Indira Gandhi National Open University (IGNOU)

Bachelor of Library and Information Science (BLIS)

STUDY MATERIALS

Course code: BLI-224
ICT Fundamentals

JATINDER SINGH
BLIS (JULY-2018)

www.jatinderjyoti.in
jatinderjyoti84@gmail.com
fb/insta: jatinderjyoti.raina

Indira Gandhi
National Open University
School of Social Sciences

BLI-224
ICT Fundamentals

Block

1

BASICS OF ICT

Programme Design Committee

Prof. Uma Kanjilal (Chairperson)
Faculty of LIS, SOSS, IGNOU

Prof. B.K.Sen, Retired Scientist, NISCAIR
New Delhi

Prof. K.S. Raghavan, DRTC
Indian Statistical Institute, Bangalore

Prof. Krishan Kumar, Retired Professor
Dept. of LIS, University of Delhi, Delhi

Prof. M.M. Kashyap, Retired Professor
Dept. of LIS, University of Delhi, Delhi

Prof. R.Satyanarayana
Retired Professor, Faculty of LIS, SOSS, IGNOU

Dr. R.Sevukan (Former Faculty Member)
Faculty of LIS, SOSS, IGNOU

Prof. S.B. Ghosh, Retired Professor
Faculty of LIS, SOSS, IGNOU

Prof. T. Viswanathan, Retired Director
NISCAIR, New Delhi

Dr. Zuchamo Yanthan
Faculty of LIS, SOSS, IGNOU

Conveners:

Dr. Jaideep Sharma
Faculty of LIS, SOSS, IGNOU

Prof. Neena Talwar Kanungo
Faculty of LIS, SOSS, IGNOU

Programme Coordinators

Prof. Jaideep Sharma and
Prof. Neena Talwar Kanungo

Course Coordinator

Prof. Uma Kanjilal

Programme Editor

Prof. Jaideep Sharma

Course Preparation Team

Unit No(s).	Contributor(s)
1-3	Dr. Naveen Kumar Singh
4	(Late) Prof. T. Viswanathan

Course Editor
Prof. Uma Kanjilal

Internal Faculty:
Prof. Uma Kanjilal
Prof. Neena Talwar Kanungo

Material Production

Mr. Manjit Singh
Section Officer (Pub.)
SOSS, IGNOU

Secretarial Assistance

Ms. Sunita Soni
SOSS
IGNOU

Cover Design

Ms. Ruchi Sethi
Web Designer
E Gyankosh
IGNOU

July, 2014

ISBN : 978-81-266-6771-0

Further information on Indira Gandhi National Open University courses may be obtained from the University's office at Maidan Garhi, New Delhi-110 068 or visit University's Website http://www.ignou.ac.in.

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by Director, SOSS.

Lasertypesetted at Graphic Printers, 204, Pankaj Tower, Mayur Vihar, Phase-I, Delhi-110091.

Printed at :

UNIT 1 BASICS OF COMPUTER TECHNOLOGY

Structure

1.0 OBJECTIVES

After completing this Unit, you will be able to:

- describe the basic architecture of the computer;
- discuss the use of various hardware and peripheral devices;
- explain the functions of major hardware components; and
- understand the basic working of Ubuntu Linux operating system.

1.1 INTRODUCTION

The utilisation of computer technology is well recognised, in almost all fields of governance, businesses, libraries, education, etc. Today, computer technology has infused every sphere of common man. From banking to shopping, railway reservation to medical prescription. In this unit, we will introduce you to the basic computer technology and how it works? Knowing internal and external components of a computer system is important for a beginner of computer technology. This unit will provide you the basic knowledge about the devices associated with computer system particularly in the context of library systems. In addition we will also try to discuss some of the basic functions of an operating system, which actually creates an interface between the computer hardware and you. Please note that this unit covers only the basics of computer technology. For further details you may refer to suggested further readings given at the end of the unit.

1.2 OVERVIEW OF COMPUTER SYSTEM

Before starting an introduction about computer, let us understand the need and reason why computers are so much important for us. Generally computers help us in performing tasks that are repetitive, involve calculation or manipulation of data and perform tasks that involve storage of large quantities of information.

Computer can be defined as "An electronic machine that is capable of interpreting and executing stored program (sequence of instructions), input data, perform operation on data (calculation and logical operations), and output the results". The data which normally we deal with consists of numeric and characters such as decimal digits 0 to 9, alphabets A to S, operators (e.g. +, -, >, =, etc.) and many other special characters or images (e.g.; @,{,],etc.). However, computers cannot understand such form of data and hence it has to be transformed into binary-form (also known as machine language) by using two symbols 0 and 1, which are called binary digits or bits. For Example: If we want a computer to do ADD (2, 3); computer needs to map this particular instruction into a binary form. Lets assume the mapping is "01= ADD, 2=10, 3=11", for a computer this instruction will be "01 10 11". 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 the instructions supplied to the computer?

b) How are they interpreted and executed?

Most of today's computers look like what is illustrated in Figure 1.1. The designs are based on concepts developed by John Von Neumann referred to as the Von Neumann architecture (shown in Figure 1.2).

Fig. 1.1: A Modern Computer

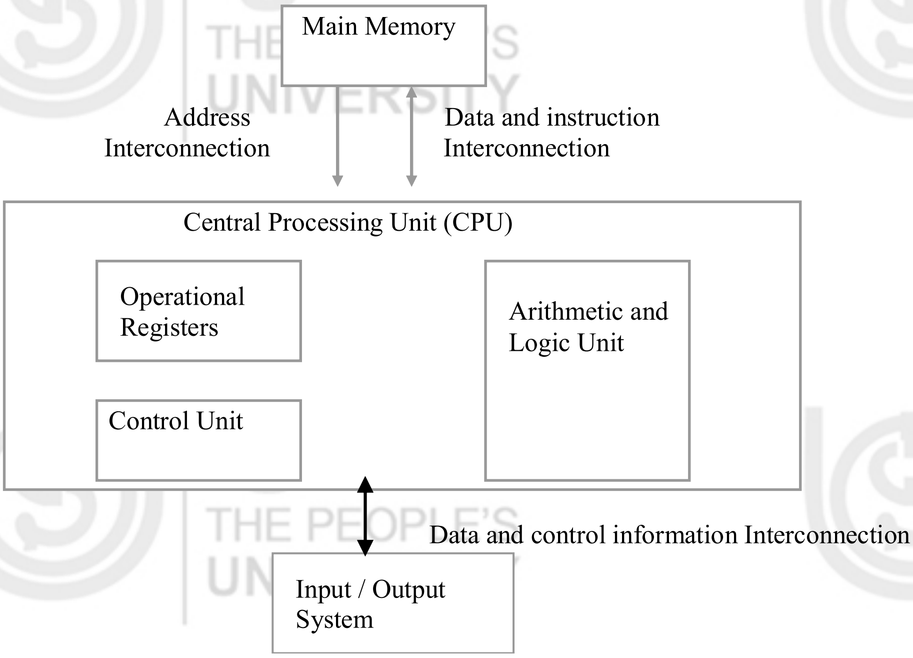


Fig. 1.2: Von Neumann Computer Architecture

Von-Neumann proposed that there should be a unit performing arithmetic and logical operation on the data. This unit is termed as Arithmetic Logic Unit (ALU). One of the ways to provide instruction to such computer will be by connecting various logic components in such a fashion, that they produce the desired output for a given set of inputs. The process of connecting various connections is termed as Hardwired. But this is a very inflexible process of programming. Let us have a general configuration for arithmetic and logical functions. In such a case there is a need of a control signal, which directs the ALU to perform a specific arithmetic or logic function on the data. Therefore, in such a system, by changing the control signal the desired function can be performed on data. Any operation, which needs to be performed on the data, can be obtained by providing a set of control signals. Thus, for a new operation one only needs to change the set of controls.

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 addition, 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. How can the instructions and data be put into the computers? The instruction and data needs to be supplied by external environment. For this an input module is needed. The main responsibility of input module is to put the data in the form of signals that can be recognised by the system. Similarly, we need another component, which 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.

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 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 Kilo Bytes (KB) or Mega Bytes (MB). One kilobyte stands for 2^{10} bytes which 1024 bytes (or approximately 1000 bytes). A mega byte stands for 2^{10} Kilobytes, which is approximately little over one million bytes.

Computers are therefore, made of four major components (parts):

1) **Input units:** Humans interface with the computer through devices like the mouse and keyboard.

2) **Central Processing Unit**: Referred as "Brain" of the computer that controls all computer operations, processes information and computes results. It is the logical component of a computer system that interprets and executes program instructions. This component comprises:

   - Control Unit- that interprets program instruction, directs the internal operation, and controls the flow of data in and out of primary memory.
   - Arithmetic and Logical Unit- that performs the arithmetic and logical operations.
   - Memory- also called primary memory holds the input/output data and program instructions. Memory is made of Integrated Circuits or Chips that are tiny silicon chips. In each chip thousands of electronic components are engraved.

3) **Secondary Storage** or Auxiliary Storage is a means for permanently storing data. Do not confuse it with the memory. Memory is a small size storage, which is used by CPU and operating system for its own use. Due to this reason it is sometimes not considered as major components of computer system, instead it is considered as an input/output Device.

4) **Output Units:** data is transferred from a computer to an output device like the printer and monitor where the results are displayed.

1.3 COMPUTER PERIPHERALS AND HARDWARE

There are two terms commonly used in computer world, computer hardware or simply hardware and peripherals. Let us discuss what we mean by these terms. As you know computer is an electronic device that has the ability to store, retrieve, and process data. The physical parts that make up a computer (the central processing unit, input, output, and memory) are called **hardware**. Any hardware device connected to a computer and any part of the computer outside the CPU and working memory is known as **Peripherals**. Some examples of peripherals are keyboards, mouse, monitors, printers, scanners, barcode reader, tape drives, microphones, speakers, joysticks, plotters, and cameras.

1.3.1 Computer Peripherals

When you look on a simple computer you can see Central Processing Unit (CPU) cabinet, Monitor, Keyboard and Mouse. Your computer and you interact through these peripherals. The keyboard and monitor are the minimum peripherals you should have with your computer. Your choice of peripherals depends on personal preference and the complexity of the interactions you intend to have with your computer.

Computer Monitors

The computer monitor is an output device that displays input on a screen and is very similar to a television monitor. When the computer wants to display something, it calculates how it needs to change the color and brightness of the different pixels, and changes the values in the video memory. The smaller the pixels, the clearer and sharper the picture appears on the monitor. Earlier CRT monitors were used as given below in the figure 1.3a. Nowadays the LCD monitors (as shown in figure 1.3b) are in common use these are thinner in size and uses lesser electricity.

Fig. 1.3a: CRT Monitors

Fig. 1.3b: LCD Monitors

Keyboard

The computer receives most of its input from the user via the keyboard that is very similar to the typewriter keyboard as shown in figure 1.4 given below. This input device is connected by a cable to the keyboard port on the back of the computer. There are extra keys on the computer's keyboard that are not found on a normal typewriter. The exact manner in which the keys function depends on the software program. We have different type of keyboards available nowadays which can support multimedia activities and Internet browsing.

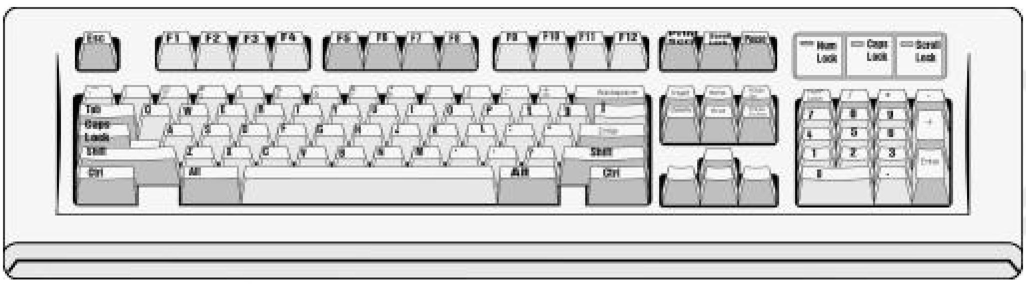


Fig. 1.4: A Computer Keyboard

As shown below in the figure 1.5, mouse is a pointing input device attached to the computer that controls the movement of the cursor on the screen. It allows the user to execute commands using point & click and click & drag techniques. As the user moves the mouse across the pad, the cursor moves across the screen. The mouse should always be used with a mouse pad to provide a smooth surface for mouse movement and to help keep the mouse from damage. If you "run out of room" on the pad, simply pick up the mouse and move it to the opposite edge and continue movement. Generally, mouse has two buttons (left and right) and newer mice have a scroll wheel between the two. Mouse commands are executed by "clicking". The term "click" refers to the left mouse button.

Fig. 1.5: Mouse a Pointing Device Attached to the Computer

Printer

A printer is an output device to transfer images and text from a computer to a printed page as given below in figure 1.6. Printers take information from the CPU and transfer it to paper, provide a hard copy (permanent human-readable text and/or graphics) of documents stored in electronic form, usually on physical print media such as paper or transparencies. Many printers are primarily used as local peripherals, and are attached by a printer cable or, in most new printers, a USB cable to a computer, which serves as a document source. Some printers, commonly known as network printers, have built-in network interfaces (typically wireless or Ethernet), and can serve as a hardcopy device for any user on the network. Individual printers are often designed to support both local and network connected users at the same time.

There are a number of different printer technologies available: Dot Matrix, Ink Jet or Laser. Inkjet printers use small cartridges full of different colour inks. They squirt a tiny drop of ink onto the paper using a bubble or pressure to form a dot. All the dots are layered on top of each other to form the desired colour. Laser printers are more expensive to run but produce a crisper image.

Fig. 1.6: Laser Printer

Scanner

An input device that is becoming less used with the advent of digital cameras. A scanner (shown below in the figure 1.7) is an optional extra rather than a necessity peripheral device connected to the Personal Computer (PC) via the parallel port or USB port. It allows you to digitise a flat, paper image and transfer it to your computer where the image can be manipulated. For instance you may use it for digitizing old photographs that can be stored on the hard drive, repaired, enhanced and printed out on a printer.

Fig. 1.7: A Scanner

Barcode Readers

Barcode reader is an input device for reading printed barcodes. We can also say it's a special scanner for barcodes. You all must have seen barcodes, that are printed on most of the products we use in our daily life. It looks like thin and thick series of black lines as shown below in figure 1.8:

1 25002 74135 0

Fig. 1.8: Example of a Barcode

Barcode reader consists of a light source, a lens and a light sensor interpreting optical impulses into electrical signals. Furthermore, almost all barcode readers contain decoder circuitry analysing the barcode's image data provided by the sensor and sending the barcode's content to the scanner's output port.

The barcode reader types can be classified based on the scanning methods; some of these types are described below:

Pen Readers

It requires the operator to swipe the pen over the code. Pen type readers consist of a light source and a photodiode that are placed next to each other in the tip of a pen or wand. A pen reader is shown below in figure 1.9.

Fig. 1.9: Pen Type Barcode Readers

To read a barcode, the tip of the pen moves from corner to corner the bars in a stable shift. The photodiode measures the intensity of the light reflected back from the light source and produces a waveform that is used to measure the widths of the bars and white spaces in the bar code. Black bars in the bar code absorb light and white spaces reflect light so that the voltage waveform produced by the photo diode is a representation of the bar and space pattern in the bar code. Further, the scanner decodes this waveform.

Laser scanners

Laser scanners as shown below in figure 1.10, work similar like a pen readers except that they use a laser beam as the light source and normally utilise either a reciprocating mirror or a rotating prism to scan the laser beam back and forth across the barcode.

Fig. 1.10: Laser type Barcode Readers

CCD Readers

Charge Coupled Device readers also called as LED scanner, use an array of hundreds of tiny light sensors queued up in a row in the head of the reader, it is shown in figure 1.11. Every sensor measures the intensity of the light immediately in front of it.

Fig. 1.11: LED Barcode Readers

The important difference between a LED reader and a pen or laser scanner is that the LED reader is measuring emitted ambient light from the bar code whereas pen or laser scanners are measuring reflected light of a specific frequency originating from the scanner itself.

Camera-Based Readers

Two dimensional imaging scanners or camera based barcode readers are the newest types of barcode reader currently available. They use a small video camera to capture an image of a bar code as depicted in a figure 1.12. The reader then uses digital image processing techniques to decode the barcode.

Fig. 1.12: Camera based Barcode Readers

Speaker

Computer speakers shown in figure 1.13 or multimedia speakers, are external speakers, commonly equipped with a low-power internal amplifier.

Fig. 1.13: A Set of Computer Speaker

Modem

A modem enables a computer to transmit data over telephone or cable lines. Computer information is stored digitally, whereas information transmitted over telephone lines is transmitted in the form of analog waves. A modem converts between these two forms. Modems come in two forms: Internal and External; Internal modems are PCI cards that you can connect as shown in the figure 1.14a. In figure 1.14b External modem is shown, that are kept outside of your computer, connected either by a USB or Serial Port.

Fig. 1.14a: Internal Modem

Fig. 1.14b: External Modem

Radio-Frequency Identification (RFID)

RFID architecture works like a small area (generally in few meters) client-server network based on radio frequencies for communication between client and server. It has two component RFID Client Tag and RFID server or antenna or receiver. RFID tag is applied or incorporated into an object or a product for the purpose of identification and tracking using radio waves. RFID tag is an integrated circuit for storing and processing information, modulating and demodulating a radio-frequency (RF) signal, and other specialised functions. The receiver or antenna is for receiving and transmitting the signal. There are generally three types of RFID tags as shown below in the table 1.1:

Table 1.1: Classification of RFID tags

ACTIVE RFID TAGS	Contain a battery and can transmit signals autonomously.
PASSIVE RFID TAGS	Have no battery and require an external source to provoke signal transmission.
BATTERY ASSISTED PASSIVE (BAP)	Require an external source to wake up but have significant higher forward link capability providing great read range.

Among the many uses of RFID equipments, one is its implementation in libraries. This technology has gradually begun to replace the conventional barcodes on library items. The RFID tag can contain many identifying information, such as a book's title, author's name, publication year, publisher name, book type, etc. This information is read by an RFID reader, which replaces the standard barcode reader. The RFID tag used on products or library materials typically measures 50 mm X 50 mm. It may replace or be added to the barcode to provide an innovative and easier way for other inventory management also. It can also act as a security device, taking the place of the more traditional electromagnetic security strip and not only the books, but also the membership cards could be fitted with an RFID tag.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) What is a significance of ALU in CPU?

....

....

....

....

2) Calculate the number of bits in two-kilobytes.

....

....

....

....

3) How inkjet printers work?

....

....

....

....

4) Differentiate between LED and Pen type barcode readers.

....

....

....

....

5) How RFID technology can be useful for library management?

....

....

....

....

1.3.2 Computer Hardware

In general when we use the term hardware it refers to nuts and bolts available at the hardware shops. Similarly the nuts and bolts of a computer is called hardware, which is a physical, electrical, and mechanical parts of the computer. In the last decade or so, the hardware technologies have seen many changes and also have advanced remarkably. From few kilobytes of RAM (Random Access Memory), now the PCs have couple of Gigabytes of RAM, hard-disks have also improved from few Gigabytes to hundreds of Gigabytes memory space, and also other drives and monitor have also emerged with better performance and cost. In this section we will study about the different internal hardware components of a computer.

CPU (Central Processing Unit)

The CPU is the computer's control center, it appears as shown in figure 1.15. Think of it as the brain that does all the thinking (computation). It reads instructions from your software and tells your computer what to do. The actual CPU is about 1.5 inches square, yet it is the most critical part of the computer. The speed at which the CPU processes information internally is measured in Megahertz (MHz) and Gigahertz (GHz). 1 GHz is equal to 1,000 MHz. Generally; processors with higher MHz or GHz enhance your ability to run creative, entertainment, communication, and productivity applications.

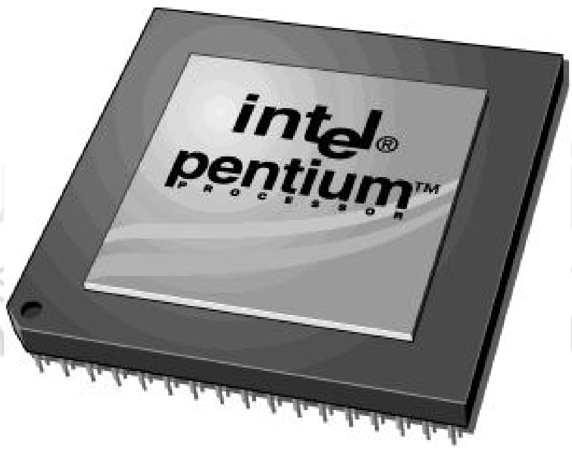


Fig. 1.15: Central Processing Unit

Motherboard

Sometimes called the system board or main board, the motherboard is the main circuit board of a PC. The motherboard (shown in figure 1.16) is the central nervous system and circulatory system, plus much more, all rolled into one. The motherboard typically contains the processor (or CPU), BIOS (basic input/output system), memory, mass storage interfaces, serial and parallel ports, expansion slots, and all the controllers required to communicate with standard peripheral devices, such as the display screen, mouse, keyboard and disk drive. Collectively, some of the chips, which reside on the motherboard are known as the motherboard's chipset.

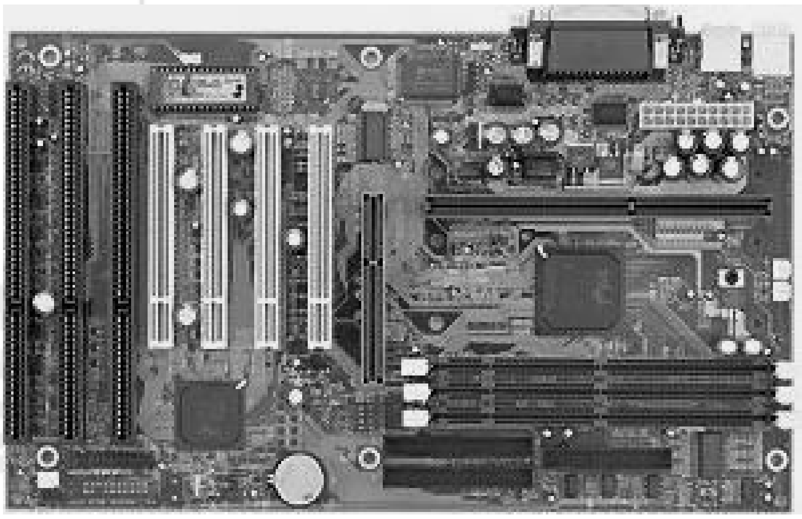


Fig. 1.16: The Motherboard

Power Supply Unit

The Power Supply Unit appears as shown in figure 1.17, is the most important component in a PC. This is the core of the system and is needed to supply power to the motherboard, which in turn, supplies power to all the other components inside and sometimes outside the computer case. Its fundamental function is to convert available power from the wall socket, to the type that a computer can use. It converts 230 volt current and splits it into useable +3.3v, +5v and +12v DC current.

Fig. 1.17: Power Supply Unit

ROM (Read Only Memory)

A type of memory chip which does not lose information even when the power is turned off. Once data is programmed into the ROM chip, its contents cannot be altered. For example, ROM BIOS chips are used to store information for starting up your computer.

RAM (Random Access Memory)

RAM as shown in figure 1.18 are chips that allow data to be written to, however unlike hard disks, the data is short term and is only used for the duration the PC is switched on. Data can be accessed very quickly using this method of temporary storage. Memory comes in many different types and sizes depending on the type of motherboard you have.

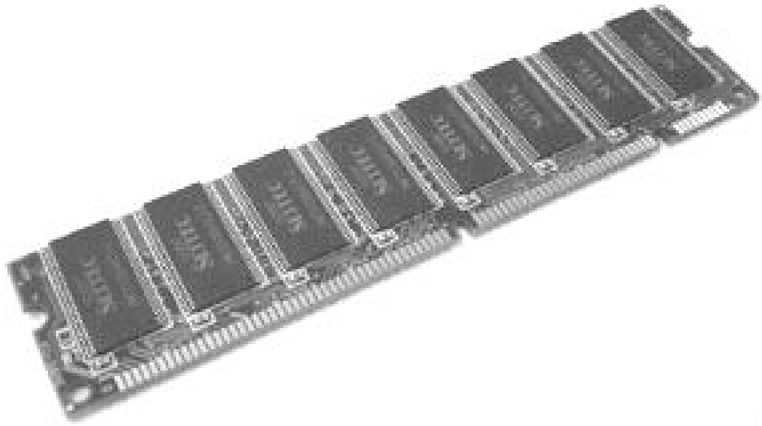


Fig. 1.18: Random Access Memory

Storage Devices

Computers are also known, as data processing machine, to store and process data computer must have memory. There are different kinds of memories like primary memory and secondary memory, primary memories like RAM (random access memory) and ROM (read only memory) is computer personal memory used to process its own data. However, the secondary memory is an off-line memory, which is used to store and carry the data like floppy disk, compact disk, hard disk etc.

Hard disk: It is normally large size memory that a computer uses to store information. Most computers come with one or two hard drives, situated inside the computer cabinet. The terms hard drive and hard disk are used interchangeably. Today's hard disks provide fast access and can hold several gigabytes of information as compared to megabytes on floppy disks. But hard disks are permanently fixed with screws in the computer cabinet; rarely we plug it out because detaching it frequently is not safe. Hence, to carry data we need some removable memory disk, further in this section we have listed some of these removable disks.

Floppy Disk: These are removable disk that stores information magnetically. You can use a floppy disk to read/write information between computers, or to make a backup of your files. Floppy disks are 3.5 inches in diameter and have a storage capacity of 1.44 MB. These are now obsolete.

Zip disks: This is also a removable disk, which can store 100 - 250 MB of data. A special 3.5" removable disk drive is needed to retrieve the information from the computer and write to the zip disk. An external zip drive can be moved from one computer to another.

Cartridge tapes: These are magnetic tapes similar to cassette tapes used for the purpose of storage and backup. Since the information is stored sequentially, in the cartridge tapes, backup and retrieval of stored information is slower. The advantage of tape is that they can be purchased with large storage capacities (1– 4 GB) allowing the entire contents of the hard drive to easily fit on one tape.

CD-ROMS: Compact disks can store approximately 650-800 MB of data or 74-80 minutes of music. These drives are read only and cannot be used for recording data.

Read/Write CD-ROMS: CD-ROM drives that write, rewrite and record data. Two types of CD-ROM disks are used: CD-Recordable (CD-R) and CD-Read/Write (CD-R/W). CD-R/W's can only be "read" by CD Read/Write drives, while CD-R disks can be read by most CD-ROM drives.

Digital Video Disc (DVD): A digital video disc (DVD) is a type of optical disc storage technology that looks like a CD-ROM disc, but can store greater amounts of data. DVDs are often used to store full-length movies and other multimedia content that requires large amounts of storage space. But to run the DVD we should have a DVD drive.

USB/Flash Drive: USB drives are very popular removable disks in current time. It is a plug-and-play portable storage device that uses flash memory and is lightweight enough to attach to a key chain. These drives can be used in place of a floppy disk, zip drive disk, or CD. When the user plugs the device into their USB port, the computer's operating system recognises the device as a removable drive. Unlike most removable drives, a USB drive does not require rebooting after it is attached, do not require batteries or an external power supply, and is platform independent. Several USB drive manufacturers offer additional features such as password protection, and downloadable drivers that allow the USB drive to be compatible with older systems that do not have USB ports. USB drives are available in capacities ranging from 1GB to 32 gigabytes GB.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

6) What is BIOS?

....

....

....

....

7) How DVDs are different from CDs?

....

....

....

....

1.4 OPERATING SYSTEM

In contrast to the hardware, software are the non-physical components of the computer system. A set of instructions is known as programs and a set of programs, which gives a finite output, is called software. Operating system is the first software that you work on the first time or whenever you start your computer. Computers need something to manage all the hardware components and give an interactive interface to control these hardware devices, which are taken care by the computer operating system. There are different operating systems available like Microsoft Windows based operating systems like Windows 7 or UNIX or Linux variants etc. In this unit we will discuss Linux based operating system Ubuntu. Linux is a freely available, open source, Unix-like operating system. Written originally for the PC by Linus Torvalds, with the help of many other developers across the Internet, Linux now runs on multiple hardware platforms. Because of its speed, stability, and low cost, Linux became the fastest growing operating system for servers. Today, Linux is widely used for both basic home and office uses. It is the main operating system used for high performance business and in web servers. Linux has made a huge impact in this world. Now days there are many Linux based operating systems available. Some of its most popular flavours are Red Hat Linux, Fedora and Ubuntu. Most of the Linux based operating system has the following given features, however these may vary from version to version:

- **Low Cost/Free:** Linux and much of its software come with the GNU General Public License and hence it is free and open source.
- **Stability:** Linux has high stability compared with other operating systems. There is no need to reboot the Linux system to maintain performance levels.
- **Performance:** Linux provides high performance on various networks. It has the ability to handle large number of users simultaneously.
- **Networking:** Linux provides a strong support for network functionality; client and server systems can be easily set up on any computer running Linux. It can perform tasks like network backup faster than other operating systems.
- **Flexibility:** Linux is very flexible and can be used for high performance server applications, desktop applications, and embedded systems. You can install only the needed components for a particular use. You can also restrict the use of specific computers.
- **Compatibility:** It runs all common Unix software packages and can process all common file formats.
- **Fast and Easy Installation:** Linux distributions come with user-friendly installation.
- **Better use of Hard Disk:** Linux uses its resources well enough even when the hard disk is almost full.
- **Multitasking:** Linux is a multitasking operating system. It can handle many things at the same time.
- **Open Source:** Linux is an Open source operating systems. You can easily get the source code for Linux and edit it to develop your personal operating system.

1.4.1 Ubuntu Operating System

Ubuntu is an operating system based on Linux that is also developed by a worldwide community of programmers. Ubuntu is based on the concept of free or open-source software, meaning that you do not pay any licensing fees for Ubuntu, and you can download, use, and share the operating system free of charge. Being a Linux-based operating system, Ubuntu has a well-deserved reputation for stability and security. Ubuntu is generally acknowledged to be the most widely used version of Linux available.

Comparing Ubuntu with Windows

Before starting with Ubuntu, it is very important for us to understand the differences between Windows and Ubuntu. The most noticeable way is licensing and distribution terms can be used to differentiate between Ubuntu operating systems and Microsoft Windows. Ubuntu is open software and it's completely "free software" free means the freedom to run, use, modify, redistribute copies, and release your improvements to the public. In addition to this it includes many of the software's used for everyday computing at no cost, unlike Windows. Some of these are given below:

Office Suite: A full office suite with a word processor, spreadsheet, and presentation software that can read and write in .doc, .xls, and .ppt formats and can also output to PDF.

Desktop Email Client: Evolution, an email program with a similar interface to Microsoft Outlook.

Web Browser: Firefox, the increasingly popular Web browser.

Others Software: Ubuntu's online Applications Guide lists some Ubuntu-compatible applications that allow you to edit images, listen to and manage music, edit and watch videos, read PDFs.

Updates and bug fixes: Frequent security updates and bug fixes for applications and the operating system are managed by Ubuntu which makes its working robust and bug free.

Start Linux Ubuntu

The following are few important steps and procedures will be helpful to install Ubuntu operating system, however it is important for you to check the minimum hardware requirements for Ubuntu installation as it may varies from version to version. As shown below in the figure 19 this screen appears as first boot screen, as you can see it has many options with an advice that for desktop or laptop installation default installation is suitable, however it can show some more option on press of F1 key.

Fig. 1.19: Ubuntu the First Boot Screen

As we are doing this installation for a desktop, let's prefer default option and press enter. We advise you not to go for server install unless you completely know about it. After you Press Enter, you may find a cluster of internal installation line scroll on the screen as shown below in figure 1.20, but it is normal processing so ignore it.

Fig. 1.20: Ubuntu Internal Installation Process on the Screen

After this Ubuntu will ask few questions to complete your personalised setting for language, time zone, username and password. In addition to these questions there are a few other questions that may cause some confusion. During installation it shows the status of the process being completed as shown below in figure 1.21.

Fig. 1.21: Status of the Process

When it will ask for Partitioning choose the Erase Entire Hard Drive option, if you are not installing Ubuntu with any other operating system on your computer. The boot loader determines which operating system the computer boots to. Hence, It is recommend to install Grub a boot loader to the MBR (Master Boot Record). Now you need to reboot your computer to install the other packages, if you are asked to drive out the installer disk, you can take it out as for further this disk is no more needed. If boot loader Grub is not installed in MBR as advised earlier you have to use boot-floppy or some other boot loader to finish the installation process. Once that finishes, you should get to a login screen as shown below in figure 1.22. Now you need to enter your recently created username and password during the installation process.

Fig. 1.22: Ubuntu Login Screen

1.4.2 Ubuntu File System

Before working on the operating system it is always important to understand the file system i.e. how files and folders or directories are organised in Linux.

Boot Directory: It has the boot information, including the Grub configuration file.

etc Directory: This has a lot of settings for software repositories you use and what other partitions or drives you have "mounted".

Home Directory: This is the only directory you will have access to without using your password to gain temporary administrator privileges. All of your files reside here, along with your settings and preferences, inside a folder called /home/username unless are not installed in a particular folder. So if your username is naveen, your files and settings will be in the /home/naveen folder.

Media Directory: or /mnt are where your media (CDs, DVDs, USB drives, etc.) and mounted partitions would go.

Root Directory: is the /home folder for root and has its own settings.

Usr Directory: is where a lot of stuff is stored that users will be using.

1.4.3 Common Commands and Utilities

The following are some of the basic commands you should know for working on any linux based operating system:

startx	If you happen to end up at a command-prompt without any graphics, you can log in and try typing this command to get back to the graphical (or "x") system.
xkill	Kills a misbehaving application. Once this command is run, the mouse cursor will become a skull and crossbones. Any window you click on after that will close immediately.
alias	Alias is used to substitute a small or more familiar name in place of a long string. It is commonly used for a long strings that are frequently used.
awk	awk utility is powerful data manipulation/scripting programming language (In fact based on the C programming Language). Use awk to handle complex task such as calculation, database handling, report creation etc.
cd	The cd sets the working directory of a process.
chmod	Chmod is a utility that changes the permission of a file.
chown	Chown is a utility that is also used to change file ownership.
cp	The cp command is used to copy files.
date	An essential command to set the date and time. Also a useful way to output current information when working in a script file.
df	The df command reports filesystem disk space usage. With no arguments, 'df' reports the space used and available on all currently mounted filesystems (of all types). Otherwise, 'df' reports on the filesystem containing each argument file.
pwd	To know the current working directory
ln	The ln command makes new, alternate file names for a file by hard linking, letting multiple users share one file. The ln command creates pseudonyms for files which allows them to be accessed by different names. These pseudonyms are called links. There are two different forms of the command and two different kinds of links that can be created.
ls	The ls command shows information about files. It lists the contents of a directory in order to determine when the configurations files were last edited.
man	Short for "manual," man displays information about commands and a keyword search mechanism for needed commands.

passwd	A quick and easy way to change passwords on a system.
Shutdown	Shutdown is a command that turns off the computer and can be combined with variables such as -h for halt or -r for reboot.
top	Top provides an ongoing look at processor activity in real time. It displays a listing of the most CPU-intensive tasks on the system, and can provide an interactive interface for manipulating processes.
vmstat	The vmstat command is used to get a snapshot of everything in a system, helping admins determine whether the bottleneck is CPU, memory or I/O. Run this command to get virtual memory statistics. vmstat reports information about processes, memory, paging, block IO, traps, and cpu activity.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

8) What are various features of Linux based operating system? List them.

....

....

....

....

9) How Ubuntu is different from Windows operating system?

....

....

....

....

1.5 SUMMARY

ICT is changing all most all aspects of our daily life and lifestyle. If we just see lot of computation and communication is happening around us, either when we check the examinations result on mobile or when we withdraw money from the banks ATM. The influences of ICT revolution are also felt in our Library system. In this unit, we introduced you to the computer hardware technology, how it works and what it is. The peripherals devices like RFID and Barcode reader; those are particularly used in libraries are also discussed. In addition, the fundamentals of Ubuntu- Linux based operating systems have been covered. We have given few basic steps, which will be useful for a learner during the installation of Ubuntu operating system. Further this unit explained about the file system of Ubuntu and few commands and utilities, which will be useful for working on this operating system.

1.6 ANSWERS TO SELF CHECK EXERCISES

1) 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 addition, subtraction, multiplication and division, and the logical operations. The control unit interprets instructions and produces the respective control signals.

2) One kilobyte stands for 2^{10} bytes which 1024 bytes so 2 kilobytes will be 2048.

3) Inkjet printers use small cartridges full of different colour inks. They squirt a tiny drop of ink onto the paper using a bubble or pressure to form a dot. All the dots are layered on top of each other to form the desired colour.

4) The important difference between a LED reader and a pen or laser scanner is that the LED reader is measuring emitted ambient light from the bar code whereas pen or laser scanners are measuring reflected light of a specific frequency originating from the scanner itself.

5) RFID technology has gradually begun to replace the conventional barcodes on library items. The RFID tag can contain bibliographic data of a particular book, which replaces the standard barcode reader. It may be used to provide an innovative and easier way for inventory management as well. It can also act as a security device, taking the place of the more traditional electromagnetic security strip. Apart from the books, membership cards could be fitted with an RFID tag.

6) The Basic Input Output System (BIOS) is software stored on a small memory chip on the motherboard. BIOS instructs the computer on how to perform a number of basic functions such as booting and keyboard control. It is also used to identify and configure the hardware in a computer such as the hard drive, optical drive, CPU, memory, etc.

7) CDs and DVDs are optical discs that are used to store data. Both are similar in composition and usage, however DVDs offer more data storage capacity compared to CDs. CDs are commonly used for audio and program files, while DVDs are used for video and program files.

8) Most of the Linux based operating system has the following given features:

- **Low Cost/Free:** Linux and much of its software come with the GNU General Public License and hence it is free and open source.
- **Stability:** Linux has high stability compared with other operating systems. There is no need to reboot the Linux system to maintain performance levels.
- **Performance:** Linux provides high performance on various networks. It has the ability to handle large number of users simultaneously.
- **Networking:** Linux provides a strong support for network functionality; client and server systems can be easily set up on any computer running Linux. It can perform tasks like network backup faster than other operating systems.
- **Flexibility:** Linux is very flexible and can be used for high performance server applications, desktop applications, and embedded systems. You can install only the needed components for a particular use. You can also restrict the use of specific computers.
- **Compatibility:** It runs all common Unix software packages and can process all common file formats.

- **Fast and Easy Installation:** Linux distributions come with user-friendly installation.
- **Better use of Hard Disk:** Linux uses its resources well enough even when the hard disk is almost full.
- **Multitasking:** Linux is a multitasking operating system. It can handle many things at the same time.
- **Open Source:** Linux is an Open source operating systems. You can easily get the source code for Linux and edit it to develop your personal operating system.

9) The most noticeable difference is licensing and distribution terms of Ubuntu operating systems and Microsoft Windows. Ubuntu is open software and it's completely "free software" free means the freedom to run, use, modify, redistribute copies, and release your improvements to the public. In addition to this it includes many of the software's used for everyday computing at no cost, unlike Windows.

1.7 KEYWORDS

ALU : Abbreviation of **a**rithmetic **l**ogic **u**nit, is one component of the CPU (central processing unit) that performs all arithmetic computations, such as addition and multiplication, and all comparison operations.

Barcode : The machine-readable representation of the Universal Production Code (UPC). Bar codes are read by a scanner that passes over the code and registers the UPC. The width of each black line and the subsequent white space between each line coincides with the numbers of the UPC.

Bit : Short for **bi**nary digi**t**, the smallest unit of information on a machine. The term was first used in 1946 by John Tukey. A single bit can hold only one of two values: 0 or 1.

Byte : Abbreviation for binary term, a unit of storage capable of holding a single character. A byte is equal to 8 bits.

CPU : CPU is the abbreviation for **c**entral **p**rocessing **u**nit. Sometimes referred to as the central processor or simply processor, is the brains of the computer where most calculations take place. In terms of computing power, the CPU is the most important element of a computer system.

CU : Short for **c**ontrol **u**nit, it is a typical component of the CPU that implements the microprocessor instruction set. It extracts instructions from memory and decodes and executes them, and sends the necessary signals to the ALU to perform the operation needed.

GRUB : Short for GNU GRand Unified Bootloader is a boot loader package from the GNU Project which provides a user the choice to boot one of multiple operating systems installed on a computer or select a specific kernel configuration available on a particular operating system's partitions.

RFID : Short for **r**adio **f**requency **id**entification, a technology similar in theory to bar code identification. An RFID system consists of an antenna and a transceiver, which read the radio frequency and transfer the information to a processing device, and a transponder, or tag, which is an integrated circuit containing the RF circuitry and information to be transmitted. RFID systems can be used where a unique identification system is needed.

1.8 REFERENCES AND FURTHER READING

Basandra, Suresh K. *Computer Today*. New Delhi: Galgotia Publication, 2003. Print

Leon, Alexis and Leon, Mathews. *Fundamentals of Information Technology*. New Delhi: Leon TechWorld publication, 1999. Print

Long, Larry. *Computer Fundamental*. New Delhi: Wiley Dreamtech Publication, 2004. Print

Saxena, Sanjay. *MS-Office for Everyone*. New Delhi, Vikas Publication House, 2001. Print

Websites:

http://www.ubuntuforums.org/

http://www.ubuntulinux.org/

UNIT 2 BASICS OF COMMUNICATION TECHNOLOGY

Structure

2.0 OBJECTIVES

After going through this Unit, you will be able to:

- understand the concept of Data Transmission terminology;
- differentiate between Serial and Parallel communication;
- differentiate between Analog and Digital Data Transmission;
- know different types of transmission modes;
- compare the different Transmission Media and their characteristics;
- explain the characteristics and working of devices used for data communication; and
- understand communication protocols/standards and realise its importance.

2.1 INTRODUCTION

We all are familiar with the term "Communication", its meaning is contextual and can be interpreted by different words. In general sense we can say the process of sharing ideas, information, and messages with others at a particular time and place is communication. Communication is a vital part of personal life and is also important in business, education, and any other situation where people encounter each other. Communication between two people is an outgrowth of methods developed over centuries of expression. Gestures, the development of language, and the necessity to engage in joint action all play a part. Communication, as we see it today, has evolved a long way. In this unit we will discuss the primitive modes of computer communication used for data transfer in networks and Internet. Also, this unit will cover the devices used for computer communication. All communication between devices requires that the devices agree on the format of the data. The set of rules defining a format is known as a protocol. In the end of this unit we have briefly given the details of communication protocols/standards used for data transfer between computers particularly for Internet.

2.2 ANALOG AND DIGITAL COMMUNICATION

Communication from a source to a destination, that is, from one computer to another or from one device to another, involves the transfer of information from the sender to the receiver. The transfer of data from one machine to another machine such that, the sender and the receiver both interpret the data correctly is known as Data Communication. Broadly the data communication can be divided into two types Analog and Digital. As you may know that signals carry the data for communication, hence this classification of Analog and Digital is based on the signals which are used for data transfer.

2.2.1 Analog Signal

We mostly use analog signals in our day-to-day life. Whatever we are talking meaning voice signals, Radio and TV broadcasting signals, Electricity signals, Sunlight or other lights signals, all are nothing but examples of Analog signal. We can say Analog signals are Nature-Signals.

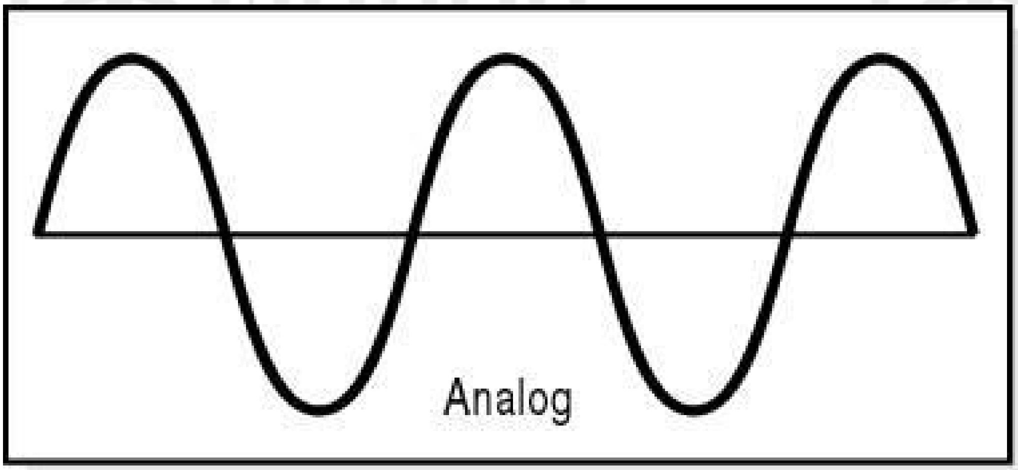


Fig. 2.1: Analog Signal

Technically, Analog signals vary constantly in one or more values; these changes in values can be used to represent data as depicted in the figure 2.1. An analog signal is continuous and can be represented by using sine waves. Human voice generates an analog (continuously varying) signal containing multiple frequencies that is transmitted as an analog signal over the medium.

It is important to note the limitations of analog signals. We know that signals are propagated with the energy and it get weaker with the distance therefore, in long distance communication like telephones, radio and TV communication we use a device named "Amplifiers". Amplifiers are used to overcome the attenuation that the signal suffers on its way. However, amplifiers amplify noise along with the original signal and hence, if the signal gets distorted, it cannot be reconstructed and it is a permanent loss. Due to this reason, this type of transmission is not used where a high level of accuracy is needed. This is used in telephony where a slight distortion in human communication does not matter.

The ability to capture the slight nature of the real world is the single advantage of analog techniques. However, once captured, modern electronic equipment, no matter how advanced, cannot copy analog signals perfectly. By converting analog signals into digital, the original audio or video data can be preserved indefinitely within the specified error bounds and copied over and over without deterioration. Once continuously varying analog signals are measured and converted into digital form, they can be stored and transmitted without loss of integrity due to the accuracy of digital methods.

2.2.2 Digital Data Transmission

Digital data transmission describes any system based on discontinuous data or events. Computers are digital machines because at their most basic level they can distinguish between just two values, 0 and 1, or off and on, or high and low voltage as depicted in figure 2. There is no simple way to represent all the values in between, such as 0.25. All data that a computer processes must be encoded digitally, as a series of zeroes and ones.

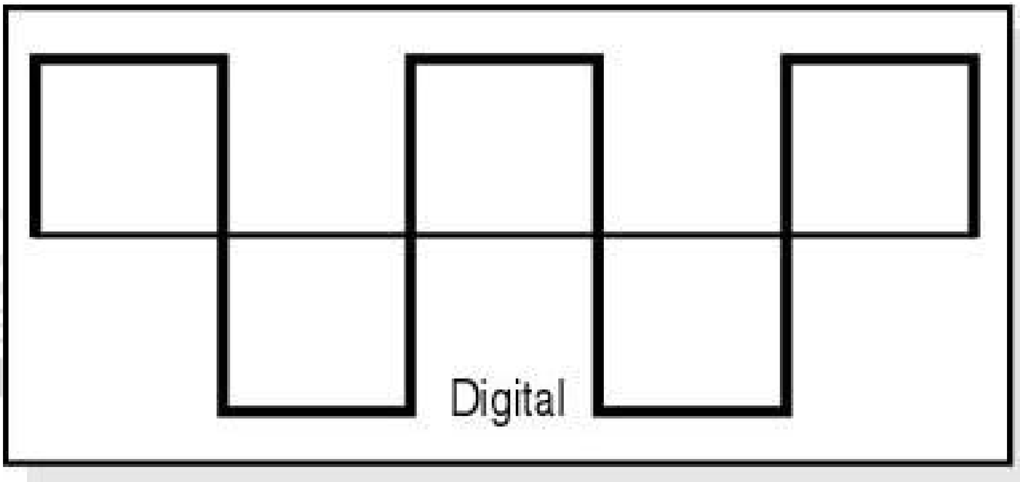


Fig. 2.2: Digital Signal

Information coming out of the computer is in the form of digital signals. The bandwidth of a digital signal is infinite as compared to any medium, which has a limited bandwidth. Therefore, as the signal is generated and enters the medium, at that point of entry, only limited frequencies are permissible on the medium and this depends upon the bandwidth. As the signal traverses over the medium it gets distorted and beyond a certain distance, the signal becomes unrecognizable from the original one. A hardware device called Repeater is used to regenerate the digital signal. The repeater measures the signal values at regular intervals to recognise the 0's and 1's in the signal and regenerates them. Hence, there is no loss of information. The number of repeaters to be used depends on the distance between the source and the destination. Any line with repeaters placed at appropriate distance is called a digital line.

When information, music, voice and video are turned into binary digital form, they can be electronically manipulated, preserved and regenerated perfectly at high speed. The millionth copy of a computer file is exactly the same as the original. This is, nevertheless, a major advantage of digital processing.

2.3 DATA COMMUNICATION MODES

Data can be transmitted from Source to Destination in a number of ways. The different modes of data transmission are outlined as follows:

- Asynchronous, Synchronous and Isochronous Communication.
- Simplex, Half Duplex and Full Duplex Communication.

2.3.1 Asynchronous, Synchronous and Isochronous Communication

When we communicate to each other, what is a basic fundamental principle? It is synchronisation of both persons, means when I am talking to you, you should listen to me and once I complete my words then you should deliver your speech otherwise we both will not be able to communicate. The same principle applies to computer or device communication and perhaps it is very important for computers are not intelligent enough to handle the cross talks. The three mechanisms used for synchronisation are Asynchronous, Synchronous and Isochronous Communication.

Asynchronous Communication

Asynchronous communication sends individual characters one at a time framed by a start bit and 1 or 2 stop bits. Each frame begins with a start bit that enables the receiving device to adjust to the timing of the transmitted signal. The message can begin at any time. Here, messages are kept as short as possible because, the sending and receiving devices should not draft out of synchronisation, when the message is being transferred. Asynchronous communication is most frequently used to transmit character data and is ideally suited for characters that are transmitted at irregular intervals, such as, when users are typing in character data from the keyboard.

Start Bit	Data Bits (7 or 8 bits)	Parity Bit (Optional)	Stop Bit or Bits

Fig. 2.3: A Typical Frame Used to Transmit a Character

A typical frame used to transmit a character data has four components as shown in figure 2.3. Start bit is used to signals the starting a frame and enables the receiving device to synchronise itself with the message. Data Bits consists of 7 or 8 bits when character data is being transmitted. Parity Bit is an optionally used as a crude method for detecting transmission errors. Stop bit or bits signals the end of the data frame. You must be curious to know about Parity Bit. Error detection in asynchronous transmission makes use of the parity bit. Parity techniques can detect errors that affect only one bit and if two or more bits are affected by errors, the parity techniques may not be able to detect them. Asynchronous transmission is simple, inexpensive and is ideally suited for transmitting small frames at irregular intervals (e.g., typing from a keyboard). As each individual character is complete in itself, if a character is corrupted during transmission, its successor and predecessor will not be affected.

We would like to tell you that this type of transmission is not suitable for transferring large amounts of data or Audio and Video. Do you want to know, why? As we have discussed above that start, stop and parity bits need to be added to each character that is to be transmitted. This adds a high overhead to transmission and wastes the bandwidth; as a result, asynchronous transmission is undesirable for transmitting large amounts of data. Successful transmission inevitably depends on the recognition of the start bits.

These bits can be easily missed or occasionally spurious, as start bits can be generated by line interference, the transmission may be unsuccessful. Due to the effects of distortion the speed of asynchronous transmission is limited. If asynchronous transmission has these problems, then what are the alternatives do we have? Let us discuss this.

Synchronous Communication

An alternative way to transfer large data is, instead of sending single character or few bit we can send a group or block of data at once, and this methods is called synchronous communication.

Here, transmission begins at a predetermined regular time instant. A synchronous signal is used to tell the receiving station that a new frame is arriving and to synchronise the receiving station. In simple language you can say it is something like we inform to our friend that at particular time you will send some books, so be at home at that time to receive books. This information is nothing but a "synchronous signal". It is important to note that in computer communication Synchronous signals can be sent along with the message therefore, it generally utilises a bit pattern that cannot appear elsewhere in the messages, ensuring that they will always be distinct and easy for the receiver to recognise. As the transmitter and receiver remain in synchronisation for the duration of the transmission, frames can be of longer length.

As frames are longer the parity method of error detection is not suitable because, if multiple bits are affected, then, the parity technique will not report error accurately. Hence, the technique used with synchronous transmission is the Cyclic Redundancy Check (CRC). The transmitter uses an algorithm to calculate a CRC value that summarises the entire value of data bits. This CRC value is appended to the data frame. The receiver uses the same algorithm, recalculates the CRC and compares the CRC in the frame to the value that it has calculated. If these values match then, it is sure that the frame was transmitted without error.

An end bit pattern indicates the end of the frame. Like sync the bit pattern for end is such that, it will not appear elsewhere in the messages, ensuring that they will always be distinct and easy for the receiver to recognise at the end of the frame.

Synchronous transmission is more efficient because, only 4 additional bytes (for start and end frames) are required to transmit upto 64 k bits. Synchronous transmission is not really prone to distortion; as a result, it can be used at high speeds. However, even with these advantages synchronous transmission has some other limitations, first it is expensive as complex circuitry is required and it is difficult to implement. If an error occurs during transmission, rather than just a single character the whole block of data is lost. The sender cannot transmit characters simply, as they occur, but has to store them until it has built up a block. Thus, this is not suitable where characters are generated at irregular intervals.

Isochronous Communication

This is another alternative mechanism for data transmission, which combines the approaches of asynchronous and synchronous communications. As in the asynchronous method, each character has both the start and stop bits. The idle period (where no transmission takes place) between the two characters is not random but an exact multiple of one character time interval. If, the time to transmit a character (including its parity, start, stop bits) is **t**, the time interval between characters cannot be random as in the asynchronous method. It is also not 0 as in the synchronous method. It has to be t, 2t, 3t………nt , where n is any positive integer. Here, the signal is expected to be received within certain delay bounds say Tmin to Tmax. Isochronous transmission guarantees transmission rates, and it is almost deterministic. It has low overheads and high speed. However, in isochronous transmission it is necessary to ensure that the clocking device is fault tolerant.

