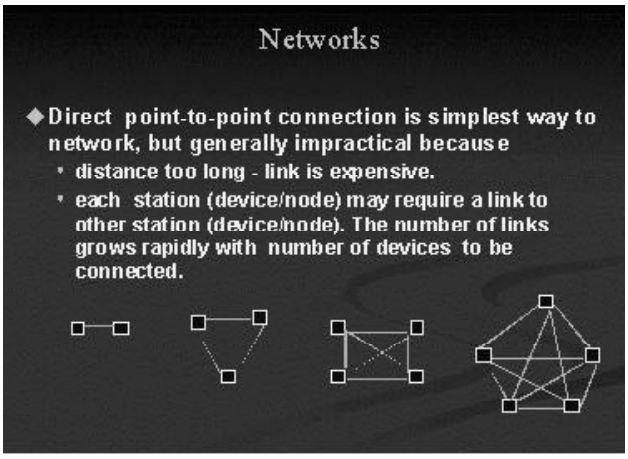
Networks are broadly classified into two types: local area networks (LANs) and wide-area networks (WANs). The WANs are also referred to as long-haul networks. They key characteristics that are often used to differentiate between these two types of networks are as follows:

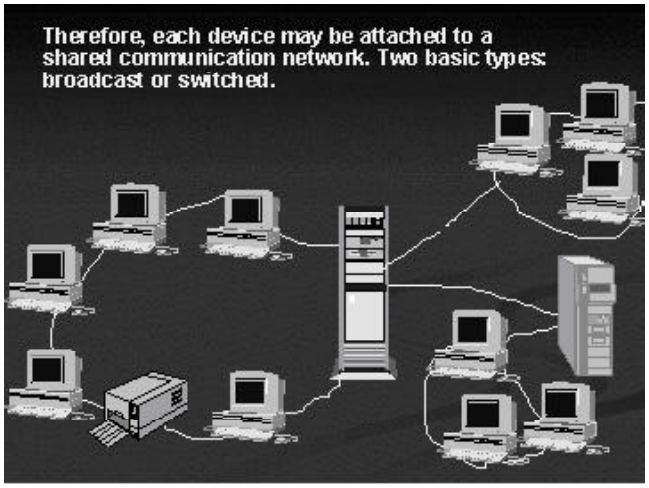
1. Geographic distribution. The main difference between the two types of networks is the way in which they are geographically distributed. A LAN is restricted to a limited geographic coverage of a few kilometers, but WAN spans greater distances and may extend over several thousand kilometers. Therefore, LANs typically provide communication facilities within a building or a campus, whereas WANs may operate nationwide or even worldwide.
2. Data Rate. Data transmission rates are usually much higher in LANs than in WANs. Transmission rates in LANs usually range from 0.2 Megabit per second (Mbps) to 1 Gigabit per second (Gbps). On the other hand, transmission rates in WANs usually range from 1200 bits per second to slightly over 1 Mbps.
3. Error Rate. Local area networks generally experience fewer data transmission errors than WANs do. Typically, bit error rates are in the range of 10-8 to 10-12 with LANs as opposed to 10-5 to 10-7 with WANs.
4. Communication Link. The most common communication links used in LANs are twisted pair, coaxial cable, and fiber optics. On the other hand, since the sites in a WAN are physically distributed over a large geographic area, the communication links used are by default relatively slow and unreliable. Typical communication links used in WANs are telephone lines, microwave links, and satellite channels.
5. Ownership. A LAN is typically owned by a single organization because of its limited geographic coverage. A WAN, however, is usually formed by interconnecting multiple LANs each of which may belong to a different organization. Therefore, administrative and maintenance complexities and costs for LANs are usually much lower than for WANs.
6. Communication Cost – the overall communication costs of a LAN is usually much lower than that of a WAN. The main reasons for this are lower error rates, simple (or absence of) routing algorithms, and lower administrative and maintenance costs. Moreover, the cost to transmit data in a LAN is negligible since the transmission medium is usually owned by the user organization. However, with a WAN, this cost may be very high because the transmission media used are leased lines or public communication systems, such as telephone lines, microwave links, and satellite channels.

Isn't there a network for all purposes?

Switching Techniques

1. Direct point-to-point connection is simplest way to network, but generally impractical because
   - Distance too long - link is expensive.
   - Each station (device/node) may require a link to other station (device/node). The number of links grows rapidly with number of devices to be connected.





1.Broadcast Networks –
No intermediate switching nodes. Transmission from one node sent over medium shared by all other nodes.•E.g.

- Bus topology LANs, all nodes attached to common cable
- Satellite networks sharing same frequency channel for transmission
- Packet radio sharing same frequency

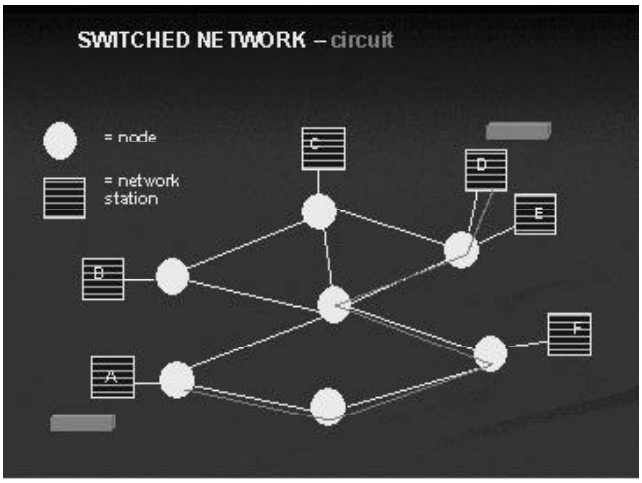
2.Switched Network

Means data sent via intermediate nodes, which are connected

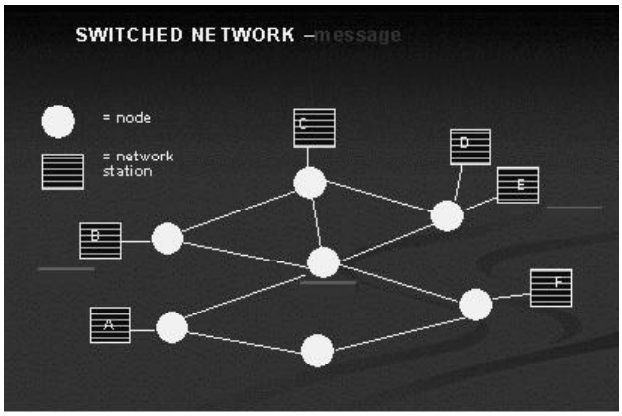
- Point-to -point. Used when data needs to be transferred beyond local area, in WANs and MANs.
- Four major types:

a) Circuit switched: dedicated circuit between source & destination nodes maintained for duration of communication. Relatively inefficient as entire capacity of channel dedicated for duration of connection. Suitable for light, intermittent transmission load. E.g.

Give example of circuit switching

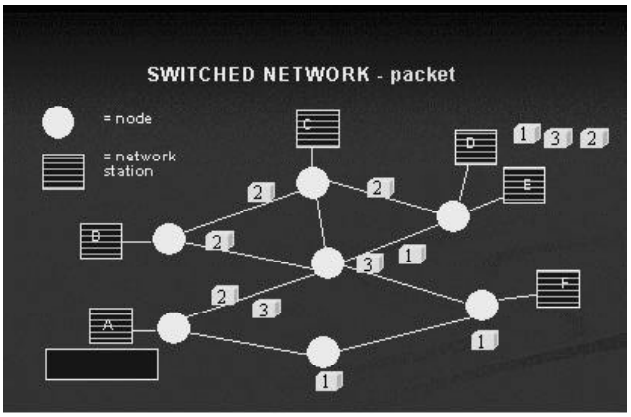


b) Message Switched: entire message is sent but it can be Stored and forwarded through intermediate nodes. Sender & Receiver do not have to be simultaneously available. E.g. email.

