

CIT 611 Computer Fundamentals

Course Developer: *Professor R. O. Ayeni*

*Ladoke Akintola University of Technology,
Ogbomoso.*

Programme Leaders: *Dr Sam. I. Ogunrinde*

Dr S. A. Reju

Course Co-ordinators: *A. Balogun*

C. Okonkwo (Mrs.)



National Open University of Nigeria

Headquarters

National Open University of Nigeria
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Annex Office

245 Samuel Adesujo Ademulegun Street
Central Business District
Opposite Arewa Suites
Abuja

E-mail: centrainfo@nou.edu.ng

URL: www.nou.edu.ng

National Open University of Nigeria, 2004

First published 2004

ISBN 978-058-216-9

All Rights Reserved. I

Table of Contents

CourseGuide	-----	1
Module 1 Computer Fundamentals: Hardware	-----	3
Unit1	What is a Computer? -----	4
1.0	Introduction -----	4
1.1	Objectives -----	4
1.2	What is Computer? -----	4
	1.2.1 The Computer and Integrated Circuit Technology -----	7
	1.2.2 Classification of Computers -----	10
1.3	Conclusion -----	11
1.4	Summary -----	11
1.5	Tutor-Marked Assignment (TMA) -----	11
1.6	Further Reading. -----	11
Unit 2	Memory System	12
2.0	Introduction -----	12
2.1	Objectives -----	12
2.2	Memory System -----	12
2.3	Characteristics Terms for Various Memory Devices -----	13
2.4	Main Memory or Primary Storage -----	15
2.5	Conclusion -----	16
2.6	Summary -----	16
2.7	Tutor-Marked Assignment (TMA) -----	16
2.8	Further Reading. -----	16
Unit 3	External/Auxiliary Memory -----	17
3.0	Introduction -----	17
3.1	Objectives -----	17
3.2	External/Auxiliary Memory -----	17
	3.2.1 Magnetic Disk -----	17
	3.2.2 Winchester Disk -----	18
	3.2.3 Magnetic Tape -----	19
	3.2.4 Optical Memories -----	19
3.3	High Speed Memories -----	20
3.4	Conclusion -----	22
3.5	Summary -----	22
3.6	Tutor-Marked Assignment (TMA) -----	22
3.7	Further Reading -----	22
Unit 4	Input/Output Organization and New Technologies -----	23
4.0	Introduction -----	23
4.1	Objectives -----	23
4.2	Input /Output (Peripherals) -----	23
	4.2.1 Input Devices -----	23

	4 2.2 Output Device -----	25
4.3	Conclusion -----	28
4.4	Summary -----	28
4.5	Tutor-Marked Assignment (TMA) -----	28
4.6	Further Readings -----	28
Unit 5	Interfaces, Parallel Processing, and Vector Processing, Pipelining and Introduction to RISC -----	29
5.0	Introduction -----	29
5.1	Objectives -----	29
5.2	Input/Output Module Interface -----	29
5.3	External Interfaces -----	30
5.4	What is Parallel Processing? -----	31
5.5	Pipelining -----	32
5.6	Vector Processing -----	33
5.7	Introduction to RISC -----	34
	5.7.1 Reasons for Increase Complexity -----	34
	5.7.2 Principles of RISC -----	35
5.8	Conclusion -----	36
5.9	Summary -----	36
5.10	Tutor-Marked Assignment (TMA) -----	36
5.11	Further Reading -----	36
Module 2 Computer Fundamentals: Software -----		37
Unit 1	Software Concepts and Terminology -----	38
1.0	Introduction -----	38
1.1	Objectives -----	39
1.2	Computer Software -----	39
	1.2.1 System Software -----	40
	1.2.2 Application Software -----	40
1.3	Conclusion -----	42
1.4	Summary -----	42
1.5	Tutor-Marked Assignment (TMA) -----	42
1.6	Further Reading -----	42
Unit 2	Categories of Languages -----	43
2.0	Introduction -----	43
2.1	Objectives -----	43
2.2	Categories of Languages -----	43
	2.2.1 Machine Language -----	43
	2.2.2 Assembly Language -----	44
	2.2.3 High Level Language -----	45
	2.2.4 Fourth Generation Language -----	46
2.3	Conclusion -----	47
2.4	Summary -----	47
2.5	Tutor-Marked Assignments (TMA) -----	47
2.6	Further Reading -----	47

Unit 3	Elements of Programming Language	48
3.0	Introduction -----	48
3.1	Objectives -----	48
3.2	Elements of a Programming Language -----	48
	3.2.1 Variables, Constants, Data Type, Array and Expressions -----	48
	3.2.2 Input and Output Statement -----	51
	3.2.3 Conditional and Looping Statement -----	52
	3.2.4 Subroutine and Functions -----	53
3.3	Conclusion -----	53
3.4	Summary -----	54
3.5	Tutor-Marked Assignment (TMA) -----	54
3.6	Further Reading -----	54
Unit 4	Operating System Concepts --	55
4.0	Introduction -----	55
4.1	Objectives -----	55
4.2	What is an Operating System? -----	56
4.3	Evolution of Operating Systems -----	57
	4.3.1 Serial Processing -----	57
	4.3.2 Batch Processing -----	57
	4.3.3 Multiprogramming -----	59
4.4	Conclusion -----	61
4.5	Summary -----	61
4.6	Tutor-Marked Assignment (TMA) -----	61
4.7	Further Reading -----	61
Unit 5	Types of Operating System -----	62
5.0	Introduction -----	62
5.1	Objectives -----	62
5.2	Types of Operating System -----	62
	5.2.1 Batch Operating System -----	62
	5.2.2 Multiprogramming Operating System -----	63
	5.2.3 Network Operating System -----	64
	5.2.4- ¹ Distributed Operating System -----	64
5.3	Conclusion -----	65
5.4	Summary -----	65
5.5	Tutor-Marked Assignment (TMA) -----	65
5.6	Further Reading -----	65
Module 3	Communication and Networking -----	66
Unit 1	Fundamentals of Data Communication -----	67
1.0	Introduction -----	67
1.1	Objectives -----	67
1.2	Definitions -----	67
1.3	Concept of Data Communication -----	67
	1.3.1 Data Communication Codes (Data Encoding) -----	68
	1.3.2 Communication Speed or Rate -----	68
	1.3.3 Private vs. Switched Channels -----	68
	-1.3.4 Analog and Digital Transmission -----	68

	1.3.5 Parallel and Serial Transmission -----	69
	1.3.6 Tariff -----	69
	1.3.7 Communication Protocols/Standard -----	69
14	Conclusion -----	70
15	Summary -----	70
16	Tutor-Marked Assignment (TMA) -----	70
17	Further Reading -----	70
Unit 2	Data Communication and Hardware -----	71
2.0	Introduction -----	71
21	Objectives -----	71
22	Data Communication -----	71
	2.2.1 Synchronous and Asynchronous Transmission -----	71
	2.2.2 Simplex, Half Duplex and Full Duplex Communication -----	72
23	Communications Hardware -----	73
	2.3.1 Sender and Receiver Hardware -----	73
	2.3.2 Communication Devices -----	74
	2.3.3 Communication Channels -----	76
2.4	Conclusion -----	77
2.5	Summary -----	77
2.6	Tutor-Marked Assignment (TMA) -----	77
2.7	Further Reading -----	77
Unit 3	Introduction to Computer Networks and Emerging Trends -----	78
3.0	Introduction -----	78
31	Objectives -----	78
32	Network Concept and Classification -----	78
33	Local Area Network (LAN) -----	79
	3.3.1 LAN Topology -----	79
	3.3.2 LAN Access Method -----	81
	3.3.3 Communication Architecture for Networks -----	82
	3.3.4 LAN Hardware and Software -----	84
	3.3.5 LAN Software/Operating System -----	85
3.4	Conclusion -----	85
3.5	Summary -----	85
3.6	Tutor-Marked Assignment (TMA) -----	85
3.7	Further Reading -----	85
Unit 4	Wide Area Network -----	86
4.0	Introduction -----	86
41	Objectives -----	86
42	Wide Area Network -----	86
	4.2.1 Communication Switching Techniques -----	86
	4.2.2 WAN Devices/Hardware -----	87
	4.2.3 Types of Wide Area Networks -----	88
	4.2.4 Public Networks -----	88
4.3	Conclusion -----	89
4.4	Summary -----	89
4.5	Tutor-Marked Assignment (TMA) -----	89

4.6	Further Reading -----	90
Unit 5	Networking Applications	91
5.0	Introduction -----	91
5.1	Objectives - -----	91
5.2	Networking Applications -----	91
5.2.1	E-mail (Electronic Mail) - -----	91
5.2.2	EDI -----	92
5.3	Networking Scenario -----	94
5.3.1	Internet -----	94
5.3.2	BITNET -----	94
5.3.3	Compuserve -----	95
5.3.4	ISDN -----	95
5.3.5	NICNET -----	95
5.4	Conclusion -----	95
5.5	Summary -----	95
5.6	Tutor-Marked Assignment (TMA)-----	95
5.7	Further Reading-----	95
Module 4	The Management of Computer Security and Principles of Cryptography-----	96
Unit 1	The Management of Computer	97
1.0	Introduction -----	97
1.1	Objectives -----	97
1.2	Definitions -----	97
1.3	Security Status on PC -----	97
1.4	Breaches of Security -----	98
1.5	Conclusion -----	99
1.6	Summary -----	99
1.7	Tutor-Marked Assignment -----	99
1.8	Further Reading -----	99
Unit 2		100
2.0	Introduction -----	100
2.1	Objectives -----	100
2.2	Security Measures -----	100
2.2.1	Physical Security -----	101
2.2.2	Software Security -----	102
2.2.3	Network Security -----	102
2.2.4	Password Security -----	103
2.3	Conclusion -----	103
2.4	Summary -----	104
2.5	Tutor-Marked Assignment (TMA)-----	104
2.6	Further Reading -----	104
Unit 3	Cryptography	105
3.0	Introduction -----	105
3.1	Objectives -----	105
3.2	Cryptography: A Brief History -----	105
3.3	Cryptography -----	106

3.3.1	Cipher Systems -----	106
3.3.2	Data Encryption Standard (DES) -----	107
3.3.3	RSA Approach to Encryption -----	107
3.4	Cryptanalysis -----	108
3.5	Conclusion-----	108
3.6	Summary-----	109
3.7	Tutor-Marked Assignment (TMA) -----	109
3.8	Further Reading -----	109
Unit 4	Computer Virus -----	110
4.0	Introduction -----	110
4.1	Objectives -----	110
4.2	The Evolution of Virus-----	110
4.3	The Menace -----	111
4.4	The Process of Infection -----	111
4.5	Classification of Viruses-----	112
4.5.1	BOOT Infectors -----	112
4.5.2	SYSTEM Infectors -----	112
4.5.3	GENERAL.COM or EXE Infectors -----	112
4.6	Conclusion-----	113
4.7	Summary-----	113
4.8	Tutor-Marked Assignment (TMA) -----	113
4.9	Further Reading -----	113
Unit 5	Viruses and Prevention -----	114
5.0	Introduction -----	114
5.1	Objectives -----	114
5.2	Some Viruses -----	114
5.3	Prevention -----	116
5.4	The Cure -----	116
5.5	Conclusion-----	117
5.6	Summary-----	117
5.7	Tutor-Marked Assignment (TMA) -----	117
5.8	Further Reading -----	117

Course Guide

Introduction

Computer Fundamental is a one semester and one credit foundation level course for first year students in Postgraduate Diploma in Information Technology and Postgraduate Diploma in Information Technology for Teachers.

The course will consist of four modules. The course has been written in such a way that you can work at your own pace subject to the requirement of National Open University of Nigeria.

Course Introduction

This course presents an overview of the technology relating to the computer systems and exposes you to operating systems and application packages. The course covers aspects on computer hardware, computer software, data communication and computer security.

The first and second modules discuss hardware and software of computer system. It includes discussions on data presentation, memory system of computers, input/output devices, parallel organisation, pipelining and reduced instruction set computers. It also covers aspects relating to the computer software such as the components of programming languages and operating system concepts.

The third and fourth modules present an introduction to data communication, networking and related technology. The field of data communication has progressed tremendously during the past decade. Therefore, the concepts relating to it are of utmost importance for a computer user.

What you will learn in this Course

In this course you will be exposed to Operating Systems and Application packages. The course covers aspects of computer hardware and computer software. It includes discussions on data presentation, memory system of computers, input/output devices, parallel organisation, pipelining and reduced instruction set computers. This course also covers data communication, computer security and virus.

Course Aims

The aim of the course can be summarized as follows: to give you an overview of the technology relating to computer systems, operation systems and application packages.

Course Objectives

On successful completion of the course you should be able to:

- explain hardware and software of computer system.
- explain data communication, networking and related technology.

Working through this Course

To complete this course you are required to study the course materials carefully and attempt all selfassessment items and then answer all *TutorMarked Assignments*. Materials on further reading are not compulsory.

The materials could be of added advantage but you may do without them. At the end of the course is a final examination.

Course materials

The major components of the course are:

1. Course Guide
2. Modules
3. Assignment File
4. Presentation Schedule

Assessment

There are two aspects to the assessment of the course. First are the Tutor-Marked Assignments. The second is the final examination.

How to get the most from this Course

In distance learning, the study units replace the university lecturer. This is one of the great advantages of distance learning; you can read and work through specially designed study materials at your own pace, and at a time and place that suit you best. Think of it as reading the lecture instead of listening to the lecturer. Just as a lecturer might give you an in-class exercise, your study units provide exercises for you to do at appropriate points.

Each of the study units follows a common format. The first item is an introduction to the subject matter of the unit. Next is a set of learning objectives. These objectives let you know what you should be able to do by the time you have completed the unit. When you have finished the unit you must go back and check whether you have achieved the objectives. This habit will significantly improve your chances of passing the course.

The main body of the unit guides you through the required reading materials. The following is a practical strategy for working through the course.

1. Read this course guide thoroughly.
2. Organize a study schedule. Note how the assignments relate to the units.
3. Once you have created your own study schedule, do everything you can to stick to it.
4. Turn to a unit and read the introduction and the objectives for the unit.
5. Work through the unit.
6. Review the objectives for each study unit to confirm that you have achieved them.
7. When you are confident that you have achieved a unit objectives, do the Tutor-Marked Assignments.
8. Proceed unit by unit.

Tutors and Tutorials

You will be notified of the date, times and location of where to submit the tutorials. Your tutor will mark and comment on your assignment, keep a close watch on your progress and on any difficulties you might encounter and provide assistance to you during the course.

Summary

Computer Fundamentals—intends to introduce you to the fundamentals of computer science. Upon completing the course, you will be equipped with basic knowledge of computer hardware and software.

Module 1

Computer Fundamentals: Hardware

Introduction

This is the first module of the course. In this module many terms and concepts related to computers have been introduced. The module is divided into five units.

In unit 1, we define a computer and concepts relating to Von Neuman Architecture.

In unit 2, you are introduced to memory system and characteristics terms for various memory devices.

Unit 3 defines external /auxiliary memory and explains high speed memories.

Unit 4 deals with input/output organization and new technologies.

Unit 5 explains important concepts of interfaces, parallel processing, vector processing and pipelining.

Unit 1

What is a Computer?

1.0 Introduction

The use of Information Technology is well recognised. It has become a must for the survival of business houses with growing information needs. Computer is one of the major components of an information technology network. Today, computer technology has permeated every sphere of existence of modern man. From railway reservations to medical diagnosis, from TV programmes to satellite launching, from match-making to criminal catching— everywhere, we witness the elegance, sophistication and efficiency possible only with the help of computers.

In this unit, we will introduce you to the computer hardware technology, how does it work and what it is. In addition we will also, discuss some of the terminology closely linked, with information technology and computers. More details on these terms can be obtained from further readings. In this unit, we shall discuss about basic computer structure.

1.1 Objectives

This unit being the first unit of the block introduces you to the world of computers. At the end of the unit, you will be able to:

- define the term computer
- define Von Neumann architecture
- describe key characteristics of memory system.

1.2 What is a Computer?

Let us first define the term "Computer". In Oxford Dictionary, Computer is defined as "An automatic electronic apparatus for making calculations or controlling operations that are expressible in numerical or logical terms".

The basic function performed by a computer is the execution of a program. A program is a sequence of instructions, which operate on data to perform a certain task. In modern digital computers, data is represented in binary form by using two symbols 0 and 1 which are called binary digits or bits. Computers use eight bits to represent a character internally. This allows up to $2^8 = 256$ different items to be represented uniquely. This

collection of eight bits is called a byte. Thus, one byte is used to represent one character internally. One of the most common codes to represent characters in computers is ASCII (American Standard Code for Information Interchange). Most computers use two bytes or four bytes to represent numbers (positive and negative) internally. Another term that is commonly used in computer is a Word, A word may be defined as a unit of 1

; information that a computer can process or transfer at a time. A word must be equal to the number of bits transferred between the central processing unit and the main memory in a single step or it may be defined as 2^N bits, where N is the number of bits transferred in a single step.

the basic unit of storage of integer data in a computer. Normally, a word may be equal to 8, 16, 32 or 64 bits. The terms 32 bit computer, 64 bit computers etc basically point to the word size of the computer.

One of the key aspects in program execution is the execution of an instruction. The key questions that can be asked in this respect are (a) how are instructions supplied to the computer? and (b) how are they interpreted and executed? We will answer these questions along with the discussion on the basic structure of the computer system.

Most of today's computer designs are based on concepts developed by John Von Neumann referred to as a Von Neumann Architecture. Von Neumann proposed that there should be a unit performing arithmetic operations on the data. This unit is termed Arithmetic Logic Unit (ALU). A control unit directs the ALU to perform specific arithmetic and logical functions on the data. Therefore, in such a system by changing the control signal, the desired operation can be performed on data.

But, how can these control signals be supplied? Let us try to answer this from the definition of a program. A program consists of a sequence of steps. Each of these steps is either a certain or logical or input/

output operations to be performed on data. Therefore, each step may require a set of control signal. Is it possible for us to provide a unique code for each set of control signals? Well, the answer is Yes. But what do we do with these codes? What about adding a hardware segment that accepts the code and generates control signals? The unit that interprets a code to generate respective control signal is termed as Control Unit (CU). Thus, a program now consists of a sequence of codes. This machine is quite flexible, as we only need to provide a new sequence of codes for a new program. Each code is, in effect, an instruction, for the computer. The hardware interprets instructions and produces the respective control signals.

The Arithmetic Logic Unit (ALU) and the Control Unit (CU) together are termed as the Central Processing Unit (CPU). The CPU is the most important component of a computer's hardware. The ALU performs the arithmetic operations such as, subtraction, multiplication and division, and the logical operations such as: "Is A = B? (where A and B are both numeric or alphanumeric data). Is a given character equal to M (for male) or F (for female)?" The control unit interprets instructions and produces the respective control signals. All logical operations are performed in the CPU in several areas called registers. One of the important considerations in determining the processing capabilities of the CPU. Register size refers to the amount of information that can be held in a register at a time for processing.

1

The speed of processing. A CPU's processing power is measured in Million Instructions Per Second (MIPS).

- How can instructions and data be put into computers? The instruction and data are to be supplied by external environment. Therefore, an input module is needed. The main responsibility of input module will be to put the data in the form of signals that can be recognized by the system. Similarly, we need another component that will report the results in proper format and form. This component is called output module. These components are referred together as input/output (I/O) components.

Are these two components sufficient for a working computer? No, because, input devices can bring instructions or data only sequentially and a program may not be executed sequentially as jump instructions are normally encountered in programming. In addition, more than one data element may be required at a time. Therefore, a temporary storage area is needed in a computer to store instructions and data temporarily. This component is referred to as memory. It was pointed out by Von Neumann that the same memory can be used for storing data and instructions. In such case, the data can be treated as data on which processing can be performed, while instructions can be treated as data, which can be used for the generation of control signals.

The memory unit stores all the information in a group of memory cells, also called memory locations, as binary digits (bits). Each memory location has a unique address and can be addressed independently. The contents of the desired memory locations are provided to the central processing unit by referring to the address of memory location. The amount of information that can be held in the main memory is known as memory capacity. The capacity of the main memory is measured in Kilobytes (KB) or Megabytes (MB). One kilobyte stands for 2 bytes, which is approximately little over one million bytes.

common interconnection structure is the Bus Structure. A bus is a set of Wires (Lines) which you can visualize on the motherboard of a computer. It is a shared media4_121.a_corinecirtt,U memory and I/O components is called a system bus. A system bus may consist of 50 to 100 separate lines.

Let us summarise the key features of a Von Neumann machine. 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 uses 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. is very difficult as they were entered manually by setting switches and plugging and unplugging.
- Each location of the main memory of Von Neumann machine is carried out in a sequential fashion (unless explicitly altered by the program itself) from one instruction to the next.

Figure 1 shows the basic structure of a conventional Von Neumann machine.

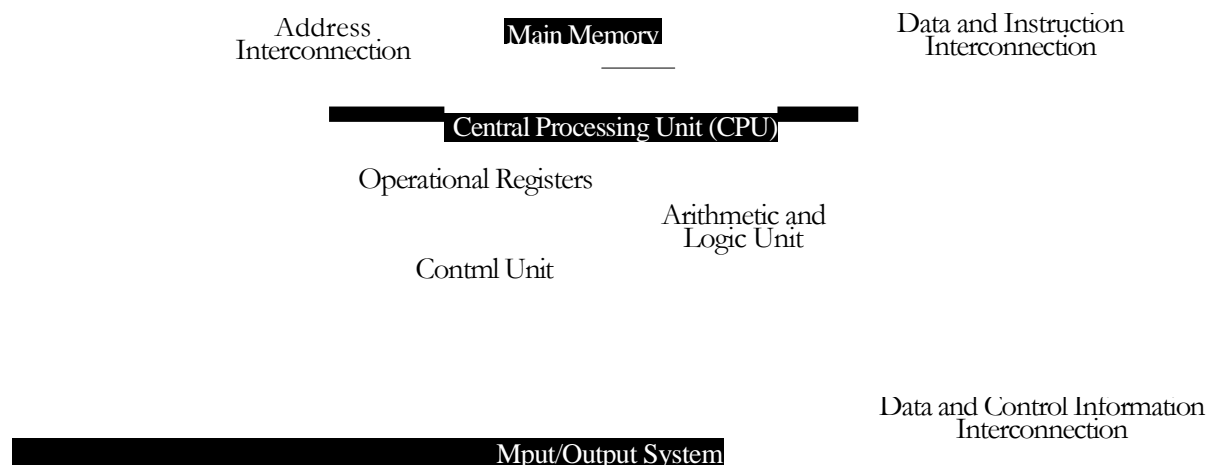


Fig. 1: Structure of a Computer

A Von Neumann machine has only a single path between main memory and control unit (CU). This feature/constraint is referred to as Von Neumann bottleneck. Several other architectures have been suggested for modern computers.

Exercise 1.1

State True or False

1. A byte is equ to 8 bits and can represent a character internally.

True _____ False _____

2. A word of PC386 is equal to one byte

True _____ False _____
 ...mkr_11--a _____

G o o c c o o h e d ' /
 tr33-Cc etaicalenis -a' a n St eft 1/0 tb-.-
 /3 CIA-Ze3 Aitc.41R Prim R942-c>,
 re-P\ 0-eg- g; Gect-7 iNcied- SA-n/ 1
 6.NI>CAmclA 2.3 a s eivvi co n AAA r
 frt!..4t.0¹⁻¹+A 4-412,21 im faie-,7_osci'tt
 oc -etcaarib, c/ Jrr

3. Von Neumann architecture specifies separate memory for data and instructions. The memory which stores data is called data memory and the one that stores instructions is called instruction memory.

True ☐ False ☒

4. In Von Neumann architecture each bit of memory can be accessed independently.

True ☒ False ☐

5. A program is a sequence of instructions designed for achieving a task/goal.

True ☒ False ☐

7. One MB is equal to 1024 KB.

True False

1.2.1 The Computer and Integrated Circuit Technology

Before we discuss the relation of computers to integrated circuit technology, let us explore more about the term Integrated Circuit (IC). In an integrated circuit, the components such as transistors, resistors and conductors are fabricated on semi-conductor material such as silicon. Thus, a desired circuit can be fabricated in a tiny piece of silicon rather than assembling several discrete components into the same circuit. Hundreds or even thousands of transistors could be fabricated on a single wafer of silicon. In addition, these fabricated transistors can be connected with a process of metalization on the same chip they have been produced.

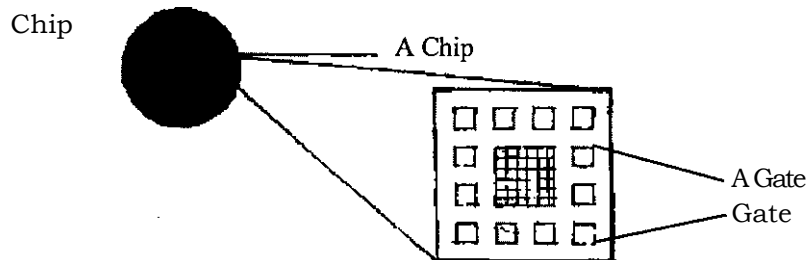
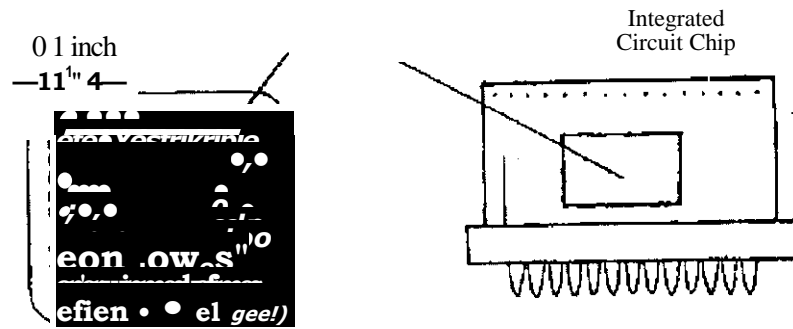


Fig. 2: Wafer, chip and Gate

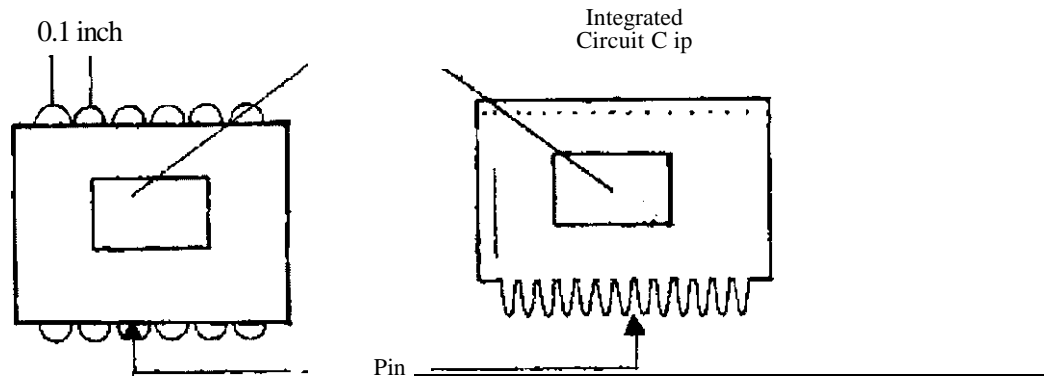
An integrated circuit is constructed on a thin wafer of silicon that is divided into a matrix of small areas (size of the order of a few millimeter-squared). An identical circuit pattern is fabricated on each of these areas and, the wafer is then broken into chips (Refer to figure 2). Each of these chips, consist of several gates, a useful logic component, and a number of input and output connection points. Each of these chips, then, can be packaged separately in a housing to protect it. In addition, this housing provides a number of pins for connecting this chip with other devices or circuits. The pins on these packages can be provided in two ways:

- In two parallel rows with 0.1 inch spacing between two adjacent pins in each row. This package is called dual in line package (DIP) (Refer Figure 3a.)
- In case, more than hundred pins are required. then single arrays of rows and columns, with spacing between two adjacent pin of 0.1 inch (Refer to Figure 3b)

Chia_4 cf calk cell, q. n 9 n



(B) 144 Pin Grid Array (PGA) Package



(A) A 24 pin dual in-line package (DIP)

(Contains dual rows which are in-line)

Fig. 4: Integrated Circuit Packages

Different circuits can be constructed on different wafers. All these packaged circuit chips then can be interconnected on a printed-circuit board to produce several complex electronic circuits such as computers.

Initially, only a few gates were integrated reliably on a chip and then packaged. This initial integration was referred to as small-scale integration (SSI). Later, with the advances in micro-electronics technologies, the SSI gave way to Medium Scale Integration where 100s of gates were fabricated on a chip. Then came Large Scale Integration (1,000 gates) and Very Large Scale Integration (VLSI 100,000,000 components) are expected to be 10,000,000 components will be fabricated on a single chip.

What are the advantages of having densely packed Integrated Circuits? These are:

Low cost: The cost of a chip has remained almost constant while the chip density (number of gates per chip) is ever increasing. It implies that the cost of computer logic and memory circuit has been reducing rapidly.

Greater Operation Speed: More is the density, and closer are the logic or memory elements, which implies shorter electrical paths and hence higher operating speed.

Smaller computers: Better portability, reduction in power and cooling requirements.

Reliability: The integrated circuit interconnections are much more reliable than soldered connections. In addition, densely packed integrated circuits enable fewer inter-chip connections. Thus, the computers are more reliable.

One of the major milestones in this technology was the Very Large Scale Integration (VLSI) where thousands of transistors can be integrated in, a single chip. The main impact of VLSI was that, it was possible to produce as many as CPU or main memory or other similar devices on a single IC chip. Let us discuss some of the important breakthroughs of VLSI technologies.

Semiconductor Memories

Initially, the IC technology was used for constructing processor, but soon it was realized that same technology can be used integrated on a single chip. The first memory chip was constructed in 1970 and could hold 256 bits. Although the cost of this chip was high, but gradually the cost of semiconductor memory came down. The memory capacity per chip has increased e.g. 1K, 4K, 16K, 64K, 256K and 1M bits.

Microprocessors

Keeping pace with electronics as more and more components were fabricated on a single chip, fewer chips were needed to construct a single processor. Intel in 1971 achieved the breakthrough of putting all the components on a single chip. The silicon processor is known as a microprocessor. The Intel 4004 was the first microprocessor. It was a primitive microprocessor designed for a specific application. Intel 8080 that came in 1974 was the first general purpose microprocessor. It was an 8 bit microprocessor. Motorola is another manufacturer in this area. At present 32 and 64 bit general purpose microprocessors are already in the market. For example, Intel Pentium is a 32 bit processor, similarly Motorola's 68000 is a 32 bit microprocessor. The Pentium that is announced by Intel in 1995 can process 64 bit data at a time. Figure 4 shows the Intel family of microprocessors.

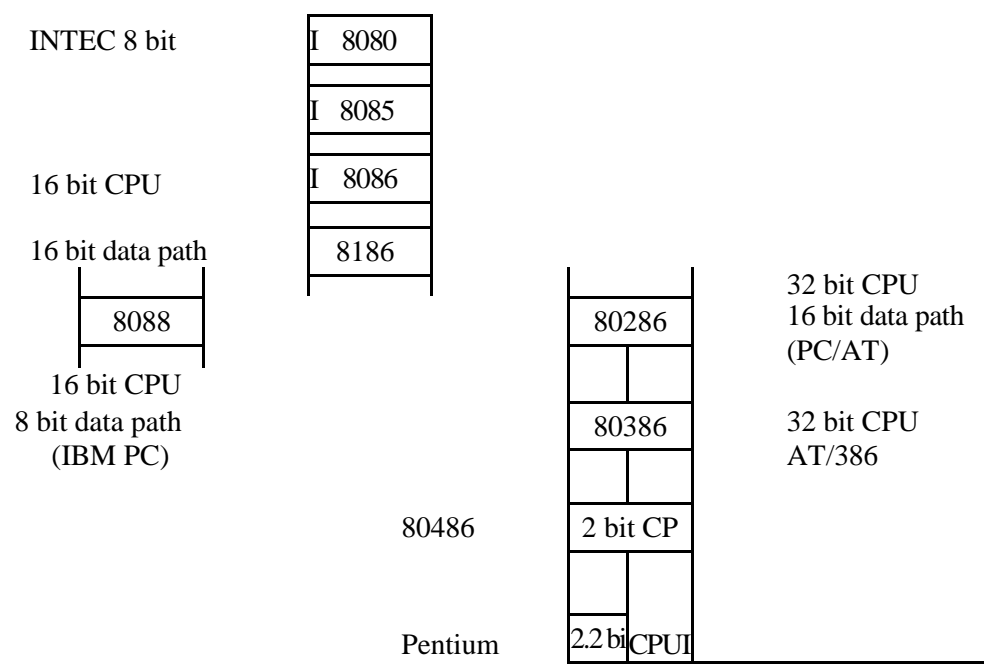


Fig. 4: Intel Microprocessor Families

The VLSI technology is still evolving and more and more powerful microprocessor and more storage space is being put in a single chip.

1.2.2 Classification of Computers

One question which we have still not answered is, is there any classification of computers? Well for quite sometime computers have been classified under three main classes. These are:

- Microcomputers
- Minicomputers
- Mainframes

Although with development in technology the distinction between them is becoming blurred, yet it is important to classify them, as it is sometimes useful to differentiate the key elements and architecture among the classes.