2.3.2 Simplex, Half Duplex and Full Duplex Communication

Based on "what point of time" sender or receiver can transfer the data, and what is a role of sender or receiver in the data communication; we have classified data transmission into three categories Simplex, Half Duplex and Full Duplex.

Simplex

This is one of the simplest techniques for data communication, in which we have one permanent sender and permanent receiver. In Simplex transmission, at any point of time one of the communicating devices can only send data, whereas the other can only receive it. Can you think about some example around?

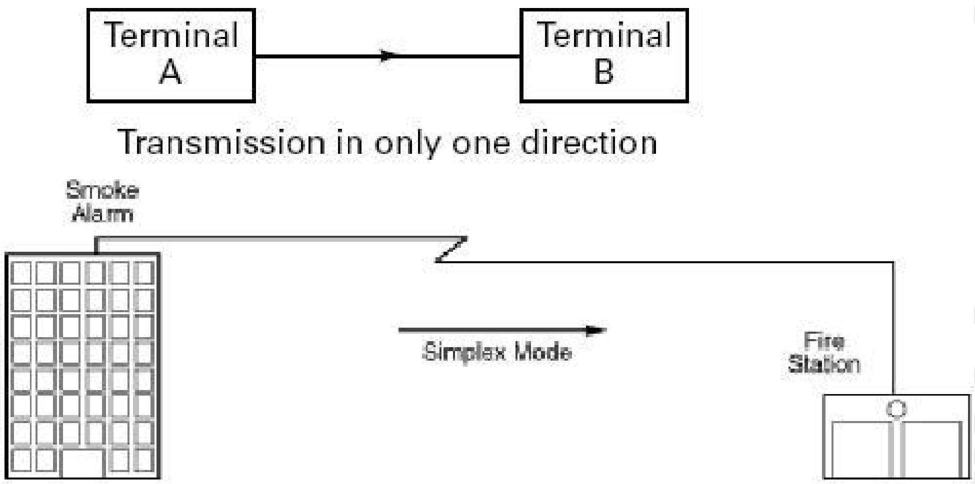


Fig. 2.4: Simplex Communication

It is like Sun, which permanently send light to us; it is also used in Broadcasting TV and Radio signals where, you can receive data from stations but can't transmit data back. These signals are unidirectional, where one party is the transmitter and the other is the receiver as shown in the Figure 2.4, where smoke alarm is shown which is connected with fire station, it is also an example of unidirectional communication. This type of channel design is easy and inexpensive to set up.

Half Duplex

As we know data communication is more about exchanging and sharing data, which may not be fulfilled by the simplex communication. Now, let's take an example where, communication is from both ways but at any point of time it is not simultaneous. This type of two-way communication where, only one party can transmit data at a time is called Half Duplex. Unlike, the Simplex mode here, both devices can transmit data though, not at the same time. Half duplex provides Simplex communication in both directions in a single channel as shown in figure 2.5, with an example of walky-talky communication.

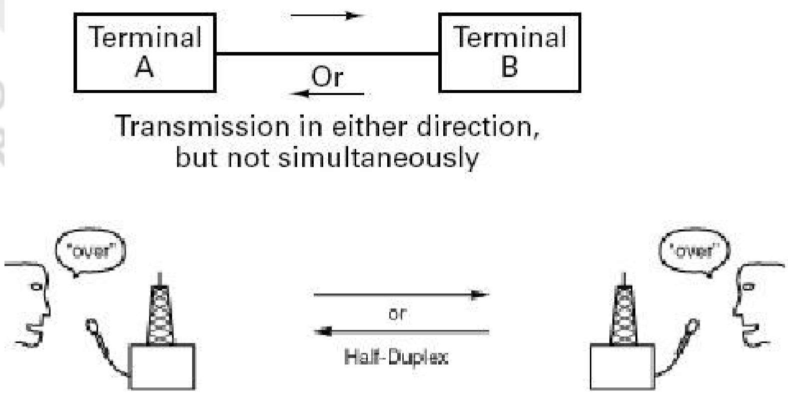


Fig. 2.5: Half Duplex Connection

When one device is sending data, the other device must only receive it and vice versa. Thus, both sides take turns at sending data. This requires a definite turnaround time during which, the device changes from the receiving mode to the transmitting mode. Due to this delay, half duplex communication is slower than simplex communication. However, it is more convenient than simplex communication as both the devices can send and receive data.

The difference between simplex and half-duplex refers to two-way communication where, only one party can transmit data at a time. Simplex refers to one-way communication where, one party is the transmitter and the other is the receiver. For example, a walkie-talkie is a half-duplex device because only one party can talk at a time. Let's try to establish an analogy with our real world regarding this communication mechanism. We can say that the half-duplex communication is something like rail tracks, where at one point of time only one train will pass through track. It can be used for both coming and going but not together.

Full Duplex

As the half duplex communication inherit the delay in data communication. Another mechanism of data communication mechanism is full duplex, refers to the transmission of data in two directions simultaneously. Here, both the devices are capable of sending as well as receiving data at the same time. As you can see from figure 2.6, that simultaneously bi-directional communication is possible, as a result, this configuration requires full and independent transmitting and receiving capabilities at both ends of the communication channel, one of its important examples is our network communication and Internet.

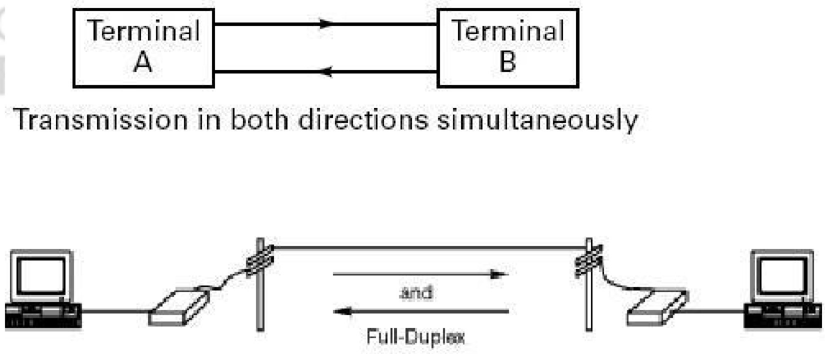


Fig. 2.6: Full Duplex Connection

Sharing the same channel and moving signals in both directions increases the channel throughput without increasing its bandwidth. For example, a telephone is a full-duplex device because both parties can talk to each other simultaneously. In contrast, a walkie-talkie is a half-duplex device because only one party can transmit at a time. It can also be view as bi-directional Road tracks. Most modems have a switch that lets you choose between full-duplex and half-duplex modes. The choice depends on which communications program you are running.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) What is Digital Communication? How it is better than Analog Communication?

....

....

....

....

2) Bring out the difference between Synchronous, Asynchronous and Isochronous transmission.

....

....

....

....

2.4 COMMUNICATION HARDWARE

A smallest computer network should comprise minimum of at least two computers (sender and receiver), communication channel that connect the sender and receiver to each other, a network interface device on each computer which is also known as NIC, some intermediate communication devices like Switch, which is used to switch the data from one point to another, In this section we will cover all the hardware components used for designing a computer network.

2.4.1 Communication Channels

One of the most common ways to transport data from one computer to another is to write them onto magnetic tape or floppy disks, physically transport the tape or disks to the destination machine, and read them back in again. While, this method is not practical, imagine the data communication delay, cost require to transportation of data between countries and continents. Normally, Cable is the medium through which information usually moves from one network device to another. There are several types of cable, which are commonly used with LANs. In some cases, a network will utilize only one type of cable; other networks will use a variety of cable types. The type of cable chosen for a network is related to the network's topology, protocol, and size, these concepts of networks you will study in the next Unit. Understanding the characteristics of different types of cable and how they relate to other aspects of a network is necessary for the development of a successful network.

Twisted Pair

Twisted pair cable comes in two varieties: unshielded twisted pair and shielded twisted pair. Unshielded twisted pair (UTP) is the most popular and is generally the best option for small networks UTP cables contain four twisted-pairs (Orange, green, Blue & Brown) as shown in the figure 2.7, enclosed in a common sheath.

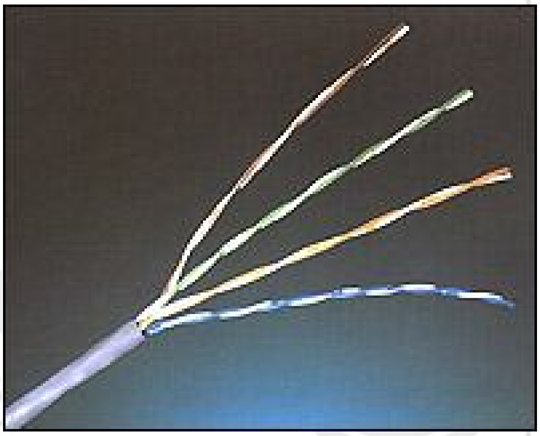


Fig. 2.7: Unshielded Twisted Pair Cable

The most common application of the twisted pair is in the telephone system. A twisted pair connects nearly all telephones to the telephone company office. Twisted pairs can run several kilometers without amplification, but for longer distances, repeaters are needed. Twisted pairs can be used for either analog or digital transmission. The bandwidth depends on the thickness of the wire and the distance traveled. A disadvantage of UTP is that it may be susceptible to radio and electrical frequency interference. Shielded twisted pair (STP) is suitable for environments with electrical interference; however, the extra shielding can make the cables quite bulky. Shielded twisted pair is often used on networks using Token Ring topology.

Coaxial Cable

Another common transmission medium is the coaxial cable. It has better shielding than twisted pairs, so it can span longer distances at higher speeds. A coaxial cable consists of a stiff copper wire as the core, surrounded by an insulating material. The insulator is encased by a cylindrical conductor, often, as a closely woven braided mesh. The outer conductor is covered with a protective plastic sheath. A cutaway view of a coaxial cable is shown in figure 2.8.

Fig. 2.8: Coaxial Cable

The construction and shielding of the coaxial cable gives it a good combination of high bandwidth and excellent noise immunity. The bandwidth possible depends on the cable length. For 1-km cables, a data rate of 1 to 2 Gbps is feasible. Longer cables can also be used, but only at lower data rates or with periodic amplifiers. Coaxial cables used to be widely used within the telephone system but have now largely been replaced by fiber optics on long-haul routes. Coax is still widely used for cable television and some local area networks.Although coaxial cabling is difficult to install, it is highly resistant to signal interference. In addition, it can support greater cable lengths between network devices than twisted pair cable. The two types of coaxial cabling are:

Thick coaxial: Thick coaxial cable is also referred to as thicknet. 10Base5 refers to the specifications for thick coaxial cable carrying Ethernet signals.

Thin coaxial: Thin coaxial cable is also referred to as thinnet. 10Base2 refers to the specifications for thin coaxial cable carrying Ethernet signals.

Optical Fiber

Fiber optic cabling consists of a center glass core surrounded by several layers of protective materials as depicted below in figure 2.9. It transmits light rather than electronic signals eliminating the problem of electrical interference. This makes it ideal for certain environments that contain a large amount of electrical interference. It has also made it the standard for connecting networks between buildings, due to its immunity to the effects of moisture and lighting.

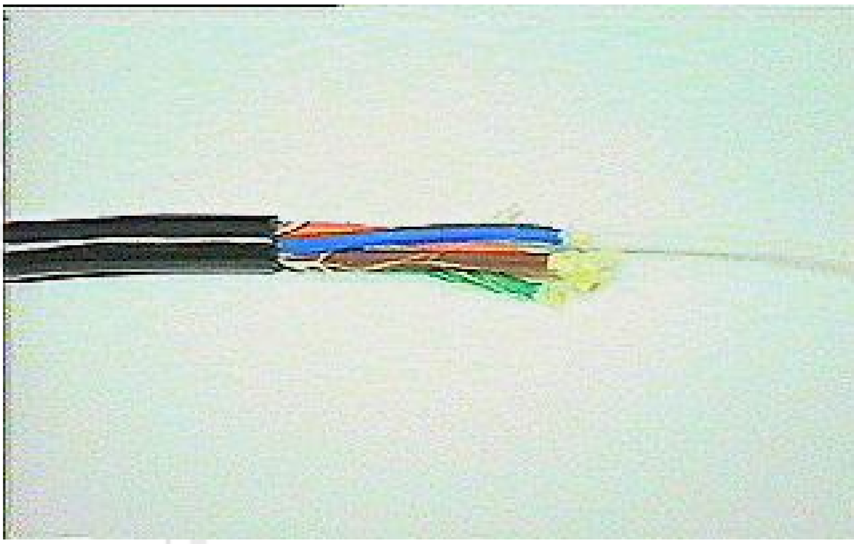


Fig. 2.9: Optical Fiber

Fiber optic cable has the ability to transmit signals over much longer distances than coaxial and twisted pair. It also has the capability to carry information at vastly greater speeds. This capacity broadens communication possibilities to include services such as video conferencing and interactive services. The cost of fiber optic cabling is comparable to copper cabling; however, it is more difficult to install and modify. 10BaseF refers to the specifications for fiber optic cable carrying Ethernet signals.

2.4.2 Communication Devices

As discussed in previous section there are several communication devices used in data communication. In the sender and receiver computers, we should have Network Interface Card and Modem (which we have already studied in the Unit Basics of Computer Technology) before connected to the Internet connection. There are many other devices being used at different levels and in different types of networks, those are explained in this section.

Network Interface Card (NIC)

A NIC is also known as a network card as shown in figure 2.10. It connects the computer to the cabling, which in turn links all of the computers on the network together as shown in figure. Each computer on a network must have a network card. Most modern network cards are 10/100 NICs and can operate at either 10 Mbps or 100 Mbps. It access to a networking medium and often provides a low-level addressing system through the use of MAC addresses. It allows users to connect to each other either by using cables or wirelessly.

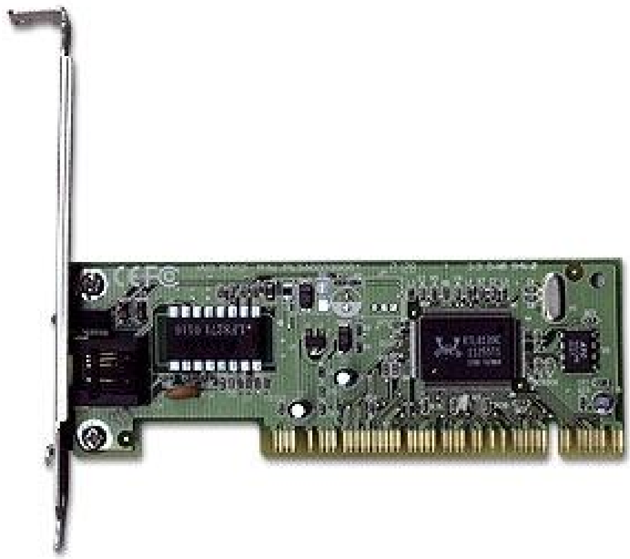


Fig. 2.10: A Network Interface Card (NIC)

Repeater

A **repeater** is an electronic device that receives a signal and retransmits it at a higher level or higher power, or onto the other side of an obstruction, so that the signal can cover longer distances without degradation, an example is shown in the figure 2.11.

Fig. 2.11: Examples of Repeater

Hub

A hub is a device used to connect a PC to the network. The function of a hub is to direct information around the network, facilitating communication between all connected devices, means A hub contains multiple ports as shown in figure 2.12, which is used to connect devices in a star topology. When a packet arrives at one port, it is copied to all the ports of the hub.

Fig. 2.12: Example of Hub

Hubs can improve performance, especially for bursty traffic and large file transfers, it also enables optimum performance of computers. However, total bandwidth remains fixed; as network traffic grows, performance suffers and it does not reduce collisions.

Switch

A switch is a data-link layer network device that forwards frames using MAC addresses in the header of frames. A common type of switch is shown in the figure 13. It is used to improve network performance by segmenting the network and creating separate collision domains. It can also reduce competition for bandwidth.

Fig. 2.13: Switch

During new installations switches should be used instead of hubs as they are more effective and provide better performance. A switch, which is often, termed a 'smart hub'. Switches and hubs are technologies or 'boxes' to which computers, printers, and other networking devices are connected. Switches are the more recent technology and the accepted way of building today's networks. With switching, each connection gets "dedicated bandwidth" and can operate at full speed. In contrast, a hub shares bandwidth across multiple connections such that activity from one PC or server can slow down the effective speed of other connections on the hub. It Supports VLAN's (virtual local area network (VLAN) is a logical grouping of hosts on one or more LANs that allows communication to occur between hosts as if they were on the same physical LAN.

Howver, switches are not as good as a router in limiting Broadcasts, also for handling Multicast packets needs quite a bit of configuration & proper designing.

Bridges

The main network device found at the data link layer is a bridge, mainly used for connecting two network or local area networks. This device works at a higher layer than the repeater and therefore is a more complex device. It has some understanding of the data it receives and can make a decision based on the frames it receives as to whether it needs to let the information pass, or can remove the information from the network. This means that the amount of traffic on the medium can be reduced and therefore, the usable bandwidth can be increased. Bridges are store and forward devices to provide error detection; a common type of bridge is shown in the figure 2.14.

Fig. 2.14: A Common Bridge

Routers

Router is a device which is use to create an internetworking in our WANs and LANs. The main purpose of router is routing, which means taking a packet from one device and sending it through the network to another device on a different network. Packet is often a confusion term; it is nothing but the unit of data at any layer of the protocol stack, prior to, or after transmission.

A packet normally contains Data Type, Packet Count, Recipient's IP address, Sender's IP address, and Data. Actually routers need to transfer the data from source machine to destination machine but other information in packets it need during its transportation process. A common type of router is shown in the figure 2.15.

Fig. 2.15: A Common Router with its Front and Back View

Router has nothing to care about the host machine; it need to care about networks and the best path to each network. The logical network address of the destination host is used to get packets to a network through a routed network, and then the hardware address of the host is used to deliver the packet from a router to the correct destination host.

2.5 COMMUNICATION PROTOCOLS/STANDARD

Before going through the origin of the Internet, let us examine what is communication. **Communication** is a process of sharing ideas, information, and messages with others at a particular time and place. Communication is a vital part of personal life and is also important in business, education, and any other situation where people encounter each other. Communication between two people is an outgrowth of methods developed over centuries of expression. Gestures, the development of language, and the necessity to engage in joint action all play a part. Communication, as we see it today, has evolved a long way. We will discuss the primitive modes of communication briefly.

2.5.1 Communication Systems

Early societies developed systems for sending simple messages or signals that could be seen or heard over a short distance, such as drumbeats, fire and smoke signals, or lantern beacons. Messages were attached to the legs of carrier pigeons that were released to fly home. Further,

The postal system are being developed to transfer written documents to destinations around the world. Even after implementing different electronic communication mediums, postal system is still one of the popular communication systems available. The first truly electronic medium for communication was developed first in the form of telegraph, which sent and received electrical signals over long-distance wires. After telegraph the most important invention is telephone systems use to transmitting human speech. The telephone network has also provided the electronic network for new computer-based systems like the Internet facsimile transmissions, and the World Wide Web.

The development of computer networks was started when computers became faster, more-powerful and smaller. In the 1960's the Advanced Research Projects Agency (ARPA) of the U.S. Department of Defence, along with researchers working on military projects at research centres and universities across the country, developed a network called the ARPANET, for sharing data and processing time of uniform computer connection over specially equipped telephone lines and satellite links. The network was designed to survive the attack or destruction of some of its parts and continue to work. Soon, however, scientists using the ARPANET realised that they could send and receive messages as well as data and programs over the network.

Today, the Internet is the widely known computer network. It uses interconnection of computer system by both wired and wireless. Smaller networks of computers, called Local Area Networks (LANs), can be installed, in a single building or for a whole organisation. Wide Area Networks (WANs) can be used to span a large geographical area. LANs and WANs use telephone lines, computer cables, and microwave and laser beams to carry digital information around a smaller area, such as a single college campus. Internet can carry any digital signals, including video images, sounds, graphics, animations, and text, therefore it has became very popular communication tool.

Need of Internet

The main reason is that each computer network is designed with a specific purpose. For example, LAN is used to connect computers in a smaller area, and it provides fast communication. As a result, networks become specialised identify. In many cases, these networks do not use the same hardware and software technology. It means that, a computer can communicate with the computers attached to the same network, because they are inter-compatible. As more and more organisations had multiple computer networks, this became a major issue. As a result, the concept of internetworking (internet) came into being. This means that there should be a network of all physically separate networks.

Need of Protocols/Standards

All methods of communication described above follow a protocol. Protocol is nothing but a convention or standard. To signify that "Everything is ok and the train can start by a green flag is also a protocol. When we write a letter we follow a protocol. If we look at them carefully, we will find that protocols normally have hidden layers. A good example is human conversation over the phone which can be used as an analogy for communication using computers.

Assume that X and Y, want to have conversation over the telephone about a cricket match. We call this an idea. Assume that each person is taking down what other has to say. Thus, the conversation takes place in the form of several messages. A message is a block of sentence. It could also consist of one world such as OK, yes denoting a positive **acknowledgement** (ACK). It could also mean a **negative acknowledgement** (NAK) or a request to repeat such as come again, pardon me. All this happens both ways.

At the level of idea, X and Y feel that they are discussing a cricket match. However, in reality, the conversation consists of a number of messages.

A message could be too long. It may not be wise for X to speak for half an hour, only to receive a request to repeat. It is therefore necessary to send/receive acknowledgements after each sentence like 'ok', 'come again' etc. A sentence is analogous to a packet in computer world. The sender X will not speak until s/he hears some form of acknowledgement, or will repeat the sentence if s/he receives a negative acknowledgement. An alternative is timeout strategy. The speaker speaks a sentence and waits for some time for any acknowledgement. If s/he does not hear anything, s/he repeats the sentence.

Apart from this **error control** we take care of **flow control**. Flow control refers to the speed mismatch between the listener and speaker. If the speaker speaks too fast, the listener will say go-slow. In computer world, if the receiving computer is not fast enough, and cannot hold any more data, it requests the sender to wait or control the transfer by slowdown.

Therefore, in computer communication, both speaker and listener should agree on the communication language/syntax, scheme of acknowledgement, during flow control, machine error control mechanism, etc. Thus, we can say that the conversation is governed by some set of rules known to both the parties. This set of rules is called protocol and it necessary for disciplined manner of conversation/communication.

2.5.2 TCP/IP Layers and Protocols

To provide the set of rules and standardization among the different computer networks, the International Standards Organization (ISO) in 1984 defined the Open System Interconnection (OSI) model. OSI model is a set of protocols that attempt to define and standardize the data communications process; we can say that it is a concept that describes how data communications should take place. The OSI model has the support of most major computer and network vendors, many large customers, and most governments in different countries. The following are the seven layers of the Open System Interconnection (OSI) reference model, depicted in the figure 2.16:

Layer 7 — Application layer

Layer 6 — Presentation layer

Layer 5 — Session layer

Layer 4 — Transport layer

Layer 3 — Network layer

Layer 2 — Data Link layer

Layer 1 — Physical layer

Fig. 2.16: OSI Layers

The OSI model is modular. Each successive layer of the OSI model works with the one above and below it. Although, each layer of the OSI model provides its own set of functions, it is possible to group the layers into two distinct categories. The first four layers i.e., physical, data link, network, and transport layer provide the end-to-end services necessary for the transfer of data between two systems. These layers provide the protocols associated with the communications network used to link two computers together. Together, these are communication oriented. The top three layers i.e., the application, presentation, and session layers provide the application services required for the exchange of information. That is, they allow two applications, each running on a different node of the network to interact with each other through the services provided by their respective operating systems. Together, these are data processing oriented.

Though OSI model is very important for data communication but multiple layers slow down the communication process. Practically OSI was never fully implemented; to provide robustness and better efficiency another but similar model was implemented named as TCP/IP model (its named was given because of its two most important protocols in it). The TCP/IP model is made up of four layers: interface layer, network, transport, and application. The first layer of TCP/IP (Application layer) is similar to the first three layers (Application, presentation and Session layer) of the OSI model. The services of transport layers of both the models are similar. Further, the services of network layers in both models are also similar, while some time network layer is also known as Internet layer. The last layer of TCP/IP is interface layer, which includes the services of data link layer and physical layer of OSI model. In OSI model, each layer takes the services of the lower layer. Whereas the layers of TCP/IP protocol suite contain relatively independent protocols as shown in figure 2.17.

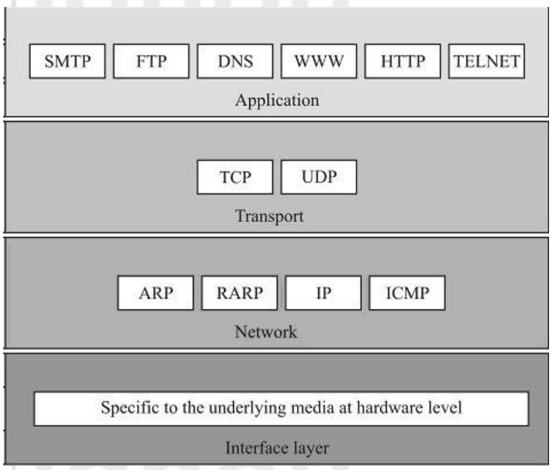


Fig. 2.17: TCP/IP and its Protocols

Layers of TCP/IP Protocol Suite

As we know TCP/IP contains four layers and each layer has its specific functions, in the following section lets find out the functions of each layer of TCP/IP.

Interface layer or (Physical + Data Link Layer)

The physical layer deals with the hardware level like, transmission media, connections and the voltage for digital signals. The data link layer deals with media access and control strategies, frame format etc.

Internet Layer or Network Layer

The Internet layer is an important layer in the protocol suite. At this layer, TCP/IP supports Internetworking Protocol (IP). IP is a host-to-host protocol. This layer is responsible for the format of datagrams as defined by IP, and routing a datagram or packet to the next hop, but is not responsible for the accurate and timely delivery of datagrams to the destination in proper sequence. IP allows raw transmission functions allowing user to add functionalities necessary for given application. Ensuring maximum efficiency, TCP/IP supports four other protocols: ARP, RARP, ICMP and IGMP in this layer.

1) **Address Resolution Protocol (ARP)**

On a LAN, each machine is identified with a unique physical address imprinted on the network interface card. ARP is used to find the physical address of a machine when its IP address is known.

2) **Reverse Address Resolution Protocol (RARP)**

It is used to find the IP address of a machine when its physical address is known. It is used when a diskless computer is booted or a computer is connected to the network for the first time.

3) **Internet Control Message Protocol (ICMP)**

IP is unreliable are best effort delivery. In case of failures ICMP is used to send notifications to the sender about packet problems. It sends error and query messages.

4) **Internet Group Message Protocol (IGMP)**

It is used for multicasting, which is transmission of a single message to a group of recipients.

Transport Layer

At this layer, TCP/IP supports two protocols: TCP, UDP, IP is host-to-host protocol, which can deliver the packet from one physical device to another physical device. TCP, UDP, are transport level protocols, responsible for delivering a packet from one process on a device to another process on the other device.

1) **User Datagram Protocol (UDP)**

It is simpler of the two protocols. It does not provide reliability. It is, therefore faster, and using for applications in which delay is intolerable (in case of audio and video).

2) **Transmission Control Protocol (TCP)**

TCP is reliable, connection-oriented protocol. By connection oriented, we mean that a connection must be established between both ends before either can transmit data. It ensures that communication is error-free and in sequence.

Application Layer

As said earlier, it is closer to combined session, presentation, and application layer of OSI model. It allows the user to run various applications on Internet. These applications are File Transfer Protocol (FTP), remote login (TELNET), email (SMTP), WWW (HTTP). The session layer of OSI model is almost dropped in TCP/IP. Whenever we work on the Internet we use most of these protocols. For example when use are surfing or browsing the Internet websites actually you are using WWW.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) Write any three important advantages of Optical fibers.

....

....

....

....

4) Differentiate between Switches and hubs.

....

....

....

....

2.6 SUMMARY

In this unit, we have studied the basic concepts of Data Transmission. As it is essential to know how data can be transmitted from the source to the destination, the different types of transmission were discussed like Asynchronous, Synchronous and Isochronous communication. Further we have explained different modes like Simplex, Half duplex and full duplex communication. How the network are designed and for this are the different components are required, to explain it we have discussed the characteristics and working of various communication channels and devices in this unit. Towards the end, we have focused about the communication protocols and standards used for communication in the networks and Internet.

2.7 ANSWERS TO SELF CHECK EXERCISES

1) Digital communication is information transmitted electronically and is encoded digitally. Digital signals are signals that are represented by binary numbers, “1” or “0”. It is stored and processed by computers. The main difference between analog and digital signals is that an analogue signal is continuous and a digital signal is discrete. Analog technologies record waveforms as they are, while digital technologies convert analog signals into numbers. Digital communication systems offer much more efficiency, better performance, and much greater flexibility than analog.

2) Asynchronous communication sends individual characters one at a time framed by a start bit and 1 or 2 stop bits. Transfer of large data sent in a group or block at once instead of sending single character or few bit at a time is known as synchronous communication. Isochronous communication combines both approaches.

3) Three advantages of Optical fibre are: a) since it transfer light instead of electronic signal, it eliminates electrical interference, b) it is able to transmit signals over much longer distance compared to coaxial or twisted pair cable, and c) cost is much lower compared to copper cabling.

4) A hub is a device used to connect a PC to the network whereas a switch is a data link layer network device.

2.8 KEYWORDS

Coaxial Cable : A type of electric cable that is used to send telegraph, telephone, and television signals.

MAC Addresses : A media access control address is a unique identifier assigned to network interfaces for communications on the physical network segment. MAC addresses are used as a network address for most IEEE 802 network technologies, including Ethernet.

Open System Interconnection (OSI) Reference Model : A set of seven layers that define the different stages that data must go through to travel from one device to another over a network. OSI model is a set of protocols that attempt to define and standardize the data communications process.

Shielded Twisted Pair (STP) : Cable construction that includes an external grounded shield as well as twisting on a regular basis to help minimize noise interferences.

Unshielded Twisted Pair (UTP) : Copper wiring used in small-to-large networks to connect host devices to hubs and switches.

2.9 REFERENCES/FURTHER READING

Tenenbaum, Andrew S. *Computer Networks*. PHI: New Delhi, 2006. Print

Stalling,William. *Data and Computer Communication*. PHI: New Delhi, 2007. Print

Godbole, Achyut S. *Web Technoligies*. New Delhi: TATA McGrawHill, 2009. Print

Forouzan, Berhouz. *TCP/IP Protocol Suite*. New Delhi: TATA McGraw Hill, 2003. Print

Websites:

http://compnetworking.about.com/lr/

http://en.wikipedia.org/w/index.php

http://www.cisco.com

http://www.tcpipguide.com

UNIT 3 BASICS OF NETWORK TECHNOLOGY

Structure

3.0 OBJECTIVES

After going through this Unit, you will be able to:

- explain the concept of computer networks;
- understand different application of networks;
- differentiate between different types of computer networks based on size, connection and functioning;
- compare the different network topologies used in LAN and WAN;
- understand the working of LAN access methods;
- explain the working of networking devices used in WAN;
- know the importance of using networked system; and
- understand the concept of wireless technologies and standards.

3.1 INTRODUCTION

With the ICT revolution the functioning of organisations has changed drastically. In a networked scenario organisations often need several people (may be at different locations) to input and process data simultaneously. In order to achieve this, a computer-networking model in which a number of separate but interconnected computers do the job has replaced the earlier standalone-computing model. By linking individual computers over a network their productivity has been increased enormously. A most distinguishing characteristic of a general computer network is that data can enter or leave at any point and can be processed at any workstation. In this unit you will be learning about the basic concepts regarding Computer Networks. Here, you will learn about different types of Networks, their applications, and various network topologies used in LAN and WAN. Also, the networking devices and access methods in LAN and WAN will be discussed. At the end of the Unit we will discuss the wireless communication technologies and standards (Wi-Fi and WiMAX) used in wireless communications.

3.2 NETWORK CONCEPT AND CLASSIFICATION

As we have seen in the last unit that a computer network consists of two or more computers that are linked (connected) together to share resources, application, allow communication. In Fig. 3.1 a networked environment is illustrated. The Computers on a network may be linked through cables, telephones lines or wireless points.

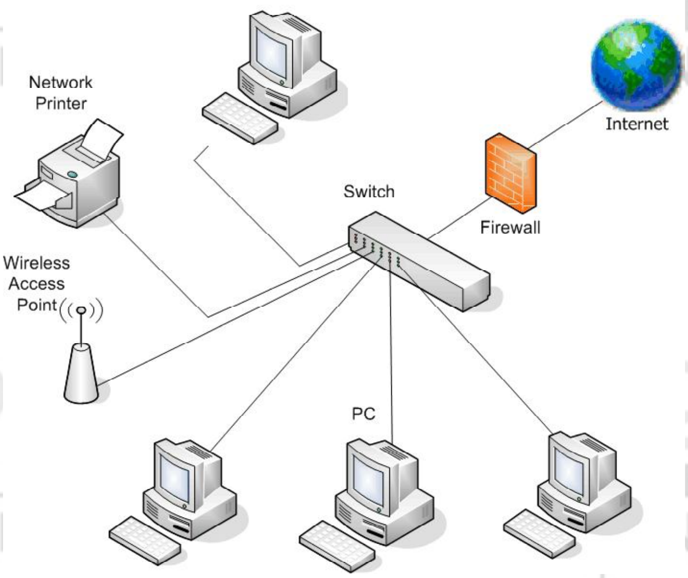


Fig. 3.1: A Simple Computer Network

A Computer network includes the network operating system in the client and server machines, the cables, which connect different computers and all supporting hardware in between such as bridges, routers and switches. In wireless systems, antennas and towers are also part of the network. As you can see in Figure 3.1, between Internet and network a firewall is depicted which provides the security to the network.

3.2.1 Advantages of Networks

Computers in a networked environment provide numerous advantages when compared to computers in a stand-alone environment. The immense benefits that the computer networks provide are in the form of excellent sharing of computational resources, computational load, increased level of reliability, economy and efficient person-to-person communication. Following are some of the major advantages of using computer networks.

Sharing: The main aim of a computer network is to make all programs, equipment, and data available to anyone on the network without regard to the physical location of the resource and the user.

Reliability: Reliability is always an important requirement of all system but especially for military, banking, air traffic control, nuclear reactor safety, and many other applications, the ability to continue operating in the face of hardware problems is of utmost importance. Computer networks provide high reliability by having alternative sources of supply. For example, all files could be replicated on two or three machines, so, if one of them is unavailable (due to hardware failure), the other copies could be used.

Saving Money: Small computers have a much better price/performance ratio than larger ones. Mainframes are roughly a factor of ten faster than personal computers but they cost much more. This imbalance has caused many systems designers to build systems consisting of personal computers, one per user, with data kept on one or more shared file server machines. In this model, the users are called clients, and the whole arrangement is called the client-server model.

Scalability: The ability to increase the system performance gradually as the workload grows just by adding more processors. With centralised mainframes, when a system is full, it must be replaced by a larger one, usually at great expense and even greater disruption to the users. With client-server model, new clients and new servers can be added when needed.

Communication Medium: A computer network can provide a powerful communication medium among widely separated users. Using a computer network it is easy for two or more people who are working on the same project and who live far apart to write a report together. When one worker, makes a change to an online document, others can see the change immediately, instead of waiting several days for a letter. This facilitates real time collaboration among far-flung groups of people which was previously impossible.

Increased Productivity: Networks increase productivity as several people can enter, share, evaluate and process data at the same time cutting down on the time. Networks facilitate handling of multiple tasks simultaneously from different locations increasing productivity and cutting down cost and time.

3.2.2 Network Classification

We see different types of networks available like mobile networks, computer networks and TV networks around us. Let us now learn about different types of computer networks available around us and try to classify them. Table 3.1 shows classification of computer networks based on size, type of connection and functional relationships.

Table 3.1: Classifications of Networks

According to the size of the Network	LAN
	MAN
	WAN
	PAN
According to the type of connection Used	Wired Network
	Wireless Network
According to the functional relationship (Network Architecture)	Peer to Peer network
	Client-Server Network

A) **Size of the Network**

According to size networks are classified as LAN (Local Area Network), MAN (Metropolitan Area Network), WAN (Wide Area Network) and Personal Area Network (PAN).

LAN (Local Area Network)

The network that spans a relatively small area that is, in the single building or campus is known as LAN as depicted in Fig. 3.2. For example, a library will have a wired or wireless LAN for users to interconnect local devices (e.g., printers and servers). Local area networking uses switches, bridges and/or repeaters, and hubs to interconnect LANs and increase overall size.

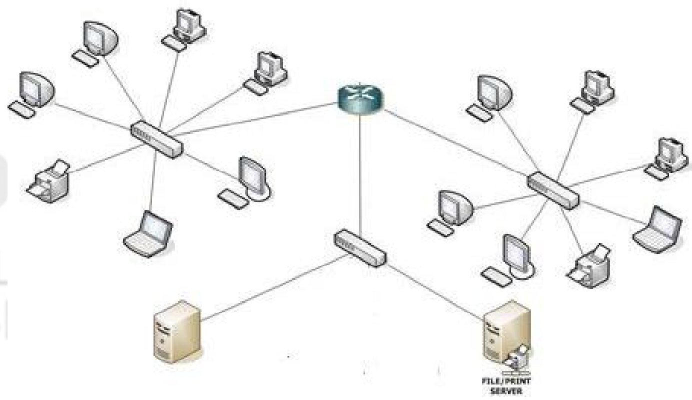


Fig. 3.2: An example of LAN

MAN (Metropolitan Area Network)

Metropolitan Area Network is a Computer network connecting two or more Local Area Networks or Campus Area Networks together but within a town or city as shown in Fig. 3.3. In terms of geographic area MANs are larger than local-area networks (LANs), but smaller than wide-area networks (WANs). MANs are usually characterised by very high-speed connections using fiber optical cable or other digital media. MAN can range anywhere from 5 to 50km in diameter.

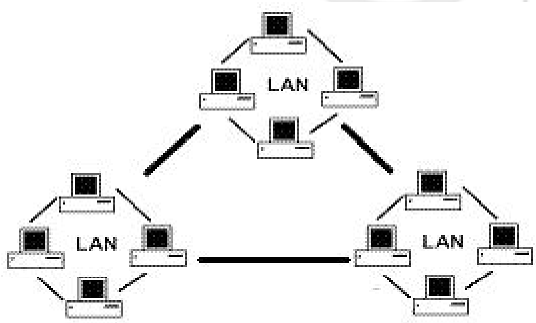


Fig. 3.3: An example of MAN

WAN (Wide Area Network)

Wide Area Network is a computer network that spans a relatively large geographical area. Typically, a WAN consists of many local-area networks (LANs), which are depicted, in Fig. 3.4. They can connect networks across cities, states or even countries. Computers connected to a wide-area network are often connected through public networks, such as the telephone system. They can also be connected through leased lines or satellites. They are generally connected with each other using routers. They have main characteristics like multiple interconnected LANs are used, generally more expensive technology are used for implementation.

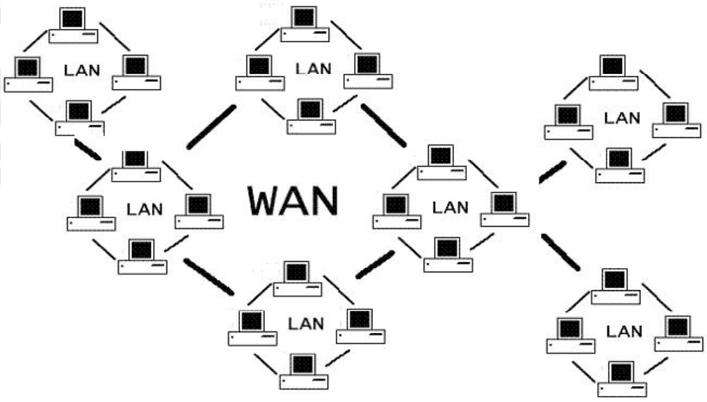


Fig. 3.4: An Example of WAN

Personal Area Network (PAN)

A personal area network (PAN) is a computer network used for communication among computer devices close to one person. Some examples of devices that may be used in a PAN are printers; fax machines, telephones, Personal Digital Assistants (PDAs) or scanners.

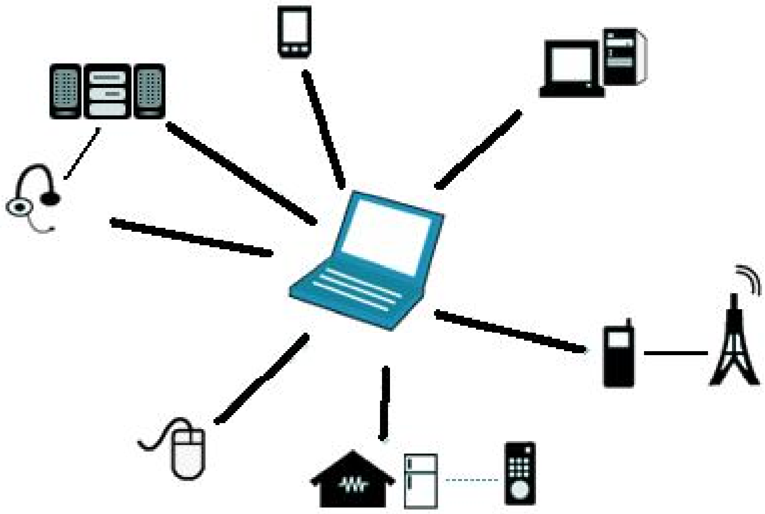


Fig. 3.5: An Example of PAN

The reach of a PAN is typically within about 20-30 feet (approximately 6-9 Meters) as depicted in the Figure 3.5. PANs can be used for communication among the individual devices (intrapersonal communication), or for connecting to a higher-level network and the Internet.

B) **Type of Connection**

According to the connection-type, networks can be classified as using wire connection or wireless connection:

- **Wired Network:** A network that connects devices using cables (wires) like Coaxial Cable, Twisted pair Cable, Optical Fiber Cable etc.

- **Wireless Network:** A network that connects devices using wireless technologies like Bluetooth, infrared, radio frequency etc.

C) Functional Relationship

Functional Relationship or Network architecture could be another important criteria to classify the networks:

Peer to peer network

Peer-to-peer networks are more commonly implemented where less than ten computers are involved and where strict security is not necessary (see Fig. 3.6). Generally, computers in this kind of network are configured to share common resources and responsibilities. Workgroups provide easy sharing of files, printers and other network resources.

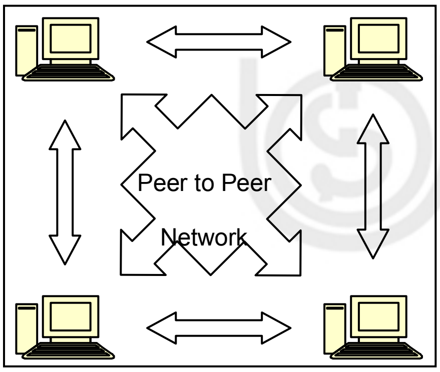


Fig. 3.6: Peer to Peer Networking

It is designed for small LANs in homes, schools, and small businesses. As the number of computers in a workgroup grows, workgroup LANs eventually become too difficult to administer and should be replaced with alternative solutions like client/server approaches.

Client-Server Network

In this architecture each computer or process on the network is either a client or a server. **Client/server networks** are more suitable for larger networks as shown in Figure 3.7. A central computer, or 'server', acts as the storage location for files and applications shared on the network. Usually the server is a higher than average performance computer. Clients are PCs or workstations on which users run applications. Clients rely on servers for resources, such as files, devices, and even processing power.

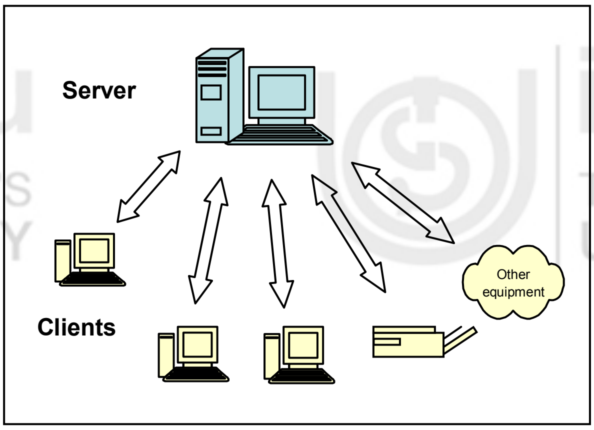


Fig. 3.7: Client - Server Networking

It provides high-level security features through server, which controls the network access of the other computers and processes.

3.3 LOCAL AREA NETWORK (LAN) OVERVIEW

As you know the computer network that spans a relatively small area that is, in the single building or campus is known as LAN, according to standard it can be in 1-kilometer radius. We have studied above the classifications of network based on different criteria. Now in this section we are going to explain about the various LAN topologies and access methods through which network computers can be arranged and configured to communicate.

3.3.1 LAN Topologies

Topology refers to the shape of a network, or the network's layout. How different nodes in a network are connected to each other and how they communicate with each other is determined by the network's topology. Topologies are either physical or logical. Some of the most common network topologies are: Bus, Star, Ring, Tree and Mesh topology. These are explained further in this section.

Bus Topology

In Bus topology, all devices are connected to a central cable, called the bus or backbone. The bus topology connects workstations using a single cable. Each workstation is connected to the next workstation in a point-to-point fashion. All workstations connect to the same cable. Fig. 3.8 shows computers connected using Bus Topology.

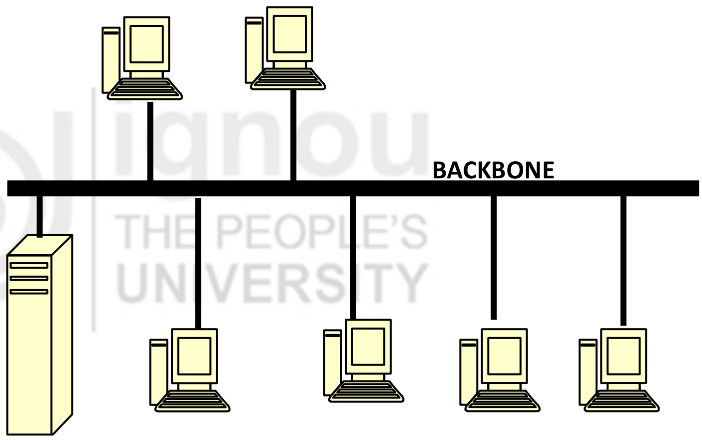


Fig. 3.8: Bus Topology

In this type of topology, if one workstation goes faulty all workstations may be affected as all workstations share the same cable for sending and receiving of information. The cabling cost of bus systems is the least of all the different topologies. Each end of the cable is terminated using a special terminator. The common implementation of this topology is Ethernet. Here, message transmitted by one workstation is heard by all the other workstations. Bus installation is simple, easy and cheap in comparison to other topologies also it requires less cable. But, this topology is used only in comparatively small networks. As all computers share the same bus, the performance of the network deteriorates when we increase the number of computers beyond a certain limit. Also, fault identification is difficult and if it occurs in the cable all transmission can be stopped.

Star topology uses a central hub through which, all components are connected. In a Star topology, the central hub is the host computer, and at the end of each connection is a terminal as shown in Fig. 3.9.

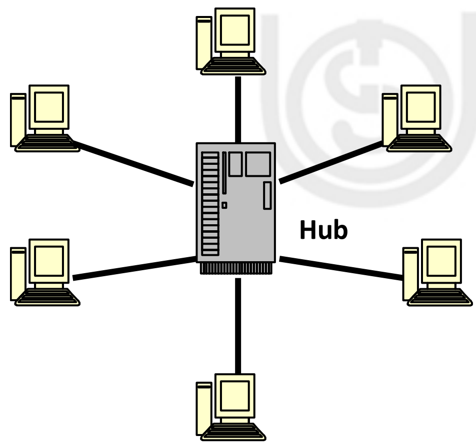


Fig. 3.9: Star Topology

Nodes communicate across the network by passing data through the hub. A star network uses a significant amount of cable as each terminal is wired back to the central hub, even if two terminals are side by side but several hundred meters away from the host. The central hub makes all routing decisions, and all other workstations can be simple.

An advantage of the star topology is that, failure in one of the terminals does not affect any other terminal; however, failure of the central hub affects all terminals. This type of topology is frequently used to connect terminals to a large time-sharing host computer. It is less expensive to mesh topology. Compared to bus topology, faults in this type of network can be easily traced. One of the main advantage of star network is that expansion and modification is easy and it can support multiple cable types like shielded twisted pair cable, unshielded twisted pair cable, ordinary telephone cable etc. However, if central hub fails it brings the entire network to a halt.

Ring Topology

In Ring Topology all devices are connected to one another in the shape of a closed loop, so that each device is connected directly to two other devices, one on either side of it, i.e., the ring topology connects workstations in a closed loop, which is depicted in Figure 3.10. Each terminal is connected to two other terminals (the next and the previous), with the last terminal being connected to the first. Data is transmitted around the ring in one direction only; each station passing on the data to the next station till it reaches its destination.

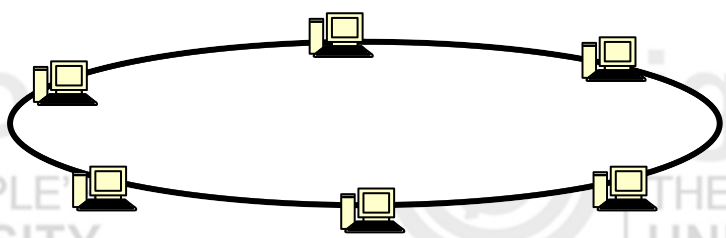


Fig. 3.10: Ring Topology

Information travels around the ring from one workstation to the next. Each packet of data sent on the ring is prefixed by the address of the station to which it is being sent. When a packet of data arrives, the workstation checks to see if the packet address is the same as its own, if it is, it grabs the data in the packet. If the packet does not belong to it, it sends the packet to the next workstation in the ring.

Faulty workstations can be isolated from the ring. When the workstation is powered on, it connects itself to the ring. When power is off, it disconnects itself from the ring and allows the information to bypass the workstation. The common implementation of this topology is token ring. A break in the ring causes the entire network to fail. Individual workstations can be isolated from the ring. In this topology fault isolation is simplified. Unlike Bus topology, there is no signal loss in Ring topology because the tokens are data packets that are re-generated at each node. But Adding or removing computers disrupts the entire network. A break in the ring can stop the transmission in the entire network.

Tree Topology

Tree topology is a LAN topology in which only one route exists between any two nodes on the network. The pattern of connection resembles a tree in which all branches spring from one root. Figure 3.11 shows computers connected using Tree Topology.

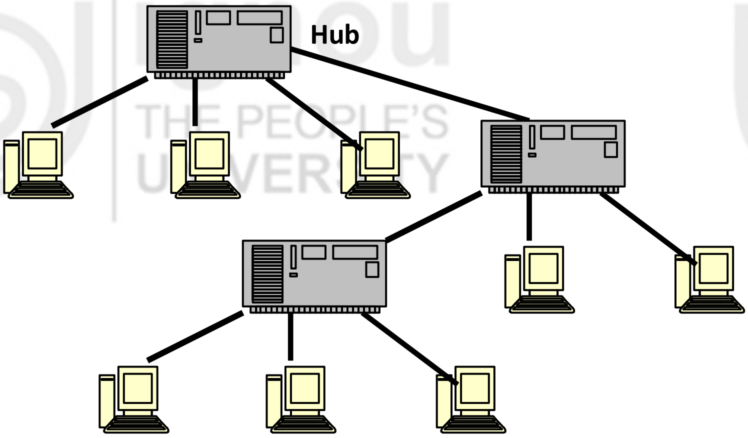


Fig. 3.11: Tree Topology

Tree topology is a hybrid topology, it is similar to the star topology but the nodes are connected to the secondary hub, which in turn is connected to the central hub. In this topology groups of star-configured networks are connected to a linear bus backbone. Similar to star topology Installation and configuration of network is easy and Supports multiple cable types like shielded twisted pair cable, unshielded twisted pair cable, ordinary telephone cable etc. However, failure in the central hub brings the entire network to a halt. Also it needs more cabling when compared to bus topology.

Mesh Topology

Devices are connected with many redundant interconnections between network nodes. In a well-connected topology, every node has a connection to every other node in the network. The cable requirements are high, but there are redundant paths built in. Failure in one of the computers does not cause the network to break down, as they have alternative paths to other computers.

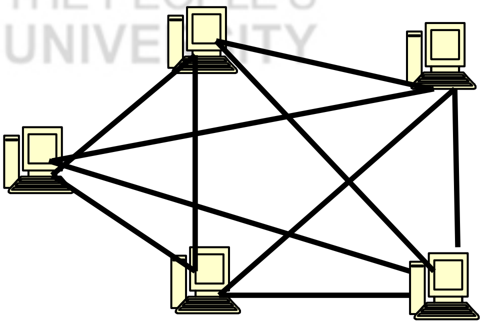


Fig. 3.12: Mesh Topology

Mesh topologies are used in critical connection of host computers (typically telephone exchanges). Alternate paths allow each computer to balance the load to other computer systems in the network by using more than one of the connection paths available. A fully connected mesh network therefore, has n (n-1)/2 physical channels to link n devices. To accommodate these, every device on the network must have (n-1) input/output ports. The main advantage of this topology is its robustness and use of dedicated links eliminates traffic and security problems. Failure in one of the computers does not affect the entire network. However, due to each dedicated cabling it requires high cost, time and I/O (input/output) ports for configuration.

3.3.2 LAN Access Methods

Access Method means how network devices will handle data transmission with one another. For data transmission in LAN, various protocols/standards are defined by IEEE (Institute of Electrical and Electronics Engineers) called IEEE 802, to enable intercommunication between equipment from a variety of manufacturers. It has set IEEE 802.3, 802.4 and 802.5 standards alternatively named as Ethernet, Token bus and Token Ring for LAN. This section will explain you about these methods and their importance in the LAN.

Ethernet

The term Ethernet refers to the family of local-area network (LAN) protocol defined by the IEEE 802.3 standard that is commonly known as the CSMA/CD protocol. Among the other technologies and protocols Ethernet is widely used in LAN-connected PCs and workstations because it is easy to understand, implement, manage, and maintain. Along with that it allows low-cost implementations and provides topological flexibility.

As you know general standard for the data link layer defined by the OSI Model. The IEEE divides this layer into two sub-layers, the logical link control (LLC) layer and the media access control (MAC) layer. IEEE 802.3 supports a LAN standard originally developed by Xerox and later extended by a joint venture between Digital Equipment Corporation, Intel Corporation and Xerox. This was called Ethernet.