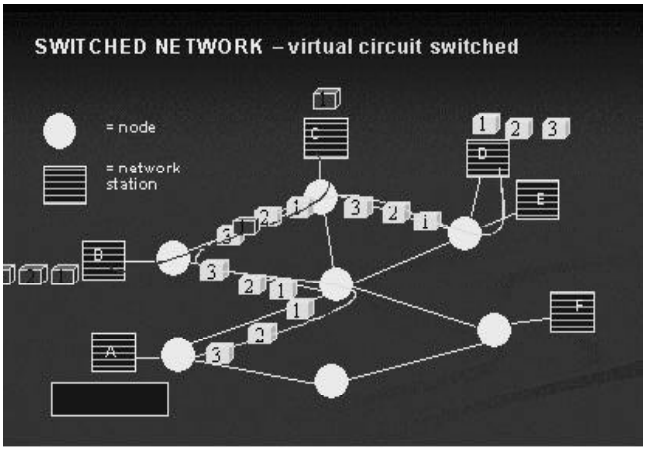


c) Packet Switched: (also called data gram) each packet is Allocated an independent pathway from source to

Destination. Packets can be stored and forwarded later if node not available. e.g. File transfer using TCP/IP. Much more efficient use of network capacity, but considerable Complexity with creating packets, headers, sequencing, recovery from errors in individual packets



d) Virtual circuit switched: creates route from source to destination before transmission starts, & all packets sent via same route.Packets arrive in correct order but route can still be shared with other messages = More efficient than circuit, less overhead than packet switched



Note:

LESSON 33
NETWORK TOPOLOGIES

Outline

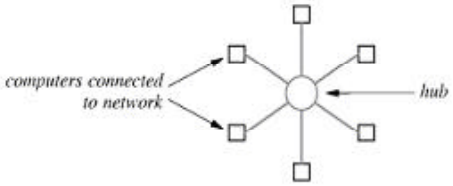
- Topologies
- Different types of Topologies
  - Star
  - Ring
  - Bus
  - Mesh
  - Tree

Network Topologies

Network is a logical extension of a data communication system. In a computer network, two or more computers (often referred to as nodes) are linked together with carriers and data communication devices for the purpose of communicating data and sharing resources. The term network topology refers to the way in which the nodes of a work are linked together. It determines the data paths that may be used between any pair of nodes in the network. Although the number of possible network topologies is seemingly limitless, the four major ones are the _ network, the ring network, the completely connected network, and the multi-access bus network.

Star Topology

All computers attach to a central point:

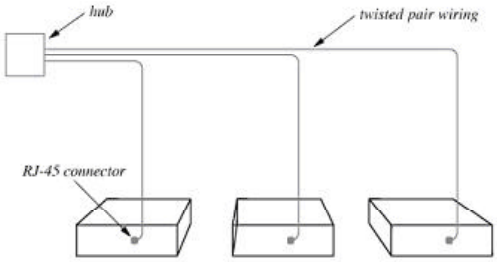


- Center of star is sometimes called a hub

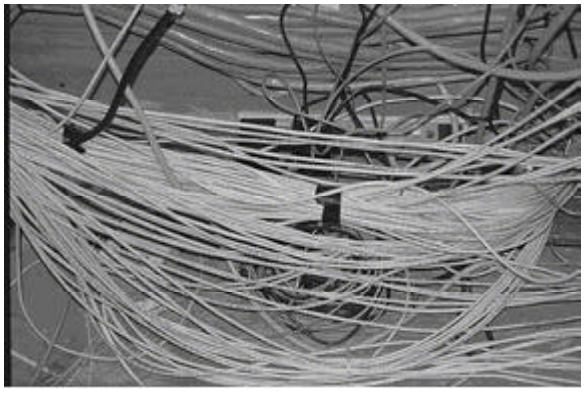
Star Topology in Practice

Previous diagram is idealized; usually, connecting cables run in parallel to computers:

All traffic emanates from the hub of the star. The central site is in control of all the nodes attached to it. The central hub is usually a fast, self-contained computer and is responsible for routing all traffic to other nodes.



- Result is:



Advantages

1. Star topology has minimal line cost because only n-lines are required for connecting n nodes.
2. Transmission delays between two nodes do not increase by adding new nodes to the network because any two nodes may be connected via two links only.
3. If any of the local computers fails, the remaining portion of the network is unaffected.

Disadvantage

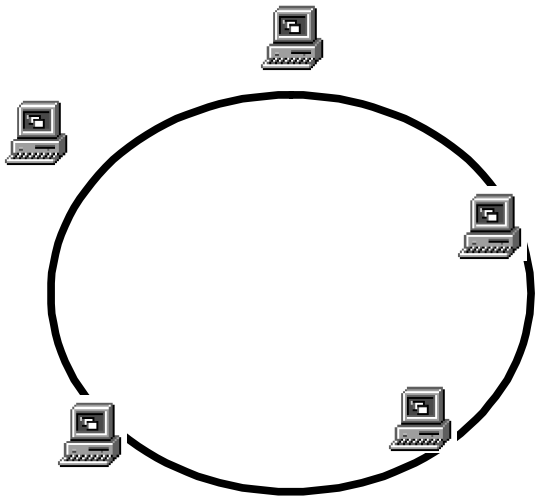
1. The system crucially depends on the central node. If the host computer fails, the entire network fails.

Ring Topology

Computers connected in a closed loop

- First passes data to second, second passes data to third, and so on
- In practice, there is a short connector cable from the computer to the ring

- Ring connections may run past offices with connector cable to socket in the office:



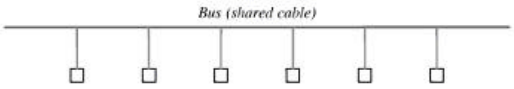
Advantages

1. The ring network works well where there is no central-site computer system.
2. It is more reliable than a star network because communication is not dependent on a single host computer. If a link between any two computers breaks down, or if one of the computers breaks down alternate routing is possible.

Disadvantages

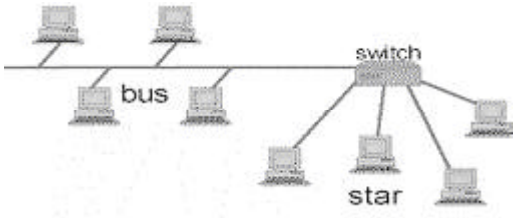
1. In a ring network, communication delay is directly proportional to the number of nodes in the network. Hence addition of new nodes in the network increases the communication delay.
2. The ring network requires more complicated control software than star network.

Bus Topology

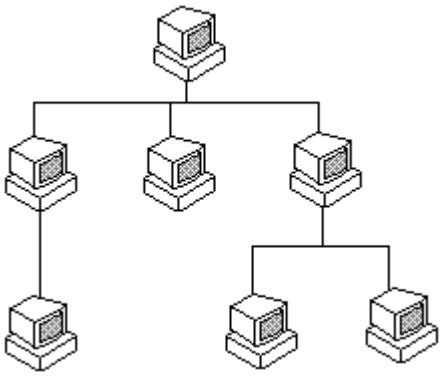


- Single cable connects all computers
- Each computer has connector to shared cable
- Computers must synchronize and allow only one computer to transmit at a time

Tree Topology



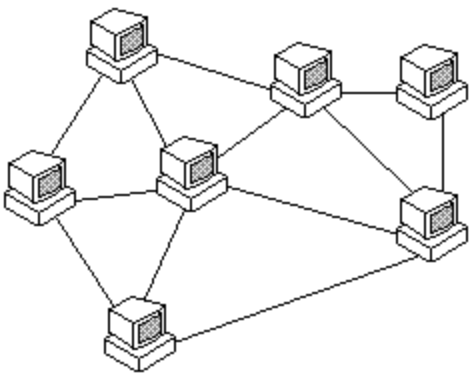
Also known as the 'Hierarchical topology', the tree topology is a combination of bus and star topologies. They are very common in larger networks. A typical scenario is: a file server is connected to a 24 port switch. A cable goes from the switch to a computer room where it connects to another switch. Many cables pass from this switch to the computers in the computer room. The node at the highest point in the hierarchy - usually a file server - controls the network.



In the topologies shown above, there is only one possible path from one node to another node. If any cable in that path is broken, the nodes cannot communicate.

Mesh Topology

Mesh topology uses lots of cables to connect every node with every other node. It is very expensive to wire up, but if any cable fails, there are many other ways for two nodes to communicate. Some WANs, like the Internet, employ mesh routing. In fact the Internet was deliberately designed like this to allow sites to communicate even during a nuclear war.



Advantages

1. This type of network is very reliable, as any link breakdown will affect only communication between the connected computers.
2. Each node of the network need not have individual routing capability.
3. Communication is very fast between any two nodes.

Disadvantage

1. It is the most expensive network from the point of view of link cost. If there are n nodes in the network, then n (n-I)/2 links are required. Thus, the cost of linking the system grows with the square of the number of nodes.
2. Which topology is widely accepted by the Industry?

Note:

LESSON 34
NETWORK PROTOCOLS & OSI REFERENCE MODEL

Outline

- Protocols/standards
- OSI model-history & definition
- Role of each OSI layer
  - Physical
  - Data
  - Network
  - Transport
  - Session
  - Presentation
  - application
- TCP/IP example

Protocols/ Standards

A standard way of communicating across a network. A protocol is the "language" of the network. A method by which two dissimilar systems can communicate. TCP is a protocol which runs over a network.