Microcomputers

A microcomputer's CPU is a microprocessor. The first microcomputers were built around 8-bit microprocessor chips. What do we mean by an 8-bit chip? It means that the chip can retrieve instructions/data from storage, manipulate, and process, an 8-bit data at a time or we can say that the chip has a built in 8-bit data transfer path. An improvement on 8-bit chip technology was in early 1980s, when a series of 16-bit chips namely 8086 and 8088 were introduced by Intel Corporation, each one with an advancement over the other. 8088 is a 8/16 bit chip i.e. an 8-bit path is used to move data between chip and primary storage (external path), at a time, but processing is done within the chip using a 16 bit path (internal path) at a time. 8086 is a 16/16 bit chip i.e. the internal and external paths both are 16 bit wide. Both these chips can support a primary storage capacity of up to 1-megabyte (MB).

Most of the popular microcomputers are developed around Intel's chips, while most of the minis and super -minis are built around Motorola's 68000 series chips. With the advancement of display and VLSI technology now a microcomputer is available in very small size. Some of these are laptops, notebook computers etc. Most of these are of the size of a small note but equivalent capacity of an older mainframe.

Minicomputer

The term minicomputer originated in 1960s when it was realized that many computing tasks do not require an expensive contemporary mainframe computers but can be solved by a small, inexpensive computer. Initial minicomputers were 8 bit and 12 bit machines but by 1970s almost all minicomputers were 16 bit machines. The 16 bit minicomputers have the advantage of large instruction set and address field; and efficient storage and handling of text, in comparison to lower bit machines. Thus, 16 bit minicomputer was more powerful machine which could be used in variety of applications and could support business applications along with the scientific applications.

With the advancement in technology the speed, memory size and other characteristic developed and the minicomputer was then used for various stand alone or dedicated applications. The minicomputer was then used as a multi-user system, which can be used by various users at the same time. Gradually the architectural requirement of minicomputers grew and a 32-bit minicomputer, which was called super mini, was introduced. ; The supermini had more peripheral devices, large memory and could support more users working simultaneously on the computer in comparison to previous minicomputers.

Mainframes

Mainframe computers are generally 32-bit machines or on the higher side. These are suited to big organizations, to marriage high volume applications. Few of the popular mainframe series are IBM, HP, etc.

Mainframes are also used as central host computers in distributed systems. Libraries of applications programs developed for mainframe computers are much large than those of the micro or minicomputers because of their evolution over several decades as families of computing. All these factors and many more

make the mainframe computers indispensable even with the popularity of microcomputers.

Supercomputer

The upper ends of the state of the art mainframe machine are the supercomputer. These are among the fastest machines in terms of processing speed, and use multi-processing techniques, where a number of processors are used to solve a problem. Lately ranges of parallel computing products, which are multiprocessors sharing common buses, have been in use in combination with the mainframe supercomputers.

The supercomputers are reaching up to speeds as well over 25000 million arithmetic operations per second.

India also has its indigenous supercomputer.

Supercomputers are mainly being used for weather forecasting, computational fluid dynamics, remote sensing, image processing, biomedical applications, etc. In India, we have one such mainframe supercomputer system — CRAY XMIP — 14, which is at present, being used by Meteorological Department.

13 Conclusion

In this unit you have learned a number of important issues that relate to a computer. In particular you learned the main units of a computer. You also learned the relationship between a computer and integrated circuit technology.

1.4 Summary

What you learned in this unit concerns general definitions of a computer and what Von Neumann Architecture is. What you learned here will be useful in the next units.

1.5 Tutor-marked Assignment (TMA)

1. What is a general purpose machine?
3. Define the following terms briefly
 - (i) Microprocessor
 - (ii) Laptop
 - (iii) Supercomputer

chip processor
laptop computer

r

1.6 Further Reading

Milan Milenkovic, *Operating Systems Concepts and Designs*, Second Edition. Tata McGraw Hill.

A supercomputer is a type of computer that is designed to perform a large number of calculations at a very high speed. It is used for a wide range of applications, including weather forecasting, scientific research, and data processing. Supercomputers are typically composed of many processors working together to solve a problem. They can perform millions of operations per second.

In a nutshell, a supercomputer is a computer that is designed to perform a large number of calculations at a very high speed. It is used for a wide range of applications, including weather forecasting, scientific research, and data processing. Supercomputers are typically composed of many processors working together to solve a problem. They can perform millions of operations per second.

Unit 2

Memory System

2.0 Introduction

In this unit we will introduce you to the memory system and characteristics of various memories. We will also define terms such as the main memory, cache memory, secondary storage and optical memories.

2.1 Objectives

This unit introduces you to memory of a computer. At the end of the unit, you will be able to:

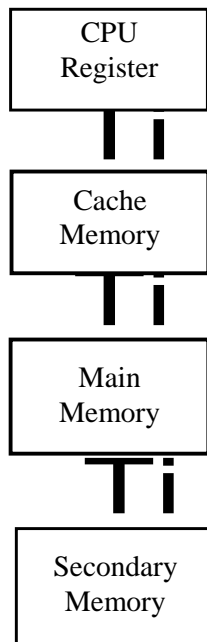
- describe key characteristics of a memory system
- distinguish various types of memories

2.2 Memory System

Memory in a computer system is required for storage and subsequent retrieval of instructions and data. A computer system uses variety of devices for storing the instructions and data, which are required for its operations. Normally we classify the information to be stored on computer in two basic categories, data and the instructions.

The storage device along with the algorithm or information on how to control and manage these storage devices constitute the memory system of a computer. A memory system is a very simple system yet it exhibits a wide range of technology and types. But unfortunately, faster memory technology is more expensive. In addition, fast memories require power supply till the information need to be stored. Both these things are not very convenient, but on the other hand the memories with less cost have very high access time, that is the time taken by CPU to access a location in the memory is high, which will result in slower operation of the CPU. Thus, the cost versus access time anomaly has led to a hierarchy of memory where we supplement fast memories with larger, cheaper and slower memories. These memory units may have very different physical and operational characteristics, therefore, making the memory system very diverse in type, cost, organization, technology and performance. This memory hierarchy will be fruitful if the frequencies of access to slower memories are significantly less than the faster memories.

Memory hierarchy diagram showing the relationship between different types of memory (Cache, Main Memory, Secondary Storage) and their access times.



Thus, a memory system can be considered to consist of three groups of memories. These are:

- Internal Processor Memories:** These consist of the small set of high speed registers which are internal to a processor and are used as temporary locations where actual processing is done.
- Primary Memory or Main Memory:** It is a large memory, which is fast but not as fast as internal processor memory. This memory is accessed directly by the processor. It is mainly based on integrated circuits.
- Secondary Memory/Auxiliary Memory/Backing Store:** Auxiliary memory in fact is much larger in size than main memory but is slower than main memory. It normally stores system programs (programs which are used by system to perform various functions), other instructions, programs and data files. Secondary memory can also be used as an overflow memory in case the main memory capacity has been exceeded. (How? the answer is not supplied in the block. You need to refer to further readings to get this answer). Secondary mem_orie cannot be accessed directly by a processor. First the information of these memories_ is transferred to the main memory and then the information can be accessed as the information of main memory.

There is another kind of memory, which is increasingly being used in modern computers and this is called *Sacittuemery*. It is logically positioned between the internal memory (registers) and main memory. It stores or catches some of the content of the main memory, which is currently in use of the processor.

Before discussing more about these memories let us first discuss the technological terms commonly used in defining and accessing the memory.

2.3 Characteristics Terms for Various Memory Devices

The following terms are most commonly used for identifying comparative behaviour of various memory devices and technologies:

Storage Capacity: It is a representative of the size of the memory. The capacity of internal memory and main memory can be expressed in terms of number of words or bytes. The storage capacity of external memory is normally measured in terms of bytes.

Access Modes: A memory is considered to consist of various memory locations. The information from these memory locations can be accessed in the following ways:

- **Random Access Memory (RAM):** It is the mode in which any memory location can be accessed

in any order in the same amount of time. Ferrite and Semiconductor memories, which generally constitute main memory, are of this nature. The storage locations can be accessed independently and there exist separate access mechanism for each location.

- **Sequential Access:** On the other hand we have memories which can be accessed in a pre-defined sequence. For example, the songs stored on a cassette can be accessed only one by one. The example of sequential access memory is Magnetic Tape. Here the access mechanism need to be shared among different locations. Thus, either the location or the location or the read/write head or both should be moved to access the desired location.
- **Direct Access:** In certain cases the information is neither accessed randomly nor in sequence but something in between. In direct access, a separate read/write head exist for a track and on a track - \ the Information be accessed serially: this semi-random mode of operations exists in magnetic disks.

Access Time: The access time is the time required between the request made for a read or write operation till the time the data is made available or written at the requested location. Normally it is measured for read operation. The access time depends on the physical characteristics and access mode used for that device.

Permanence of Storage: Some memories loose information over a period of time. For example, there can be some memories where the stored data bit value 1 loses its strength to become 0 over a period of time. These kinds of memories require refreshing. The memories, which require refreshing, are termed as dynamic memories. In contrast, the memories, which do not require refreshing are called static memories. Another factor, which can destroy the contents, is the presence and absence of electricity. The memories which lose their content on failure of power are termed as volatile memories those which do not are called non-volatile. Magnetic memories are non-volatile and semiconductor main memories are volatile in nature.

Physical Characteristics: In this respect, the memory devices can be categorized into four main categories viz, electronic, magnetic, mechanical and optical. One of the requirements for a storage device is that it should exhibit two well-defined physical states, such that 0 and 1 can be represented in those two states. The Data transfer rate of the memory depends on the how quickly the state can be recognized and altered. The following table lists some of the memory technologies along with their physical and other important characteristics.

Technology	Access time (in seconds)	Access Mode	Performance of Storage	Physical nature of storage medium	Average cost(Rs/bit) (Approx.)
Semiconductor	10^{-8}	Random	Volatile	Electronic	10^{-2}
Magnetic disk	10^{-2}	Direct	Non-volatile	Magnetic	10-s
Magnetic tape	10^{-1}	Sequential	Non-volatile	Magnetic	10-s
Compact disk ROM	Approx. 10^{-1}	Direct	Non-volatile	Optical	10^{-1}

Table I: Characteristics of Some Memory Technologies.

The physical size of memories should be small and it must consume less power. Higher power consumption may result in more costly equipment for internal cooling of computer. The storage devices, which require mechanical motion e.g. hard disks are more prone to failure rather than the semiconductor memories which

With low cost semiconductor memory devices, the cost of memory has decreased significantly.

The physical size of memories should be small and it must consume less power. Higher power consumption may result in more costly equipment for internal cooling of computer. The storage devices, which require mechanical motion e.g. hard disks are more prone to failure rather than the semiconductor memories which

With low cost semiconductor memory devices, the cost of memory has decreased significantly.

The physical size of memories should be small and it must consume less power. Higher power consumption may result in more costly equipment for internal cooling of computer. The storage devices, which require mechanical motion e.g. hard disks are more prone to failure rather than the semiconductor memories which

are totally electronic in nature. Very high speed semiconductor memories are also prone to failure as technology is moving towards its limits.

2.4 Main Memory or Primary Storage

Er.imarnetacuy consists of semiconductor memory chips and is used to store the data and program urrently in_use—Each storage element of memory is directly (randomly) accessible and can be examined and thbdfdd without affecting other cells and hence primary memory is also called Random Access Memory (RAM). Main memory stores a variety of critical information required for processing by the CPU. How does it store the information? Please answer it yourself.

The memory unit stores all the information in memory cells also called memory locations, in binary digits. Each memory location has a unique address. The contents of the desired memory locations are provided to the central processing unit by referring to the address of the memory location. The amount of information that can be held in the main memory is known as memory capacity. The capacity of the main memory is measured in kilobytes (KB) or megabytes (MB).

All modem computers have semiconductor memory as its main memory. Semiconductor memory is known i as Random Access Memory (RAM) since any part of the memory can be accessed for reading and writing. ATC6IWEVait of main memory is Read Only Memory (ROM). ROMs (Read Only Memories) are the memories on which it is not possible to write the data when they are on-line to the computer. They can only...12e reid.The ROMs can be used in storing. programs provided by the manufacturer of computer for basic operations. ROMs are non-volatile in nature and need not be loaded in a secondary storage device. ROMs are fabricated in large number in a way where there is no room for even a single error.

ROMs can be written only at the time of manufacture. Another similar memory is PROM. PROMs are also non-volatile and can be programmed only once by a special write device hence the name Programmable ROM (PROM). The writing process in PROM can be performed eletrically by the supplier or the customer.

Special equipment is needed to perform this writing operation. Therefore, PROMs are more flexible and convenient ROMs.

The ROMs/PROMs can be written just once (in ROMs at the time manufacture and PROMs at any time later), but in both the cases once whatever is written on, cannot be changed. But what about a case where you read mostly but write only very few times. This lead to the concept of Read mostly memories and the best examples of these are EPROMs (Erasable PROMS) and EPROMS (Electrically erasable PROMS). The EPROMS can be written electrically. But, the write operation is not simple. It requires erasure of whole storage cells by exposing the chip to ultra violet light, thus bringing them to same initial state. This venture is a time consuming process. Once all the cells have been brought to same initial state, then the EPROM can be written electrically. EPROMs becoming increasingly popular as they do not require prior erasure of previous contents.

However, in EPROMs, the writing time is considerably higher than reading time. The biggest advantage of EPROM is that it is non-volatile memory and can be updated easily, while the disadvantages are the high cost and at present they are not completely non-volatile and the write operation takes considerable time. Figure 7 summarize the features of these read only and read mostly memories.

Memory Type	Write Time	Order of Read Time	Number of Cycles allowed
ROM	Once at the time of manufacture	Nano seconds	ONE
PROM	Hours	Nano seconds	ONE
EPROM	Minutes (including time of erasure	Nano seconds	HUNDREDS
EPROM	Milliseconds	Nano seconds	THOUSANDS

Common Features

- Non-destructive
- Long data life
- Non-volatile

2.5 Conclusion

In this unit you have learned about the memory system of a computer. We have also introduced you to the characteristics terms of various memory devices. In particular you learned the functions of the main memory.

2.6 Summary

You are now able to describe the characteristics of memory system. You are also in a position to distinguish types of memories. You will need knowledge that you have required in this unit in the subsequent units.

2.7 TutorMarked Assignment

State True or False

1.Memory hierarchy is built in a computer system, as the main memory cannot store very large data

True

False

2.The secondary memory is slower than that of main memory but has larger capacity.

True

False

3.In Random access memory any memory location can be assessed independently.

True

I

False

4.What are the difference in:

- (a) Volatile and Non-volatile memory
- (b) Static and Non-dynamic memory

2.8 Further Reading

Milan Milenkovic, *Operating Systems Concepts and Designs*, Second Edition, Tata Mcgraw Hill.

Unit 3

External / Auxiliary Memory

3.0 Introduction

In this unit we will introduce you to External/Auxiliary memory. You will also learn about high speed memories.

3.1 Objectives

At the end of this unit, you will be able to:

- differentiate various external memories
- define the importance of cadre memory.

3.2 External/Auxiliary Memory

As discussed earlier the cost of RAM is very high and the semiconductor RAMS are mostly volatile, therefore, it is highly important that a secondary cheap media should be used which should show some sort of permanence of storage and should be relatively inexpensive. The magnetic material was found to be inexpensive and quite long lasting. Therefore, it became an ideal choice. Magnetic tape and magnetic disks are commonly used as I storage media. With the advancements in the optical technology now the optical disks are trying to make in roads as one of the major external memory. We will discuss about the characteristics of these memories in this section.

3.2.1 Magnetic Disk

A magnetic disk is a circular platter of plastic, which is coated with magnetized material. One of the key components of a magnetic disk is a conducting coil named as Head which performs the job of reading or writing operation.

Organization and Format

A head of disk is a small coil and reads or writes on the position of the disk rotating below it, therefore, the data is stored in concentric set of rings (refer to figure 4). These are called tracks. The width of a track is equal to the width of the head, the adjacent tracks are separated by inter track gaps. As we go towards the outer tracks the size of a track increase but to simplify electronics same number of bits are stored in each track.

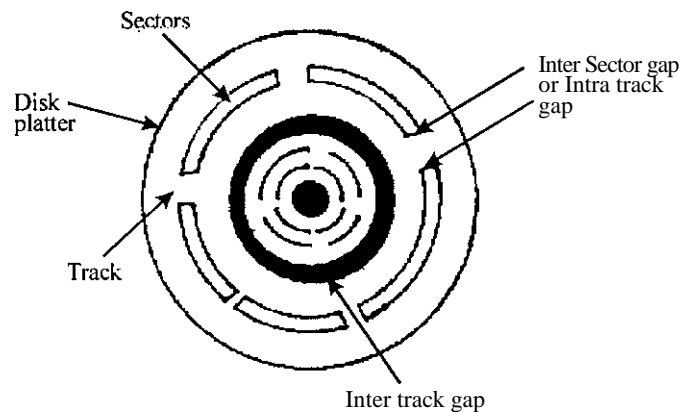


Fig. 8: Logical layout of Magnetic Disk

The data is transferred from and to the disks in blocks. Block is a section of disk data and is normally equal to a sector. A track is divided into 10-100 sectors and these sectors should be either fixed or variable length sectors. Two adjacent sectors are separated by intra-track gaps. This helps in reducing the precision requirements of sectors. To identify the sector position normally there may be a starting point on a track or a starting and end point of each sector.

Floppy disks

A floppy disk is made of a flexible thin sheet of plastic material with a magnetic coating and grooves arranged in concentric circles with tracks. Floppy disk becomes a convenient recording medium to transport information from one location to another. Disk is removable from the reading device attached to the computer and therefore provides unlimited storage capacity. The floppy disks of today are available in two sizes 5.25 inches and 3.5 inches and their capacity ranges from 360 KB to 1.44 MB per disk.

3.2.2 Winchester Disk

This is a sealed rigid magnetic oxide medium disk, which typically holds 101\413 to 10 GB of data.

Winchester disks are not removable from the drives and since they are sealed, dust and other contaminations, which are likely in a floppy disk, are minimized. These provide substantially faster data access compared to floppy disk and provide very large data storage for on-line retrieval.

Sides: The magnetic coating if applied to both the sides of the platter is called a double sided disks. The data can be recorded on either side of these disks. Some inexpensive disks were initially single sided.

Platters: Some disks have single platter e.g. floppy disks while some disks have multiple platters which are stacked vertically, normally at a distance of an inch. This is known as disk pack. In disk pack one additional term cylinder is defined which is the ring of all co-centric tracks (figure 9). A disk pack can contain multiple heads mounted with the same arm.

Access Time on Disk

Disk operates in semi-random mode of operation and normally is referenced block wise. The data access time on disk consist of two main components.

Seek time: Time to position the head on a specific track. On a fixed disks it is the time taken by electronic circuit to select the require head while in movable head disk it is time required to move the head to a particular track.

Latency time: The time required by a sector to reach below the read/write head. On an average it is half of the time taken for a rotation by the disk.

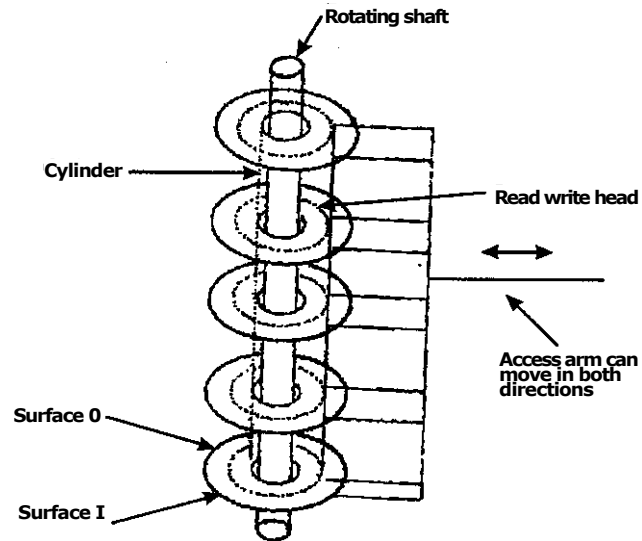


Fig. 8: The Disk Pack

In addition to these two times the time taken to read block of word can be considered but normally it is too small in comparison to latency and seek time and in general theaugas.tame.is considered to be sum of sedilime and latency time. Since access time to disks is large, therefore, it is advisable to read a sizeable portion of data in a single go and that is why the disks are referenced block wise.

3.2.3 Magnetic Tape

Magnetic tapes are mounted on reels or a cartridge or a cassette of tape to store large volumes or backup data. These are cheaper and since these are removable from the drive, they provide unlimited storage capacity.

Information retrieval from tapes is sequential and not random. These are not suitable for on-line retrieval of data, since sequential searching will take long time. These are convenient for archival storage, or for I-backup. The tapes are one of the earliest storage devices. They are low cost, low speed, portable and are still widely used because of their low cost.

3.2.4 Optical Memories

Optical memories are alternate mass storage devices with huge capacity. The advent of compact disk digital audio system, a non-erasable optical disk, paved the way for the development of a new low cost storage technology. In optical storage devices the information is written using a laser beam. These devices which are memories can store large amount of data. We will discuss here three optical memory devices, which are now becoming increasingly popular in various computer applications.

CD-ROM

The CD-ROM (Compact Disk Read-Only Memory) is a direct extension of audio CD. CD-ROM players are more rugged and have error-correction facility. This ensures proper data transfer from CD-ROM to the main memory of the computer. CD-ROM is written into during the process of manufacture by a high power laser beam. Information is retrieved from a CD-ROM using a low power laser, which ingenerate in an optical disk drive unit. In CD-ROMs the information is stored evenly across the disk in segments of the same size.

Therefore, in CD-ROMs data stored on a track increases as we go towards outer surface of disk. Thus, the CD-ROMs are rotated at variable speeds for the reading process.

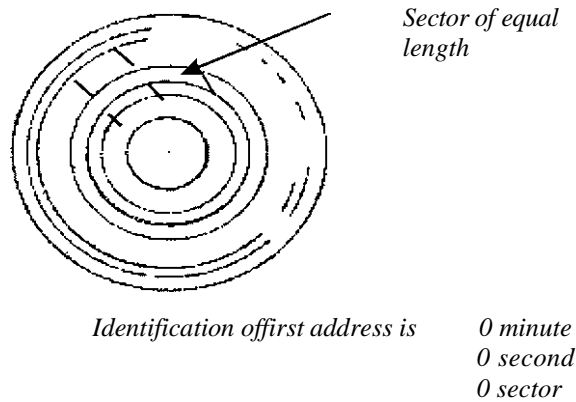


Fig. 7: A CD-ROM's disk layout

Figure 7 indicates the layout used for CD-ROMs. As discussed earlier, the data is stored sequentially along a spiral track. In this disk random access becomes more difficult locating a desired address involves first moving the head to the specific area then adjusting the rotation speed and then reading the address, and then to find and access the specific sector.

CD-ROMs are very good for distributing large amount of information or data to large number of users.

The three main advantages of CD-ROMs are:

- large data/information storage capacity
 - mass replication is inexpensive and fast
- these are removable disks, thus are suitable for archival storage

The disadvantages of CD-ROMs are:

- It is read-only, therefore, cannot be updated
- Access time is longer than that of a magnetic disk

WORM

In certain applications only few copies of compact disks are to be made which makes CD-ROMs production economically unviable. For such cases write-once, read-many CD has been developed. WORM disks are prepared in such a way that they can be written only once subsequently by a laser beam of modest intensity. The disk controller of WORM is more expensive than that of CD-ROM. WORM uses sector structures same as that of magnetic disks. High power laser first prepares the WORM disk. A CD writer can write them once.

33 High Speed Memories

The Need: Why the high speed memories? Is the main memory not a high-speed memory? The answer to second question is definitely "No", but why so, well for this we have to go to the fundamentals of semiconductor technology which is beyond the scope of the unit, if the memories are slower then how slow are they? It has been found that the access of main memories is slower than the speed of the processor. Since each instruction require several memory accesses therefore, faster memories will be of tremendous help in increasing the overall through out of a computer.

1. Magnetic Ink Character Recognition (MICR) devices are generally used by the banking industry to read the account numbers on cheques directly and does the necessary processing.
2. Optical Mark Recognition (OMR) devices can sense marks on computer readable papers. This kind of device is used by academic and testing institutions to grade aptitude tests where candidates mark the correct alternatives on a special sheet of paper. The optical mark recognition devices then directly read these answer sheets and the information sent to a computer for processing.
3. Optical Bar Code Reader (OBR) scans a set of vertical bars of different widths for specific data and are used to read tags and merchandise in stores, medical records, library books, etc. These are available as hand held devices.

4.2.2 Output Devices

The output normally can be produced in two ways: either on a display unit/device or on a paper. Other kinds of output such as speech output, mechanical output is also being used in Certain applications. In this section, we will discuss only the display and printing output devices.

Display Devices

One of the most important peripherals in computer is the graphic display device. Conventional computer display terminals known as alphanumeric terminals, display characters (images) from a multi-dot array (normally 5 x 7 or 7 x 9). These are used to read text information displayed on the screen. However, there are increasing demand for display of graphics, diagrams and pictures to make the visual presentation of information more effective for user interaction and decision making.

Graphic display is made up of a series of small 'pixels' (picture elements) whose pattern produces the image. Each dot on the screen is defined as a separate unit, which can be directly addressed. Since each dot can be controlled individually there is much greater flexibility in drawing pictures. There are three categories of display screen technology:

1. Cathode Ray Tube (CRT)
2. Liquid Crystal Display (LCD)
3. Projection Displays

CRT Displays

The main components of a CRT terminals are the electron gun, the electron beam controlled by an electromagnetic field and a phosphor coated display screen.

The electromagnetic field is used to create an image (refer to figure 5). There are two types of CRT displays

- (a) Vector CRT displays in which the electron beam directly hits the phosphor to be created.
- (b) Raster displays in which the image is projected on to the screen by directing the electron beam across each row of picture elements. This type of display provides a high dynamic capability since image is continuously refreshed and it allows for continual user input and output. It offers full colour display at a relatively low cost and is becoming very popular.

Liquid Crystal Displays (LCD)

First introduced in watches and clocks in 1970s, LCD is now applied to display terminals. The major advantage of LCD is the low energy consumption. The CRT (Cathode Ray Tube) is replaced by liquid crystal to produce the image. This also has colour capability.

These are commonly used

of compactness and low energy consumption.

Projection Displays

The personal size screen of the display is replaced by a screen upon which images are normally used in the group presentation. These systems can be connected to computer and whatever appears on the computer terminal gets enlarged and projected on a large screen. Another popular method is to connect computer to a LCD flat screen, and to project the LCD image using Overhead Projector. These are popularly used for seminars, classrooms, marketing presentations, etc.

Printers

Printers are used for producing output on a paper. There are a large variety of printing devices, which can be classified according to the printing speeds. Current estimates indicate that about 1500 types of printers are commercially available conforming to about 15 different printing technologies. The following categories of printers are identified.

(a) Printing Technology — impact printers vs non-impact printers

Impact printers use variations of standard typewriter printing mechanism where a hammer strikes paper. Impact printers use variations of standard typewriter printing mechanism through inked ribbon. Non-impact printer transmits electrical signals to etch or induce symbols on paper. Many of these require special coated or treated paper.

(b) Character forms — fully formed characters vs dot matrix

Fully formed characters are constructed from solid lines and curves like the characters of a typewriter whereas a dot matrix character is made up from a carefully selected cluster of dots to each other. Obviously print quality of a dot matrix printer will be poorer compared to that from fully formed characters.

(c) Printing sequence — serial vs page

This indicates the amount of information at

• • • • • input within a • • • • • Serial printing

is done character by character whereas line printing forms an entire line at a time. A page printer outputs a whole page of character and images simultaneously during one cycle.

Clearly the speed of output will depend upon the printing sequence incorporated in the printing device. We will now look at three of the most popular printers.

Dot Matrix Printers

This is one of the most popular printers used for personal computing systems. These printers are relatively cheaper compared to other technologies. This uses impact technology and a print head containing banks of wires moving at high speeds against inked ribbon and paper. Characters are produced in matrix format. The speeds range from 40 characters per second (CPS) to about 1,000 cps. A disadvantage of this technology is that the print quality is low.

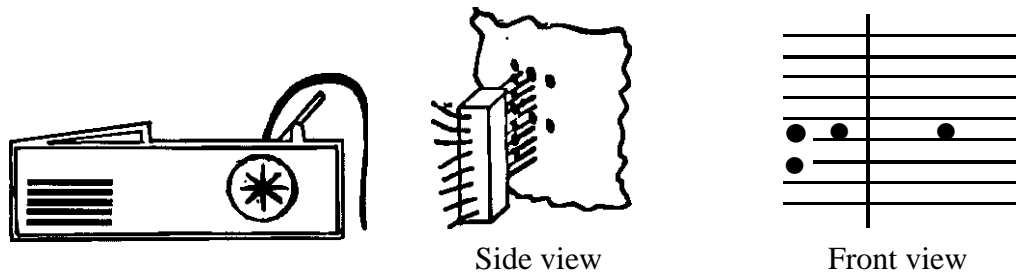


Fig. 8: Dot Matrix Printer

Ink Jet Printers

These print by spraying a controlled stream of tiny ink droplets-accurately on the paper forming either dot matrix or solid characters. These are non-impact and hence relatively silent and high quality printers. The typical speeds range from 50 cps to above 300 cps and this technology has been used well for production of colour printing and elaborate graphics.

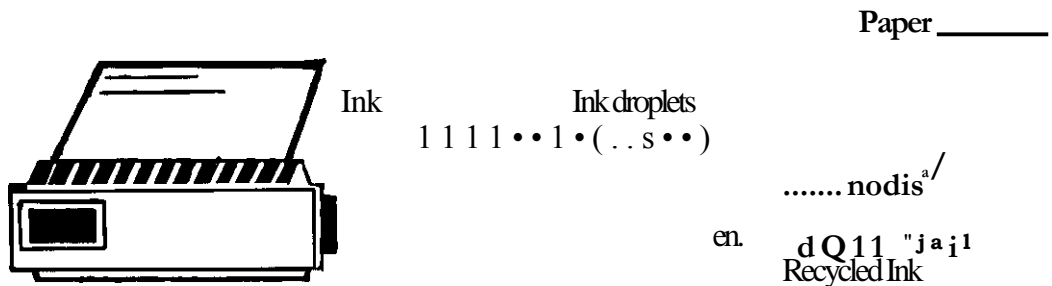


Fig. 9: Inkjet Printing

Laser Printers

This is a high quality, high speed, and gaithumeiechnology, which works in non-impact fashion on plain paper or pre-printed forms. Printing is achieved by deflectingl_a_es Meant ol_____Itp_the,photosensitive surface of a drum and the latent image attracts the toner to the image areas. The toner is then electrostatically transferred to the paper and fixed into a permanent image. Speeds can range from 10 pages a minute to about 200 pages per minute. This technology is relatively expensive but is becoming very popular because of the quality, speed and noiseless operations.

These three printers are compared in Table 2.

Plotters

To produce graphical output on paper the plotters are used.

After discussing so much about the I/O devices, let us come back to one of the basic question: How are I/O devices connected to computers? We will try to answer the question in the next section.

Style of Printing	Print Quality
Dot	Low quality but steadily improving. Very useful for high-volume work (low quality) because of great speeds and low printing cost. Lack of descenders on letters (many models rectify inability to reproduce well on a photocopier are major disadvantage) and the problems. Dot Matrix Printers have the ability to produce multiple copies of documents in a single print.
Inkjet	High quality print character sets: choice of type set is controlled by program; print quality quite good than dot matrix; cheap; useful for low volume applications; high printing cost;
Laser	Very high quality, prints at speeds unapproachable by other technologies.Price is affordable. Excellent for high volume high quality works.

Table 2: Some Printer Technologies

Exercise 4.1

State True or False

(a) In a case where graphical user interfaces are common mouse should not be used.

True ☐ False ☐

(b) Keyboard is one of the most common input device.

True ☐ False ☐

(c) Scanners are devices used for outputting pictures.

True ☐ False ☐

(d) Projection displays can be used for classroom teaching.

True ☐ False ☐

(e) Keyboard, VDU and printers are essential for computers.

True ☐ False ☐

4.3 Conclusion

By now you are in position to identify some of the input/output devices. You have also learned how to identify the set: and parallel interfaces. This unit is important for unit 5.

4.4 Summary

This unit introduces you to some important devices of a computer. Knowledge beyond what you require for this unit may be gained in the specified further reading.

4.5 Tutor marked Assignment (TMA)

1. Differentiate between the CRT display and LCD display
2. Compare and contrast the Laser and Dot Matrix Printers.
3. Define the term "Source Data Automation" Give two paxmple. •

4.6 Further Reading

William Stallings, *Computer Organization and Architecture*

1-k-etc-el cif et 4 "4-411' ref
if a et - (A-10-r t -c c'c'-4 J e-e3-0 Cit tfrfrt

CRC

0-t42, A-A-10-4-cfrO GO- arty -1)fi-e-th-1A bc9A-0-1, c,, -Lhcp
A t t i t e r 6 . - p - t r f ? : < -
c r % e - - " A I I I - C L / V r f r 6 5 8 6 , 0 2 , - i

UL

c •GLapH lett 0.4-P-A crono--1
cc O-re ructMt, 04-efil ec-Abk

d'e■I GS 0(1r

(A3 Lif 'Q S / Lai-84J -

1)51;J 01,C/\

cr`Vt prA-ttat4 CO- Wei

Q

(s1 for e-13.

crye

A/ As I- it-fG

Cs—e-Te VOL() oil

(um/ et 4,) 4

It 6

/Ait--CL-Cc_a_13:21- .1 Ct

017144

C/sre (1-'44

L O Q 3 pri-c t

1,1 Q),j ccro •

tirti AA.A-e)

/4,36 Seed Get-4

k1,Ati.2e4 CI' bbα4X fur e'rl

9314(63Le 9,61

Iki-frc(1611-1-

Unit 5

Interfaces, Parallel Processing, and Vector Processing, Pipelining and Introduction to RISC

5.0 Introduction

In this unit we will discuss more about the input/output organization and newer technologies in computer hardware.