IEEE 802.3 defines two types of data transmissions Broadband and Baseband. Broadband transmissions enable two or more communication channels to share the bandwidth of the transmission media. Broadband networks can simultaneously accommodate video, voice and data. Most cable modem providers use broadband communications. On the other hand, Baseband transmissions enable digital signals over a single frequency. With Baseband transmission, the entire communication channel capacity is used to transmit a single date signal. Most LAN's today use Baseband technology.

The MAC layer has two primary responsibilities first is Data encapsulation, including frame assembly before transmission, and frame parsing/error detection during and after reception and second is Media access control, including initiation of frame transmission and recovery from transmission failure. Let's discuss MAC frame structure for IEEE 802.3 with the help of following Figure 3.13.

Preamble	Start Delimiter of frame	Destination Address	Source Address	Length of Data Field	Data	Pad	Frame Check Sum

Fig. 3.13 : Ethernet Frame Format

Each frame has seven fields explained as follows:

Preamble: The first field of 802.3 frame is 7 byte (56 bits) long with a sequence of alternate 1 and 0 i.e., 10101010. This pattern helps the receiver to synchronise and get the beginning of the frame.

Starting Delimiter (SD)/ Start-of-Frame (SOF): The second field start delimiter is 1 byte (8 bit) long. It has pattern 10101011. Again, it is to indicate the beginning of the frame and ensure that the next field will be a destination address. Address, here, can be a single address or a group address.

Destination Address (DA): This field is 6 byte (48 bit) long. It contains the physical address of the receiver.

Source Address (SA): This filed is also 6 byte (48 bit) long. It contains the physical address of the sender. These source and destination addresses are actually MAC address to uniquely identify Each Ethernet network interface card (NIC) that is assigned by the card manufacturer. Each manufacturer that complies with IEEE standards can apply to the IEEE Registration Authority for a range of numbers for its cards. Each MAC address is a 48-bit number, of which the first 24 bits identify the manufacturer. This part of the MAC address (manufacturer ID or organizational unique identifier is assigned by the registration authority. The second half of the address (extension of board ID) is assigned by the manufacturer. The number is usually programmed into the hardware so that it cannot be changed.

Length of Data Field: It is 2 byte (16 bit) long. It indicates the number of bytes in the information field. The longest allowable value can be 1518 bytes.

Data: This field size will be a minimum of 46 bytes long and a maximum of 1500 bytes as will be explained later.

Pad: This field size can be 0 to 46 bytes long. This is required if, the data size is less than 46 bytes as a 802.3 frame must be at least 64 bytes long.

Frame Checksum (FCS): This field is 4 bytes (32 bit) long. It contains information about error detection. This sequence contains a 32-bit cyclic redundancy check (CRC) value, which is created by the sending MAC and is recalculated by the receiving MAC to check for damaged frames. The FCS is generated over the DA, SA, Length/Type, and Data fields. Here it is CRC-32.

Minimum and Maximum Length of Frame

Minimum frame length is 64 bytes and Maximum frame length is 1518 bytes. Minimum length or lower limit for frame length is defined for normal operation of CSMA/CD. This is required so that, the entire frame is not transmitted completely before its first bit has been received by the receiver. If, this happens then the probability of the occurrence of collision will be high (the same has been explained earlier in the previous section CSMA/CD).

Hence, Ethernet frame must be of 64 bytes long. Some of the bytes are header and trailer parts of the frame. If, we consider 6 bytes destination address, 6 bytes source address, 2 bytes length and 4 bytes FCS (6+6+2+4=18) then, the minimum length of data will be 64-18= 46 bytes. If, frame is less than 46 bytes then, padding bits fill up this difference. As per 802.3 standard, the frames maximum length or upper limit of frame is = 1518 bytes (excluding preamble and SD). If we subtract 18 bytes of header and trailer then, the maximum length will be 1500 bytes.

Frame Transmission

Whenever the end station MAC receives a transmit frame request with the associated address and data information from the LLC layer, the MAC starts the transmission sequence by transferring the LLC information into the MAC frame buffer. First, the preamble and start delimiter are inserted in the PRE and SD fields, next both destination and source addresses are inserted into the address fields. Similarly, LLC data bytes are counted, and the number of bytes is inserted into the Length/Type field. Next, the LLC data bytes are inserted into the Data field. If the number of LLC data bytes is less than 46, padding is added to bring the Data field length up to 46. An FCS value is generated over the DA, SA, Length/Type, and Data fields and is appended to the end of the Data field. After the frame is assembled, actual frame transmission will depend on whether the MAC is operating in half-duplex or full-duplex mode. The IEEE 802.3 standard currently requires that all Ethernet MACs support half-duplex operation, in which the MAC can be either transmitting or receiving a frame, but it cannot be doing both simultaneously. Full-duplex operation is an optional MAC capability that allows the MAC to transmit and receive frames simultaneously.

Token Bus

In a token passing system, a small data frame is passed from device to device across the network in pre-determined order. The device that has control of the token frame has the ability to transmit data across the network. When a device has data to send, it must wait until it has the token and then sends its data. When the data transmission is complete, the token is released so that other devices may use the network media. The main advantage of token-passing networks is that they are deterministic. In other words, it is easy to calculate the maximum time that will pass before a device has the opportunity to send data. Both IEEE 802.4 (token bus) and 802.5 (token ring) are based on this approach. However, token bus is suitable in few situations instead of token ring (token ring we will discuss in next section of this unit).

As we know LANs are implemented in a factory automation and process control situation, where the nodes are computers controlling the manufacturing process. In this type of application, real-time processing with minimum delay is needed. Processing must occur at the same speed as the objects moving along the assembly line. Ethernet is not a suitable protocol for this purpose because the number of collisions is not predictable and the delay in sending data from the control center to the computers along the assembly line is not a fixed value. Token ring is also not a suitable protocol because an assembly line resembles a bus topology and not a ring.

Token bus combines features of both Ethernet and token ring. It combines the physical configuration of bus topology and the predictable delay feature of token ring.

Token bus is a physical bus that operates as a logical ring using tokens. Stations are logically organised into a ring. When a device wants to transmit across the bus, it has to determine whether the media is in use. If no other device is transmitting, the signal is sent. Each device receives the signal and then determines whether its address matches that of the recipient. Messages that weren't addressed to the device are disregarded. When dealing with bus networks, it is important to pay careful attention to termination. Each end of the trunk cable needs to be properly terminated. Without termination the signal will bounce back down the cable causing collisions. Generally, bus topologies use coaxial cable.

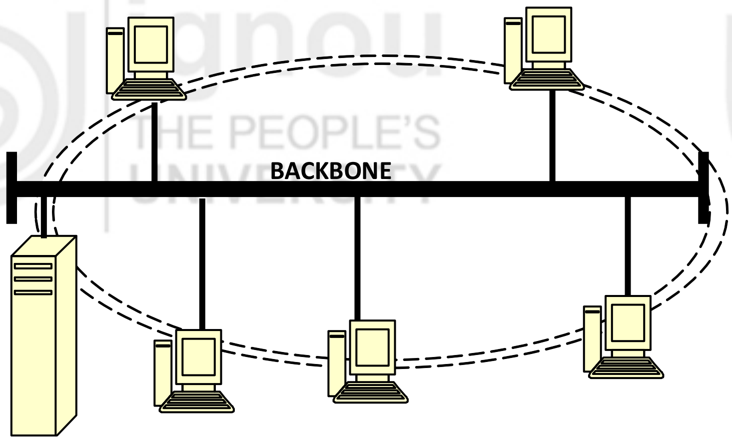


Fig. 3.14: Token Bus

Token bus is limited to factory like automation and process control and has limited commercial application in data communication. It has some other disadvantages like it lead to network performance degradation, weakened signal and difficult troubleshooting.

Token Ring

Token Ring uses a logical ring topology whereby the data is sent from one machine to the next and so on around the ring until it ends up back where it started. It also uses a token passing protocol, which means that a machine can only use the network when it has control of the Token; this ensures that there are no collisions because only one machine can use the network at any given time.

The computers on the LAN are connected so that data is passed around the network in a logical ring (see Figure 3.15). The token ring configuration calls for the computers to be wired to a central hub, where physically it might not look like a ring, but the hub is wired so that the data passes from one computer to the next in a circular motion. The computers pass a packet of data called a token around the network. Only the computer that holds the token can transmit a message on to the ring. At the start, a free Token is circulating on the ring, to use the network, a device first has to capture the free Token and replace the data with its own message. It should write its data and the recipient's address onto the Token and pass it to the next machine in the ring. Next machine that reads the address, realises it is not its own, so passes it on to another machine in the ring sequence. The same process continues till the token reaches to the intended machine.

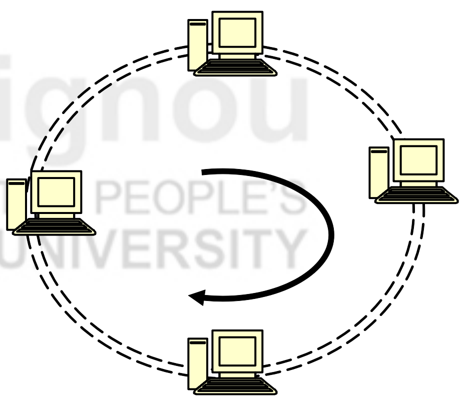


Fig. 3.15: Token Ring

The intended machine read the message and first sends the message back with an acknowledgement to say that it has received the data. This acknowledgement receipt is then sent to machine next in the ring sequence (in the same direction) who checks the address, realises that it is not it's own and so forwards it on to the next machine in the ring, consequently it is forwarded to sender machine. Sender Machine recognises the address, reads the acknowledgement from number 4 (represented by the purple flashing screen) and then releases the free Token back on to the ring ready for the next machine to use. It is basics of Token Ring and it shows how data is sent, received and acknowledged, but Token Ring also has a built in management and recovery system that makes it very fault tolerant. Token ring is technically more sophisticated than Ethernet, and it includes a number of built-in diagnosis and correction mechanisms that can help troubleshoot network problems. Also, because data is transmitted in a more orderly fashion, token ring does not suffer as badly under heavy data traffic. Almost everything about token ring is more expensive than Ethernet by comparison he cable, the network adapter cards, and the other components as well.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) How computer networks provides reliability and scalability? Explain.

2) What is a main limitation of peer-to-peer networking?

3) In real time processing environment Token Bus should be preferred over Token Ring. Justify the statement.

....

....

....

....

....

....

....

....

3.4 WIDE AREA NETWORK (WAN) OVERVIEW

As we know wide area network (WAN) is a network linking geographically distinct locations, which may or may not belong to the same organisation. Over the last few years the web-based applications and wireless networking have changed our outlook about computer networks. Today's corporate networks are accessible virtually anytime from anywhere, and these remote access solutions typically involve a combination of varied WAN services.

In this section we will explain the WAN connection methods, recognize the various WAN physical topologies and hardware devices used for building WANs.

3.4.1 WAN Topologies

WAN topologies use both LAN add enterprise-wide topologies as building blocks, but add more complexity because of the distance they must cover, the larger number of users they serve, and the heavy traffic they often handle. Here we will explain about the WAN topologies. These are almost similar like LAN. However, instead of single node or machine in WAN we connect one network (which may or may not be LAN). These networks may be based on different types of technologies, connections and topologies. Also, the different network may use different types of connections to interconnect with the WAN.

Single link

As Bus topology may not be possible in case of WAN or very costly to implement WAN. We can have another technique similar to peer-to-peer connection to design WAN, where single interconnection points for each location is arranged as depicted in Figure 3.16. Each network site depends on every other site in the network to transmit and receive its traffic. However, the peer-to-peer LANs use computers with shared access to one cable, whereas the WAN peer-to-peer topology uses different locations, each one connected to another one through dedicated channel.

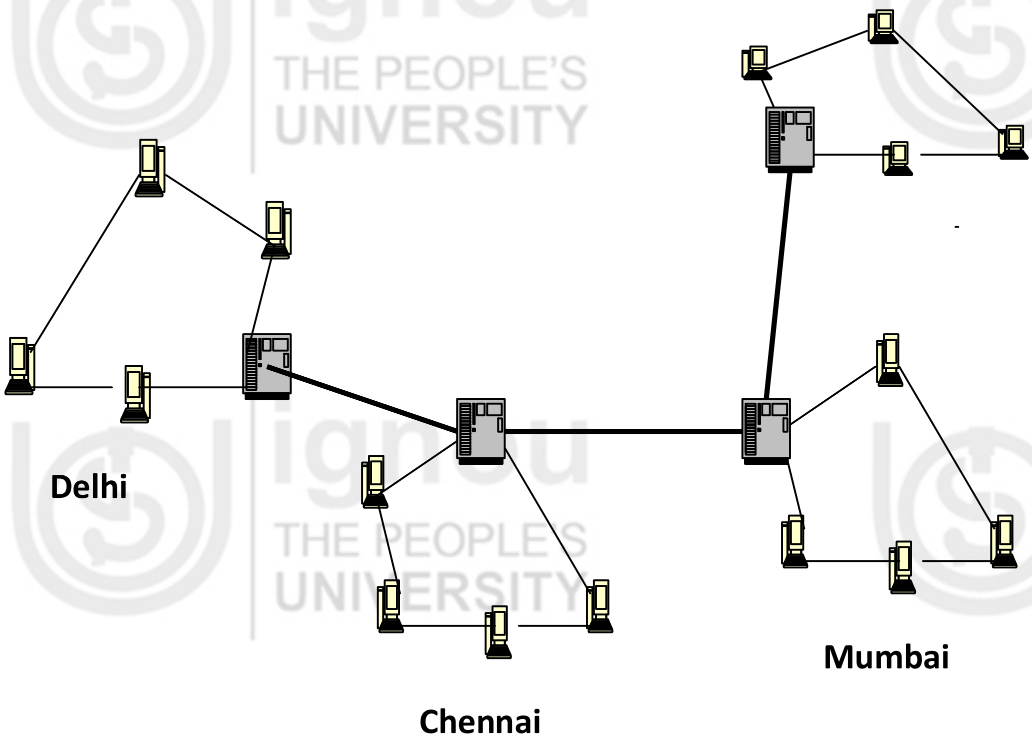


Fig. 3.16: Single Link WAN

Ring

Similar to LAN Ring topology, in WAN Ring topology, each site is connected to two other sites so that the entire WAN forms a ring pattern (see Figure 3.17), except that a ring WAN topology connects locations rather than local nodes. Unlike the Single-Link WAN, routers at any network site can redirect data to another route if one route becomes too busy, however a single cable (as in single link WAN) can break the interconnect of entire network.

Also, the single Link WAN at the time of expanding requires at least one additional link. Therefore, WANs that use the ring topology are only practical for connecting less than four or five network sites.

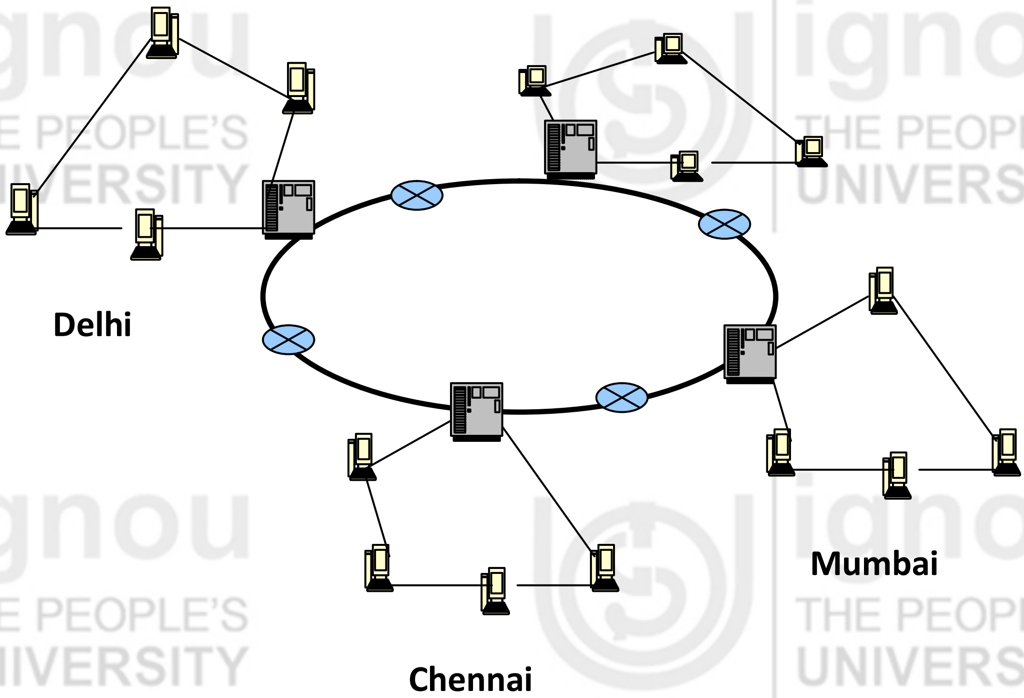


Fig. 3.17: Ring based WAN

Star

The star WAN topology also follows the arrangement of a star LAN. A single network site acts as the central connection point for several other points as shown in figure 18. This arrangement provides separate routes for data between any two sites. As a result, star WANs are more reliable than the Single link or Ring WANs. Another advantage of a star WAN is that when all of its dedicated circuits are functioning, a star WAN provides shorter data paths between any two sites.

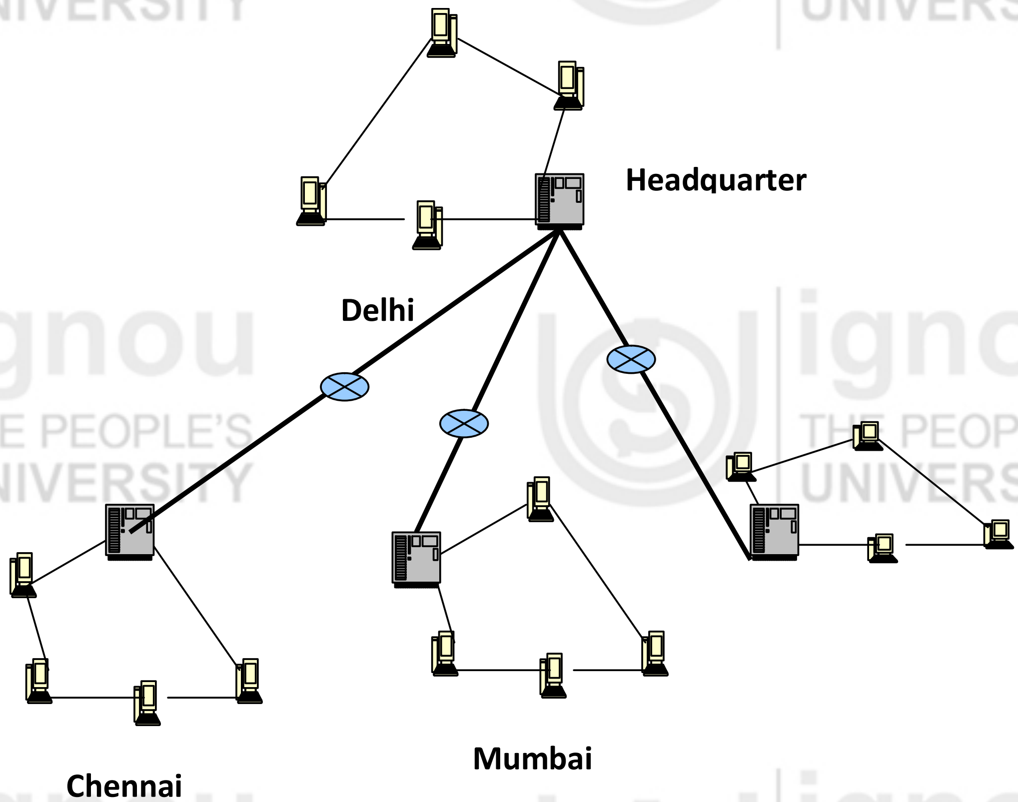


Fig. 3.18: Star WAN

Mesh

As shown in Figure 3.19, Mesh WAN topology incorporates several directly interconnected network sites. Because every site is interconnected, data can be transferred directly from its source to its destination. If one connection suffers a problem, routers can redirect data easily and quickly. One disadvantage to a mesh WAN is the implementation cost; connecting every network site on a network to every other involves leasing a large number of channels. To decrease costs, we can choose to implement a partial mesh, in which critical WAN networks are directly interconnected.

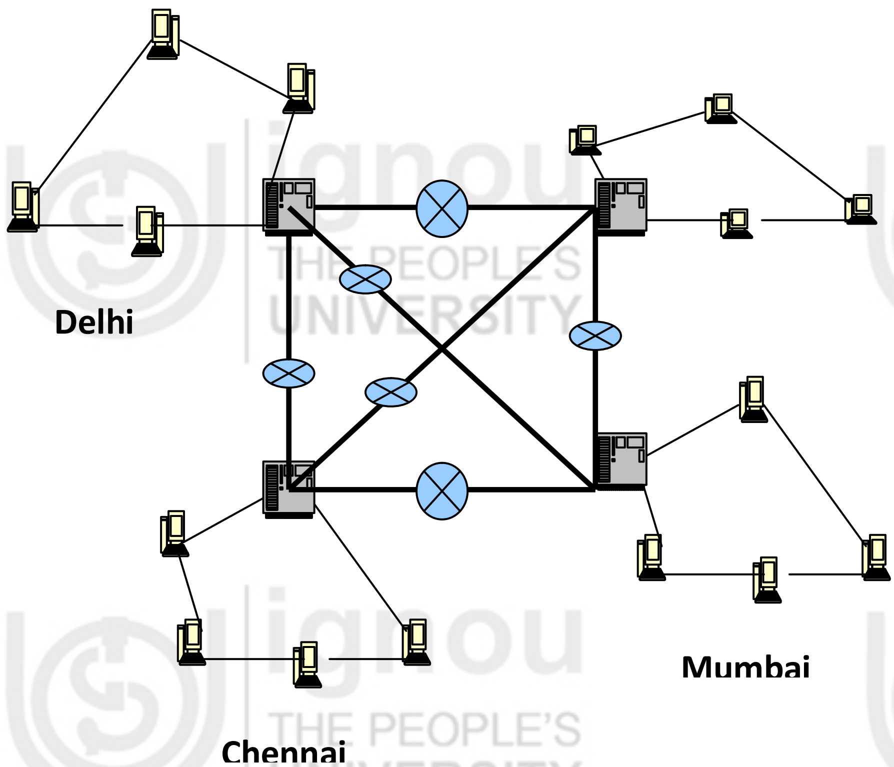


Fig. 3.19: Mesh WAN Topology

3.4.2 WAN Switching Methods

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. There are three switching techniques commonly employed and these are Circuit, Message and packet switching.

Circuit Switching

In circuit switching there is a dedicated communication path between the sending and receiving devices as discussed above in single link WAN. The dedicated path is a connected sequence 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 in. circuit switching involves three steps-circuit establishment; data transfer; and circuit termination. Circuit switching is mainly for voice telephone network, but is not all that effective for data communication networks, as channel capacities are not fully utilised, as data communication equipments do not generate data continuously.

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 network; 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 transmit it to the next node. Examples of a message are electronic mails, 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.

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 are 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 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.

3.4.3 WAN Devices/Hardware

The switching techniques utilise 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. As we have already discussed about LAN devices like bridge and routers in the last unit, in this section briefly we have given the role of these devices in the WAN.

Bridges

Bridges are used to connect two LANs that use identical LAN protocols over a wide area as depicted in the Figure 3.20.

The bridge acts as an address filter, which picks up packets from one LAN that is 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. As all devices use the same protocols, the amount of processing required at the bridge is minimal.

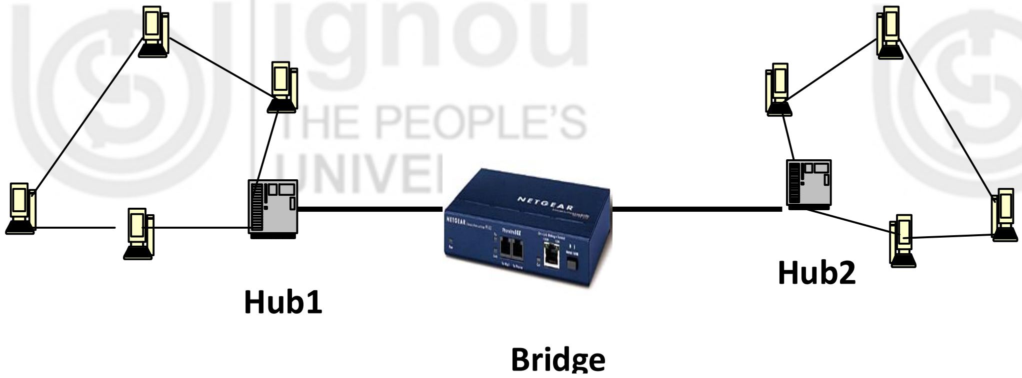


Fig. 3.20: Bridge Connectivity in a Network

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 necessary header and trailer.

Routers

Routers can be used to connect networks that may not be similar. Routers provide connectivity between two LANs or two WANs over large geographical distances as shown in the Figure 3.21. Routers operate at the network layer 3 of the OSI model. All routers 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.

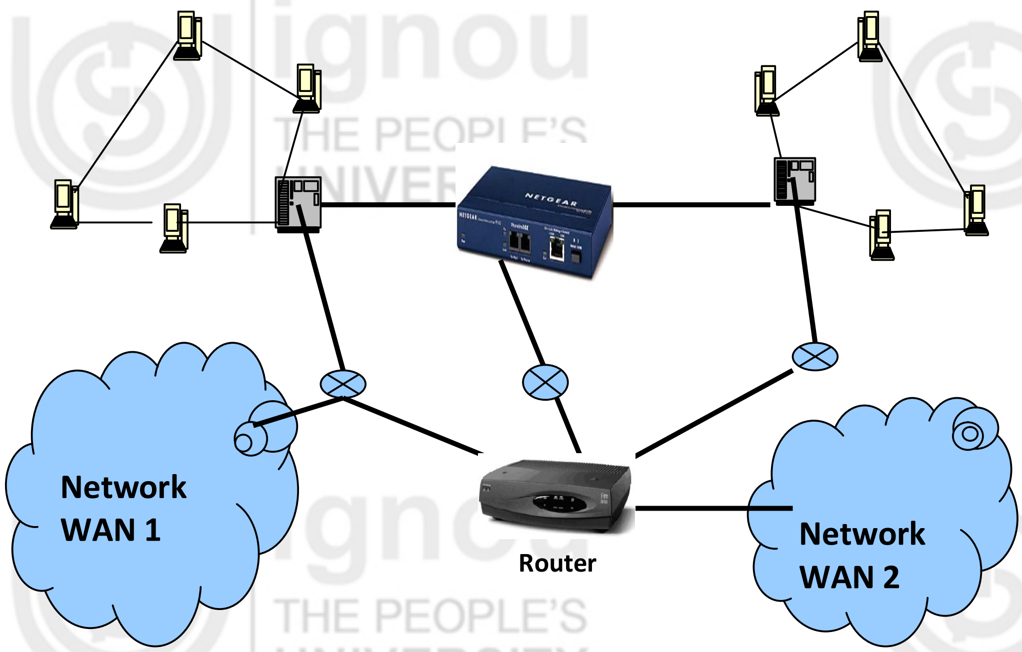


Fig. 3.21: Router Connectivity in a Network

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 as given the following Figure 3.22. The term gateways and 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 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 networks.

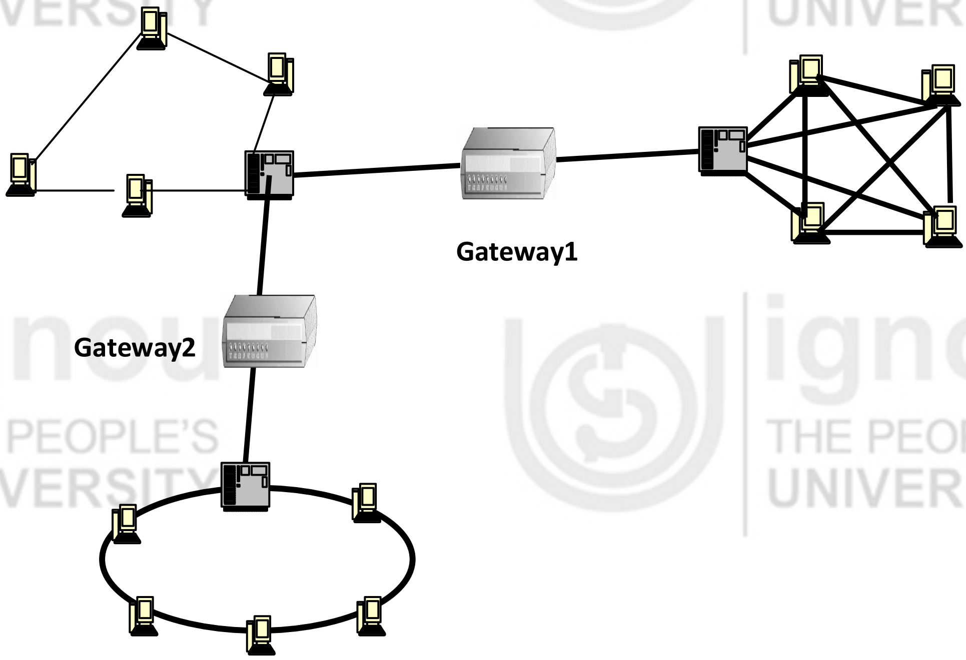


Fig. 3.22: Gateway Connectivity in a Network

3.5 WIRELESS TECHNOLOGY

As we know wireless transmission means where data transmission happen without wires. Here, there are no physical connectors between the two communicating devices. It is highly useful in difficult geographical areas (mountains, jungles, swamps, air, sea, etc.) where building wired networks and fixed devices is very difficult or not possible. In wireless transmission data is sent through the atmosphere or air. The three main types of Wireless Media are Radio, Microwave, and Infrared as explained below:

- The radio portion of the electromagnetic spectrum extends from 10 KHz to 1GHz. Within this range there are numerous bands or ranges of frequencies that are designated for specific purposes.
- Microwave is a radio system, which uses very high frequencies to send and receive data. Because of the high frequencies involved, microwave stations are located about 30 kilometres apart and in line of sight (visible to each other).
- Unguided infrared and millimetre waves are widely used for short-range communication. The remote controls used on televisions, stereos etc. all use infrared communication.

Wireless technologies represent a rapidly emerging area of growth and importance for providing wireless networking. Wireless networks are a system of portable computers that communicate using radio transmission. Nowadays everyone wants an uninterrupted omnipresent network access from everywhere and every time. There is interest in creating mobile computing labs utilizing laptop computers equipped with wireless Ethernet cards. Recently, wireless industry has made important progress in resolving some restraints to the widespread acceptance of wireless technologies. Some of the constraints have included disparate standards, low bandwidth, and high infrastructure and service cost. Wireless technologies can both support the institution mission and provide cost-effective solutions. Wireless is being adopted for many new applications: to connect computers, to allow remote monitoring and data acquisition, to provide access control and security, and to provide a solution for environments where wires may not be the best solution.

Wireless networks are usually configured with the base stations strategically placed around the networking site. Optical fiber cable can be used to cable all base stations together. The transmission power of the base stations and portables must be properly adjusted, so that wireless access is omnipresent to the users. Here unlike the cellular telephone system, each cell has only one channel that covers the entire bandwidth, which is available. Typically its bandwidth is 1 to 2 Mbps.

3.5.1 Wi-Fi

Wi-Fi is a trademark of the Wi-Fi Alliance for certified products based on the IEEE 802.11 standards. It is often suggested that Wi-Fi means Wireless Fidelity (compared with the long-established audio recording term High Fidelity or Hi-Fi.). However, officially the term Wi-Fi does not mean anything.

This certification warrants interoperability between different wireless devices. Actually, when wireless networking technology first entered the market, due to lack of standardization interoperability problem arose between different products from different vendors, consequently IEEE 802.11 standards were developed to guide the vendors. But it is important to note that not every IEEE 802.11 compliant device is certified by the Wi-Fi Alliance, which may be because of certification costs that must be paid for each certified device type. Also, that the lack of the Wi-Fi logo does not imply that a WLAN-device is incompatible to certified Wi-Fi devices.

Wi-Fi Enabled Devices

Nowadays we see that most of the personal computer operating systems, many game consoles, laptops, smart phones, printers, and other peripherals support Wi-Fi. Many consumer devices use Wi-Fi. Amongst others, personal computers can network to each other and connect to the Internet, mobile computers can connect to the Internet from any Wi-Fi hotspot, and digital cameras can transfer images wirelessly.

In addition to restricted use in homes and offices, Wi-Fi can make access publicly available at Wi-Fi hotspots provided either free of charge or to subscribers to various providers. Organizations and businesses such as airports, hotels and restaurants often provide free hotspots to attract or assist clients. Wi-Fi technology is also used to set up mesh networks. Wi-Fi also allows connectivity in peer-to-peer (wireless ad-hoc network) mode, which enables devices to connect directly with each other. This connectivity mode can prove useful in consumer electronics and gaming applications. Wireless routers which incorporate modem and a Wi-Fi access point, often set up in homes and other premises, provide Internet-access and internetworking to all devices connected (wirelessly or by cable) to them.

Advantages and Limitations

In addition to resolving the implementation problems of wired networking at difficult geographical areas, wireless networks typically reduce the costs of network deployment and expansion also. You may be thinking about the wireless networking cost, but gradually it will be more economical. As you can see wireless network adapters are now built into most laptops soon it will be with other networking devices.

Wi-Fi now extensively used in corporate infrastructures. At a basic level of services most of the vendors and brands of access points and client network interfaces has interoperability. Wi-Fi has a global set of standards some time interchanged with name of IEEE 802.11. Any standard Wi-Fi device will work anywhere in the world. One of the major concerns of wireless technology is security; to address the security requirements current version of Wi-Fi Protected Access encryption (WPA2) is considered as secure. It can defeat many security threats, provided strong passwords are used.

Another important limitations of wireless technology are slow data rate and poor quality of services. However, some new protocols for Quality of Service (WMM) make Wi-Fi more suitable for latency-sensitive applications (such as voice and video), and power saving mechanisms (WMM Power Save) improves battery operation.

3.5.2 WiMAX

The meaning of WiMAX is Worldwide Interoperability for Microwave Access, the name "WiMAX" was created by the WiMAX Forum, which was formed to promote conformity and interoperability of the standard.

It is a telecommunications technology that provides wireless transmission of data using a variety of transmission modes, from point-to-multipoint links to portable and fully mobile Internet access. Less data rate is major problem of wireless technology however through WiMAX we can reach up to 3 Megabits/second broadband data rate. The IEEE standard available for WiMAX is 802.16 standard. The WiMAX technology can be used for connecting different Wi-Fi hotspots to the Internet and to enable the wireless broadband access. It can also support all data and telecommunications services through portable connectivity. As it can provide broadband wireless transmission of data using a variety of transmission modes it can be highly useful in the disaster management. It is interesting to note that WiMAX is a possible replacement candidate for cellular phone technologies such as GSM and CDMA, or can be used as an overlay to increase capacity. It has also been considered as a wireless backhaul technology for 2G, 3G, and 4G networks in both developed and poor nations.

IEEE 802.16 (WiMAX)

WiMAX is a term coined to describe standard, interoperable implementations of IEEE 802.16 wireless networks, similar to the way the term Wi-Fi is used for interoperable implementations of the IEEE 802.11 Wireless LAN standard. However, WiMAX is very different from Wi-Fi in the way it works.

Data Link Layer

In Wi-Fi the media access controller (MAC) uses contention access, all subscriber stations that wish to pass data through a wireless access point (WAP) are competing for the WAP's attention on a random interrupt basis. This can cause subscriber stations distant from the WAP to be repeatedly interrupted by closer stations, greatly reducing their throughput. In contrast, the 802.16 MAC uses a scheduling algorithm for which the subscriber station needs to compete only once (for initial entry into the network). After that it is allocated an access slot by the base station. The time slot can enlarge and contract, but remains assigned to the subscriber station, which means that other subscribers cannot use it.

Physical Layer

The original version of the standard on which WiMAX is based (IEEE 802.16) specified a physical layer operations and frequency range. The operating frequency range was updated and modified afterwards in new versions to 2 to 11 GHz range. More higher versions, including 802.16e, also provides multiple antenna support, which are expected to fetch possible benefits in terms of coverage, self installation, power consumption, frequency re-use and bandwidth efficiency.

WiMAX subscriber or client units are available in both indoor and outdoor versions from different manufacturers. Self-install indoor units are suitable, but radio losses mean that the subscriber must be significantly closer to the WiMAX base station than with professionally installed external units. Indoor units are comparable in size to a cable modem. However, outdoor units are roughly the size of a laptop PC, and their installation is comparable to the installation of a residential satellite dish.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) How Ring based WAN is better than single Link WAN?

....

....

....

....

3.6 SUMMARY

In this unit we have learned about networking concept and its utilisation. Here we discussed the different types of networks and the difference between them. We have seen the immense benefits that the computer networks provide in the form of excellent sharing of computational resources, computational load, increased level of reliability, economy and efficient person-to-person communication. Computer networks are basically classified based on size like LAN, MAN, WAN, based on connection like wired and wireless and based on functional requirements like peer-to-peer and client-server. Network topologies means ways of geographical interconnecting in networking nodes like Star, Bus, Ring, Tree, and Mesh topologies, these topologies are discussed for LAN and WAN. The access methods, techniques and networking devices used for developing LAN and WAN are also discussed. In the last of this unit wireless technology is discussed briefly with its two important standards Wi-Fi and WiMAX .

3.7 ANSWERS TO SELF CHECK EXERCISES

1) Computer networks provide high reliability by having alternative sources of supply. For example, all files could be replicated on two or three machines, so, if one of them is unavailable (due to hardware failure), the other copies could be used. Scalability can be achieved by increasing the system performance gradually as the workload grows just by adding more processors. With centralised mainframes, when a system is full, it must be replaced by a larger one, usually at great expense and even greater disruption to the users. With client-server model, new clients and new servers can be added when needed.

2) Peer to peer network is designed for small LANs in homes, schools, and small businesses. As the number of computers in a workgroup grows, workgroup LANs eventually become too difficult to administer and should be replaced with alternative solutions like client/server approaches.

3) Token bus combines features of both Ethernet and token ring. It combines the physical configuration of bus topology and the predictable delay feature of token ring. LANs are implemented in a factory automation and process control situation, where the nodes are computers controlling the manufacturing process. In this type of application, real-time processing with minimum delay is needed. Processing must occur at the same speed as the objects moving along the assembly line. Token ring is not a suitable for this purpose because the number of collisions is not predictable and the delay in sending data from the control center to the computers along the assembly line is not a fixed value. Token ring is also not a suitable protocol because an assembly line resembles a bus topology and not a ring.

4) A ring WAN topology connects locations rather than local nodes. Unlike the Single-Link WAN, routers at any network site can redirect data to another route if one route becomes too busy, however a single cable (as in single link WAN) can break the interconnect of entire network. Also, the single Link WAN at the time of expanding requires at least one additional link.

3.8 KEYWORDS

Ethernet : A system for connecting a number of computer systems to form a local area network, with protocols to control the passing of information and to avoid simultaneous transmission by two or more systems.

MAC Address : A media access control address is a unique identifier assigned to network interfaces for communications on the physical network segment. MAC addresses are used as a network address for most IEEE 802 network technologies, including Ethernet.

Routers : Small physical devices that join multiple networks together. Technically, a router is a Layer 3 gateway device.

WAP : Short for the **W***ireless* **A***pplication* **P**rotocol, a secure specification that allows users to access information instantly via handheld wireless devices such as mobile phones, pagers, two-way radios, smartphones and communicators.

Wi Fi : 1. A facility allowing computers, smartphones, or other devices to connect to the Internet or communicate with one another wirelessly within a particular area.

WiMAX : WiMAX is a wireless communications standard designed to provide 30 to 40 megabit-per-second data rates, with the 2011 update providing up to 1 Gbit/s for fixed stations.

3.9 REFERNCES ANDFURTHER READING

Forouzan, Berhouz. *TCP/IP Protocol Suite.* TATA McGraw Hill. Print

Godbole, Achyut S. *Web Technoligies*. TATA McGrawHill. Print

Stalling, William. *Data and Computer Communication*. PHI: New Delhi. Print

Tenenbaum,Andrew S. *Computer Networks*. PHI: New Delhi. Print

Websites

http://en.wikipedia.org/w/index.php

http://www.angelfire.com/mech/phattony/

http://technet.microsoft.com/en-us/library/

UNIT 4 TECHNOLOGY CONVERGENCE

Structure

4.0 OBJECTIVES

After going through this Unit, you will be able to understand and appreciate:

- what is meant by convergence;
- convergence as a phenomenon;
- the goal of convergence;
- important convergence objectives;
- organisations responsible for convergence standardisation;
- the origin and progress of convergence;
- the factors that drive convergence phenomenon;
- connectivity, capacity and content as the main focus of convergence;
- layered convergence architecture;
- convergence in different layers of the architecture;

- technology and techniques convergence;
- network, access and services convergence; and
- convergent applications like internet radio and interactive TV.

4.1 INTRODUCTION

As students of BLIS, you must be aware that the world has witnessed the phenomenon of information explosion in the last century. The advent of industrial age accompanied by an increase in world population has brought about significant growth in information generation and dissemination. All round developments in areas like transportation, power, communications, entertainment and education have led to exponential growth in information. Today, information generated doubles every five years. By the middle of last century, computers were pressed into service to store and manage large volumes of information in electronic form. The electronic storage media is capable of storing large volumes of data while occupying small physical space. When a complete document is stored in electronic form, we call it as **electronic document**. We use digital technology to store information electronically. Hence, an electronic document is also known as **digital document**. Digital technology is very rugged and reliable and uses only two signal values to represent information. Numerically, we represent the two signal values by binary digits (bits) '1' and '0'.

The world is evolving towards a Networked Electronic Information Society (NEIS). As you know, a networked society means one in which a large proportion of the world population is interconnected or networked by some form of telecommunication system and the people carry out their day-to-day activities using the network predominantly. Day-to-day activities may involve tasks such as banking, ticket booking for travel or entertainment programmes, product ordering, financial transactions, exchange of mails, retrieving of information from a database, downloading of music files, simple telephone conversation etc. Electronic information is central to all these tasks. Activities that are carried out in the electronic domain using networks are usually denoted with a prefix 'e-' such as e-banking. We now list some of the e-activities that are evolving in the networked society:

- e-education
- e-governance
- e-library
- e-health
- e-banking
- e-commerce
- e-mail
- e-entertainment
- e-journal
- e-newspaper
- e-marketing
- e-procurement

E-activities are often referred to as online-activities. Examples include online-ticketing, online-banking etc. To illustrate what e-activities mean, we elaborate two of the above activities. First, we illustrate e-education. This means sign up with a university, pay fees, obtain lessons, submit assignments give examination and obtain results, all via network without having to visit the university actually. Counselling is also available via network. A student may book time with a counsellor and interact with him/her via Internet. In India, IGNOU is offering e-education in many areas. Second, we look at an e-library. The resources and services of an e-library can be accessed on the network without having to physically go to the library. For this purpose, it is necessary to store all the library resources in digital form. A library whose resources are available in digital form is known as **digital library**. When accessed over networks from remote places, a digital library becomes an **e-library**. What is underlying the NEIS or the so-called e-world? It is the concept of convergence.

4.2 WHAT IS CONVERGENCE?

Last few decades have witnessed significant development in the field of telecommunications, computer networking, information technology and Internet. Optical fibre, wireless and satellite communications have seen many advances. Communication satellites and remote sensing satellites are used extensively today. Mobile communication has become a common place with close to half of the world population connected by mobile networks. Internet has grown leaps and bounds from a few interconnected computers in 1980 to over a billion computers in 2009. This is often termed as Internet Revolution. Usefulness of these developments to the society increases significantly when there is a synergy of purpose amongst different technologies and techniques. In other words, the use of technologies, the design of networks and the development of services must all be directed towards a common goal. Such a common goal is NEIS. The process of bringing together different technologies, techniques, networks and services to realise the common objective of evolving towards NEIS is known as **convergence**. E-activities are central to NEIS. And, convergence is a fundamental requirement to make e-activities possible in a networked society. Without convergence no networking or e-applications are possible. It is because of convergence concept the networked society is becoming a reality. We study the goal and objectives of convergence in detail in Section 4.3.

In fact, convergence is not limited to NEIS. It is much wider phenomenon covering many technologies and even human resource. For example, convergence is much talked about in automobile industry. There it refers to the use of electronics, communications and computers in automotive controls and operations. For example, use of radar to sense vehicles or other obstacles ahead on the road and automatically adjusting the speed is possible only when the relevant technologies converge. Another interesting example is the use of intelligence in air conditioners. Conventional air conditioners maintain the room ambience according to certain set parameters. Air conditioners with intelligence determine the ideal room ambience and maintain the same. They take into account factors like human biological rhythm, the present heat load, atmospheric temperature etc. in determining the ideal ambience. Artificial intelligence, biological science, instrumentation and computer engineering are some of the disciplines that contribute to producing such an air conditioner. There is little doubt that the world is witnessing a phenomenon of convergence in almost all the fields.

Convergence was born when the computer and communication engineers realised the tremendous potential of bringing the two technologies together. From then on, every development relating to technology, switching techniques, protocols, services and applications has kept in mind the convergence aspect. We trace the genesis and development of convergence in Section 4.4. Connectivity, inter-operability, content access and capacity of communication and computer systems lie at focus of convergence. We discuss this in Section 4.5. The fact that the convergence principle is applicable from basic communication systems that move bits and bytes to sophisticated network applications has led to the study of convergence architecture. We cover this in Section 4.6. The individual components of the architecture are discussed in Sections 4.7 – 4.10. Convergence applications are becoming commonplace these days. These days, we hear a lot about access to Internet from mobiles in third generation (3G) mobile communication systems. This is an excellent example of convergence between mobile communications and Internet. With this, we would have a host of new applications like mobile banking, m-commerce etc. Literally, every application that is emerging today is based on the convergence concept. We discuss convergent applications in Section 4.11.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Define convergence.

....

....

....

....

4.3 GOAL AND OBJECTIVES OF CONVERGENCE

As discussed earlier, convergence is an essential fundamental requirement for NEIS to emerge. The degree of success of convergence phenomenon would determine the speed with which mankind would move towards NEIS. In the context of NEIS, the goal of convergence may be stated as:

To achieve the capability wherein any network can deliver any service to any platform, and any user with an appropriate consumer device can access any application that runs on the networked world.

In some sense, the goal as stated above is reflection of the famous idiom "Unity in Diversity" in the context of NEIS. Multiplicity of networks and computer platforms make up the basic infrastructure of the networked world. For example, we have landline and radio networks. We have PCs, servers and mainframe computers. These networks and platforms are designed by using a variety of different technologies. For example, two different technologies, viz. GSM and CDMA are used for designing mobile networks. GSM stands for Global System for Mobile communications and CDMA for Code Division Multiple Access. Landline networks may use copper cables or optical fibres. There would be a multiplicity of applications and a multiplicity of consumer devices. In the midst of all these multiplicity, there exists the unity of purpose that any user connected to any network is given the capability for accessing any application that runs on any platform anywhere in the networked world. Of course, the user must have the appropriate consumer device that is required by an application. Obviously, a user cannot expect to access an application that delivers pictures by using a plain telephone set. Integrated consumer devices are capable of transmitting and receiving audio, video and data. Such devices enable the users to access multimedia applications on the global network. We are witnessing the concept of integrated consumer device being applied in mobile handsets extensively. As you may be aware, we have mobile handsets that have built-in camera, FM radio, MP3 audio player, Internet access feature download facility etc.

There are only three generic forms of information in NEIS: Audio, Video and Data. Audio includes speech and music. Video includes television, movies pictures, photos etc. Data includes all computer-generated information such as text, computer graphics, computer animation, digitised documents etc. Every application in NEIS is built around these three fundamental forms of information.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

2) What is an integrated consumer device?

3) What are the fundamental constituent components of information?

....

....

....

....

4.4 GENESIS OF CONVERGENCE

The genesis of convergence can be traced to mid 1950s. At that time, the computers were expensive resources and required highly controlled environment for their operations. They were housed in big computer centres that were air-conditioned with temperature, humidity and dust control. Entry was restricted to these centres and people were required to remove their shoes and wear aprons to work inside the centres. Computer centres were few in numbers and often located in faraway places. Users had to travel long distances to reach the centres. India had only three or four computers in the first half of 1960s located in the major cities like Kanpur, Delhi, Kolkata (the then Calcutta) and Mumbai (the then Bombay). In effect, working with computers was difficult for prospective users.

Around the same time, computer manufacturers were keen on increasing the clientele for their computers for economic reasons. Being expensive resources, high usage was essential to justify the cost associated with the establishment of computer centres. On the one hand, the prospective users could not get easy access to computers and on the other, computer resource was being under utilised. A solution was needed to resolve this issue for the benefit of both users and the owners of the computer systems.

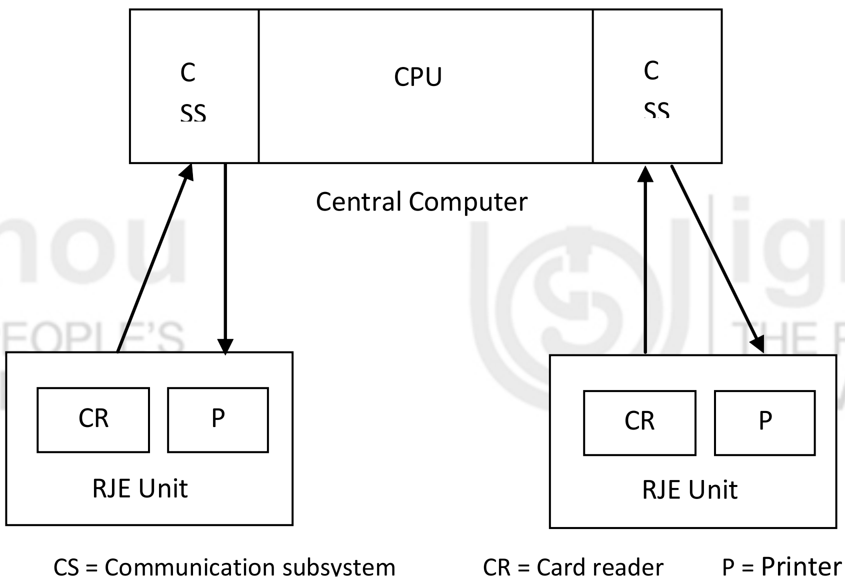


Fig. 4.1: Remote Computing Set-up

The solution was found in the famous saying that goes as: If Prophet Mohammad cannot go to the mountain, the mountain comes to him. Computer manufacturers took an approach reflected by this saying in solving the problem of low usage of computers. They decided to take the access to computers to the doorsteps of users. The approach is known as **remote computing**, i.e. the users could access a computer and perform computations from a remote location. Implementation of remote computing required a suitable communication link between the remote site and the computer centre as shown in Fig. 4.1.

In those days, computer programs were used to be prepared on cards by punching holes and submitted to a computer centre for execution. A program thus submitted was called a job. Accordingly, in remote computing, the equipment at the remote site was called Remote Job Entry (RJE) terminal. This usually consisted of a card reader as an input device and a printer as output device suitably interlinked to the main central computer via dedicated communication lines. Many RJE terminals could be connected to one central computer thus allowing users from different locations to access the same central computer. This solution turned out to be win-win solution for both users and the computer owners. Users could access the computer easily and at the same time the utilisation of the computer went up significantly. Remote computing is the first application that brought computer and communication technologies together laying the foundation for technology convergence. Thus, the genesis of convergence lies in remote computing that started in late 1950s.

RJE centres were non-interactive online centres. They were permanently connected to the main computer centre via communication lines; i.e. online. They were non-interactive because users could not access the computer online. They submitted their jobs on cards and their outputs became available many hours later. Thus, RJE centres were offline as far as the users were concerned. However, RJE centres demonstrated the successful remote online connectivity.

Once remote computing was successful, suitable software systems were developed to allow user interact directly with the computer giving birth to what is known as **interactive computing.** In interactive computing, the users interact directly with the main computer. I/O is done via visual display units (VDU) or **monitors,** as they are popularly known, with keyboard acting as the input device and the screen as the output device, a configuration that we are very familiar with these days. The VDUs, i.e. the monitors are connected online to the computer. It was just a matter of time that a cluster of VDUs was placed in RJE centres, laying the foundation for network computing.

The idea of sharing data, information and other resources emerged in the late 1960s. Focus shifted from communication between terminal and computer to communication between computers, which marked the emergence of **computer networks**. In fact, the ideas developed in remote computing form the basis of access to Internet even today. The convergence of computer and communication technologies took deep roots by the end of 1960s. This was most aptly summarised in the words of a famous information scientist R. M. Fano as quoted in the following:

The marriage of computers and communications has been celebrated and consummated. The honeymoon is over and the two partners are realising the hard realities of inter-dependence on each other. - R.M. Fano, 1972.

After 38 years of this marriage, the bond is growing from strength to strength and today we are witnessing the phenomenon of convergence in a variety of areas to realise the common goal of networked society.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) Why is remote computing a win-win solution for users and computer owners?

....

....

....

....

4.5 CONVERGENCE FOCUS

Quite often, a question arises as to what should be the focus of convergence efforts? There are three important aspects towards which convergence efforts in NEIS may be directed:

- Connectivity
- Capacity
- Content

At present, less than one-sixth of the world population is connected to telecommunication infrastructure for some service or the other. Over five billion people in the world still have no network connectivity. The existing connection may be for a simple service like telephone, dial-up Internet access or a multitude of services offered by ISDN. Should the convergence efforts be directed towards increasing the connectivity level and thereby make a larger percentage of the world population as part of NEIS? Bringing in more people on the network has been considered important by major international organisations. In fact, International Telecommunication Union (ITU) has suggested that every village in the world be brought on to the connectivity map by the end of year 2015 as part of its global multi-stakeholder initiative called Connect the World launched in June 2005.