- Computer communications require a high degree of cooperation between parties involved in the transaction. Implementation of this cooperation is called a communications protocol/standard.
- A data communications protocol may include rules for:
  - Signal definition;
  - Flow control;
  - Formatting of data;
  - Addressing;
  - Error handling; and
  - Sequencing

Data communications standards may arise from 3 sources:

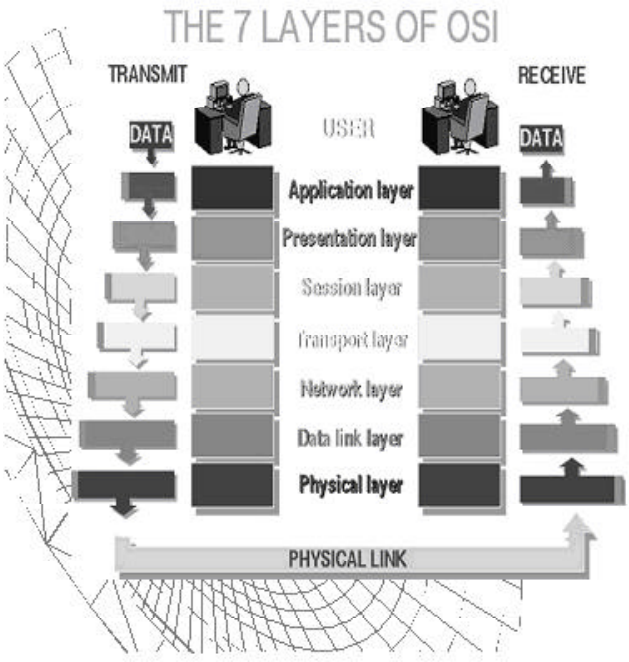
- By Use - Standards that become popular by everyone adopting them. Eg Ethernet, TCP/IP
- Proprietary - Standards developed by a vendor for use by that vendor's devices, typically used only in vendor's own networks eg IBM SNA, Apple AppleTalk.
- By Law - Standards developed/adopted by standards organizations such as IEEE, ISO, eg IEEE 802 protocols for LANs.

Examples :

- HTTP (Hypertext Transfer Protocol) allows transfer of Web pages via a browser.
- FTP (File Transfer Protocol) – Allows the transfer of one or more files from one machine to another across the Internet
- TCP/IP Transmission Control Protocol / Internet Protocol) - The protocols, or conventions, that computers use to communicate over the Internet.
- SMPTP Simple Mail Transfer Protocol – A protocol used to send e-mail on the Internet. SMTP is a set of rules regarding the interaction between a program sending e-mail and a program receiving e-mail.

OSI Model-History & Definition

- Older computer networks were closed (proprietary) systems. Definition & publication of standards allows development of hardware /software which operate in open systems.
- Open systems interconnection (OSI) model an attempt to present a complete conceptual model for network hardware and software.
- Developed by ISO in late 1970's, original idea was for complete & universal protocol specification – commercial pressures prevented this.
- Model organizes network functions into seven layers.



Layers of OSI

- As frame is being built, each of the layers adds its own header to the data from the layer above (encapsulation).
- At the receiving end of the communications link, each layer removes its header (decapsulation) and passes the data portion up to the next layer.

- Peer layer communication occurs since corresponding layers process only the data that is contained in the specific layer header.

1. Physical Layer

- Physical layer provides mechanical, functional and electrical interface specification for connection to the media.
- Implemented largely in hardware (eg media & network interface card) & determines what connectors are used, how many pins, what voltage defines a 1 or a 0, etc.
- RS232, example of protocol with physical layer components.

2. Data Link

- Data Link layer compiles 1's and 0's into frames by defining beginning & end of data field & adding error correction bits.
- Provides basic error detection and correction.
- Deals with physical addresses. Information in the header usually includes the destination and source addresses of the frame. Eg. to Computer 8, from computer 12.
- May be connection (HDLC) or connectionless (Ethernet).

3. Network Layer

- Network Layer routes packets one hop at a time from source, across network, to destination. Layer decides how to best forward packet, based on either dynamically determining best route, or by looking up a routing table.
- Devices that operate on network layer forward packets from one segment to the next based on the destination network address of the packet.
- May be connection oriented (X.25) or connectionless (IP).

4. Transport Layer

- Objective of transport layer is to ensures reliable end to end delivery of data across network.
- Performs sequencing and flow control , ensures recipient can reassemble a stream of packets in right order.
- Message multiplexing, fragmentation and reassembly. May be reliable (TCP) or best efforts (UDP).

5. Session layer

- Session establishes, maintains and synchronizes dialog between communicating applications on remote computers.
- Responsible for regulating the flow of information between applications.
- takes care of security and handling errors
- Eg Remote Procedure Calls (RPC).

6. Presentation Layer

- Presentation layer is translator. Eg Translates ACSII into EBCIDIC, and vice versa;
- Performs compression, decompression; encryption and decryption.
- Presentation layer transforms data into the form that the Application layer can accept.

7. Application layer

- The Application layer provides services such as printing, file transfer, remote terminal services, and directory browsing.
- Some user applications exist directly at the Application layer, such as Telnet and FTP.
- Other user applications have Application layer functions built into them.
- Eg word processing program that can print to a network printer has Application layer functions built into it
- (Application Programming Interfaces (API)) .

Note:

LESSON 35
DATA COMMUNICATION HARDWARE

Outline

- Network Interface Unit
- Connectors
- Servers
- Extending the Network
  - Repeaters
  - Hubs
  - Bridges
  - Routers
  - Gateways
- WAN Hardware

Network Interface Unit

A NIU (or network card) is a circuit board that is installed into the expansion slot of computer. It consists of two functional parts:

- a bus interface unit (provides physical connection to workstation's I/O bus) and
- a communications interface unit which provides connection to medium.

Network interface Card

Functions of NIC:

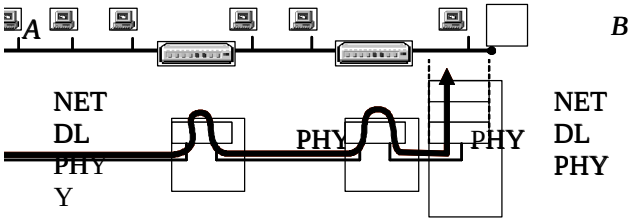
- Implements MAC protocol including monitoring medium – exact role depends on which MAC;
- Serve as buffer between medium & workstations memory;
- Give unique address to a workstation; and
- Receive and transmit data.

Network Interconnection Devices

- Connections to other networks may be LAN to LAN or LAN to WAN or sometimes LAN to Host.
- Network differences determine the connection method to be used. For example:
  - Ethernet to Ethernet - pass message.
  - Ethernet to SNA - translate and pass message.
- Devices require extra intelligence to manage these interconnections. Bridges, routers and gateways all used to interconnect networks and devices.

a) Repeater

- A certain type of cables can only transport signals to a certain distance. Beyond the distance, the signals may be too weak to be decoded accurately. This limits the coverage of a LAN.
- Repeaters are used to extend the coverage of a LAN due to its limitation in transmissions.
- A repeater is operated in physical layer, it mostly deals with signal regeneration. They do not deal with the MAC frame format.



b) Hub

- Used to interconnect workstations
- Hubs can be:

  i) Active - provide signal regeneration.

  ii) Passive - simply provide station connection.

c) Switches

- A switch creates a "virtual LAN" for every packet – an instantaneous connection between the sending and receiving nodes.
- Switches can greatly improve network performance as switching done in hardware.
- Can be used to connect multiple LANs.

1. Bridges

- Bridges connect two LANS using same protocol – can be different media.
- "Dumb" bridges - just pass the message (rather like a repeater).
- Intelligent bridges - understands addressing element of the message (I.e. operates at data link layer of OSI). Examine the address on the message and makes a decision about which segment to pass the message to (filter).
- Bridges maintain tables of node addresses which can be set up by an administrator or learned by the bridge.

Advantages :

- Bridges buffer messages and error check before passing them on;
- Reduce size constraints - can have large number of segments and workstations; and Keep performance statistics.

Disadvantages

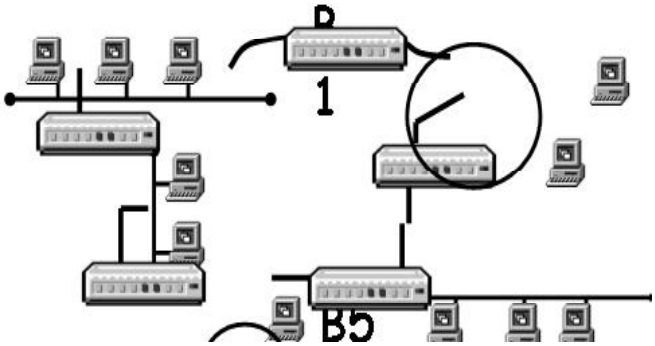
- Delay due to buffering; and
- Buffers may get overloaded.

Bridges operate at the bottom two layers of the OSI model (physical and data link). Therefore they are used to connect

networks that use the same communication protocols above the data-link layer but may or may not use the same protocols at the physical and data –link layers. For eg , bridges may be used to connect two networks, one of which uses fiber-optic cables and other uses coaxial cables or one of which uses Ethernet technology and the other uses token ring technology. But both networks must use same high –level protocols (eg TCP/IP or XNS) to communicate.