5.1 Objectives

At the end of this unit you will be able to:

- define input/output module interface
- define an external interface
- define parallel processing
- define pipelining
- distinguish vector processing

5.2 Input/Output Module Interface

The input/output module interfaces (I/O module) is normally connected to the computer system on one end and one or more input/output devices on the other. Normally, an I/O module can control one or more peripheral devices. An I/O module is needed because of:

- (a) Diversity and variety of I/O devices makes it difficult to incorporate all the peripheral device logic (i.e. its control commands, data format etc.) into CPU. This in turn will also reduce flexibility of using any new development.
- (b) The I/O devices are normally—slower than that of memory and CPU, therefore, it is suggested not to use them on high speed bus directly for communication purpose.
- (c) The data format and word length used by the peripheral may quite be different from that of a CPU. An I/O module is mediator between fl_sp_Locessonid I/O device/devices. Its onicalsibe-clata-exchange be eel_ j_the_externaisleyices_animain_memory_or_exatices_and_CPURegisters. Therefore, an I/O module an interface internal to the computer which connect it to CPU and main memory and an interface external to the computer connecting it to external device or peripheral. The I/O module should not only communicate the information from CPU to VO device, but it should also co-ordinate these two.

1/1g:

CO-04R Pale

t.%¹ b/W29 Pct_{iX} m_pa° ffr,

cin^r a_r^A, t_c, cfr c_sflt 2e-e⁻R2, ix eft teA-c—v
F P - F I R - P - a o c d e L i t c k a k a
(1/4IVc Cil d C' (A ilo

In addition, since there are speed differences between CPU and I/O devices, the I/O module should have facilities like buffer (storage area) and error detection mechanism.

If an I/O module takes more processing burden then it is termed as I/O channel or processor. The primitive I/O modules, which require detailed control by processor, are called I/O controller or device controller.

These I/O controllers are normally used in micro-computers, while I/O processors are mainly used for Mainframe, because the VU work for microcomputer is normally limited to single user's job. Therefore, we do not expect a very 'huge amount of I/O to justify the investment in I/O processor, which are expensive.

53 External Interfaces

Our discussion on I/O system will not be complete if we do not discuss about external interfaces. External interface is the interface between the I/O module and the peripheral devices. This interface can be characterized into two main categories:

- (a) parallel interface
- (b) serial interface

Figure 10 gives a block diagram from these.

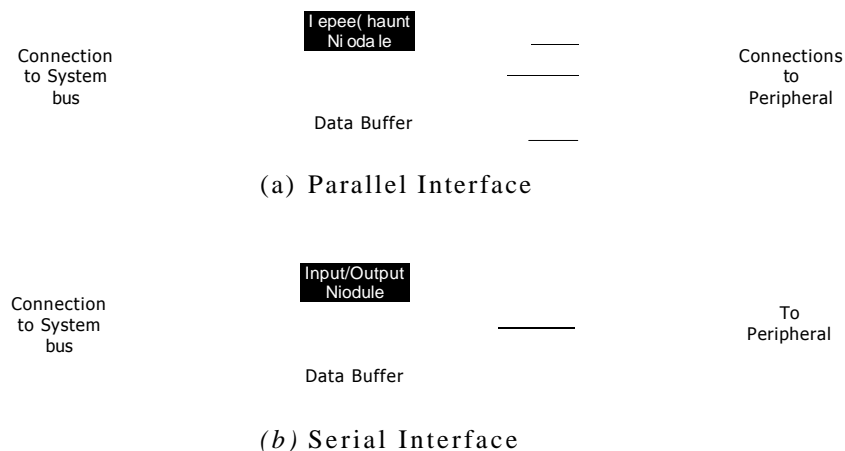


Figure 10: Parallel and Serial Interface

In parallel interface multiple bits can be transferred simultaneously. The parallel interfaces are normally used for high-speed peripherals such as tapes and disks. The dialogue that takes place across the interface includes the exchange of control information and data. A common parallel interface is centronics.

In serial interface only one line, is used to transmit data, therefore, only one bit is transferred at a time. Serial interface is used for serial printers and terminals. The most common serial interface is RS-232-C.

Exercise

State True or False

- (a) Devices are normally connected directly to system bus.

True ☐ False ☒

- (b) Input/output Module is needed only for slower I/O devices.

True ☐ False ☒

urre:(2---t-- 5

(tf\--9'

(1--19-64

n410

inct,,,,menoni

5beOO

bdo 3 Oinla."-- .x bit c

r

la4-C ifo detlic'csi Yer-A drat a-c easoc at e

a-kt Li-cle't4etchb

6⁶ p₂-1 12 4r-e-r ri-enod t ecCero-cege-i-f Pteetiej incre.A4Zi

d Feetep

tc?' cesg⁻ci₆

(c) Data buffering is helpful for smoothing out the speed difference between CPU and input/output devices.

True ☒ False ☐

(d) Parallel interfaces are commonly used for connecting printers to a computer.

True ☐ False ☒

What is a device controller?

An I/O module that req. detailed control by a processor. They are called I/O controller or device controller. They are commonly used in micro-computers.

54 What is Parallel Processing?

In the previous section, we have defined the input/output interfaces of a computer. Let us now, explore the newer technologies and trends in computers.

Computers can be made faster by increasing their clock speed or by using improved VLSI technology. But there is a practical limitation to both of them. The only way, that seems possible, theoretically that can increase the speed in an unlimited fashion is to have multiple modules, that is break the original problems into independent modules and execute them simultaneously. The more the number of processors doing the job simultaneously the more the speed will be.

Parallel processing is a term used to denote a large class of techniques that are used to provide simultaneous data processing tasks for the purpose of increasing the computational speed of a computer.

One of the most popular classifications of parallel computer is M. J. Flynn's classification and is based on the multiplicity of instruction streams and data streams in computer system. The sequence of instruction read from the memory constitute the instruction stream, and the data they operate on in the processor constitute the data stream. Flynn's classification divides the computers into four categories:

- single instruction single data (SISD)
- single instruction multiple data (SIMD)
- multiple instruction single data (MISD)
- multiple instruction multiple data (MIMD)

SISD organization is available in most serial computers today. Instructions are executed sequentially, though may be overlapped in their execution stage (pipelining). All the functional units are under the supervision of one control unit. The Von Neumann architecture falls under this category.

SIMD organization consists of multiple processing elements supervised by the same control unit. All processing units receive the same instruction broadcast but works on different data streams. One of the very common example is the execution of a 'for' loop in which the same set of instructions are executed for, may be a different set of data.

MISD organization implies interactions between N processors because all memory streams are derived from the same data stream shared by all processors. If the interaction between the processors is high, it is called a tightly coupled system, or else it is called a loosely coupled system. Most multiprocessors fit into this category.

Parallel processors can be categorized into several categories. These include:

- Array Processors
- Distributed Architecture
- Multiprocessors

—re) in Cre-o-St -itz en> ceio--r 201 tiliar 4 '---i pc.

in (re"-it -rt t cioci-c rtk'

t i s k v p a v e d t v i L s 1 4 - e - a - n - A t i i t v e t ' i 1 - 7 1 - e - C ' 4 4 - i n i t r a l - i t i - e c - A n c l
St itl' uttaneseresatje wed. te 'Lek 3i n4 4 4 - a 1 4 : E n t a l l % - t i 0 1 6 1
A c r e g g . r ' h i t e l ' f i e s * f c t i f i t i f t i k i z i
P A n a k e F r . r r o c e s s ' e t c - c / c i a l t e - C -

- Data flow Architecture

Array processors are parallel architectures which deal with repetitive operations. These are an example of SIMD architecture with applications varying from mathematical array operations and in structure in which data objects are known well in advance.

Distributed Architecture are composed of re.t_h_Lely_autoll3-s us_which are capable of handling complete system and execution functions, and which cooperate together to run a large application.

Multiprocessors are architectures composed ann__S,cs,ainputi.cfg-tanits, which operate in a synchronous mode. Generally, both shared and local memory can be available for these processors. The communication between the processors takes place either through the shared memory area or through the inter processor messages.

Dataflow Architectures are functionally distributed architectures, in which the operations are triggered with the arrival of data at these processors. They may be viewed as very general MIMD parallel architectures.

Most of these parallel processor computers follow an important concept called pipelining.

In the subsequent sections, we will discuss some of the newer technological concepts.

5.5 Pipelining

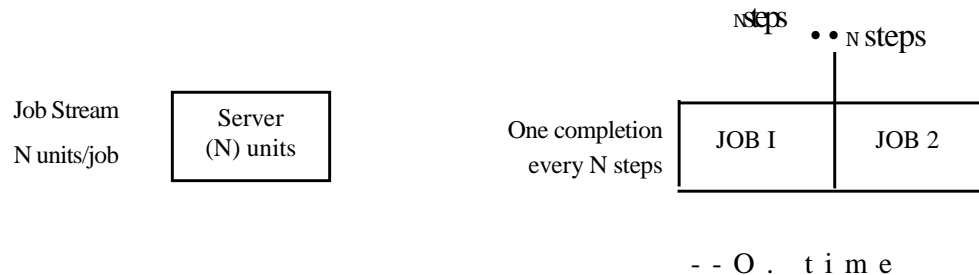
The basic idea behind pipeline design is quite natural; it is not specific to computer technology. In fact the name pipeline stems from the analogy with petroleum pipelines in which a sequence of hydrocarbon products iiPiUmped through a pipeline. The last product might well be entering the pipeline before the first product

een removed from the terminus. The key contribution of pipelining is that it provides a way to start a new ask before an old one has been completed. The concept can be better understood by taking an example of an industrial plant. To achieve pipelining, one must subdivide the input task into a sequence of subtasks, each of high can be executed by a specialized hardware stage that operates concurrently with other stages in the hardware. Successive tasks are streamed into the pipe and get executed in an overlapped fashion at the subtask level. Hence the completion rate is not a function of total processing time, but rather of how soon a new process can be introduced.

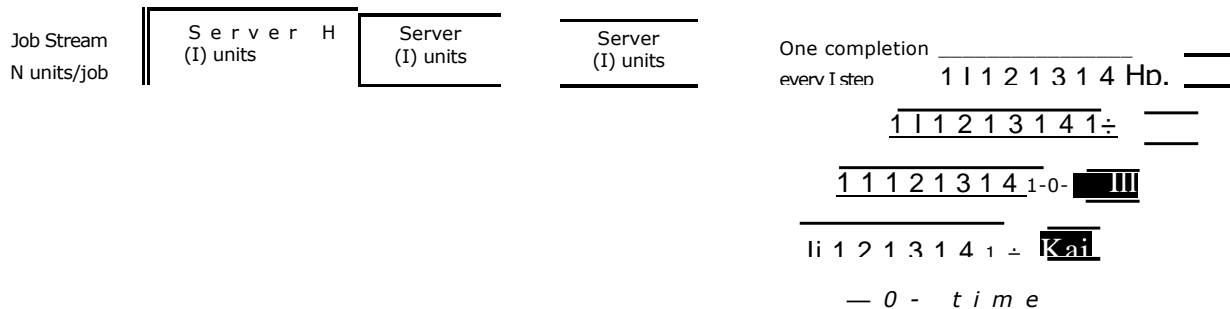
The subdivision of labour in assembly lines has contributed to the success of mass production in modern industry. By the same token, pipeline processing has led to the tremendous improvement of system throughput in modern digital computers.

Now consider for execution of an instruction a computer performs N processing steps. A server unit as represented in figure 5(a) can perform any one of the N steps in one unit time. If we process M instructions by this server, then the rate of completion is one instruction of every N steps, and the time behaviour of the job stream is as described in figure 5(a). Compare figure 5(a) with figure 5(b) which shows N servers concatenated in a sequence each performing only a single step, the job flows through the collection of servers by visiting server 1, then server 2, and soon and finally emerging from server N after N steps. The time behaviour of this system is described by figure 5(b).

Figure 5(b) is an ideal model assembly line, such as an automobile assembly plant.



(a) Sequential execution with Nunit server



(b) Pipelined execution with 1-unit servers

Fig II: Two ways of executing N-unit jobs in a stream

Thus, any arithmetic task is divided into steps. Thus, a sort of concurrent operation can be obtained in the pipeline. This will help in increasing the output of the system.

Though, superficially it seems that it should be possible to execute any application in various pipeline stages, it is not actually so. Only those applications, which can be broken into independent subtasks, can take the advantage of pipelining, it is most efficient for those applications that need to repeat the same task many times with different sets of data.

56 Vector Processing

Some of the applications such as:

- weather forecasting;
- petroleum explorations;
- seismic data analysis;
- flight simulations;
- artificial intelligence and expert systems;
- image processing and its applications;

are so complex that without the use of sophisticated computers it would not be possible to complete these in reasonable amount of time.

These computational problems are beyond the capabilities of a conventional computer, as these problems require a vast number of computations that may take a conventional computer days or even weeks to be completed. The vector computers have emerged as the most important high performance architecture for numerical problems. it has two key qualities of efficiency and wide applicability.

These scientific problems may involve manipulations of large arrays, where the same computation% are carried out with different sets of data. These numbers can be easily formulated as vectors and matrices of floating point number (real numbers) vector an be defined as an ordered .gt of a one-dimensional array of data items. Each element in a vector is a scalar quantity, which may be a floating point number, an integer, a logical value, or a character (byte).

Supercomputers

A commercial computer with vector instructions and pipelined floating point arithmetic operations is referred to as a supercomputer. Supercomputers are very powerful, and high performance machines used mostly for scientific computations. To speed-up the operation, the components are packed tightly together to minimise the distance that the electronic signals have to travel.

A supercomputer is a computer best known for its high computational speed, fast and larger memory systems, and extensively used for parallel processing. It is equipped with multiple functional unit each unit has its own pipeline configuration. Although the supercomputer is capable of general purpose applications

- Decrease the memory access time, use a faster but expensive technology for main memory probably it will be feasible after few years.
- Access more words in a single memory access cycle. That is instead of accessing one word from the memory in a memory access cycle, access more words. This is termed as memory interleaving.
- Insert a high speed memory termed as Cache between the main memory and processor.
- Use associative addressing in place of random access.

Cache Memory

The time taken with cache	(100+20) For the first Read operation	+20 x 3 for last three read operation
	120+6	180nos
Time taken without cache	100 x 4	400 nos

1. Cache on the edge of the cache.
 2. Cache on the edge of the cache.
 3. Cache on the edge of the cache.
 4. Cache on the edge of the cache.
 5. Cache on the edge of the cache.
 6. Cache on the edge of the cache.
 7. Cache on the edge of the cache.
 8. Cache on the edge of the cache.
 9. Cache on the edge of the cache.
 10. Cache on the edge of the cache.

Exercise 3.1

State True or False

- (a) High speed memories are needed to bridge the gap between I/O device and memory.

True

False

☒

- (c) Cache memory increases load on main memory.

True

☐

False

☒

- (c) The principle of locality says that all the references to data have to be to the same memory location.

True

☐

False

☐

3.4 Conclusion

We have introduced you to magnetic disks and winchester disks. You know much about magnetic tapes and optical memories. This unit completes our discussion on the introductory concepts of computers.

3.5 Summary

In this unit you learned key characteristics of memories and the technologies which are used for constructing these memories.

3.6 Tutor-Marked Assignment

1. Compare and contrast RAM and ROM.
2. What is the importance of read mostly memories?
3. What is the head of disk?
4. Match the following pairs
 - (i) Variable rotation speed
 - (ii) Low cost, low speed devices
 - (iii) Double sided double density
 - (a) Magnetic tape
 - (b) Floppy disks
 - (c) CD-ROM

3.7 Further Reading

W. Stallings, *Computer Organization and Architecture*, Third Edition, Maxwell Macmillan International Edition.

M. M. Mano, *Computer System Architecture and Organization*, Second Edition. McGraw Hill international Editions. 1988.

Unit 4

Input/Output Organization and New Technologies

4.0 Introduction

Till now we have discussed about various components and memory system for a computer. We have also introduced one interconnection structure called the Bus. In this unit we will discuss more about the input/output organization and newer technologies computer hardware. We will discuss briefly about input/output devices, two popular device interfaces and the concepts of parallel organization. We will also discuss about the concepts of pipelining, vector processing and the concepts relating to Reduced Instruction Set Computers (RISC) architecture.

4.1 Objectives

At the end of this unit you should be able to:

- identify some of the input/output devices
- identify the serial and parallel interfaces
- define the Flynn's classification of computers
- list the characteristics of multiprocessor systems
- define pipelining and vector processing
- define the concepts relating to RISC.

4.2 Input/Output Devices (Peripherals)

The computer will be of no use if it is not communicating with the external world. Thus, a computer must have a system to receive information from outside world and must be able to communicate results to external world. Thus, a computer consists of an I/O system. This system includes two basic components, one is the I/O devices and another called I/O module, which not only connects an I/O device with the system bus, but plays a very crucial role in between. A device, which is connected to an I/O module of computer, is called a peripheral. Let us discuss few input/output devices in this section and we will define the I/O module in the

subsequent sections.

K

4.2.1 Input Devices --- 3"1-• t A r - L i

Input technologies are rapidly developing. These, as the name suggests, are used for transferring user command or choice or data to the computer.

Penphone - ÷ o-re device

illare owl/meted ft
23

Alecialt cr-f Pc.

Keyboard

The keyboard is one of the most common input devices for computers. The layout of the keyboard is like that of the traditional WARTY typewriter, although there are some extra commands and function keys provided for. Substantial development has taken place in the ergonomics of keyboard design to ensure that operator strain is minimal.

Pointing Devices

The keyboard provides facility to input data and commands to computer in text form. We find that, while working with a display based packages, we are mostly pointing to some area in the display to select an option and move across on the screen to select subsequent options. For such cases pointing devices are very useful. There are several pointing devices, some of them are:

- (a) Mouse: Mouse is a handy device which, can be moved on a smooth surface to simulate the movement of cursor that is desired on the display screen. Mouse could be optical; offering quite and reliable operation, or mechanical. which is cheaper but noisier. User can move the mouse, stop it at a point where the pointer is to be located and, with the help of buttons, make selection of choices.
- (b) Light Pen: This is a pen shaped device allowing natural movement on the screen. The pen contains the light receptor and is activated by pressing the pen against the display screen. Receptor is the scanning beam which helps in locating the pen's position. Suitable system software is provided to initiate necessary action when we locate an area on the display surface with the help of the light pen.

There are a few other pointing devices known as track balls and joysticks, which are used more on entertainment like games. We will not discuss them in this section.

Voice/Speech Input

One of the most exciting areas of research is in recognizing human voices/speech so that this could be inputted into computer directly. This approach will eliminate the need for keying in data and will facilitate casual users to use the computer very easily. There are several problem areas for research since speech recognition system should be able to identify who is speaking and what the message is. Voice recognition techniques along with several other techniques to convert the voice signals to appropriate words and derive the correct meaning of words are required for a commercially viable comprehensive speech recognition system. We have found limited success in this area and today devices are available commercially to recognize and interpret human voices within limited scope of operation.

Scanners

Scanners facilitate capturing of the information and storing them in a graphic format for displaying back on the graphical screen. Scanner consists of two components, the first one to illuminate the page so that the optical image can be captured and the other to convert the optical image into digital format for storage by computer.

The graphic image scanned can now be seen and processed directly by the computer. Substantial research work is going to establish methods by which the scanned image can be automatically converted into equivalent text for further processing.

Source Data Automation

Most recent trends for data input are towards source data automation. The equipment used for source data automation capture data as a by-product of a. business activity thereby completely eliminating manual input of data. Some examples are:

found in all other computers, it is specifically optimized for the type of numerical calculations involving vectors and matrices of floating point numbers.

A measure used to evaluate computers in their ability to perform a given number of floating point operations per second is referred to as flops. The term mega flop is used to denote million flops and giga flops to denote billion flops. A typical supercomputer has to a basic cycle time of 4 to 20 nsecs. If the processor can calculate a floating-point operation through a pipeline each cycle time, it will have the ability to perform 50 to 250 mflops. This would be sustained from the time the first answer is and does not include the initial setup time of

(C1 1115024
vft

5.7 Introduction to RISC

Aim of the computer architect is to design computers which are more cost effective than their predecessors. Cost effectiveness includes

1. cost of hardware to manufacture the machine
2. cost of programming
3. cost incurred in correcting both the initial hardware and subsequent programs on account of the new/old compatible architecture.

5.7.1 Reasons for Increased Complexity

Let us see what are the reasons for increased complexity that leads to complex instruction set computers (CISC) and what exactly we mean by this.

1. *Speed of memory vs the Speed of CPU*

The CPU, in general, was about 10 times as fast as main memory. Thus, more and more instructions were put into hardware. Since then many higher level instructions have been added to machines in an attempt to improve performance. This trend began towards the imbalance in speed.

The advances in semiconductor memory have made several changes to the assumptions about the relative difference in speed between CPU and the main memory. Semiconductor memories are both fast and relatively inexpensive. Cache memories have further reduced the difference between the CPU and the memory speed.

2. *Code density*

With early computers memory was very expensive, it was therefore cost effective to have very compact programs. It was assumed that more instructions would result in small programs. While code compaction is important, however, the cost of 10 percent more memory is often far cheaper than the cost of squeezing code by 10 percent out of the CPU architecture innovations.

3. *Support for high level language*

With the increasing use of more and more high level languages, manufacturers have become eager to provide more powerful instructions would result in small programs. While code compaction is important, however, the cost of 10 percent more memory is often far cheaper than the cost of squeezing code by 10 percent out of the CPU architecture innovations.

Their research shows that integer constants appeared almost as frequently as arrays or structures; more than 80 percent of the scalars were local variables to subroutines; more than 90 percent of the arrays or structure were global variables.

In addition it was found that procedure calls and returns are an important aspect of high level language programs. These are also the most time consuming operations in compiled high level language programs.

Thus, it will be profitable to consider ways of implementing these operations efficiently. Two aspects are significant in this respect; the number of parameters and variables that procedure dealt with; and the depth of

LC4Lvippg Lcir— co-rk c-rxs-1/4 cift a4(f- pa_rf ci e7s7rn,_,(0,gmit_s
ck &to-, ru't-ti 34 di-9-6P 74-
gLocsig-c ./.(4al et — vaPib-Ozs 0.) 1., poi-75774

nesting.

Further, it has been statistically found that 98% of dynamically called procedures pass lower than six arguments, and 92% of them use fewer than six scalar variables.

Thus, in the 1980s, a new philosophy evolved called RISC. These machines would have fewer instructions (reduced set) and the instructions would be simple and would generally execute in one cycle, hence, the reduced instruction set computers (RISCs). Optimising compilers are used in these machines to compile the programs into simple machine instructions.

5.7.2 Principles of RISC

The following are the characteristics of RISC architecture:

1. Hardware instructions should be kept simple unless there is a very good reason to do otherwise.
2. Microcode, a programming way of implementing control unit, is not magic. Moving software into macrocode does not make it better, it just makes it hard to change.
3. Simple decoding and pipelined execution are more important than program size. At the peak rate a new instruction is started every cycle. Pipelined execution achieves a peak performance of one.

Laser printer	Dot Matrix printer (DMP)
(i) High cost	Low cost
(ii) Form characters using laser beam	Use dot matrix to print characters
(iii) Non impact printer	Impact printer
(iv) High quality output	Output is of low quality
(v) Noiseless operation	Noisy
(vi)	Produces multiple copies in a single print.

4. In source data automation data is entered essentially where it is being produced through automatic devices. Examples are OMR & OBR.

Exercise 5.1

Write True or False

- (a) Devices are normally connected directly to system bus.

True ☐ False ☒

- (c) Input/Output Module is needed only for slower I/O devices

True ☐ False ☐

- (e) Data buffering is helpful for smoothing out the speed difference between the CPU and Input/Output devices.

True ☐ False ☐

- (c) Parallel interfaces are commonly used for connecting printer to a computer

True ☐ False ☒

What is a device controller?

.....

.....

.....

5.8 Conclusion

You learned input/output devices, two popular device interfaces and the concept of parallel organization. You learned concepts of pipelining and vector processing.

5.9 Summary

The unit introduced you to input/output devices and important concepts of parallel organization, pipelining and vector processing.

5.10 TutorMarked Assignment (TMA)

1. Explain the term "Input/Output module interfaces"
2. Compare and contrast the parallel interface and a serial interface
3. Explain the terms
 - (i) Pipelining
 - (ii) Vector processing
 - (iii) Parallel processing and
 - (iv) RISC architecture

5.11 Further Reading

M. Moris Mano, *Computer System Architecture and Organization*, Second Edition, McGraw-Hill International Editions (1988).

Module 2

Computer Fundamentals: Software

Introduction to Module 2

This module covers aspects relating to computer software such as components of programming languages and operating system concepts.

Unit 1

Software Concepts and Terminology

1.0 Introduction

A computer contains two basic parts: (i) Hardware and (ii) Software. In the first five units we touched hardware issues in detail. In this unit and also in the rest of the units of this module we will discuss topics related to software. Without software a computer will remain just a metal. With software, a computer can store, retrieve, solve different types of problems, create friendly environment for software development etc.

The process of software development is called programming. To do programming one should have knowledge of (i) a particular programming language, (ii) set of procedures (algorithm) to solve a problem or develop software.

The development of an algorithm is basic to computer programming and is an important part of computer science studies. Developing a computer program is a detailed process, which requires serious thought, careful planning and accuracy. It is a challenging and an exacting task, drawing on the creativity of the programmer. Once an algorithm is obtained, the next step for a solution using a computer would be to program the algorithm using mathematical and data processing techniques. Programming languages constitute the vehicle for this stage of problem solving. The development of programming languages is one of the finest intellectual achievements in Computer Science. To use a computer, it is necessary to understand a programming language. Understanding them does not really mean only being able to use them. A lot of people can use them without really fully understanding them.

An Operating System is system software, which may be viewed as an organized collection of procedures for operating a computer and providing an environment for execution of programs. It acts as an interface between users and the hardware of a computer system. There are many important reasons for studying operating systems. Some of them are:

- (1) User interacts with the computer through operating system in order to accomplish his task since it is his primary interface with a computer.
- (2) It helps users to understand the inner functions of a computer very closely.
- (3) Many concepts and techniques found in operating system have general applicability in other applications.

In this unit, we will discuss about the concepts relating to a programming language and in the next unit we will deal with the operating system concepts.

SC 964-7-1 --t cci--74-44--i .4_ atthehc_
firt-ockit)--0v.9 oir et pc 94—Gv:at ern- ;rte7x`rtt4 1/4 1 -
roi-c-entts
as
C!; r~t .?

11 Objectives

After completion of this unit, you should be able to:

- define the terms system and application software

12 Computer Software

Computer software consists of sets of instructions that mould the raw arithmetic and logical abilities of the computer units to perform.

In order to communicate with each other, we use natural languages like English, Bengali, Tamil, Marathi, Gujarati etc. In the same way programming languages of one type or another are used in order to communicate instructions/commands to a computer for solving problems. Learning programming language requires learning the symbols, words, and rules of the language.

Program and Programming A computer can neither think nor make any judgements on its own. Also it is impossible for any computer to independently analyze a given data and follow its own method of solution. It needs a program to tell it what to do. A program is a set of instructions that are arranged in a sequence that guides the computer to solve a problem.

The process of writing a program is called Programming. Programming is a critical step in data processing. If the system is not correctly programmed, it delivers incorrect results that cannot be used. There are two ways in which we can require program. One is to purchase an existing program, which is normally referred to as packaged software the other is to write a program from the scratch in which case it is called customized software.

A computer software can be broadly classified into two categories:- System Software and Application Software.

Today, there are many languages available for developing programs software. These languages are designed keeping in mind some specific areas of applications. Thus, some of the languages may be good for writing system programs/software while some other for application software. Since a computer can be used for writing various types of application system software, there are different programming languages.

- (1) System Programming Languages: System Programs are designed to make the computer easier to use. An example of system software is an operating system, which consists of many other programs, managing input/output devices, memory, processor &c. To write an operating system, the programmer needs instructions to control the computer's circuitry (hardware part). For example, instructions that move data from one location of storage to a register of the processor. C and C++ languages are widely used to develop system software.
- (2) Application Programming Language/Application programs are designed for specific applications, such as payroll processing, inventory control etc. To write programs for payroll processing or other applications, the programmer does not need to control the basic circuitry of a computer. Instead the programmer needs instructions that make it easy to input data, produce output, do calculations, store and retrieve data. Programming languages that are suitable for such application programs support these instructions but not necessarily the types of instructions needed for development of system programs.

There are two main categories of application programs: business programs and scientific application programs. Most programming languages are designed to be good for one category of applications but not necessarily for the other, although there are some general purpose languages that support both types. Business applications are characterized by processing of large inputs and large outputs, high volume data storage and retrieval but call for simple calculations. Languages, which are suitable for business program development, must support high volume input, output and storage but do not need to support complex calculations. On the other hand, programming languages that are designed for writing scientific programs contain very powerful instructions for calculations but rather poor instructions for input, output etc. Amongst traditionally used

System
Date: 15th April 2021
Page: 133
Page No: 461.4:5
Page No: 441-442

programming languages, COBOL (Commercial Business Oriented Programming Language) is more suitable for business applications whereas FORTRAN (Formula Translation Language) is more suitable for scientific applications. Before we discuss more about languages, let us briefly look at the categories of software viz, system and application software.

1.2.1 System Software

Language Translator

A language translator is a system software which translates a computer program written by a user into a machine understandable form. We will discuss more about them in the next section.

Operating System

An Operating System (OS) is the most important system software and is a must to operate a computer system. An operating system manages a computer's resources very effectively takes care of scheduling multiple jobs for execution and manages the flow of data and instructions between the input/output units and the main memory. Advances in the field of computer hardware have also helped in the development of more efficient operating systems. More details on operating systems are given in unit 4 of this module.

Utilities

Utilities are often used to perform small tasks. A few examples are: Sort/Merge utilities which are used for sorting large volumes of data and merging them into a single sorted list, formatting etc.

1.2.2 Application Software

Application software is used to enable the computer to solve a specific data processing task. A number of powerful application software packages, which does not require significant programming-knowledge, have been developed. These are easy to learn and use as compared to the programming languages. Although these packages can perform any general and special functions, there are applications where these packages are not found adequate. In such cases, application program is written to meet the exact requirements. A user application program may be written using one of these packages or a programming language. The most important categories of software packages available are:

- Data Base Management Software
- Spreadsheet Software
- Word Processing Desktop Publishing (DTP) and presentation software
- Graphics Software
- Data Communication Software
- Statistical and Operational Research Software.

Data Base Management Software

Databases are very useful for storing and maintaining queries databases and reports. Many of today's Database Management System are Relational Database Management Systems. Many RDBMS packages provide smart assistants for creation of simple databases for invoices, orders and contact lists. Many database management systems are available in the market these days. You can select any one based on your needs, for example, if you have only few databases then package like database, FoxPro etc may be good.

If you require some additional features and moderate work load then Lotus Approach, Microsoft Access are all right. However, if you are having high end database requirements which requires multi-user environment and data security, access right very good user interface etc. then you must go for professional RDBMS package like Ingress, Oracle, Integra etc.

Accounting Package

The accounting packages are one of the most important packages for an office. Some of the features, which you may be looking on an accounting, may be:

- tax planner facility
- facility for producing charts and graphs
- finding accounts payable
- simple inventory control facility
- payroll functions
- on-line connection to stock quotes
- creation of invoices easily

One of the good packages in this connection is Quicken for windows.

Communication Package

The communication software includes software for fax. The fax-software market is growing up. Important fax software is Delrina's WinFax PRO 4.0. Some of the features such as Remote Retrieval and Fax Mailbox should be looked into under fax software. These features ensure that irrespective of your location you will receive the fax message. Another important feature is Fax Broadcast. This allows you to send out huge numbers of faxes without tying up your fax machine all day.

If you have to transfer files from your notebook computer to a desktop computer constantly then you need a software program that coordinates and updates documents. On such software is Lap Link for Windows. This software offers very convenient to use features. For example, by simply dragging and dropping a file enables file transfer. This software can work if a serial cable or a Novell network or a modem connects you.

Desktop Publishing Packages

Desktop Publishing Packages are yet popular in Indian context. Newer publishing packages also provide certain in-built formats such as brochures, newsletters, flyers etc which can be used directly. Already created text can be easily put in these packages, so are the graphics placements. Many DTP packages for English and languages other than English are available. Microsoft Publisher, PageMaker, Corel Ventura are few popular names. Desktop Publishing Packages in general, are better equipped in Apple-Macintosh computers.

Information Providers

One of the very interesting information provider which will become popular in India also is Automap Road Atlas by Microsoft. This package may provide city-to-city driving instructions and maps. You may also get the best route, calculate the time it will take.

Many information provided are the Internet access programs. Today, the Internet access packages comes as a part of operating system however, many other packages can be used for accessing information on the World Wide Web. One very simple to use popular tool of browsing is Netscape.

Organizers Contact Managers, PIMs

Some of the tasks of an office manager can be able to:

- track contacts
- balance schedules
- manage projects, and
- prioritize tasks.