As mentioned earlier, interactive services are likely to be dominating NEIS. Interactive services, restricted to text transfer like e-mail, information access and simple graphics transfer can be supported by limited bandwidth connections that exist today. The current information transfer rates are in the range of 64-128 kbps. The rates are proving to be inadequate even for the current level of usage. For example, Internet response for Web access at times becomes so slow that it made someone to expand the acronym WWW as World Wide Wait! Multimedia interactive services envisaged in NEIS demand transmission and distribution of high fidelity voice, high quality video, 3-D graphics and other forms of information. This calls for a broadband communication pipe supporting data rates of the order of 100 Mbps or more to be extended to customer premises. Broadband services have the potential of increasing the revenue earnings for the network operators as the users may be prepared to pay high charges for the services. Should then the convergence efforts be directed towards enhancing the capacity of the existing communication pipes instead of increasing the connectivity?

A network infrastructure without contents and applications to run on it is like 5-star hotels without guests to occupy them. The importance of content creation has been emphasised by many experts. For example, a CEO of a networking company once remarked something to the following effect "We do not want to build dumb pipes. If we make only racks and servers, that is dumb. What we should do is to meld contents and networks" Norio Ohga, an ex-chairman of Sony once remarked "Without content, network is nothing". Yet the debate of content versus connectivity is hot. One of the main reasons for this is that the statistics show that consistently connectivity applications earn much more revenue than content delivery systems. In connectivity applications like telephone, the end users create the content. The highest revenue is earned from voice conversation in both landline and mobile networks. The other source of major revenue is the fax usage, which is again a connection-oriented service. Another pointer is a comparison between Small Messaging System (SMS) and Wireless Access Protocol (WAP) on the mobile networks. SMS is connectivity oriented where the end user creates the content. WAP is a content delivery protocol for wireless devices. Statistics show that SMS is more widely used than WAP and brings in more revenue. On the Internet too, it is the e-mail that is the 'killer' application and not the wide area information system (WAIS). In fact, it appears that even historically the connectivity has been more important for people. For example, in the postal system the main revenue earner is the letter communication rather than the newspaper distribution. It appears that the people prefer to create their own content and all that they need is an efficient connectivity infrastructure that is affordable.

So, the debate goes on. In the opinion of this author, all the three aspects, viz. connectivity, capacity and content are important. Small and medium network operators must concentrate on connectivity. Large operators must focus on capacity. Content creation must be left to end users and specialists like movie producers.

4.6 CONVERGENCE ARCHITECTURE

Applications Convergence
Services Convergence
Access Convergence
Networks Convergence
Protocol Convergence
Technology Convergence

Fig. 4.2: Convergence Architecture

The phenomenon of convergence in the context of networked society is occurring at many levels. Different levels of convergence may be represented in the form of a layered architecture as shown in Fig. 4.2. At a very broad level, convergence is seen at sectoral level of infrastructure such as telecommunication sector, power sector, automobile sector, air conditioning sector etc. At the second level is the convergence in technology. Technology convergence is studied at two levels. First, how different technologies in the same sector exhibit synergy to achieve certain common goals? Second, how technologies in different sectors come together to produce better products? Our main concern is the first question in Information Communication Technology (ICT) sector to achieve the goals of NEIS. Technology convergence is fundamental to the entire process of convergence in NEIS. Without technology convergence, other layers of convergence are not possible. Once the technologies are able to interwork, techniques that can be applied across all technologies need to be evolved. In particular, switching and signalling techniques need to converge in ICT. Techniques convergence layer deals with this aspect. Once converging techniques are available, the different networks and systems can be made to interwork with each other. Then comes the need for access convergence. From any one interface, the user must be able to access all the services available in NEIS. In the architecture shown in Fig. 4.2, a distinction is made between services and applications. Services are provided by the network operators and the applications are designed and run by end users. Applications make use of the network services.

4.7 TECHNOLOGY CONVERGENCE

There are four major communication technologies that form part of modern telecommunications:

1) Electrical communication system
2) Optical communication system
3) Radio or wireless communication system
4) Satellite communication system.

Now, let us see how each one of these technologies converges to support the emergence of NEIS. The four communication technologies complement each other to improve **connectivity** and **capacity**.

Electrical communication uses copper cables of different types that have different information carrying capacities. Copper cables have been in use for a long time for establishing networks worldwide. Information carrying capacity of copper cables is being continuously improved and even today copper cable is a favourite medium in communication systems. Optical fibres are replacing the copper wires in certain segments of telecommunications because of their extremely high information carrying capacity. However, laying copper or optical cables to every village, home and office in the world is a formidable task and it may take many centuries if this were to be achieved. The enormity of the task becomes clear when we realise that we have been able to connect via landline only about 16% of the world population in the last 130 years. Further, laying optical fibres is more difficult than laying copper cables because of special considerations required for optical fibres in bending and routing. With such constraints, large-scale connectivity cannot be achieved in a short time frame by using cables. However, only high speed copper cables and optical fibres can offer high bandwidths. The bandwidth capacity of optical fibres is, of course, orders of magnitude higher than that of copper cables.

Our interest in wireless communication lies in short-haul radio technology that may be placed under four categories: mobile communication, wireless LANs, piconet, and wireless in local loop (WLL or WILL). Since the introduction of cellular mobile telephony in 1990, over one billion mobile subscriptions have been registered throughout the world as of 2004. There is a discernible trend that indicates that new users prefer mobile connectivity to landline connectivity. Today over 90% of the countries in the world have mobile networks and over 100 countries have more mobile subscriptions than fixed landline subscriptions. Clearly, mobile communication is raising connectivity levels significantly. Data and video carrying capacity of mobile systems are very limited.

Wireless LANs are useful in such cases of office and home networks where wired infrastructure may not be available. They can also provide high capacity. Their main problem is in electromagnetic interference and lack of privacy and security. Wireless LANs may be affected by or cause interference from/to neighbouring networks. Eavesdropping is very easy on wireless LANs.

WLL short-haul radio technology has the excellent potential of making interactive services a reality. Residential premises may be connected to nearby communication centres via WLL radio technology. The centres may be equipped to offer large bandwidth using optical fibres. This principle may be extended to provide limited interactive facility to individual users via mobile communication technology. However, WLL is capable of providing much higher capacity than mobile technology. The use of short haul radio implies the establishment of radio network, which can be done more easily in urban areas than in rural areas.

In the context of remote area connectivity, satellite communication presents a practical solution for connecting remote areas. Establishing connectivity via satellite links is by far the quickest. Even remotest villages can be brought on to the world network map in a matter of few hours by using satellite links. Local cabling may be used to carry information between the satellite terminal and the houses in the village. Satellite links do not need any landline infrastructure except for local distribution. At present, efforts are on to develop satellite mobile systems wherein a hand-held device may connect to a satellite directly without the help of a terminal. This also offers an elegant solution to rural area connectivity even though the capacity of the hand-held terminals may be very limited. As mentioned earlier, the main problem with the satellite is the limited capacity. The bandwidth of a typical present day communication satellite is at least a few hundred times lower than that offered by a single optical fibre. Limited bandwidth of satellites does not permit high-capacity two-way interactive communication to individual homes.

Trained manpower is an important consideration in selecting technologies. Copper cable communication is well known and a large number of trained technicians are available to install and maintain such systems. This is the reason why a large number of cable operators have sprung up in the market. Optical fibres are new and satellite communication is highly specialised. Manpower is limited in both areas. Radio communication is well understood but mobile communication involves new concepts that are not taught extensively at present.

Table 4.1 summarises the strengths and limitations of different technologies that we have discussed so far. The convergence of the four communication technologies, inter-operating with each other, can provide a suitable solution to the problems of connectivity, capacity and trained manpower in the networked information society of tomorrow.

Table 4.1: Technology Convergence

Technology	Connectivity	Capacity	Manpower
Copper cables	Fair	Reasonable	Very Good
Optical fibres	Difficult	Very High	Limited
Satellite	Rural areas	Limited	Limited
Short haul radio	Urban areas	Limited	Low

We now have a paradoxical situation. Short-haul radio and satellite communication enhance connectivity but have limited bandwidth capabilities. High speed copper cables and optical fibres can offer very high bandwidth but have serious connectivity limitations. Hopefully, these technologies are poised to play complementary roles in shaping the networked information society of tomorrow. Clearly, the goal of telecommunications in the years to come is to establish a wide band multi-services network reaching out as large a population of the world as possible.

4.7.1 Bluetooth Technology

Bluetooth is the codename for a short-range wireless convergent technology. It is a low power radio technology covering a small range of distances up to 10 metres. The purpose of the technology is to make bluetooth enabled devices that are in the vicinity of a master device to communicate in a wireless mode. The technology is very simple to use. All that one needs to do is to bring a bluetooth device close to bluetooth-enabled computer and a communication can start. There is no cable or modem or driver to be installed. The simplicity of its use is very attractive to users. For example, a mobile phone may connect to a laptop computer and send and receive electronic mail if the two devices have bluetooth interface incorporated in them. Bluetooth communication calls for no additional infrastructure like mobile communication that calls for a radio network to be in place. Bluetooth devices can communicate directly without any network support.

Bluetooth functions in master-slave configuration. Usually a bluetooth-enabled computer acts as a master. Slave devices are dumb, basically doing whatever the master tells them to do. Typical slave devices include digital cameras, mobile handsets, scanners and personal digital assistants (PDA). Up to 7 slave devices can be active at a time. If more than seven slave devices come in the vicinity, the master computer can place these devices in 'parked' mode. There can be up to 255 parked nodes in the net. The master can wake up a parked node and allow it to communicate with it. All communication is between the master and the slave. Direct slave-to-slave communication is not possible. The technical name for the Bluetooth network is **piconet**. Multiple piconets can be interconnected to form what is called a **scatternet**.

Bluetooth has potential for many interesting applications. These include:

- Receiving and sending email from mobiles
- Check in at airports
- Check in at movie theatres
- Sending and receiving fax from PDAs and mobiles
- Digital walkie-talkie
- Make e-payments
- Transferring picture files from digital cameras to computers
- Connecting a scanner to a computer without a cable.

In general, wherever there is a queue-based activity, bluetooth technology can be used. In other words, bluetooth technology has the potential of eliminating queues from this world!

4.7.2 3G and WiMAX Technologies

Mobile communication has undergone three generations of development. The first generation (1G) mobile communication was basically analog voice communication. The second (2G) generation systems are designed for digital voice communication and low speed data transmission. In India, we have mainly 2G mobile systems prevalent and 3G is picking up fast. There are two technologies and standards associated with 2G systems: GSM and CDMA. GSM stands for Global System for Mobile communications. CDMA stands for Code Division Multiple Access systems. Both technologies operate in India. Low speed data feature of 2G systems supports SMS (Small Messaging System) service. Since higher data rate can support services like MMS (Multimedia Messaging System), the 2G technology was enhanced using two techniques: General Packet Radio Service (GPRS) and Enhanced Data rates for GSM Evolution (EDGE). The resulting systems were called 2.5G mobile systems. Much higher-level services can be offered in 3G systems. These include video telephony, Internet browsing, mobile television and high quality music. The main difference from one generation to the next is the amount of bandwidth available for services. Higher the bandwidth, the larger is the variety of services that can be offered. 3G systems offer voice, video and high-speed data services taking convergence one step further.

4.8 PROTOCOL CONVERGENCE

Wi Max is the wideband wireless data access technology. This technology permits mobile operator to offer very high-speed data services to users. The data rates may vary from 2 Mbps to 100 Mbps.

Switching, signalling and transmission techniques are the main ones in networks. Signalling includes addressing and routing apart from monitoring sessions and connections. In this section, we shall bring out the divergence that exists in these areas and the convergence efforts that are currently being pursued.

As you are aware, the principle of switching is central to all networks. The switching technique adopted by data networks is different from the one used by telephone networks. Telephone networks use the technique of **circuit switching** whereas data networks use **packet switching**. In circuit switching, a communication path is established between the source and destination before the actual information transfer takes place. This path and the associated resources remain dedicated to the communicating pair until any one of the two entities initiates a disconnect request. The path is used effectively if there is continuous traffic during the service period. However, if the traffic between the source and destination is not continuous, the path remains partly idle during the service, leading to inefficient use of network resources. In packet switching, information transfer takes place in a store and forward fashion from node to node in the form of packets. At every node, the destination address is processed before the next forward move is made. The packets move from one location to another in the general direction of the destination. There is no connection between the source station and the destination station.

In the context of multi services network, it is recognised that packet switching is not suitable for real time services like telephony and video distribution. And, circuit switching is not efficient for bursty non-continuous traffic from the point of view of utilising network resources. The realisation that both circuit switching and packet switching are required for different purposes on a common digital network led to the search for new switching techniques. The two techniques have now converged and a new switching technique called **cell switching** is now accepted as the standard switching technique for future broad band infrastructure. Cell is a small-sized packet that moves very fast on a digital infrastructure. Cell switching is capable of supporting both real time and non-real time services efficiently. However, it is foreseen that for sometime to come, all the three switching techniques would co-exist.

The importance of signalling in networks is reflected by statements such as 'If we cannot signal, we cannot make a call' and 'If we cannot signal, we cannot transfer a packet'. Attempts to convert existing single-service networks to multi-service ones has led to the emergence of a plethora of signalling techniques. Specific convergence questions are how do we support real time services like telephony on IP-based packet networks and how do we transport packets efficiently on circuit switched networks? Circuit-switched and packet-switched networks use entirely different forms of signalling. The signalling protocols are also very different. Convergence signalling protocols on the IP-based packet networks include media gateway control protocol (MGCP), session initiation protocol (SIP) and real time control protocol (RTCP). These are protocols brought out by the Internet Society. ITU-T has brought out recommendation H.323 that allows interworking of circuit switched networks and IP-based packet networks. H.323 refers to a number other signalling protocols used in circuit switched networks. H.323 is the key convergence protocol that is used at present to offer real time services like IP Telephony on IP-based packet networks.

At the transmission level, two transmission hierarchies are emerging. One is ITU recommendation known as synchronous digital hierarchy (SDH) and the other evolved by U.S. in the context of optical networks known as optical carrier (OC) hierarchy. OC transmission systems run on synchronous optical networks (SONET). SDH is applicable to both copper and optical networks. Luckily, the transmission speeds are identical in both SDH and OC at different levels of hierarchy. The speeds at which SONET runs are specified as OC1, OC2, OC3 etc. The lowest speed of SONET is OC1 at 51.84 Mbps. SDH speeds start at OC3 level that corresponds to 155.52 Mbps. The basic speed of ATM network is also OC3.

4.9 ACCESS CONVERGENCE

The economic considerations do not apply only to network providers. It is equally important for the users. A user may not invest in different types of access interfaces for different services. A single telecommunication connection should fetch all types of services to a user subject only to the limitations of speed and capacity of his interface. It is this philosophy that has led to the idea of access convergence. The trend of access convergence is being seen in the following:

1) Access to Internet from mobile handset and vice versa
2) Access to global satellite services from mobile and vice versa
3) Access to Internet from telephone network and vice versa
4) Access to global satellite services from telephone network and vice versa
5) Automatic access among bluetooth devices
6) Access to Internet via television cables and vice versa.

Access convergence provides tremendous flexibility to end-users. They are able to access a variety of services from whatever device they have on hand. For example, a person with a mobile, a tablet or a laptop can literally access all Internet services without physically having a connection to the Internet. Satellite telephony allows telephone conversations directly via satellites from anywhere in the world.

4.10 SERVICE CONVERGENCE

The idea of convergence in different areas that we have discussed so far is to carry multiple services using one common infrastructure. The trend of service convergence, i.e. offering multiple services in one interface, is seen in the following:

1) Real time services like telephony on the Internet
2) E-mail via mobile phones
3) Web access via bluetooth interface and mobile phones
4) E-mail and Web access via television cables
5) Internet services delivered on TV sets
6) Broadcasting radio and TV programmes on the Internet
7) Normal telephone services via cable TV networks.

Broadcasting on the Internet is sometimes called webcasting. The above are examples of some of the recently introduced services using the concept of convergence. For the future, ISDN and broadband ISDN are being designed to carry every conceivable service. Let us now see how service convergence issue is being addressed in broadband ISDN. In broadband ISDN, the services are placed under some generic categories as shown under:

- Audio services
  - Telephone quality (speech)
  - Broadcast quality (music)
- Motion video services
  - Studio quality
  - Broadcast quality
  - High definition television (HDTV)
- Still video services
  - Low resolution (fax)
  - High resolution (pictures and photographs)
- Text services
- Computer graphics services
- Computer animation services
- Satellite imageries

Each of the above services, when digitised, demand different communication capacities, i.e. different speeds. To carry a variety of services with different capacity requirements, the architecture of broadband ISDN uses the ATM information transport mechanism. ATM is structured in a layered manner with different functionality stacked one over another. One of the layers is known as **ATM adaptation layer**. It has a sublayer called **convergence sublayer**. This sublayer has been designed to multiplex a variety of services and carry them on the underlying network infrastructure. The underlying infrastructure for ATM is by and large an optical network like SONET.

4.11 CONVERGENT APPLICATIONS

In this section, we shall discuss some important upcoming convergence applications: Music on demand, Internet radio, Internet telephony and Interactive Television.

Music on demand, known as streaming audio in technical terms deals with listening to music on Internet. As you know, Internet is a data network capable of transporting 1s and 0s. Once digitised, music is in the form of 1s and 0s. Therefore, theoretically digital music can be transported over Internet. Listening to music happens in real time. But, the Internet is not designed to provide real time services. Therefore, the convergence question is 'how to deliver music in real time on Internet?'

Digital sound is compressed before storage. One of the most popular standards used for digitisation and compression is MP3 standard and format. There are many web sites on the Internet that store music in MP3 digital form. The most popular application architecture on the Internet is client-server architecture. The server runs web server software and the client web browser software. The interaction between the web browser and the web server takes place via the application protocol HTTP that runs on TCP. HTTP is not a real time protocol. One way to listen to music under HTTP is to download the entire music file and play it from the client machine. A typical song in MP3 format takes about 4MB and it would take about 8 minutes to download the entire file at 64 kbps speed. Internet rarely runs at that speed. It could sometimes take as much as 15 minutes to download the entire file. The user will have to sit idle during download of each song. This clearly is an unacceptable solution. Streaming audio overcomes this deficiency and provides a near real time solution.

Streaming audio application is designed under client-server architecture supplemented with a media server on the server side and a media player on the client side. The media server and the media player interact via the real time protocol (RTP) that runs on UDP. Initially, the user interacts with the server via the browser for selecting the music bit. Thereafter, the media server and the media player come into picture to play the music in real time. The streaming audio configuration is depicted in Figure 4.3.

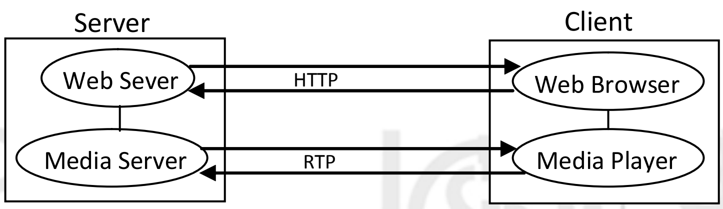


Fig. 4.3: Streaming Audio Configuration

RTP runs on UDP. UDP runs on IP which is a packet protocol. A packet in the case of streaming audio comprises an accumulation of digital samples of music/speech for certain duration, usually 5-8 msec. In other words, each sample is not transmitted in real time but about 40-64 samples are grouped together and transmitted. The grouping of samples and making them available in bursts lead to unnatural audio. However, grouping up to 6 msec is not noticeable by the human ear. As you know, in packet switching, each packet takes a different time to reach the destination depending on the traffic conditions on the network. While an average delay is computable, a packet may arrive much later than the average delay or even be lost on the way. Such delay variation from packet to packet causes jitters in the music affecting the quality. Streaming audio takes care of these problems by providing a music buffer at its input. At the start of a listening session, the buffer is filled with music samples before the music is played out to the user. The buffer holds about 15-20 seconds of music. The user has to wait only for a short while until the buffer is full before he can start listening to music. The music packets continue to arrive from the server while the media player is playing out the music and emptying the buffer contents. This is a streaming operation and hence the name streaming audio. On the one side the buffer is being emptied and on the other it is being filled. The provision of the buffer ensures that continuous music is available even in the presence variable packet delays.

In order to provide flexibility to user in listening, certain control functions as available in a cassette player need to be implemented by the media player-server combination. The control functions include controls like pause, resume, stop, fast forward and fast backward. Some flow control feature is also required. For example, if the packets arrive too fast from the server, then the buffer would become full and the packets would be discarded at the player end. To prevent this from happening, the media player may ask the media server to pause the flow momentarily. All such control functions are implemented using a protocol known as real time control protocol (RTCP) which is used in conjunction with RTP. There are two other protocols associated with streaming services. They are protocols used for establishing a real time session before streaming operation starts. The protocols are real time streaming protocol (RTSP) and session initiation protocol (SIP).

Internet radio is based on the same principles as streaming audio. There are some differences between music-on-demand and Internet radio. In music-on-demand, the music bit is selected by the user whereas in Internet radio, the programme is played out by the station. There is no user interaction for selection. This means that HTTP interaction is absent. User also has no control like pause, resume etc. Some radio stations play a second channel which is delayed by about 10 minutes to allow the users to take a break. After the break, the user may switch over to the delayed channel and listen to the programme from where he/ she left. Another important difference is that music-on-demand is a unicast service whereas Internet radio is a multicast service. Many persons listen to Internet radio at a time. Hence, the same audio stream has to be sent to different destinations.

The architecture of streaming audio is the building block of **streaming video** too. Video bandwidth is much larger and hence requires higher speeds of transmission on the Internet. The buffer size on the client side is also much larger as the sampling rate for digitisation of video is much higher. For example, 10 seconds of audio requires 80 kB of buffer whereas video for the same duration would require 2 MB of buffer. Another important requirement of video streaming is the synchronisation required between the video signal and the accompanying audio signal. Video streaming is the basis for a variety of applications such as distance learning, digital libraries, home shopping and video-on-demand.

Internet telephony is also known as **Voice over Internet Protocol** or simply **Voice over IP** (VoIP). Internet telephony is built around a set of protocols and standards defined by ITU. IT also uses the set of protocols we discussed as part of streaming audio. It was in 1996 that first attempts were made to build Internet telephony gateways when ITU first proposed architecture suitable for voice services in LANs in its standard H 323. This standard was revised in 1998 and forms the basis for Internet telephony. The architecture is depicted in Figure 4.4. Basically, the Internet and the telephone network are interconnected by a gateway. The gateway implements H.323 protocols on the Internet side and PSTN protocols on the telephone network side. The end devices on the Internet are terminals and on the PSTN they are telephones. H.323 refers to a number of other ITU protocols. The telephone speech is digital and should conform to G.711 ITU recommendation that specifies 64 kbps PCM stream. G.711conformity is mandatory for VoIP. Some other compressed audio streams that run at lower speeds like 6.4 kbps or 5.3 kbps are permitted optionally. The digital stream is transported over RTP under the supervision of RTCP. Both RTP and RTCP run on UDP. Since multiple compression schemes are permitted optionally, there is a need to negotiate and agree on the scheme before start of the session. This is done by call control protocol specified in H.245. Call establishment and release and call signalling, like sending out ring tone etc. are done according to ITU signalling protocol Q.931. The call control and call signalling protocols run on TCP. On the LANs, ITU suggests the use of a gatekeeper device that organises telephone call for LAN terminals. A protocol, named registration, admission and status (RAS) has been defined for the gatekeeper to interact with LAN terminals. The RAS protocol runs on UDP.

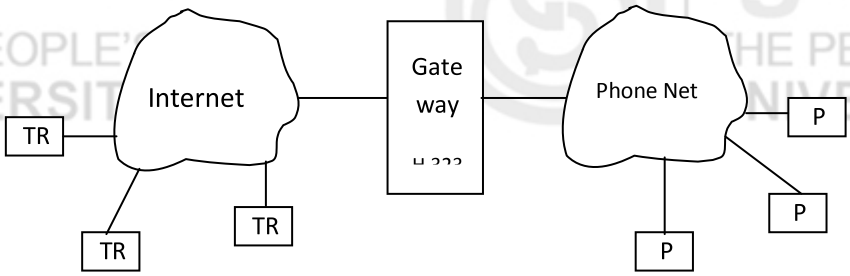


Fig. 4.4: Internet Telephony Architecture

Interactive television (ITV), also called two-way television is a convergence application where a one-way video broadcasting system is turned into a two-way system capable of handling media, i.e. voice, video and data. ITV is a television set with a return path to the broadcaster or the ITV service provider. The return path makes it interactive. Information flows not only from the ITV service provider to the viewer but also from the viewer to the ITV service provider and/or his computer. When fully developed, ITV would offer video-on-demand service. For the present, ITV systems are providing broadcast programmes with facility for interaction. ITV is evolving in three stages:

1) Programmes are broadcast and the return path is provided via PSTN.

2) Programmes are broadcast and the return path is provided via set-top box.

3) Video-on-demand programmes and return path via set-top box.

The set-top box is more than a TV tuner. It has a computer with a phone, coaxial cable or satellite link to the ITV service provider and the Internet. There is a phone modem or a cable modem or a network card, which connects the set-top box to a public network. Unlike broadcast service where the viewer can only watch one of the many programmes currently being broadcast, ITV gives the viewer an individual choice of content that may be exclusive to that viewer.

The potential use of ITV is enormous. To name a few, ITV may be used for marketing, advertising, child counselling, public relations, education and even politics. For the user, there is the promise of choice, fun, convenience and empowerment. Sitting at home, one will be able to get literally any product or service delivered at the touch of a button. This is known as t-commerce, television commerce. Users may click on advertisements to know more about the product. Viewers may choose camera angles while watching their favourite sport events. This indeed is exciting. Users may pause and resume programmes that they are viewing, provided the programmes are not live. With live programmes, the users may record the programmes for later viewing. And, of course, e-mail can be sent through ITV.

There is one serious aspect of ITV, which is now spreading to Internet. All ITV systems have a feature called click stream analysis. This feature creates a complete record of the clicks that a user performs on his/ her set-top box. This record is later analysed to build profiles of users. In a positive sense, the purpose is to provide the user information that he/she is interested in. The negative aspect is that the service provider is actually treading into the private life of individuals.

ITV systems call for wide band infrastructure to transport individual programmes to different viewers. This infrastructure is an ATM network with synchronous optical network (SONET) as the transport infrastructure.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

5) Differentiate between streaming audio and Internet radio.

....

....

....

....

4.12 SUMMARY

In this Unit the convergent aspects in the context of NEIS have been discussed. It introduces you to the concept of convergence and discusses the driving factors viz. societal need, economics of operation and technology support. The convergent phenomena in technology, network, access and service areas are explained in detail. The Unit finally covers the convergent applications like internet radio and interactive TV.

4.13 ANSWERS TO SELF-CHECK EXERCISES

1) The process of bringing together different technologies, techniques, networks and services to realise the common objective of evolving towards Networked Electronic Information Society (NEIS) is known as convergence.

2) Integrated consumer devices are capable of transmitting and receiving audio, video and data. Such devices enable the users to access multimedia applications on the global network. We are witnessing the concept of integrated consumer device being applied in mobile handsets extensively. As you may be aware, we have mobile handsets that have built-in camera, FM radio, MP3 audio player, Internet access feature download facility etc.

3) There are only three generic forms of information in NEIS: Audio, Video and Data. Audio includes speech and music. Video includes television, movies pictures, photos etc. Data includes all computer-generated information such as text, computer graphics, computer animation, digitised documents etc. Every application in NEIS is built around these three fundamental forms of information.

4) Remote computing has turned out to be win-win solution for both users and the computer owners as users could access the computer easily and at the same time the utilisation of the computer is optimised.

5) Internet radio is based on the same principles as streaming audio. However, there are some differences between them. In streaming audio, the music bit is selected by the user whereas in Internet radio, the programme is played out by the station. There is no user interaction for selection. This means that HTTP interaction is absent. User also has no control like pause, resume etc. Some radio stations play a second channel which is delayed by about 10 minutes to allow the users to take a break. After the break, the user may switch over to the delayed channel and listen to the programme from where he/ she left. Another important difference is that streaming audio is a unicast service whereas Internet radio is a multicast service. Many persons listen to Internet radio at a time. Hence, the same audio stream has to be sent to different destinations.

4.14 KEYWORDS

Bluetooth : A wireless networking technology that uses short-wave radio frequencies to interconnect cell phones, portable computers, and other wireless electronic devices.

Cell Switching : A convergent switching technique that combines the strengths of circuit and packet switching.

Circuit Switching : The process of establishing a dedicated connection between two end points via a network.

Distributive Services : Services that are broadcast or multicast in nature i.e. services that are distributed to a large number of users without a request from a user.

On- demand Services : Services that can be obtained by users by placing a demand on the network.

Packet Switching : The process of transferring information from a source to a destination on a node-by-node basis.

RJE Terminal : Remote job entry terminal, an equipment set up to use computers from remote places.

4.15 REFERENCES AND FURTHER READING

Douskalis, B. *IP Telephony*. New Delhi: Pearson Education Asia, 2001. Print

Hersent, O., Gurle, D. and Petit, J.P. *IP Telephony: Packet-Based Multimedia Communication System*s. New Delhi: Pearson Education Limited, 2001. Print

Johnston, C. B. Communication Conglomerates. In: *Encylopedia of Information Systems*, Volume 2, USA: Elsevier Science. Print

Pavlik J. V. And Powell, A. C. New Media. In: *Encylopedia of Information Systems*, Volume 2, USA: Elsevier Science. Print

Rao, K. R., Bojkovic, Z. S. and Milovanovic, D. A. *Multimedia Communication Systems*. New Delhi: Prentice Hall of India, 2002. Print

Viswanathan, T. *Convergence of the Telecommunications, Information Technology, Media and Power Sectors.* New Delhi: Centre of Publications, 2001. Print

Viswanathan, T. *New Telecom Policy*. New Delhi: Government of India, 1999. Print

BLOCK 2 MIDDLEWARE TECHNOLOGIES

Introduction

Middleware refers to a software that acts as a bridge between an operating system or database and applications. Three basic types of middleware are: i) communication middleware, ii) database middleware, and iii) system middleware. Middleware has two separate but related connotations: i) software that enables two separate programs to interact with each other and ii) software layer inside a single application that allows different aspects of the program to work together.

In this block we will discuss some of the commonly used software tools like office tools and database management systems. The block also discusses the multimedia tools covering different aspects of multimedia production and delivery.

This block has three Units:

Unit 5 on **Office Tools: Word Processing, Presentation and Spreadsheets** deals with open source Office productivity tool LibreOffice covering its three components, Writer, Impress and Calc.

Unit 6 on **Database Management Systems** introduces the database approach, database management system (DBMS), database environment, database security and some popular DBMS packages.

Unit 7 covers **Multimedia** discussing characteristics of Multimedia systems, design and storyboarding aspects, tools for multimedia development, storing and retrieval and data integrity issues.

UNIT 5 OFFICE TOOLS: WORD PROCESSING, PRESENTATION AND SPREADSHEETS

Structure

5.0 OBJECTIVES

After going through this Unit, you will be able to:

- understand the features of LibreOffice as a freely available, open source, fully-featured office productivity suite;

- use LibreOffice applications such as Writer (word processor or text editor), Impress (presentation) and Calc (spreadsheet); and
- explain how different applications wizards in the office tools help you in handling tasks for its various applications.

5.1 INTRODUCTION

Office tools are applications that allow viewing, creating and modifying of general office documents (e.g letters, memos, reports, presentations, spreadsheets, image editing, etc.). Office productivity tools also include applications for managing employee tasks.

Choosing free and open source software over proprietary software does not mean that you compromise on features and support. Open source office tools are as feature-rich as proprietary tools. They also provide ample online documentation and have large communities of users and developers as support system. These office tools are generally referred to as office suite, productivity suite, applications tools because they come as a collection of applications mainly consisting of word processor, spreadsheet and presentation bundled together and quite often sharing a common user interface.

With an **open-source licence**, **LibreOffice** can be freely used and distributed with no licence worries. **LibreOffice** is a leading open-source office software suite for word processing, spreadsheets, presentations, graphics, databases and more. It is available in many languages and works on all platforms. It stores all your data in an international open standard format and can also read and write files from other common office software packages. It can be downloaded and used completely free of charge for any purpose.

In this Unit you will be introduced to the LibreOffice applications: Writer (word processing), Impress (presentation) and Calc (spreadsheet). This Unit has been adapted as a summary from the 'Getting Started Guide 4.2' – the official documentation issued by The Document Foundation, the body responsible for developing and launching LibreOffice suite. The material is available under the terms of either the GNU General Public License (http://www.gnu.org/licenses/gpl.html), version 3 or later, or the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), version 4.0 or later for distribution and modification. The documents used are listed in the References and Further Readings section for further reference.

5.2 GETTING STARTED WITH LIBREOFFICE SUIT

LibreOffice is a freely available, fully-featured office productivity suite. It is a powerful open-source office suite that provides tools for all types of office tasks such as writing texts, working with spreadsheets, creating graphics and presentations, or designing scientific formulas. With LibreOffice, you can use the same data across different computing platforms. You can also open and edit files in other formats, including Microsoft Office, then save them back to LibreOffice format. Its native file format is Open Document Format (ODF), an open standard format that is being adopted by governments worldwide as a required file format for publishing and accepting documents.

LibreOffice comprises several application modules designed to interact with each other. These modules feature the same graphical user interface and similar functionality. LibreOffice includes the following components:

Writer (word processor) – it is a full-featured word processor with page and text-formatting capabilities. It also includes some features that are usually found only in expensive desktop publishing applications.

Calc (spreadsheet) – it has all of the advanced analysis, charting, and decision making features expected from a high-end spreadsheet.

Impress (presentations) – Impress provides all the common multimedia presentation tools, such as special effects, animation, and drawing tools. It is integrated with the advanced graphics capabilities of LibreOffice Draw and Math components.

Draw (vector graphics)

Draw is a vector drawing tool that can produce everything from simple diagrams or flowcharts to 3D artwork.

Base (database) – Base provides tools for day-to-day database work within a simple interface. It can create and edit forms, reports, queries, tables, views, and relations, so that managing a relational database is much the same as in other popular database applications.

Math (formula editor)

Math is the LibreOffice formula or equation editor. You can use it to create complex equations that include symbols or characters not available in standard font sets. While it is most commonly used to create formulae in other documents, such as Writer and Impress files, Math can also work as a standalone tool.

5.2.1 Advantages of LibreOffice

The advantages of LibreOffice over other office suites are:

- **No licensing fees required** – LibreOffice is free for anyone to use and distribute at no cost. Many features like PDF export are available as extra cost add-ins in other office suites whereas in LibreOffice it is free. There are no hidden charges involved at present and not expected in the future as well.
- **Open source** – You can distribute, copy, and modify the software as much as you wish, in accordance with the LibreOffice Open Source license policy.
- **Cross-platform**- LibreOffice runs on several hardware architectures and under multiple operating systems, such as Microsoft Windows, Mac OS X and Linux.
- **Extensive language support** – The LibreOffice user interface is available in over 40 languages and the LibreOffice project provides spelling, hyphenation, and thesaurus dictionaries in over 70 languages and dialects. LibreOffice also provides support for both Complex Text Layout (CTL) and Right to Left (RTL) layout languages (such as Urdu, Hebrew, and Arabic).
- **Consistent user interface** – All the components have a similar "look and feel," making them easy to use and master.
- **Integration** – The components of LibreOffice are well integrated with one another. They share a common spelling checker and other tools. The drawing tools for example available in Writer are also found in Calc and with similar but enhanced versions in Impress and Draw. Another advantage is that there is no need to know which application was used to create a particular file. For instance one can open a Draw file from Writer.
- **Granularity** – LibreOffice options can be set at a component level or even at document level. One may change options at component level so as not to disturb the other components.

- **File compatibility** – In addition to its native OpenDocument formats, LibreOffice includes PDF and Flash export capabilities, as well as support for opening and saving files in many common formats including Microsoft Office, HTML, XML, WordPerfect, and Lotus 1-2-3 formats. An extension (included) provides the ability to import and edit some PDF files.
- **No vendor lock-in** – LibreOffice uses OpenDocument, an XML (eXtensible Markup Language) file format developed as an industry standard by OASIS (Organization for the Advancement of Structured Information Standards). These files can easily be unzipped and read by any text editor, and their framework is open and published.
- **You have a voice** – Enhancements, software fixes, and release dates are community driven. You can join the community and affect the course of the product you use.

You can read more about LibreOffice and The Document Foundation on their websites at: http://www.libreoffice.org/ and http://www.documentfoundation.org/.

The software can be downloaded from http://www.libreoffice.org/download. For Information on installing and setting up LibreOffice on the various supported operating systems visit: http://www.libreoffice.org/get-help/installation/.

The most common way to launch any component of LibreOffice is by using the system menu, the standard menu from which most applications are started. Clicking on the LibreOffice menu entry or tile opens the LibreOffice Start Center as shown in Fig. 5.1. From the Start Centre individual components of LibreOffice may be selected to open. You can also select to open an existing file or use a template from the Start Centre.

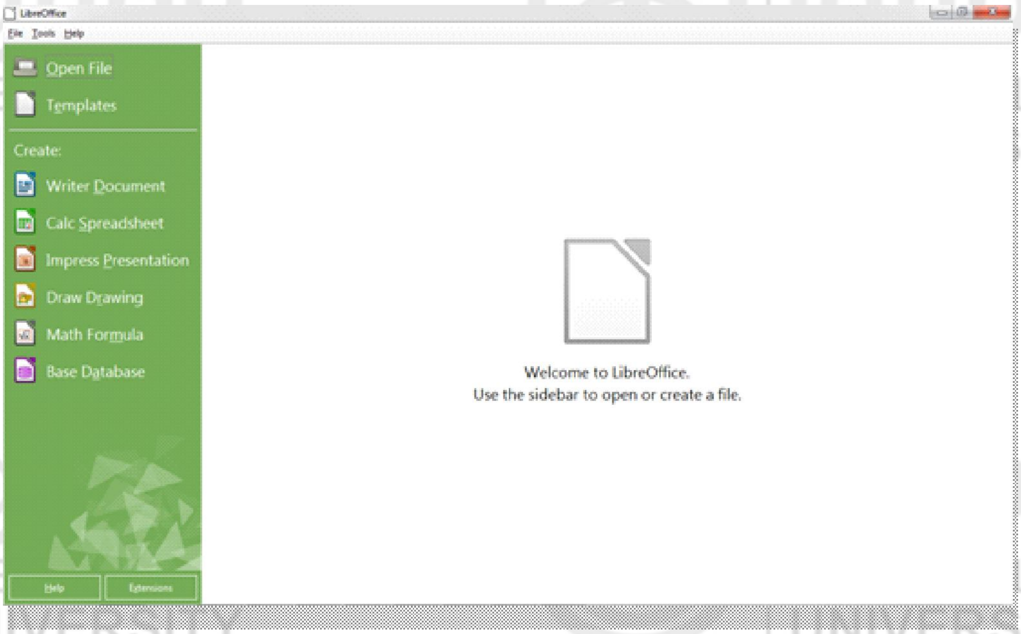


Fig. 5.1: Start Centre of LibreOffice

For each component of LibreOffice, the main window is almost similar. Common features include the menu bar, standard toolbar, and formatting toolbar at the top of the window and the status bar at the bottom as shown in Fig. 5.2.

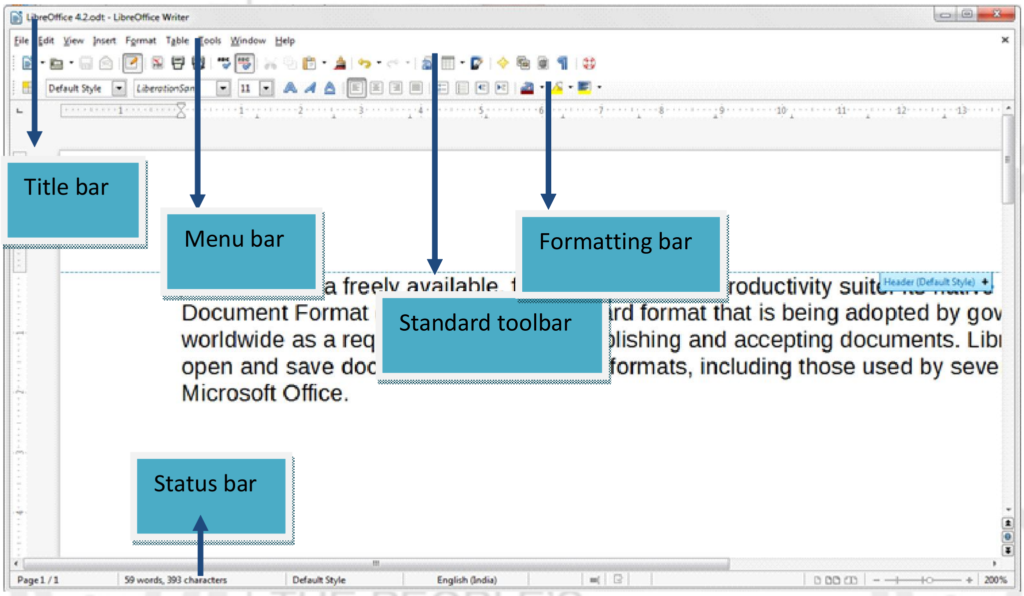


Fig. 5.2: LibreOffice Standard Interface

Menu bar

The menu bar is located just below the title bar on top of the LibreOffice window. The menu options available are: File, Edit, View, Insert, Format, Table, Tools, Window and Help as common in other Office applications. On selecting one of the menus, a sub-menu drops down to show commands.

Toolbars

LibreOffice has two types of toolbars: docked (fixed in place) and floating. Docked toolbars can be moved to different locations or made to float, and floating toolbars can be docked. In a default LibreOffice installation, the top docked toolbar, just under the menu bar, is called the Standard toolbar.

The second toolbar at the top, in a default LibreOffice installation, is the Formatting bar. It is context-sensitive i.e., it shows the tools relevant to the current position of the cursor or the object selected. For example, when the cursor is on a graphic, the Formatting bar provides tools for formatting graphics.

Status bar

The status bar is located at the bottom of the workspace. It provides information about the document. It is similar in Writer, Calc, Impress, and Draw, but each LibreOffice component includes some component-specific items. An example of the Writer status bar is shown in Fig. 5.3

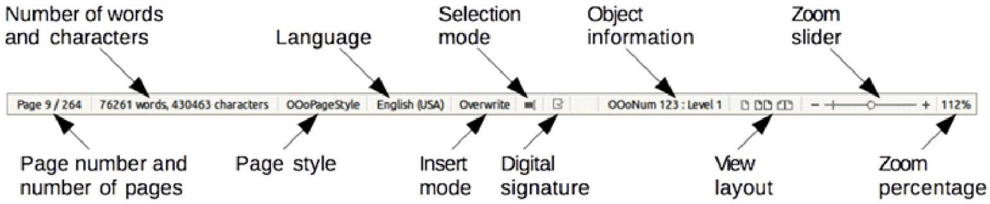


Fig. 5.3: Status bar in Writer

5.3 WORD PROCESSING WITH LIBREOFFICE: WRITER

LibreOffice Writer is the word processor application. In addition to the usual features of a word processor (spelling check, thesaurus, hyphenation, autocorrect, find and replace, automatic generation of tables of contents and indexes, mail merge and others), Writer provides the following important features:

- Templates and styles.
- Page layout methods, including frames, columns, and tables.
- Embedding or linking of graphics, spreadsheets, and other objects.
- Built-in drawing tools.
- Master documents—to group a collection of documents into a single document
- Change tracking during revisions.
- Database integration, including a bibliography database.
- Export to PDF, including bookmarks.

Apart from these above mentioned features there are many more which are relevant for all levels of users.

5.3.1 Working with Text

Working with text (selecting, copying, pasting, moving) in Writer is similar to working with text in any other program. LibreOffice also has some convenient ways to select items that are not next to each other, select a vertical block of text, and paste unformatted text.

Selecting items that are not consecutive:

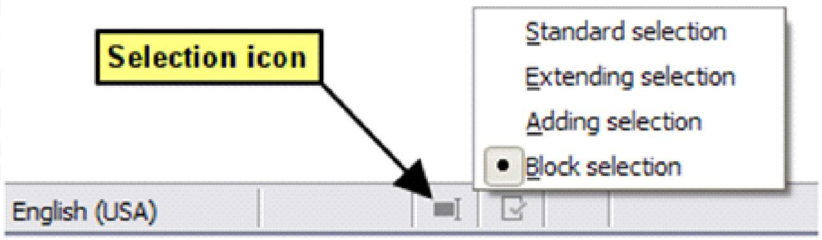
1) Select the first piece of text.
2) Hold down the *Ctrl* key and use the mouse to select the next piece of text.
3) Repeat as often as needed.

To select nonconsecutive items using the keyboard:

1) Select the first piece of text. (For more information about keyboard selection of text, see the topic "Navigating and selecting with the keyboard" in the Help.)
2) Press *Shift+F8*. This puts Writer in "Adding selection" mode.
3) Use the arrow keys to move to the start of the next piece of text to be selected. Hold down the *Shift* key and select the next piece of text.
4) Repeat as often as required. Now Press *Esc* to exit from this mode.

Selecting a vertical block of text

You can select a vertical block or "column" of text that is separated by spaces or tabs (as you might see in text pasted from e-mails, program listings, or other sources), using LibreOffice's block selection mode. To change to block selection mode, use **Edit > Selection Mode > Block Area**, or press *Ctrl+F8*, or click on the **Selection** icon in the Status Bar and select **Block selection** from the list.



Now highlight the selection, using mouse or keyboard, as shown below.

January	February	March
April	May	June
July	August	September
October	November	December

Fig. 5.4: Selecting a vertical block of text

Cutting, copying, and pasting text

Cutting and copying text in Writer is similar to cutting and copying text in other applications. You can use the mouse or the keyboard for these operations. You can copy or move text within a document, or between documents, by dragging or by using menu selections, toolbar buttons, or keyboard shortcuts. You can also copy text from other sources such as Web pages and paste it into a Writer document.

To *move* (drag and drop) selected text using the mouse, drag it to the new location and release it. To *copy* selected text, hold down the *Ctrl* key while dragging. The text retains the formatting it had before dragging.

To *move* (cut and paste) selected text, use *Ctrl+X* to cut the text, insert the cursor at the paste-in point and use *Ctrl+V* to paste. Alternatively, use the buttons on the **Standard** toolbar.

When you paste text, the result depends on the source of the text and how you paste it. If you click on the **Paste** button, any formatting the text has (such as bold or italics) is retained. Text pasted from Web sites and other sources may also be placed into frames or tables. If you do not like the results, click the **Undo** button or press *Ctrl+Z*.

To make the pasted text take on the formatting of the surrounding text where it is being pasted:

- Choose **Edit > Paste Special**, or
- Click the arrow button of the combination **Paste** button, or
- Click the **Paste** button without releasing the left mouse button. Then select **Unformatted text** from the resulting menu.

The range of choices on the Paste Special menu varies depending on the origin and formatting of the text (or other object) to be pasted.

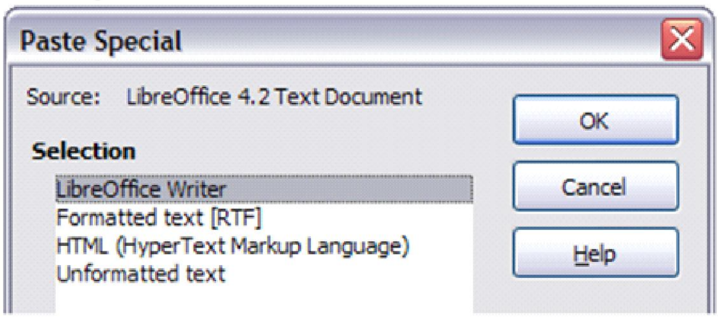


Fig. 5.5: Paste Special menu

Finding and replacing text and formatting

Writer has two ways to find text within a document: the Find toolbar for fast searching and the Find & Replace dialog. In the dialog, you can:

- Find and replace words and phrases
- Use wildcards and regular expressions to fine-tune a search
- Find and replace specific attributes or formatting
- Find and replace paragraph styles

Using the Find & Replace dialog

To display the Find & Replace dialog, use the keyboard shortcut *Ctrl+H* or choose **Edit > Find & Replace** from the Menu bar. If the Find toolbar is open, click the Find and Replace button ( ) on the toolbar. Once opened, optionally click the **Other Options** symbol to expand the dialog. Click the button again to reduce the dialog options.

To use the Find & Replace dialog:

1) Type the text you want to find in the **Search for** box.
2) To replace the text with different text, type the new text in the **Replace with** box.
3) You can select various options such as matching the case, matching whole words only, or doing a search for similar words.
4) When you have set up your search, click **Find**. To replace the found text, click **Replace.**

Inserting special characters

A *special character* is one not found on a standard English keyboard. For example, © ¾ æ ç ñ ö ø ¢ are all special characters. To insert a special character:

1) Place the cursor where you want the character to appear.
2) Choose **Insert > Special Character** to open the Special Characters dialog.
3) Select the characters (from any font or mixture of fonts) you wish to insert, in order, then click **OK**. The characters selected for insertion are shown in the lower left of the dialog. As you select a character, it is shown on the right, along with its numerical code.

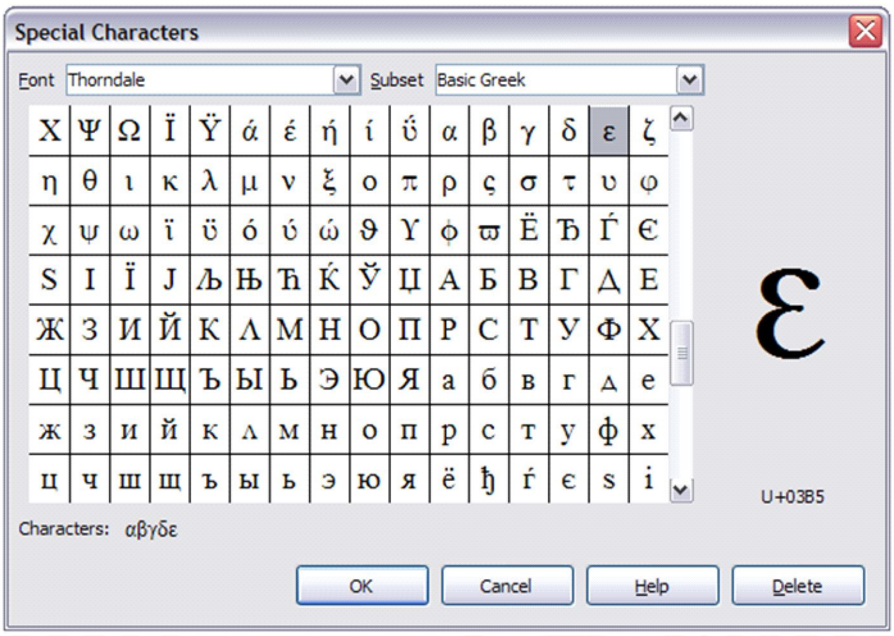


Fig. 5.6: The Special Characters dialog, where you can insert special characters

Checking spelling and grammar

Writer provides a spelling checker, which can be used in two ways.

 **AutoSpellcheck** checks each word as it is typed and displays a wavy red line under any unrecognised words. When the word is corrected, the line disappears.

 To perform a combined spelling and grammar check on the document (or a text selection) click the **Spelling and Grammar** button. This checks the document or selection and opens the Spelling and Grammar dialog if any unrecognised words are found. In order to use this, the appropriate dictionaries must be installed. By default, four dictionaries are installed: a spellchecker, a grammar checker, a hyphenation dictionary, and a thesaurus.

Using built-in language tools

Writer provides some tools that make your work easier if you mix multiple languages within the same document or if you write documents in various languages.

The main advantage of changing the language for a text selection is that you can then use the correct dictionaries to check spelling and apply the localised versions of Autocorrect replacement tables, thesaurus, grammar, and hyphenation rules.

You can also set the language for a paragraph or a group of characters as **None (Do not check spelling)**. This option is especially useful when you insert text such as web addresses or programming language snippets that you do not want to check for spelling.

You can also set the language for the whole document, for individual paragraphs, or even for individual words and characters, from **Tools > Language** on the Menu bar.

Another way to change the language of a whole document is to use **Tools > Options > Language Settings > Languages**. In the *Default languages for documents* section, you can choose a different language for all the text that is not explicitly marked as a different language.

The spelling checker works only for those languages in the list that have the symbol () next to them. If you do not see the symbol next to your preferred language, you can install the additional dictionary using **Tools > Language > More Dictionaries Online**.

The language used for checking spelling is also shown in the Status Bar, next to the page style in use.

Using AutoCorrect

Writer's AutoCorrect function has a long list of common misspellings and typing errors, which it corrects automatically. For example, "hte" will be changed to "the".

AutoCorrect is turned on when Writer is installed. To turn it off, uncheck **Format > AutoCorrect > While Typing**.

Choose **Tools > AutoCorrect Options** to open the AutoCorrect dialog. There you can define which strings of text are corrected and how. In most cases, the defaults are fine.

To stop Writer replacing a specific spelling, go to the **Replace** tab, highlight the word pair, and click **Delete**.

To add a new spelling to the list, type it into the *Replace* and *With* boxes on the Replace tab, and click **New**.

See the different tabs of the dialog for the wide variety of other options available to fine-tune AutoCorrect.

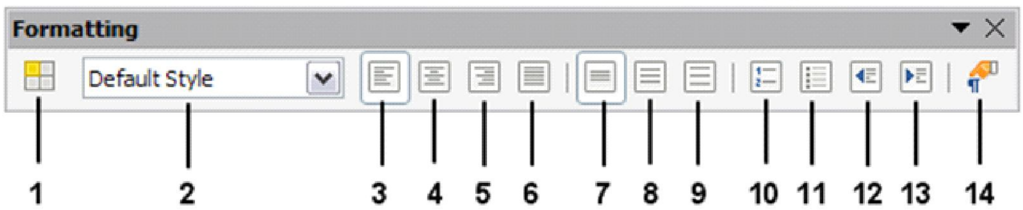
5.3.2 Formatting Text

Styles are central to using Writer. Styles enable you to easily format your document consistently, and to change the format with minimal effort. A style is a named set of formatting options. When you apply a style, you apply a whole group of formats at the same time. In addition, styles are used by LibreOffice for many processes, even if you are not aware of them. For example, Writer relies on heading styles (or other styles you specify) when it compiles a table of contents.

Writer defines several types of styles, for different types of elements: characters, paragraphs, pages, frames, and lists.