Why Bridges?

- To link two LANs in two far locations beyond the limit that a repeater can support.
- To link two different types of LANs.
- To split a LAN into two separate LANs to improve the performance, reliability and security.
- To expand the coverage of a LAN.



2. Routers

- Routers perform similar tasks to bridge, but also operate at network level of OSI model, so make higher level decisions.
- Select best possible path from one point in the network to a selected destination.
- Scan network to monitor traffic patterns, additions/ deletions of nodes, & keep routing tables up-to-date.
- Exchange routing table info with other routers on network.
- They can be used to prevent traffic from moving from/to certain portions of the network and to collect statistics.
- There are a number of routing protocols in existence: e.g. next hop, OR plan entire trip based on distance or load.
- Router may be special purpose computer, or routing functions may be embedded in ordinary computer or other devices e.g. HUB.

Routers operate at the same network layer of the OSI model. Therefore , routers do not care what topologies or access level protocols the interconnected network segments use. Routers uses the bottom three layers of the OSI model, they are usually used to interconnect those networks that use same high-level protocols above the network layer, the protocols of data-link layers are transparent to routers. I f two network segments use different protocols at these two layers, a bridge must used to connect them. Bridges know the ultimate destination of a data, but routers only know which is the next router for the data being transferred across the network. Routers are smarter than bridges in the sense that they not only copy a data from one network segment to another, but they choose the best route for the data by using information in routing table to make this decision.

3. Gateways

- A gateway generally connects LAN to WAN, i.e. networks that do not have a common network protocol. It is essentially a protocol converter.
- Gateway must understand both protocols in use in order to do higher-level conversions.

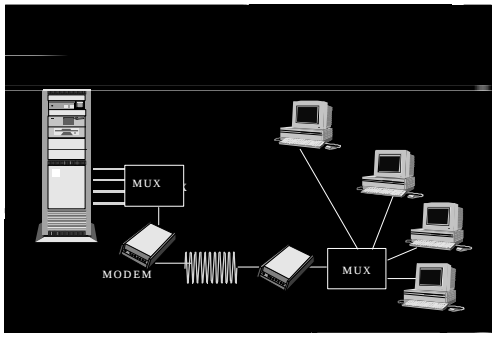
It operates at top three layers of OSI model (session, presentation and application). They are the most sophisticated networking tools and are used for interconnecting dissimilar networks that use the different communication protocols. Thus gateways are used to interconnect networks that are built on totally different communication architectures. For e.g.- it is used to interconnect two networks, one of which uses the IP suite and the other uses the SNA protocol suite.

Internetworking: interconnecting two or more networks to form a single network. Is called an Internet work. Therefore, a WAN of multiple LAN's is an Internet work. Inter networks are often heterogeneous networks composed of several network segments that may differ in topology and protocol. For e.g. , an internet work may have multiple LAN's , some of which may have ring topology ; some of these LAN's may be using Ethernet technology etc.

Interconnecting two networks having the same topology and protocol is simple because two networks can easily communicate with each other., but connecting two dissimilar networks that have different topologies and protocols requires an interconnecting scheme that provide some common point of reference for the two networks to communicate with each other. There are some tools available that provides interconnection of similar and dissimilar network to form a single network system.

WAN Hardware

- Multiplexers - allow sharing of communication line by multiple devices by combining many inputs into one and back again. Use either time or frequency division multiplexing.
- Concentrator - more complex type of multiplexer. Mainly used to connect multiple terminals to host computer.
- Modem



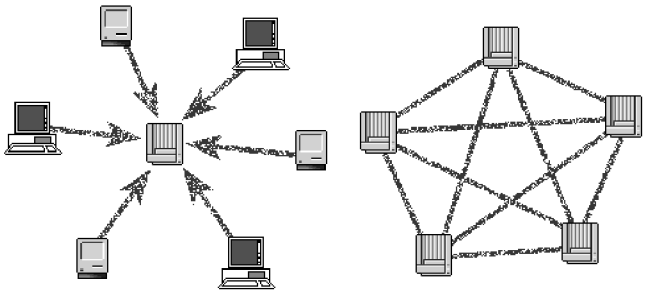
LESSON 36
EXTERNAL LINKS - INTERNET-I

Outline

- History
- Introduction to Internet
- Uses of Internet
- Connection of personal computer to Internet
- Equipments and programs to connect the Internet
- Computers on the Internet work together to transfer the information

History

The Internet started as a governmental / military endeavor - it was an attempt to create a system of communication that could withstand a nuclear attack. The idea was simple: instead of creating a centralized communication system, which could be easily taken out in a war, they wanted to create a system composed of individual, equal nodes. In the centralized system, all lines went to a central data center where they were routed out to specific places. In the decentralized system, data went from node to node until it made it to its destination. If a few nodes were lost - no problem - you'd just use others.



Centralized System

Decentralized System

Although the Internet has grown significantly since its inception, and is used by more than just the military and government researchers, it retains this fundamental structure. There is no centralized hub and anyone can become a node.

Give an example of decentralized system.

Give an example of Centralized system.

HOw Can Completely Different Computer Systems All Link Up to Form The Internet?

Although each computer may be different, they all use a common protocol (a language) called TCP/IP. The development of this protocol allowed the Internet to grow into what it is today. Instead of being a CLOSED network (like a local area network) the TCP/IP protocol allows any person or group to form a gateway and connect to the network.

Brief Time-Line:

1969	ARPANET formed to allow electronic communication between defense, universities and defense contractors.
1973	Research begun on a TCP/IP protocol that would allow different computers to access the ARPANET.
1980	First separate network connected to the ARPANET using the new TCP/IP Gateway protocol.
1983	Official beginning of the Internet: ARPANET split into MILNET (for the military) and ARPANET (for everybody else).
1986	NSFNET created a supercomputer 'backbone' for the Internet using new high- speed links.
1994	With so many new networks connected to the Internet, NSFNET ceases to be the backbone of the net, just a major player.

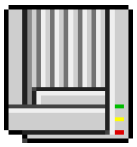
Speed at Which the Number of Internet Hosts (Nodes) Increased by a Factor of TEN:

DATE	# of Hosts:
1977	
100+	
1984	
1000+	
1987	
10,000+	
1989	
100,000+	
1992	
1,000,000+	
1996	
10,000,000+	

Just imagine the number in 2008

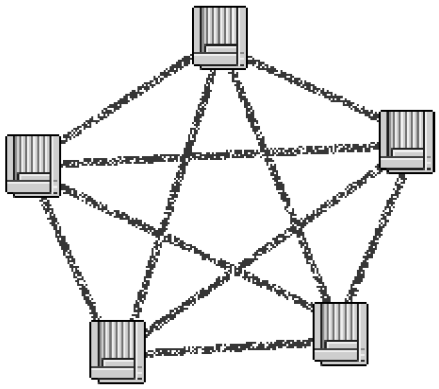
What Isthe Internet?

1) The Server



The server is really the building block of the Internet. It is what constitutes a 'Node' on the net. A server is generally a Unix workstation, multi-tasking machines that can do more than one thing (maybe millions) at a time. Since a single user can't possibly use a multi-tasking computer to its full potential, these kinds of computers have been designed to be shared by many people. When multiple people use the same computer, it is necessary to work out a system for sharing the resources (storage, CPU, software, etc.) of the computer. This is, therefore, what a server is: a multi-tasking computer which can accommodate many people at the same time and share its resources with them.

2) The Domain



A Domain is a collection of inter-connected servers. All of the CED servers are grouped in a single domain: "CED". You can also group domains: all of the domains at Berkeley are grouped into another domain: "Berkeley."

Servers can be connected in a variety of ways. Personal Computers are often linked to the Internet using a Dial-Up connection - the computer's modem uses the phone line to connect with a server. Most often, the phone line is also used for incoming calls when the modem is not on. Most servers are inter-connected using Leased Lines, that is, lines that are dedicated to the computers. These lines are usually fiber-optic or some other material that moves data faster than a phone line. It is also possible to use 2 other devices for making links: microwaves and satellites. These devices allow signals to travel far further. (Since the Internet uses satellites it has been said to truly encircle the globe.)

Names

Server names include both their individual names, and the names of all of the domains that they are in. Like your postal address, which goes from particular to general, server names begin with the particular name of the machine ('chicheemaun') and then move from the most particular domain to the most

general. Chicheemaun is in the domain 'ced', which is in turn in the domain 'berkeley.' Each name is separated with a dot: "." The highest level domain indicates how the server is being used. If it is used by a commercial enterprise, it will have '.com' at the end. Since chicheemaun is used by a university, it ends with '.edu' Chicheemaun's full name, therefore, is: chicheemaun.ced.berkeley.edu

Other first level domains (in the US) are:

.com = Commercial

.edu = Educational

.gov = Governmental

.mil = Military

.net = Major Network Support Centers

.org = Other Organizations

Name Some Educational Sites

Name some network support centers sites

First level domains outside of the US include the country name (example, Canada = .ca; Japan = .jp, etc.)