These things can be easily done using organizers programs, which have a phone book model for maintaining lists of contacts. They also have a calendar for entering appointments and to-dos. Some of these packages

are Okna's DeskTop Set for Windows, Lotus Organizer, Microsoft Outlook, etc.

If you are interested in knowing more, then only names and addresses about your contacts such as details like the industry they are working with, the product they are manufacturing, their business with you last year, when you last spoke to them etc, then you must look to a contact management software. One such software is "Symantec Act for Windows".

If you want to go even further then you can look for a Personal Information Manager (PIM). PIM is a tool that stores virtually any information such as reference materials, project details etc. The PIM document contains outlines, folders and links. Most of the data in the PIMs is presented as an outline, for example, the clients may represent the top level, the date of an appointment with him at the next level, and the details of the meeting indented further below. This item can be linked to any other, allowing data to be entered only once and linked up to all other appropriate places.

Suites

Suites are a set of programs sold as a group package mainly for the business user. The suite package includes programs for Word Processing, Electronic Spreadsheet, Databases, and Presentation Graphics Software and may be a Mail Software. For example, Microsoft Office Professional for Windows includes programs as Microsoft Word, Microsoft Excel and Microsoft Access, and a license for Microsoft Mail etc. The word processing, spreadsheet, and presentation graphics software interfaces in a suite are well-integrated allowing easy data transfer among these applications. Today there is a growing family of office compatible products, which will be included in suites.

In the fast developing software era the list discussed above cannot be complete. Please refer to latest PC journals for most recent software trends.

Exercise 1.1

1 What is computer software?

1.6 4444411/24.-thir Cete-2 fnalbin-s oc 'gat leiti
He computer wi ho o-re atiAti
ore used to Go chl) 4(-9 rcunpv-s c chancet/fe<

13 Conclusion

Here you are introduced to topics related to software. It was also explained that the process of software development is called programming.

1.4 Stunmary

The topics which are covered here include system software, application software.

1.5 Tutormarked Assignment (TMA)

I Explain the following terms

- (a) Operating systems
- (b) Database management software

1.6 Further Reading

R. J. Dromey. *How to Solve it by Computers*, Prentice Hall of India.
Pratt, *Programming Languages*, Prentice Hall of India.

Unit 2

Categories of Languages

2.0 Introduction

In this unit you will be introduced to categories of languages.

21 Objectives

After completing this unit you should be able to:

- define various categories of language
- define the terms such as computers interpreters, fourth generation languages

22 Categories of Languages

We can choose any language for writing a program according to the need. But a computer executes programs only after they are represented internally in binary form (sequences of 1s and 0s). Programs written in any other language must be translated to the binary representation of the instruction before the computer can execute them.

Programs written for a computer may be in one of the following categories of languages:

2.2.1 Machine Language

1 This is a sequence of instructions written in the form of binary numbers consisting of 1s, 0s to which the
 2 computer responds directly. The machine language was initially referred to as code, although now the term
 3 code is used more broadly to refer to any program text.

An instruction prepared in any machine language will have at least two parts. The first ar.anmat lid operation, which tells the computer what functions is to be performed. All computers have an operation code for each of it functions aglaconiipart of the instruction is the operand or it tells the computer when to find or store the data that has to be manipulated.

Just as hardware is classified into generations based on technology, computer languages also have a generation classification based on the level of interaction with the machine. Machine language is considered to be the first fir neration language.

Advantage of Machine Language

It is faster in execution since the computer directly starts executing it.

t - c c, „ 4 4, - p a cf->i
(Aaryst s t
L—t-rt fry- arn **b**_e
cr **cr che** en) rheezt3

Disadvantage of Machine Language

It is difficult to understand and develop a program using machine language. Anybody going through this program for checking will have a difficult task in understanding what will be achieved when this program is executed. Nevertheless, the computer hardware recognizes only this type of instruction code.

The following program is an example of a machine language program for adding two numbers.

0011	1110	Load A register with value 7
0000	0111	value 7
0000	0110	Load B register with 10
0000	1010	A = A + B
1000	0000	Store the result
0011	1010	into the memory location
0110	0110	
0000	0000	whose address is 100 (decimal)
0111	0110	Halt processing

2.2.2 Assembly Language

When we employ symbols, (letter, digits or special characters) for the operation part, the address, register, operand and other parts of the instruction code, this representation is called an assembly language program. This is considered to be the second generation language

Machine and Assembly languages are referred to as low level languages since the coding for a problem is at the individual instruction level.

Each machine has got its own assembly language, which is dependent upon the internal architecture of the processor.

An assembler is a translator, which takes its input in the form of an assembly language program and produces object code as its output.

The following program is an example of an assembly language program for adding two numbers X and Y and storing the result in some memory location.

LDA, 7	Load register A with 7
LDB, 10	Load register B with 10
ADD A, B	A = A + B
LD(100), A	Save the result in the location 100
HALT	Halt process

From this program, it is clear that usage of mnemonics in our example (LD, ADD, HALT are the mnemonics) has improved the readability of our program significantly.

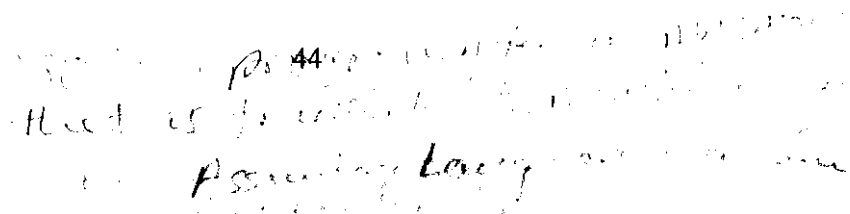
A machine cannot execute an assembly language program directly, as it is not in a binary form. An assembler is needed in order to translate an assembly language program into the object code executable by the machine. This is illustrated in the figure 1.

Assembly
Language
Program

—01 Assembler

Object Code in machine
language

Figure 1: Assembler



Advantage of Assembly Language

Writing a program in assembly language is more convenient than in machine language. Instead of binary sequence, as in machine language, it is written in the form of symbolic instructions. Therefore, it gives a little more readability.

Disadvantages of Assembly Language

Assembly language (program) is specific to particular machine architecture. Assembly languages are designed for specific make and model of a microprocessor. It means that assembly language programs written for one processor will avork on a different processor if it is architecturally different. That is why the assembly language is not portable.

Assembly language program is not as fast as machine language. It has to be first translated into machine (binary) language code.

2.2.3 High level Language

We have talked about programming languages as COBOL, FORTRAN and BASIC. They are called high level programming languages. The program shown below is written in BASIC to obtain the sum of two numbers.

```
10 LET X = 7
20 LET Y = 10
30 LET sum = X + Y
40 PRINT SUM
50 END
```

The time and cost of creating machine and assembly languages was quite high. And this was the prime motivation for the development of high level languages.

Since a high level source program must be translated first into the form the machine can understand, this is done by a software ailloteampihm—which, _s_tAes thc_source sodp..asit_iptand pnalucts_as output-the-machine language code of the machine on which it is to be executed. This is illustrated in figure 2.



Figure 2: Compiler

During the process of translation, the Compiler reads the source programs statementwise and checks the syntax (grammatical) errors. If there is any error, the computer generates a printout of the errors it has detected. This action is known as diagnostics.

- There is another e of software, which also does the translation. This is called in Interpreter. The
- .iler and Internreter ave different approaches to translation. The following table lists the differences
 - etween a Compiler and an Interpreter.

Compiler	Interpreter
:1) Scans the entire program first and then translates it into machine code	Translates the program linuy line
:2) Converts the entire program to machine code; when all the syntax errors are removed execution 901 takes place.	Each time the program is executed, every line is checked for syntax error and then convertee to equivalent machine code.

<i>Compiler</i>	<i>Interpreter</i>
:3) Slow for debugging (removal of mistakes from a program) 14) Execution time, is less	Good for fast debugging Execution time is more.

Advantages of High Level Programming Language

There are four main advantages of high level programming languages. These are:

- 1 Readability: Programs written in these languages are more readable than assembly and machine language.
- 2 Portability: Programs could be run on different machines with little or no change. We can, therefore, exchange software leading to creation of program libraries.
- 3 Easy debugging: Errors could easily be removed (debugged).
- 4 Easy Software Development: Software could easily be developed commands of programming language are similar to natural languages (ENGLISH)

High level languages are also called third generation language.

2.2.4 Fourth Generation Language

The fourth generation of programming languages is not as clearly defined as are the other earlier generations. Most people feel that a fourth generation language, commonly referred to as 4GL is a high level language that requires significantly fewer instructions to accomplish a particular task than a third generation language does. Thus, a programmer should be able to write a program faster in 4GL than in third generation language. Most third generation language are procedural languages. This means that the programmer must specify the steps, that is the procedure, the computer has to follow in a program. By contrast, in a fourth generation language the programmer does not have to give the details of procedure in the program, but instead, specifies what is wanted. For example, assume that a programmer needs to display some data on a screen, such as the address of a particular employee (SANTOSH) from the personal file. In a procedural language, the programmer would have to write a series of instructions in the following steps:

- | | |
|--------|--|
| Step 1 | Get a record from the personal file |
| Step 2 | If this is the record for SANTOSH, display the address |
| Step 3 | If this is not the record for SANTOSH, go to Step 1 |

In a non-procedural language (4GL), however, the programmer would write a single instruction that says:
Get the address of Santosh from personal file.

Major fourth generation languages are used to get information from files and data bases, as in the previous example and to display or print the information. These fourth generation languages contain a query language which is used to answer queries or questions with data from a database. For example the following figure shows a query in a common query language SQL,

```
SELECT ADDRESS FROM PERSONNEL
WHERE NAME = "SANTOSH"
```

Some fourth generation languages are used to produce complex printed reports. These languages contain tailor-made programs called generators. With a report generator available, the programmer specifies the headings, detailed data and totals needed in a report. Thus, the report generator produces the required report using data from a file. Other fourth generation languages are used to design screens to be used for data input

and output and for menu. These languages contain certain types of programs called screen painters. The programmer designs how the screen is to look and, therefore, we can say that the programmer paints the screen, using the screen painter program. Fourth generation languages are mostly machine independent.

Usually they can be used on more than one type of computer. They are mostly used for office automation or business applications, but not for scientific programs. Some fourth generation languages are designed to be easily learnt and used by end users.

2.3 Conclusion

We choose a language for writing a program according to the need but a computer executes programs only after they are represented internally in binary form.

2.4 Summary

Categories of languages are machine language, assembly language, high-level languages and fourth generation languages.

2.5 Tutor marked Assignment (TMA)

- What are the differences between a compiler and an interpreter.
- State True and False
 - Assembly language programs are machine independent

True ☐ False ☒

- Machine language programs are independent.

True ☒ False ☐

- High level language are machine independent

True ☒ False ☐

- All programs are machine independent.

True ☐ False ☒

- 4th generation languages are dependent or databases they are using.

True ☐ False ☐

2.6 Further Reading

Dromey, *How to Solve It by Computers*, Prentice Hall of India
Pratt, *Programming Language*, Prentice Hall of India

Handwritten notes in the right margin:

- 1. Compiler is a program that converts a high level language into machine language.
- 2. Interpreter is a program that converts a high level language into machine language and executes it.
- 3. Compiler is a program that converts a high level language into machine language and produces an object code.
- 4. Interpreter is a program that converts a high level language into machine language and executes it.

01714 P Ti c'' a- ri Alqc-119ce cject l) craw e div
 La"t.eAreA 1.-16-11 A
 Ma*gwberiq 2nEI --
 117_11) tet'ei to-si -31-r-e)
 I. ' et% ...e
 --SaiX'--tegest 13 I. 1— An- Starter+ p ft.1145
 4.4¹¹¹-144tt.

Unit 3

Elements of Programming Language

3.0 Introduction

In this unit you will be introduced to elements of programming language.

3.1 Objectives

After completing this unit, you should be able to:

- define the elements of programming languages such as arrays, expressions, variables, input/output statements;
- Define conditional and looping structures; and
- Define subroutines and functions

3.2 Elements of a Programming Language

Learning a programming language requires understanding of concepts such as representation of different types of data in the computer, various methods of expressing mathematical and logical relationship among data elements and the mechanics for controlling the sequence in which operations can be executed for inputting, processing and outputting of data.

3.2.1 Variables, Constants, Data type, Array and Expressions

These are the smallest components of a programming language.

Variable: The first thing we must learn is how to use the internal memory of a computer in writing a program. Memory may be pictured as a series of separate memory cells as shown in figure 3. Computer memory is divided into several locations. Each locations has got its own address.

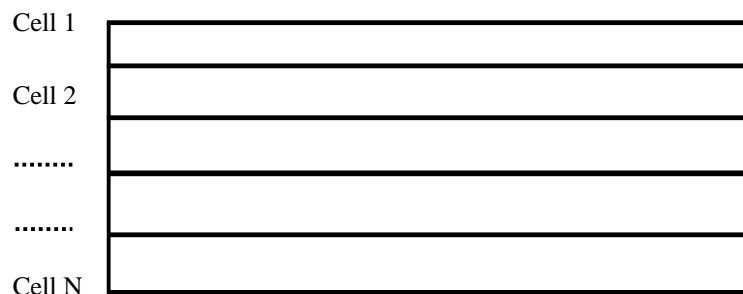


Figure 3: Memory Organization

Each storage location holds a piece of information. In order to store or retrieve information from a memory location, we must give that particular location a name. Now study the following definition.

Variable: It is a group of characters assigned by the programmer to a single memory location and used in the program as a name of that memory location to access the value stored in it. For example in expression $A = 5$, A is a name of memory location i.e. a variable where 5 is stored. **Constant:**

It has fixed value in the sense that two cannot be equal to four. String constant is simply a sequence of

characters such as "computer" which is a string of 8 characters. The numeric constant can be

integer representing whole quantities or a number with a decimal point to represent numbers with fractional part. Constant would probably be the most familiar concept to us since we have used it in doing everything that has to do with numbers. Numeric constants can be added, subtracted, multiplied, divided, and also compared to say whether two of them are equal, less than or greater than each other.

As string constants are a sequence of characters, a related string constant may be obtained from a given one, by chopping off some characters from beginning or end or both or by appending another string constant at the beginning or end. For example, from 'Gone with Wind', we can get one with 'Gone with Wind' and so on. String constants can also be compared in a lexicographic (dictionary) sense to say whether two of them are equal, not equal, less than or greater than each other.

" a A N

f) 'oe, 5 rk ' C

Data type: In computer programming, the data type refers to anything that is processed by the computer. There are different types of data processed by the computer, numbers are one type of data and words are another type. In addition, the operations that are performed on data differ from one type of data to another type. For example, multiplication applies to numbers and not words or sentences.

Data type defines a set of related values/integers, number with fraction, characters and a set of specific operations that can be performed on those values.

In BASIC a statement $LET A = 15$ denotes that A is a numeric data type because it contains numbers but in a statement $LET AS = "BOMBAY"$ AS is a variable of character data type. Data type also defines in terms of contiguous cells should be allocated for a particular variable.

Array: In programming we deal with large amount of related data. To represent each data element we have to consider them as separate variables. For example, if we have to analyze for the sales performance of a particular company for the last 10 years, we can take ten different variables (names) each one representing sales of a particular year. If we analyze sales information for more than 10 years, then number of variables will further increase accordingly. It is difficult to manage with large number of variables in a program. To deal with such situation an array is used.

An array is a collection of same type of data (either string or numeric), all of that are referenced by the same name.

For example, list of 5 years sales information of a company can be referred to by same array name A

A(1)	A(2)	A(3)	A(4)	A(5)
50,000	1,00,000	5,00,000	8,00,000	9,00,000

A(1) specifies sales information of a first year

A(2) specifies sales information of a second year

A(3) specifies sales information of a fifth year

VRatable 4 s ofe lota3 sinned hs Want e, -414 Q gu1

•bri

ek reel Si.

e'r" tvre- S - l k 7 4 R 0 . 1 1 4 - t r t C 4
 CMS-ft/VT:- (0.4a givecl vativ.e • 11- 114)-y bt a V 0-11-.j01-e ft a &in own e,a'c
 57A /Atc X/ers- comedie, Skri;4/ 8' \$itai 9-re cA-Kradd
 p g ecrifec cri swtie tip?efdab; (06- / Sdnerj' ar flume-etc) ht
 ca-eirsvrecrns. 445-n-14154C 12) e1/4. nelnagrate erre:M^M •

Da97■9-reArr h o s vacia etlice) j Ceetrzt c'a ferrn-4,
 fast./ fruit' Va-ttist.

Expression: we know that we can express intended arithmetic operations using expressions such as $X + Y + 7$ and so on. Several simple expressions can even be nested together using parentheses to form complex expressions. Every computer language specifies an order by which various arithmetic operators are evaluated in a given expression. An expression may contain operators such as:

Parentheses)
Exponentiation	A
Negation	
Multiplication, division	
Addition, subtraction	+, -

The operators are evaluated in the order given above. For example, the expression

$2 + 8 * (4 - 6/3)$ Sub expression $(4 - 6/3)$ taken up first

$2 + 8 * (4 - 2)$ division $6/3$ within $(4 - 6/3)$ has higher priority than $4 - 6$

$2 + 8 * 1$ subtraction $(4 - 2)$ is performed next $(4 - 6/3)$ is now complete.

$2 + 8 *$ will be executed first then its result will be added with 2 that is $16 + 2 = 18$.

It is useful to remember the order of priority of various operators. But it is safer to simplify expressions and enclose them in parentheses to avoid unpleasant surprises. So far we have focused on arithmetic expressions. But expression is a very general concept. We mentioned earlier that apart from arithmetic operations we could compare numbers or strings. We do it by using relational operators in expressions.

The following is a list of relational in expressions.

- equal to
- not equal to

- less than
- greater than

$<=$ less than or equal to

$>=$ greater than or equal to

These operations have the same level of priority among themselves but a lower priority than arithmetic operators mentioned earlier. The relational expressions result in one of the truth values, either TRUE or FALSE. When a relational expression such as $(3 > 5)$ is evaluated to be FALSE by such languages, a value 0, that is false, is assigned, whereas $(5 < 7)$ will be evaluated to be TRUE and value 1 will be assigned.

Note that relational expressions are capable of comparing only two values separated by appropriate relational operator. If we want to express an idea such as whether number 7 happens to be within two other numbers 4 and 10, we may be tempted to write relational expression $4 <= 7 <= 10$. Such reasonable expectation from us may be a bit too complex for a computer language. In such cases, we need to explain our idea in terms of simple relational expressions such as $(4 <= 7)$ AND $(7 <= 10)$ which means that 7 is between 4 and 10. To combine several relation expressions for expressing complex conditions, we use logical operators such as AND or operators. Among other logical operators NOT simply negates a truth value in the sense that NOT TRUE is FALSE. and NOT FALSE is TRUE. The logical operators have lower priority than relational operators. Among themselves they have the following order of priority during evaluation of logical expressions.

<i>Operator</i>	<i>Meaning</i>
NOT	Simply negates a truth-value. NOT (2 > 5) is TRUE
AND	TRUE is both adjoining expressions are TRUE otherwise it is FALSE. For example (4 < 7) and (7 < 10) is TRUE whereas (4 > 7) and (7 < 10) is FALSE.
OR	FALSE is both adjoining expressions' are FALSE otherwise it is TRUE for example (4 < 2) OR (7 > 2) is TRUE.
XOR	TRUE only if one of the adjoining expressions is TRUE and other is FALSE. The XOR has same priority as OR. (4 < 7) XOR (7 < 10) is FALSE.

Of the four logical operators NOT, AND, OR are sufficient to express any logical condition.

3.2.2 Input and Output Statement

Syntax: input list of variables

Input statements are normally used for supplying values to a list of variables through an input device such as a keyboard. On execution, a question mark? (in BASIC) would be displayed on the screen, in response to which we are expected to supply a sequence of desired values matching appropriately the type (i.e. integer, floating-point or string type) and the number of variables in the list. For example, in BASIC, INPUT A,B,C is an input statement with list of 3 variables A,B,C.

Assignment Statement

Syntax: variable = expression

This is the most obvious statement used, for assigning value to an identifier (i.e. variable). In BASIC language this is also known as LET statement though the key word LET is optional in most of them. Note the use of = as assignment operator. You would recall that it was also used as relational operator. Some purists justifiably object to ambiguous use of= . In Pascal, := is used for assignment instead of just = which is used as relational operator. In language C = is used for assignment but for relational operator = is used. Few examples of assignment statement in BASIC are given below.

```
LET SUM = X + Y
LET X = 5
```

SUM and X are called variables whereas A + B is called an expression.

Print Statement

Syntax: print list of variables.

This is again an obvious way to display values of specified list of variables on display screen. We shall make use of print to represent generic output statement independent of any particular output device such as display screen, printer and so on. Computer languages normally provide very rich facilities for making the display attractive as possible. An example of print statement in BASIC is PRINT A,B,C which will display values of A,B,C on display screen.

3.2.3 Conditional and Looping Statement

Conditional Statement

If then else structure

Syntax: if(condition)

 Then statement(s) for Task A,
 Else statement(s) for Task A
 End if

This is one of the most important program structures that enable a modern programmer to develop well structured programs. Note that we are calling it as structure rather than statement. As structure it accommodate statements corresponding to alternate tasks to be performed depending upon truth value of the condition. The structures has three components: a conditional expression the truth value of which determines the action, a then – clause denoting task to be performed if the condition is TRUE (Task A), and an optional else clause with task to be performed if the condition is FALSE.

Looping Statement

Obviously, the purpose of a loop structure is to repeat certain tasks until no more repetition is desired. We shall be presenting several variations of loop structure appropriate to handle different situations.

While-end While:

Syntax: WHILE (While Condition)

statement(s) for task being repeated

End while

A loop structure is characterized by the body of statements to be repeated, and the condition for termination of loop. In while-end-while version the condition is tested before every execution of the body. The body is executed if the condition is found to be TRUE. The loop is active as long a condition evaluates to be TRUE and terminates when the condition is evaluated to be FALSE. In case the condition is found to be FALSE for the first time, the body will be skipped.

For-end for structure

Syntax: for var = Start to Finish step iner do

Body of Statement

End for

This loop structure is distinctly different from the while condition — end while we have discussed so far. Unlike it, the end for structure counts the number of times the body is to be executed on the basis of the number of times a control variable is incremented from start to finish in specified steps. Managing the control variable is the responsibility of the for-end-for structure. So we need not and should not worry about assigning any value to it in the body of statements. Any attempts to do so may land us in trouble. This loop structure is so famous that it is available in all languages, structured or not. As usual, key words provided may be

different. The Microsoft BASIC calls it for-next, and optionally provides for specification of control variable after next. Many other languages insist on mention of the control variable after next.

We have discussed program structures quite elaborately. Let us turn our attention to data structures that hold our values. So far we have been dealing with simple variables that hold only one value of a desired type. So we have integer variable, floating point variable, and string variable holding only one value on the corresponding type at any moment. Can we have data structures that are capable of holding more than one value? Programming languages do offer complex data structures required for advanced programming. Array and record are two of them. A detailed discussion of these topics is beyond this unit, however you can refer to further readings for more details.

3.2.4 Subroutine and Functions

When writing a program sometimes, it is necessary to repeat a statement or group of statements at several points. For example, it may be necessary to perform the same computation several times. Repeating the same statement in a program each time makes a program lengthy and also gives poor readability. These problem would be sorted out if the necessary statements could be written (coded) once and then referred to each time they are needed.

This is the purpose of subroutine and function. Subroutine and function consist **a** one or more basic statements that can be referred to at different poitain, kprogram. Each time a subroutine and function is riferaiia, we say it is called.

A subroutine is a group statement that can be executed from different points in a program. Syntax for defining subroutine varies from one language to another language.

Functions: The word function in programming means essentially what it does in mathematics. Function is a rule or series of rules that assign one and only one value Y to a given value X. Thus, X and Y is called variables. Value of Y is dependent upon value of X. Variable X is also called the argument of the function. The notation used for function should already be quite familiar to you.

$$Y = F(X)$$

In this notation 'F' is the name of the function. Consider a specific example

$$Y = F(X) = 3X + 3$$

If we let X equal 3, then this rule or function directs us to compute 12 for Y if we assign the value of X = 3, then Y is 12. Using function notation, $F(3) = 12$ and $F(9 - 3) = 12$. Each language has different rules to define a function.

Library functions: These are functions supplied with you language compiler, interpreter. The code for a library function does not appear in the program. The computer inserts it when the program is translated into machine language. Some built in functions in C programming language are Printf, Scan etc.

3.3 Conclusion

You have been introduced to many aspects of computer programming languages. These concepts are essential to acquire for any programming language.

3.4 Summary

In this unit you learned many concepts such as variables, constants, arrays, expressions, input/output statements, assignment statements, conditional and looping structures and subroutine functions.

3.4 Tutor marked Assignment (TMA)

- I What is the purpose of a looping statement in a programming language?
2. What is the advantage of using an array in a programming language?
3. What is the difference between subroutine and function?

3.6 Further Reading

R.J. Dromey, *How to Solve It by Computers*. Prentice Hall of India

Pratt, *Programming Languages*, Prentice Hall of India

Unit 4

Operating System Concepts

4.0 Introduction

An Operating System is a sytgin softwAre _whig-t_h _may be viewed as an organised collection of software corSting_ of nrocedures for opeisatia computer and providing an environment for execution of programs.

It acts as an interface between users aid_t12e1 hardware ofSonsel-sysiii. There are many important reasons for studying operating systems. Some of them are:

- (1) User interacts with the computer through operating system in order to accomplish his task since it is his primary interface with a computer.
- (2) It helps users to understand the inner functions of a computer very closely.
- (3) Many concepts and techniques found in operating system have general applicability in other applications.

In the previous units, we mainly introduced different types of software and categories of programming languages. All these software have been developed under a particular operating system environment. The introductory concepts and principles of an operating system will be the main issues for the discussion in this unit. Evolution and types of operating systems will also be broadly covered here.

4.1 Objectives

After going through this unit, you should be able to:

- list and interpret the basic functions of operating systems
- to make comparison among different types of operating systems
- display status information regarding various Read/Write error conditions. For example, the program may find that the file has been reached or that there was a hardware failure. The write operation may encounter various errors, depending upon the output device (no more disk space, physical end of tape, printer out of paper and so on).
- close both files after the entire file is copied.

As we can observe, a user program takes heavy use of the operating system. All interaction between the program and its environment must occur as the result of requests from the program to the operating system.

Operating System Commands: Apart from system calls, users may interact operating system directly by means of operating commands.

For example, if you want to list files or sub-directories in MS-DOS, you invoke direct command. In either

case, the operating system acts as an interface between users and the hardware of a computer system. The fundamental goal of computer systems is to solve user problems. Towards this goal computer hardware is designed. Since the bare hardware alone is not very easy to use, programs (software) are developed. These programs require certain common operations, such as controlling peripheral devices. The command function of controlling and allocating resources are then brought together into one piece of software, the operating system.

To see what operating systems are and what operating systems do, let us consider how they have evolved over the years. By tracing that evolution, we can identify the common elements of operating systems and examine how and why they have developed as they have.

4 What is an Operating System?

An operating system is an essential component of a computer system. The primary objective of an operating system is to make computer system convenient to use and utilize computer hardware in an efficient manner.

An operating system is a large collection of software, which manages resources of the computer system. such as memory, processor, file system and input/output devices. It keeps track of the status of each resource and decides who will have a control over computer resources, for how long and when. The positioning of operating system in overall computer system is shown in figure 4.

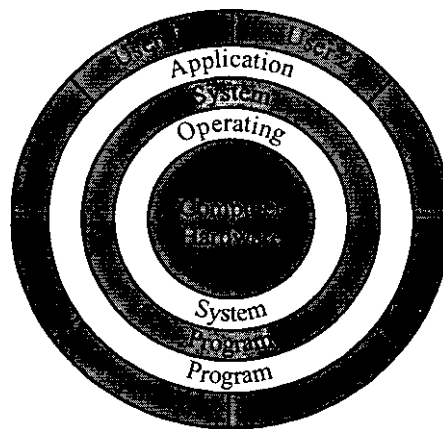


Fig. 4: Component of computer system

From the figure, it is clear that operating system directly controls computer hardware resources. Other programs rely on facilities provided by the operating system to gain access to computer system resources. Here are two ways one can interact with operating system.

1. By means of Operating System Call in a program
2. Directly by means of Operating System Commands

System Call: System calls provide the interface to a running program and the operating system. User program receives operating system services through the set of system calls. Earlier these calls were available in assembly language instructions but now a days these features are supported through high-level languages like C, Pascal etc, which replaces assembly language for system programming. The use of system calls in C or Pascal programs very much resemble pre-defined or sub-routine calls.

As an example of how system calls are used, let us consider a simple program to copy data from one file to another. In an interactive system, the following system calls will be generated by the operating system:

- Prompt messages for inputting two file names and reading it from terminal.
- Open source and destination file.

99004 Coil LS S Ccta Cr-r7 eT aft tz) qUQf ci 04'30k-¹3 far⁻¹- b•-
s (fin:yvickl kr(c_{ctb} rterin,, cNe Cr herviiera_c-e.

- Prompt error messages in case the source file cannot be open because it is protected against access or destination file cannot be created because there is already a file with this name.
- Read the source file.
- Write into the destination file.

The next development was the replacement of card-decks with standard input/output and some useful library programs, which were further linked with user program through a system software called linker. While there was a definite improvement over machine language approach, the serial mode of operation is obviously not very efficient. This results in low utilization of resources.

4.3 Evolution of Operating Systems

AuJapeatiag,system may process its task serially (sequentially) or concurrently (several tasks simultaneously). It means that the resources of the computer system may be dedicated to a single program until its completion or they may be allocated among several programs in different stages of execution. The feature of operating system to execute multiple proas in an interleaved fashion or in different time cycles, is caltal as multiprogramming systems. In this section, we will try to trace the evolution of operating system. In particular, we will describe serial processing, batch processing and multiprogramming.

4.3.1 Serial Processing

Programming in Is and Os (machine language) was quite common for early computer systems. Instruction and data used to be fed into the computer by means of console switches or perhaps through a hexadecimal keyboard. Programs used to be started by loading the program counter register (a register which keeps track of which instruction of a program is being executed) with the address of the first instruction of a program and its result (program) used to be examined by the contents of various registers and memory locations Of the machine. Therefore, programming in this style caused a low utilization of both users and machine.

Advent of inputs/output devices, such as punched cards paper tapes and language translators (Compiler/Assemblers) brought a significant step in computer system utilization. Program, being coded into programming language, is first changed into object code (binary code) by translator and then automatically gets loaded into memory by a program begins and its results gets displayed or printed. Once in memory, the program may be re-run with a different set of input data.

The process of development and preparation of a program in such environment is slow and cumbersome due to serial processing and numerous manual processing. In a typical sequence first the editor i's called to create a source code of user program written in programming language translator is called to convert a source code into binary code and then finally loader, and then finally loader is called to load executable program into main memory for execution. If syntax errors are detected, the whole process must be restarted from the beginning.

4.3.2 Batch Processing

Utilization of computer resources and improvement in programmer's productivity was still a major prohibition. During the time that tapes were being mounted or programmer was operating the console, the CPU was sitting idle.

The next logical step in the evolution of operating system was to automate the sequencing of operations involved in program execution and in the mechanical aspects of program development. Jobs with similar requirements — were batched together and run through the computer as a group. For example, suppose the operator received one FORTRAN program, one COBOL program and another FORTRAN program. If he runs them in that order, he would have to set up for FORTRAN program environment (loading the FORTRAN compiler tapes) then Set up COBOL program and finally FORTRAN program again. If he runs the two

FORTRAN programs as a — batch, however he could set up only once for FORTRAN, thus saving operator's time.

Batching similar jobs brought utilization of system resources quite a bit. But there were still problems. For example, when a job is stopped, the operator would have to notice that fact by observing the console, determine why the program stopped and then load the card reader or paper tape reader with the next job and restart the computer. During this transition from one job to the next the CPU sat idle.

To overcome this idle time, a small program called a resident monitor was created which is always resident in the memory. It automatically sequenced one job to another job. Resident monitor acts according to the directives given by a programmer through control cards which contain information like marking of jobs beginnings and endings, commands for loading and executing programs, etc. These command belong to job control language. These job control language commands are included with user program and data. Here is an example of job control language commands.

\$COB	-Execute the COBOL compiler
\$JOB	-First card of a job
\$END	-Last card of a job
\$LOAD	-Load program into memory
\$RUN	-Execute the user program

\$ END

```

Data from Program
$ RUN
$ LOAD
USER'S
$ COB
$ JOB
```

Fig. 5: Card deck for Cobol Program for a simple batch system

With sequencing of program execution mostly automated by batch operating system, the speed discrepancy between fast CPU and comparatively slow input/output devices such as card readers, printers emerged as a major performance bottleneck. Even a slow CPU works in the microsecond range, with millions of instructions per-second. But, fast card reader, on the other hand, might read 1200 cards per minute. Thus, the difference in speed between the CPU and its input/output devices may be three orders of magnitude or more.

The relative slowness of input/output devices can mean that CPU is often waiting for input/output. As an example, an Assembler or Compiler may be able to process 300 or more cards per second. A fast card reader, on the other hand, may be able to read only 1200 cards per minute. This means that assembling or compiling a 1200 card program would require only 4 seconds of CPU time out 60 seconds to read. Thus, the CPU is idle for 56 out of 60 seconds or 93.3 percent of the time. The resulting CPU utilization is only 6.7 per cent. The process is similar for output operations. The problem is that while an input/output is occurring, the CPU is idle, waiting for the input/output to complete; while the CPU is executing, input/output devices are idle.