Formatting paragraphs

You can apply many formats to paragraphs using the buttons on the Formatting toolbar and by using the Paragraph panel of the Sidebar's Properties deck. Fig. 5.7 shows the Formatting toolbar as a floating toolbar, customised to show only the buttons for paragraph formatting. The appearance of the buttons may vary with your operating system and the selection of icon size and style in **Tools > Options > LibreOffice > View**.



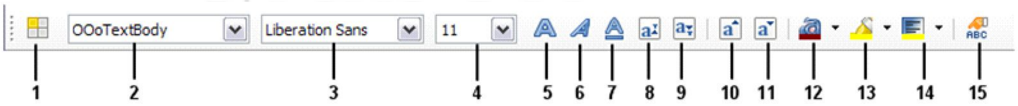
1 Open Styles and Formatting Window	**5** Align Right	**10** Numbering On/Off
	6 Justified	**11** Bullets On/Off
2 Apply Style	**7** Line Spacing: 1	**12** Decrease Indent
3 Align Left	**8** Line Spacing: 1.5	**13** Increase Indent
4 Centered	**9** Line Spacing: 2	**14** Paragraph format dialog

Fig. 5.7: Formatting toolbar, showing buttons for paragraph formatting

Formatting characters

You can apply many formats to characters using the buttons on the Formatting toolbar. Fig. 5.8 shows the Formatting toolbar, customised to include only the buttons for character formatting. The Character panel of the Sidebar's Properties deck also provides buttons for character formatting.

The appearance of the buttons may vary with your operating system and the selection of icon size and style in **Tools > Options > LibreOffice > View**.



1. Open Styles and Formatting Window	**6. Italic**	**11. Reduce Font**
2. Apply Style	**7. Underline**	**12. Font Colour**
3. Font Name	**8. Superscript**	**13. Highlighting**
4. Font Size	**9. Subscript**	**14. Background Colour**
5. Bold	**10. Increase Font**	**15. Open Character Format Dialog**

Fig. 5.8: Formatting toolbar, showing buttons for character formatting

Autoformatting

You can set Writer to automatically format parts of a document according to the choices made on the Options page of the AutoCorrect dialog (**Tools > AutoCorrect Options**).

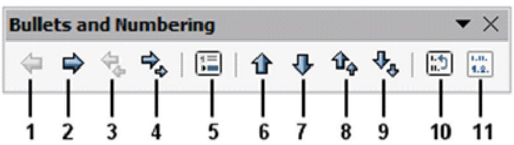
Creating numbered or bulleted lists

There are several ways to create numbered or bulleted lists:

- Use autoformatting, as described above.
- Use list (numbering) styles.
- Use the **Numbering** and **Bullets** buttons on the Formatting toolbar

Using the Bullets and Numbering toolbar

You can create nested lists (where one or more list items has a sub-list under it, as in an outline) by using the buttons on the Bullets and Numbering toolbar (Fig. 5.9). You can move items up or down the list, create sub-points, change the style of bullets, and access the Bullets and Numbering dialog, which contains more detailed controls. Use **View > Toolbars > Bullets and Numbering** to see the toolbar.



1 Promote One Level	**5** Insert Unnumbered Entry	**8** Move Up with Subpoints
2 Demote One Level	**9** Move Down with Subpoints	**10** Restart Numbering
3 Promote One Level with Subpoints	**6** Move Up	**11** Bullets and Numbering
4 Demote One Level with Subpoints	**7** Move Down	

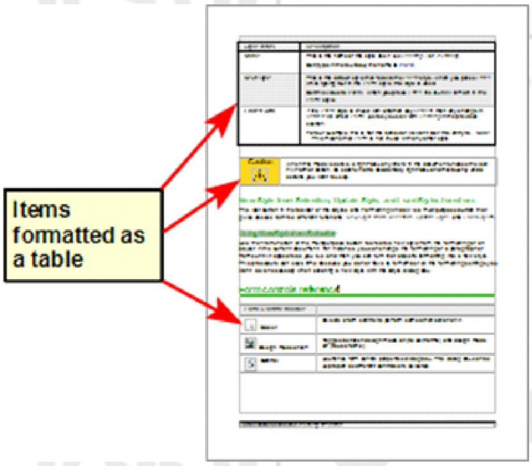
Fig. 5.9: Bullets and Numbering toolbar

5.3.3 Formatting Pages

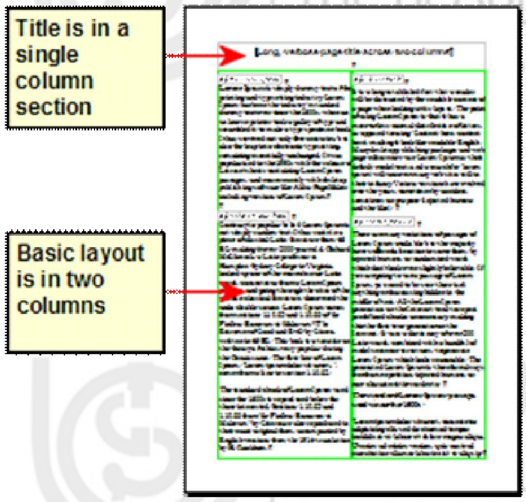
Writer provides several ways for you to control page layouts: page styles, columns, frames, tables, and sections.

The best layout method depends on what the final document should look like and what sort of information will be in the document. Here are some examples.

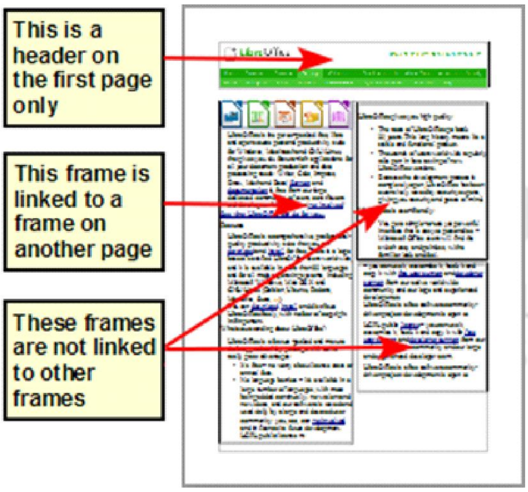
For a book with one column of text, some figures without text beside them, and some other figures with descriptive text, use page styles for basic layout, and tables to place figures beside descriptive text when necessary.



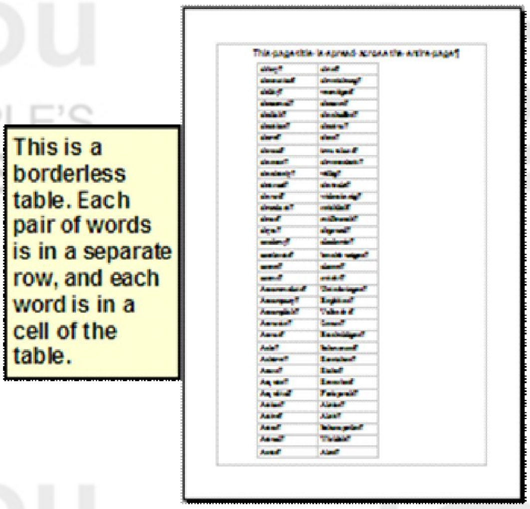
For an index or other document with two columns of text, where the text continues from the left-hand column to the right-hand column and then to the next page, all in sequence (also known as "snaking columns" of text), use page styles (with two columns). If the title of the document (on the first page) is full-page width, put it in a single-column section.



For a newsletter with complex layout, two or three columns on the page, and some articles that continue from one page to some place several pages later, use page styles for basic layout. Place articles in linked frames and anchor graphics to fixed positions on the page if necessary.



For a document with terms and translations to appear side-by-side in what appear to be columns, use a table to keep items lined up, and so you can type in both “columns”.



Creating headers and footers

A header is an area that appears at the top of a page above the margin. A footer appears at the bottom of the page below the margin. Information such as page numbers inserted into a header or footer displays on every page of the document with that page style.

To insert a header, you can either:

- Choose **Insert > Header > Default Style** (or some other page style, if not Default Style), or
- Click above the top margin to make the Header marker appear (Fig. 5.10), and then click on the +.

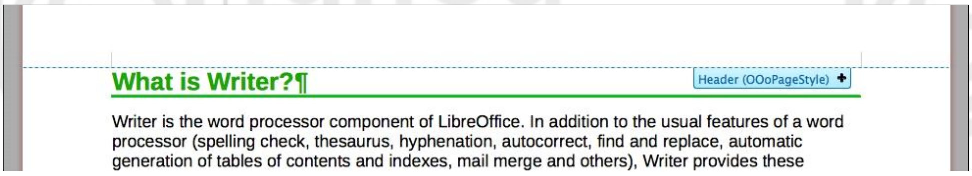


Fig. 5.10: Header marker at top of text area

After a header has been created, a down-arrow appears on the header marker. Click on this arrow to drop down a menu of choices for working with the header (Fig. 5.11).

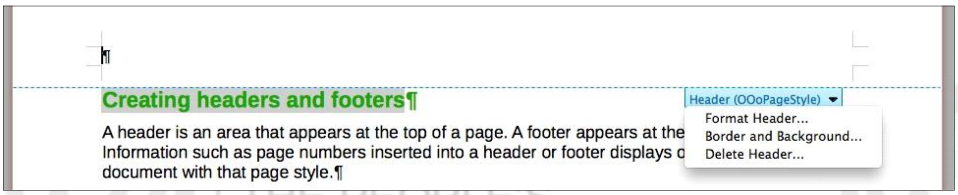


Fig. 5.11: Header menu

To format a header, you can use either the menu item shown in Figure 26 or **Format > Page > Header**. Both methods take you to the same tab on the Page Style dialog.

Other information such as document titles and chapter titles is often put into the header or footer. These items are best added as fields. That way, if something changes, the headers and footers are updated automatically. Here is one common example.

To insert the document title into the header:

1) Choose **File > Properties > Description** and type a title for your document.

2) Add a header (**Insert > Header > Default**).

3) Place the cursor in the header part of the page.

4) Choose **Insert > Fields > Title**. The title should appear on a gray background (which does not show when printed and can be turned off).

5) To change the title for the whole document, go back to **File > Properties > Description**.

Changing page margins

You can change page margins in three ways:

- Using the page rulers—quick and easy, but does not have fine control
- Using the Page Style dialog—can specify margins to two decimal places
- Using the Page panel on the Properties deck of the Sidebar

To change margins using the rulers:

1) The gray sections of the rulers are the margins. Put the mouse cursor over the line between the gray and white sections. The pointer turns into a double-headed arrow and displays the current setting in a tool-tip.

2) Hold down the left mouse button and drag the mouse to move the margin.

Fig. 5.12: Moving the margins

To change margins using the Page Style dialog:

1) Right-click anywhere in the text area on the page and select **Page** from the context menu.

2) On the **Page** tab of the dialog, type the required distances in the Margins boxes.

To change margins using the Page panel of the Properties deck of the Sidebar:

1) On the open Sidebar (**View > Sidebar**) select the **Properties** tab.

2) Open the Page panel if is not open by clicking the plus (+) symbol in the panel title

3) Click the **Margin** button to open the sub-panel and enter the required dimensions in the **Custom** size boxes (clicking the **More Options** button will open the Page Style dialog).

5.3.4 Using Mail Merge

Writer provides very useful features to create and print:

- Multiple copies of a document to send to a list of different recipients (form letters)

- Mailing labels
- Envelopes

All these facilities use a registered data source (a spreadsheet or database containing the name and address records and other information).

5.3.5 Tracking Changes to a Document

You can use several methods to keep track of changes made to a document.

1) Make your changes to a copy of the document (stored in a different folder, or under a different name, or both), then use Writer to combine the two files and show the differences. Choose **Edit > Compare Document**. This technique is particularly useful if you are the only person working on the document, as it avoids the increase in file size and complexity caused by the other methods.

2) Save versions that are stored as part of the original file. However, this method can cause problems with documents of non-trivial size or complexity, especially if you save a lot of versions. Avoid this method if you can.

3) Use Writer's change marks (often called "redlines" or "revision marks") to show where you have added or deleted material, or changed formatting. Choose **Edit > Changes > Record** before starting to edit. Later, you or another person can review and accept or reject each change. Choose **Edit > Changes > Show**. Right-click on an individual change and choose **Accept Change** or **Reject Change** from the context menu, or choose **Edit > Changes > Accept or Reject** to view the list of changes and accept or reject them.

5.3.6 Linking to another Part of a Document

If you type in cross-references to other parts of a document, those references can easily get out of date if you reorganize the order of topics, add or remove material, or reword a heading. Writer provides two ways to ensure that your references are up to date, by inserting links to other parts of the same document or to a different document:

- Hyperlinks
- Cross-references

The two methods have the same result if you *Ctrl+click* the link when the document is open in Writer: you are taken directly to the cross-referenced item. However, they also have major differences:

- The text in a hyperlink does **not** automatically update if you change the text of the linked item (although you can change it manually), but changed text does automatically update in a cross-reference.
- When using a hyperlink, you do not have a choice of the content of the link (for example text or page number), but when using a cross-reference, you have several choices, including bookmarks.
- To hyperlink to an object such as a graphic, and have the hyperlink show useful text. You need to give such an object a useful name (instead of a default name like *Graphics6)*, or use the Hyperlink dialog to modify the visible text. In contrast, cross- references to figures with captions automatically show useful text, and you have a choice of several variations of the name.

- If you save a Writer document to HTML, hyperlinks remain active but cross-references do not. (Both remain active when the document is exported to PDF.)

Using hyperlinks

The easiest way to insert a hyperlink to another part of the same document is by using the Navigator:

1) Open the document containing the items you want to cross-reference.

2) Open the Navigator by clicking its button, choosing **View > Navigator**, or by pressing *F5*.

3) Click the arrow part of the combination **Drag Mode** button, and choose **Insert as Hyperlink**.

4) In the list at the bottom of the Navigator, select the document containing the item that you want to cross-reference.

5) In the Navigator list, select the item that you want to insert as a hyperlink.

6) Drag the item to where you want to insert the hyperlink in the document. The name of the item is inserted in the document as an active hyperlink.

You can also use the Hyperlink dialog to insert and modify hyperlinks within and between documents.

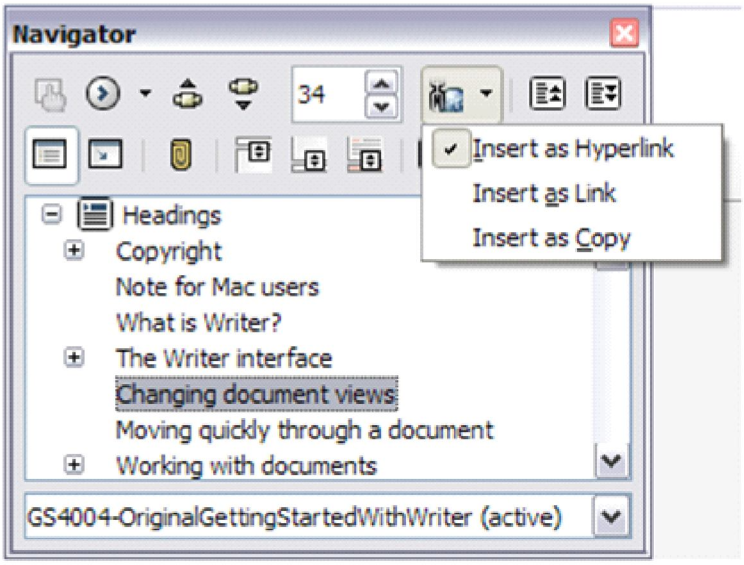


Fig. 5.13: Inserting a hyperlink using the Navigator

Using cross-references

If you type in references to other parts of the document, those references can easily get out of date if you reword a heading, add or remove figures, or reorganize topics. Replace any typed cross- references with automatic ones and, when you update fields, all the references will update automatically to show the current wording or page numbers. The *Cross-references* tab of the Fields dialog lists some items, such as headings, bookmarks, figures, tables, and numbered items such as steps in a procedure. You can also create your own reference items.

To insert a cross-reference to a heading, figure, bookmark, or other item:

1) In your document, place the cursor where you want the cross-reference to appear.

2) If the Fields dialog is not open, click **Insert > Cross-reference**. On the *Cross-references* tab (Fig. 5.14), in the *Type* list, select the type of item to be referenced (for example, *Heading* or *Figure*). You can leave this page open while you insert many cross-references.

3) Click on the required item in the *Selection* list, which shows all the items of the selected type. In the *Insert reference to* list, choose the format required. The list varies according to the Type. The most commonly used options are **Reference** (to insert the full text of a heading or caption), **Category and Number** (to insert a figure number preceded by the word *Figure* or *Table*, but without the caption text), **Numbering** (to insert only the figure or table number, without the word "Figure" or "Table"), or **Page** (to insert the number of the page the referenced text is on). Click **Insert**.

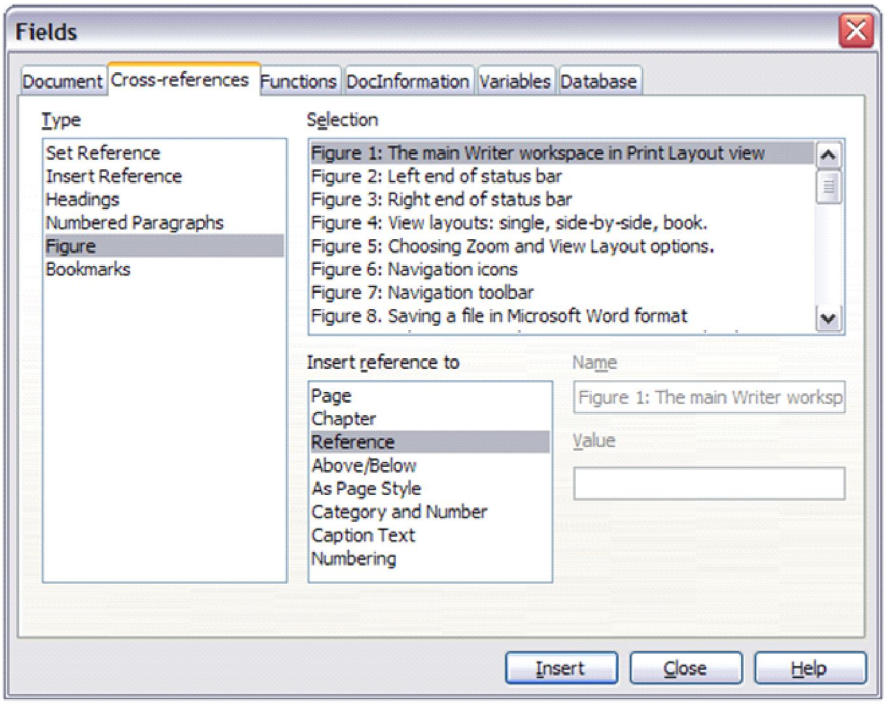


Fig. 5.14: The Cross-references tab of the Fields dialog

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Enumerate important features of LibreOffice Writer.

....

....

....

....

5.4 PRESENTATIONS WITH LIBREOFFICE: IMPRESS

Impress is the presentation (slide show) program included in LibreOffice. You can create slides that contain many different elements, including text, bulleted and numbered lists, tables, charts, and a wide range of graphic objects such as clipart, drawings and photographs. Impress also includes a spelling checker, a thesaurus, text styles, and background styles.

To use Impress for more than very simple slide shows requires some knowledge of the elements which the slides contain. Slides containing text use styles to determine the appearance of that text. Creating drawings in Impress is similar to the Draw program included in LibreOffice.

When you start Impress for the first time, the Presentation Wizard is shown. Here you can choose from the following options:

- Empty presentation gives you a blank document
- From template is a presentation designed with a template of your choice
- Open existing presentation
- Click Create to open the main Impress window.

If you prefer not to use the Presentation Wizard in future, you can select Do not show this wizard again. You can enable the wizard again later in Tools > Options > LibreOffice Impress > General > New document and select the Start with wizard option.

5.4.1 Main Impress Window

The main Impress window (Fig. 5.15) has three parts: the Slides pane, Workspace, and Sidebar. Additionally, several toolbars can be displayed or hidden during the creation of a presentation.

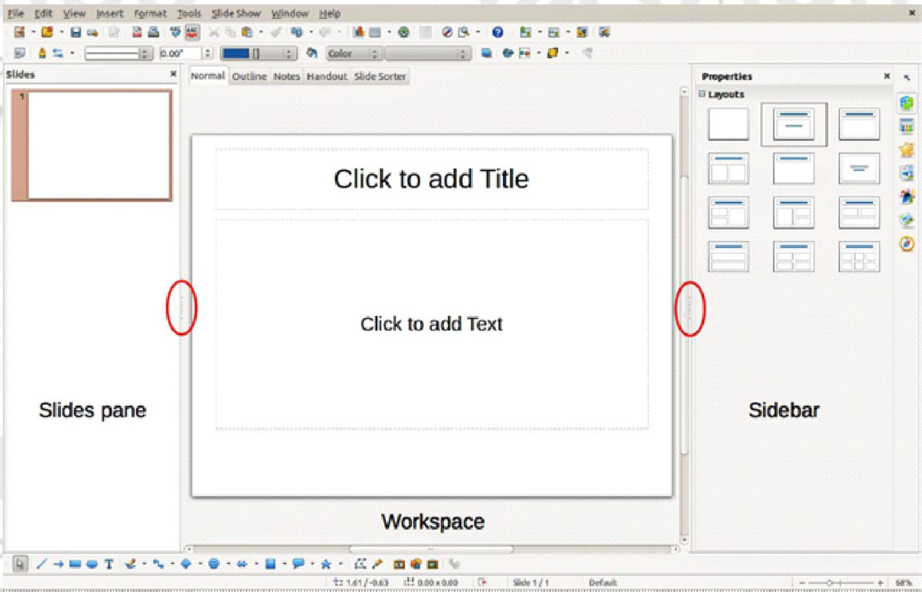


Fig. 5.15: Main window of Impress; ovals indicate the Hide/Show markers

Slides pane

The Slides pane contains thumbnail pictures of the slides in your presentation, in the order they will be shown unless you change the slide show order. Clicking a slide in this pane selects it and places it in the Workspace. When a slide is in the Workspace, you can make changes any way you like.

Several additional operations can be performed on one or more slides simultaneously in the Slides pane:

- Add new slides to the presentation.
- Mark a slide as hidden so that it will not be shown as part of the presentation.
- Delete a slide from the presentation if it is no longer needed.
- Rename a slide.
- Duplicate a slide (copy and paste) or move it to a different position in the presentation (cut and paste).

It is also possible to perform the following operations, although there are more efficient methods than using the Slides pane:

- Change the slide transition following the selected slide or after each slide in a group of slides.
- Change the sequence of slides in the presentation.
- Change the slide design.
- Change slide layout for a group of slides simultaneously.

Sidebar

The Sidebar has seven sections. To expand a section you want to use, click on its icon or click on the small triangle at the top of the icons and select a section from the drop down list. Only one section at a time can be open.

Properties

Shows the layouts included within Impress. You can choose the one you want and use it as it is, or modify it to meet your own requirements. However, it is not possible to save customised layouts.

Master Pages

Here you define the page (slide) style for your presentation. Impress includes several designs of Master Pages (slide masters). One of them – Default – is blank, and the rest have background and styled text.

Custom Animation

A variety of animations can be used to emphasize or enhance different elements of each slide. The Custom Animation section provides an easy way to add, change, or remove animations.

Slide Transition

Provides a number of slide transition options. The default is set to No Transition, in which the following slide simply replaces the existing one. However, many additional transitions are available. You can also specify the transition speed (slow, medium, fast), choose between an automatic or manual transition, and choose how long the selected slide should be shown (automatic transition only).

Styles and Formatting

Here you can edit and apply graphics styles, but you can only edit presentation styles.

When you edit a style, the changes are automatically applied to all of the elements formatted with this style in your presentation. If you want to ensure that the styles on a specific slide are not updated, create a new master page for the slide.

Gallery

Opens the Impress gallery where you can insert an object into your presentation either as a copy or as a link. A copy of an object is independent of the original object. Changes to the original object have no effect on the copy. A link remains dependent on the original object. Changes to the original object are also reflected in the link.

Navigator

Opens the Impress navigator, in which you can quickly move to another slide or select an object on a slide. It is recommended to give slides and objects in your presentation meaningful names so that you can easily identify them when using the navigator.

Workspace

The Workspace (normally in the center of the main window) has five tabs: Normal, Outline, Notes Handout, and Slide Sorter (Fig.5.16). These five tabs are called View buttons. The Workspace below the View buttons changes depending on the chosen view.

Fig.5.16: Workspace tabs

Workspace views

Each of the workspace views is designed to ease the completion of certain tasks; it is therefore useful to familiarize yourself with them in order to quickly accomplish those tasks. Note Each Workspace view displays a different set of toolbars when selected. These toolbar sets can be customised by going to **View > Toolbars** on the Menu bar, then check or uncheck the toolbar you want to add or remove.

Normal view

Normal view is the main view for working with individual slides. Use this view to format and design and to add text, graphics, and animation effects.

Outline view

Outline view (Fig. 5.17) contains all the slides of the presentation in their numbered sequence. It shows topic titles, bulleted lists, and numbered lists for each slide in outline format. Only the text contained in the default text boxes in each slide is shown, so if your slide includes other text boxes or drawing objects, the text in these objects is not displayed. Slide names are also not included.

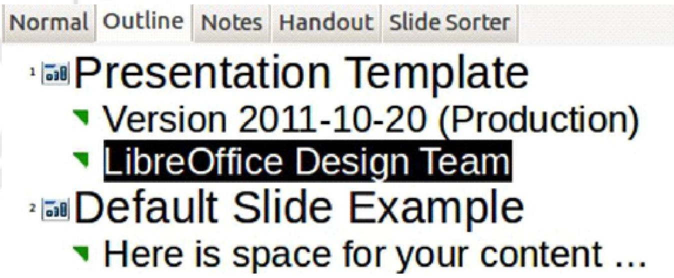


Fig. 5.17: Outline view

Notes view

Use the Notes view (Fig. 5.18) to add notes to a slide. These notes are not seen when the presentation is shown.

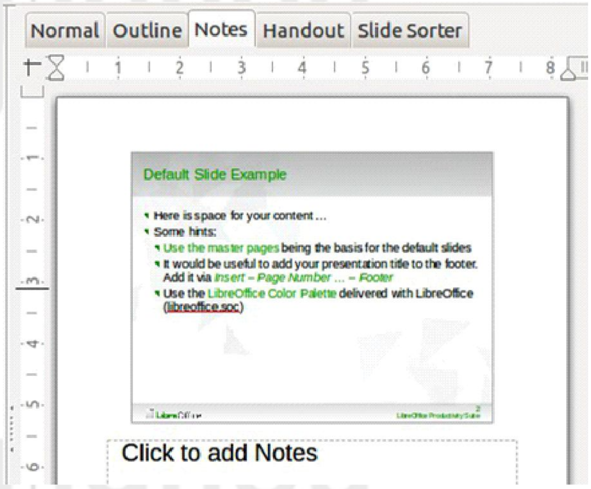


Fig. 5.18: Notes view

1) Click the Notes tab in the Workspace.

2) Select the slide to which you want to add notes.

3) Click the slide in the Slides pane, or double-click the slide name in the Navigator.

4) In the text box below the slide, click on the words Click to add notes and begin typing.

You can resize the Notes text box using the coloured resising handles which appear when you click on the edge of the box. You can also move the box by placing the pointer on the border, then clicking and dragging. To make changes in the text style, press the F11 key to open the Styles and Formatting dialog or click on the Styles and Formatting icon on the Sidebar.

Handout view

Handout view is for setting up the layout of your slide for a printed handout. Click the Handout tab in the workspace and the Layouts section opens on the Sidebar (Fig. 5.19) where you can then choose to print 1, 2, 3, 4, 6, or 9 slides per page. If the Layouts section does not open, then click on the Properties icon at the side of the Sidebar. Use this view also to customize the information printed on the handout.

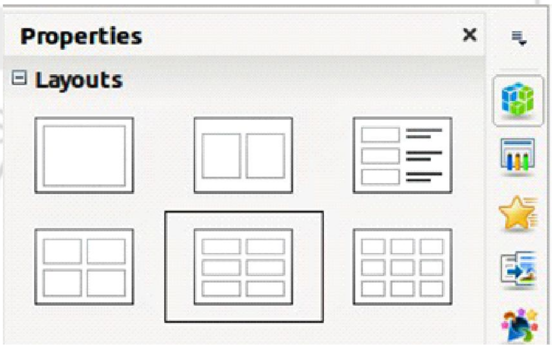


Fig. 5.19: Handout layouts

Slide Sorter view

Slide Sorter view (Fig. 5.20) contains all of the slide thumbnails. Use this view to work with a group of slides or with only one slide.

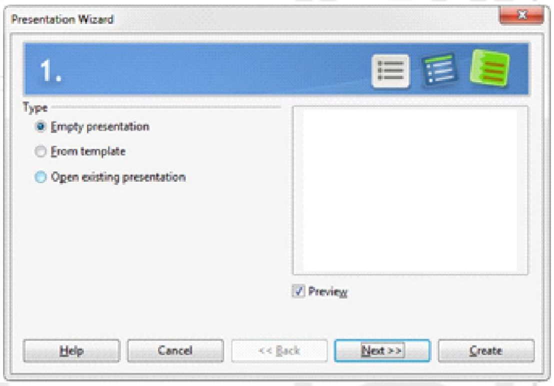


Fig. 5.20: Slide Sorter view

You can work with slides in the Slide Sorter view just as you can in the Slide pane. To make changes, right-click a slide and choose any of the following from the context menu:

- New Slide – adds a new slide after the selected slide.
- Duplicate Slide – creates a duplicate of the selected slide and places the new slide immediately after the selected slide.
- Delete Slide – deletes the selected slide.
- Rename Slide – allows you to rename the selected slide.
- Slide Layout – allows you to change the layout of the selected slide.
- Slide Transition – allows you to change the transition of the selected slide.
  - For one slide, select a slide and add the desired transition.
  - For more than one slide, select a group of slides and add the desired transition.
- Hide Slide – any slides that are hidden are not shown in the slide show.
- Cut – removes the selected slide and saves it to the clipboard.
- Copy – copies the selected slide to the clipboard without removing it.
- Paste – inserts a slide from the clipboard after the selected slide.

5.4.2 Creating a New Presentation

This section describes how to start a new presentation using the Presentation Wizard.

When you start Impress, the Presentation Wizard appears (Fig. 5.21).

1) Under Type, choose one of the options. These options are covered in the Impress Guide.

   – Empty presentation creates a blank presentation.

   – From template uses a template design already created as the basis for a new presentation.

The wizard changes to show a list of available templates. Choose the template you want.

– Open existing presentation continues work on a previously created presentation. The wizard changes to show a list of existing presentations. Choose the presentation you want.

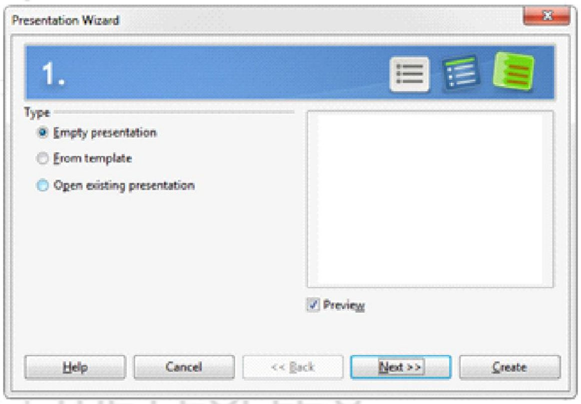


Fig. 5.21: Presentation Wizard

2) Click Next. Fig. 5.22 shows the Presentation Wizard step 2 as it appears if you selected Empty Presentation at step 1. If you selected From template, an example slide is shown in the Preview box.

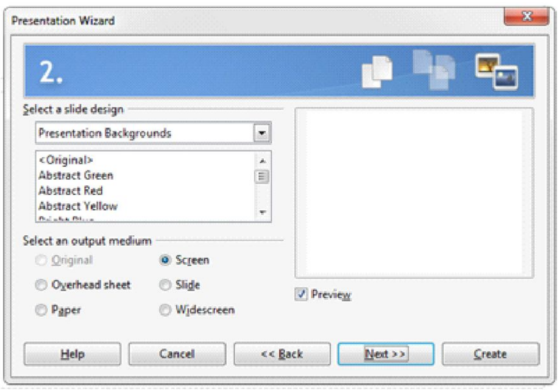


Fig. 5.22: Selecting a slide design

3) Choose a design under Select a slide design. The slide design section gives you two main choices: Presentation Backgrounds and Presentations. Each one has a list of choices for slide designs. If you want to use one of these other than **<Original>**, click it to select it.

– The types of Presentation Backgrounds are shown in Fig. 5.22. When you select a presentation background, you will see a preview of the slide design in the Preview window.

– **<Original>** is for a blank presentation slide design.

4) Select how the presentation will be used under Select an output medium. Majority of presentations are created for computer screen display. It is recommended to select Screen. You can change the page format at any time.

5) Click Next and step 3 of the Presentation Wizard appears (Fig. 5.23).

   – Choose the desired slide transition from the Effect drop-down menu.

   – Select the desired speed for the transition between the different slides in the presentation from the Speed drop-down menu. Medium is a good choice for now.

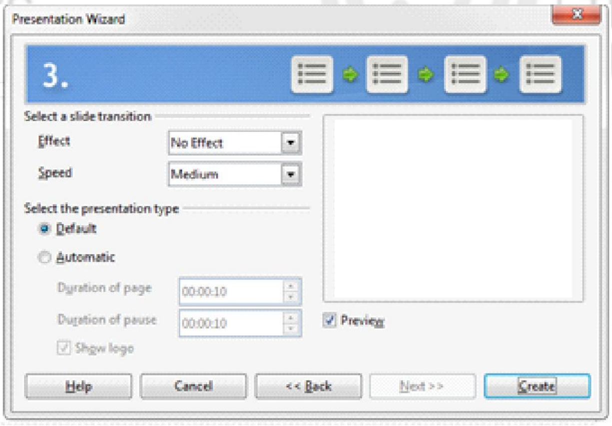


Fig. 5.23: Selecting a slide transition effect

6) Click **Next** and step 4 of the Presentation Wizard appears allowing you to enter information about your company and the presentation you are creating.

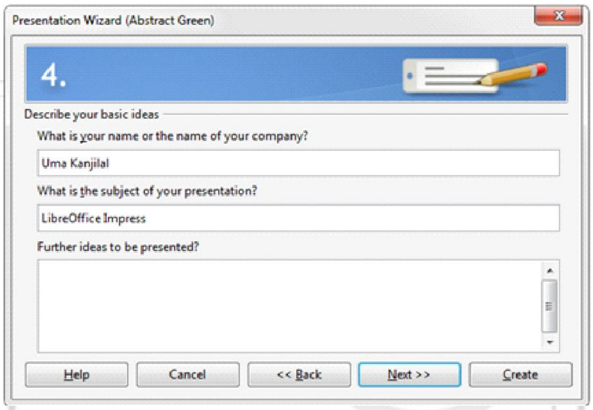


Fig. 5.24: Entering information about your presentation

7) Click Next and step 5 of the Presentation Wizard appears showing a preview of what your presentation will look like (Fig. 5.25). If the preview does not appear, select Preview.

8) If you want to create a summary of your presentation, select **Create summary**.

9) Click **Create** and your new presentation is created.

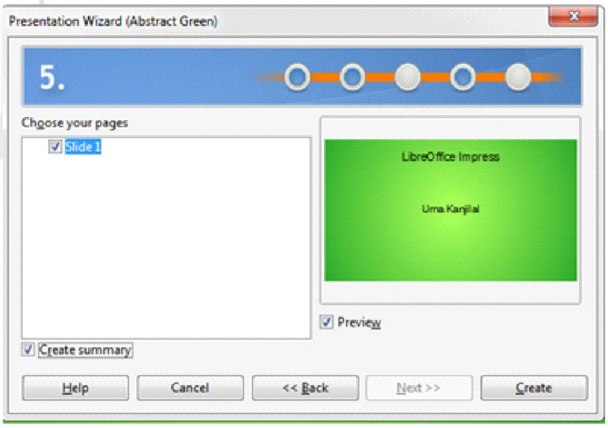


Fig. 5.25: Presentation preview

5.4.3 Formatting a Presentation

A new presentation contains only one empty slide. In this section we will start adding new slides and preparing them for the intended contents.

Inserting slides

New slide

A new slide can be inserted into a presentation as follows:

1) Go to Insert on the Menu bar and select **Slide**.

2) Or right-click on a slide in the **Slides Pane** or Slide Sorter view and select **New Slide** from the context menu.

3) Or, right click in an empty space in the Workspace and select **Slide > New Slide** from the context menu.

4) Or click the Slide icon in the **Presentation toolbar**. If the Presentation toolbar is not visible, go to **View > Toolbars** on the Menu bar and select Presentation from the list.

5) A new slide is inserted after the selected slide in the presentation.

Duplicate slide

Sometimes, rather than starting from a new slide you may want to duplicate a slide already included in your presentation. To duplicate a slide:

1) Select the slide you want to duplicate from the **Slides Pane**.

2) Go to Insert on the **Menu bar** and select Duplicate Slide.

3) Or, right click on the slide in the **Slides Pane** or **Slide Sorter** view and select **Duplicate Slide** from the context menu.

4) Or, right click on a slide in the Workspace and select Slide > Duplicate Slide from the context menu.

5) Or click on the triangle to the right of the Slide icon in the **Presentation toolbar** and select **Duplicate Slide** from the context menu. If the Presentation toolbar is not visible, go to **View > Toolbars** on the Menu bar and select Presentation from the list.

6) A duplicate slide is inserted after the selected slide in the presentation.

Selecting slide layout

When creating a presentation, the first slide is normally a title slide. You can use either a blank layout or one of the title layouts as your title slide. Click on the Properties icon at the side of the Sidebar to open Layouts section and display the available layouts (Fig. 5.26). The layouts included in LibreOffice range from a blank slide to a slide with six contents boxes and a title.

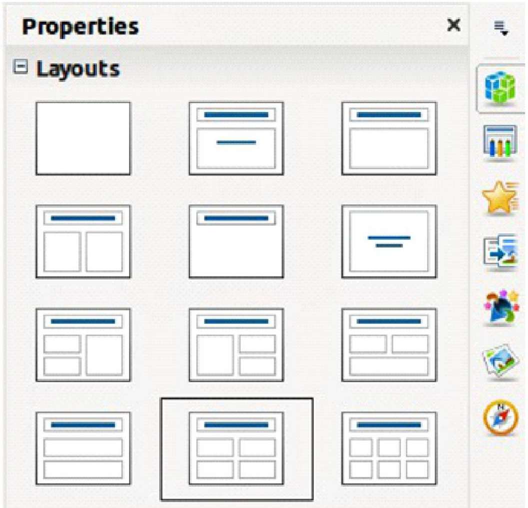


Fig. 5.26: Available slide layouts

To create a title, if one of the title layouts has been selected, click on Click to add title and then type the title text. To add text content, depending on the slide layout selected, click on Click to add text.

To select or change the layout of a slide, select the slide in the Slides Pane so that it appears in the Workspace and select the desired layout from the Layouts section in the Sidebar. Several layouts contain one or more content boxes. Each of these content boxes can be configured to contain text, movies, images, charts or tables. You can choose the type of contents by clicking on the corresponding icon that is displayed in the middle of the contents box as shown in Fig.5.27. If you intend to use the contents box for text, click on Click to add text.

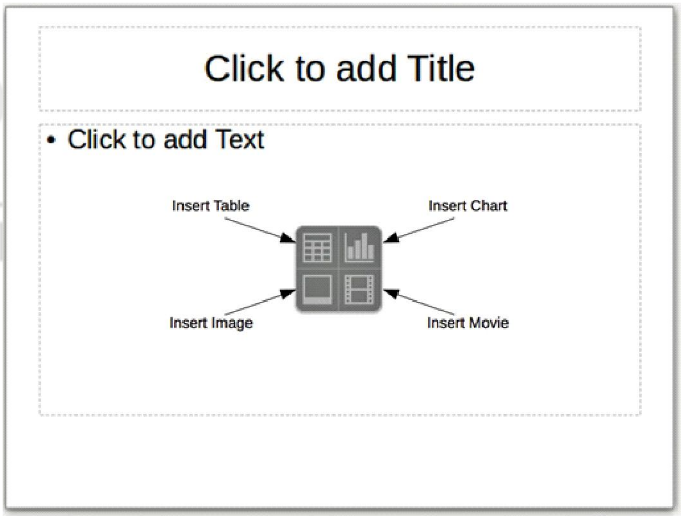


Fig.5.27: Selecting contents type

Modifying slide elements

A slide contains elements that were included in the slide master, as well as those elements included in the selected slide layout. However, it is unlikely that the predefined layouts will suit all your needs for your presentation. You may want to remove elements that are not required or insert objects such as text and graphics.

Although Impress does not have the functionality to create new layouts, it allows you to resize and move the layout elements. It is also possible to add elements without being limited to the size and position of the layout boxes.

To resize a contents box, click on the outer frame so that the resising handles are displayed. To move it, place the mouse cursor on the frame so that the cursor changes shape. You can now click and drag the contents box to a new position on the slide.

To remove any unwanted elements:

1) Click the element to highlight it. The resising handles show it is selected.

2) Press the Delete key to remove it.

Adding text

To add text to a slide that contains a text frame, click on Click to add text in the text frame and then type your text. The Outline styles are automatically applied to the text as you insert it. You can change the outline level of each paragraph as well as its position within the text by using the arrow buttons on the Text Formatting toolbar.

Adding objects

To add any objects to a slide, for example a picture, clipart, drawing, photograph, or spreadsheet, click on Insert then select from the drop down menu what type of object you want to insert.

Modifying appearance of all slides

To change the background and other characteristics of all slides in the presentation, you need to modify the master page or choose a different master page. A Slide Master is a slide with a specified set of characteristics that acts as a template and is used as the starting point for creating other slides. These characteristics include slide background, objects in the background, formatting of any text used, and any background graphics.

5.4.4 Setting up a Slide Show

Impress allocates reasonable default settings for slide shows, while at the same time allowing for customising many aspects of the slide show experience.

Most of the tasks are best done in Slide Sorter view where you can see most of the slides simultaneously. Go to **View > Slide Sorter** on the Menu bar or click the **Slide Sorter** tab at the top of the Workspace.

One slide set – multiple presentations

In many situations, you may find that you have more slides than the time available to present them or you may want to provide a rapid overview without dwelling on the details. Rather than having to create a new presentation, you can use two tools that Impress offers: hiding slides and custom slide shows.

Hiding slides

1) Select the slide you want to hide in the Slide Pane or Slide Sorter view on the Workspace area.

2) Go to **Slide Show > Hide Slide** on the **Menu bar** or right-click on the slide thumbnail and select Hide Slide from the context menu. Hidden slides are marked by a diagonal bars across the slide.

Custom slide shows

If you want to create a custom slide show from the same presentation:

1) Select the slides you want to use in your custom slide show.

2) Go to **Slide Show > Custom Slide Show** on the Menu bar.

3) Click on the **New button** to create a new sequence of slides and save it with a different name. You can have as many slide shows as you want from a single presentation.

Slide transitions

Slide transition is the animation that is played when a slide is changed for the next slide in your presentation. You can configure the slide transition from the Slide Transition section in the Tasks Pane.

1) Go to **Slide Show > Slide Transition** on the **Menu bar** or click on the **Slide Transition icon** on the Sidebar to open the options available for slide transitions.

2) Select the desired transition, the speed of the animation, and whether the transition should happen when you click the mouse (preferred) or automatically after a certain number of seconds.

Running a slide show

To run a slide show, do one of the following:

- Click **Slide Show > Start** from first Slide on the Menu bar.
- Click the Start from first Slide icon on the Presentation toolbar.
- Press F5 on the keyboard.

If the slide advance is set to *Automatically after X sec*, let the slide show run by itself. If the slide advance is set to *On mouse click*, do one of the following to move from one slide to the next:

- Use the arrow keys on the keyboard to go to the next slide or to go back to the previous one.
- Click the mouse to move to the next slide.
- Press the spacebar on the keyboard to advance to the next slide.

Right-click anywhere on the screen to open a context menu where you can navigate through the slides and set other options.

To exit the slide show at any time including when the slide show has ended, press the Esc key.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

2) Discuss the steps in running a slide show.

....

....

....

....

5.5 SPREADSHEETS WITH LIBREOFFICE: CALC

Calc is the spreadsheet component of LibreOffice. You can enter data (usually numerical) in a spreadsheet and then manipulate this data to produce certain results.

Other features provided by Calc include:

- Functions, which can be used to create formulas to perform complex calculations on data.
- Database functions, to arrange, store, and filter data.
- Dynamic charts; a wide range of 2D and 3D charts.
- Macros, for recording and executing repetitive tasks; scripting languages supported Include LibreOffice Basic, Python, BeanShell, and JavaScript.
- Ability to open, edit, and save Microsoft Excel spreadsheets.
- Import and export of spreadsheets in multiple formats, including HTML, CSV, PDF, and PostScript.

Calc works with elements called spreadsheets. Spreadsheets consist of a number of individual sheets, each sheet containing cells arranged in rows and columns. A particular cell is identified by its row number and column letter. Cells hold the individual elements – text, numbers, formulas, and so on – that make up the data to display and manipulate. Each spreadsheet can have several sheets, and each sheet can have several individual cells. In Calc, each sheet can have a maximum of 1,048,576 rows (65,536 rows in Calc 3.2 and earlier) and a maximum of 1024 columns.

5.5.1 Calc Main Window

When Calc is started, the main window opens as shown in Fig. 5.28.

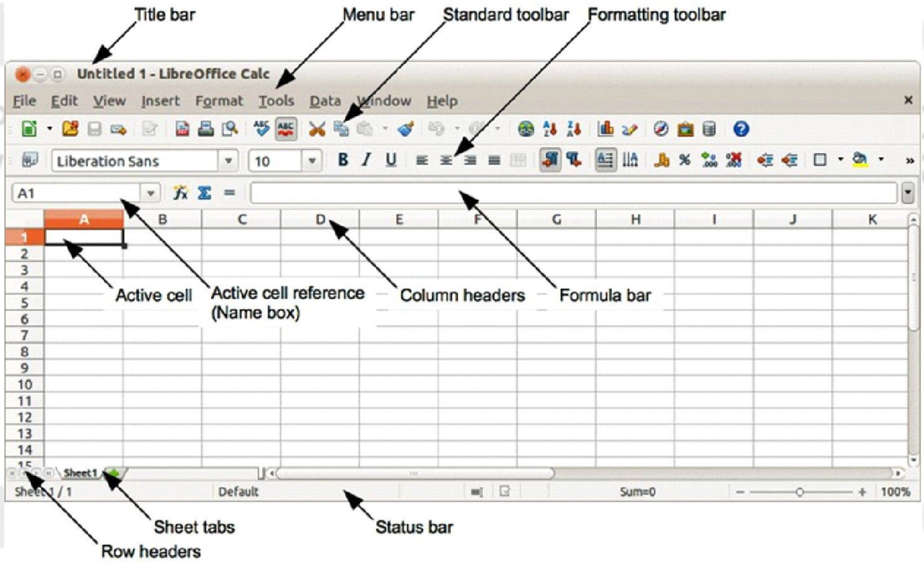


Fig. 5.28: Calc main dialog, without Sidebar

Formula bar is a special component of the Calc which is not common with other LibreOffice applications is described below.

Formula bar

The Formula Bar is located at the top of the sheet in the Calc workspace. The Formula Bar is permanently docked in this position and cannot be used as a floating toolbar. If the Formula Bar is not visible, go to **View** on the Menu bar and select **Formula Bar**.

Fig. 5.29: Formula bar

Going from left to right and referring to Figure 2, the Formula Bar consists of the following:

- **Name Box** – gives the cell reference using a combination of a letter and number, for example A1. The letter indicates the column and the number indicates the row of the selected cell.
- **Function Wizard** – opens a dialog from which you can search through a list of available functions. This can be very useful because it also shows how the functions are formatted.
- **Sum** – clicking on the Sum icon totals the numbers in the cells above the selected cell and then places the total in the selected cell. If there are no numbers above the selected cell, then the cells to the left are totalled.

Function = – clicking on the Function icon inserts an equals (=) sign into the selected cell and the **Input line**, allowing a formula to be entered.

- **Input line** – displays the contents of the selected cell (data, formula, or function) and allows you to edit the cell contents.
- You can also edit the contents of a cell directly in the cell itself by double-clicking on the cell. When you enter new data into a cell, the Sum and Function icons change to **Cancel** and **Accept** icons .

5.5.2 Spreadsheet Layout

Individual cells

The main section of the workspace in Calc displays the cells in the form of a grid. Each cell is formed by the intersection of the columns and rows in the spreadsheet.

At the top of the columns and the left end of the rows are a series of header boxes containing letters and numbers. The column headers use an alpha character starting at A and go on to the right. The row headers use a numerical character starting at 1 and go down.

These column and row headers form the cell references that appear in the Name Box on the Formula Bar (Figure 2). If the headers are not visible on your spreadsheet, go to **View** on the Menu bar and select **Column & Row Headers**.

Sheet tabs

In Calc you can have more than one sheet in a spreadsheet. At the bottom of the grid of cells in a spreadsheet are sheet tabs indicating how many sheets are there in your spreadsheet. Clicking on a tab enables access to each individual sheet and displays that sheet. An active sheet is indicated with a white tab (default Calc setup). You can also select multiple sheets by holding down the *Ctrl* key while you click on the sheet tabs.

5.5.3 Opening a CSV File

Comma-separated-values (CSV) files are spreadsheet files in a text format where cell contents are separated by a character, for example a comma or semi-colon. Each line in a CSV text file represents a row in a spreadsheet. Text is entered between quotation marks; numbers are entered without quotation marks.

To open a CSV file in Calc:

1) Choose **File > Open** on the Menu bar and locate the CSV file that you want to open.

2) Select the file and click **Open**. By default, a CSV file has the extension .csv. However, some CSV files may have a .txt extension.

3) The **Text Import** dialog (Fig. 5.30) opens allowing you to select the various options available when importing a CSV file into a Calc spreadsheet.

4) Click **OK** to open and import the file.

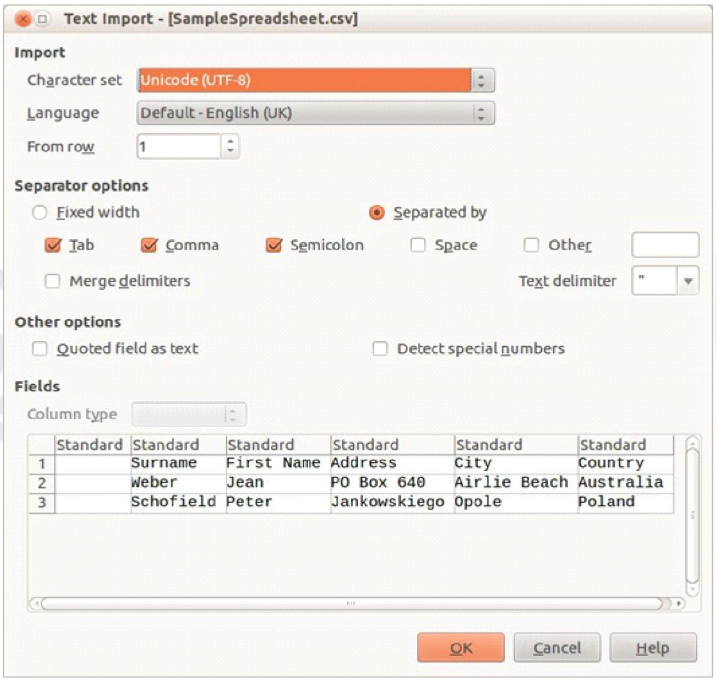


Fig. 5.30: Text Import dialog

The various options for importing CSV files into a Calc spreadsheet are as follows:

- **Import**

– *Character Set* – specifies the character set to be used in the imported file.

– *Language* – determines how the number strings are imported.

If Language is set to Default for CSV import, Calc will use the globally set language. If Language is set to a specific language, that language will be used when importing numbers.

– *From Row* – specifies the row where you want to start the import. The rows are visible in the preview window at the bottom of the dialog.

- **Separator Options** – specifies whether your data uses separators or fixed widths as delimiters.

– *Fixed width* – separates fixed-width data (equal number of characters) into columns. Click on the ruler in the preview window to set the width.

– *Separated by* – select the separator used in your data to delimit the data into columns.

When you select *Other*, you specify the character used to separate data into columns. This custom separator must also be contained in your data.

– *Merge delimiters* – combines consecutive delimiters and removes blank data fields.

– *Text delimiter* – select a character to delimit text data.

- **Other options**

– *Quoted fields as text* – when this option is enabled, fields or cells whose values are quoted in their entirety (the first and last characters of the value equal the text delimiter) are imported as text.

– *Detect special numbers* – when this option is enabled, Calc will automatically detect all number formats, including special number formats such as dates, time, and scientific notation.

The selected language also influences how such special numbers are detected, since different languages and regions many have different conventions for such special numbers.

When this option is disabled, Calc will detect and convert only decimal numbers. The rest, including numbers formatted in scientific notation, will be imported as text. A decimal number string can have digits 0-9, thousands separators, and a decimal separator. Thousands separators and decimal separators may vary with the selected language and region.

- **Fields** – shows how your data will look when it is separated into columns.

– *Column type* – select a column in the preview window and select the data type to be applied the imported data.

– *Standard* – Calc determines the type of data.

– *Text* – imported data are treated as text.

– *US English* – numbers formatted in US English are searched for and included regardless of the system language. A number format is not applied. If there are no US English entries, the *Standard* format is applied.

– *Hide* – the data in the column are not imported.

5.5.4 Navigating within Spreadsheets

Calc provides many ways to navigate within a spreadsheet from cell to cell and sheet to sheet. You can generally use the method you prefer.

Cell navigation

When a cell is selected or in focus, the cell borders are emphasised. When a group of cells is selected, the cell area is coloured. The colour of the cell border emphasis and the colour of a group of selected cells depend on the operating system being used and how you have set up LibreOffice.

- **Using the mouse** – place the mouse pointer over the cell and click the left mouse button.

  To move the focus to another cell using the mouse, simply move the mouse pointer to the cell where you want the focus to be and click the left mouse button.

- **Using a cell reference** – highlight or delete the existing cell reference in the Name Box on the Formula Bar (Fig. 5.29). Type the new cell reference of the cell you want to move to and press *Enter* key. Cell references are case-insensitive: for example, typing either a3 or A3 will move the focus to cell A3.