Naming Systems:

Domain Name:	The naming system that we just described is called the Domain Name. It uses words, separated by dots, to describe the Internet address of a particular node. For example, chicheemaun.ced.berkeley.edu
IP (Internet Protocol):	Computers do not use names, they use numbers. The names are for people who can't memorize a string of numbers. When you ask the computer to contact another one and you give it a Domain Name, the computer duly translates that name into an Internet Protocol address. For example, 128.95.10.207 There is no difference between a Domain Name and an IP except the way in which the information is represented.

Domain Name

The naming system that we just described is called the Domain Name. It uses words, separated by dots, to describe the Internet address of a particular node. For example, chicheemaun.ced.berkeley.edu

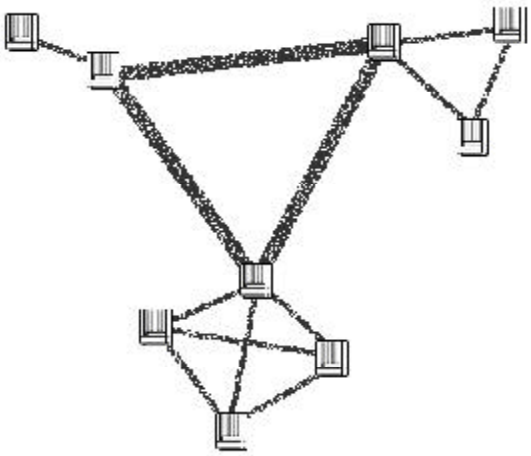
IP (Internet Protocol)

Computers do not use names, they use numbers. The names are for people who can't memorize a string of numbers. When you ask the computer to contact another one and you give it a Domain Name, the computer duly translates that name into an Internet Protocol address. For example, 128.95.10.207 There is no difference between a Domain Name and an IP except the way in which the information is represented.

The Internet is the largest computer network in the world. It consists of thousands of connected networks world-wide. A network is a collection of computers that are connected to share information. Each organization on the Internet is responsible for maintaining its own network.

And some of the information on the Internet is free. Governments, universities, colleges, companies and individuals provide free information to educate and entertain the public.

The Net



So what is the Internet?

It is a group of domains. Each domain is itself a network, sometimes including sub-networks and sub-sub-networks. The Internet is just the sum total of all of the networks. How big is it - NOBODY KNOWS!! There are so many different networks all connected up, now, that nobody really has any idea of its size, let alone how to measure it. The closest count is in terms of the number of hosts which is now in the tens of millions and in almost every country in the world. The only thing in common amongst all of the different networks is that they use the TCP/IP protocol as a common language so that they can all communicate. However, there are other protocols out there, and other vast networks that run parallel to the Internet! The only thing that seems for certain is that the Internet continues to grow and change at a phenomenal rate

What The Internet Offers ?

A. The Ability to Share Information

The first function of a network is to share information. When I say information I mean bits and bytes, data stored on hard drives, floppy disks, CD ROMs, etc. When you log into a network you will be able to use that network to get data from a different (remote) computer or to put data from your computer

into that remote computer. Never mind where the remote computer is - it may be in the room next to you, or on the other side of the world.

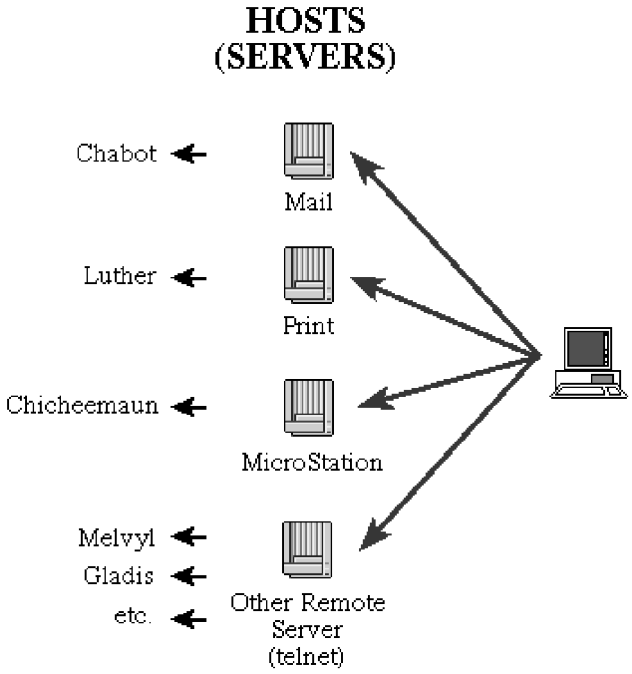
Write some sites on which you can get the information about astrology.

b. The Ability to SHARE FUNCTIONS (Hardware & Software)

This is a little harder to grasp. The second function of networks is to allow users to log in to remote computers and to use that computer's processing power - its CPU - or any hardware or software attached to it!! If you have a program that I want to use on your computer, and our computers are networked (it also depends on the type of computer or network), I could log in to your computer from my own, send you a File that I wanted tinkered with by the program sitting in your computer, get your computer to do that for me, and then have it send it back to me! I could also use your computer to print a document or to search a database sitting on its hard drive.

Another network function is to allow users to log into different computers and to use the remote computer's hardware or software. This is referred to as a Client - Server relationship. The client asks the server to do something for it. On the Internet your computer becomes a client when you log into other machines to run programs.

Sharing Function



You can also move consciously into another machine. There are two commands that let you do this, rlogin and telnet. When this happens nothing will change on your monitor except that the messages that appear at the prompt will be generated by a different machine.

LESSON 37
EXTERNAL LINKS -INTERNET – II

Outline

- Day to day advantages of internet
  - Electronic mail
  - Information
  - Entertainment
  - Programs
  - Online shopping
  - Chat
- How the personal computers attached to the network ?
- What Are Different Equipments And Programs To Connect The Internet ?
  - Computers
  - Programs
  - Modem
  - Internet service providers
- How All The Computers On The Internet Work Together To Transfer Information?
  - Packets
  - TCP/IP
  - Router

Day To Day Advantages Of Internet

Electronic Mail

Exchanging electronic mail (e-mail) is the most popular feature on the Internet. You can exchange electronic mail with people around the world, including friends, colleagues, family members, customers and even people you meet on the Internet. Electronic mail is fast, easy, it saves money and paper.

Now a days it is not necessary to have home address but you should have your email id.

Write ten site names on which you can have your email-ID.

How much free space is provided to you if you create an email-id on Yahoo.com ?

Information

The Internet gives you access to information on any subject imaginable. You can review newspapers, magazines, academic papers, government documents, television show transcript, famous speeches, recipes, job listings, works by Shakespeare, airline schedules and much more.

Entertainment

The Internet also lets you review current movies, hear television theme songs, read movie scripts and have interactive conversations with people around the world even with celebrities.

Write 10 site names on which you can play an online games and you can win some excited prizes (may be a smart bike)

Programs

Thousands of programs are available on the Internet. These programs include word processors, spreadsheets, games and much more.

Online Shopping

You can buy items such as books, computer programs, flowers, music CDs, pizza, stocks, second-hand cars and much more.

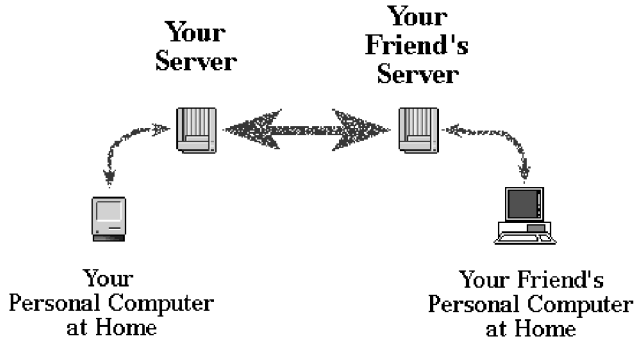
Write five site names on which you can do online shopping in India

Chat

The chat features allows you to exchange typed messages with another person on the internet. A message you send will instantly appear on the other person's computer. You can chat with one person at a time or with a group of people.

Write any Five Chat Messengers

How The Personal Computer Connects to The Internet ?



How do individuals and their computers get on the Internet? By using an account on a server. The account has its own name on the server so that information can be sent to it. The computer at home connects to its account on the server using a modem and a dial-up phone line. When you dial the number and your computer connects to the server you must enter a password and account name. You then send and receive information by first sending it to the server, then letting the server send it to other computers on the net.