Over the years, of course, improvements in technology resulted in faster input/output devices. But CPU speed increased even faster. Through its use and resource utilization by overlapping input/output and processing - it IBM's, and later dedicated input/output processors brought a major improvement in this direct (Direct Storage Memory Access) chip which directly transfers the entire block of data from its own buffer to main memory without intervention by CPU was a major development. While CPU is executing, DMA can transfer data between high speed input/output devices and main memory. CPU need not be interrupted per block only by DMA. Apart from DMA, there are two other approaches to improving system performance by overlapping input, output and processing. These are buffering and spooling.

If the CPU is on the average much faster than an input device, buffering will be of little use. If the CPU is always faster, then it always finds an empty buffer and has to wait for the input device. For output, the CPU can proceed at full speed until, eventually system buffers are full then the CPU must wait for output device. This situation occurs with input/output bound jobs where —the amount of input/output device controls the speed of execution, not by the speed of the CPU.

More sophisticated form of input/output buffering called SPOOLING, (simultaneous peripheral operation on line) essentially use the disk as a very large buffer (figure 6) for reading and for storing output files.

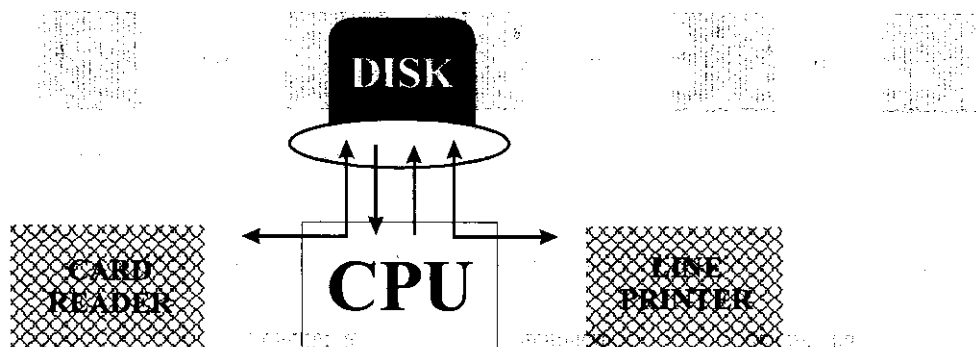


Fig d: Spooling

Buffering overlap input output and processing of a single job where a Spooling allows CPU to overlap the input/output job with the computation and input of other jobs.

Therefore this approach is better than buffering. Even in a simple system, the spooler may be reading the input of one job while printing the output of a different job.

4.3.3 Multiprogramming

(program) to utilize CPU time at any moment. More number of programs competing for system resources, better will be resource utilization.

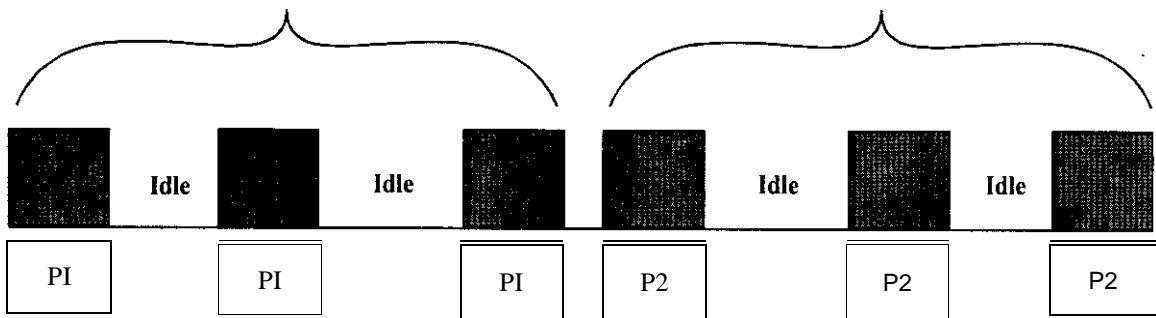
The idea is implemented as follows. The main memory of a system contains more than one program

(Figure 7).

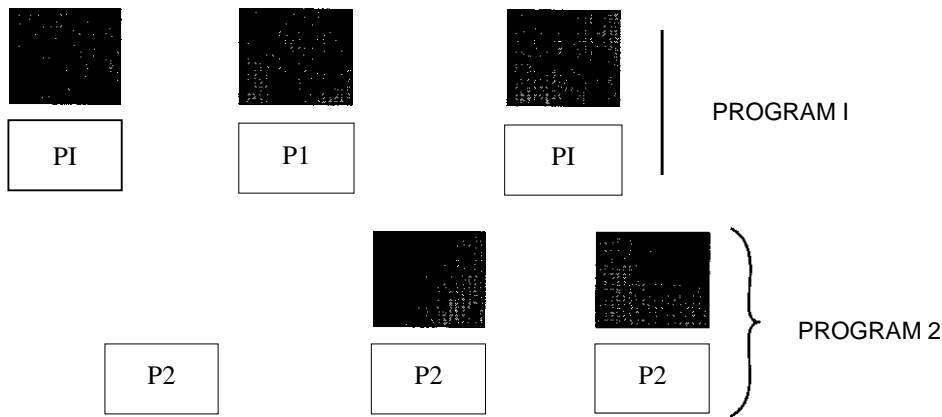


Fig. 7: Memory layout in multiprogramming environment

The operating system picks one of the programs and start executing. During execution process (program) I may need some I/O operation to complete. In a sequential execution environment (Figure 8a), the CPU would sit idle. In a multiprogramming system, (Figure 8b) operating system will simply switch over to the next process (program2).



(a) Execution in sequential programming environment



(b) Execution in multiprogramming environment

Fig. 8: Multiprogramming

When that program needs to wait for some I/O operation, it switches over to Program 3 and so on. If there is no other new program left in the main memory, the CPU will pass its control back to the previous programs.

Time-sharing has traditionally been employed to increase the resource utilization of a computer system and to support multiple simultaneously interactive users (terminals).

Compared to operating system which supports only sequential execution, multiprogramming system requires some form of CPU and memory management strategies which is beyond the scope of this block.

4.4 Conclusion

You are now able to list and interpret the basic functions of operating systems you should also be able to compare different types of operating systems.

4.5 Summary

Operating System is an essential component of a system software, which consists of procedures for managing computer resources. Batch systems allowed automatic job sequencing by a resident monitor and improve the overall utilization of systems greatly.

4.6 Tutor Marked Assignment (TMA)

(1.400-7j⁵ Pirkr' 4'(i⁰/5 •

1. What is a system call? -- *inierec*—e h/0¹ a

2 Define the essential difference between

- (a) spooling ----.4 -, . / ' ,
- (b) buffering— C 0, a spix , IN r " + e / ' ,

i re) 01 3"n CI⁶/ —"

3. What are the main advantages' of multiprogramming ...,

4.7 Further Reading

Milan Milenkovic, *Operating Systems, Concepts and Design*, Second Edition, Tata Mcgraw-Hill.

Unit 5

Types of Operating System

5.0 Introduction

This unit continues from where we stopped in unit 4. Here we introduce you to types of operating system.

5.1 Objectives

The reasons for studying operating system will be stated here again for emphasis as:

- user interacts with the computer through operating system in order to accomplish his task since it is his primary interface with a computer.
- it helps users to understand the inner functions of a computer closely.
- many concepts and techniques found in operating systems have general applicability. 52

Types of Operating System

In this section, we will discuss general properties, types of application advantages, disadvantages and basic requirements of different types of operating systems.

5.2.1 Batch Operating System

As discussed earlier during batch processing environment it requires grouping of similar jobs, which consist of programs, data and system commands.

The suitability of this type of processing is in programs with large computation time with no need of user interaction or involvement. Some examples of such programs include payroll, forecasting, statistical analysis, and large scientific number crunching programs. Users are not required to wait while the job is being processed. They can submit their programs to operators and return later to collect them.

But it has two major disadvantages:

- (i) Non-interactive environment.
- (ii) Off-line debugging.

Non-interactive environment: There are some difficulties with a batch system from the point of view of user. Batch operating systems allow little or no interaction between users and executing programs. The time taken between job submission and job completion in batch operating system is very high. Users have no control over intermediate results of a program. This type of arrangement does not create flexibility in software development.

The second disadvantage with this approach is that programs must be debugged which means a programmer cannot correct bugs the moment it occurs.

5.2.2 Multiprogramming Operating System

Multiprogramming compared to batch operating systems are fairisophisticated. As illustrated in figure 5, multiprogramming has a significant potential for improving systems throughout and resource utilization with very minor differences. Differen ammin o ,eratin system are intaltsrnthl yltiprocessor and multi-user operating systems. In this section, we will briefly discuss the main q ery m,cre_ pit'cre95-1 carreeL5 d"Scica-re. &AAR features and functions of these systems.

Multi-tasking Operating Systems: A running state of a program is called a process or a task. A multi-tasking operating system supports two or more active processes simultaneously. Multiprogramming operating system is the operating system which in addition to supporting multiple concurrent process (several processes in execution states simultaneously) allows the instruction and data from two or more separate process to resale in primary memory simultaneously. Note that multiprogramming implies multiprocessing or multitasking operation, but multiprocessing operation (or multitasking) does not imply multi-programining. Therefore, multitasking operation is one of the mechanism that multiprogramming operating system employs in managing the totality of computer related resources like CPU, memory and I/O devices.

Multi-useV operating systemallows sit 3311/0,1neous access to a computer system through two or more terminals. Although frequently associated with multiprogramming, multi-user operating system does not imply multiprogramming or multitasking. A dedicated transaction processing system such as railway reservation system that supports hundreds of terminals under control of a single program is an example of multi-user operating system. On the other hand, general purpose time sharing systems (discussed later in this section) incorporate features of both multi-user and multiprogramming operating system. Multiprocessor operation without multi-user can be found in the operating system of some advanced personnel computers and in real systems (discussed later).

"4514⁴¹ CS-Ser StAa•rit Cnotee C p CA FiAtot.e. ats-

Time Sharing System: It is a form of multi programmed operating system, which operate. In an interactive mode with a quick response time. The user types a request to the computer through a keyboard. The computer processes it and a response (if any) is displayed on the user's terminalcA time sharinglysteinyllows the many users to simultaneously share the computer resources. Since each action or commarin a time-shared

The term multi-tasking is described as any system that runs or appears to run more than one application program one time An effective multi-tasking environment must provide many services both to the user and to the application program it runs. The mosi important of these are resource management which divides the computers time, memory ancl peripheral devices among competing tasks and inter process communication, which lets- tasking co-ordinate their activities by exchanging information.

Real Time Systems: It is another form of o • erati • 1 e used in environments where a large

number of events mostly external to comnuter systems, must be accepted and processed in a short time or 17%24hin certain deadlines. Examples of such applications are flight control, real time simulations etc. Real time systems are also frequently used in military application. .

A primary objective of real-time system is to provide quick response times. U • nd resource itilizipp_areaf-stemllaconcem to real-time system. Inihe real-time system each process is assigne a certain level of priority according relative importance of the event procetses. The processor is normally i located to the highest priority process among those, which are ready to execute. Higher priority process usually pre-empt execution of lower priority processes This form of scheduling called

ulti pro r o m m i c e O iticati:613-et,24,2f ots- —;,e, arencentlf sfgAn-

"rctietp_____ "2t,, - ' a,4, 44210 Al-a-DP_ irrenrif era 723fl_f of dean °"45/71

Lte-3- 5 re °en a cart m aCS a-e_l-cote) en Pitn-v17 ⁴4Csarey 9-1 ,e4tabufeta44,

pre-emptive scheduling, is used by a majority of real-time systems.

5.2.3 Network Operating System

A network operating system is a collection of software and associated protocols that allows a set of autonomous computers which are interconnected by a computer network to be used together in a convenient and cost-effective manner. In a network operating system, the users are aware of existence of multiple computers and can log into remote machines and copy files from one machine to another machine.

Some of typical characteristics of network operating systems which may be different from distributed operating system (discussed in the next section) are the followings: •

Each computer has its own private operating system instead of running part of a global system wide operating system.

Each user normally works on his/her own system; using a different system requires some kind of remote I/O, instead of having the operating system dynamically allocate processes to CPUs.

- Users are typically aware of where each of their files are kept and must move file from one system to another with explicit file transfer commands instead of having file placement managed by the operating system.

The system has little or no fault tolerance; if 5% of the personnel computers crash, only 5% of the users are out of business.

Network operating system offers many capabilities including:

- Allowing users to access the various resources of the network hosts.
- Controlling access so that only users in the proper authorization are allowed to access particular resources.
- Making the use of remote resources appear to be identical to the use of local resources.
- Providing up-to-the minute network documentation on-line.

5.2.4 Distributed Operating System

A distributed operating system is one that looks to its users like an ordinary centralized system funsp multiple independent CPUs. The key concept here is transparency, On other words, the use of multiple processors should be invisible to the user. Another way of expressing the same idea is to say that user views the system as virtual unprocessed but not as a collection of distinct machines. In a true distributed system, users are not aware of where their programs are being run or where their files are residing; they should all be handled automatically and efficiently by the operating system.

Distributed operating systems have many aspects in common with centralized ones but they also differ in certain ways. ecastal-gti, system, for example often allow programs to 'non several processors at the same time, thus requiring more complex processor scheduling (scheduling refers to a set of policies and

mechanisms built into the operating systems that controls 'the order in which the work to be done is completed) algorithms in order to achieve maximum utilization of CPUs'time.

Fault-tolerance is another area in which distributed operating systems are different. Distributed systems are considered to be more reliable than uniprocessor based system. They perform even if certain part of the hardware is malfunctioning. This additional feature supported by distributed operating system has enormous implications for the operating system. *Advantages of Distributed Operating System*

There are three important advantages in the design of distributed operating system.

- I Major breakthrough in microprocessor technology: Micro-processors have become very much powerful and cheap, compared with mainframes and minicomputers, so it has become attractive to think about designing large systems consisting of small processors. These distributed systems clearly have a price/performance advantages over more traditional systems.

2. Incremental Growth: The second advantage is that if there is a need of 10 percent more computing power, one should just add 10 percent more processors. System architecture is crucial to the type of system growth, however, since it is hard to give each user of a personal computer another 10 percent.
3. Reliability: Reliability and availability can also be a big advantage; a few parts of the system can be down without distributing people using the other parts; On the minus side, unless one is very careful, it is easy for the communication protocol overhead to become a major source of inefficiency.

Exercise

- I Define the essential differences between the time-sharing and real time operating systems.

.....

.....

.....

- 2 List the main differences between network operating systems and distributed operating system.

.....

.....

5.3 Conclusion

You now know some types of operating systems. You are also now in a position to enumerate the advantages and disadvantages of each system.

5.4 Summary

Operating Systems are now almost always written in higher level languages (C, PASCAL etc). UNIX was the first operating system developed in C language. Operating Systems provide a number of services. At the lowest level, they are system calls which allow a running program to make a request from operating system directly. At a higher level, there is a command interpreter which supports a mechanism for a user to issue a request without writing a program.

5.5 Tutor Marked Assignment (TMA)

- I Define the essential differences between the line sharing and real time operating systems.
2. List the main differences between network operating systems and distributing of operating systems.

5.6 Further Reading

Milan Milenkovic, *Operating Systems and Design*, Second Edition. Tata McGraw — Hill.

Module 3

Communication and Networking

Module Introduction

The trend now is of a greater communication between computers. It is expected that with better communication infrastructure becoming available in this country, network based computers will be present everywhere. A computer could be connected to other computers in the same room, in the same building, in the same campus, in the same town, in the same country or even globally. As organizations become more dependent on computers, it is necessary to plan for contingencies.

This module provides a brief glimpse of Fundamentals of Data Communication and various networking technologies and some of the emerging trends.

Unit 1

Fundamentals of Data Communication

1.0 Introduction

A standalone computer in today's context is not very useful, but coupled with communication technology, it opens up an enormous repository of information to its users. Information is carried in data communication systems as signals between two or more points, which could be at a distance of a few inches or several thousand kilometers.

These signals are subject to various effects while they are in transit which alter their characteristics to some degree. The data communication technology ensures that information between two communication entities is transferred in a reliable and orderly manner.

To understand transmission one needs to study the electrical form that messages take while they are in transit and of media and communication technologies that ensure error-free transmission. In our subsequent units we would be discussing the basic concepts of electronic communications. Data Communication is also well known as DATACOMM.

1.1 Objectives

At the end of this unit, you would be able to:

- define the basic concepts of Data Communication
- define various terminology of Data Communication

1.2 Definitions

In data communications, four basic terms - Data, Signal, Signalling and Transmission are frequently used. Data are entities that convey meaning. Information is obtained by processing data using desired functions.

Signals are electric or electromagnetic encoding of data, and signaling is propagation of signal along suitable

communication medium.

Transmission is Communication of data achieved by the propagation and processing of signals.

1.3 Concept of Data Communication

The Concept of Data Communication evolved from sharing the computation power of a computer along with various resources available in a computer environment such as printer, hard-disk etc. With increasing demand for exchange of information across the globe, the need for data communication has increased in many folds. Due to physical constraints involved in connecting two remote points physically, data communication has emerged as an instant solution.

Safq Gannaiticakto cr cen4,14 Prtsral cn-LAt P6i_ci. A 75f4 cZar-
i4-ebt.,ft— pc yak"- tecindirtaiy, ki4-3-a-re4 ett wit-at t(os `l. ceretioicsfr_
edlAhei lg traScifeind 4o a re (cadhce or orders fr iuuers
in Olt estatin lent 49 Otareliegi ad ittingted 20 - ngenrt _*
catru44 it? set indio4 err ki 11 • Teke&rs-- t_Agiolgeti cseg*

1.3.1 Data Communication Codes (Data Encoding)

The first step towards understanding communications is to look at computer data at its base level. As all of us know computer and computer device manage store and exchange data using electronic pulses or digital signals, that come in two varieties, the binary digit '0' indicates the absence ("OFF") and '1' indicates ("ON") the presence of electric current. A series of ONs and OM in various combinations can be sent on the communication channels to represent any character.

Every character (letter, numeral, symbol, or punctuation mark) is composed of a group of bits called codes. To avoid incompatibility between system, the computer industry has created a number of standards that establish relationships between bit combinations and their corresponding characters.

The most widely used codes are the American Standards Code for Information Interchange (ASCII) and the IBM standard Extended Binary Coded Decimal Interchange Code (EBCDIC). Some of the characteristics of these codes are tabled below:

Data Codes

	ASCII	EBCDIC
Controlling authority	ANSI	IBM
Number of bits	7	8
Number of Character	128	256

To ensure successful transmission of data between two points, the sender and receivers should use the same code, or they can use translators to aid in communication. Translation is data communication software, which translates the sender's data into the codes receiver understands.

1.3.2 Communication Speed or Rate

The speed at which two computers exchange or transmit data is called communication rate on transmission speed. The unit of measurement of the speed is measured using bps (bit per second) or baud (both are not same, for more details refer to further readings). Normal PC based communication station transferred, using 300 to 9600 bps, whereas a main frame computer uses 19,200 baud or more.

E.3.3 Private vs. Switched Channels

Private leased lines are permanently connected circuits between two or more points. These lines available for private use by the leasing party. Whereas, switched lines or dial up lines are available to any subscriber. For example, telephone connection in our home or office is a dial up lines as we dial up a number to use a particular circuit.

1.3.4 Analog and Digital Transmission

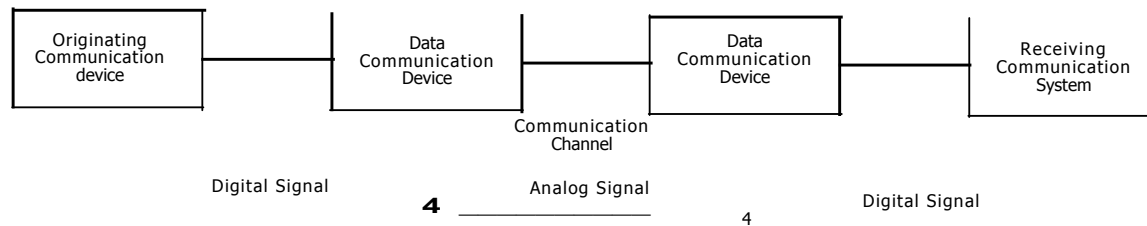
One of the fundamental concepts in data transmission is to understand the difference between analog and digital signals.

An analog signal is one that is continuous with respect to time, and may take quAny_value within a given range of values. Human voice, and music when converted to electrical signal using suitable devices produce analog signals.

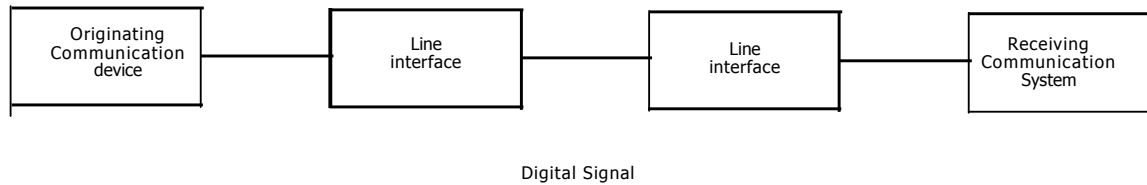
A digital signal may take on only a discrete set of values within a given range. Most computers and computer-related equipment are digital.

Irrespective of original form of the message, the actual transmission of signal could be either in analog or digital form. Figure 1 shows the two possible transmission methods:

21212.52a—tail 0let pnuacj Jed dp etnca-n entby Gooffecied c;rala
ervita, Sou (klAosi ItS3 u-rt 1y Liety Ltd--4 itirdr4;e J-
eAf ctArcui_u_bk 0-94780-Si3 Cel?heri e f telephone.



(a) Analog Transmission



(c) Digital Transmission

Fig 1: Analog and Digital Transmission

1.3.5 Parallel and Serial Transmission

In parallel data transmission, there are multiple parallel lines connecting the transmitting and receiving units.

Each wire carries a bit of information. Normally, one character is transferred in one go.

In serial data transmission, each bit is sent sequentially one after another and it requires only one pair of wire conductors for connecting the receiving and transmitting units.

Serial transmission is slower than parallel transmission, which is used primarily for transferring data between devices at the same site, most common example is Personal Computer to Printer. Communication between computers is almost always serial.

1.3.6 Tariff

As most of the data communication is being done using public channels such as telephone lines, a user has to pay for the use of channels. The tariff is generally calculated taking various data communication parameters into consideration such as volume of data transferred, communication rate or transmission rate and time utilized on the communication channel.

1.3.7 Communication Protocols/Standard -

There are several manufacturers of computer hardware and software across the globe. For successful data communication these products should be compatible with each other or they should conform to certain set of rules so that any one can use them. These sets of rules are known as communication protocols or communication standards.

In other words, protocols are technical customs or guide lines that govern the exchange of signal transmission and reception between equipments.

Communication protocols are usually defined and approved by some international body such as ISO, CCITT or IEEE. At times protocols defined by certain manufacturer of computers or networking products become so widely accepted, that they become 'defacto' standard even though these protocols may not have approval of any standardization body. Some of the functions that communication protocol regulates are:

- Control of information transfer
- Structure and formats of data

- Error recoveries
- Re-transmission control
- Interface management.

Only communication devices using same protocols can communicate with each other.

To make sure that the bits (or characters) are transmitted through the channel or line, we need a link level protocol between two computers. Some of the main functions to be performed by such a link level protocol are:

- Assuring the data to be transmitted is split into data blocks with beginning and ending markers. This is called framing and the transmission block is called a frame.
- Achieving data transparency. This allows a link to treat a bit pattern, including normally restricted control characters, just as pure data.
- Controlling the flow of data across the Link. It is essential not to transmit bits faster than they can be received at the other end. Otherwise, the receiver overflows and the data is overrun, or all buffering capacity is used up, leading to loss of data.
- Controlling errors. This involves detection of errors using some kind of redundancy check. It also involves acknowledgement of correctly received messages and requests for retransmission of faulty messages.

There are basically two classes of link protocols. They are, Binary SYNChronous protocols (BISYNC) and High Level Data Link Control (HDLC) protocols. BISYNC is based on character control, whereas HDLC is a bit oriented protocol. In fact, HDLC is widely used in most link protocols of computers. You can refer to further readings for more details.

14 Conclusion

By now you are able to define basic concepts of Data Communication and various terminology of Data Communication.

1.5 Summary

In this unit we discussed the concept of data communication and defined the process involved in data communication. Various hardware and software concepts were also discussed along with the digital and analog signals. It also introduced you to various types of communication devices and also explained what is transmission protocol.

1.6 Tutor Marked Assignment (TMA)

I Match the following

- | | | |
|---------------------------------------|---------------------------------|-------------------|
| (i) Protocols-- | (a) Baud | |
| (ii) Translators.— | (b) Formalized set of rules | |
| (iii) | | Communication (c) |
| Translate data into same type of code | | |
| (iv) Communication tariff— | (d) Volume of data transferred. | |

2. Fill in the blanks

- (a) allow 4 one single communication channel to send many transmission at the same time.
Any subscriber can use —t,— ' lines
..... (c).L..... is a unit of measurement for communication rate.

1.7 Further Reading

William Stallings, *Data and Computer Communication*, Prentice-Hall of India.

Unit 2

Data Communication and Hardware

2.0 Introduction

In this unit you will learn the functions of nodes, workstations, modem and codes, multiplexer and communication channels.

2.1 Objectives

At the end of this unit, you would be able to:

- state the functions of the following Data Communication hardware:
Nodes and Workstations, Modem and codes, multiplexes and communication channels
- Describe the implication of Data Communication

2.2 Data Communication Modes

So far, we have discussed a method of transferring information electronically and various aspects involved in it such as data characters in a computer system are represented using codes such as ASCII or EBCDIC. The exchange of information can take place only if a facility exists to send information from one end, character by character and also to receive it at the other end in the same manner or sequence. Hence we must understand how the characters are transmitted over the transmission medium.

2.2.1 Synchronous and Asynchronous Transmission

The mode of transmission is the way in which coded characters are assembled for the process of transmission and permits the receiving devices to identify where the coding for each character begins and ends within the torrent of bits. When two computers communicate, they must have a way to synchronize the flow of data so that the receiving computer can read at the same speed at which the sending computer transmits. The principal modes are asynchronous and synchronous.

In synchronous transmission characters are transmitted as groups, with control characters in the beginning and at the end of the bit train. The transmission and receiving intervals between each bit are precisely timed permitting the grouping of bits into identifiable characters. In synchronous modes, intervals between characters are uniform with no space between consecutive bytes.

In asynchronous transmission each character is transmitted separately, that is, one character at a time. Each character begins with a start bit, which tells the receiving device where the character coding begins and ends with a stop bit, which tells the receiving device where the character coding ends. Then, the next character is sent, with start and stop bits. The start and stop bits and the interval of time between consecutive characters allow the receiving and sending computers to synchronize the transmission. The parity bit is used

in S r

for error checking while transmitting data.

Asynchronous communication is slower than synchronous communication, it is typically used at

(communication at 2400 bits per second. Asynchronous communication does not require complex and costly hardware as required by synchronous communication and is the mode most widely used with microcomputers.

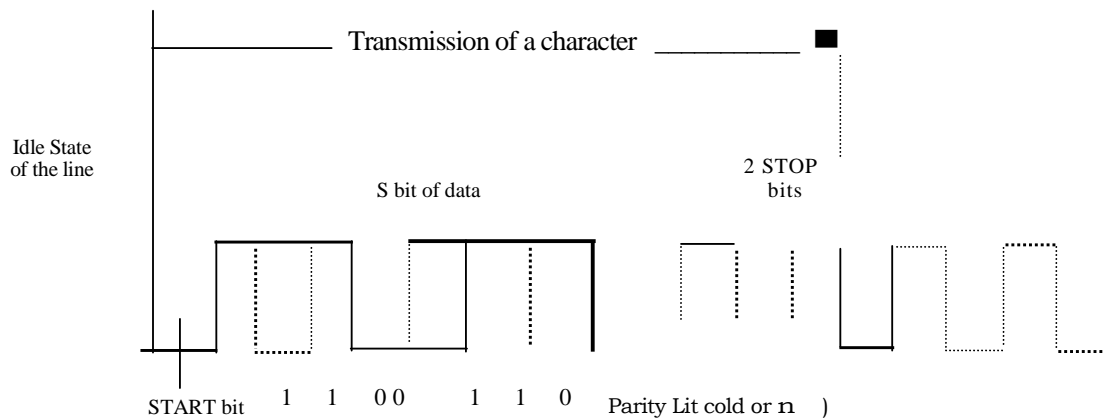


Fig. 2(a): Asynchronous Serial Transmission

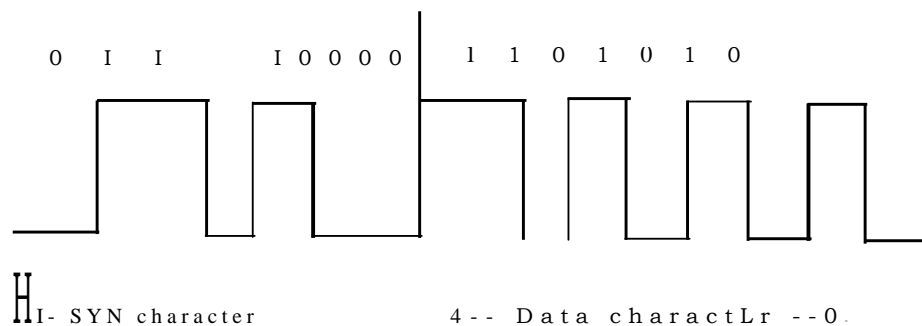


Fig. 2(b): Synchronous data character format

2.2.2 Simplex, Half-duplex and Full Duplex Communication

The direction in which information can flow over a transmission path is determined by the properties of both the transmitting and the receiving devices. There are three basic options.

In simplex mode, the communication channel is used in one direction. The receiver receives the signals from the transmitting device. A typical use is to gather data from a monitoring device at a regular interval.

The simplex mode is rarely used for data communication.

In half duplex mode, the communication channel is used in both directions, but only in one direction at a time. This requires the receiving and transmitting devices to switch between send and receive modes after each transmission. The analogous example of this mode is the old wireless system (walkie-talkie) which is used in either transmit mode or receive mode.

In a network, the communication channel is used in both directions at the same time. Typical example of this mode of transmission is the telephone in which parties talk to each other at the same time.

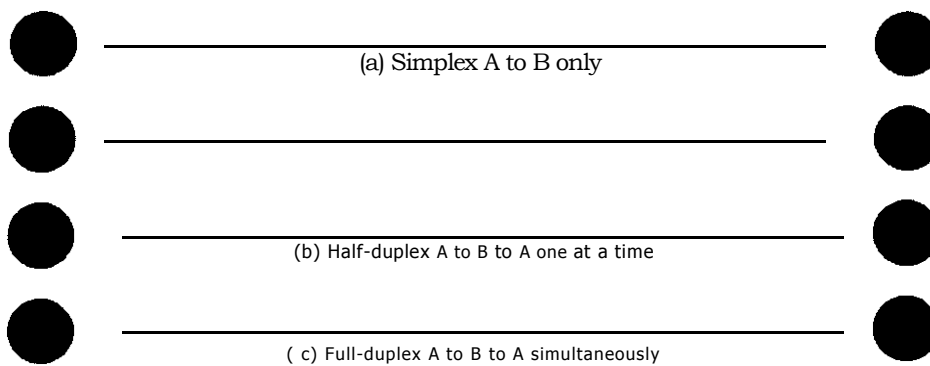


Fig. 3: Modes of Transmission

2.3 Communication Hardware

We have already discussed about communication codes, units or measurement; data transmission methods such as analog and digital, parallel and serial; data transmission modes such as synchronous and asynchronous simplex, half duplex and full duplex; the transmission error control and protocols. All this was intended to lay the foundation for all that is going to follow.

For any basic data communication process, the hardware required are:

- sender and receiver hardware
- communication devices
- communication channels

2.3.1 Sender and Receiver Hardware

Sender and Receiver Hardware are used for handling communication messages, data transfer etc.

Nodes and Workstations

Data Communication is done using various communication devices and software interconnected for information exchange. The devices used to communicate in a data communication network are called workstations.

These workstations may include computer, terminal, printer, and any other communication devices. Each workstation is connected to something called a data communication network node.

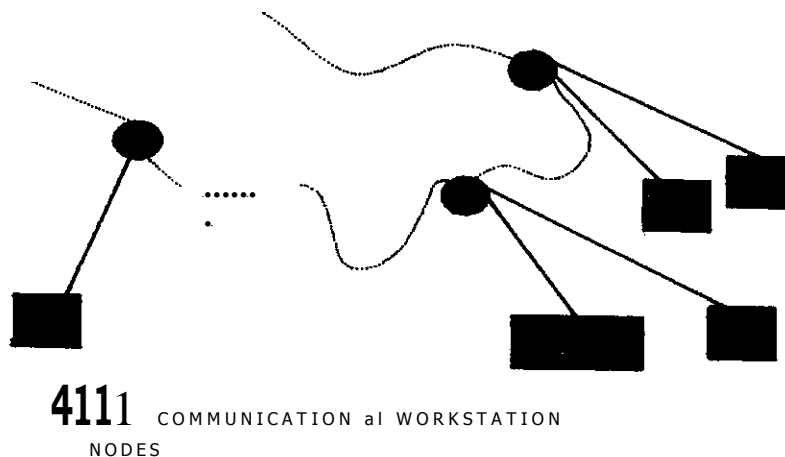


Fig. 4: Communication nodes and workstations

tAitCri__51thn 5_{ae} **dentf** La_ud 1-0 Co /nanny icric. p^re ce- tight c-Conx
n ettm *ark-* ta..cet,a..ok . pc fernatial, en.;11e.4⁻
NODE LA a *ztatft* O_{zyksez}t *arkSialktAis*
Pc(i/f..c

h> 31-ushre Box,

tic dt(ote ree, ne .9f) e cc's' Pc- scPer'''c're -

Multiplexer/

As the name suggests, patiplexing is a forpAclarnsInger in which communication channels are used for ssral trAncrniccions For example, the telephone lines that we use for our daily conversation can carry hundreds or even thousands of conversation using multiplexing. In other words multiplexing is a type of network mechanism, which allows a number of simple, low cost terminals to share each communication line, introduces almost no delay and requires no special computer software.