Using the Navigator – click on the **Navigator** icon  on the Standard toolbar or press the *F5* key to open the **Navigator** dialog (Fig. 5.31) or click the **Navigator Tab** button in the open Sidebar. Type the cell reference into the Column and Row fields and press the *Enter* key.

- **Using the Enter key** – pressing *Enter* moves the cell focus down in a column to the next row. Pressing *Shift+Enter* moves the focus up in a column to the next row.

- **Using the Tab key** – pressing Tab moves the cell focus right in a row to the next column.

Pressing *Shift+Tab* moves the focus to the left in a row to the next column.

- **Using the arrow keys** – pressing the arrow keys on the keyboard moves the cell focus in the direction of the arrow pressed.

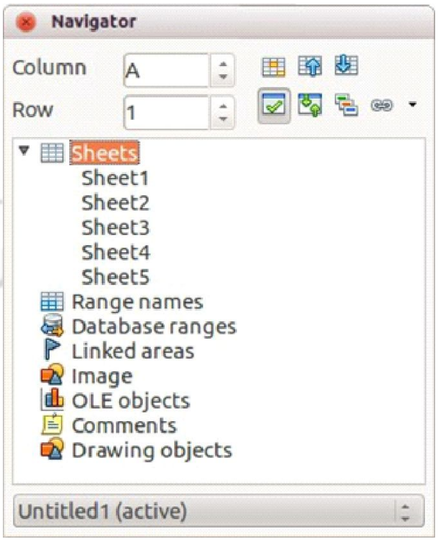


Fig. 5.31: Navigator dialog in Calc

- **Using Home**, **End**, **Page Up** and **Page Down**
  - *Home* moves the cell focus to the start of a row.
  - *End* moves the cell focus to the last cell on the right in the row that contains data.
  - *Page Down* moves the cell focus down one complete screen display.
  - *Page Up* moves the cell focus up one complete screen display.

Sheet navigation

Each sheet in a spreadsheet is independent of the other sheets in a spreadsheet, though references can be linked from one sheet to another sheet. There are three ways to navigate between different sheets in a spreadsheet.

- **Using the Navigator** – when the Navigator is open (Fig. 5.31), double-clicking on any of the listed sheets selects the sheet.
- **Using the keyboard** – using key combinations *Ctrl+Page Down* moves one sheet to the right and *Ctrl+Page Up* moves one sheet to the left.
- **Using the mouse** – clicking on one of the sheet tabs at the bottom of the spreadsheet selects that sheet.

If your spreadsheet contains a lot of sheets, then some of the sheet tabs may be hidden behind the horizontal scroll bar at the bottom of the screen. If this is the case:

- Using the four buttons to the left of the sheet tabs can move the tabs into view (Fig. 5.32).
- Dragging the scroll bar edge to the right may reveal all the tabs.
- Right-clicking on any of the arrows opens a context menu where you can select a sheet (see Fig. 5.33).

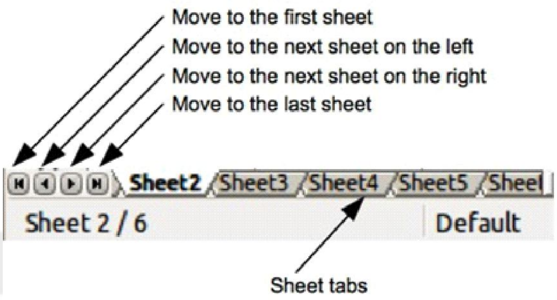


Fig. 5.32: Navigating sheet tabs

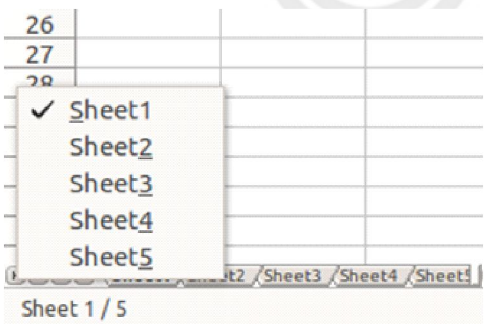


Fig. 5.33: Right-click any arrow button

5.5.5 Working with Columns and Rows

Inserting columns and rows

When you insert a column, it is inserted to the *left* of the highlighted column. When you insert a row, it is inserted *above* the highlighted row. When you insert columns or rows, the cells take the formatting of the corresponding cells in the next column to left or the row above.

Single column or row

Using the **Insert** menu:

1) Select a cell, column, or row where you want the new column or row inserted.

2) Go to **Insert** on the Menu bar and select either **Insert > Columns** or **Insert > Rows.**

Using the mouse:

1) Select a column or row where you want the new column or row inserted.

2) Right-click the column or row header.

3) Select **Insert Columns** or **Insert Rows** from the context menu.

Multiple columns or rows

Multiple columns or rows can be inserted at once rather than inserting them one at a time.

1) Highlight the required number of columns or rows by holding down the left mouse button on the first one and then dragging across the required number of identifiers.

2) Proceed as for inserting a single column or row above.

Deleting columns and rows

Single column or row

To delete a single column or row:

1) Select a cell in the column or row you want to delete.

2) Go to **Edit** on the Menu bar and select **Delete Cells** or right-click and select **Delete** from the context menu.

3) Select the option you require from the **Delete Cells** dialog (Fig. 5.34).

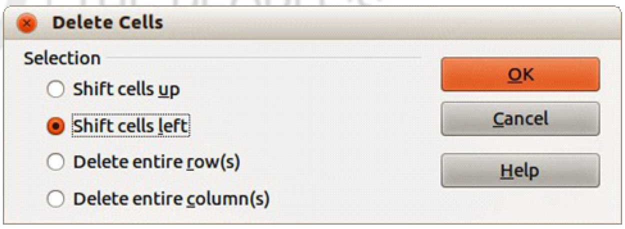


Fig. 5.34: Delete Cells dialog

Alternatively:

1) Click in the column or header to select the column or row.

2) Go to **Edit** on the Menu bar and select **Delete Cells** or right-click and select **Delete Columns** or **Delete Rows** from the context menu.

Multiple columns or rows

To delete multiple columns or rows:

1) Select the columns or rows

2) Go to **Edit** on the Menu bar and select **Delete Cells** or right-click and select **Delete Columns** or **Delete Rows** from the context menu.

5.5.6 Working with Sheets

Click on the **Add Sheet** icon . This inserts a new sheet after the last sheet in the spreadsheet without opening the **Insert Sheet** dialog. You can move or copy sheets within the same spreadsheet by dragging and dropping or using the **Move/Copy Sheet** dialog. To move or copy a sheet into a different spreadsheet, you have to use the **Move/Copy Sheet** dialog.

Using a dialog

Use the **Move/Copy Sheet** dialog (Fig. 5.35) to specify exactly whether you want the sheet in the same or a different spreadsheet, its position within the spreadsheet, the sheet name when you move or copy the sheet.

1) In the current document, right-click on the sheet tab you wish to move or copy and select **Move/Copy Sheet** from the context menu or go to **Edit > Sheet > Move/Copy** on the Menu bar.

2) Select **Move** to move the sheet or **Copy** to copy the sheet.

3) Select the spreadsheet where you want the sheet to be placed from the drop-down list in **To document**. This can be the same spreadsheet, another spreadsheet already open, or you can create a new spreadsheet.

4) Select the position in **Insert before** where you want to place the sheet.

5) Type a name in the **New name** text box if you want to rename the sheet when it is moved or copied. If you do not enter a name, Calc creates a default name (Sheet 1, Sheet 2, and so on).

6) Click **OK** to confirm the move or copy and close the dialog.

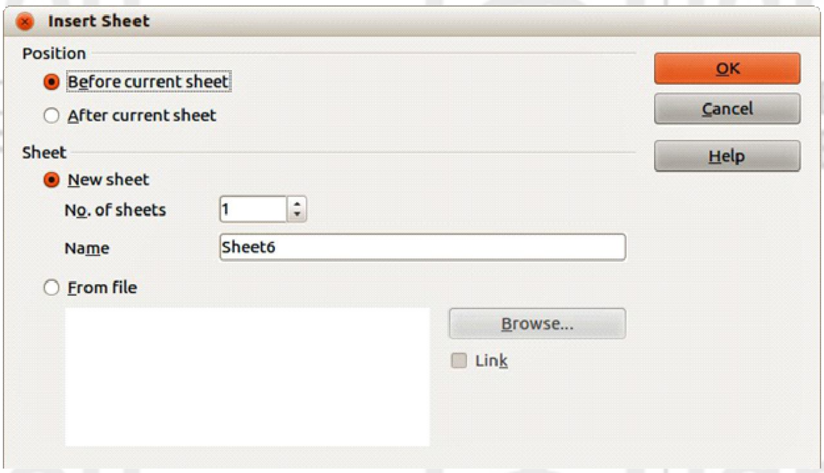


Fig. 5.35: Insert Sheet dialog

Deleting sheets

To delete a single sheet, right-click on the sheet tab you want to delete and select **Delete Sheet** from the context menu, or go to **Edit > Sheet > Delete** from on the Menu bar. Click **Yes** to confirm the deletion.

5.5.7 Working with a Spreadsheet

Entering Numeric Data

Click in the cell and type in the number using the number keys on either the main keyboard or the numeric keypad. To enter a negative number, either type a minus (–) sign in front of it or enclose it in parentheses (brackets), like this: **(1234)**. By default, numbers are right-aligned and negative numbers have a leading minus symbol.

Entering Text

Click in the cell and type the text. Text is left-aligned by default. If a number is entered in the format *01481*, Calc will drop the leading 0. To preserve the leading zero, for example for telephone area codes, type an apostrophe before the number, like this: **'01481**. The data is now treated as text and displayed exactly as entered.

Using the Fill Tool on cells

At its simplest, the Fill tool is a way to duplicate existing content. Start by selecting the cell to copy, then drag the mouse in any direction (or hold down the Shift key and click in the last cell you want to fill), and then choose Edit > Fill and the direction in which you want to copy: Up, Down, Left or Right. A more complex use of the Fill tool is to add a fill series to a spreadsheet, select the cells to fill, choose Edit > Fill > Series.

Formatting Numbers

Several different number formats can be applied to cells by using icons on the Formatting toolbar. Select the cell, then click the relevant icon. Some icons may not be visible in a default setup; click the down-arrow at the end of the Formatting bar and select other icons to display.

5.5.8 Sorting Records

Sorting within Calc arranges the cells in a sheet using the sort criteria that you specify. Several criteria can be used and a sort applies each criteria consecutively. Sorts are useful when you are searching for a particular item and become even more useful after you have filtered data.

Also, sorting is useful when you add new information to your spreadsheet. When a spreadsheet is long, it is usually easier to add new information at the bottom of the sheet, rather than adding rows in their correct place. After you have added information, you then carry out a sort to update the spreadsheet.

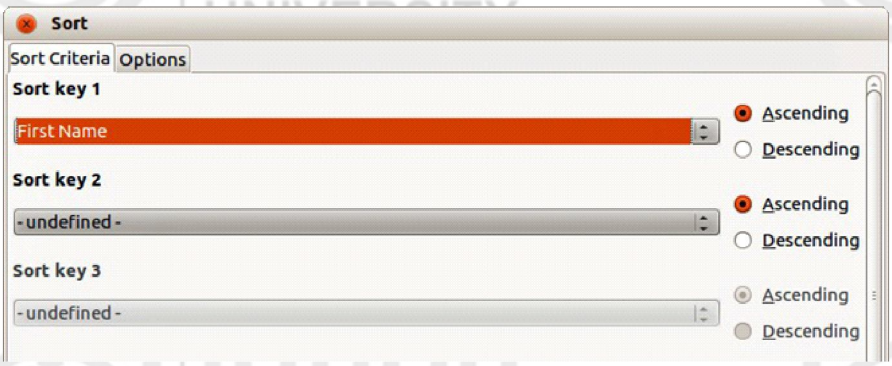


Fig. 5.36: Sort Criteria dialog

To sort cells in your spreadsheet:

1) Select the cells to be sorted.

2) Go to **Data > Sort** on the Menu bar to open the **Sort** dialog (Fig. 5.36).

3) Select the sort criteria from the drop down lists. The selected lists are populated from the selected cells.

4) Select either ascending order (A-Z, 1-9) or descending order (Z-A, 9-1).

5) Click **OK** and the sort is carried out on your spreadsheet.

5.5.9 Using Formulas and Functions

You may need more than numbers and text on your spreadsheet. Often the contents of one cell depend on the contents of other cells. Formulas are equations that use numbers and variables to produce a result. Variables are placed in cells to hold data required equations.

A function is a predefined calculation entered in a cell to help you analyze or manipulate data. All you have to do is enter the arguments and the calculation is automatically made for you. Functions help you create the formulas required to get the results that you are looking for.

5.5.10 Analysing Data

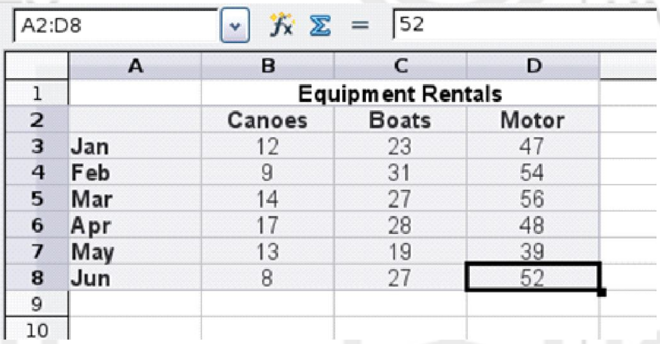
Calc includes several tools to help you analyze the information in your spreadsheets, ranging from features for copying and reusing data, to creating subtotals automatically, to varying information to help you find the answers you need. These tools are divided between the Tools and Data menus.

One of the most useful of these tools is the PivotTable, which is used for combining, comparing, and analysing large amounts of data easily. Using the PivotTable, you can view different summaries of the source data, display the details of areas of interest, and create reports, whether you are a beginner, an intermediate or advanced user.

5.5.11 Creating a Chart

Chart is a useful tool for quantitative analysis and representation of a data series in a performance report or a presentation. It is also visually appealing to the laymen for understanding various data elements.

To create a chart, you will first highlight (select) the data to be included in the chart. The selection does not need to be in a single block, as shown in Fig. 5.37; you can also choose individual cells or groups of cells (columns or rows).



A2:D8 | 52

	A	B	C	D
1		Equipment Rentals		
2		Canoes	Boats	Motor
3	Jan	12	23	47
4	Feb	9	31	54
5	Mar	14	27	56
6	Apr	17	28	48
7	May	13	19	39
8	Jun	8	27	52
9				
10				

Fig. 5.37: Selecting Data for Plotting

Next, open the Chart Wizard dialog using one of two methods.

- Choose **Insert > Chart** from the menu bar.
- Or, click the **Chart** icon on the main toolbar.

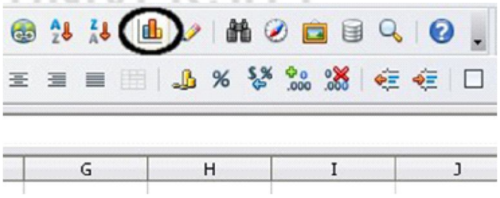


Fig. 5.38: Insert Chart from Main Toolbar

Either method inserts a sample chart on the worksheet, opens the Formatting toolbar, and opens the Chart Wizard, as shown in Fig. 5.39.

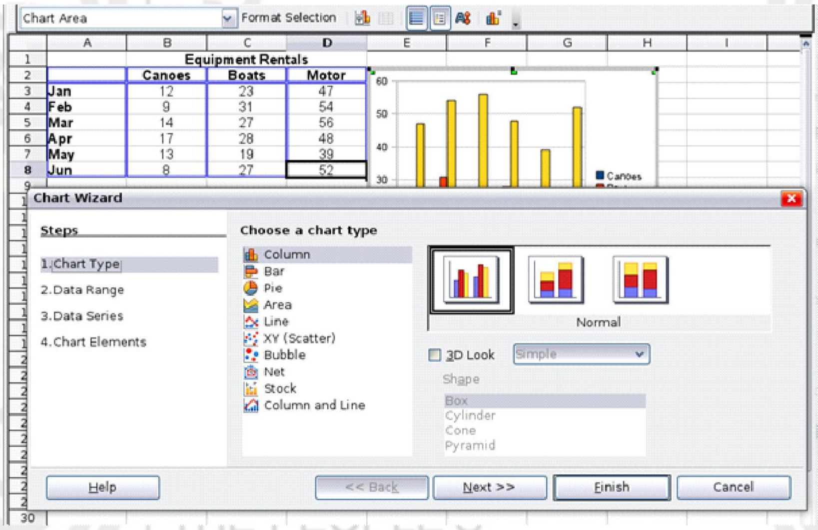


Fig. 5.39: Chart Wizard

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) Explain steps in sorting data in Calc.

....

....

....

....

5.6 SUMMARY

In this Unit you have been introduced to an open source office productivity tool LibreOffice. The special features of LibreOffice have been discussed in this Unit. In the section on Word processing in with Writer you were introduced to working with text, formatting using mail merge, tracking changes and linking to other parts of a document. In Impress you learned about the structure, creating and formatting presentation and setting up a slide show. Finally in the Calc section you were introduced to Calc main window and the spreadsheet layout. Importing data in CSV format, navigating within spreadsheet, working with columns and rows as well as spreadsheets, sorting records, using formula and functions, analysing data and graphical presentation of data in Calc has also been covered. The Unit has been presented in a manual format to help you in getting started with LibreOffice.

5.7 ANSWERS TO SELF CHECK EXERCISES

1) Writer provides the following important features:

   - Templates and styles.
   - Page layout methods, including frames, columns, and tables.
   - Embedding or linking of graphics, spreadsheets, and other objects.
   - Built-in drawing tools.
   - Master documents—to group a collection of documents into a single document
   - Change tracking during revisions.
   - Database integration, including a bibliography database.
   - Export to PDF, including bookmarks.

2) To run a slide show, do one of the following:

   - Click **Slide Show > Start** from first Slide on the Menu bar.
   - Click the Start from first Slide icon on the Presentation toolbar.
   - Press F5 on the keyboard.

If the slide advance is set to *Automatically after X sec*, let the slide show run by itself. If the slide advance is set to *On mouse click*, do one of the following to move from one slide to the next:

- Use the arrow keys on the keyboard to go to the next slide or to go back to the previous one.
- Click the mouse to move to the next slide.
- Press the spacebar on the keyboard to advance to the next slide.

Right-click anywhere on the screen to open a context menu where you can navigate through the slides and set other options.

3) To sort cells in your spreadsheet:

   a) Select the cells to be sorted.

   b) Go to **Data > Sort** on the Menu bar to open the **Sort** dialog.

   c) Select the sort criteria from the drop down lists. The selected lists are populated from the selected cells.

   d) Select either ascending order (A-Z, 1-9) or descending order (Z-A, 9-1).

   e) Click **OK** and the sort is carried out on your spreadsheet.

5.8 KEYWORDS

BeanShell : Java-like scripting language, invented by Patrick Niemeyer. It runs in the Java Runtime Environment (JRE) and uses Java syntax, in addition to scripting commands and syntax.

Complex Text Layout : Typesetting of writing systems in which the shape or positioning of a grapheme depends on its relation to other graphemes. The term is used in the field of software internationalization, where each grapheme is a character.

CSV : A common, relatively simple file format that is widely supported by consumer, business, and scientific applications.

JavaScript : An object-oriented computer programming language commonly used to create interactive effects within web browsers.

Open source software : Software for which the original source code is made freely available and may be redistributed and modified.

Python : A high-level general-purpose programming language.

5.9 REFERENCES AND FURTHER READING

The Document Foundation. Introducing LibreOffice *In* LibreOffice 4.2 Getting Started Guide, Chapter 1 <https://wiki.documentfoundation.org/images/2/20/GS4201-IntroducingLibreOffice.pdf>

The Document Foundation. Setting up LibreOffice *In* LibreOffice 4.2 Getting Started Guide, Chapter 2 <https://wiki.documentfoundation.org/images/8/8b/GS4202-SettingUpLibreOffice.pdf>

The Document Foundation. Using Styles and Templates *In* LibreOffice 4.2 Getting Started Guide, Chapter 3 < https://wiki.documentfoundation.org/images/6/6b/GS4203-StylesAndTemplates.pdf>

The Document Foundation. Getting Started with Writer *In* LibreOffice 4.2 Getting Started Guide, Chapter 4 <https://wiki.documentfoundation.org/images/1/11/GS4204-GettingStartedWithWriter.pdf >

The Document Foundation. Getting Started with Calc *In* LibreOffice 4.2 Getting Started Guide, Chapter 5 <https://wiki.documentfoundation.org/images/c/c0/GS4205-GettingStartedWithCalc.pdf>

The Document Foundation. Getting Started with Impress *In* LibreOffice 4.2 Getting Started Guide, Chapter 6 <https://wiki.documentfoundation.org/images/e/ea/GS4206-GettingStartedWithImpress.pdf>

https://doc.opensuse.org/documentation/html/openSUSE_121/opensuse-startup/art.oofficequick.html

UNIT 6 DATABASE MANAGEMENT SYSTEMS

Structure

6.0 OBJECTIVES

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

- appreciate the limitations of the traditional approach to application system development;
- give reasons why the database approach is now being increasingly adopted;
- discuss different views of data;
- list the components of a database management system;
- enumerate the features and capabilities of a database management system; and
- discuss the security threats for the DBMS and their counter measures.

6.1 INTRODUCTION

We all live in an age where the world is full of data and information. Everyone is aware of Internet, which has become a huge source of information and also growing everyday. The amount of data is actually very large to access and maintain.

For example, if we go to the bank to deposit or withdraw funds, if we make booking in a hotel or railways reservation, even purchasing items from a supermarket or Mall, all these activities involves an automatic update of the database which keeps the inventory of the store room or showroom items.

Relational database systems have become increasingly popular since the late 1970's. They offer a powerful method for storing data in an application-independent manner. This means that for many enterprises the database is at the core of the IT strategy. Developments can progress around a relatively stable database structure which is secure, reliable, efficient, and transparent.

In early systems, each suite of application programs had its own independent master file. The duplication of data over master files could lead to inconsistent data. Efforts to use a common master file for a number of application programs resulted in problems of integrity and security. The production of new application programs could require amendments to existing application programs, resulting in *unproductive maintenance*.

Data structuring techniques, developed to exploit random access storage devices, increased the complexity of the insert, delete and update operations on data. As a first step towards a DBMS, packages of subroutines were introduced to reduce programmer effort in maintaining these data structures. However, the use of these packages still requires knowledge of the physical organisation of the data. In the traditional database applications, most of the information was stored and accessed in either textual or numeric form. But nowadays, video clips, pictures files, weather data, satellite images and sound image files are stored in separate and specialised databases such as multimedia databases.

In this unit we will introduce basic concepts of file oriented approach, drawbacks with file oriented approach, database approach, use of DBMS, levels of abstraction of DBMS, types of DBMS and security features.

6.2 FILE ORIENTED APPROACH

A file based system is a collection of application programs that perform services for the users wishing to access information. For each application a separate master file and its own set of personal files were present. Each program within a file based system defines and manages its own data. Because of this, there are certain limits as to how that data can be used or transported. File based systems were developed as better alternatives to paper based filing systems. By having files stored on computers, the data could be accessed more efficiently. It was common practice for larger companies to have each of its departments looking after its own data. The problems that arise with this type of file based system are listed below:

- Data separation and isolation
- Limited Data Sharing – No centralised control of data.
- Program Data dependence – All programs maintain metadata for each file they use.
- Data Redundancy (Duplication of data) – Different systems / programs have separate copies of the same data. The same piece of information may be stored in two or more files. For example, the particulars of an individual who may be a customer or an employee may be stored in two or more files. Some of this information may be changing, such as the address, the pay drawn, etc. It is therefore, quite possible that while the address in the master file for one application has been updated the address in the master file for another application may have not been. It may not also be easy for the computer-based system to even find out as to in how many files the repeating items such as the address is occurring. The solution therefore, is to avoid this data redundancy and the keeping of multiple copies of the same information and replace it by a system where the address is stored at just one place physically, and is accessible by all applications from this itself.

The disadvantages of file based systems are:

- Incompatible data (different file formats).
- Lack of flexibility in organising and querying the data.
- Increased number of different application programs
- Lengthy Development times – Programmers must design their own file formats.
- Excessive Program Maintenance.
- Each Application programmer must maintain their own data.
- Each Application program needs to include code for the metadata of each file.
- Each application program must have its own processing routines for reading, inserting, updating and deleting data
- Lack of coordination and central control.
- Problems with Data Redundancy.
- Waste of space to have duplicate data.
- Causes more maintenance headaches.
- When data changes in one file, could cause inconsistencies.
- Compromises data integrity.

These disadvantages of file based system motivate a database approach, which will be taken in the next section for discussion.

6.3 DATABASE APRROACH

Having pointed out some difficulties that arise in a straight forward file-oriented approach towards information system development, it is useful to see how the problems stated in the earlier section can be mitigated by using the database approach.

Organisations involved in manufacturing and business or public utility services such as hospitals, hotels, government departments, etc. would be in a position to rely into the database approach. Some of the advantages of database approach are:

- Central Repository of shared data
- Data is managed by a controlling agent
- Stored in a standardised, convenient form
- Consistency of data
- Integrity of data
- Security of data (Unauthorised access is restricted)
- Data independence
- Allows for more analysis of the same amount of data
- Improves data access and system performance
- Potentially increases productivity
- Increases concurrency
- Provides data backups and recovery.

Some potential disadvantages of database systems are the cost of implementing them, the amount of effort needed to transfer data into the database from a current system, and also the impact on the whole company if the database fails (even if only for a relatively short period). It requires DBMS software for implementation.

6.4 DATABASE AND DBMS

You have seen in the previous section the purpose for which a DBMS approach is preferred over the conventional file based approach. Since the DBMS of an organisation will in some sense reflect the nature of activities in the organisation, some familiarity with the basic concepts, principles and terms used in this field are important. This section concentrates on those components which are relevant in the context of a database approach.

6.4.1 Database

A database is a logically coherent collection of data with some inherent meaning, representing some aspect of real world and which is designed, built and populated with data for a specific purpose.

6.4.2 Why Use a Database?

A database gives users access to data, which they can view, enter, or update, within the limits of the access rights granted to them. Databases become all the more useful as the amount of data stored continues to grow.

A database can be local, meaning that it can be used on one machine by one user only, or it can be distributed, meaning that the information is stored on remote machines and can be accessed over a network. The primary advantage of using databases is that they can be accessed by multiple users at once.

6.4.3 Database Management System (DBMS)

It is a collection of programs that enables user to create, store, modify and extract information from a database. In other words it is general-purpose software that provides the users with the processes of defining; constructing and manipulating the database for various applications. There are different types of DBMSs, ranging from small systems that run on personal computers to huge systems that run on mainframes.

The major operations that can, be performed on the database by using database software are:

- Database Definition
- Database creation (storing data in a defined database)
- Retrieval (query and reporting)
- Update(Changing the contents of the database)
- Programming User Facilities for system development)
- Database revision and restructuring
- Database integrity control
- Performance Monitoring

The following are some of the examples of database applications:

- Computerised Library Systems
- Automated Teller Machines (ATMs)
- Flight Reservation Systems
- Railway Reservation Systems
- Inventory Management Systems
- Application Systems meant for Government Organisations
- Banking Systems
- Insurance Systems
- Hospitals and Health Sector
- Retails Markets
- Transport
- Education
- Manufacturing
- IT and BPO
- Share Business

6.4.4 Characteristics of Data in a Database

The data in a database should have the following characteristic features:

- ***Shared*** – Data in a database are shared among different users and applications.
- ***Persistence*** – Data in a database exist permanently in the sense the data's life time is beyond the scope of the process that created it.
- ***Validity/Integrity/Correctness*** – Data should be correct with respect to the real world entity that they represent.
- ***Security*** – Data should be protected from unauthorised access.
- ***Consistency*** – Whenever more than one data element in a database represents related to real-world values, the values should be consistent with respect to the relationship.
- ***Non-Redundancy*** – No two data items in a database should represent the same real-world entity.
- ***Independence*** – The three levels in the schema: internal, conceptual and external (discussed in the later sections), should be independent of each other so that the changes in the schema at one level should not affect the other levels.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Define a DBMS. Give some examples.

....

....

....

6.5 LEVELS OF ABSTRACTION IN A DBMS

The three-level architecture forms the basis of modern database architectures. This is in agreement with the ANSI/SPARC study group on Database Management Systems. ANSI/SPARC is the American National Standards Institute/Standard Planning and Requirement Committee). As shown in the figure, the architecture for DBMSs is divided into three general levels: ***External, Conceptual*** and ***Internal***.

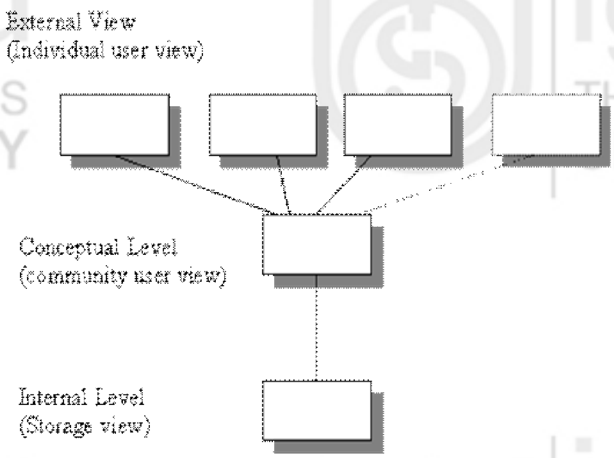


Fig. 6.1: Three-level architecture

External Level: This level is concerned with the way individual users see the data.

Conceptual level: This level can be regarded as a community user view - a formal description of data of interest to the organisation, independent of any storage considerations.

Internal level: This level is concerned with the way in which the data is actually stored.

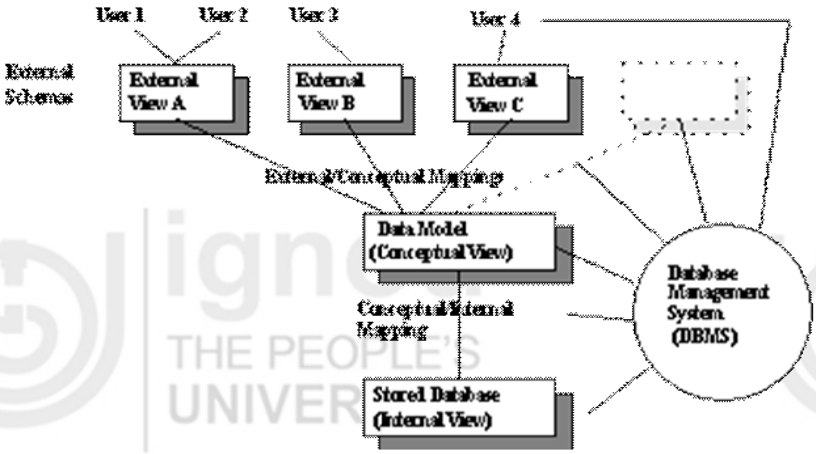


Fig. 6.2: Working of the three- level architecture

6.5.1 External View

A user is anyone who needs to access some portion of the data. They may range from application programmers to casual users with ad -hoc queries. Each user has a language at his/her disposal.

The application programmer may use any high level language or a front-end interface while the casual user will probably use a query language.

Regardless of the language used, it will include a Data Sub- Language DSL which is that subset of the language which is concerned with storage and retrieval of information in the database and may or may not be apparent to the user.

A DSL is a combination of two languages:

Data definition language (DDL) – provides for the definition or description of database objects.

Data manipulation language (DML) – supports the manipulation or processing of database objects.

Each user sees the data in terms of an external view: Defined by an external schema, consisting basically of descriptions of each of the various types of external record in that external view, and also a definition of the mapping between the external schema and the underlying conceptual schema.

6.5.2 Conceptual View

- An abstract representation of the entire information content of the database.
- It is in general a view of the data as it actually is, that is, it is a "model" of the "real world"
- It consists of multiple occurrences of multiple types of conceptual record, defined in the conceptual schema.

- To achieve data independence, the definitions of conceptual records must involve information content only.
- storage structure is ignored
- access strategy is ignored
- In addition to definitions, the conceptual schema contains authorisation and validation procedures.

6.5.3 Internal View

The internal view is a low-level representation of the entire database consisting of multiple occurrences of multiple types of internal (stored) records.

It is however at one remove from the physical level since it does not deal in terms of physical records or blocks nor with any device specific constraints such as cylinder or track sizes. A detail of mapping to physical storage is highly implementation specific and is not expressed in the three-level architecture.

The internal view described by the internal schema:

- defines the various types of stored record
- what indices exist
- how stored fields are represented
- what physical sequence the stored records are in

In effect the internal schema is the storage structure definition.

Mappings

The conceptual/internal mapping:

- defines conceptual and internal view correspondence
- specifies mapping from conceptual records to their stored counterparts
- An external/conceptual mapping:
- defines a particular external and conceptual view correspondence
- A change to the storage structure definition means that the conceptual/internal mapping must be changed accordingly, so that the conceptual schema may remain invariant, achieving physical data independence.

A change to the conceptual definition means that the conceptual/external mapping must be changed accordingly, so that the external schema may remain invariant, achieving logical data independence.

6.5.4 Data Independence

The three schema architecture further explains the concept of data independence, the capacity to change the schema at one level without having to change the schema at the next higher level.

There are two types of data independence:

i) **Logical data independence**

- The ability to change the conceptual schema without having to change the external schemas or application programs. When data is added or removed, only the view definition and the mappings need to be changed in the DBMS that support logical data independence.

- If the conceptual schema undergoes a logical reorganisation, application programs that reference the external schema constructs must work as before.

ii) Physical data independence

- The ability to change the internal schema without having to change the conceptual schema. By extension, the external schema should not change as well.
- Physical file reorganisation to improve performance (such as creating access structures) results in a change to the internal schema. If the same data as before remains in the database, the conceptual schema should not change.
- For example, providing an access path to improve retrieval speed of section records by semester and year, should not require a query to be changed, although it should become more efficient by utilising the access path.

With a multi-level DBMS, the catalogue must be expanded to include information on how to map requests and data among the levels. The DBMS uses additional software to accomplish the mappings.

Data independence occurs because when the schema is changed at some level, the schema at the next higher level remains unchanged. Only the mapping between the levels is changed.

6.5.5 Database Languages and Interfaces

Because a database supports a number of user groups, as mentioned previously, the DBMS must have languages and interfaces that support each user group.

DBMS Languages

- DDL – the **data definition language**, used by the DBA and database designers to define the conceptual and internal schemas.
- The DBMS has a DDL compiler to process DDL statements in order to identify the schema constructs, and to store the description in the catalogue.
- In databases where there is a separation between the conceptual and internal schemas, DDL is used to specify the conceptual schema, and SDL, **storage definition language**, is used to specify the internal schema.
- For true three-schema architecture, VDL, **view definition language**, is used to specify the user views and their mappings to the conceptual schema. But in most DBMSs, the DDL is used to specify both the conceptual schema and the external schemas.
- Once the schemas are compiled, and the database is populated with data, users need to manipulate the database. Manipulations include retrieval, insertion, deletion and modification.
- The DBMS provides operations using the DML, **data manipulation language**.
- In most DBMSs, the VDL, DML and the DML are not considered separate languages, but a comprehensive integrated language for conceptual schema definition, view definition and data manipulation. Storage definition is kept separate to fine-tune the performance, usually done by the DBA staff.
- An example of a comprehensive language: SQL, which represents a VDL, DDL, DML as well as statements for constraint specification, etc.

DBMS Interfaces

Types of interfaces provided by the DBMS include:

Menu-Based Interfaces for Web Clients or Browsing

- Present users with list of options (menus)
- Lead user through formulation of request
- Query is composed of selection options from menu displayed by system.

Forms-Based Interfaces

- Displays a form to each user.
- User can fill out form to insert new data or fill out only certain entries.
- Designed and programmed for naïve users as interfaces to canned transactions.

Graphical User Interfaces

- Displays a schema to the user in diagram form. The user can specify a query by manipulating the diagram. GUIs use both forms and menus.

Natural Language Interfaces

- Accept requests in written English, or other languages and attempt to understand them.
- Interface has its own schema, and a dictionary of important words. Uses the schema and dictionary to interpret a natural language request.

Interfaces for Parametric Users

- Parametric users have small set of operations they perform.
- Analysts and programmers design and implement a special interface for each class of naïve users.
- Often a small set of commands included to minimise the number of keystrokes required. (I.e. function keys)

Interfaces for the DBA

- Systems contain privileged commands only for DBA staff.
- Include commands for creating accounts, setting parameters, authorising accounts, changing the schema, reorganising the storage structures etc.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

2) Describe the three levels of data abstraction?

3) What is Data Independence?

....

....

....

....

6.6 DATABASE ENVIRONMENT

The following components form the Database System Environment:

Stored Data Manager

- The database and the database catalogue are stored on disk
- Access to the disk is handled by the Operating System.
- A higher-level stored data manager controls access to DBMS information that is stored on disk, whether part of the database or the catalogue.
- The stored data manager may use basic OS services for carrying out low-level data transfer, such as handling buffers.
- Once data is in buffers, the other DBMS modules, as well as other application programs can process it.

DDL Compiler

A DDL Compiler processes the schema definitions and stores the descriptions (meta-data) in the catalogue.

Runtime Database Processor

It handles database access at runtime. It receives retrieval or update operations and carries them out on the database. Its access to the disk goes through the stored data manager.

Query Compiler

It handles high-level queries entered interactively. It parses, analyses and interprets a query, then generates calls to the runtime processor for execution.

Precompiler

Precompiler extracts DML commands from an application program written in a host language. The commands are sent to DML compiler for compilation into code for database access. The rest is sent to the host language compiler.

Client Program

This program accesses the DBMS running on a separate computer from the computer on which the database resides. It is called the client computer, and the other is the database server. In some cases a middle level is called the application server.

Database System Utilities

DBMSs have database utilities that help the DBA manage the system. Functions include:

- **Loading** - used to load existing text/sequential files into the database. Source format and desired target file are specified to the utility, and the utility reformats the data to load into a table.
- **Backup** – creates a backup copy of the database, usually by dumping database onto tape. Can be used to restore the database in case of failure. Incremental backup can be used which records only the changes since the last backup.
- **File Reorganisation** – reorganise database files into different file organisations to improve performance.

- **Performance Monitoring** – monitors database usage and provides statistics to the DBA. DBA uses the statistics for decision-making.

Tools, Environment and Communication Facilities

- **CASE Tools** – used in the design phase to help speed up the development process.
- **Data dictionary system** – stores catalogue information about schemas and constraints, as well as design decisions, usage standards, application program descriptions, user information. Also called an information repository. Can be accesses directly by DBA or users when needed.
- **Application development environments** – (i.e., JBuilder) provide environment for developing database applications, and include facilities to help in database design, GUI development, querying and updating and application development.
- **Communication software** – allow users at remote locations to access the database through computer terminals, workstations or personal computers. Connected to the database through data communications hardware such as phone lines, local area networks etc.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) What is a DML Compiler?

....

....

....

6.7 TYPES OF DBMS SYSTEMS

There are several common types of databases; each type of database has its own data model (how the data is structured). They include***; Flat File Model, Hierarchical Model, Relational Model, Network Model, Object Oriented Relational Model***. These approaches differ in the way in which they permit the user to view and manipulate relationships. As far as practical implementation is concerned, nowadays, only Relational and Object Oriented Relational models are the most preferred models by the software industry.

6.7.1 Flat-file Database

A flat file database is a database designed around a single table. The flat file design puts all database information in one table, or list, with fields to represent all parameters. A flat file may contain many fields, often, with duplicate data that are prone to data corruption. If you decide to merge data between two flat files, you need to copy and paste relevant information from one file to the other. There is no automation between flat files. If you have two or more flat files that contain client addresses, for example, and a client moved, you would have to manually modify the address parameters in each file that contains that client's information. Changing information in one file has no bearing on other files. Flat files offer the functionality to store information, manipulate fields, print or display formatted information and exchange information with others, through email and over the Internet. Some flat files may be attached to external files, such as text editors, to extend functionality and manage related information.

6.7.2 Hierarchical Database Management System

As its name implies, the Hierarchical Database Model defines hierarchically-arranged data.

Perhaps the most intuitive way to visualise this type of relationship is by visualising an upside down tree of data. In this tree, a single table acts as the "root" of the database from which other tables "branch" out.

You will be instantly familiar with this relationship because that is how all windows-based directory management systems (like Windows Explorer) work these days.

Relationships in such a system are thought of in terms of children and parents such that a child may only have one parent but a parent can have multiple children. Parents and children are tied together by links called "pointers" (perhaps physical addresses inside the file system). A parent will have a list of pointers to each of their children as shown in the *figure 6.3*.

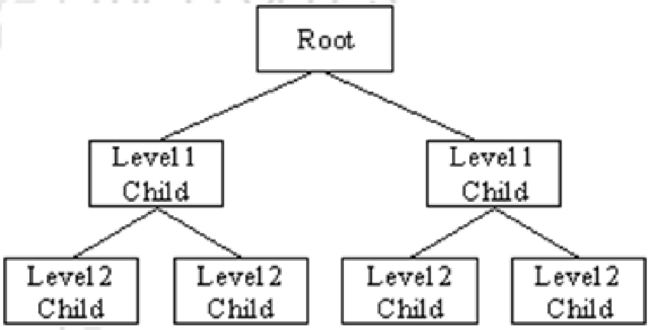


Fig. 6.3: Hierarchical Database Model

This child / parent rule assures that data is systematically accessible. To get to a low-level table, you start at the root and work your way down through the tree until you reach your target. Of course, as you might imagine, one problem with this system is that the user must know how the tree is structured in order to find anything!

The hierarchical model however, is much more efficient than the flat-file model we discussed earlier because there is not as much need for redundant data. If a change in the data is necessary, the change might only need to be processed once. Consider the student flat-file database example as given below:

Table 6.1: Student flat-file

Name	Address	Course	Grade
Mr. Eric Tachibana	123 Kensigton	Chemistry 102	C+
Mr. Eric Tachibana	123 Kensigton	Chinese 3	A
Mr. Eric Tachibana	122 Kensigton	Data Structures	B
Mr. Eric Tachibana	123 Kensigton	English 101	A
Ms. Tonya Lippert	88 West 1st St.	Psychology 101	A
Mrs. Tonya Ducovney	100 Capitol Ln.	Psychology 102	A
Ms. Tonya Lippert	88 West 1st St.	Human Cultures	A
Ms. Tonya Lippert	88 West 1st St.	European Governments	A

As we mentioned before, this flat-file database would store an excessive amount of redundant data. If we implemented this in a hierarchical database model, we would get much less redundant data. Consider the following hierarchical database scheme:

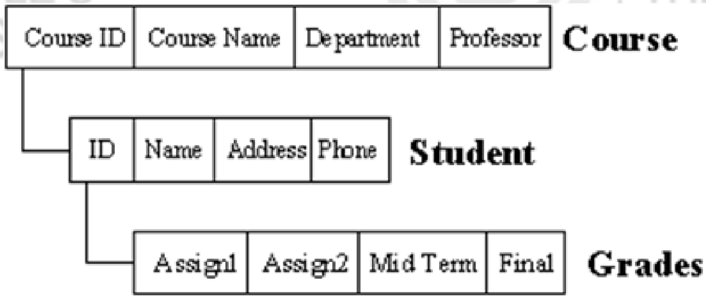


Fig. 6.4: Hierarchical database Model for Student flat-file

However, as you can imagine, the hierarchical database model has some serious problems. For one, you cannot add a record to a child table until it has already been incorporated into the parent table. This might be troublesome if, for example, you wanted to add a student who had not yet signed up for any courses.

Worse, yet, the hierarchical database model still create repetition of data within the database. You might imagine that in the database system shown above, there may be a higher level that includes multiple courses. In this case, there could be redundancy because students would be enrolled in several courses and thus each "course tree" would have redundant student information.

Redundancy would occur because hierarchical databases handle one-to-many relationships well but do not handle many-to-many relationships well. This is because a child may only have one parent. However, in many cases you will want to have the child be related to more than one parent. For instance, the relationship between student and class is a "many-to-many". Not only can a student take many subjects but a subject may also be taken by many students. How would you model this relationship simply and efficiently using a hierarchical database? The answer is that you wouldn't.

Though this problem can be solved with multiple databases creating logical links between children, the fix is very kludge and awkward. Faced with these serious problems, the computer brains of the world got together and came up with the network model.

Disadvantage

Links are unidirectional i.e., parent to child only. The model does not support many to many relationships. So, in this model if we need any information from lower level then we have to follow the complete hierarchy of the system.

6.7.3 Network Database Management System

In this many users can access and share one or more databases located on a network. The network model is the entity relationship model with all relationship restricted to be either binary or many-one relationships.

The data is represented by collection of records and relationships among data are represented as links, which can be diagrammatically viewed as pointers. The pointers are used to locate a particular record.

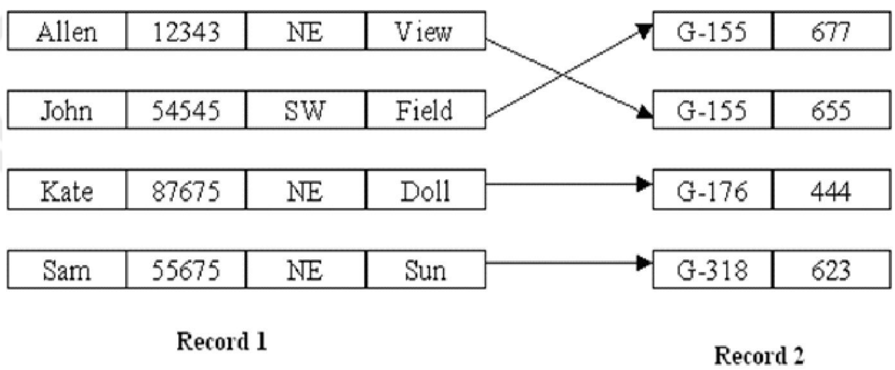


Fig. 6.5: A Network Scheme for a Bank Database

Disadvantage

The use of pointers leads to complex structures, which makes mapping of related data very difficult.

6.7.4 Relational Database Management System

Relational database, on the other hand, incorporates multiple tables with methods for the tables to work together. The relationships between table data can be collated, merged and displayed in database forms. Most relational databases offer functionality to share data:

- Across networks
- Over the Internet
- With laptops and other electronic devices, such as palm pilots
- With other software systems

Designing a relational database takes more planning than flat file databases. With flat files, you may add information, as you deem necessary. With relational databases, you must be careful to store data in tables such that the relationships make sense. Building a relational database is dependant upon your ability to establish a relational model. The model must fully describe how the data is organised, in terms of data structure, integrity, querying, manipulation and storage.

Relational databases allow you to define certain record fields, as keys or indexes, to perform search queries, join table records and establish integrity constraints. Search queries are faster and more accurate when based on indexed values. Table records can be easily joined by the indexed values. Integrity constraints can be established to ensure that table relationships are valid. If you are able to establish a one-to-many relationship in your data tables, you should be using a relational database because a flat file is not sufficient to handle your data processing needs.

Relational databases offer more robust reporting with report generators that filter and display selected fields. Relational databases offer the capability to build your own reporting modules. Most relational databases also offer the capability to import and export data from other software. Let us see the terminology that is used in RDBMS:

Relation -> table; Tuple -> Row, Record;

Cardinality -> No. of rows; Attribute -> column, field

Degree -> No. of columns;   Primary Key -> unique identifier

Domain -> Set of legal values;

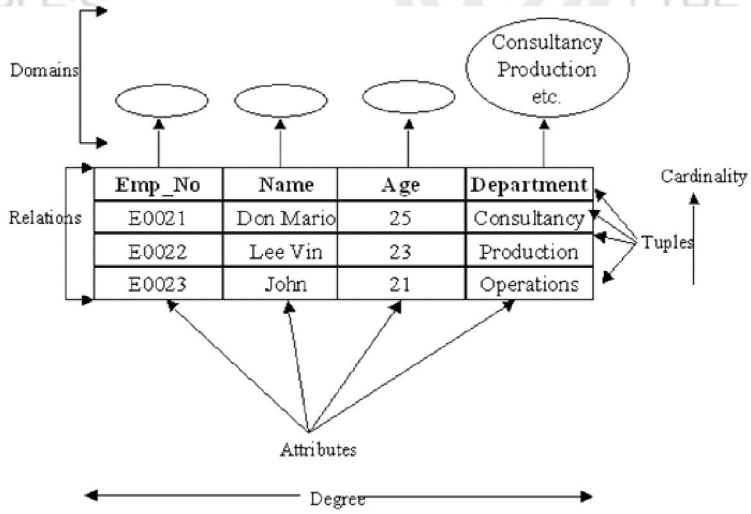


Fig. 6.6: Components of RDBMS

Example: A database file that consists of more than one tables and the relationship can be created between tables on the common field is called relational database. For example a database "student" is created having two tables "Address" and "Marks" as shown in figure below.

Address	Marks
Code	Code
Name	Math
City	Physics
Telephone	English

In the above tables, "Code" is common field in both tables and these two tables are linked together on this field.

The relational database is very popular and it is mostly used to create and maintain the large amount of related data. In business, a relational database contains the following types of tables.

- Table to store customer information
- Table to store employee information
- Table to store order information
- Table to store inventory information etc.

The customer, order and inventory table can be linked together to process or to generate reports of orders and billing etc.

6.7.5 Object Oriented Relational Database Management System

Object oriented databases are also called Object Database Management Systems (ODBMS). Object databases store objects rather than data such as integers, strings or real numbers. Objects are used in object oriented languages such as Smalltalk, C++, Java, and others. Objects basically consist of the following:

- **Attributes** – Attributes are data which defines the characteristics of an object. This data may be simple such as integers, strings, and real numbers or it may be a reference to a complex object.
- **Methods** – Methods define the behaviour of an object and are what was formally called procedures or functions.

Therefore objects contain both executable code and data. There are other characteristics of objects such as whether methods or data can be accessed from outside the object. We don't consider this here, to keep the definition simple and to apply it to what an object database is. One other term worth mentioning is classes. Classes are used in object oriented programming to define the data and methods the object will contain. The class is like a template to the object. The class does not itself contain data or methods but defines the data and methods contained in the object. The class is used to create (instantiate) the object. Classes may be used in object databases to recreate parts of the object that may not actually be stored in the database. Methods may not be stored in the database and may be recreated by using a class.

Comparison to Relational Databases

Relational databases store data in tables that are two dimensional. The tables have rows and columns. Relational database tables are "normalised" so data is not repeated more often than necessary. All table columns depend on a primary key (a unique value in the column) to identify the column. Once the specific column is identified, data from one or more rows associated with that column may be obtained or changed.

To put objects into relational databases, they must be described in terms of simple string, integer, or real number data. For instance in the case of an airplane. The wing may be placed in one table with rows and columns describing its dimensions and characteristics. The fuselage may be in another table, the propeller in another table, tires, and so on.

Breaking complex information out into simple data takes time and is labor intensive. Code must be written to accomplish this task.

Object Persistence

With traditional databases, data manipulated by the application is transient and data in the database is persisted (Stored on a permanent storage device). In object databases, the application can manipulate both transient and persisted data.

When to Use Object Databases

Object databases should be used when there is complex data and/or complex data relationships. This includes a many to many object relationship. Object databases should not be used when there would be few join tables and there are large volumes of simple transactional data.

Object databases work well with:

- CAS Applications (CASE-computer aided software engineering, CAD-computer aided design, CAM-computer aided manufacture)
- Multimedia Applications
- Object projects that change over time.
- Commerce.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

5) What is Data Model?

....

....

....

....

6.8 VARIOUS DBMS ARCHITECTURES

DBMS may not be similar for all the applications. Let us discuss various types of DBMS architectures in this session.

6.8.1 Centralised DBMS Architecture

- Used mainframes to provide main processing for user application programs, user interface programs and DBMS functionality
- User accessed systems via "dumb" computer terminals that only provided display capabilities, with no processing capabilities.
- All processing was performed remotely on the computer system, and only display information was sent to the terminals, connected via a network.
- Dumb terminals were replaces with workstations, which lead to the client/server architecture.

6.8.2 Client / Server Architecture

- Define specialised servers with specific functionalities (file servers, print servers, web servers, database servers)
- Many client machines can access resources provided by specialised server.
- Client machines provide user with the appropriate interfaces to utilise servers, as well as with local processing power to run local applications.
- Some machines are client sites, with client software installed and other machines are dedicated servers.
- **Client** – a user machine that provides user interface capabilities and local processing.
- **Server** – machine that provides services to client machines such as file access, printing, and database access.

Two-tier Client / Server Architecture

- In relational DBMSs, user interfaces and application programs were first moved to the client side.
- SQL provided a standard language, which was a logical dividing point between client and server.
- Query and transaction functionality remained on server side. In this architecture, the server is called a query server, or transaction server.

- In relational DBMSs, the server is called an SQL server, because most RDBMSs use SQL.
- In such systems, the user interface and application programs run on the client, when DMBS access is needed, the program establishes a connection to the DBMS on the server side. Once the connection is created, the client can communicate with the DBMS.
- ODBC (Open Database Connectivity) is a standard that provides and application processing interface which allows client side programs to call the DBMS as long as both sides have the required software. Most database vendors provide ODBC drivers for their systems.
- Client programs can connect to several RDBMS and send query and transaction requests using the ODBC API
- Query requests are sent from the client to the server, and the server processes the request and sends the result to the client.
- A related Java standard is JDBC, which allows Java programs to access the DBMS through a standard interface.
- These systems are called two tier architectures because the software components are distributed over two systems, the client and server.

Three-tier Client/Server Architecture for Web Applications

- Many web applications use three-tier architecture, which adds an intermediate layer between the client and the database server.
- The middle tier is called the application server, or the web serrver. Plays an intermediate role, by storing business rules (procedures/constraints) used to access data from database.
- Can improve database security by checking the client's credentials before forwarding request to database server.
- Clients contain GUI interfaces and application specific rules.
- The intermediate server accepts the requests from the client, processes the request and sends the database commands to the db server, then passes the data from the database server to the client, where it may be processes further and filtered.
- The three tiers are: user interface, application rules, and data access.