There are many commercial Internet services which provide this service. As a student with a computer at home you can choose to either create a Home IP account which allows you to connect, via modem, to your CED account, or to go to a commercial provider who will give you the same thing (for a price). Either way, once you have logged in, you will have full access to the net.

What Are Different Equipments And Programs to Connect The Internet ?

Equipments And Programs To Connect The Internet

Computer

You can use any type of computer, such as an IBM-compatible or Macintosh computer, to connect the Internet.

Programs

You need special programs to use the Internet. Most companies that connect you to the Internet provide the programs you need free of charge.

Modem
You need a modem to connect to the Internet. Choose a modem with a speed of at least 14,000 bps, although a modem with a speed of 28,800 bps is recommended..

What is a Modem?

Internet Service Provider
An Internet Service Provider (ISP) is a company that gives you access to the Internet for a free.

Cost
Many providers offer you a certain number of hours per day or month for a set fee. If you exceed the total number of hours, you are usually charged for every extra hour you use the provider. And some providers offer unlimited access to the Internet for a set fee. Make sure you are aware of any hidden charges or restrictions.

Write any five internet service providers in India .

Commercial Online Service
A commercial online service is a company that offers a vast amount of information and access to the Internet for a fee.

Cost
Most commercial online services let you try their service free of charge for a limited time. After the trial period, most online services offer a certain number of hours per day or month for a set fee. And if you exceed the total number of hours, you are usually charged for every extra hour you use the online service.

How All The Computers On The Internet Work Together to Transfer Information
All computers on the Internet work together to transfer information around the world.

Packets
When you send information over the Internet, the information is broken down into smaller pieces, called packets. Each packet travels independently over the Internet and may take a different path to arrive at the intended destination, the packets are reassembled.

Download Information
When you receive information from another computer on the Internet, you are downloading the information. When you send information to another computer on the Internet, you are uploading the information.

TCP/IP
Transmission Control Protocol/Internet Protocol (TCP/IP) is a language of a computer on the Internet use to communicate with each other. TCP/IP divides information you send into packets and sends the packets over the Internet. When information arrives at the intended destination, TCP/IP ensures that all the packets arrived safely.

Backbone
The backbone of the Internet consists of high-speed data lines that connect major networks all over the world.

Router
A router is a specialized computer device that regulates traffic on the Internet and picks the most efficient route for each packet. A packet may pass through many routers before reaching its intended destination.

Distingush between Internet and Intranet

LESSON 38
SECURITY & ELECTRONIC INFECTION

Outline

- Security
  - Physical security
  - Logical security
  - User Authentication
  - Access Control
  - Cryptography
- Electronic Infections
  - Viruses
  - Worms
  - Email viruses
  - Trojan horses
  - Prevention

Physical Security

Physical security includes but is not limited to:

- Controlling access to computer hardware.
- Preventing service interruptions (power, hardware failure).
- Planning for disaster and other contingencies.

Logical security includes but is not limited to:

- Controlling access to and use of software and data.
- Recovering data, transmissions and software.
- Archiving data, software and documentation.

Security in computer systems deals with the protecting the various resources and information of a computer system against destruction and unauthorized access. A total approaches to computer security involves both external and internal security. External security deals with securing the computer system against external factors such as fires, floods, earthquakes, stolen disks/tapes, leading out of stored information by a person who has access to the information, and so on. For external security, the commonly used methods include maintaining adequate backup copies of stored information at place far away from the original information, using security guards to allow the entry of only authorized persons into the computer centre, allowing the access to sensitive information to only trusted employee/users, and so on. Internal security, on the other hand, mainly deals with the following aspects.

1. User Authentication

Once a user is allowed physical access to the computer facility, the user's identification must be checked by the system before the user can actually use the facility. This requirement is taken care of by user authentication mechanisms.

2. Access Control

A computer system contains many resources and several types of information. Obviously, not all resources and information are meant for all user. Therefore, even when a user passes the authentication phase and is allowed to use the computer facility, a way is needed to prohibit the user from accessing those resources/information that he/she is not authorized to access. This requirement is taken care of by access control mechanisms.

3. Cryptography

A third type of internal security is often used to protect unauthorized access of information that is highly critical in nature. That is, even if a user somehow manages to gain access to some information that he/she is not authorized to access, a way is needed to ensure that the user cannot make use of that information. This requirement is taken care of by cryptography mechanisms.

Below we will discuss about the commonly used mechanisms for providing these three different types of internal security in computer systems.

User Authentication

User authentication deals with the problem of verifying the identify of a user (person or program) before permitting access to the requested resource. That is, an authentication mechanism prohibits the use of the system (or some resource of the system) by unauthorized users by verifying the identity of a user making a request.

Authentication basically involves identification and verification. Identification is the process of claiming a certain identify by a user, while verification is the process of verifying the user's claimed identity. Thus, the correctness of an authentication process relies heavily on the verification procedure employed.

The three basic approaches to user authentication are as follows:

1. Proof by Knowledge. In this approach, authentication involves verifying something that can only be known by an authorized user. Authentication of a user based on the password supplied by him/her is an example of proof by knowledge. Authentication methods based on the concept of proof by knowledge are again of two types – direct demonstration method and challenge – response method. In he direct and challenge – response methods. In the direct demonstration method, a user claims his/her identity by supplying information (like typing in a password) that the verifier checks against pre-stored information. On the other hand in the challenge-response method, a user proves his or her identity by responding correctly to the challenge questions asked by the verifier. For instance, at the time of initially registering in a system as a user, the user picks a function, for example, x + 18. When the user logs in, the system randomly selects and displays a number, say 105, in which case the user must type 123 for authentication to be successful.

2. Proof by Possession. In this approach, a user proves his/her identity by producing some item that can only be possessed by an authorized user. The system is designed to verify the produced item to confirm the claimed identity. For example, a plastic card with a magnetic strip on it that has a user identifier number written on it in invisible, electronic form may be used as the item to be produced by the user. The user inserts the card in a slot meant for this purpose in the system's terminal, which then extract the user identifier number from the card and checks to see if the card produced belong to an authorized user.
3. Proof by Property. In this approach, the system is designed to verify the identity of user by measuring some physical characteristics of the use that are hard to forge. The measured property must be distinguishing, that is, unique among all possible users. For example, a special device (known as a biometric device) may be attached to each terminal of the system that verifies some physical characteristic of the user, such as the person's appearance, fingerprints, hand geometry, voice, and signature. In deciding the physical characteristic to be measured, an important factor to be considered is that the scheme must be psychologically acceptable to the user community. Biometric systems offer the greatest degree of confidence that a user actually is who he/she claims to be, but they are also generally the most expensive to implement. Moreover, they often have user acceptance problems because user see biometric devices an unduly intrusive.

In practice, a system may use a combination of two or more of these authentication methods. For example, the authentication mechanism used by automated cash-dispensing machines in banks usually employs a combination of the first two approaches. That is, a user is allowed to withdraw money only if he or she produces a valid identification card and specifies the correct password corresponding to the identification number on the card.

To provide good security with password-based authentication, it is important that passwords are kept secret and passwords are chosen in such a manner that they are difficult to guess.

Access Control

Once a user or a process has been authenticated, the next step in security is to devise ways to prohibit the use or the process from accessing those resources/information that he/she/it is not authorized to access. This issue is called authorization and is dealt with by using access control mechanisms.

When talking about access control in computer system, it is customary to use the following terms:

1. Objects. An object is an entity to which access must be controlled. An object may be an abstract entity, such as a process, a file, a database, a tree data structure, or a physical entity, such as a CPU, a memory segment, a printer, and a tape drive.

   Each object has a unique name that differentiates it from all other objects in the system. An object is referenced by its unique name. In addition, associated with each object is a "type" that determines the set of operations that may be performed on it. For example, the set of operations possible on objects belonging to the type "data file" may be Open, Close, Create, Delete, Read and Write, whereas for objects belonging to the type "program file," the set of possible operations may be Read, Write, and Execute.
2. Subjects. A subject is an active entity whose access to objects must be controlled.

   That is, entities wishing to access and perform operations on objects and to which access authorizations are granted are called subjects. Example of subjects is processes and users.
3. Protection Rules. Protection rules define the possible ways in which subjects and objects are allowed to interact. That is, protection rules govern the subjects' access to objects. Therefore, associated with each (subject, object) pair is an access right that defines the subset of the set of possible operations for the object type that he subject may perform on the object. The complete set of access rights of a system defines which subjects can perform what operations on which objects. At any particular instance of time, this set defines the protection state of the system at that time.

   With the above-mentioned concepts, access to objects by subjects by subjects is controlled in the following manner. Suppose subject S requests for an access on object O, where r belongs to the set of operations that may be performed on O. To validate this access request, the access controls modules for the security system check whether access r is permitted for subject S on object O in the current protection state of the system. If yes, the access is permitted; otherwise the access is denied.