«Multiplexing is used in two ways:

(a) By dividing a communication channel into various smaller segments of different frequencies.

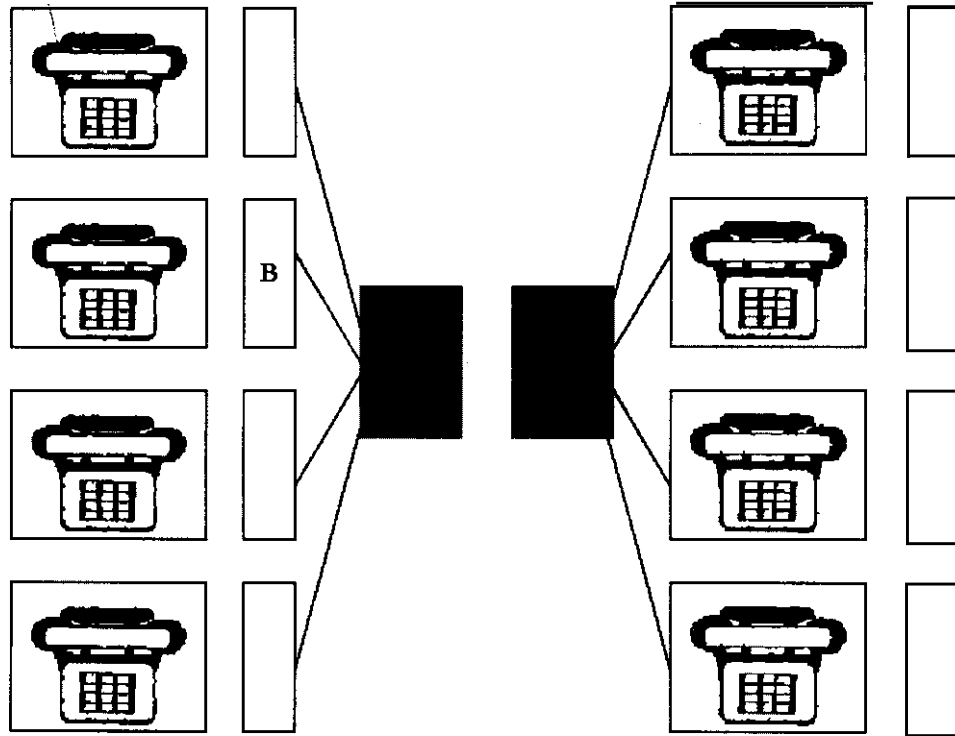


Fig. 5: Multiplexing

(b) By taking groups of bytes from each sender and send or transmit them over the channel one after another. Each group of data bytes are tagged at the beginning and end with start and end control bytes. These, by data bytes, are then separated at the receiving end and sent to their respective places. This type of multiplexing occurs so fast that the transmission seems continuous.

2.3.2 Communication Devices

There are several types of communication devices or interface used in data communication. These interfaces are connection between receiver and sender hardware involved in data communication and the communication network. Some of these interface or communication devices are discussed below.

(a) **Modems or modulator:** demodulators are devices which convert digital signals into analog for transmission facilities such as telephone lines. At the receiving end, a modem performs the reverse function and converts analog signal into digital form.

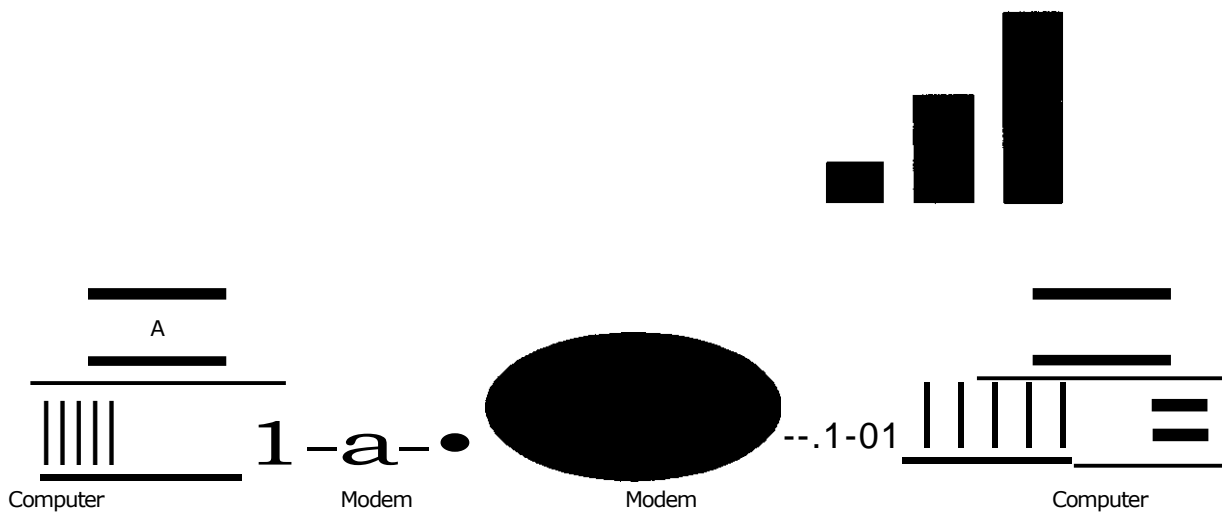
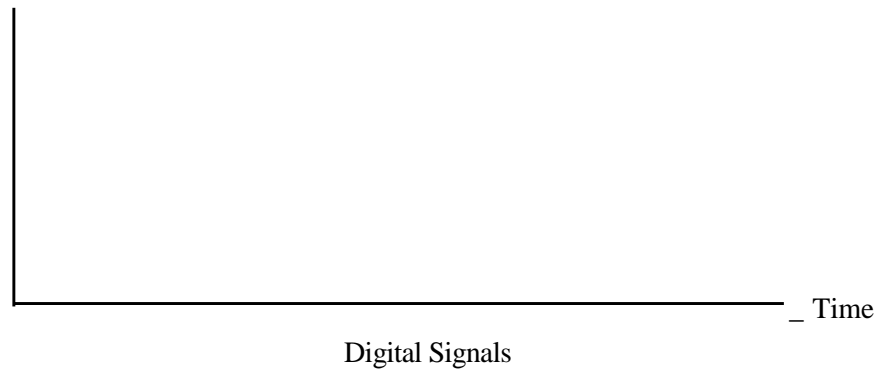
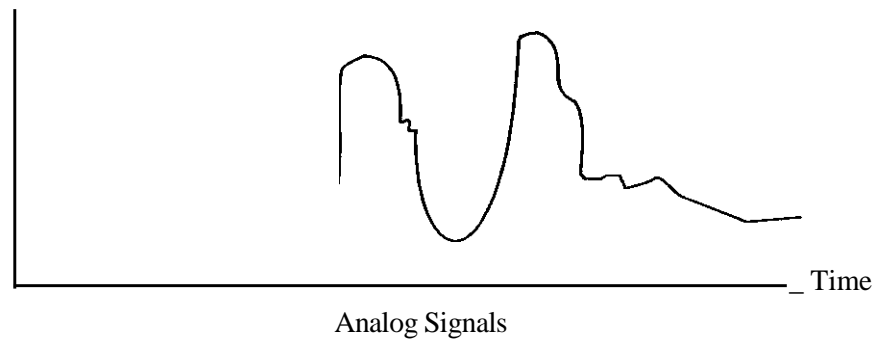


Fig. 6: Data Communication using Modem

Line driver are simple devices used to transmit digital signals over short distances. These devices do not modulate or demodulate signals, but use filters to reduce the high frequency component and the modified signal is transmitted directly over the media.

These devices are used for distances of up to 1 km and can achieve data transfer rates of up to 19200 bps. It is mainly used for connecting VDU terminals with a computer at a distance of more than 100 feet or so.

(b) Codec performs the opposite function of a modem

A communication port/adaptor is a connector on the computer which is used as an entry or departure point for data involved in data Communication. A common type of communication port is RS 232C. This adaptor is used for transfer of data between the computer and the telephone line. During transmission process it adds start and stop bits and controls the communication rate or it strips the start and stop bits where it receives data.

2.3.3 Communication Channels

The most basic hardware required for communication is the media through which data is transferred. There are several types of media, and the choice of the right media depends on many factors such as cost of transmission media, efficiency of data transmission and the transfer rate. We describe some of these transmission media.

Two Wires Open Line

This is the simplest of all the transmission media. It consists of a simple pair of metallic wire made of copper or sometimes aluminum of between 0.4 and 1 mm diameter, and each wire is insulated from the other. There are variations to this simplest form with several pairs of wire enclosed in a single protected cable called a multi core cable or molded in the form of a flat ribbon.

This type of media is used for communication within a short distance, up to about 50 meters, and can achieve a transfer rate of up to 19200 bits per second.

Twisted Pair Cable

A twisted pair consists of a pair of insulated conductors that are twisted together. The advantages of a twisted pair cable over the two wire open lines is, it provides better immunity from spurious noise signals. As the two wires are close to each other, both pick up equal interferences caused by extraneous signal sources and this reduces the differential signal added by the noise.

Twisted pair cable is used for communications up to a distance of 1 Km and can achieve and transfer rates of up to 10 megabytes per second. But, as the speed increased the maximum transmission distance is reduced, and may require repeaters.

Twisted pair cable is widely used in telephone networks and is increasingly being used for data. The hub based Ethernets normally use UTP (Unshielded Twisted Pair) cable.

Coaxial Cable

A coaxial cable consists of a solid conductor running coaxially inside a solid or braided outer annular conductor. The space between the two conductor is filled with a dielectric insulating material. The Larger the cable diameter, the lower is the transmission loss, and higher transfer speeds can be achieved. A coaxial cable can be used over a distance of about 1 Km and can achieve a transfer rate of up to 100 megabytes per second.

A coaxial cable is of two types — a 75-ohm cable, which is used by the cable TV operators and the 50-ohm cable which is used in high-speed broadband networks and are low loss cables.

Fibre Optic Cables

A fibre optic cable carries signals in the form of fluctuating light in a glass or plastic fibre. An optical fibre cable consists of a glass or plastic core surrounded by a cladding made of a similar material but with a lower refractive index. The core transmits the light while the change in refractive index between the core and the cladding causes total internal reflection, thus minimizing the loss of light from fibre.

As light waves have a much wider bandwidth than the electrical signal and are immune from electromagnetic

interferences, this leads to high data transfer rates of about 1000 megabytes per second and can be used for long and medium distance transmission links.

Radio, Microwave and Satellite Channels

Radio, microwave and satellite channels use electromagnetic propagation in open space. The advantage of these channels lies in their capability to cover large geographical areas and being inexpensive than the wired installation.

The demarcation between radio, microwave and satellite channels lie in the frequencies in which they operate. Frequencies below 1000 MHZ are radio frequencies and higher are the microwave frequencies.

Radio frequency transmission may be below 30 MHZ above 30 MHZ and thus the techniques of transmission are different. Owing to the characteristics of the ionosphere, frequencies below 30 MHZ are reflected back towards the surface of the earth. Above 30 MHZ propagation is on line of line of sight paths.

Antennas are placed in between the line of sight paths to increase the distance Radio frequencies are prone toStiott and, thus, they require repeaters along the path to enhance the signal. Radio frequencies can achieve7Fraia transfer rates of 100 ISpbs to 400 Kbps.

Microwave links would invariably use line of sight transmission with repeaters placed at every 100 - 200 kms. Microwave links can achieve data transfer rates of-about 1000 mbps.

Satellite links use microwave frequencies is of the order of 4 -12 Ghz with the satellite as a repeater. They can achieve data transfer rates of about 1000 Mbps.

2.4 Conclusion

You have learnt the functions of nodes, workstations, modem, codes, multiplexers and communication channels.

2.5 Summary

This unit introduced you to various types of communication devices and transmission protocols.

2.6 Tutor Marked Assignment (TM1 , ,

1. Explain data communication modes
2. Mention three communication hardware and enumerate their functions.—) 100% 100%

2.7 Further Reading

William Stalling, *Data and Computer Communication*, Prentice - Hall of India.

Unit 3

Introduction to Computer Networks and Emerging Trends

3.0 Introduction

Today, Computer networks form the backbone of most enterprises big or small around the world. Computer networks allow people remote to the computer to access the information available to that computer. Computer networks are being used to provide resource sharing between systems separated from a few feet to thousands of kilometers. This technology is leading many corporations to take advantage of the reduced process and increased performance in the workplace. In the two units of this module, we have discussed about the data communication, in this unit we will discuss LAN networking and how it plays an important role in information exchange.

31 Objectives

At the end of the unit, you would be able to:

- define LAN
- define topology and its characteristics
- define a Network Interface Unit (NIU) and communication architecture

32 Network Concept and Classification

Communication using computer has brought a revolution in the world of Information Technology, particularly in the field of personal computer. We have always heard of networking or the term network. A network is a way or means of transmitting or receiving (exchange) information from one or more sources.

As an example, car salesmen, after years in the business have developed a network of associates. When the car salesman needs to locate a car to make a sale, the car salesman calls out to his network to retrieve information on the location of the car. Employment agents will frequently keep in touch with their clientele for possible openings or to locate a candidate for an opening. Without the capability of networking, these two people would have a difficult time. It is same in computing. Networks provide the means for locating and transporting information.

In computing networks, the origin of the information request utilizes the services of a network to locate and return the information. This is done with addresses. In the two previous example of the car salesman and the employment agent, a telephone number can be considered the address of their associate or client.

Addresses in computer networking are used in the same manner. These addresses identify the network resource. There are two popular architectures for networking — hierarchical and peer.

Peer networking does not need pre-defined network addressing. Instead, each resource on the network is seen as a peer. Each network resource is a peer to the other network resources. When a new network resource joins the network it introduces itself and notifies its peer of any other network resources that it knows about peer networks are open and share network information.

The entire computer network can be classified into two broad categories (However, elaborate categorization exists). They are:

- (a) LAN (Local Area Network)
- (b) WAN (Wide Area Network)

3.3 Local Area Network (LAN)

As number of systems grows within an organization, a need is felt for sharing expensive resource and exchanging data and information between systems. This need of information exchange and resource sharing within an organization has resulted in development of Local Area Network or LAN.

A LAN is a data communication network, which connects many computers or workstations (computers terminal, printer, etc.) and permits exchange of data and information among them, within a localized area, inside a building, or a cluster of buildings. The distance between two communication points

connected on the same LAN channels is usually up to 02-05 kma.

LANs are not rigidly defined but tend to share most of all of the following characteristics:

- (a) All the connected devices in the network share the transmission media.
- (b) Each device connected in the network can either operate standalone or in the network.
- (c) Area covered is small.
- (d) Data transfer rates are high, usually 1 Mbps-100 Mbps. (Million of bits per second).
- (e) Each device connected in the network can communicate with any other device in network.
- (f) Cost of setting up the network is usually low.

3.3.1 LAN Topology

A network topology refers to the physical layout of the network in which all the devices are connected.

includes all the hardware that makes up the network. The points of connection to the network by the stations are called Nodes or link stations. There are several types of topographical design and strategies used to implement LAN. The majority of these are based on three types of topologies:

- (a) S t a r
- (b) B u s
- (c) R i n g

Each topology has its advantages and disadvantages.

Star Topology

A star topology is shown in Figure 7. In this topology, a number of stations are connected directly to a central station or controller. Communications on the connecting links between the stations and the central station of a star topology can be bi-directional and are point-to-point. A station on this type of network passes an information frame to the central controller, which then forwards the information to the destination station. The central controller handles all communication between stations on the network.

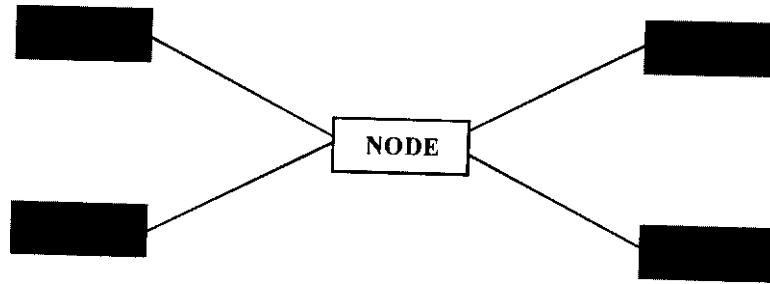


Fig 7: Star Network

Failure of a station on a star network is easy to detect and can be removed from the network. However, failure of the central controller will disable communication throughout the whole network.

Bus Topology

A bus topology is shown in Figure 8. All stations are connected to a single communication line. This single communication line is referred to as a bus. Information frames originating at a station are propagated away from the station in both directions on the bus. Each station on the bus interrogates the information frame destination address field for its own address. If the destination field does not match the station's address, the station discards the information frame back on to the bus. If the destination address matches the station address, it accepts the information frame and processes the frame.

An extension to the bus topology is tree topology and is depicted in Figure 8. Tree topology extends the branches of the bus topology allowing more stations to access the bus.

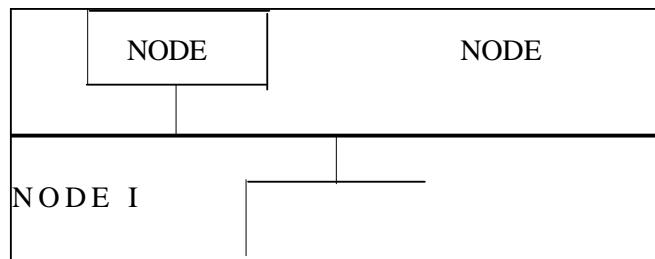


Fig. 8(a)

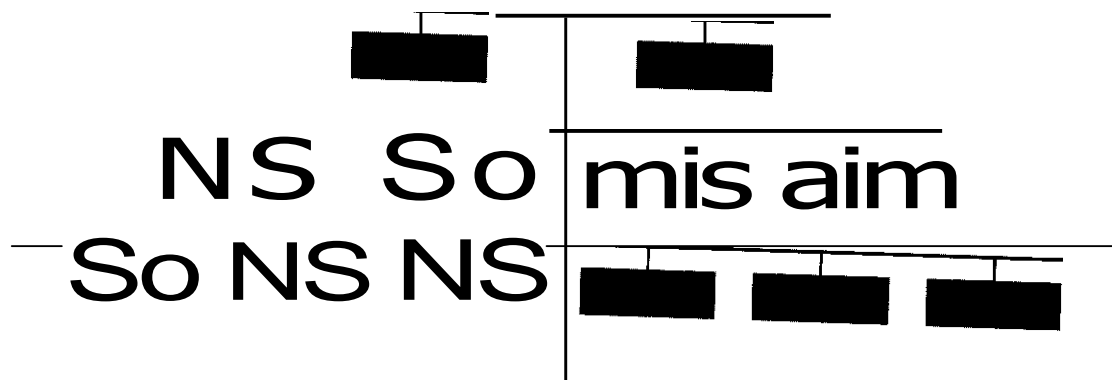


Fig 8: (a) Bus Network (b) Tree Network

On a bus or tree network, there is no central point for management and control. These functions are distributed to each station on the bus. A break in the bus can be difficult to locate but limits the range of communications between stations that traverse the broken point.

Ring Topology

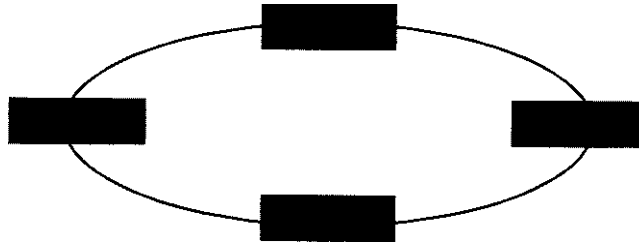


Fig. 9: Ring Network

A ring topology is shown in figure 9. Local area networks that have each station attached to an adjacent station using point-to-point links form a physical ring. Each station attached and active to the ring regenerates the information frame, then re-transmits the information frame on the ring. The ring itself is logically circular and the information travels in one direction.

Failure of a station in a ring topology disrupts the ring because the information frame is not regenerated.

Additions or deletions of stations to the ring can be disruptive, if the changes are not managed properly.

3.3.2 LAN Access Method

A discipline must be imposed on devices connected to the network to ensure a controlled access to the media. Access methods are the means or ways by which stations actually gain the use of the common channel to transmit messages. The right to transmit is an issue only in broadcast where workstations share a single channel.

Many techniques have been proposed, but two of these are commonly used

- (i) Carrier Sense Multiple Access with Collision Detection (CSMA/CD)
- (ii) Token

CSMA/CD

CSMA/CD access method is used with bus networks. The bus operates in a Multiple Access (M.A.) mode. A node is allowed to transmit on the bus, if it senses that the medium is free (carrier sense). Occasionally two or more nodes may simultaneously sense that the medium is free and begin to transmit. This creates a collision, as the contents of transmitted information frames will collide resulting in corruption of the information frame. This collision is detected (collision detect) by the transmitting node. The two (or more) nodes involved then wait for a further short random time interval before trying to retransmit a frame once again.

Token Passing

Another way of controlling access to a shared medium is by the use of a control (permission) token. The control token is passed from one node to another according to a defined set of rules understood and adhered to by all nodes. A node may transmit a frame when it is in possession of the token and after it had transmitted the frame, it passes the token to the next device in a predetermined sequence.

Fig. 10: Token bus

In token passing, a logical ring of all nodes connected to the physical medium is first established and a single token is generated; the control token passes from one node to another traversing the logical ring. The token keeps on circulating the logical ring until it is received by a node waiting to send an information frame. After receipt of the token, waiting station transmits the waiting frames on the physical medium after which it passes the control token to the next node in the logical ring.

For token passing, the physical medium need not be a ring topology, it can be used to control access to a bus network also.

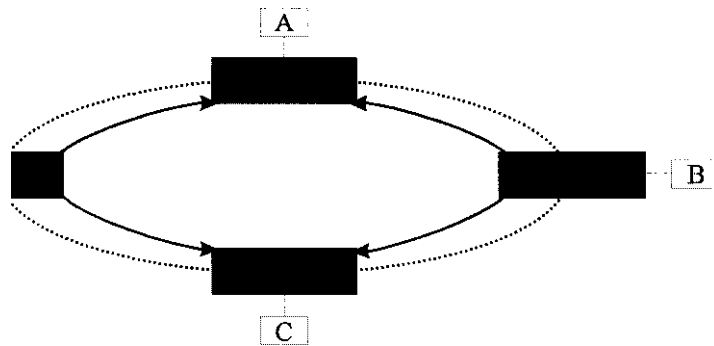


Fig. 11: Ring Token

3.3.3 Communication Architecture for Networks

The task of achieving data communication is a herculean one. Instead of using entire data communication as a single module, the task is broken into various subtasks. And each subtask operates separately by different layers of the communication architecture.

Systems that follow the OSI (Open System Interconnecting) standard are said to be open to one another at different layers and thus called open systems. The communication architecture specifies independent layers that contain modules for performing defined functions. The architecture defines the functions of the module and relationship between them. Each layer in an open system, which communicates with its equal in another open system by using protocols defined in OSI standard

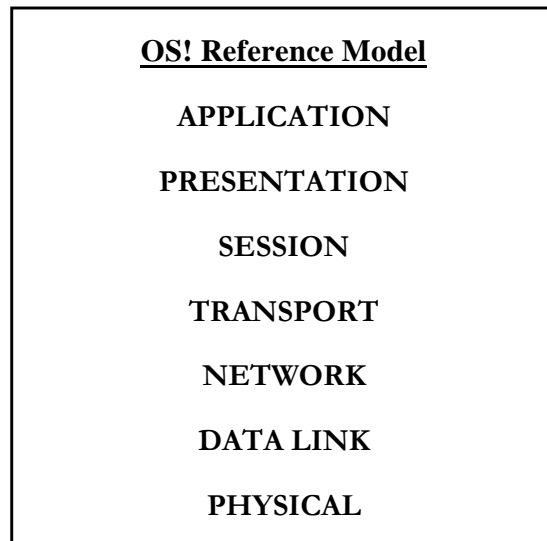


Fig. 12: Seven layers of the OSI Reference Model

The OSI reference model was accepted by OSI as an international standard in 1983. The details on these layers will be dealt with in later courses. However, a brief introduction of these layers is presented here.

Application Layer: This layer supports semantic exchanges between applications existing in open systems.

This layer also provides access to the lower OSI functions and services.

Presentation Layer: This layer concerns itself with the representation of the data to the end user or application. This includes data conversions and code translation (e.g. ASCII to EBCDIC).

Session Layer: This layer provides the mechanism for organizing and structuring interaction between applications and/or devices.

Transport Layer: This layer is responsible for transparent and reliable transfer of data. The lower layers handle the attributes of the transfer medium.

Network Layer: This layer is the agent for establishing connections between networks. The standard also includes operational control procedures for inter-network communication as well as routing information through multiple networks.

7(

Data Link Layer: This layer provides the functions and protocols to transfer data between network resources and to detect errors that may occur in the physical layer.

c

Physical Layer: This layer defines the mechanical, electrical, functional and procedural standards for the physical transmission of data over the communications medium.

As we have seen communication architecture is a common set of rules that define the rules on which all connected nodes in a network should communicate with each other. The communication architecture defines two kinds of relationship between functional modules, interfaces and protocols.

ipjstssfaces aruulef for communicating or exchanging information between dissimilar modules or process,

whereas rules for communication between interfaces of modules of a process.

The following figure explains the difference between interface and protocols.

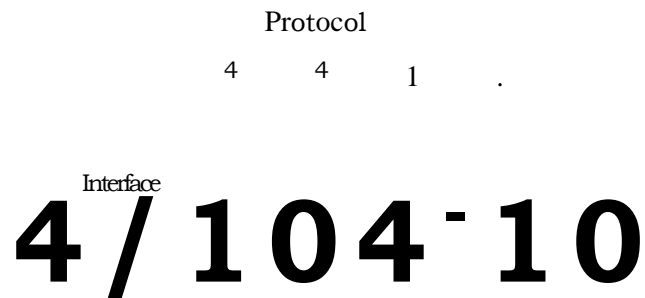


Fig. 13: Interfaces and Protocols

3.3.4 LAN Hardware and Software

As we have seen so far, to realize a LAN process, several functions are to be performed. These are so specialized in nature that they require hardware specially built for such purpose. Here we will discuss briefly the basic hardware components of LAN, these are:

- (a) Transmission channel
- (b) Network Interface Unit (NIU)
- (c) Servers
- (d) Workstations

Transmission Channel

Generally following four types of channels are used for data communication in a LAN. They are already discussed in section 1.5.3. — Cckkk;-4.--) cetsi 01.4 ccotuaArt. RaciCC

Network Interface Units (NIU)

Network interface units connect each device in the LAN network to share transmission device. It contains the rules or logic to access the LAN. NIU is also used to implement LAN protocols and for device attachments. Its function depends on the type of topology used in LAN. In microcomputers, NIU may be installed as an add-on card.

Servers

One of the major benefits of implementation of LAN is sharing expensive resources such as storage devices, printer etc. This is achieved through providing servers in the LAN. It is a dedicated computer, which controls one or more resources. This contains both hardware and software interface for LAN. Three major categories of servers used in LANs are:

- (i) File Server
- (iii) Printer
- (iii) Modem Server

A networking file server is used to share storage space for files. Besides providing storage space for files in a LAN environment, it is used for taking periodical backup, and also provide gateway to other se-rvers within and between LANs. '

Similarly, printer server is used to handle printing works of all workstation connected in the network. In LAN environment also modem is required to get connected to other network or simply to use a telephone. A modem server is used to share few telephone lines and modems by all connected workstations in a network.

3.3.5 LAN Software/Operating System

As the name suggests, LAN Operating System is required to operate on the LAN system, manage the tremendous work load with a number of various Types of server attached to it. It has basically two aspects (i) Server software (ii) Workstation software. As in case of other multi-user operating systems, LAN operating system also facilitates the sharing of expensive resources such as printer, storage space etc. among all LAN users, provides security for data and permits connection to other networks.

There are various types of LAN operating system for example Novel Netware, WINDOWS NT, etc.

3.4 Conclusion

By now you should be able to define LAN, topology and its characteristics. You should also know what a Network Interface Unit and Communication architecture are.

3.5 Summary

In this unit, you learned about LAN concept and its utilization. It defines what is LAN, its topology and hardware and software for LAN.

3.6 Tutor Marked Assignment (TMA)

I. Data transfer rate for LAN is normally:

- (a) 0 1-100Mbps
 - (b) 1 —2 Mbps
 - (c) 1 - 1 0 M b p s
 - (d) None of the above
2. Explain two important features of LAN.

/ ta . • I , 4 j / inn 'TC'
 An at/rt.::
 Cente-C
 L <
 c C t . r-c 14 2008/05/11 14:11:11

3.7 Further Reading

Andrew Tanenbaum, *Computer Networks*, Third Edition, Prentice Hall of India.

Unit 4

Wide Area Network

4.0 Introduction

Impact of networking and data communication have been felt across the globe in various sectors such as education, medicine and transport.

Information sharing has the concept of communication. It has increased access to people in different fields.

4.1 Objectives

At the end of this unit, you would be able to:

- list and distinguish devices used in WAN.
- define the trends and impact of data communication and networking in various fields.
- describe E-mail, EOI and explain the term internet.

4.2 Wide Area Network

As the name suggests, WAN spread across countries and continents, satellite being one of the transmission media.

A Wide Area Network or WAN, is a network that links separate geographical locations. A WAN can be a public system such as the Public Switched Telephone Network (the PSTN) or one of the various packet switched services provided by the public telecommunication authorities. WANs can also use most other types of circuit including satellite networks, ISDN, Value Added Networks (VANs/NADs).

The network can be a private system made up from a network of circuits leased from the local telephone company or set up using public systems as virtual private networks. Private Networks are one which operates in the same way as a private networks but which uses remote services for the transmission of information.

The main distinguishing feature between a WAN and LAN is that, the LAN is under the complete control of the owner, whereas the WAN needs the involvement of another authority like the Telephone Company.

LANs are also able to handle very high data transfer rates at low cost because of the limited area covered. LANs have a lower error rate than WANs.

4.2.1 Communication Switching Techniques

In a WAN, two computing devices are not connected directly. A network of switching nodes provides a transfer path between the two devices. The process of transferring data blocks from one node to another is called data switching.

OA'

86

efrit° 4-7-7 onOur"

it err "(CM_ erg/ 0(ref ntt-e-dttrkt-) lac Ni

Ve.nj Atel 664°3 trbilr-tc cq

QJ

tcio. ced-f-

1000 1000 0000-00 0000-00 0000-00 0000-00 0000-00 0000-00 0000-00 0000-00 0000-00 0000-00

There are three switching techniques commonly employed, and these are:

1. Circuit Switching

In circuit switching, there is a dedicated communication path between the sending and receiving devices. The decision path is a connected series of links between switching nodes. A conventional telephone network, where a dedicated path is set between the caller and the called party for the duration of a telephone call is an example of circuit switching.

Communication via circuit switching involves three steps: Circuit establishment; data transfer; and circuit termination.

Circuit switching is mainly used in telephone networks but is not all that effective for data communication networks, as channel capacities are not fully utilized, as data communication equipments do not generate data continuously.

2. Message Switching

Message switching is an alternative switching technique, where it is not necessary to establish a dedicated path between the sending and receiving devices. In Message Switching, the sending device appends the destination address to the message and passes it to the next switch. The message is then passed through the network from one node to another till it reaches the intended destination. Each switching node receives a message, stores it briefly and then transmits it to the next node. Examples of a message are electronic mail, computer files, telegrams and transaction queries and responses. A complete exchange may consist of several messages.

The basic disadvantage of message switching is the variable delay at intermediate switching nodes.

3. Packet Switching

Packet Switching combines the advantages of message and circuit switching. Packet Switching is functionally similar to message switching, in which data is transmitted in block, stored by the first switching node it meets in the network and is forwarded to the next and subsequent downstream nodes until it reaches the destination. The length of data block is limited in a packet switching network. Typical maximum length of packets is between 128 bytes to 4096 bytes. There are two approaches to packet switching.

- Datagram
- Virtual circuit

In datagram approach, each packet is treated independently and may follow a different path through the network. Packets may be re-ordered, dropped or delivered in wrong sequence. The communication protocols provide the error recovery and sequencing of packets at the receiving device.

In virtual circuit approach, a fixed logical path through the network from the sender to the receiver is established before any packets are sent. This path remains unchanged for the duration of the session. This is quite like circuit switching, but no resources are reserved along the path. Packets are buffered at intermediate nodes awaiting transmission.

4.2.2 WAN Devices/Hardware

The switching techniques utilize the routing technology for data transfer. Routing is responsible for searching a path between two computing devices that wish to communicate and for forwarding the data packets on this path. Devices such as bridges, router and gateways provide this routing function.