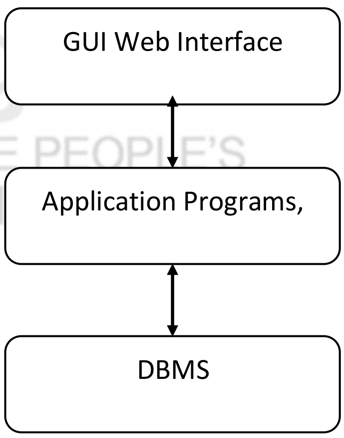


Fig. 6.7: Three-tier client/server Architecture

Let us study the security features of the database in the next section.

6.9 DATABASE SECURITY

Databases need to have level of security in order to protect the database against both malicious and accidental threats. A threat is any type of situation that will adversely affect the database system. Some factors that drive the need for security are as follows:

- Theft and fraud
- Confidentiality
- Integrity
- Privacy
- Database availability

Threats to database security can come from many sources. People are a substantial source of database threats. Different types of people can pose different threats. Users can gain unauthorised access through the use of another person's account. Some users may act as hackers and/or create viruses to adversely affect the performance of the system. Programmers can also pose similar threats. The Database Administrator can also cause problems by not imposing an adequate security policy.

Some threats related to the hardware of the system are as follows:

- Equipment failure
- Deliberate equipment damage (e.g. arson, bombs)
- Accidental / unforeseen equipment damage (e.g. fire, flood)
- Power failure
- Equipment theft

Threats can exist over the communication networks that an organisation uses. Techniques such as wire tapping, cable disruption (cutting / disconnecting), and electronic interference can all be used to disrupt services or reveal private information.

Countermeasures

Some countermeasures that can be employed are outlined below:

- Access Controls (can be Discretionary or Mandatory)
- Authorisation (granting legitimate access rights)
- Authentication (determining whether a user is who they claim to be)
- Backup
- Journaling (maintaining a log file - enables easy recovery of changes)
- Encryption (encoding data using an encryption algorithm)
- RAID (Redundant Array of Independent Disks - protects against data loss due to disk failure)
- Polyinstantiation (data objects that appear to have different values to users with different access rights / clearance)
- Views (virtual relations which can limit the data viewable by certain users).

6.10 POPULAR DBMS PACKAGES

There are many popular DBMS available in market, some of them are:

- Borland Paradox
- Filemaker
- IBMDB2
- Informix
- Ingres
- Interbase
- Microsoft SQL Server
- Microsoft Access
- Microsoft FoxPro
- Oracle
- Sybase
- MySQL
- PostgreSQL
- mSQL
- SQL Server

Let us discuss the database project development in the next section.

6.11 DATABASE PROJECT DEVELOPMENT

The conventional Systems Life Cycle consists of:

- Analysis
- Design
- Development
- Implementation
- Maintenance

In practice these phases are not always sharply distinguished; for small projects it may not be necessary to go formally through every one. The move from one phase to the next is essentially a move from the general to the specific.

At each stage, particularly where a DBMS is involved, we shall be concerned both with information and with processes to be performed using that information.

6.11.1 Analysis

The outputs from this stage should be:

A **CONCEPTUAL DATA MODEL** describing the information which is used within the organisation but not in computer-related terms. This level of data analysis will be considered in more detail later. One of the problems with any systems design in a large organisation is that it must proceed in a piecemeal manner – it is impossible to create a totally new Global system in one fell swoop, and each sub-system must dovetail with others which may be at quite a different stage of development. The conceptual data model provides a context within which more detailed design specifications can be produced, and should help in maintaining consistency from one application area to another.

A **CONCEPTUAL PROCESS MODEL** describing the functions of the organisation in terms of events (e.g. a purchase, a payment, a booking) and the processes which must be performed within the organisation to handle them. This may lead to a more detailed functional specification - describing the organisational requirements which must be satisfied, but not how they are to be achieved.

6.11.2 Design

This stage should produce:

A **LOGICAL DATA MODEL**: a description of the data to be stored in the database, using the conventions prescribed by the particular DBMS to be used. This is sometimes referred to as a SCHEMA and some DBMSs also give facilities for defining SUB-SCHEMA or partitions of the overall schema. Logical data models supported by present day DBMSs will be considered later.

A **SYSTEM SPECIFICATION**, describing in some detail what the proposed system should do. This will now refer to COMPUTER PROCESSES, but probably in terms of INPUT and OUTPUT MESSAGES rather than internal logic, describing, for instance, the effect of selecting an item from a menu, or any option within a command driven system. Program modules are defined in terms of the screen displays and/or reports which they generate. Note that the data referred to here has a temporary existence, in contrast with what is stored in the database itself.

6.11.3 Development

Specification of the database itself must now come down another level, to decisions about **PHYSICAL DATA STORAGE** in particular files on particular devices. For this a knowledge of the computer operating system, as well as the DBMS, is required. Conventional program development - coding, testing, debugging etc. may also be done. If a totally packaged system has been purchased this may not be necessary - it will simply be a matter of discovering how to use the command and query language already supplied to store and retrieve data, generate reports and other outputs. Even here an element of testing and debugging may be involved, since it is unlikely that the new user of a system will get it exactly right the first time. It is certainly inadvisable for this sort of experimentation to take place using a live database!

6.11.4 Implementation

This puts the work of the previous three phases into everyday use. It involves such things as loading the database with live rather than test data, staff training, probably the introduction of new working practices. It is not unusual to have an old and a new system running side by side for a while so that some back-up is available if the new system fails unexpectedly.

6.11.5 Maintenance

Systems once implemented generally require further work done on them as time goes by, either to correct original design faults or to accommodate changes in user requirements or operating constraints. One of the objectives of using a DBMS is to reduce the impact of such changes - for example the data can be physically re-arranged without affecting the logic of the programs which use it. Some DBMSs provide utility programs to re-organise the data when either its physical or logical design must be altered.

6.12 DATABASE ADMINISTRATOR (DBA)

The database administrator (DBA) is the person (or group of people) responsible for overall control of the database system. The DBA's responsibilities include the following:

- Deciding the information content of the database, i.e. identifying the entities of interest to the enterprise and the information to be recorded about those entities. This is defined by writing the conceptual schema using the DDL.
- Deciding the storage structure and access strategy, i.e. how the data is to be represented by writing the storage structure definition. The associated internal/conceptual schema must also be specified using the DDL.
- Liaising with users, i.e. to ensure that the data they require is available and to write the necessary external schemas and conceptual/external mapping (again using DDL).
- Defining authorisation checks and validation procedures. Authorisation checks and validation procedures are extensions to the conceptual schema and can be specified using the DDL.
- Defining a strategy for backup and recovery. For example periodic dumping of the database to a backup tape and procedures for reloading the database for backup. Use of a log file where each log record contains the values for database items before and after a change and can be used for recovery purposes.
- Monitoring performance and responding to changes in requirements, i.e. changing details of storage and access thereby organising the system so as to get the performance that is "best for the enterprise".

6.13 SUMMARY

A database system is an integrated collection of related files along with the details about their definition, interpretation, manipulation and maintenance. A DBMS is a major software component of database system. It consists of collection of interrelated data and programs to access that data. The primary goal of a DBMS is to provide an environment which is both convenient and efficient to use in retrieving information from and storing information into the database.

The DBMS not only makes the integrated collection of reliable and accurate data available to multiple applications and users but also controls from unauthorised, users to access the data.

A DBMS is a major software system consisting of a number of elements. It provides users DDL for defining the external and conceptual view of the data and DML for manipulating the data stored in the database. The database manager is the component of DBMS that provide the interface between the user and the file system. The database administration defines and maintains the three level of the database as well as the mapping between levels to insulate die higher levels from changes that take place in the lower levels. The DBA is responsible for implementing measures for ensuring the security, integrity and recovery of the database.

6.14 ANSWERS TO SELF CHECK EXERCISES

1) **DBMS** short for Database Management System is a software system that uses a standard method of cataloging, retrieving, and running queries on data. The **DBMS** manages incoming data, organises it, and provides ways for the data to be modified or extracted by users or other programs. Some **DBMS** examples include MySQL, PostgreSQL, Microsoft Access, SQL Server, FileMaker, Oracle, RDBMS, Clipper, and Foxpro.

2) There are three levels of abstraction:

   a. ***Physical level:*** The lowest level of abstraction describes how data are stored.

   b. ***Logical level:*** The next higher level of abstraction, describes what data are stored in database and what relationship among those data.

   c. ***View level:*** The highest level of abstraction describes only part of entire database.

3) Data independence means that the application is independent of the storage structure and access strategy of data?. In other words, The ability to modify the schema definition in one level should not affect the schema definition in the next higher level.

   Two types of Data Independence:

   Physical Data Independence: Modification in physical level should not affect the logical level.

   Logical Data Independence: Modification in logical level should affect the view level. Logical Data Independence is more difficult to achieve

4) It translates DML statements in a query language into low-level instruction that the query evaluation engine can understand.

5) It is a collection of conceptual tools for describing data, data relationships data semantics and constraints.

6.15 KEYWORDS

Data definition language (DDL)	:	The subset of SQL used for defining and examining the structure of a database.
Data manipulation language (DML)	:	1. The subset of SQL used for inserting, deleting, updating and fetching data in a database.
Data Sub Language (DSL)	:	A computer language used to define or manipulate the structure of a relational database management system (DBMS)
SQL	:	SQL is short for Structured Query Language, an international standard language for manipulating relational databases.

6.16 REFERENCES AND FURTHER READING

Desai, Bipin C. *An Introduction to Database Systems*. New Delhi: Galgotia, 1999. Print

Leon, Alexis and Leon, Mathews. *Database Management Systems.* New Delhi: Vikas Publishing, 2002. Print

Navathe, Elmasri. *Fundamental of Database Systems*. 3rd ed. Pearson Education, 2003. Print

Ramakrishna, Raghu and Gehrke, Johannes. *Database Management Systems*. 2nd ed. Mc Graw Hill, 2000. Print

Silberschats, Abraham, F.Korth, Henry and Sudarshan S. *Database System Concepts*. 5th ed. Mc Graw Hill International Edition, 2006. Print

Umanath, Narayan S. and Scamell, Richard W. *Data Modeling and Database Design.* Cengage Learning, 2009. Print

UNIT 7 MULTIMEDIA

Structure

7.0 OBJECTIVES

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

- define a multimedia system and its characteristics;
- discuss the various types of media associated with the multimedia systems;
- differentiate between multimedia and print media;
- describe various storage techniques, access media and formats;
- suggest various s/w and h/w requirements for a Multimedia System; and
- discuss security, copyright and other issues pertaining to the Multimedia Systems.

7.1 INTRODUCTION

With the advent of multimedia, the computer has evolved into a distinctive medium that is uniquely capable of juxtaposing text, images, audio, and video. Multimedia permits an extraordinary flexibility in conveying concepts – through words, pictures, and sounds, as something that can be built or played as well as read or watched. New genres, such as simulation games, are emerging that challenge the user or player to build some complex creation – a city, species, business, or world- out of some given set of resources, or that put the person into a simulated environment or through a scenario to meet a challenge or learn a skill. The computer thereby turns the passive reader into a participant; it cues the person of a need to do something, but not necessarily what to do. With multimedia the computer draws on more of the senses, and more dimensions of intelligence, enlarging the opportunity to learn for those who have been less able to learn from conventional teaching materials. Some uses of the new media are genuinely inspired, provocative, and engaging, and these examples suggest that multimedia have opened an important new chapter in the history of the imagination.

7.2 MULTIMEDIA

A multimedia system is a system which combines different types of media such as text, web pages, hyperlinks, audio, video and graphics. The use of multimedia gives the user a more full experience than a single media such as text. The result of a multimedia system will include some kind of multimedia presentation.

Multimedia systems are information systems that combine a variety of different media types, including text, hypertext, audio, images and video. Professional multimedia systems, especially at the time of creation, involve many participants with a wide breadth of experience.

The creation of multimedia systems encompasses each of the information processes.

Multimedia primarily focus' on the information process of displaying, however, it also investigates a number of the other elements involved, including processing, organising, storing and retrieving and collecting.

7.3 CHARACTERISTICS OF MULTIMEDIA SYSTEMS

Multimedia consists of information presented in various forms, including text, numerical, audio etc.. It will often include some kind of hypermedia. The hypermedia may contain clickable links or may be disguised as a menu such as at the start of a DVD.

Multimedia is the use of several different media to convey information (text, audio, graphics, animation, video, and interactivity). Multimedia also refers to computer data storage devices, especially those used to store multimedia content.

As the information is presented in various formats, multimedia enhances user experience and makes it easier and faster to grasp information. Presenting information in various formats is nothing new, but multimedia generally implies presenting information in various digital formats. It is also used in visual arts to describe works created using more than one medium.

Multimedia finds its application in various areas including, but not limited to, art, education, entertainment, engineering, medicine, mathematics, and scientific research. In education, multimedia is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopaedia and almanacs. A CBT lets the user go through a series of presentations, text about a particular topic, and associated illustrations in various information formats.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Define Multimedia and Multimedia Application.

....

....

....

....

7.3.1 Types of Media

The following are the main types of media:

Text refers to any typed material. Text includes numbers and characters whose meaning must be read to gain meaning. It may be typed in through a word processor but is also often typed into the multimedia software.

Numbers: Numerals which allow calculations. Includes currency

Hypertext is a term created by Ted Nelson to describe non-linear writing in which you follow linked paths through a world of textual documents. The most common use of hypertext is found in the links on World Wide Web pages.

Audio: Sound which has been digitised.

Images (Graphics)

Images are pictures, graphic elements which are displayed on the screen. They maybe of two types:

Bit-mapped graphics (also known as raster images): A bit-mapped image is one composed of pixels. This means that there is a relationship between the graphic on the screen and the bits in memory. Bit-mapped images can also be called raster images and would normally be composed in a paint program.

Bitmap images come from paint style programs and are the format used for digital photos. Bitmap images can be very large but can be compressed down to a small file size. Bitmaps pixelate when they are enlarged. ***Pixelation*** refers to that block and grainy look that photos and images get when they are enlarged to much. Jpegs are one example of raster images.

Vector graphics: These are made of objects such as straight lines, curves or shapes. It is created through a mathematical formula.

Animation: Animation begins with still images which are then given the illusion of movement is the movement of a graphic. It is the result of making a sequence of drawing called frames and showing them rapidly one after another. Animation maybe cel (also called cell) based or path based

Video: Video combines pictures and sound displayed over time. It is important to know that here are differences between an animation and a video. An animation begins with a still images which is then given the illusion of movement but a video records a continuous event such as a rock concert where movement and sound are of course a natural part of life; and very much so at a rock concert where a person's eardrums may be moving as much as anyone on stage.

Images have proven a very powerful means of communicating messages, and can evoke a strong response from the user or viewer. However, video, like sound, is very processor and memory hungry. It also requires special hardware and software to enable it to be recorded and, to a lesser extent, played back.

Piped/ streamed video is an effective method of displaying video footage on a computer screen. It involves a suitable video card to receive and translate the video data into a form that can be displayed on the screen. However, the computer has no control over the displayed signal.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

2) List the various types of media commonly used with multimedia?

....

....

....

....

7.3.2 H/W Requirements for Multimedia Systems

Multimedia technologies continue to develop, and are thus still in an evolutionary process. However below is an indication of what you would need to do some basic multimedia.

RAM: 1 GB minimum

SPEED: 1 GHz

STORAGE: Depending on the work your doing you may require anything from a 10GB available space upwards

The above specifications are only for minimal graphics work. If video editing or heavy graphics is required than you will need to at least double the minimal RAM and speed of the processor. The amount of RAM, ROM and hard drive space that a computer can utilise is a limiting factor in its multimedia capabilities. This is due to the high storage requirements demanded by high quality images and sound files.

7.3.3 Components of a Multimedia System

Consider the Components (Hardware and Software) required for a multimedia system:

- **Capture devices** – Video Camera, Video Recorder, Audio, Microphone, Keyboards, mouse, graphics tablets, 3D input devices, tactile sensors, VR devices. Digitising Hardware
- **Storage Devices** – Hard disks, CD-ROMs, DVD-ROM, etc
- **Communication Networks** — Local Networks, Intranets, Internet, Multimedia or other special high speed networks.
- **Computer Systems** – Multimedia Desktop machines, Workstations, MPEG/ VIDEO/DSP Hardware
- **Display Devices** – CD-quality speakers, HDTV,SVGA, Hi-Res monitors, Colour printers etc.

7.3.4 Key Issues for Multimedia Systems

Some of the key issues multimedia systems need to deal are:

- How to represent and store temporal information?
- How to strictly maintain the temporal relationships on play back/retrieval?
- What process are involved in the above Data has to represented **digitally** – Analog–Digital Conversion, Sampling etc
- Large Data Requirements – bandwidth, storage, data compression is usually mandatory.

7.3.5 Desirable Features for a Multimedia System

Given the above challenges the following feature a desirable for a Multimedia System:

- ***Very High Processing Power*** – needed to deal with large data processing and real time delivery of media.
- ***Multimedia Capable File System*** – needed to deliver real-time media – e.g. Video/Audio Streaming.
- ***Special Hardware/Software needed*** – e.g. RAID technology.
- ***Data Representations*** – File Formats that support multimedia should be easy to handle yet allow for compression/decompression in real-time.
- ***Efficient and High I/O*** – input and output to the file subsystem needs to be efficient and fast. Need to allow for real-time recording as well as playback of data.
- for example, Direct to Disk recording systems.
- ***Special Operating System*** – to allow access to file system and process data efficiently and quickly. Need to support direct transfers to disk, real-time scheduling, and fast interrupt processing, I/O streaming etc.

- ***Storage and Memory*** – large storage units (of the order of hundreds of Tb if not more) and large memory (several GB or more). Large Caches also required and high speed buses for efficient management.
- ***Network Support*** – Client-server systems common as distributed systems common.
- ***Software Tools*** – user friendly tools needed to handle media, design and develop applications, deliver media.

7.4 PRINT VS MULTIMEDIA

There are large differences between print and multimedia and advantages and disadvantages associated with both.

The Mode of Display

Print media is limited to hard copy such as books, newspapers, magazines and posters while multimedia addresses a number of senses and can be used on computers, mobile phones, PC/TV combinations and a host of other diverse applications.

The Interactive Nature of Multimedia

Interactivity means that the user is able to make an immediate response to what is happening and modify the processes. This means that multimedia is more than just a one way communication. Effectively, there is communication between the user and the application.

The Dynamic Nature of Multimedia

Once a book is published it cannot change, it is static. An interactive multimedia product can change depending on actions taken by both the author and the user. The multimedia author can easily and cheaply change the multimedia presentation whereas this is an expensive exercise for the traditional author. 'Smart' products can change the way they work based on characteristics the program 'learns' about the user.

Data Structure

Computers give us many different ways of accessing data and multimedia systems need to be structured so the user can easily find the required data. Structures include: A hierarchical structure or it can be sorted into an order such as alphabetical.

Indexing

We are familiar with the indexes provided in many reference books but these indexes are always very selective. An interactive multimedia product can index every occurrence of every word.

Ease of Distribution

Unlike print media which requires printing presses, ink, paper, a willing publisher, distribution networks in stores, advertising to sell, sales staff to sell the book and a whole lot of money to cover all of these costs multimedia just requires a computer, a person to create the material and a distribution media whether that be over the internet, disc or a combination. Clearly it is much simpler and cheaper to produce most multimedia then it is to produce books. However it would be simplistic to say that it is always easier. Movies such as Avatar are multimedia productions which require far more expenditure than any book and probably far more staff, but this is a particular example.

Authority of Document

Books are often granted more authority than a website. A website may have very little checking and editing by a third party whereas most books will have many people who read through the manuscripts and check sources. However this is not always the case. There have been many books published which have very doubtful facts and authority.

7.5 MAJOR AREAS OF MULTIMEDIA USE

Multimedia systems may be designed for a mass audience or for a specific target audience. Major areas include: education, leisure and entertainment information. It is possible to combine a number of these areas into one application. Educational games, or edutainment software, for example, consist of programs that look like games but are designed to teach in an entertaining manner. Each of the above examples deals with the presentation of information in a variety of forms. The significant advantage of multimedia is its interactive nature, which gives control to the participant, rather than the designer.

The variety of presentation methods makes multimedia a useful tool across all the age groups. Multimedia removes the limitation of forcing everyone to move at the same pace by providing an interactive approach to learning with more advanced users moving ahead. Government and businesses also use interactive multimedia programs to train staff. This type of training is not constrained by time and can be programmed to meet individual needs.

The main uses of Multimedia are:

- Multimedia In Education and Training
- Leisure and Entertainment
- Information Kiosks
- Virtual Reality and Simulations

Multimedia in Education and Training

The "Reader Rabbit" series is an educational multimedia program which provides interactive assistance in developing reading skills. Participants include school students, teachers, parents and siblings. The direct users (students) control activities as they make choices from a menu.

Data includes text, images and video combined to produce structured activities.

The information technology required is a multimedia computer with a CD-Rom, sufficient RAM and processing speed.

Leisure and Entertainment

Electronic games, 3D adventure games, sporting games and interactive movies are extremely popular forms of multimedia applications. The key to their popularity lies in their interactive nature. The new generations of games provide ingenious levels of interactivity and realism to captivate the user of the product. The attraction of this type of application is realism, fast action and user input through peripherals such as mouse, track-pad, keyboard and joystick. Computer-based games have led to many developments in interactive computing. This type of application requires a high level of graphics computing power and hence the impetus to develop more efficient algorithms for display movement and more powerful graphics cards.

Information Kiosks

Multimedia has been successfully used to provide guidance through the use of information centres in locations such as museums, airports, shopping precincts and libraries.

An information centre or kiosk allows the user to search for specific information in his or her own language without the need to approach a stranger. Location details, product availability, arrivals and departures, opening and closing times and so on are all examples of information that can be presented in this format. Touch screens are often an integral part of an information centre because they reduce the technology literacy requirements normally associated with the use of computers.

Virtual Reality and Simulations

Virtual reality is a computer-generated artificial reality that projects a person into a sensation of three dimensional space. Virtual reality is becoming more popular in arcade type games such as Atlantis. You may have even tried to tee off on a virtual golf driving range or driven a virtual reality race car. A far more important use of virtual reality is a simulator. A simulator is a device that represents the behaviour of physical or abstract systems. Simulators are used very effectively in training applications for aeroplane pilots or bus drivers to help them deal with various real-life situations in a safe environment. The medical area has also seen increases in the' application of simulation technology. For example, surgeons can rehearse a particular operation on a 'digital' patient, and some phobias, such as a fear of crowds, can be treated. One ethical issue for implementers of virtual reality equipment is how to keep users from suffering uncomfortable effects from using virtual reality systems. Known as cyber-sickness or simulator sickness, symptoms include eyestrain, nausea and confusion and even visual or audio flashback for some users.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) List some of the applications of Multimedia?

....

....

....

....

7.6 ADVANCES IN TECHNOLOGY

Technological advances have continued to influence the development of multimedia including: accessibility, communication speeds, compression rates, CDROM, DVD and firewire technologies. Ease of access to the Internet and World Wide Web has made it very easy for large numbers of people to access multimedia systems stored as a series of linked web pages.

The main categories are as follows:

- Storage Capacity
  - o Hard drives

- o CDROM
- o DVD
- Firewire
- Internet Bandwidth

Storage Capacity

Multimedia is very memory hungry but with the explosion in memory storage capacity this is becoming less and less of an issue. However it is important to note that without the large increase in storage capacity multimedia would still be difficult to use.

Hard Drives

Hard drive storage capacity continues to explode. Hard drives are now measured in gigabytes with plans underway for terabyte drives.

CDROM

CD-Rom, which stands for compact disk-read only memory, is an optical disk format that is used to hold pre-recorded text, graphics and sound. In this type of disk the content is recorded during manufacture and then read many times by the purchaser. Initially, a CD-Rom drive was only a single speed drive, with the ability to access data at 150 kilobytes per second. Through improved technology it is now possible to purchase drives that are 40 or 60 times (40x, 60x) this speed. The faster the drive spins, the more quickly it is able to deliver information to the processor. Each CD-Rom disk can store up to 783 megabytes of information. The development of this technology effectively began the multimedia revolution due to the sudden availability of high capacity random access data storage.

DVD

DVD (digital versatile disk) is an optical disk technology that is expected to rapidly replace the CD-Rom disk (as well as the audio compact disc) over the next few years. Increased storage means increased flexibility for multimedia developers. A DVD holds 4.7 gigabyte of information on one of its two sides, or enough for a 133-minute movie. With two layers on each of its two sides, it can hold up to 17 gigabytes of video, audio, or other information. Compared to the current CD-Rom disk of the same physical size, holding 700 megabytes, this is more than 28 times as much information!

Firewire

FireWire is essentially a new way to connect different pieces of equipment, such as digital videos and hard drives, so they can share information rapidly and easily. Originally created by Apple, and standardised in 1995 as the specification IEEE-1394, it is a technology that:

- allows fast transfer of data (up to 400 Mbps);
- allows many devices to be connected simultaneously on the bus network;
- is hot pluggable;
- provides power through the cable;
- allows plug and play;
- has low cabling costs;

- has increased data transfer rates through the use of a fairly inexpensive option.;
- This has expanded the possible applications of multimedia into more powerful and data intensive areas.

Internet Bandwidth

Improvement in communication speed (or bandwidth) through technologies such as cable, ADSL and satellite have made the presentation of multimedia content via the Internet a viable option increasing download speeds that used to frustrate many Internet users, particularly for pages that incorporate graphics, sound and video. Compression and animation technologies such as MP3, QuickTime, Flash, Shockwave, Java and JavaScript enhance the interactive nature of the World Wide Web.

7.7 MULTIMEDIA DESIGN

It is important when designing multimedia that all elements be in harmony and that images and graphics be optimised for fast and efficient use of bandwidth.

There are 3 aspects to optimisation:

- the basic coding of the page,
- scripts that are used on the page,
- images (large can have the most significant impact on load times).

Image Types

There are many image formats in use on web sites but the three most common are GIF, JPEG, and Flash content. Our focus is on GIF and JPEG.

The gif format only uses 8-bit colour (256 colours) but it is a 'lossless' format and so detail is not reduced. Gif images are the best way to render artwork such as line drawings, charts and maps, particularly if the number of colours can be reduced. However GIFs will give a larger file size if the picture has too many colours. For an image with a large number of colours a jpeg is preferable.

Jpeg images use 24-bit colour (16,777,216 colours) but the compression algorithm is loss. This means that resolution on a computer screen is lost on compression, this may not be noticeable but a hardcopy image would show the degradation. Jpeg images are used for photographs (screen but not good printed). A reasonable quality photo in jpeg is commonly around 20-30 Kb where the equivalent in gif could be 100-300 Kb. With jpeg files the quality required is set when the image is saved.

Each image format has strengths and weaknesses. GIF or Graphics Interchange Format was developed by CompuServe before the Internet days as a way to share images on the CompuServe bulletin boards. Due to limitations with screen resolutions and colour depths at the time, GIF images were only able to show 256 colours, more colours were imitated by Dithering, a process of tricking the eye into seeing one colour by using 2 or more sets of colour dots spaced closely together.

A chessboard with black and white squares becomes a grey blob when viewed from a distance even though close up we can see each black or white square. This is the concept behind dithering. Our eyes merge the black and white squares together. On a computer screen these black and white squares would be called pixels awhile the number of black and white squares would be called the resolution. The more squares the better the definition of the image. Unfortunately if there are too many colours to be dithered the file size becomes very large.

The JPEG file format is newer and can handle up to 16 million colours. The initial problem is that JPEG images do contain many more colours, and each colour requires coding for display, possibly making larger file sizes. The compression rate of a JPEG is easily changed at the time it is saved. It is possible to get a good balance between resolution and compression rate. Lower compression rates will still give a quality print. As the compression rate increases the image on the screen will not seem to vary much but the print quality will degrade quickly as the compression rate is increased.

Improving Load Times

The main way to make an image load faster is to make the file size smaller. This can be accomplished by either:

- making the dimensions of the image smaller, or
- decrease the amount of coding that is required to display the image.

The easiest way to reduce an image's file size is to reduce the image's physical dimensions. Make the picture smaller. Smaller the image, smaller is the file size. If we reduce the image size by one half to 40 pixels by 40 pixels we then have 40x40 or 1600 pixels. So reducing the image size in half reduces the file size to one fourth of the original. The following are the principles to be followed for the Image file size reduction:

- Use the smallest image dimensions that will work with your layout. And likewise the fewer images on the page, the fewer image pixels, therefore the smaller the page size. GIF and JPEG image formats use different methods of saving image information and so tend to be better at showing some types of images and worse at showing others. GIF images, since they are limited to 256 colours per image, are better at displaying images with large solid blocks of colour and images with very small physical dimensions. The images look flatter and don't have the depth of colour of a JPEG. The GIF format will produce smaller file sizes than JPEG for these types of images.JPEG images are better at showing gradients or subtle changes from one colour to another. Therefore JPEGs reproduce photographs very well, or any other image with gradations. This means that the JPEG format will produce smaller file sizes for photo style images than the GIF format will.

- Choose the correct image format for the type of image you are using. Most web pages will contain a combination of GIF and JPEG images. Decreasing the coding is called image compression. Both GIF and JPEG images can be compressed but the process is different. In GIF images we try to limit the number of colours, in a JPEG image we use software algorithms to remove redundant information from the file. Whenever we compress a file we will loose some image quality. We have to reach a balance between a small file size and acceptable image quality.

- Find the least acceptable level of image quality. Most images can handle some compression with very little quality loss, and all images can stand more image quality loss and still be acceptable. You need to be able to decide how much decline in image quality is acceptable, but remember the lower the quality the fast the image load. GIF images can usually be reduced from 256 colours to 128 colours or less, the fewer colours used the smaller the file size. JPEG images can almost always be reduced to a quality setting of 80% and frequently can be reduced down to as little as 15-30%. So when you use a higher compression level (smaller number) the file size will be reduced. Try smaller and smaller settings until you find the smallest setting that still displays an acceptable quality. The fastest loading page will have no images and the slowest loading page will be completely filled with high resolution images.

Hardware in Multimedia Systems

A variety of hardware is required to create and display the multimedia applications that are currently available including:

- Display Devices
  - o Cathode Ray Tubes
  - o Flat Panel
    - § Touch Screens
    - § Liquid Crystal Display (LCD) Screens
- Projection Devices
- Speakers and sound
- MIDI

Display Devices

Display devices, also called screens, monitors or CRTs, are output devices

Display screens are of two types:

Cathode-ray tubes (CRTs) produce an image by firing a beam of electrons onto the inside of a phosphorous coated glass screen.

Flat panel displays are made up of two plates of glass with a substance in between them that can be activated in different ways. Flat panel displays are distinguishable by either the substance between the plates (liquid crystal, electroluminescent material or gas plasma) or the arrangement of the transistors in the screen that control each pixel.

Touch Screens

A touch screen is sensitive to human touch. Used at information kiosks, computer based training, and by those having difficulty manipulating a mouse or keyboard. 3 types are: resistive, surface wave and capacitive.

Resistive: Coated with thin, metallic, electrically conductive and resistive layer. When touched a change in electrical current is registered as a touch event and sent to controller for processing.

Surface wave: Uses ultrasonic waves that pass over the screen. When touched, a portion of the wave is absorbed. This change in the ultrasonic waves registers the position of the touch event and is sent to controller for processing.

Capacitive: Coated with material that stores electrical charges. When touched, a small charge is drawn to point of contact. Circuits located at each corner of panel measure charge and send the information to the controller for processing. Capacitive touch screens must be touched with a finger.

Liquid Crystal Display (LCD) Screens

An LCD screen is an electro-optical device commonly used in digital watches, calculators, and portable computers.

- The LCD is a liquid crystal placed between a pair of transparent electrodes. Liquid crystal changes the phase of the light passing through it and can be controlled by voltage applied between electrodes. When placed between a pair of plane polariser plates then light can pass through only if the correct voltage is applied.
- LCDs are formed by integrating a number of such cells or by using a single liquid crystal plate and a pattern of electrodes. Electrodes in computer screens are in rows and columns. By applying voltage to a row and several columns the pixels at intersections are set.
- Putting a transistor, on top of each pixel, can slow fading. It will also 'remember' the setting of that pixel. These active matrix displays are as good as CRTs but much more expensive than passive matrix displays.

Projection Devices

A data projector is a device that allows the projection of video and graphic images onto a screen. Typically, projected images were best viewed in reduced light but the newer models are much brighter (1,000 lumens or more) and this has become unnecessary.

Three main types of technology are used to project images:

- **Three guns:** Separate guns for each of the red, blue and green colours are used to project light onto a screen. Although they produce the most light output they are also the most expensive.
- **Digital light projector (DLP**) technology uses a tiny multi-faceted semiconductor mirror chip to reflect light from a light source. The chip controls tiny semiconductor-based mirrors to produce an image that is sharper than the traditional LCD version.
- **LCD-based projectors** perform in a similar way to LCD screens where the light source is passed through an LCD before being projected onto a screen.

Speakers and Sound

Audio output is relatively straightforward - you just connect speakers or headphones to the sound card of the computer.

- Sound output devices are designed to convert binary data into information by producing digitised sound.
- To do this you need the necessary software driver and sound card and a digital audio circuit board such as a Sound Blaster.
- Sound cards generally plug into an expansion slot on the motherboard or come integrated with the mother-board on newer machines.
- Speakers provide the easiest and cheapest output device though it is often the quality of the sound file rather than the speakers that contribute most to the quality of the sound.

MIDI

MIDI, or musical instrument digital interface, is used to input or output analog musical information from electronic musical instruments (known as synthesizers) in digital format ready for use by a computer. The information transmitted contains the note identification, time of play and the loudness, rather than the actual note itself. ***Midi*** files are very small compared to ***wav*** files.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) State whether the following statements are true or false.

a) A computer capable of handling text, graphics, audio, animation and video is called multimedia computer.

b) MIDI data is digitised sound.

c) Animation can be used to emphasise the three-dimensional nature of objects.

d) Most commonly used format for graphics is .BMP or bitmap pictures.

e) If the sequence and timing of these multimedia elements can be controlled by the user, then one can name it as Non-Interactive Multimedia.

7.8 SOFTWARE IN MULTIMEDIA SYSTEMS

There is a range of software used for creating a multimedia product. The popular among them are:

- Presentation Documentation
- Multimedia Application Software
- Authoring software
- Animation software
- Web browsers and Editors

7.8.1 Presentation Documentation

Computer-generated presentations are quickly replacing the more traditional paper and slide-based approach. Presentation software uses graphics and data/ information from other software tools to communicate and make presentations of data to others. Modern presentation software packages provide facilities to create presentations that may be viewed using: a standard monitor, a data projector for a large group, hard copies of slides and transparencies hard copies of paper handouts. This type of software, including packages such as Microsoft PowerPoint and Claris AppleWorks, can incorporate text, numbers, images, audio, animation and video, and allow relatively fast generation of a series of linked screens.

Presentation software creates several different types of documentation to help the presenter. These include:

On Screen Presentations

These are a series of slides which take the place of slide projectors as well as overhead projectors. They can be displayed on a monitor or projected onto a screen by way of a data projector. The timing of the slides can be either manual from the click of a mouse or other peripheral or automatically timed.

Overhead Transparencies

The slides can be printed out onto overhead transparencies if required.

7.8.2 Multimedia Application Software

Application software is a computer program used for a specific task. Programs such as Microsoft Word, used to produce word processed documents, and Adobe Photoshop, used to generate complicated graphical images, can be used to generate content that can be incorporated into a multimedia project. A wide variety of specific software is available to generate and edit the various parts of a multimedia project such as video, sound grabs, images and animation. Some word processing software even allows the direct importation of video or sound.

7.8.3 Authoring Software

Multimedia authoring packages are designed to be relatively quick to use. These types of multimedia tools are called rapid application development tools or RAD tools. Authoring software allows the user to sequence and time the occurrence of events, determining which graphics, sound, text and video files are to be utilised at any given point in the final product. The software also allows the creator to determine the level of user interaction. Examples of this type of software include eXe, Courselab, Xerte, Macromedia Authorware, Macromedia Director etc. Authoring software is going to play an ever increasing role in education and training as employers and politicians increasingly look for low cost and more flexible solutions to their training needs.

7.8.4 Animation Software

There are many programs designed to create animations and to ease the burden of drawing the hundreds of cells or frames required for even a simple animation. Each has its own strengths and weaknesses but the majority provides the tools to produce an animation using one of the different animation techniques, including the traditional frame-by-frame, path-based, morphing or warping. Examples of animation software are Blender, Flash, Director, Image soft, 3D Max. There are countless others. However simple animations can be made simply and easily by using a GIF animator.

7.8.5 Web Browsers and Web Editors

Web browsers are used to look for information on the net while editors are used to create web pages. There are many web browsers but the most common two are Microsoft Internet Explorer and Netscape Navigator. Editors are used to create web pages and they range from text based such as Note pad and BBedit to WYSIWYG systems such as Dreamweaver and Frontpage. Pages on the WWW are written primarily in html (Hypertext Mark-up Language).

HTML files are plain text files and describe text formatting and are not a programming language as such, All html documents have the same structure with different codes, as the designer requires. Software that is WYSIWYG - what you see is what you get - implies that the contents of the screen are the same as the final printer or browser versions.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

5) What is the basic H/W and S/W required for creating a simple Multimedia Presentation?

....

....

....

....

7.9 INFORMATION COLLECTION IN MULTIMEDIA SYSTEMS

Collecting information for a multimedia project generally requires the digitisation of data presented in another form and will include a range of activities including: writing the notes, digitising audio and video, editing and generally gathering information.

Digitisation of Analog Data

Audio, video and images are often represented in an analog format. To be used in a multimedia presentation they must first be converted to digital. An ADC or analog to digital converter is the tool used in this process.

Digital Text and Numbers

For use in a computer, data needs to be converted into binary numbers. This has been achieved for letters, numbers and other characters by the development of a number of coding methods. One common coding system is known as ASCII, which stands for American Standard Code for Information Interchange. ASCII is a 7-bit code that represents 128 letters, numbers and punctuation symbols as a 7-bit binary number. The eighth bit in the byte is used as an error checking parity bit. For example, the letter ‘e’ is converted to the binary number 1100101 using the ASCII code, which can be stored by the computer. Other coding systems exist which also translate characters into the digital equivalent. EBCDIC or Extended Binary Coded Decimal Interchange Code is used by IBM mainframes. This system uses an 8-bit code, and allows 256 different symbols.

Digitisation of Graphics

Digitisation of graphics refers to the process of putting images into a format that computers can read and store.

Bit mapping, or memory mapping, is the relation-ship between the image on screen and the bits stored in primary and/or secondary memory. One or more bits must be stored for each pixel (picture element) drawn in an image.

Each dot on the screen is known as a picture element or pixel. The more pixels that can be displayed by a screen, the better will be the image quality or resolution.

Pixel: At the simplest level, 1 pixel is represented by 1 bit, where 1 means on, or black, and 0 means off, or white. For example, a black and white graphic 640 by 400 would contain 256,000 (640 x 400) pixels, and would require 256,000 bits to be stored in memory. That is, 32,000 (256,000 = 8) bytes of memory or 32.25 (32,000 =1,024) Kb.

Resolution is normally quoted in terms of the number of horizontal pixels times the number of vertical pixels, for example, 800 times 600.

Tones (colours or greyscale) are a progressive series of shades from white to black, and are used in graphics to add contrast and detail. Such an image however, requires more memory because each pixel must be described by an increased number of bits. For example, 2 bits of memory per pixel (00, 01, 10, 11) would produce 4 (2s) tonal colours, 3 bits per pixel would allow 8 (23) tones (000, 001, 010, 011,100, 101, 110, 111) and so on.

The term ***Bit Depth*** refers to the number of bits that describe each pixel in the image. A 16-bit image would produce 65,536 (216) different colours or tones. It is obvious that the more detailed the image, in colour and tonality, the more memory would be required for its storage.

Digitisation of Audio

Sound and other analog data is generally represented as a transverse wave, and can be converted to digital form by a process called sampling. The two important aspects of sampling are sampling size and sampling rate.

Sampling size refers to the number of bits used to store each sample from the analog wave. For example, an 8-bit sample can represent 256 (28 = 256) possible levels in a particular sample.

A higher sample size will result in increased accuracy, but higher data storage requirements.

Sampling Rate refers to the number of samples or slices taken of the analog wave in 1 second. The higher the sampling size, the better will be the representation of the initial analog signal.

Methods of Digitising/ Capturing Video Images

Capturing full motion video requires a video capture card to digitise the signal (unless using a digital video recorder, in which case it is already digitised) before storing on disk for later editing.

The standard PAL (phase alternate line) video signal used in India displays a frame rate of 25 frames per second. One frame of medium resolution and 16-bit colour requires approximately 1 Mb of storage space per frame. This translates to 25 Mb per second of video, or a staggering 1,500 Mb per minute.

Current personal computers cannot sustain a transfer rate between secondary and primary storage of 1,500 Mb per minute, so a number of solutions are applied including:

- Video data will be compressed during recording, using a codec.
- Decreased colour depth to fewer colours or even black and white shades requires significantly less memory.
- Decreased resolution reduces number of pixels to describe in each frame.

7.10 STORYBOARD FOR MULTIMEDIA SYSTEMS

A storyboard is an illustrated scene-by-scene layout of the multimedia presentation. It generally includes text notes and sketches of the most important parts of the presentation.

The components of any multimedia production, be they text, hypertext, sound, video or graphics, have to be combined in such a way that the result is a blended product in which each element supports all others.

Different types of storyboards, as listed below, help achieve this:

Linear – sequential, movement through project.

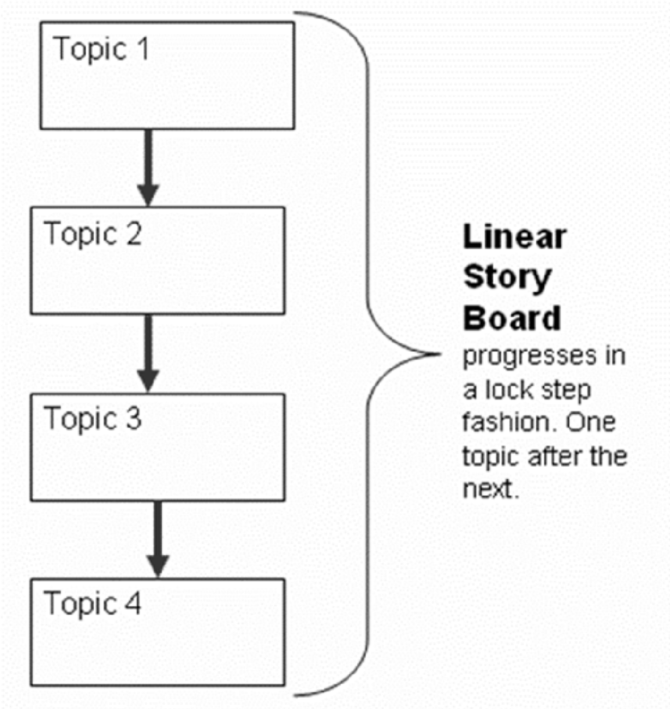


Fig.7.1: Linear Sequential Storyboard Format

Hierarchical – top down approach, which provides users with multiple choices at each stage of the project.

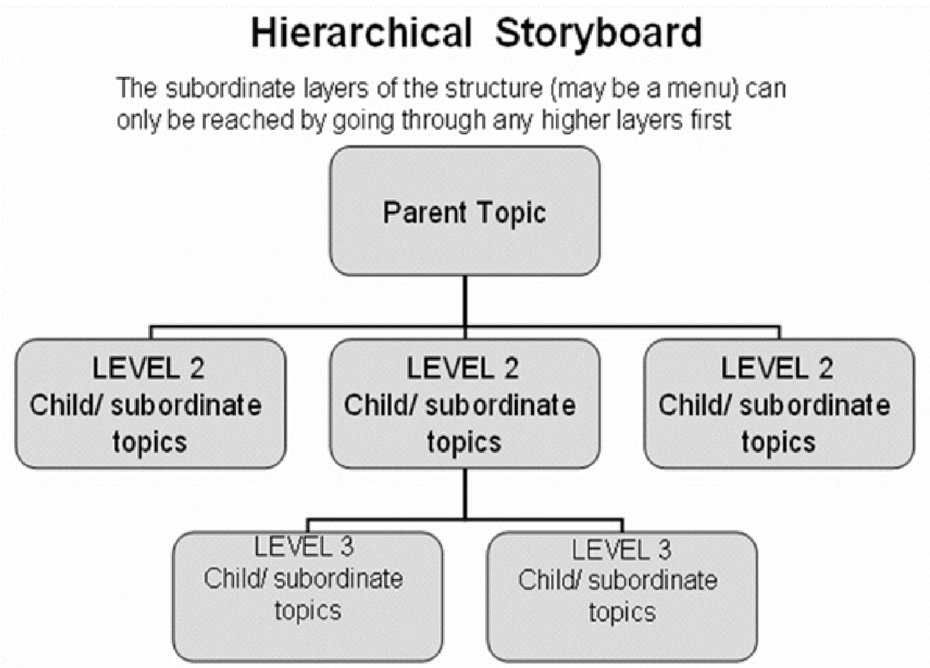


Fig. 7.2: Hierarchical – Top Down Approach Storyboard Format

Non-linear – the user can choose the path of navigation through the project without significant restriction. There is no particular structure inherent in the design.

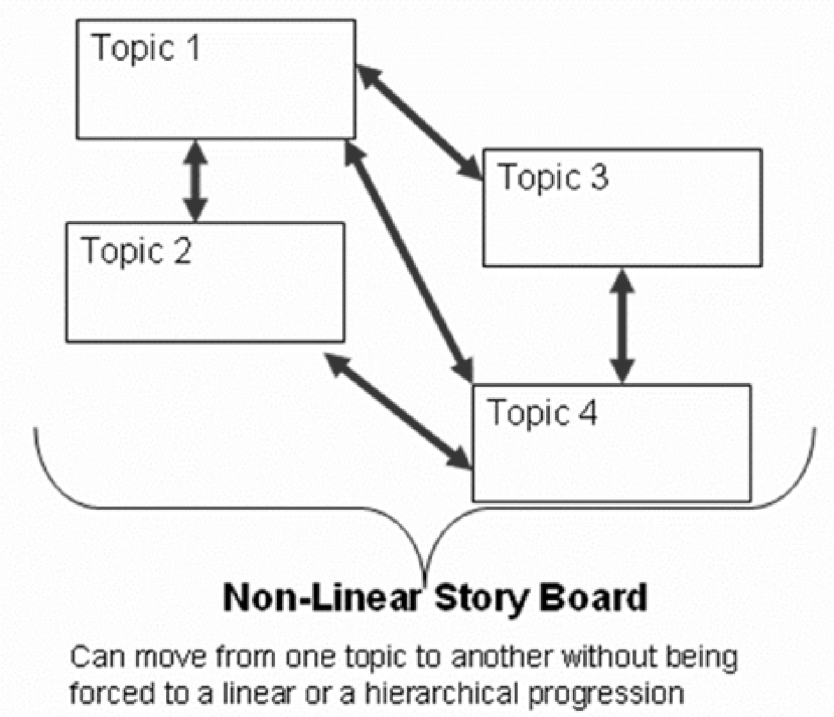


Fig.7.3: Non Linear Storyboard Format

Composite - aspects of each of the previous three types are incorporated.

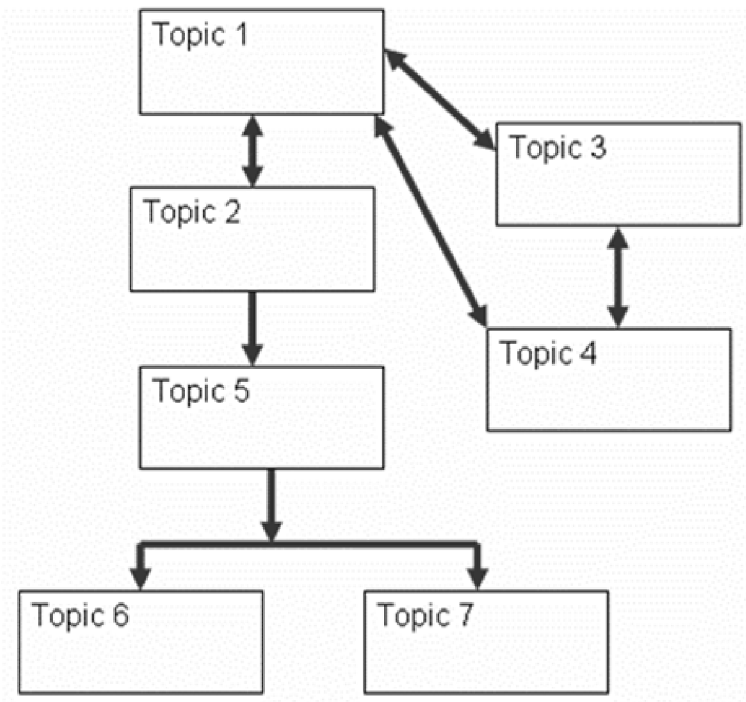


Fig. 7.4: Composite Storyboard Format

7.11 PROCESSING IN MULTIMEDIA SYSTEMS

Processing involves the modification of data in some way that makes it more useable. Processing in a multimedia system will include editing and compression. Some examples of processing include:

- Image processing (morphing, Warping)
- Animation Processing (Cell-based animation, Path-based animation)

Multimedia consists of information presented in various forms, including text, numerical, audio and so the creation of this material occurs in an integrated environment using, items from different sources, such as sound and video.

The integration of a range of forms of information leads to increased storage requirements. For example, a two hour movie contains so much sound and visual information that if stored without modification on a standard CD-Rom it would require over 300 disk changes during a single showing.

The solution to the storage dilemma comes from a mathematical process called compression. Compression is a method of removing redundant or repetitive elements from a file so that the file requires less storage space and thus less time to transmit. The algorithm that defines the process of compression and decompression is called a codec.

A concept that is often utilised in multimedia is called hypermedia. Hypermedia refers to the linking of data types. For example, clicking on a particular section of the screen may cause additional text to be displayed, or the user to 'jump' to a different section of the program. The term 'hyper' refers to a link between elements of the program, and so hypertext, for example, is text that is linked to other sections.

7.11.1 Image Processing

Image processing is another aspect of video work that demands processing power. There are two primary types that you need to be aware of:

Morphing is an image processing technique used for metamorphosis from one image to another. The idea is to get a sequence of intermediate images which when put together with the original images would represent the change from one image to the other. For example, changing a photo of yours into any another photo say, a tiger.

Warping refers to modification made to an existing image by stretching and resizing. In this case it is clear that you are still looking at the original image. For example, changing a picture of yours so that you have very large ears.

7.11.2 Animation Processing

Animation processing is required when presenting a series of graphics in rapid succession. The human eye experiences something called persistence of vision, and this is the principle upon which animation is based.

As one graphics frame is replaced by the next the screen is blank for approximately 500 milliseconds. However, the eye retains the last image until it is replaced and thus we see movement.

While all animation involves a series of images, there are several techniques that may be used:

Cel-based animation is a method of producing animation by the creation of a sequence of individual still images each produced on a separate cel. The term "cel" refers to an individual still image to be used in the animation sequence and comes from the more traditional hand-drawn images on separate sheets of celluloid, hence "cel".

Path-based Animation is a method in which the starting point, end point and the path to be followed by an object are defined. The software generates the object, as it moves along the defined path, to represent the animation. Path-based animation involves displaying the movement of objects onto a fixed background.

The background pixels do not change, only those for the moving object. This saves memory and processing time. Animation is achieved by drawing the object, wiping it and the drawing the object in a new position. This process is made easier because animation software can create objects between two objects in key position.

7.12 STORING AND RETRIEVING IN MULTIMEDIA SYSTEMS

The concept of storage and retrieval is very important for any multimedia project because whether the mode is text, video, audio or graphic there are a wide variety of formats available. Each format will have advantages and disadvantages.

Not all formats will be supported by any program so that means that you must find a suitable format from the list of supported formats. Also not just any format will do the job. you must chose the right format for the right job. For example jpegs are a compressed graphic format. They have great colours but cannot be altered without losing data. If image needs to be altered then a different format should be used.

The different modes and their formats are as follows:

Video: ***avi, mpeg, quicktime, real.*** The same as with other file types there are a large number of video formats. The main ones are avi, mpeg, quick time, real player, flash and shockwave. mpegs are quickly becoming a standard. Mpegs are used on VCD (mpeg 1) and DVDs (mpeg 3 which is a higher resolution than mpeg 1).

- AVI (Audio Video Interleave)
- MPEG (Motion Pictures Expert Group)
- Quicktime Apple digital video, also available on PC

Image: ***jpg, eps, pict, tiff, bmp, gif***: When choosing an image type two criteria are important: accessibility and file size. Below are formats:

- JPG (Joint Photographics Experts Group) A lossy compression format designed to reduce the size of full colour bit maps.
- EPS (Encapsulated Postscript) Uses vector graphics and often will only be interpreted by a printer. Therefore it cannot be displayed on screen.
- PICT Apple graphics format that is bit mapped or vector.
- TIFF (Tagged Image File Format) A bit mapped format, standard choice for scanned images.
- BMP (Bit Map) An uncompressed bit mapped format (Native to PCs).
- GIF (Graphics Interchange Format) Bit mapped format that uses a lossless compression algorithm. Limit of 256 colours.

Text and numbers: ***txt, doc, pdf, html, rtf***: There are not a large number of text formats. Below are the main ones:

- TXT ASCII text (only ASCII characters) or delimited text exported from a database
- DOC Microsoft Word file

- PDF (Portable Document Format) Proprietary digital format produced by Adobe which preserves format and look of text document.
- HTML - (Hypertext Mark-up Language) Text based language which a web browser uses to render text, images, audio and video.
- RTF (Rich Text Format) Used to produce a document that can be read by other word processors.

Audio: ***Wav, midi, mp3***. There are a wide range of audio formats. Below are three common ones.

Wav (wave file): digitised audio file of varying levels of quality which can be played directly by a sound card. This is a native PC sound file.

MIDI: (musical instrument digital interface) stores note information rather than sound which results in a small in a field size. However, it requires a synthesizer to replay.

MP3: Derived from audio layer 3 in a MPEG video. Currently requires a software player.