Electronic Infection

Types of Infection

- Viruses - A virus is a small piece of software that piggybacks on real programs. For example, a virus might attach itself to a program such as a spreadsheet program. Each time the spreadsheet program runs, the virus runs, too, and it has the chance to reproduce (by attaching to other programs) or wreak havoc.
- E-mail Viruses - An e-mail virus moves around in e-mail messages, and usually replicates itself by automatically mailing itself to dozens of people in the victim's e-mail address book.
- Worms - A worm is a small piece of software that uses computer networks and security holes to replicate itself. A copy of the worm scans the network for another machine that has a specific security hole. It copies itself to the new machine using the security hole, and then starts replicating from there, as well.
- Trojan Horses - A Trojan horse is simply a computer program. The program claims to do one thing (it may claim to be a game) but instead does damage when you run it (it may erase your hard disk). Trojan horses have no way to replicate automatically.

  Name any two famous E-mail viruses

What's a "Virus"?

Computer viruses are called viruses because they share some of the traits of biological viruses. A computer virus passes from computer to computer like a biological virus passes from person to person.

There are similarities at a deeper level, as well. A biological virus is not a living thing. A virus is a fragment of DNA inside a protective jacket. Unlike a cell, a virus has no way to do anything or to reproduce by itself -- it is not alive. Instead, a biological virus must inject its DNA into a cell. The viral DNA then uses the cell's existing machinery to reproduce itself. In some cases, the cell fills with new viral particles until it bursts, releasing the virus. In other cases, the new virus particles bud off the cell one at a time, and the cell remains alive.

A computer virus shares some of these traits. A computer virus must piggyback on top of some other program or document in order to get executed. Once it is running, it is then able to infect other programs or documents. Obviously, the analogy between computer and biological viruses stretches things a bit, but there are enough similarities that the name sticks.

An Ounce of Prevention

You can protect yourself against viruses with a few simple steps:

- If you are truly worried about traditional (as opposed to e-mail) viruses, you should be running a more secure operating system like UNIX. You never hear about viruses on these operating systems because the security features keep viruses (and unwanted human visitors) away from your hard disk.
- If you are using an unsecured operating system, then buying virus protection software is a nice safeguard.
- If you simply avoid programs from unknown sources (like the Internet), and instead stick with commercial software purchased on CDs, you eliminate almost all of the risk from traditional viruses. In addition, you should disable floppy disk booting -- most computers now allow you to do this, and that will eliminate the risk of a boot sector virus coming in from a floppy disk accidentally left in the drive.
- You should make sure that Macro Virus Protection is enabled in all Microsoft applications, and you should NEVER run macros in a document unless you know what they do. There is seldom a good reason to add macros to a document, so avoiding all macros is a great policy.
- Load some anti viruses in your system like Norton

Name any two anti viruses other than Norton.

What are Macro viruses?

Note:

LESSON 39
HACKING AND ENCRYPTION

Outline

- Hacking
  - Hacker ethics
  - How Hacker Hacks
  - True story of Hacker
  - Ethics of hacking
- Cryptography
  - Encryption
  - Decryption
  - Symmetric and asymmetric cryptosystems

Who is a Hacker?

- Originally - A person who pulls clever practical jokes.
- Generally - A good programmer.
- According to the Press - A malicious meddler who tries to discover sensitive information by snooping around where they should not be.

What is Hacking?

Unauthorized use, or attempts to circumvent or bypass the security mechanisms of an information system or network. Hacking means illegally accessing other people's computer systems for destroying, disrupting or carrying out illegal activities on the network or computer systems.

'Hackers' hack for their own amusement, challenge or sometimes financial, ideological or commercial gain. Hacking is a far less common than is generally imagined. As a private user, you are very unlikely to be affected by hacking.

Hacking is, very simply, asking a lot of questions and refusing to stop asking. This is why computers are perfect for inquisitive people -- they don't tell you to shut up when you keep asking questions or inputting commands over and over and over. But hacking doesn't have to confine itself to computers. Anyone with an inquisitive mind, a sense of adventure and strong beliefs in free speech and the right to know most definitely has a bit of the hacker spirit in them.

Samples:

- Hard-core porn stored on Lawrence Livermore National Laboratory's computers
- Windows 95 Is A Hacker's Dream Over The Internet
- Gang War in Cyberspace
- A mail bomb
- US White House Hacked?
- Croatian teen hackers break Pentagon codes.
- FTC shuts down 3 sites for phone fraud.

Hacker Ethics

- Hackers do have a sense of ethics.
- Hackers believe that their duty is to share their expertise by writing free software and sharing information.
- No damage, No Crime.
- Help Sysops and Admins by breaking in.
- 90% of hackers willingly share hacking information.

Are There Legal or Appropriate Forms of Hacking?

One of the common misconceptions is that anyone considered a hacker is doing something illegal. It's a sad commentary on the state of our society when someone who is basically seeking knowledge and the truth is assumed to be up to something nefarious. Nothing could be further from the truth.

How Hackers Hack

- Guess user passwords.
- Password cracker programs.
- Social Engineering.
- Breaking and Entering.
- Use of Bulletin Boards to find more information on how to hack systems.

True Story of a Hacker

Mr. RAD.

- 13 year old boy in 1986.
- Used a Commodore 64.
- Ran a hacker BBS
- Wrote and ran programs
- Cost innocent phone customers more than $300,000.
- Caught red handed by the F.B.I.

Ethical Hacking

- Is there such a thing?
- IBM suggests that there is!

Conclusion

- Hacking is an easy, so make sure you are secure.
- If you hack sooner or later you will be caught.
- Computer professionals must be responsible with the information they acquire and hold.
- Security professionals can learn from hackers.

Benifts of Hacking.

Comment

Cryptography

Cryptography is a means of protecting private information against unauthorized access in that situation where it is difficult to ensure access control. The basic idea behind this security

technique is that if it is not possible to ensure access control, it is better to prevent comprehension of information.

Basic Concepts and Terminologies

Two primitive operations employed by cryptography are encryption and decryption. Encryption (also called enciphering) is the process of transforming an intelligible information (called plaintext or cleartext) into an unintelligible form (called ciphertext). Decryption (also called deciphering) is the process of transforming the information back from ciphertext to plaintext.

Encryption is basically a mathematical function (encryption logarithm) having the following form:

C = E (P, Ke)

Where P is the plaintext to be encrypted, Ke is an encryption key, and C is the resulting ciphertext. Decryption of C is performed by a matching function (decryption algorithm) that has the following form.

P = D (C, Kd)

Where Kd is the decryption key. Note that the decryption function D is the inverse of the encryption function E. Therefore we have

D (E, (P, Ke), Kd) = P

To prevent the plaintext from being easily revealed, it must be possible to transform a given plaintext into a large variety of possible ciphertext selected by a specific parameter. The keys Ke and Kd serve as this parameter. That is, the function parts remain the same but he keys are changed as often as necessary.

The above-described general structure of a cryptosystem is illustrated in figure 1

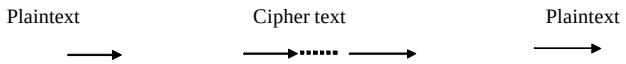


Figure 1 General structure of a cryptosystem.

Symmetric and Asymmetric Cryptosystems

There are two broad classes of cryptosystem –

- Symmetric
- Asymmetric.

Symmetric

In a symmetric cryptosystem, either both the encryption key (Ke) or decryption key (Kd) are the same or one is easily derivable form the other. Usually, a common key (K) is used for both enciphering and deciphering. For security, it is important that the key of a symmetric cryptosystem is also known as shared-key or private-key cryptosystems.

Asymmetric

In an asymmetric cryptosystem, on the other hand, the decryption key (Kd) is not equal to the encryption key (Ke). Furthermore, it is computationally impractical to derive Kd from Ke. Because of this property, only Kd needs to be kept secret and Key is made publicly know. Asymmetric cryptosystems are also known as public – key cryptosystems.

The Data Encryption Standard (DES) cryptosystem is the best known and the most widely used symmetric cryptosystem today. On the other hand, the Rivest-Shamir-Adleman (RSA) cryptosystem is the first published and practically the most satisfactory asymmetric cryptosystem today.

Note:

LESSON 40
BACK-UP AND RECOVERY

Outline

- Backups & Disaster Recovery
- The risks to your data
- Comparison of Backup Method Data Risk Coverage
- Data problem prevention
- Preventing Data Loss and System Problems At Shutdown
- Back Up and Recover Your Information

Backups & Disaster Recovery

One of the things that harp on the most about when it comes to preventive maintenance and system care is the importance of regular, reliable data backup. No matter how well you treat your system, no matter how much care you take, you cannot guarantee that your data will be safe if it exists in only one place. The risks are much greater than most people realize.

How important is your data to you? You can respond to this question with words, but the steps you take to protect your data are the real answer. I find it troubling when people lose large quantities of data--because they have no backups--and then they get very upset, talking about "how important that data was!" If it's important, why wasn't it backed up? If it matters enough to get upset over losing, it's worth protecting, and backups are an essential part of data protection.