Bridges

Bridges are used to connect two LANs that use identical LAN protocols over a wide area. The bridge acts as an address filter which picks up packets from one LAN that are intended for a destination on another LAN and passes these packets on the network. Bridges operate at the data link layer (layer 2) of the OSI model.

BAL.96ES 04-e aakz/ tb conn^s_c f- R 1.-4Ns.

04sics of sce-n-e enslaki

cso-er,c₂jcZ e_{re}...6-ct4ee apant-er? t4 Pth lin^k; /rf^{er} 031' 74.- ara)94-
4-k-4 Lie ra--tizet Prtsr-t p 4.61eV CX co7 it 64 et 'err-nrcertz cni ca,ruity,

Aiefo-ra-rk •

model. ocols, the amount of proce;g g required at the bridge is minimal. If the distance between the two LANs is large, the user would require two identical bridges at either end of the communication link.

Besides a point-to-point link, the intervening communication facility can be a network such as a wide area packet switching network in such cases the bridges need to add X.25 link layer header and trailer.

Routers

Routers can be used to connect networks that mayat,tsljagr. Routers provide connectivity between two LANs or two WANs over large geographical distances. Routers operates at the network layer (layer 3) of the OSI model. All routes participate in a routing protocol to access the network topology, and based on this information routers compute the best route from a sender to the receiver.

For large Wide Area Networks spanning thousands of kilometers, the normal practice is to put network routers at suitable locations to minimize link costs for leased lines and provide adequate reliability from link failures. Networks and other system are then connected to the nearest router.

Gateways

Gateways are used to connect two dissimilar LANs. The tern,igatways_misl_routers are used interchangeably, though there is a subtle difference between the two. A router operates at the network layer (layer 3) of the OSI model, whereas a gateway operates on, the application layer (layer 7) of the OSI model. A gateway is required to convert data packets from one protocol format to another before forwarding it, as it connects two dissimilar ctrks_

While discussing the WAN devices we referredla X.25; what is it? X.25- is a set of recommendation,,by International Telegraph and Telephone Consultative Committee for packet switched network. You can refer to further readings for more details.

4.2.3 Types of Wide Area Networks

The essential purpose of Wide Area Networks, regardless of the size or technology used, is to link separate locations in order to move data around. A WAN allows these locations to access shared computer resources and provides the essential infrastructure for developing widespread distributed computing systems. We will now discuss the different types of WAN, which are commonly used.

4.2.4 Public Networks

Public Networks are those networks, which are installed and run by the telecommunication authorities and are made available to any organization or individual who subscribe to it. Examples include Public Switched Telephone Networks (PSTN), Public Switched Data Networks (PSDN), Value Added Services (VANs/ VADs) and the Integrated Services Digital Networks (ISDN). We would be discussing the main features of these services.

Public Switched Telephone Network (PSTN)

The features of the PSTN are its low speed, the analog nature of transmission, a restricted bandwidth and its widespread availability. As PSTN is designed for telephones, modems are required when it is used for data communication.

The PSTN is most useful in wide area data communication systems as an adjunct to other mechanisms. It is seldom advisable to use PSTN as the sole communication medium for building a network system. Costs are high, as data connections last for a considerable time. Also, the links set up are unreliable and can terminate without warning.

PSTN connections are usually easy to obtain at short notice, and are widely available and cover almost every location where people live and work. PSTN is most useful for occasional user or as backup to private

q f-es Li 0-cif ail carp i ce.f 3 a Li -it-gz 0⁴⁴-44 j "czsi
eX-tat at (cover c-, 1 :1 L nite2
(g) catit's3 cr fin hico& A-int-kit Po 0 2 4 1 t- (ce j
Rc. (A.4t4.3-@it Gcfreci Ce" nec /a atilt S:frui(chr ger/ -

circuits. It is also used for facsimile (FAX) machines.

Public Switched Data Networks (PSDN)

The term PSDN covers a number of technologies, although currently it is limited to Public Packet Switched Networks available to the public. The main features of all PSDNs are their: high level of reliability and the high quality of the connections provided. They can support both low and high speeds at appropriate costs.

Like the PSTN, a PSDN is very useful and adjunct to a private network for backup and occasional access purposes. It can also be used to link computer systems and networks of one organization to several other organizations. PSDN is very popular for connecting public and private mail systems to implement electronic mail services with other companies.

Value Added Services (VANs/VADs)

In Value Added Services, the provider of such services must process, store and manipulate the data that is carried on the network, that is, add value to it. The technique can be used in specific types of business in which it is advantageous to be able to share information with other companies in the same line.

Electronic Data Interchange (EDI) is one area for Value Added Services in which two trading partners exchanging trading documents such as purchase orders, invoices, transportation etc. using electronic means. In India, Videsh Sanchar Nigam Ltd. is a service provider.

Integrated Services Digital Network (ISDN)

The ISDN is a network concept providing for the integration of voice, video and data services using digital transmission media and combining both circuit and packet switching techniques. The motivating force behind ISDN is that telephone networks around the world have been making a transition towards utilizing digital transmission facilities for many years.

Users in shops or small offices can use their digital connection to telephone company for transmitting both voice and data over the same twisted pair cable which connects their telephone. As information from the telephone/PC/Stereo/TV/PABX are all seen as bit streams by the networks switch, they can be switched and transported by the same network.

Private Networks

The basic technique used in all forms of private WAN is to use private (or more usually leased) circuits to link the locations to be served by the network. Between these fixed points the owner of the network has complete freedom to use the circuits in any way they want. They can use the circuits to carry large quantities of data or for high-speed transmissions.

43 Conclusion

By now you are able to distinguish the devices used in WAN. You are now in a position to describe in detail E-mail, EDI and Internet.

4.4 Summary

In this unit you learned about Wide Area Network. You also learned much about types of WAN. Brief description of some well known networks were presented.

4.5 Tutor-Marked Assignment (TMA)

1. Describe two types of WAN.— '—
3. Which of the following networking is suitable for networking within a building?
(a) WAN (b) LAN (c) MAN (d) None of (a) — (c)

3. Protocols are used to communicate between
 - (a) similar module
 - (b) dissimilar module
 - (c) both (a) and (b)
 - (d) None of the above
4. Which of the following topology is least affected by removal/addition of workstations?
 - (a) RING
 - (b) STAR
 - (c) BUS
 - (d) None of the above
5. The maximum length of data in packet switching method is
 - (a) 1 byte
 - (b) 128 — 4096 bytes
 - (c) 10 MB
 - (d) None of (a) — (c).

4.6 Further Reading

A.S Tanenbaum, *Computer Network*, Prentice — Hall of India

Unit 5

Networking Applications

5.0 Introduction

In this unit you will be introduced to the Electronic Mail (E-mail) and Electronic Data Interchange (EDI).

5.1 Objectives

At the end of this unit, you would be able to:

- Describe E-Mail, EDI

5.2 Networking Application

Networking revolution has swept the world and slowly but steadily more and more enterprises are beginning to rely more on various form of electronic data exchange.

5.2.1 E mail (Electronic Mail)

E-mail system is basically used for sending message electronically to individuals or group of individuals in an inter and intra office environment. It requires networks to connect them. In the world of information technology, E-mail is considered to be first experience and exposure to the field of data communication and networking.

But an E-mail system can do more than just send message back and forth. Applications that will be build on the messaging infrastructure include multimedia mail, database access, document sharing, fax routing, scheduling etc. The most promising areas are workflow, deviation support, task automation, information routing etc. Every day E-mail vendors always come up with their new version of E-mail system. Let us discuss few important aspect of a good E-mail systems.

The E-mail should contain the feature to compose and send messages. the ability to easily send and receive messages that can be quickly opened and read by the recipient are important to improve productivity and use of an E-mail system.

- (a) Automatic differentiation of text created when typing, from initial message.

An E-mail system should provide support for OLE object in an embedding, which allows user to incorporate graphics, sound and text into in messages.

Attaching a file or object to an E-mail message is the faster way to route information to a workgroup.

One of the advantages of E-mail is that it gives users the ability to review, respond to a file, and discard incoming message quickly. E-mail is fast becoming more than just a way to route electronics note. It is becoming an important communication medium and infrastructure for workgroup applications that make it easier for people to work together.

5.2.2 EDI

Electronic Data Interchange is the inter-organizational exchange of business documentation in structured, machine-processable form.

EDI is often viewed as simply a way of replacing paper documents with electronic documents, and replacing traditional methods of transmission such as mail, phone, or in-person delivery with electronic transmission. However, EDI is actually a way of replacing manual data entry with electronic data entry.

purpose-ac The purpose of EDI is not to eliminate paper, but rather to eliminate processing delays and data re-entry. Electronic data interchange can be used to electronically transmit documents such as purchase orders, invoices, shipping notices, receiving advices, and other standard business correspondence between trading partners. EDI can also be used to transmit financial information and payments in electronic form. When used in this application, EDI is usually referred to as EFT, Electronic Funds Transfer.

What is Application-To-Application EDI?

Figure shows the use of EDI in place of traditional methods for the transmission of a purchase order between a buyer and a seller and demonstrates the key concept behind EDI. Once data are entered into the buyer's computer system, the same data are electronically entered into the seller's computer, without the need for rekeying or re-entry. This is normally referred to as application-to-application EDI. When EDI is fully integrated with application programs, not only do data flow electronically between trading partners without the need for rekeying, data also flow electronically between internal application of each of the trading partners.

The repeated rekeying of identical information in the traditional paper-based method business communication creates a number of problems that can be eliminated or significantly reduced through the usage of EDI. These problems include:

- Increased time
- Low accuracy
- High labour charges
- Increased uncertainty

EDI has become a major means of business communications among large companies in the U.S. EDI can do, for communications between unrelated companies, what Electronic Funds Transfer (EFT) has done for transactions between large financial organizations-substitute electronic transactions for paper ones.

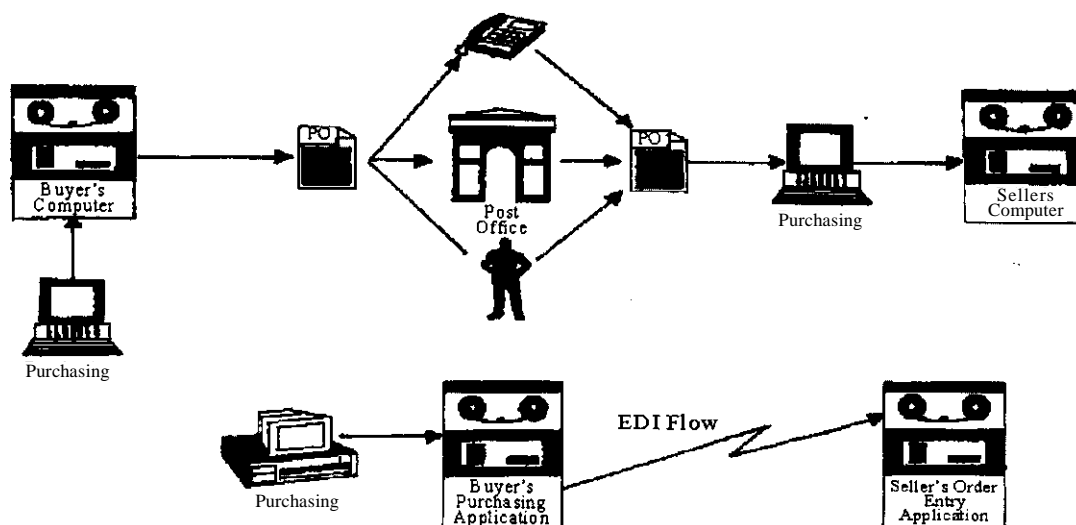


Fig. 14: EDI vs traditional methods

EDI consists of standardized electronic message formats (called transaction sets) for common business documents such as Request for Quotation, Purchase Order, Bill of Lading, Receiving Advice, Invoice, and similar documents. These electronic transaction sets enable the computer in one company/organization to communicate with the computer in another company/organization without actually producing paper documents. The human effort required to read, sort and physically transport such documents is eliminated. The documents just mentioned, for which standard EDI formats are either in existence or under development, constitute about 85% of the official communications associated with commercial transactions between business, government educational institutions, and non-profit establishments in U.S and most of the industrialized world.

To take full advantage of EDI's benefits, a company must computerize its basic accounting records. Trading partners are individual organizations that agree to exchange EDI transactions. EDI cannot be undertaken unilaterally but requires the co-operation and active participation of trading partners. Trading partners normally consist of an organization's principal suppliers and wholesale customers. Large retail stores, because they transact business with a large number of suppliers, were among the early supporters of, and participants in EDI.

Benefits of EDI

The use of EDI eliminates many of the problems associated with traditional information flow.

- The delay associated with order making are eliminated, time required to reenter data is also eliminated.
- Since data is not repeatedly keyed, the chances of error are reduced.
- As data is not re-entered at each step in the process, labour costs can be reduced.
- Because time delays are reduced, there is more certainty in information flow.

The other advantage in the use of EDI is that it generates the functional acknowledgement whenever an EDI message is received, and it is electronically transmitted to the sender. This acknowledgement states that the message is received.

Therefore, the core concept of EDI is that data are transferred electronically in machine processable form, i.e. the EDI message can be immediately processed by a receiving computer without, any human intervention, or interpretation or keying. Therefore, EDI is most suited in the areas where any of the following characteristics exist:

- Handles a large volume of repetitive standard action.
- Operates on very tight margin
- Faces strong competition requiring significant productivity improvements.
- Operates under time constraints.

EDI eliminates the paper documents associated with common business transactions. Consequently, the handling, filing, and transportation necessitated by the existence of the paper documents are also eliminated. Electronic documents (messages) can be duplicated and routed (transmitted) instantly to anyone in the organization with a need to see them. Where a hard copy of a document is desired or required by law, a paper copy can be produced. Space that normally would be occupied by files of multiple paper copies can be devoted to more productive use and the manual filing operations eliminated entirely.

All of the above benefits result in a more efficient operation and usually provide identifiable cost savings to the company that implements EDI. In addition, eliminating the time required for the preparation and physical movement of paper documents speeds up the entire process of information transfer between companies.

These benefits are so compelling that companies must soon adhere to EDI standards if they expect to sell to large U.S organizations such as Fortune 1000 companies, where the volume of these documents is always burdensome. The alternative will be to send paper documents to a third party (service bureau) where the document is converted to an EDI message acceptable to the addressed company. Such service bureaus are already in operation in U.S.A. It is this kind of computerization which is forcing India as a Country to adopt

ED! technology for international transactions.

53 Networking Scenario

We have seen what electronic data communication is and how it plays a crucial role in the success of organizations. Its impact has been felt across the globe, let us take a look at some of the networks.

5.3.1 Internet

Internet is the world's largest networks, originated out of a US department of defence funded project. It is a unique collection of networks with vast proportions of its own kind. It has evolved into one of the technologies greatest democracies, permitting the passage of all kinds of information exchange with full freedom. The first result of the project funded by US defence department for a fault tolerant networking system, is Arpanet, which in time became the largest, the most potent and definitely most uncontrollable force in the world.

The networks or computers operating in different platform are connected to Internet by a common protocol known as TCP/IP Transmission control protocol/Internet protocol. Internet provides the following services.

(a) Global Electronic Mail

As we have already discussed e-mail permits user to send and receive messages electronically to an individual or a group. Internet mail makes delivery of area of message more wide and reliable.

(b) USENET-Views and News

Views and News or USENET is the BBS (Bulletin Board Service) of Internet. The messages in this BBS are organised into thousands of News groups, which cover specific areas of interest.

(c) TELNET-Remote Login

TELNET-Remote Login

Telnet allos.ailInternet user to access a remote host. After properly connecting and logging into the remote host, the user can enter data, run programs or do any other operation.

(d) FTP — File Transfer Protocol

It permits an Internet user to move or transfer a file from one computer to another even if they are running on different platform (or operating system). The files may have data, graphics, text etc.

More than 1000 new users are joining Internet each month making it a global information ocean. This also makes difficult to find any thing specific on Internet. There are several powerful tools used on Internet for searching information. These information-tracking utilities are means to develop easy method of discovering, locating, and retrieving information on various objects freely available on Internet. Some of the well-known utilities are Gopher, Archie, Hytelnet, WAIS, and WWW. The WWW out of these is the most popular and have resulted in the massive growth of Internet resources. You can obtain more information on Internet in further readings.

5.3.2 BITNET (Because — Its Time Network)

City University of New York and University of Yale established BITNET basically to exchange information between universities. The basic services provided by BITNET are:

- (a) Exchange of electronic messages (datafile)
- (b) Electronic student admission
- (c) Remote job entry and storing research information.

A counterpart of BITNET in India is ER Net, which is established for educational and research purposes.

5.3.3 CompuServe

CompuServe is a commercial network based in US. It uses telephone lines and microwave for communication. CompuServe user can easily communicate with each other, around the world. Following are the services offered by CompuServe.

- (a) Electronic Mail box for message transfer
- (b) Bulletin Board Services
- (c) News report along with report on sports and weather
- (d) Information on computer hardware and software.

5.3.4 ISDN (Integrated Services Digital Network)

ISDN is basically used for Communication of both data and voice. ISDN was first launched in Singapore in 1990. It has the following features:

- (a) Transmission and processing of digital data
- (b) Processing of various types of information data, voice, video, etc.
- (c) Electronic mail box
- (d) Tele-conferencing
- (e) Tele fax, video fax, etc.

5.3.5 NICNET

NICNET is a satellite-based nationwide network of **NIC (National Information Centers)**. **The basic idea** of interviewing NICNET is to extract data from each village, district, ~~acittounclia~~. **It consists** of earth of stations in almost all districts, state capitals, regions and **NIC headquarter in Delhi**. **The basic objectives** of

NICNET are:

- (a) To help government in better planning administration
- (b) To help government maintain communication in times of national emergencies **and natural disasters**.

In India, many other Government owned network exist.

54 Conclusion

By now you know E-mail and EDI are operated.

5.5 Summary

Networking revolution has swept the world and slowly but steadily more and more enterprises are beginning to rely more on various form of electronic exchange. E-mail system is basically used for sending message electronically and EDI is the interorganizational exchange of business documentation in structured machine — processable form.

5.6 Tutor-Marked Assignment (TMA)

1. Write a short note on E-mail
2. Explain the Application-to-Application ED!
3. Write a short note on Internet.

5.7 Further Reading

William Stalling, *Data and Computer Communication*. Prentice-Hall of India.

Module 4

The Management of Computer Security and Principles of Cryptography

Introduction to Module 4

In this module you will learn how to manage a computer. In particular, here you will be introduced to Computer Security and Principles of Cryptography.

Unit I

The Management of Computer Security•

1.0 Introduction

In this unit we will describe the various ways in which loss or corruption of data can occur and the means by which this can be avoided. Our main thrust would be directed towards PCs, but we would also look at the measures available on larger machines.

Once more and more information is kept in digital form, the protection of data in computer systems begins to pose challenges to designers, researchers and system managers. It is clear that when such data involves financial transaction, there has to be complete reliability. Even in University examination systems, the security of the data is very important. In an information society even for business the competition would be on the basis of the information available. Therefore, Sing of information or eavesdropping when data is being communicated from one part of the organization to another, can be a serious threat.

11 Objectives

At the end of this unit, you should be able to:

- define the computer security
- give various security measures to be undertaken

12 Definitions

Generally the terms privacy, integrity and confidentiality are loosely construed to be synonymous with security. These however have different connotations with respect to data or information. They also address different areas of information systems. To better understand the measures and to ensure protection in each area, let us see their definitions.

Security ata_or information **seciarity_is-thappacctinal** or intentional, destruction, ,
•idos= or modification. Computer data security refers to the technological safeguards and managerial procedure, which can be applied to computer hardware, software and data to ensure that organizational assets and individual privacy are protected.

Privacy: Is a concept applied to an individual. It is the right of an individual to decide what information he/she wishes to share with others or is willing to accept from others.

13 Security Status On PC

Before studying the ways in which security can be compromised, let us see what some of the leading magazines have to say about PC security.

The "New Scientist" in its issue of 7 July 1983 warned: "New PC users beware! PCs are the biggest
I g \ M t 14 of n-310- aol e s,erkisuuctu s-Pfritz) de_c-Lcife castal- 11 4 At a.

tic-€

cs: cceel- farn.---.% c1f-44

threat to Computer Security. Micros are left in unguarded offices or at home, where data snoopers may steal confidential files or data, if not the machine itself".

A few years later "PC Week" in its issue of 7 May 1985, Cautioned: "The PC has NO inherent security". For a user with confidential data, the data stored in the PC is vulnerable because anyone can walk up and turn the switch ON and access information.

The same article had this to say for networked PCs: "Put your PC online and you are open to a new world of terrors. A LAN is an open opportunity for mischievous or disloyal coworkers to get confidential information. A telephone connection can invite everyone from the 10-year old down the street to the KGB to interpret your communiqués and romp through your memory banks, borrowing a file here, erasing others there, until the only memory you have of your data is an ulcer".

The article Digital Defence in the February 7-2 1,1995 issues of "Business Today" has the following to say "If you marvel at the speed, efficiency, and ease-of-use with which your computer system crunches, sorts and spews out data, remember this: it's just as simple for a digital desperado on the prowl to coax out the same data from the system".

14 Breaches of Security

The above warnings paint a fairly bleak picture for PC users. Some of the ways in which data loss or manipulation can occur have been hinted at in these articles. Let us look at the details of the manner in which losses can occur.

Thefts and Media: May sound preposterous but it is a distinct possibility. A smart person with a false calling card can take away the PC for repairs and of course never show his face again! However, electronic media like floppies and CD-ROMs are slightly safer as it is far easier to lock up floppies in a safer place.

Damage due to Breakage: Floppies are easily breakable. It is hard to visualize dropping PCs but it can happen if they are shifted from one place to another. More likely is that something may get dropped on the PC resulting in damage. Damage can also occur due to natural causes such as storms or floods, or due to electrical or other fires.

Environmental Damage: The manufacturer recommends some environmental conditions like temperature and humidity ranges, voltage limits, dust micron limits etc. If the conditions in your office remain outside — these limits the PC and media are likely to get damaged.

Inadvertent Corruption/Loss: This can occur due to:

- **Usage of Inferior Media:** If sub-standard media is used, as it would be generally cheaper, after using it for some time it may develop faults and data held on it may become unusable. One hears about frequent corruption of data on inferior floppies.
- **Erasures of Files:** Files may get erased from the media due to incorrect actions by the operator. Corruption may occur due to the PC being subjected to frequent power failures, wrong programming techniques or defective software.

Environmental Losses: Excessive dust or humidity can result in corruption of disc surfaces or read/write heads resulting in loss of data.

Malicious Damage/Leakage: We now turn to the real problem of computer installations, be they stand alone PCs or large main frames with hundred of terminals. It is not necessary that outsiders would do this; it is equally possible that some unhappy or impish insiders may wreck havoc.

Unauthorized Access: As the saying goes "Curiosity killed that Cat", but it does not stop the human trying to look at things they should not or need not. Information on personnel, finance, products or assets can be accessed and copied for malafide use.

Modifications Erasures etc: The person accessing data files may be authorized to read the data only, but he would like alter, modify or erase the data by writing into that file.

Computer Viruses: This is the latest threat to computer users. These can be introduced deliberately or unknowingly by anyone at anytime and the consequences to the user would be equally disastrous. The

problems created by viruses include:

- Destruction of file Allocation Table (FAT) — user loses everything in his media;
- Erasing of specific programs and/or data on discs;
- Alter contents of fields in the file;
- Suppress execution of RAM resident programs;
- Destroy parts of programs/data held on disc by creating bad sectors;
- Reduction of free space on disc;
- Formatting of discs or tracks on discs in a different way;
- Overwriting of entire disc directories;
- "Hang" the system at periodic intervals so that keyboards become inoperative;
- Automatic copying of results obtained by other programs into some designated areas. Data

Tapping: In large computer systems or when systems are networked, data has to travel between the processing unit and terminals, or different processors, over communication lines. Any person trying to get access to data can intercept the traffic on the circuit by tapping into the cable at convenient points. This may even enable him to send spurious channel over the network — access the computer itself by emulating terminal responses. Very sophisticated means have been developed which allow a person to 'listen' to the traffic on a line even without physically connecting into it. Thus the data following over communication lines is ever susceptible to "eavesdropping".

15 Conclusion

By now you are familiar with security status on PC and breaches of security.

1.6 Summary

This unit discussed security status on PC and breaches of security. The unit also listed some damages and data tapping.

1.7 Tutor-Marked Assignment (TMA)

1. Define the terms security and privacy
2. Make a list of four breaches of security

18 Further Reading

Milenkovic. *Operating System: Concepts and Design*, Second Edition, Tata McGraw — Hill

Unit 2

Security Measures

2.0 Introduction

In this unit, we will discuss measures for data protection. In particular we discuss physical security and software security. You will also learn about Network Security and Password Security.

21 Objectives

At the end of this unit, you would be able to:

- explain physical security and software security.

22 Security Measures

The measures for data protection taken by an organization reflect its awareness and attitude towards information and Information Technology. If top management treats computers as a dehumanized, intangible, but necessary evil, the measures taken to protect data, individual privacy and data integrity would, at best, be lackadaisical. On the other hand, if the management considers information as an important resource and computers as an aid decision making one would find a positive approach and involvement by the management towards security of information. This attitude naturally percolates down to the lower levels and the workers consider the computer correspondingly as an enemy or an ally.

One of the best and first steps in ensuring data security is to create an awareness and develop a culture within the organization towards the ways in which information can be lost/altered and what would be the consequences, of such an occurrence, to the organization and individuals. The other steps that can be taken are:

- IT Planning: The organization must decide on a policy for introduction of IT. This must be done at the highest level and should address issues such as level of protection for various aspects of information relating to the organization;
Selection of Technology: Keeping in mind obsolescence due to new innovations and necessity for keeping in step;
- Identification of points of exposure of weak links to device means to plug them;
- Physical protection of machine and media.

Control and monitoring the access to data, its usage by persons and its integrity must be clearly defined and responsibility for ensuring these must rest on persons designated for these tasks; an audit procedure would go a long way in ensuring adherence to laid down guidelines. While the above are relevant for any computer based MIS implementation, in case of PCs, the rules for acquisition and use must be unambiguously stated.

Additional points to be looked into are:

- Information classification;
- Responsibility for security;
- Using training to increase security awareness and propagation of "do's and don'ts"
- Guidelines for creation and changes to passwords etc.

There are four principles for ensuring security and recovery in case of breaches of security:

Prevent: The best method is of course—stopping all breaches of security before they occur. 'Need to know' policy is an offshoot of the principle of prevention.

Detect: However, one may try to ensure it, total security is almost impossible. The next principle, therefore, is that you must be able to detect breaches to security, whenever they occur, within the shortest possible time. This helps in damage assessment and also, in devising further preventive measures.

Minimise Damage: The aim here is to contain the damage, when losses occur, to reduce the adverse effects of such damage.

Recovery: There must be enough resilience in the system to recoup the losses/damage and become functional, by reinstating the status, at the earliest.

We would now look at the various measures available to the PC user, to ensure security of machine and data, relating to the principles enumerated above.

1.2.1 Physical Security

These measures are for PCs being used in offices. The PC may be in use by an individual or being shared between two or more users. The measures available are:

- Physically bolt down the PC to a table so that it cannot be casually lifted and taken away;
- Locate the PC in such a way that it is conveniently accessible to the user, but hidden from casual passers — by;
- Have likeable cupboards for floppies and keep them locked at all times, except when used;
- Keyboard and PC locking devices can be fitted so that the PC cannot be operated unless these locks are opened;
- Keep a record of all floppies in use; do not permit alien floppies into the organization;
- Use lockable rooms for PCs, specially those handling sensitive data. Make it a practice to lock the room when leaving it for even a short time;
- The above apply to server, gateways and the like.

Environmental Conditions: The PCs are fairly rugged and can tolerate wide ranges of temperatures, humidity and voltages. However, to ensure trouble free and prolonged life, consider the following measures:

- Have temperature and humidity gauges placed in the close proximity of PC; and keep a casual watch to ensure that conditions are within limits. Switch off if the limits are exceeded.
- If your normal electrical supply is subject to large variations of voltage and frequency or spikes, it is prudent to have voltage and frequency stabilizers for the PC;
- Ensure that excessive dust or paper scrap does not accumulate near the PC;
- The plug sockets should fit snugly and cables leading to terminals and printers should be secured properly and not left hanging;
- You may consider putting a thin transparent plastic cover on the keyboard if it does not hamper your handling the keyboard;
- The most important is the use of a vacuum cleaner at regular intervals.

2.2.2 Software Security

As is apparent from the views on security of various leading magazines provided on PCs, there is hardly any security provided on the PC. There are some measures you can take to ensure that data is not corrupted or modified by unauthorized users and to reinitiate the database to its known status in case this happens and these are:

- Use original software for Operating System, compilers or software packages. You may have to pay for it, but you can then be sure that it would be bug-free, known also as "licensed" software;
- Use correct procedures for shutting down the PC so that all files etc. would be properly closed;
- If you develop your own applications introduce passwords to access your application; these passwords should not be visible on the screen when keyed-in;
- Keep backups of all your files. Whenever you operate on any file, (especially in update/append/alter mode), if you have your own programs — they should include a "copy" procedure; this ensures that a back up of your data file would always be automatically taken.

2.2.3 Network Security

The protection required for networked systems is much more extensive as physical security measures are totally inadequate. It is also extremely difficult to know-who, when and how someone is accessing your data. In LANs generally, there would be one server which holds the shareable data on network **and** services the requests of various nodes. The normal method used is password identity for permitting access, the measures that can be adopted for additional security are:

- Keep the servers away and limit physical access to them.
- Run servers in the background mode, thus the server can be booked on, for use in the network, but, for direct use of the server, a separate password would be necessary.
- Some networks provide auditing facilities, which can be used to advantage;
- Be aware that the network cables can be tapped, so try and shield or conceal them to prevent easy access; if possible use optical fibre;
- Use codes and ciphers in data communication, remember, however that this would impose considerable overheads on your resources;
- Use fibre-optic cables for highly sensitive networks as they are difficult to tap; however here too it may be possible to steal data through sensing the perturbations of the fibre itself;
- Prohibit the use of passwords embedded in communication access scripts; if this is unavoidable, use encryption for passwords;
- Consider the use of see-through devices for any system accessed through networks or through dial up.

Protection against Virus: A number of measures are available for reducing the risk of being attacked by computer virus:

- Build employee awareness of the risk;
- Do not allow the use of outside programs on company PCs or networks;
- Do not interface company networks to outside **"Bulletin Boards"**;
- Make system/server files "Read only";
- Try and obtain source code for important software in use and compile it in house;
- If source code is difficult to follow, it should ring a warning bell in your head;
- Check executable code using "debug" or separate utilities to study code structure and check spaces for viruses.

2.2.4 Password Security

In most organization or computer systems, the only authorization for data access is giving the correct password; rightly speaking, this is only one step; the whole process would be:

Identification: An identification user code only indicates an object with a unique identity assigned to it. Thus it should not become authorization to access data without further checks, if some measure of security is desired;

Authentication: This process verifies that a person or objects is who he, she or it claims to be. This could be achieved by asking some standard questions (from a large selection) and getting answers to them. If the answers match with those held on the systems, the person or object is authenticated. Biometric and other physical authentication processes are also popular in systems where security is a primary concern.

Authorisation: This is the last step in the process. Through this you can ensure that only a given user, terminal or other resource can access data to which permission has been granted to read, write or alter. Thus a matrix can be created to indicate which users have access to which file, records or fields. If the user request passes the matrix he is allowed access, otherwise he is denied access to some parts of the database.

Other Aspects

We have had a fairly close look at the measures for data protection available on stand alone as well as networked PCs. Some of the measures that we studied can be implemented only on mini and mainframe systems easily, while trying to introduce them on PCs may incur too much of resource overheads. We would now take a quick look at the protection, detection and recovery mechanisms available on large systems. This is in order to give you pointers for discerning when to go in for a larger system rather than a PC LAN and what facilities to look for.

Database Access: Larger systems provide various mechanisms to prevent access to data. User classes can be defined automatically prohibiting access to data by user class. User can be given only "query view" of the data so that he can have only "read" access to a limited amount of data. In some systems, certain terminal numbers can display or access only some parts of database, thus, even a user, with higher access permissions cannot access some data on those terminals.

Access to Operating Systems: In some systems the operating system is written in a lower level language and users are not the use of that language. Thus, the user cannot alter any part of the operating system. Some operating systems follow the concept of access control levels. In this any program which has equal or higher access control level cannot access any routines which are below that level. The operating system routines are placed at much lower level and paths are predefined for access to these, which incidentally, are via other system routines placed at a high level. From this point of view 'UNIX' is not a secure Operating System as 'C' which is the language in which 'UNIX' is written, is also available to the user as a programming language, however, it have many good security features.

Access Control Cards: This is the latest method and is also available on PCs. Here an additional card is inserted on the PC. This card has its own memory and software. The user can program upto ten complex account codes. Anyone wanting access to a PC has first to pass through authentication routines through this card. Only when he passes, he is allowed to access the PC itself. These codes can be reprogrammed whenever required. Thus the basic problem of preventing access to the operating system of the PC can be solved to a large extent.

23 Conclusion

Security could be discussed under various headings and these include Physical Security, Software Security, Network Security and Password Security.

2.4 Summary

This unit discussed various headings of security.