1) Get your data safe

2) Get your data safe while not impacting production

3) Get it done as fast as possible

Data is generally regarded as a company's first or second most valuable asset -- right on par with people. Equally important with backing up data is the ability to recover that same data and recovery it in a timely fashion to keep a business up and running.

What is Data Backup?

Data back up is the process of creating a copy of the data from an on-line storage device to a secondary storage device such as magnetic tape, floppy disk, disk pack or WORM disc so that if by accident the data on the on-line storage device gets corrupted or gets lost, it can retrieved from the back up device and stored back on the on-line storage device.

There are many reasons that people neglect doing backups:

- They don't understand how important they are, because they haven't had a disaster happen to them (yet).
- They don't know how to do them.
- They forget to do them because they don't have a routine for doing backups.
- Doing the backup is a time-consuming chore and they can't be bothered.

The Risks to Your Data

Data can be damaged or lost in any one or more of the following ways

1. A disc crash
2. A virus attack
3. a hardware malfunction
4. an unintended accidental deletion of useful files
5. a natural disaster damaging the computer system.

If you do any sort of reasonable amount of computing, it is only a matter of time before you some day need access to backups of your data or programs. There are many different risks to your data; most people just think of the infamous, dreaded disk crash. This is a real risk due to the technology used in data storage, but there are many other ways that you can easily lose data on your PC. In fact, the list of risks below is far from exhaustive, though it covers the most common problems.

I have used the word disc crash usually it is called head crash.

a) Head Crashes

Since the read/write heads of a hard disk are floating on a microscopic layer of air above the disk platters themselves, it is possible that the heads can make contact with the media on the hard disk under certain circumstances. Normally, the heads only contact the surface when the drive is either starting up or stopping. Considering that a modern hard disk is turning over 100 times a second, this is not a good thing.

If the heads contact the surface of the disk while it is at operational speed, the result can be loss of data, damage to the heads, damage to the surface of the disk, or all three. This is usually called a head crash, two of the most frightening words to any computer user. The most common causes of head crashes are contamination getting stuck in the thin gap between the head and the disk, and shock applied to the hard disk while it is in operation.

b) Virus Attack

we will study this in next chapter because I think this is the very interesting topic for all of you and I can complete it in just few lines

c) Hardware Malfunction

The hardware in your system will work better and more reliably, and last longer, if it is maintained regularly according to the directions I outline in this Guide. But ultimately you can "get away with" poor preventive maintenance practices in most cases; at worst your hardware will fail and you will have to replace it. Expensive, but not the end of the world, usually. However, there is one thing that if lost, cannot be readily replaced: your data.

d) An Unintended Accidental Deletion Of Useful Files

e) A Natural Disaster Damaging The Computer System.

Name some natural disasters that can damage your data

Comparison Of Backup Method Data Risk Coverage

Some of the backup methods and devices described in this section do a much better job than others of protecting against the risks to your data. The table below shows a general summary of how the various methods stack up in terms of protecting you from the hazards that threaten your valuable data. Remember that this is just a general guideline; also remember that some of these risks are much more common than others are, depending on how you use your PC:

Data Risk	Floppy Disks	Tape Drives	Removable Storage Drives	Removable Hard Disks	In-Place Hard Disk Duplication	Network Backup	File Archiving
Hardware Failure	High	High	Moderate to High	Moderate to High	Moderate	High	Low
Software Failure	High	High	High	High	Moderate	High	Low
File System Corruption	High	High	High	High	Low to Moderate	Moderate to High	Moderate
Accidental Deletion	High	High	High	High	High	High	High
Virus Infection	Moderate	High	Moderate to High	Moderate	Low	Moderate to High	Low to Moderate
Theft	High	High	High	Moderate to High	None	Low to Moderate	None
Sabotage	High	High	High	High	Very Low	Low	None
Natural Disaster	High	High	High	High	None	Low to Moderate	None

Data can never be properly replaced; it can only be protected against loss. This is why it is so critical that you take positive, proactive steps to reduce your chances of catastrophic data loss.

Data Prevention

The best way to avoid data loss is of course to prevent it by using your computer wisely. There are many specific activities you can undertake, that will help protect your data and prevent it from being lost. These are examined in this section.

Data has a great disadvantage compared to PC hardware in that it can be lost, and once lost, never easily replaced. You can get your hard disk replaced under warranty, but the new one will come to you totally empty. However, data has a great advantage that compensates for this weakness: it can be readily and easily duplicated (unlike your hard disk drive :))

By creating backups of important data on a timely basis, you ensure that your data, or most of it anyway, will never be truly lost; at worst, some will be lost and you will experience the inconvenience of restoring it in the event of a hard disk failure, for example. It is absolutely critical that you create backup copies of all important data, documents and programs on your hard disk, so that you have some protection in the event that catastrophe strikes.

Backups are important enough, and have enough factors related to how they work.

Preventing Data Loss and System Problems at Shutdown

Following proper shutdown procedures can eliminate a number of different problems that occasionally plague computer users. In general, simply shutting off the PC when you decide that you are done with it is a very dangerous practice with today's multitasking operating systems. There are two main reasons for this, and both are related to the fact that the system typically has many things going on at the same time:

- Unexpected Activity: Modern operating systems, such as Windows 95 or Windows NT, have many different tasks running on them at the same time. Even when the system appears to be idle, there are various system jobs in operation. These can lie dormant for minutes or hours and then start up unexpectedly and jump into operation very quickly. Examples would include memory-resident virus-scanning programs, screen savers, automatic backup utilities, program schedulers, etc. One of these programs can decide to start up a fraction of a second before you hit the power switch on your seemingly "idle" PC, with the result being that you are cutting the power to the hard disk in the middle of a write operation. This can damage the file system and threaten the safety of the files on your disk.
- Open Files and Caches: Modern operating systems also have files that remain open whenever they are in operation. Further, they use disk caching, which can sometimes mean that writes to the hard disk are held pending in memory, for performance reasons. Turning off the PC before the system has a chance to close its open files and flush the disk cache can lead to data loss and file system corruption.
- In general, you should avoid unexpected shutdowns of the PC (unexpected either by yourself, or unexpected by your operating system). To avoid the problems described above, you should always do the following before you try to shut down your PC:
- Save Data in All Open Applications: Make sure you save any files you are working on in all open programs. Most programs will either automatically save for you when you exit them, or at least ask you if you want them to be saved, but watch out for the few brain-dead applications that will exit without saving unless you are careful.
- Close All Open Applications: Close all programs that you have opened. In Windows, using the {Alt} most easily does this+{F4} key combination in each program and responding to prompts. To close an MS-DOS prompt, type "EXIT" at the prompt and hit {Enter}.
- Shut Down the Operating System: Give the proper command to the operating system to shut it down. In Windows this is done using the {Alt}+{F4} combination,

or choosing "Shut Down..." from the Start menu. Wait until you are instructed that it is safe before hitting the power switch.

BACK-Up and Recover your Information – Windows Xp

The Backup utility in Windows XP Professional helps you protect your data in the event your hard disk fails or files are accidentally erased due to hardware or storage media failure. By using Backup you can create a duplicate copy of all of the data on your hard disk and then archive it on another storage device, such as a hard disk or a tape.

If the original data on your hard disk is accidentally erased or overwritten, or becomes inaccessible because of a hard–disk malfunction, you can easily restore it from the disk or archived copy by using the Restore or Automated System Recovery Wizards.

To start Backup or to access Restore and Automated System Recovery

- Click Start, click All Programs, click Accessories, click system Tools, and then click Backup.

Windows XP Backup, Restore, and Automated System Recovery all function when Windows XP Professional is functioning. If your computer does not start properly, you may need to use Recovery Console. Recovery Console provides a command line during Startup from which you can make system changes when Windows XP Professional doesn't start.

To learn more about Backup, Restore, and Automated System Recovery, see Help and Support Center.

Recovery Console

You can use Recovery Console to perform many tasks without starting Windows XP, including: starting and stopping services, reading and writing information on a local disk drive, and formatting drives. However, you must install Recovery Console while your computer is still functioning. The Recovery Console feature should be used only by advanced users. Before using Recovery Console, it is recommended that you back up your information on a tape drive, because your local hard disks might be reformatted-thus erased-as part of the recovery. You can also run Recovery Console from the Windows XP CD.

Note:

LESSON 41
GROUP DISCUSSION
CONTRIBUTE TO THE PRODUCTION OF A SECURITY POLICY

Note:

"The lesson content has been compiled from various sources in public domain including but not limited to the internet for the convenience of the users. The university has no proprietary right on the same."

EIILM UNIVERSITY
SIKKIM

Jorethang, District Namchi, Sikkim- 737121, India
www.eiilmuniversity.ac.in