2.5 Tutor.Marked Assignment (TMA)

1. Write a short note on Physical Security
2. Write a short note on Software Security.

2.6 Further Reading

M. Milenkovic, *Operating Systems; Concepts and Design*, Second Edition, Tata McGraw — Hill.

Unit 3

Cryptography

3.0 Introduction

In this unit you will be introduced to the science of cryptography. You will learn how to transform plain text or original information into an intelligible (for one who has not been initiated) form (cipher text) so that it may be sent over unsafe channels of communication.

31 Objectives

At the end of this unit, you would be able to:

- describe some approaches to cryptography
- define the concepts of cryptanalysis

32 Cryptography: A Brief History

In the current scenario of computer usage that predict that networks are likely to be present everywhere, security has become much talked about issue. Especially, with the relatively easy access to the Internet, the confidence in the sanctity of the data is a major concern.

However, the science of cryptology is at least 4,000 years old. Cryptology is defined as "the design and analysis of codes and ciphers." In its first 3,000 years, Cryptology was developed independently in several ancient cultures, including Egypt, India and Mesopotamia. Perhaps the oldest evidence of Cryptology is an ancient inscription, carved about 1900 BC, in the main chamber of an Egyptian tomb. It used some unusual hieroglyphic symbols in place of more ordinary symbols. Although the inscription utilized one of the important methods of cryptography transformation of the text-it was not secret writing. The intention appears to have been to impart authority and dignity to the writing.

However, rather better known technique is attributed to Julius Caesar, who used a simple cipher system more than 2,000 years ago to conceal military information. The method consisted of replacing the letters of alphabets in the original text by letters that are a fixed number of places away. For example, 'A' could be replaced by 'D' 'B' by 'E' and so on. Thus, the key for the cipher would be as follows:

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ

Cipher: DEFGHIJKLMNOPQRSTUVWXYZABC

Using this system, the secret message ZHOFRPH WR WKH FFtPSXWHU FRXUVH" would actually mean "Welcome to the Computer Course". There were of course other systems developed around the

Unit 3

Cryptography

3.0 Introduction

In this unit you will be introduced to the science of cryptography. You will learn how to transform plain text or original information into an intelligible (for one who has not been initiated) form (cipher text) so that it may be sent over unsafe channels of communication.

31 Objectives

At the end of this unit, you would be able to:

- describe some approaches to cryptography
- define the concepts of cryptanalysis

32 Cryptography: A Brief History

In the current scenario of computer usage that predict that networks are likely to be present everywhere, security has become much talked about issue. Especially, with the relatively easy access to the Internet, the confidence in the sanctity of the data is a major concern.

However, the science of cryptology is at least 4,000 years old. Cryptology is defined as "the design and analysis of codes and ciphers." In its first 3,000 years, Cryptology was developed independently in several ancient cultures, including Egypt, India and Mesopotamia. Perhaps the oldest evidence of Cryptology is an ancient inscription, carved about 1900 BC, in the main chamber of an Egyptian tomb. It used some unusual hieroglyphic symbols in place of more ordinary symbols. Although the inscription utilized one of the important methods of cryptography transformation of the text-it was not secret writing. The intention appears to have been to impart authority and dignity to the writing.

However, rather better known technique is attributed to Julius Caesar, who used a simple cipher system more than 2,000 years ago to conceal military information. The method consisted of replacing the letters of alphabets in the original text by letters that are a fixed number of places away. For example, 'A' could be replaced by 'D' 'B' by 'E' and so on. Thus, the key for the cipher would be as follows:

Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ

Cipher: DEFGHIJKLMNOPQRSTUVWXYZABC

Using this system, the secret message ZHOFRPH WR WKH FRPSXWHU FRXUVH" would actually mean "Welcome to the Computer Course". There were of course other systems developed around the

similar time in history which included approaches such as:

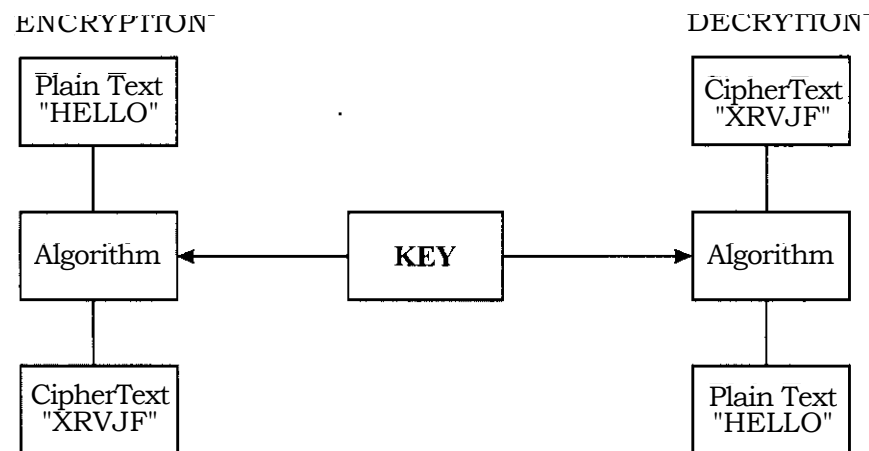
- Writing secret messages backwards.
- Writing messages vertically rather than horizontally.
- Substituting dots for vowels.
- Using alphabets of other languages such as Greek or Hebrew.
- Substituting special symbols for the normal letters of the alphabets.

A Caesar type encryption mechanism can be easily detected by trying all possible displacements till the message becomes meaningful.

3.3 Cryptography

Cryptography is the process of transforming plain text or original information into an unintelligible form (cipher text) so that it may be sent over unsafe channels of communication. A data string (key) controls the transformation process. Anyone getting hold of the cipher text while it is on the unsafe channel would need to have the appropriate key, to be able to get to the original information. The authorized receiver is assumed to have that key. This is illustrated in the following figure:

Encryption Process



/ The stage of sor_ersion of tl_ts_plain text intoaamtogram is called Encrypting or enciphering or encoding. Reconverting the cryptogram back into the original form, when done by the authorized person is called t decrypting or deciphering or decoding.

3.3.1 Cipher Systems

Although there may appear to be many variations, there are mainly two basic classes of cipher systems. These two classes are called transposition and substitution.

Transposition Cipher

A transposition involves rearrangement or change in the sequence of the letters of the plain text message without any change in their identity. However, the substitution involves a replacement of the plain text letters by other letters (or other symbols) without any change in their sequence. Transposition and substitution may be combined in a single cryptosystem.

The Caesar Cipher mechanism was described earlier and is a specific case of what may be called a

"transportation type cipher".

Code System

A code system is a specialized form of substitution in which entire words, long phrases or even sentences of the plain text are replaced by arbitrarily selected equivalents. These may be other words, groups of letters, groups of figures or some combination of these. It is only in rare cases that the substitution process is applied to elements smaller than whole words.

A code system, therefore, makes use of a codebook in which the words, phrases and sentences of the vocabulary are listed in an organized manner and accompanied by their equivalent code groups. Many large commercial firms have their own private codes, constructed especially for their use.

In modern times, communication does not take place only through the passage of text. The same general ideas of substitution and transposition, which were used for literal cryptosystems, are also used for encryption, for speech (ciphony), facsimile (ciphax) and television (civision). In literal cryptosystems, the unit of encryption is usually a single character. But in ciphony, cifax, or civision, the relevant unit is a timed portion of the continuously varying audio or image scanning signal. Ciphony, cifax, civision system are categorized as privacy system or security systems. Privacy system mainly offers the protection against direct listening or direct viewing. The security systems, on the other hand, offer greater protection which actually in some cases will attempt protection against analysis as well.

3.3.2 Data Encryption Standard (DES)

IBM developed the Data Encryption Standards (DES) algorithm in the early 1970s. DES specifies a method for encrypting 64-bit blocks of clear data plain text into corresponding 64-bit blocks of cipher text employing a user specified 56-bit key. DES may be double or triple encrypted for additional security, with the user employing a different key after each transmission.

Because 2^{56} combinations of the keying variable are possible (and these keying variables can be changed readily), the algorithm is deemed by some experts to be highly secure. Cryptography experts in industry and government agencies maintain that DES is still a reliable standard. Operating at one try per microsecond, it would require approximately 2,284 years to break the code. Another consideration is the effect on security if the length of time the key operative is shorter than the time it would take search for the key.

DES is commonly used in the design, generation, and verification of personal identification numbers (PINs). These personal passwords are at the heart of the security scheme for validating ownership of automatic teller machine (ATM) debt cards. Message authentication also uses the DES algorithm.

Some experts believe that the DES is breakable. It may be possible for a parallel processor using special integrated circuits to go through all the permutations of a single DES transmission in one day. Although DES offers a high degree of security for commercial threats, the security of DES is lower for national or military threats.

The two main components of the DES based system are an algorithm and a key. The DES algorithm is a complex interactive process comprised of substitutions, permutations, and mathematical operations. The important feature about the DES approach is that the algorithm is fixed and is **public** information. However, the actual key used is a shared secret between the originator and the receiver of a transmission.

Advances in DES include lengthening a key to 128 bits and the multi-pass DES, which involves several passes usually three of encryption and decryption using different keys.

3.3.3 RSA Approach to Encryption

In its continuing search for a truly secret code, another encryption method which is now known as the RSA after its three inventors from the Massachusetts Institute of Technology, namely, Ronald Rivest, Adi Shamir and Leonard Adelman is now considered an important standard. The principle behind the RSA method is that it is easier to multiply two numbers than to factorise their product. This is even more so if the two numbers in

question are large prime numbers. For example, it is easy to multiply the prime numbers 11,927 and 20,903 and get the number 249,310,081. But if you are given the number 249,310,081 it is rather difficult to find its prime factors. But if the number is small, such as 35, it is easy to see that its prime factors are 5 and 7. The factoring of 29083 into 127 and 229 comes somewhere in between in terms of level of difficulty. There are an infinite number of prime numbers, and there is no known pattern to them except that they are prime.

Relying on this difficulty, Rivest and his colleague in the year 1977 had proposed the system, which is now known as RSA-129. RSA-129 digit number given below which was open to challenge by anyone in the world factorise into its prime factors. This challenge stood unbroken for about 17 years. But in 1993 it was broken through a cooperative effort of academics and hobbyists using over 1500 computers working for over 8 months on the Internet. The RSA-129, 114, 381, 625,757, 888, 867, 669, 235, 779, 976, 146, 612, 010, 218, 296, 721, 242, 362, 562, 561, 842. 935, 706, 935, 245, 733, 897, 830, 597, 1239 563, 958, 705, 058, 989,075,147,599,290, 026, 879, 543, 541, factors into the following two prime numbers, one of which is the following number of 64 digits, 3,490,529, 510, 847, 650, 949. 147, 849, 619, 903, 898. 133, 417, 764, 638, 493, 387,843, 990, 820, 577 and the other 32, 769, 132, 993, 266, 709, 549, 961. 988,190, 834, 461, 413, 177, 642, 967. 992, 942, 539, 798, 288, 533 of 65 digits.

The episode about the RSA is an interesting one to show that what may be considered difficult and impenetrable today may actually be rather easily broken into by more and more powerful computers of tomorrow.

The pursuit of truly secret code, physicists have been contemplating new approaches based on quantum keys. These ideas are still at a theoretical stage but with the kind of development seen in the past in other aspect of the computer industry, it may not be difficult to imagine that this approach of quantum cryptography can become a reality in the coming decade.

3.4 Cryptanalysis

The interpretation of secret communications without any previous knowledge of the system or the key is called cryptanalysis. In the case of modern cryptosystem, this requires extensive theoretical study, unusual power of observation, inductive and deductive reasoning, great concentration and perseverance. Also necessary are vivid imagination guided by good judgment. This has to be supplemented by a special aptitude and intuition gained from long and varied practical experience. It is possible that isolated, short cryptogram may resist solution indefinitely, even if it is in a fairly simple system.

However, a large volume of material even in a very complex cryptosystem may well be solved with time and effort, and especially with the help of modern powerful computers.

In general, the art of cryptanalysis may be reduced to three basic steps:

- (a) arrangement and rearrangement of data to disclose non-random characteristics or manifestations (e.g. in frequency counts, repetitions, patterns, and symmetrical phenomena);
- (b) recognition of the non random characteristics of manifestations when disclosed;
- (c) an explanation of the non-random characteristics when recognized. The requirements for the first step are experience or ingenuity and time may be appreciably reduced by the use of machine aids; for the second step, experience or statistics; and for the third step, experience or imagination, and intelligence.

3.5 Conclusion

You are now in a position to give some approaches to cryptography and define the major concepts of cryptanalysis.

3.6 Summary

Ensuring data security on PCs and LANs is a major problem. Unless organizations create security awareness in its work force, any measures for data security are not likely to prove successful. In this unit, the major approaches to cryptography have been illustrated. These range from Caesar Cipher to RSA-129.

3.7 Tutor Marked Assignment (TMA)

1. Write: "Fundamentals of Computer Science" using Caesar cryptography.
2. Write a short note on cryptanalysis.

3.8 Further Reading

M. Milenkovic, *Operating Systems: Concepts and Design*, Second Edition, Tata McGraw — Hill.

Unit 4

Computer Virus

4.0 Introduction

The world of computers and Information Technology is going through an era of electronic terrorism, in the form of virus. It is a problem that is potentially so dangerous that it threatens the proper functioning of the computer system in today's information age.

In this unit, we would discuss about the evolution of virus, the menace caused by them, the process of infection, classification of viruses, types of viruses, the prevention and the cure against them.

4.1 Objectives

At the end of this unit you should be able to:

- define computer viruses and its evolution
- explain the damage a virus can do
- define the process of virus infection

4.2 The Evolution of Virus

The concept of virus dates back to 1949, when John Von Neumann submitted a paper putting forward the concept of a "Self Replicating" program; the idea seemed impossible and was dropped. Subsequently, the first virus like program appeared in the form of a recreational game called "CODE-WARS" at the Bell Labs of American Telephone and Telegraph Company. In "Code-wars" two players were to code a set of programs that would destroy the other players programs. Realizing the potential danger of such programs, the authors did not reveal the presence of such programs. Concurrently, at Massachusetts Institute of Technology, students were carrying out experiments with computer, which no body had ever tried. Their relatively harmless hobby of messing up with programs gave rise to the idea and concept of computer viruses.

The first commercial application of viruses was in 1985, when two Pakistani brothers, in order to keep track of software piracy used Brain Virus (also known as Pakistani virus) on their low cost software sold from their outlet in Lahore. Hidden in nearly every disk they sold, was an extra program not supplied by the original manufacturer, a snippet of computer code, self-replicating in nature that would infect an unauthorized user's computer by disrupting his operations. These self-replicating programs multiplied so fast that, today, they are a threat to the smooth operation of a computer.

4.3 The Menace

The virus, whether biological or electronic is an information disorder. Biological viruses are tiny genetic codes DNA or RNA that take over the machinery of a living cell and are capable of making thousands of replicas of the original virus. Like its biological counterpart, virus carries in it an instructional code that makes copies of itself

Lodged in a host computer, the typical virus takes temporary control of the computer disk operating system. When the infected system comes in contact with an uninfected computer, the virus passes onto the uninfected machine and, thus spreads like a forest fire, infecting machines after machine with which it comes in contact.

Computer viruses are computer programs, which are a collection of coded instructions. The basic difference between a normal program and a virus is that viruses are self-replicating, they have the capability of executing themselves without being asked for. Computer virus is a very broad term in itself and includes not only viruses, but also Worms and Trojans.

Trojans are similar to viruses. They move around as valid programs, sometimes getting executed with flashy opening screens describing them as —"Word Process" or a "Database Package". Trojans are programs that claim to do something but do something completely different and in the process damage information stored on a computer system. Trojans do not infect other software.

Worms travel longer distances, by storing themselves in critical areas of the disk from where they get loaded and have with them sufficient code to transfer themselves outward from the system they infect. Worms have been known to damage and infect entire LANs.

Apart from self-replication, another devastation caused by viruses is data loss. A virus can also take steps to avoid its detection. That makes viruses even more dangerous, because you may come to know about the infection when it has struck. Even though all viruses are developed with a specific characteristic, most of them result in data-loss. Most viruses are designed to perform simple feats but in order to do so, they:

- Corrupt the most sensitive area of the disk; the File Allocation Table (FAT) or the directory area.
- Modify the interrupt organization of the system, meaning when a read or write to screen operation takes place, it is routed through the virus code in the memory resulting in unresolvable interrupt clashes where a program opens up a file simultaneously for read-write access and the virus interrupts every operation.
- Even though the virus may have no instructions built into it to destroy data, it can nevertheless render a disk full of files absolutely useless.

4.4 The Process of Infection

To understand how a virus infects a system, we go back to the elementary working of a computer. On booting, the system carries out the ROM instructions, the first being the Power On Self Test (POST), which is followed by the bootstrap process of reading the boot record and loading of the disk operating system. In MS-DOS, it involves the loading of [IBMDOS.COM](#), [IBMBIO.COM](#) and [COMMAND.COM](#) along with some optional files like CONFIG.SYS and AUTOEXEC.BAT. The booting process culminates into the system prompt displayed on the VDU.

The infection may begin as soon as a computer system boots from a contaminated disk or executes an infected program. Whatever viruses are present gets activated, which immediately begin to spread throughout the system.

Another important aspect that needs mention here is the interrupt mechanism. All input/output activities on a PC are carried out by interrupts. The interrupt mechanism in itself is very complex. We will try to understand it with an example. Let us say, a user wishes to save his program and presses the required keys on the keyboard. This treated as an interrupt. The main memory has specific routines to handle these user requests. One such set of routines exists in the ROM-BIOS and the another is in the DOS program in the memory, loaded from [IBMBIO.COM](#). The routine that services the interrupt requests are termed as Interrupt Service

Unit 4

Computer Virus

4.0 Introduction

The world of computers and Information Technology is going through an era of electronic terrorism, in the form of virus. It is a problem that is potentially so dangerous that it threatens the proper functioning of the computer system in today's information age.

In this unit, we would discuss about the evolution of virus, the menace caused by them, the process of infection, classification of viruses, types of viruses, the prevention and the cure against them.

4.1 Objectives

At the end of this unit you should be able to:

- define computer viruses and its evolution
- explain the damage a virus can do
- define the process of virus infection

4.2 The Evolution of Virus

The concept of virus dates back to 1949, when John Von Neumann submitted a paper putting forward the concept of a "Self Replicating" program; the idea seemed impossible and was dropped. Subsequently, the first virus like program appeared in the form of a recreational game called "CODE-WARS" at the Bell Labs of American Telephone and Telegraph Company. In "Code-wars" two players were to code a set of programs that would destroy the other players programs. Realizing the potential danger of such programs, the authors did not reveal the presence of such programs. Concurrently, at Massachusetts Institute of Technology, students were carrying out experiments with computer, which no body had ever tried. Their relatively harmless hobby of messing up with programs gave rise to the idea and concept of computer viruses.

The first commercial application of viruses was in 1985, when two Pakistani brothers, in order to keep track of software piracy used Brain Virus (also known as Pakistani virus) on their low cost software sold from their outlet in Lahore. Hidden in nearly every disk they sold, was an extra program not supplied by the original manufacturer, a snippet of computer code, self-replicating in nature that would infect an unauthorized user's computer by disrupting his operations. These self-replicating programs multiplied so fast that, today, they are a threat to the smooth operation of a computer.

3.6 Summary

Ensuring data security on PCs and LANs is a major problem. Unless organizations create security awareness in its work force, any measures for data security are not likely to prove successful. In this unit, the major approaches to cryptography have been illustrated. These range from Caesar Cipher to RSA-129.

3.7 Tutor Marked Assignment (TMA)

1. Write: "Fundamentals of Computer Science" using Caesar cryptography.
2. Write a short note on cryptanalysis.

3.8 Further Reading

M. Milenkovic, *Operating Systems: Concepts and Design*, Second Edition, Tata McGraw — Hill.

4.3 The Menace

The virus, whether biological or electronic is an information disorder. Biological viruses are tiny genetic codes DNA or RNA that take over the machinery of a living cell and are capable of making thousands of replicas of the original virus. Like its biological counterpart, virus carries in it an instructional code that makes copies of itself.

Lodged in a host computer, the typical virus takes temporary control of the computer disk operating system. When the infected system comes in contact with an uninfected computer, the virus passes onto the uninfected machine and, thus spreads like a forest fire, infecting machines after machine with which it comes in contact.

Computer viruses are computer programs, which are a collection of coded instructions. The basic difference between a normal program and a virus is that viruses are self-replicating, they have the capability of executing themselves without being asked for. Computer virus is a very broad term in itself and includes not only viruses, but also Worms and Trojans.

Trojans are similar to viruses. They move around as valid programs, sometimes getting executed with flashy opening screens describing them as —"Word Process" or a "Database Package". Trojans are programs that claim to do something but do something completely different and in the process damage information stored on a computer system. Trojans do not infect other software.

Worms travel longer distances, by storing themselves in crittas of the ,disc from where they get loaded and have with them sufficient code to transfer themselves outward from the system they infect. Worms have been known to damage and infect entire LANs.

Apart from self-replication, another devastation caused by viruses is data loss. A virus can also take steps to avoid its detection. That makes viruses even more dangerous, because you may come to know about the infection when it has struck. Even though all viruses are developed with a specific characteristic, most of them result in data-loss. Most viruses are designed to perform simple feats but in order to do so, they:

- Corrupt the most sensitive area of the disk; the File Allocation Table (FAT) or the directory area.
- Modify the interrupt organization of the system, meaning when a read or write to screen operation takes place, it is routed through the virus code in the memory resulting in unresolvable interrupt clashes where a program opens up a file simultaneously for read-write access and the virus interrupts every operation.
- Even though the virus may have no instructions built into it to destroy data, it can nevertheless render a disk full of files absolutely useless.

4.4 The Process of Infection

To understand how a virus infects a system, we go back to the elementary working of a computer. On booting, the system carries out the ROM instructions, the first being the Power On Self Test (POST), which is followed by the bootstrap process of reading the boot record and loading of the disk operating system. In MS-DOS, it involves the loading of [IBMDOS.COM](#), [IBMBIO.COM](#) and [COMMAND.COM](#) along with some optional files like CONFIG.SYS and AUTOEXEC.BAT. The booting process culminates into the system prompt displayed on the VDU.

The infection may begin as soon as a computer system boots from a contaminated disk or executes an infected program. Whatever viruses are present gets activated, which immediately begin to spread throughout the system.

Another important aspect that needs mention here is the interrupt mechanism. All input/output activities on a PC are carried out by interrupts. The interrupt mechanism in itself is very complex. We will try to understand it with an example. Let us say, a user wishes to save his program and presses the required keys on the keyboard. This treated as an interrupt. The main memory has specific routines to handle these user requests. One such set of routines exists in the ROM-BIOS and the another is in the DOS program in the memory, loaded from [IBMBIO.COM](#). The routine that services the interrupt requests are termed as Interrupt Service

Routines (ISR's) and are located in the memory with their addresses. Then interrupt request activated a number and not the routine address, thus, there exists a table with the interrupt numbers and the corresponding routine address in DOS. When an interrupt request is made, the CPU looks up the table, performs the required routines and transfers the control back to the program.

The contents of ISR address table being in the RAM is vulnerable to modification by user programs and that is what a virus does-modify the ISR's address

45 Classification of Viruses

Viruses are classified on the basis of their mode of existence and there are three categories of viruses:

1. BOOT Infectors
2. SYSTEM Infectors
3. GENERAL EXECUTABLE PROGRAM Infectors.

4.5.1 BOOT Infectors

As the name suggests, they are characterized by the fact that they physically reside in the boot sector (0 zero) sector of the disk. A system infected by such a virus will have the virus residing in a particular area of the disk rather than in a program file. These viruses get loaded soon after the Power On Self Test and control the system and remains in control at all times. They sometimes have the capability to trap soft booting (i.e. CTRL ALT DEL) and remain in control even if the system is booted from a non-infected floppy, thereby contaminating the clean floppy.

Boot infectors displaces information originally residing on the location, which they occupy. While writing onto the boot sector, the virus ensures that the boot record is not deleted. Once the virus is loaded, it automatically transfers control to the area where the boot record is available. The reason behind doing this is that the boot record contains instructions to read [IBMBIO.COM](#) and [IBMDOS.COM](#) and if these files are not readable, access to the disk is not possible, and so the virus becomes ineffective.

Boot infectors typically create "Bad sectors". Boot infectors are the types, which once loaded would stay in the memory until the system is shut off, and until the disk reformats.

4.5.2 SYSTEM Infectors

This second category of viruses deals with the components of the system itself. All machines without exception require an operating system in order to create an environment in which the operator works. In MS-DOS, [COMMAND.COM](#) contains all the internal commands. If no such command file exist, commands such as COPY, DIR etc. are not loaded onto the memory when the machine is booted. The System Infectors attach themselves to a file such as [COMMAND.COM](#) or other memory resident files and manipulate these files.

System infectors differ from Boot infectors in the sense that system infectors gain control after the computer is booted: and infects the hard disk or bootable floppies, which contain the appropriate system files only. They have another peculiarity that they may activate after a given period of time or may instantly begin subtle modifications of systems error messages or information messages.

4.5.3 [GENERAL.COM](#) or EXE Infectors

From the infection point of view, these viruses are most dangerous and devastating of the three classes of viruses. They attach themselves to program files and can spread to almost any executable program in any system. These viruses change the original program instructions into a "jump" to its own code and follows that code with a return to the original program. As a result, whenever the program is executed, the virus gets loaded and executed first and then allows the original program to proceed. It remains memory resident and infects each and every program that is loaded for execution.

By attaching themselves to EXE or COM files, they alter the file size and sometimes multiple infections

renders program files too large to be accommodated in the memory.

4.6 Conclusion

In this unit, you were introduced to the evolution of virus, the menace cause by them and the classification of viruses.

4.7 Summary

This unit enumerates the process of infection, classification of viruses, prevention and cure.

4.8 Tutor-Marked Assignment (TMA)

- I** Write a short note on the evolution of virus
2. Explain the process of virus infection.
3. Classify viruses

4.9 Further Reading

S.A Reju, Virus: Lecture Notes

Unit 5

Viruses and Prevention

5.0 Introduction

In this unit you see a list of some viruses. You will also prevent viruses from attacking your computer, and when an attack occurs you will see ways of curing the virus

51 Objectives

At the end of this unit, you would be able to:

- identify different types of viruses
- explain how to prevent or cure a virus infection.

52 Some Viruses

The virus list has become a non-ending entity with new viruses joining the list every other day. We would be discussing some of the most commonly prevalent viruses in the computer industry. However, this list is incomplete.

1. Scores Virus

These viruses are prevalent in Macintosh machines. Scores virus has a built in time trigger that activates at two, four and seven days after the disk has become infected. The consequences are varied ranging from printing problems, system crashes and malfunctioning of disk operations. This virus does not directly affect data files, but erasure of this virus requires deletion of all files.

2. Brain Virus

This is one of the first viruses that came into being. Also known as the Pakistani virus, it was developed by the Pakistani brothers to keep track of low cost software, those were sold out of their outlet in Lahore. The virus pops up a screen saying "Welcome to the Dungeon". This virus is known to destroy data and are highly contagious.

3. Lehigh Virus

This virus originated at the Lehigh University Computer Center. This virus stays in the stack space of [COMMAND.COM](#). With the booting of a PC from an infected disk, the virus is spread through commands such as COPY, TYPE, DIR etc. On any other disk with [COMMAND.COM](#) the virus code gets copied to the other disk and a counter is incremented on the parent. When the counter reaches a value of 4, all files of the disk gets erased. The boot sector gets ruined and also the FAT.

4. Friday the 13th

This virus attacks not only the [COMMAND.COM](#) but also other executable files. When a COM or EXE file is executed for the first time after booting, the virus captures a specific interrupt and inserts its own code; after which, whenever any EXE file is executed, the virus code is written to the end of the file resulting in increase in size of the file by 1808 bytes. In COM files the virus code is written to the beginning of the file.

The increase in size of the EXE and COM files causes the program to become too large to be loaded into the memory. Also after a certain interval of time, delays are inserted resulting in considerable slowing down of the programs. The worst disaster occurs, if the infected EXE or COM is executed when the system data is Friday the 13th, all files get deleted.

5. Sunnyvale Slug

This does a variety of things like displaying a message "Greetings from Sunnyvale. Can you find me?" and also sometimes modifies the COPY Command resulting in deletion of files instead of copying.

6. Raindrops

This virus infects COM files. It intercepts the load and execute function of MS-DOS. It checks whether the file is EXE or not, if the file is not an EXE file, the first three bytes of the file are replaced by a jump instruction at the end of the file, where it gets attached after encryption. This results in dropping or showering of characters on the screen like raindrops and is also accompanied by appropriate sound effects.

7. Happy Birthday 30th

This virus gets activated on January 5th, if any of the infected program get executed, and will ask the user to type "Happy Birthday 30th". It might destroy all the data stored on a disk. The symptoms of this virus are that the computer memory is reported 6KB less than actual e.g. 634 KB instead of 640 KB.

Following is a list of prominent viruses that have created havoc on many machines across the globe. A comprehensive list cannot be built as every other day a new virus gets included into it.

Name	Also known as	Name	Also known as
<u>M a r i j u a n a</u>	(Stoned)	Tequila	
Joshi	(Happy Birthday Josh')	Slow	
Flip		BFD	
Eddie	(Dark Avenger)	World peace	
<u>J e r u s a l e m</u>	(Jerusalem Ver. A to E,)(Friday the 13th)	Kanishka	
Serum	(Yankee Doodle, Version A, B)	Alfa	
<u>K i n k y</u>	(Fellowship)	Trikal	
V 2 0 0 0		Mummy	
<u>Z e a l o t</u>	(Key press)	Mubank	
Changu-Mangu		Charas	
<u>M a r c h - 6 t h</u>	(Michael Angelo)	Feist	
<u>F r o d o</u>	(4096, 100 Years)	Boabob	
<u>D e s i</u>	(Made in India)	Monkey	
<u>D a t a l o c k</u>		ExeBug	
Gravity	(Raindrops, 1701, 17XX) (1701/1704 Version A,B)		
		Feist	
<u>H o n g K o n g</u>		NewBug	
Liberty		1403	
Form			
Pronto	(15XX)	G3	
Taiwan		Cansu	

Name	Also known as	Name	Also known as
Invader	(And-CAD)	Dong 2	
Generic	(Key drop)	Long-1	
Possessed	(Poss)	Khobar	
Black Monday		PFO	
Plastic:me		Die-Hard-2	
Bosh		Bloomington	
Dir-2		Fat-avenger	
Gumnam		Green-Catapillar	
		Angelina	

5.3 Prevention

Even though the computer industry has found a somewhat plausible solution to the virus problem in the form of vaccines, it is always advisable to follow the dictum "Prevention is better than cure". Moreover, the viruses are made faster than the vaccines. It is a good practice to follow some simple precautionary measures, which can reduce the possibility of a virus attack. The precautionary measures are:

- The CHKDSK command can be incorporated to the AUTOEXEC.BAT to check the disk. If the number of hidden files increases, the matter should be looked into.
- Do not copy pirated software on your system.
- Write protect tags should be used on the original software diskettes
- Proper backup of all data and program files should be kept.
- Copying of files should be done carefully, a better practice is to write the COPY command in a batch file with CHKDSK command.
- Used floppies should be reformatted.
- Avoid letting the system to be used by unauthorized users.
- Restrict the use of outside floppies.
- Do not download suspicious shareware programs.

5.4 The Cure

The viruses are not omnipotent. Viruses can be cured with anti-viral programs. The anti-viral programs perform one or more of the following functions.

- prevention
- detection
- vaccination
- inoculation
- identification, and/or
- damage control

A good anti-viral utility is one which checks whether the system has been infected or not. These programs stop the virus from infecting the system. They do not allow the modification of executable files, so that a file virus cannot get a foothold.

Some of them refuse to let any program make itself resident in RAM unless allowed by the user. Others do not allow the user to run a program unless it is on a list of approved and tested applications. The detectors warn the user of the presence of a virus after it is loaded into the machine or disk. These programs maintain a file with a list of checksum values of the executable files. The identifiers rely on the fact that when the virus replicates, it makes a copy of itself.

The vaccinators inject some code into the executable files. When the vaccinated file is run, the injected

code performs an integrity check on the program being executed and warns if any changes have been made.

The inoculators insert the virus signature into infected areas or files at appropriate locations. When the virus performs their self-detection, they find their signature and believe that the memory/disk file is already infected and so do not infect.

The better equipped anti-viral programs control damages. They may be preventive or restorative. Preventive techniques include stopping attempts at direct access such as formatting and deleting, or even write protecting the hard disk while testing unfamiliar software. The restorative process is achieved by maintaining a copy of the CMOS information, boot sector information, the file allocation table etc. in a safe area like a floppy.

As a virus can hide itself in many different ways, it is difficult to detect all viruses with just one anti-viral program. Moreover, the virus writers keep altering the viral code, so that any existing anti-viral programs cannot detect it. The point to remember is that there is not per cent foolproof anti-virus program available and, in principle there never will be.

5.5 Conclusion

Here we will discuss different types of viruses, the prevention and the cure.

5.6 Summary

This unit completes the discussion of viruses, the prevention and the cure. With the development of computer field, new viruses may appear. Therefore, one has to keep track of latest development in this field knowing how to counter the resulting problems.

5.7 Tutor Marked Assignment (TMA)

1. Make a list of at least 10 viruses with their characteristics.
2. Make a list of anti-virus packages that you are familiar with.

5.8 Further Reading

S.A. Reju: Virus Lecture Notes.