Compression and Decompression

These days compression of data is not only handy but is crucial to multimedia performance. In the following pages we are going to examine the various aspects of compression and decompression. Some of the important terms for this work are the following: compression, decompression, sip, archive, tar, mpeg, video compression, image compression, lossless, lossy, decompression, codec, coding redundancy, spatial redundancy, temporal redundancy, psycho visual redundancy.

Video Compression

Video compression is achieved by the use of a codec. A codec is a compression/ decompression algorithm. A number of video codecs are in use, including Indeo (created by Intel) and MPEG-1, 2, 3, 4 and 7, created by the Moving Pictures Expert Group.

Image Compression

Image compression involves minimising the size of a file without degrading the quality of the image to an unacceptable level. The reduction in file size allows more information to be stored in a given amount of disk or memory space. Smaller file size also reduces the time required for information to be sent over the Internet. Compression schemes can be:

Lossless, where the decompressed data is identical to the original data with nothing lost in the translation, or,

Lossy, where the decompressed data is slightly different from the original, hopefully in insignificant ways, though some data is sacrificed for the sake of compression.

Decompression

Decompression (expansion of a compressed file) is only possible if lossless compression has been utilised. It is only possible to expand a compressed file format to the original data set if the compression algorithm does not discard information in the original compression process. Compression strategies can take advantage of four kinds of redundancy:

- Coding redundancy, which relies on the fact that not all data will occur with the same probability.
- Spatial redundancy, which occurs because pixels that are near each other are likely to be similar to each other.
- Temporal redundancy, which occurs because pixels in consecutive frames of a video are likely to be similar, for example a "talking head".
- Psycho visual redundancy, which occurs because the human visual system is better at detecting changes in luminance (brightness) than chrominance (colour).

Image Format Choice

Web graphics are very different to printed graphics. They must have a small file size and may be interactive. Web graphics require to use RGB (red green blue) colour scheme designed for monitors while a print media requires CMYK. Common web formats are:

- GIF
- JPEG (also called jpg)
- PNG

Images in other formats can be very large and need to optimise for the web and changed into one of these formats.

As such it is important to consider the most appropriate compression technique to utilise when preparing an image for the World Wide Web or any other multimedia project. Consider the criteria like:

- The gif format only uses 8-bit colour (256 colours) but it's a 'lossless' format and thus detail is not reduced. Gif images are the best way to render artwork such as line drawings, charts and maps, particularly if the number of colours can be reduced. However GIFs will give a larger file size if the picture has too many colours. For an image with a large number of colours a jpeg is preferable.
- Jpeg images use 24-bit colour (16,777,216 colours) but the compression algorithm is lossy. This means that resolution on a computer screen is lost on compression, this may not be noticeable but a hardcopy image would show the degradation. Jpeg images are used for photographs (screen but not good printed). A reasonable quality photo in jpeg is commonly around 20-30 Kb where the equivalent in gif could be 100-300 Kb. With jpeg files the quality required is set when the image is saved.

7.13 ISSUES RELATED TO MULTIMEDIA SYSTEMS

There are many issues that have arisen from the development of multimedia. Some of them are to do with the violent nature of some content and the appropriate use of the internet but many of the issues stem from other areas such as copyright, the way technology has moved ahead so quickly and the general impact of those issues upon the individual and society. Lastly individual rights must also be protected along with those of companies and this is covered by concerns over data integrity (how accurate is the data that we are being fed. Is it accurate? Is it honest? Has there been a corruption of the data in any way.

The following are the important issues to be considered while developing the Multimedia Systems:

- Copyright
- Appropriate Use of The Internet
- Merging Technologies
- Data Integrity

7.13.1 Copyright and Multimedia Systems

Under the Copyright Act it is illegal to reproduce a literary work or to make an adaptation of that work without the permission of the author. "Literary work" includes software:

- The author or producer has the right to expect payment for their work. Time and effort are worth money.
- Computer programs are a particular problem because they are relatively easy to copy and modify. This becomes an ethical issue for users who have the technology to break copyright and avoid detection.
- If a multimedia developer does not acknowledge where images and sound files come from then they are breaking copyright.
- Shareware allows a user to try the software before buying. But shareware is not freeware and if you continue to use the software after the allowable trial period is over then you are breaking copyright.

Software licensing

There are a number of different license formats as listed below:

Single user - the price allows the use of one copy of the software on one computer.

Site license - the price allows the use of the software on a specified number of computers.

Freeware - software may be used freely and copied for other users

Shareware - software may be trailed for a period of time before a licensing fee is paid for a single or site license.

7.13.2 Appropriate Use of the Internet and Multimedia Systems

Appropriate levels of usage of Internet and other content involve the concepts of ethics and moral behaviour. Ethical issues relate to questions of morals. Is it right or wrong to proceed with a specific action? If you behave ethically or morally then you are behaving in a manner that most people would regard as the right way to act. Correct ethical or moral behaviour would normally reflect the norms of society. It is important to understand that norms change. Many people may feel that what was considered unethical behaviour 30 years ago is normal in current times. Widespread access to the Internet keeps the issue of ethical and moral issues active. People are being exposed to all sorts of material. Who should have access to different material? How much censorship should be imposed on the Internet? How much censorship is possible?

7.13.3 Merging Technologies and Multimedia Systems

The lines dividing computers, telephones, radio, television, voice and data are blurring. This merging of content is called "***digital convergence***" which is where content can be stored as a digital file rather than the more traditional analog form:

- It is possible to use the Internet for telephony using a PC with a sound card, microphone and a standard Internet connection.
- Desktop radio broadcasting is also available on the Internet. Sites such as Real Audio's www.real.com provide a compressed sound source that can be downloaded and played in real-time. Television has also made it to the Internet with WebTV. However, bandwidth limitations produce a poor quality picture.
- Audio and video files are available in two forms, downloadable files or live content called streaming audio or video. Streaming audio or video allows you to listen to or view the sound/ video while it is being downloaded.

7.14 DATA INTEGRITY IN MULTIMEDIA SYSTEMS

Data integrity means that the data is accurate, consistent, and up-to-date. As the amount of information available on the Internet increases, the issue of data integrity becomes more important. Multimedia systems used by educational institutions must use a data source that has integrity.

The reliability of material is not always easy to establish. There are many unreliable sources of information on the web. It is important that only reliable sources be quoted and used for research.

The publishers of the Encyclopedia Britannica could choose to use data from questionable sources but do not, instead preferring to spend the necessary time on research to ensure the integrity of their product.

It is important that those researching material on the web ensure that all source data is cross-referenced to ensure its accuracy.

7.15 CAREER PATH IN MULTIMEDIA

Business: Multimedia designers can be used in business application in a variety of ways like presentations, training, marketing, advertising, product demos, databases, catalogues and networked communications. They can be successfully engaged in video conferencing also.

Advertising: Imaginative and attractive advertisements can be made with the combination of text, pictures, audio and video. Multimedia designers have a big role in creation of advertisements. A product is well received by a customer if it is supported by a good multimedia advertisement campaign.

Gaming and Graphic Design: They are perhaps making the maximum use of multimedia. No computer game is complete without elaborate computer graphics, be it a arcade game, strategy based game or sports game. A computer game with good graphics is more enticing to play then a game with less or bad graphics. Multimedia designers have a great role in making a game successful.

Product Design: Multimedia can be used effectively for designing a product. First its prototype can be made before actually making the product. Multimedia programmers can be employed in this work.

Education and Training: It is perhaps the need of the hour requirement for the multimedia. Topics which are difficult to understand by reading the text can be made simple with the help of multimedia. Time is coming when the multimedia lessons will take place of classroom teaching. Students can repeat a lesson as many times until he understands the concept. Multimedia designers have big role in all this work.

Leisure: Multimedia can also be used for entertainment. Most of the cartoon films are made with the help of multimedia. It is used in scientific movies to give special effects like animation, morphing etc. Actors created by combining different frames with the help of multimedia can replace the actual actors.

With this Multimedia, students will gain creative skills and technological knowledge leading to many exciting career opportunities including in the fields of electronic publishing, web design, information architecture, human-computer interface, design, multimedia design and production, 3-D animation, computer games, exhibition design, scientific and medical visualisation and special effects for film and television. Escalating demand for these skills by the Creative Industries provide students with exciting options for professional placement and eventual employment nationally and internationally.

7.16 SUMMARY

Multimedia is the media that uses multiple forms of information content and information processing (e.g. text, audio, graphics, animation, video, interactivity) to inform or entertain the user. *Multimedia* also refers to the use of electronic media to store and experience multimedia content. Multimedia is similar to traditional mixed media in fine art, but with a broader scope. The term "rich media" is synonymous for interactive multimedia.

Multimedia may be broadly divided into linear and non-linear categories. Linear active content progresses without any navigation control for the viewer such as a cinema presentation. Non-linear content offers user interactivity to control progress as used with a computer game or used in self-paced computer based training. Non-linear content is also known as hypermedia content.

Audio is an important component of multimedia which can be used to provide liveliness to a multimedia presentation. MIDI is Musical Instrument Digital Interface. MIDI is a communication standard developed for electronic musical instruments and computers. A digital image is represented by a matrix of numeric values each representing a quantised intensity value. A *bitmap* is a simple information matrix describing the individual dots that are the smallest elements of resolution on a computer screen or other display or printing device.

Animation is created from drawn pictures and video is created using real time Visuals. Animation is possible because of a biological phenomenon known as *persistence of vision.* The different techniques used in animation are cel animation, and morphing.

7.17 ANSWERS TO SELF CHECK EXERCISES

1) ***Multimedia*** is the field concerned with the computer-controlled integration of text, graphics, drawings, still and moving images (Video), animation, audio, and any other media where every type of information can be represented, stored, transmitted and processed digitally.

   A ***Multimedia Application*** is an Application which uses a collection of multiple media sources e.g. text, graphics, images, sound/audio, animation, movie clips, and/or video.

2) Text, graphics, images, audio, video, animation etc.

3) Some of the major applications of Multimedia are: -

   a) To create Computer Based Training.

   b) To create interactive web contents with sound and movie files.

   c) To create Multimedia presentation for marketing and promotional aids.

   d) To create informative Kiosk etc.,

4) (a) True (b) False (c) True (d) True (e) False

5) A high end PC/ workstation with a reasonable VRAM, RAM and CD-ROM Drive and Speakers as for as Hardware is concerned. The basic required configuration is: RAM- 1 GB minimum, SPEED-1 GHz, STORAGE- Depending on the work your doing you may require anything from a 10GB available space upwards. For the software any image editing software like Adobe PhotoShop, animation software like 3D Max or Blender, some sound editing software like Sound Forge or Audacity and lastly Authoring software like Macromedia Director.

7.18 KEYWORDS

FireWire : A standard high-performance serial bus for connecting digital devices together or to a computer.

Hypermedia : An extension to hypertext providing multimedia facilities, such as those handling sound and video.

Hypertext : A software system allowing extensive cross-referencing between related sections of text and associated graphic material.

Pixel : Abbreviation of picture element refers to a minute area of illumination on a display screen, one of many from which an image is composed.

Raster graphics : A dot matrix data structure representing a generally rectangular grid of pixels, or points of colour, viewable via a monitor, paper, or other display medium.

Vector graphics : Use of geometrical primitives such as points, lines, curves, and shapes or polygons—all of which are based on mathematical expressions—to represent images in computer graphics.

7.19 REFERENCES AND FURTHER READINGS

Chapman, Nigel and Chapman, Jenny. *Digital Multimedia*. John Wiley & Sons, 2004. Print

Halsall, Fred. *Multimedia Communications: Applications, Networks, Protocols and Standards.* Addison Wesley, 2001. Print

Heath, Steve. *Multimedia and communications Technology*. Elsevier (Focal Press), 1999. Print

Hilman, David. *Multimedia Technology and Applications*. New Delhi: Galgotia, 2008. Print

Rao, Kamisetty, Bojkovic, Zoran and Milovanovic, Dragorad. *Introduction to Multimedia Communications: Applications, Middleware, Networking*. Wiley, 2005. Print

Steinmetz, Ralf and Nahrstedt, Klara. *Multimedia Applications*. Springer, 2004. Print

Weixel, Suzanne. *Multimedia Basics*. New Delhi: Thomson, 2004. Print

BLOCK 3 NETWORK FUNDAMENTALS

Introduction

The rapid development of Information Communication Technologies (ICT) has resulted in proliferation of computer networks in all walks of our life. The importance of computer networks and telecommunications has grown tremendously in the last one decade. Technological advances in communication have ushered in a new era not only of computer power but of access to information services through telecommunication networks. We are fast moving towards an information society which critically depends on networked environment for its sustenance. An understanding of concepts underlying telecommunications and networking has become essential for the LIS professionals in a networked society.

This Block has five Units:

Unit 8 on **Network Topology** deals with four basic aspects of data networks, viz. topology, media access control (MAC) protocols, address resolution and routing.

Unit 9 is **Communication Protocols and Network Addressing** which deals with two distinct but closely related aspects: Internet communication protocols and network addressing.

Unit 10 covers different aspects of **Protocol Architectures**

Unit 11 dealing with **Network Applications and Management** coveres main user applications that runs on Internet and network management. The applications discussed include non-real time ones like text and multimedia messaging and real time ones like interactive television and music on demand.

Unit 12 on **Network Security** discusses information security issues in a networked environment.

Blank Page

UNIT 8 NETWORK TOPOLOGY

Structure

8.0 OBJECTIVES

After going through this Unit, you will be able to understand and appreciate:

- What is meant by network topology;
- Difference between physical and logical topologies;
- Basic topologies like star, bus, ring, tree and hybrid;
- Why star topology is popular;
- Topology related network components like hubs and switches;
- Different logical topologies;
- Ethernet and token passing ring protocols;
- Merits and demerits of Ethernet and token passing ring protocols;
- Why address resolution is required and how it is performed;

- Purpose of domain name servers an address resolution protocol;
- The need for encapsulation;
- Router connectivity and the functioning of routers in Internet; and
- Routing algorithms and the associated performance parameters.

8.1 INTRODUCTION

As you are aware, computers world over are interconnected via the Internet. The connection to the Internet happens in a variety of ways. For example, a home computer is usually stand-alone computer connected to the Internet via a dial up telephone line. In homes where there is more than one computer, they may be interconnected to form a home computer network. In such cases, one of the computers acts as the Internet link. It is called a proxy Internet server. Other computers access the Internet via the proxy server. The home proxy server also accesses the Internet via dial up line usually. In some rare cases, a home may use a leased line to access Internet. Computers in organisations and offices are generally interconnected locally. The local network, called Local Area Network (LAN) is then connected to the Internet via a gateway using a leased line. The gateway may be a firewall or a proxy server. There are a variety of ways in which LAN computers may be interconnected.

Network topology refers to the study of geometric properties of the way in which the computers in a network are interconnected. A generalised network connection is shown in Fig. 8.1. Here four computers are attached to what is called a network cloud. The network cloud is a graphic symbol that denotes a network without specifying the geometry or other interconnection details. Network cloud is a black box that hides the interconnection details from the viewer.

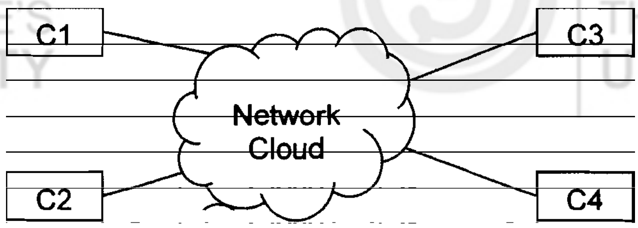


Fig. 8.1: A network cloud

There are a variety of ways in which these computers or nodes, as they are often called, can be interconnected physically inside the network cloud. For example, they may be interconnected as a daisy chain; say, C1-C2-C4-C3; or to a central switch as in the case of telephones being connected to an exchange. In a daisy chain connection, information moves from node to node in the order in which the chain is formed. In the above example, data moves from C1 to C2, C2 to C4 and C4 to C3 and vice versa for reverse flow. When a switch is used, a direct connection between two computers is established as in the case of calling and called subscribers in telephone communication. Network topology deals with the study of the way in which the computers are connected inside network cloud.

8.2 PHYSICAL AND LOGICAL TOPOLOGIES

You may be aware that in data networks like the Internet, data moves in packets. A long string of text is spilt into small packets, say 1024 bytes long, and sent over the network hopping from node to node. Accordingly, these networks are called packet switched networks or packet data networks (PDNs). Interestingly, the way in which packets move in a network may not correspond to the way in which the computers are physically connected for a variety of reasons. One important reason is traffic management and routing. The idea is very similar to the vehicular traffic management. If a road is congested and there is traffic jam, one tends to take a different route even though the alternative route may be longer. Depending on the jam, the traffic may be diverted for quite sometime. Similarly, if a link is congested in a network, an alternative route may be chosen to forward the packets. Route is chosen independently for every packet depending on the traffic conditions. The path taken by a packet for traversing form a source computer to a destination computer is known as logical path. Obviously, there may be a number of logical paths in a network. The collection of all such paths is called the logical topology of the network. The physical links constitute the physical topology of the network. In essence, the logical topology refers to the way in which packets travel from a source to a destination, whereas the physical topology refers to the actual physical interconnection of the computers in the network. Physical topology is also referred to as *real* topology and the logical one as *virtual* topology.

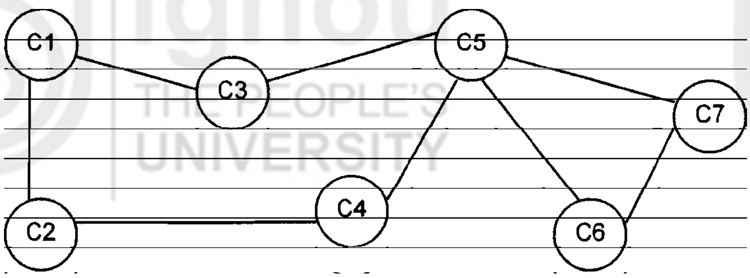
Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) A packet switched network uses a packet size of 2^{11} bytes. Determine the number of packets to be transmitted to transfer a file of size 1 MB.

2) Consider the physical topology given below:



Enumerate the number of logical paths between C1 and C7.

3) Can the packets of the same file travel via different logical paths? Do you foresee any problem in this case?

....

....

....

....

8.3 FULLY CONNECTED TOPOLOGY

Data communication involves computers in one part of the world being able to contact and communicate with computers in other parts of the world. For example, a home computer connects to different websites at different times. It is not just fixed one-to-one connection. It is multi-point communication connecting different to destinations at different times. For this purpose, a computer needs access to all the other computers that need to be contacted. One way of achieving this is to establish direct connections between the source computer and all the destination computers. In this case, we need as many links as there are destination computers. For example, if a home computer were to connect to 10 different web sites, then we would need 10 different links connecting the home computer to each web site. As you know, this obviously is not the case. However, if every computer in the world were to be connected to every other computer like this, then we need a very large number of links. A network formed this way for five computers is illustrated in Fig. 8.2. Here, every computer is directly connected to every other computer. Networks with this kind of connectivity are said to have fully connected topology. There are 10 links in Fig. 8.2. The links are assumed to be full duplex in the sense that they are capable of transporting information both ways. If the links were unidirectional (half-duplex) as in the case of optical fibres, we would need twice the number of links for two-way communication. The number of links required in a fully connected network becomes very large even with moderate number of computers. For example, we require 1225 links for fully interconnecting 50 computers.

In a general case with N computers, $N(N-1)/2$ links are required as reasoned in the following.

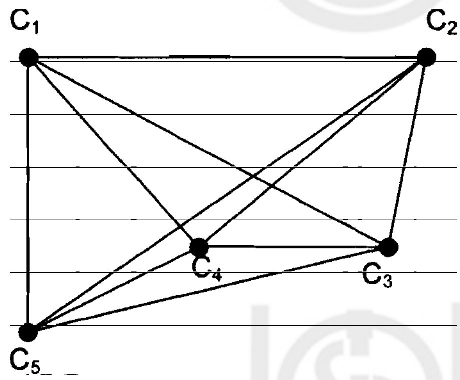


Fig. 8.2: Fully connected topology for five computers

Let us consider the N computers in some order. In order to connect the first computer to all other computers, we require $(N - 1)$ links. With this, the second computer is already connected to the first. We now need $(N - 2)$ links to connect the second computer to the others. For the third computer, we need $(N - 3)$ links, for the fourth $(N - 4)$ links, and so on. The total number of links N works out as follows:

$$L = (N-1) + (N-2) + \ldots + 1 + 0 = N(N-1)/2$$

Establishing separate and direct communication links connecting each computer to every other computer as shown in Fig. 8.2 is very expensive and is impracticable. Hence, this is just not done.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) How many half-duplex communication links are required for fully connected topology with 10 computers supporting full-duplex communication?

5) Draw a fully connected topology for two-way communication with four computers using fibre optic links. How many links are there in your network?

....

....

....

....

8.4 STAR TOPOLOGY

Star Topology is basically a physical topology. As the name implies, the topology looks like a star in the sky with rays emanating from the central point in all directions. It is a centralised topology where all computers are connected to a central point, which we call as *star point.* The topology is depicted in Fig. 8.3. This topology is easy to administer and maintain. The links can be tested and repaired from the central star point. This is one of biggest advantages of star topology. The topology is also a robust one. If a link or a computer fails, the rest of the network is not affected. As a result, this topology is very popular and is used extensively.

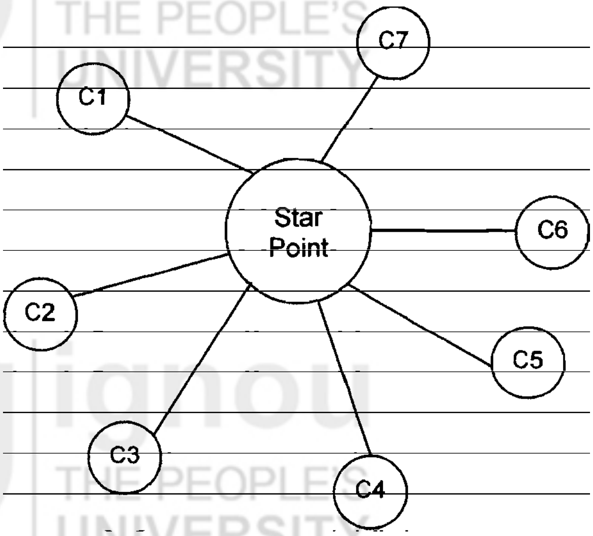


Fig. 8.3: Star topology

However, if the star point fails, the whole network fails. This is a disadvantage. Special care is taken at the time of designing the star point to make it very reliable.

Strictly speaking, physical star topology does not imply any logical topology. The logical topology is dependent on how the star point has been designed. This, in fact, makes this topology very attractive as it offers the flexibility of easy maintenance on the one hand and permits different logical topologies to be implemented on the other. Logical topology, as we mentioned earlier, defines the way in which a packet traverses form a source computer to a destination computer. Usually, there is set of rules that govern the exchange of packets between computers. Such a set of rules is called a protocol. There are many protocols used in networks. We learn more about protocols later in this unit. The star point of star topology can be designed to implement a variety of protocols such as Ethernet protocol, token ring protocol and a switch. Usually, either a hub that implements Ethernet protocol or a switch that permits switched connections is used as the star point. We learn about hubs and switches in the next section.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

6) Enumerate the advantages of star topology.

....

....

....

....

8.5 HUBS AND SWITCHES

As mentioned in the previous section, hubs and switches are used to implement logical topology in a physical star topology. Hub is a terminology used in several contexts in networks. In satellite networks, hub means a special ground station with which small satellite terminals communicate. In the context of LAN, hubs are used to implement two different logical topologies: Ethernet and token ring. But in most of the textbooks, the word hub is used to denote Ethernet hub, i.e. the hub that implements Ethernet logical topology. You must, however, be aware of the existence of different types of hubs. A hub is sometimes loosely called a concentrator as it connects to all computers in a network in star configuration. Ethernet is also implemented in bus physical topology that we discuss in the next section. Discussions on Ethernet in this section also apply to bus topology implementation. In fact, Ethernet was first implemented using bus topology. Later, it was implemented using hubs in star topology. At present, most of the Ethernet implementations are based on star physical topology using hubs.

Recall that logical topology corresponds to the way packets are transported within a network. Ethernet transports packets in much the same way information is exchanged in a group discussion among people. What happens in a group discussion? One who has something to say starts speaking and others listen. In the same way, in Ethernet topology, a computer starts transmitting whenever it has a packet to send. Other computers listen. Since any computer in the network can start a transmission Ethernet is called multiple access (MA) scheme. As happens in a group discussion, sometimes more than one computer may start a transmission simultaneously. What happens when two persons start speaking? Both of them go quiet and one among them starts afresh? Similar thing happens in Ethernet. When more than one computer start transmission simultaneously, we say that a collision has occurred. The computers that transmit simultaneously detect the collision, go quiet and follow a predetermined procedure to start the transmission afresh. Hence, Ethernet is a collision detection (CD) scheme. In a group discussion, if someone is already speaking, another person does not start to speak. Similar thing happens in Ethernet where detection of an ongoing transmission is called carrier sense (CS). On the whole Ethernet is a CSMA/CD scheme.

Let us now see how a hub helps implement Ethernet. Hubs come in 4, 8, 16, 24 and 48-port configurations. One computer can be attached to each port. Each port has provision for input/output and power connections. At the computer end there is a network interface card (NIC) that connects to the hub port. One of the ports is specially designed to be able to attach to another hub, thus allowing cascading of hubs. Cascading is useful when clusters of computer are located in nearby geographical areas. For example, an organisation spread over multiple floors of a multi-storey building, may use one hub per floor and cascade them so that computers in different floors can communicate with each other. Fig 8.4 shows a schematic of cascaded hubs with four ports each. Port 4 is specially designed to connect to another hub. Port 4 of the hub in Floor 1 is connected to Port 1 of the hub in Floor 2. Port 4 of the hub in Floor 2 is connected to Port 1 of the hub in Floor 3. Thus all the three hubs are cascaded in a daisy chain fashion. You may note that only two computers can be connected to the hub in Floor 2. If no cascading is used, the special port can be used to connect a computer. This is shown in the hub in Floor 3. Hubs that have provision for cascading are also called *stackable hubs*.

The internal mechanism of an Ethernet hub forwards any incoming packet from any computer to the output lines of all other computers as well as to the output line of the sending computer. In this sense, the hub acts as a broadcaster.

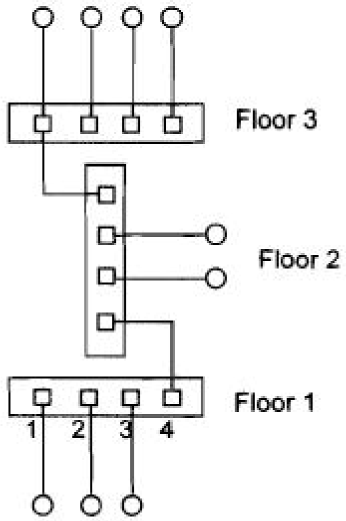


Fig. 8.4: 4-port cascaded hubs

This broadcast allows all the other computers and the transmitting computer to listen to the ongoing transmission. The transmitting computer is able to detect a collision by monitoring its own output line. If a bit received on its output line is not the same as the one sent by it, the computer knows that a collision has occurred. The Ethernet NIC in the computer senses carrier and detects collision. The hub enables multiple access feature as all the computers are connected to it and anyone can start a transmission. This is Ethernet hub implements CSMA/CD protocol. An important requirement of a hub-based design is that all computers connected to the hub must operate at the same speed.

The star point of star topology could be a switch. A switch is like a telephone exchange. The switch examines the destination address in an incoming packet and routes the packet to the appropriate outlet much as the telephone exchange examines the dialled number and routes the call to the appropriate destination. Much as the way telephone exchanges are interconnected, switches can be interconnected to route packets to computers that are not local. In fact, this is how most of the Internet connections work. An important advantage of the switch when compared to hub is that different computers can operate at different speeds. Of course, the source and the destination computer pair must operate at the same speed.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

7) What are the different logical topologies that can be implemented by hub?

....

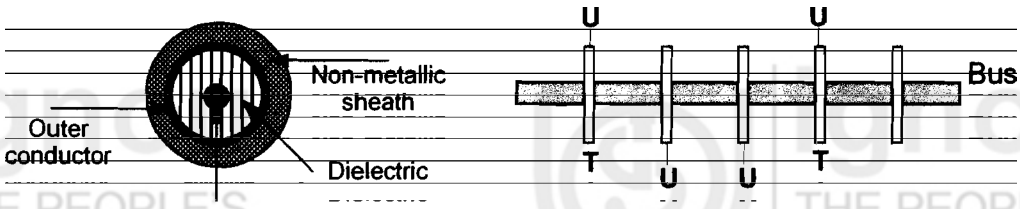
....

....

....

8.6 BUS TOPOLOGY

Bus is a cable laid linearly. Imagine a coil of cable unrolled, stretched and laid from end-to-end in a linear fashion and it becomes a bus. The cable is of coaxial type. Coaxial cable has one central conductor with a surrounding metallic shield. The central conductor and the shield are separated by a dielectric medium. Dielectric, as you may know is an insulator that electrically separates the inner conductor and the outer metallic shield. A thick non-metallic sheath further protects the central conductor and the shield. Fig 8.5(a) depicts a coaxial cable. The outer conductor (the metallic shield) is usually grounded and acts as an electromagnetic shield.



Inner central conductor

U = User node T = Tap

(a) Coaxial cable

(b) Bus topology

Fig. 8.5: Coaxial cable and the bus topology

With the outer conductor grounded, the cable essentially has one single conductor that carries current. When laid out as bus, the central conductor is like broadcast medium. Since the central conductor acts as a broadcast medium, it is called ether, the space, and the network as Ethernet. The bus topology is shown in Fig. 8.5(b). The bus cable is open at both ends and needs to be terminated with suitable terminations for proper electrical operation. Taps are essentially screws driven halfway into the central conductor. Taps are connected to the NICs of user nodes or computers. Thus, all computers are connected to the central conductor. Tap connection to the central conductor is passive and failure, if any, in one of the computers does not affect others in the network. But if the cable breaks at any point, the entire system is affected. It is not easy to detect and locate the cable break and a special device called *time domain reflectometer* is used for this purpose.

The NIC implements CSMA/CD protocol described in the previous section. Since all NICs are connected to the central conductor, there is multiple access (MA). They are able to sense (Carrier Sense) and listen to the transmission that is taking place on the bus. Any collision can be detected (CD) and necessary action taken to resolve the same.

The first version of Ethernet was implemented using a thick coaxial cable specified as 10Base5. The specification 10Base5 means the cable operates at 10 Mbps speed, uses what is called baseband modulation and can have a maximum length of 500 meters. This Ethernet version was called thick Ethernet. True to its name, the cable was thick and difficult to handle. Later another version called thin Ethernet was introduced with the cable specification as 10Base2. This cable was more flexible and easier to handle. Bus based Ethernet is the forerunner of hub based Ethernet. Because of the problems faced in the maintenance of thick and thin Ethernet, hub based solution was invented. The cable used in hub based system is a twisted pair and the specification is 10BaseT. Twisted pair cables cover a maximum length of 100 m. They are like telephone cables and are easier to handle. Some high-speed implementations of Ethernet used optical fibre. Then, the system specification is like 100BaseF. Optical fibres cover a distance of about 2000 m.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

8) What is the operating speed and the maximum distance covered by 20Base3 Ethernet system?

....

....

....

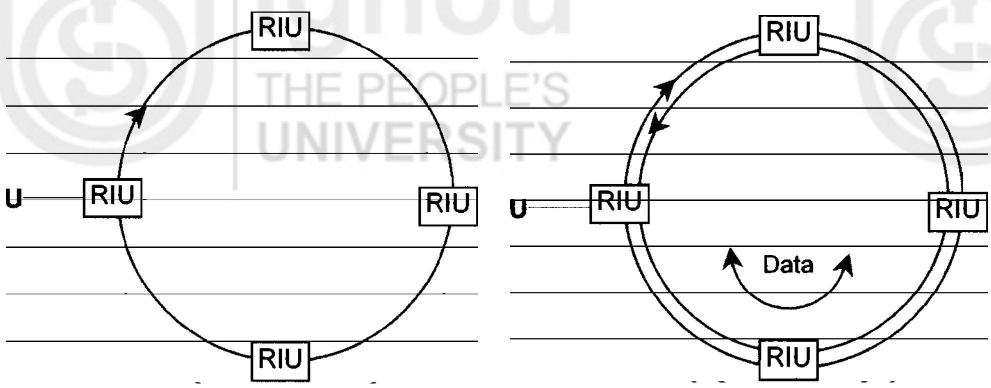
8.7 RING TOPOLOGY

In ring topology, a physical ring is formed by making point-to-point connection between computers. The computers themselves may not physically appear to be in the form of a ring, but electrically they form a ring. For example, two computers placed in adjacent rooms may be part of the ring. There is a circular communication path. Ring topology may be built around a single ring or two rings (dual ring).

Ring topologies are depicted in Fig. 8.6. The equivalent of tap in bus topology is Ring Interface Unit (RIU) in ring topology. User computers are attached to the RIUs. Unlike passive bus taps, RIUs are active units. Being active units, their failure rate is higher than passive taps. If a bus tap fails, only the concerned computer is affected. But if a RIU fails, the entire ring operation is affected. Hence, special considerations are required in ring topology to handle failures. In fact, failure management complicates the ring design and for this reason ring topology is not very popular.

Single ring topology shown in Fig. 8.6(a) usually uses bi-directional medium like copper wire. In case a segment of the ring or a RIU fails, the ring is folded back by the two end RIUs and the ring form of functioning continues. Dual ring configuration shown in Fig. 8.6(b) is generally adopted in the case of optical fibre design. As you may know, optical communication is naturally unidirectional as light that acts as the carrier of information is launched at one end of the fibre and received at the other. In dual ring, information travels in opposite directions in the two rings.

As in the case of bus topology, it is difficult to implement and maintain a ring structure physically. Hence, logical ring structure is often implemented using a ring hub in physical star topology. The logical topology in a ring network is called token passing ring or simply token ring.



(a) Single ring (b) Dual ring

Fig. 8.6: Ring Topology

A token circulates on the ring when it is idle. Token is a special bit pattern like 01111111 (zero followed by 7 ones). It is ensured that this pattern does not occur in the data transmitted by the computers. A special coding scheme is used to prevent the occurrence of token pattern in the data. A computer that has data to transmit seizes the token and starts transmitting the data. Imagine a circular formation of people around a round table participating in a group discussion. A ball is given to the group. The ball is passed around from one person to the next. The rule is that a person can speak only when he/she holds the ball. When a person wants to speak, he/she retains the ball when received and starts speaking. Once finished speaking, he/she passes the ball to the next person in the circular formation. The token ring scheme functions in a similar fashion. Two observations are important. First there is no collision in this scheme. Hence, the token ring functions more efficiently than Ethernet. However, Ethernet is more popular. Second, every station gets a chance to transmit as the token goes around. In Ethernet this is not the case. An unlucky station may keep on colliding again and again and may get a chance to transmit for quite some time.

On the ring, when a station starts transmission, the stations (computers) downstream listen to the transmission and monitor the destination address. If the destination address does not match one's own address, it passed to the next station as it is. When the data reaches the destination computer, the station copies and drains (takes away) the data from the ring. Thus a connection is established between the source and destination stations. After the source station has transmitted all data, it reintroduces the token on the ring. The token may now be seized by another station that has data to send. Obviously, the station next to the source station is the first one that can seize the token. Because of the use of token, the logical topology is called token passing ring. In optical fibre implementation, the logical topology is called Fibre Distributed Data Interface (FDDI). Ring networks operate at speeds of 10 Mbps to 1000 Mbps.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

9) Discuss the problem that would arise if a station on the ring seizes the token but fails to reintroduce the same on the ring after completing the data transmission. Suggest a mechanism to overcome such a problem.

....

....

....

....

8.8 MESH TOPOLOGY

Mesh is a complex interlaced structure realised by a bunch of point-to-point interconnections. Computers are interconnected without any geometric shape in mind. They are connected depending on the demand. They are somewhat like fully connected networks with some links missing. Therefore, they are sometimes referred to as partially connected topology. Mesh topology is illustrated in Fig. 8.7. In this topology, some of the nodes of the network are connected to more than one node by point-to-point links.

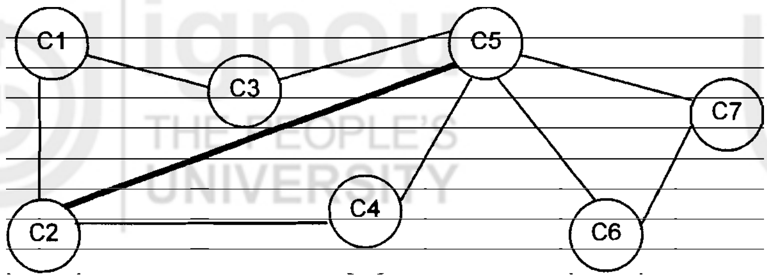


Fig. 8.7: Mesh Topology

Mesh is the actual topology that connects the wide area Internet all over the world. The nodes on the Internet invariably have mesh connectivity, i.e. multiple point-to-point links. Each node is connected to more than one node. The point-to-point links are established based on two considerations:

- Traffic between nodes
- Redundant routes.

Wherever heavy traffic is envisaged between two nodes, a direct link is established. Such links are sometimes called high usage routes. For example in Fig. 8.7, the link between node 2 and 5 is a high-capacity link to cater for heavy traffic between the two nodes. In addition, there must be at least one alternative route between a given source and destination i.e. a primary route and a secondary route. For example, the secondary route between C3 and C5 is via C3-C1-C2-C5. The secondary route may be used when C3-C5 link is broken. Often, there are many alternative routes between a source and destination. In such a case, the source and intermediate nodes must have some kind of intelligence to make a routing decision to select the best route at a given time. Most of the nodes on the Internet are routers that are capable of selecting the best possible route. Routing decisions take a definite amount of time. Hence, more is the number of intermediate nodes more is the time taken for the information to reach the destination. Usually, the shortest path with minimum number of intermediate nodes is chosen as the primary route. Only when that route is heavily loaded with traffic or unavailable for some other reason, an alternative route is chosen. We discuss routing and routing algorithms in more details in Sections 8.13 and 8.14 respectively.

8.9 TREE TOPOLOGY

Tree topology is also called hierarchical topology. In this topology, there is a clear hierarchy amongst the nodes. This is similar to a hierarchical structure in an organisation where there is a Chief Executive Officer (CEO) at the top, many senior level executives under him/her, junior executives reporting to seniors and so on. There are levels of responsibility and a clear reporting structure. Tree topology is modelled along the same lines. It is shown in Fig. 8.8. Strictly speaking, the structure is an inverted tree with the root node at the top and the branch and leaf nodes below the root.

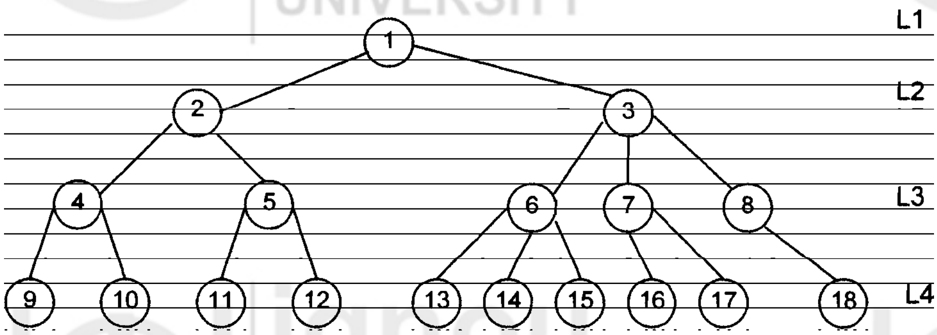


Fig. 8.8: Tree Topology

Four levels are shown in Fig. 8.8. The number of levels in a tree is called its depth. The top level node (Level 1) is called the root node and the bottom level nodes (level 4) as leaf nodes. The intermediate level nodes are called branch nodes at the designated levels. Each node except the root node has one point-to-point link connecting itself to the higher level node. Each node except leaf nodes has as many point-to-point links as there are braches attached to it. Leaf nodes, being at the bottom level have no branches. There is an interesting relationship between the number of nodes and the number of point-to-point links in a tree. The number of links is always one less than the number of nodes. Verify this in Fig. 8.8.

The number of branches that emanate from a node is called the branching factor (BF) of that node. If the branching factor is uniformly two for all the nodes, then the tree is called a binary tree. In Fig. 8.8, the left portion of the tree is shown to be binary. The tree itself is not binary as there are portions with branching factors that are not two. If all the nodes (except leaf nodes) in a tree have the same branching factor, then the same may be called the BF of the tree. If the BF of a tree is one, then the tree reduces to linear topology. The extreme right portion of Fig. 8.8 comprising nodes 8 and 18 represents the linear topology.

There is strict hierarchy of interaction amongst the nodes. Two nodes at the same level emanating from the same branch node above interact through that branch node. For example, nodes 6 and 7 in Fig. 8.8 communicate via node 3. If the nodes are attached to different branches, then the communication proceeds by traversing up the tree as much as required. For example, the communication between nodes 9 and 12 takes place via the route 9-4-2-5-12.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

10) What is the branching factor of a leaf node?

11) How many nodes and point-to-point links are there in a binary tree of depth 5?

12) Why is it a tree topology is also called a hierarchical topology?

....

....

....

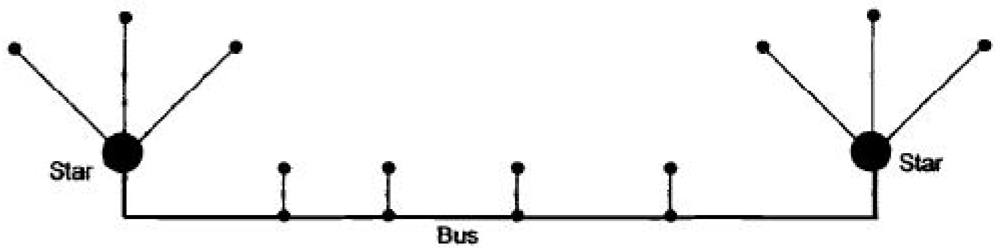
....

8.10 HYBRID TOPOLOGY

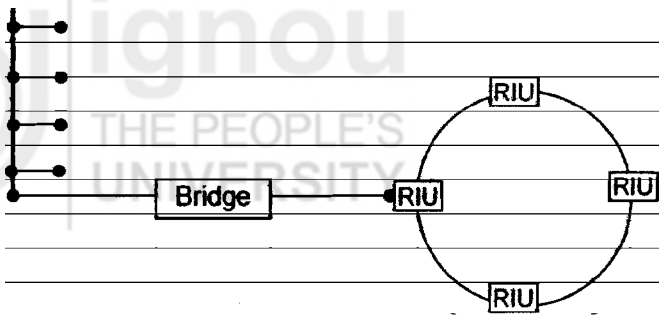
A hybrid topology is a combination topology in which two or more of the topologies discussed above coexist and work together. Figure 8.9 shows two example hybrid topologies. A large variety of hybrid topologies are possible. In Fig. 8.9(a), two star topologies are interconnected by a bus topology. This implementation is typical in campus networks like in a university. Each department may have star implementation while a bus or ring network may interconnect the departments. In general, such an implementation is adopted wherever the facilities that need to be interconnected are dispersed. For example, an office that is situated in different floors of a multi-storey building may use a hybrid structure. Computers in each floor may use Ethernet hub based star structure and the different floors may be interconnected by a bus structure. It is important to note that the entire set up works on the logical topology of Ethernet. Hence, interconnection of star with bus is seamless with no intermediate device.

Figure 8.9(b) shows a hybrid topology interconnecting a bus topology and a ring topology. In this case an intermediate device called a bridge is required. This is because the two topologies implement different logical protocols, viz. Ethernet and token ring. Bridge is an intelligent device. It implements two different logical topologies or protocols. In Fig. 8.9(b) it implements Ethernet on one side and token ring on the other. It appears like Ethernet NIC for the bus topology and as a token ring station attached to a RIU on the ring topology side. It is capable of recognising addresses at the logical level and mule *packets 1mm* one LAN to the other, it converts (reformats) packets of one protocol to that of another.

In some organisations, same topology is implemented in segments in geographical locations that are far apart All segments together are considered as one LAN.



Bus (a) Star-Bus-Star Hybrid Topology



(b) Bus-Ring Hybrid Topology

Fig. 8.9: Hybrid Topologies

In such cases, the signal from one segment to another traverses a long distance. In the process the signal level may be attenuated and may become too low to be recognised properly at the destination. To avoid this, devices called repeaters are used in between the segments. Repeaters are non-intelligent devices. They just amplify the signal level and make the signals strong so mat destination may recognise them property. In some cases, bridges may be used in place of repeaters for better management of LAN segments. Since, the bridges transform packets from one topology level to another, they automatically amplify the signals.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

13) In a university, LIS department has implemented star Ethernet LAN and the computer science (CS) department bus Ethernet LAN. The two departments are

in two different buildings that are far apart. How would you interconnect the two LANs?

14) In Question 13, how would your interconnection strategy change if the CS department had ring LAN? Explain the function of the device used.

15) Compare the features of bridges and repeaters.

....

....

....

....

8.11 MEDIA ACCESS CONTROL PROTOCOLS

Ethernet and token ring are two media access control (MAC) protocols that we have already studied briefly. In this section, we learn more details about them. Bus is a medium and so is ring, single or dual. Many computers are connected to these media. Any of the computers can access the medium that it is connected to. Since many computers access the same medium, we need some kind of protocol (a set of rules) so that there is an orderly access to the medium. Actually the access takes place in a controlled manner. Hence, the nomenclature MAC protocol is used.

In bus topology, Ethernet protocol is used. In ring topology, token passing ring protocol is used. Both these protocols are called multiple access protocols as they define sets of rules for multiple computers to access a medium. In hub based star implementations, the protocol depends on the type of hub used. The hub may be Ethernet hub or token ring hub.

Let us first see more details of Ethernet. As you already know, the Ethernet protocol rules are summarised in the acronym CSMA/CD. Consider the case when there is an ongoing transmission on the bus. The protocol must ensure that no new transmission starts at this stage. If it does, it will collide with the ongoing transmission and both transmissions will fail. No new transmission while there is an ongoing transmission is ensured by carrier sense (CS) mechanism. You may be aware that carrier refers to a high frequency transmission that carries the information signal. We have carrier frequencies in AM/FM radio broadcast. For example, 92.5 MHz is the carrier frequency of a FM radio station. The music signal is superimposed on this frequency and broadcast. Similarly, in LANs, information bits are superimposed over a carrier. The presence of carrier on the bus implies that there is a live transmission on the bus. Hence, any computer that has data to transmit will first sense the bus to see if the carrier is present. This process is aptly described as 'Listen before talking'. The rule specifies that a station can transmit only if there is no carrier present on the bus, i.e. the bus is idle or free.

There may be more than one station ready to transmit at any point of time. All such stations will start transmitting as soon as they find the bus idle. There is multiple access (MA) that results in collision. Let us what a collision is and how it is detected. If a computer is transmitting bit '0' and that gets changed to bit T on the bus or vice versa (bit T changes to bit '0'), then a collision is said to have occurred. The transmitting station is continuously listening to its own transmission on the bus and detects such a collision. When it finds a T changed to '0' or vice versa, it knows that a collision has occurred. Continuous comparison of what is transmitted and received on the bus is the collision detection (CD) mechanism. This process is often called as 'Listen while talking'.

Once a collision is detected, what do the stations do? They wait for random times and retransmit again. One station may wait for one millisecond, another for two and so on. Since the wait time is random for each station, it is likely that each station waits for different time and then attempts retransmission. The collision is resolved in this manner. Since the wait time is random for each station, it is possible that two or more stations wait for the same random time. In such a case, there will be a collision again during the retransmission attempt. If this happens, the same process of waiting for random times is repeated until the collision is finally resolved.

Now let us look at the details of token passing ring protocol. This protocol is relatively simple when compared to Ethernet. However, certain types of failures need to be taken care of in this protocol. As mentioned earlier, a token circulates on the ring whenever there is no data transmission on the ring. A token is a particular bit pattern and is recognised by this pattern. When a station has data to send it seizes the token and starts its own transmission. By seizing we mean that the station changes the token bit pattern such that it is no longer recognised as token. Instead, the pattern corresponds to one that indicates the beginning of transmission of data. Following this pattern, the destination and source addresses are sent. Whenever a station sees the beginning pattern, it examines the destination address to determine if the data is destined for itself. If so, it copies the data. When the data transmission is complete, the source station reintroduces a token on the ring. If any other station has data to send, it follows a similar procedure. To avoid a station holding the ring for a very long time, an upper limit is set for the size of data packet that can be transmitted at a time. If a station has large data to send, it needs to break down the same into a number of packets and transmit. After sending one packet, the station will have to wait for its turn to get the token. Only when it gets the token again, it can transmit another packet. This ensures that every station gets a fair chance to transmit.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

16) What is a carrier in LAN?

17) Find out and record the carrier frequency of a nearby AM radiobroadcast station.

18) What are the reasons due to which an Ethernet station experiences a collision during a retransmission attempt?

19) What is a token in token ring protocol?

20) In some token ring implementation, the destination station, instead of source station, reintroduces the token. What difference does it make?

....

....

....

....

....

....

8.12 ADDRESS RESOLUTION

In data networks, destination addresses have different formats at different levels. This is required for easy implementation of a complex system. For example, at user level, we need easy to remember addresses like names. Such user addresses are not transmittable as such. The network needs numerical addresses specified in bits. The name addresses provided by the user are decoded and the user data packet is encapsulated with decoded numerical addresses. The process of converting the addresses from one format to another is known as address resolution. Encapsulation takes place at several levels in a hierarchical structure for packet transmission. Three addresses and two levels of encapsulation are important to understand although there may be as many as six addresses encapsulated often. We study the important ones now.

Consider the case of a user sending e-mail. He/she uses a destination address something like *james_bond@mgm.co.uk.* The sending computer cannot use this address as it is because every computer on the Internet is addressed by a 32/128-bit number. The server mgm.co.uk is known on the network by a number assigned to it and not by its alphabet description. The 32/128-bit number is called Internet Protocol (IP) address. IP address has two versions: IPv4 and IPv6. IPv4 uses 32-bit address and IPv6 128-bit. IPv4 has been in use for a very long time, over 30 years, and most of the computers on the Internet have IPv4 addresses as of now. IPv6 has been introduced recently. Over the years, IPv6 is expected to replace IPv4 addresses.

The first step in packet transmission is to resolve the string address to numerical IP address. This is done with the help of Domain Name Servers (DNS) that are located in a hierarchical structure throughout the Internet. DNS have a table of string addresses with the corresponding numerical IP address. Usually, the table in one DNS is only partial and the entire set of addresses is covered by the complete hierarchical structure of the DNS. To start with, the sending computer accesses the nearest DNS by sending it the character string address provided by the user. If the DNS has the particular string stored in its table, it returns the numerical IP address. Otherwise it accesses another DNS that is in the hierarchy. This process is continued until a DNS is found that has the particular string address and its corresponding numerical address in its table. This process constitutes the first level of address resolution.

Once the sending computer receives the numerical IP address, it encapsulates the user message with numerical addresses. The numerical address received from DNS is used as the destination address and its own numerical address as the source address. This is the first level of encapsulation. The next level takes place in LANs.

You are aware of the use of NIC in bus LANs and RIU in ring LANs. These interface units have their own unique addresses assigned by the manufacturer. They are accessed by these addresses only. These addresses are 48-bit long. The destination and source stations are identified by 48-bit interface addresses on the LAN. The computers in which the interface units are housed are identified by their IP addresses. IP addresses are not recognised by the interface units. We now have two addresses: 32/128-bit IP address for the computer and the 48-bit interface address. We need to resolve the destination IP address to destination interface address before the data transmission can take place on the LAN. This is the second level of address resolution, in bus LAN, this address resolution is done by using a protocol called Address Resolution Protocol (ARP). Let us now see how ARP works. In all LANs, there is provision to broadcast information. This is usually done by reserving a special broadcast address. On bus LANs using ARP, the sending computer broadcasts the destination IP address received as part of the first-level encapsulated packet. All other stations (NIC) read this broadcast. Whichever NIC is attached to the computer that has this IP address responds in reply. The sending computer now knows the NIC address to which the user information should be forwarded. It now encapsulates the user packet with the received NIC address as destination and its own NIC address as source and transmits the packet. This is the second level of encapsulation.

Thus address resolution and encapsulation are two important functions carried out in data networks at different levels.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

21) Why do we need address resolution in data networks?

22) What is the function of DNS?

23) How does ARP resolve IP addresses?

....

....

....

....

8.13 ROUTERS

A router is a device that forms one of the basic building blocks of Internet. Internet cannot function without routers. You are already familiar with repeaters and bridges. Router is a higher-level device that performs the functions of a bridge and a repeater and more. The primary function of router is to direct the user packets encapsulated with IP addresses in the direction of the destination. In this sense, the routers are much like telephone exchanges for the data networks. Telephone exchanges route the phone calls to the appropriate destination. Similarly, routers forward the data packets towards the destination. The telephone exchanges examine the number dialled to determine the destination. Routers examine the destination IP address in the incoming packets to decide the destination route. For this purpose, routers maintain what are called routing tables that contain entries relating to a destination addresses and the corresponding output that should be used to route the incoming packet.

Consider a user connected to a bus LAN in an institution in Delhi wanting to send a file to a user connected to a bus LAN in an institution in New York. Both LANs are connected to Internet via routers. The routers are connected to the LAN on one side and to the internet on the other as shown in Fig. 8.10. They are recognised both as a local machine and an Internet machine. They have unique Internet IP addresses as wet as local LAW addresses.

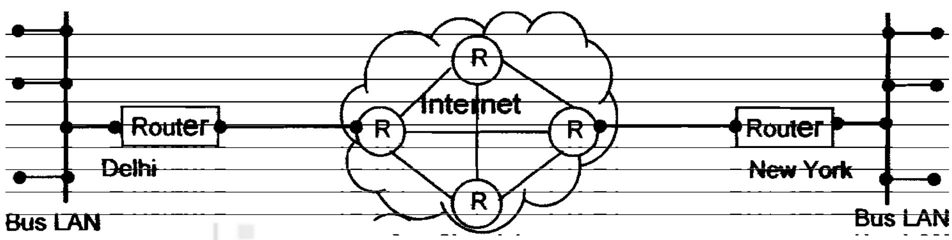


Fig. 8.10: Routers Use

There are also routers on the Internet that connect to other Internet routers in a mesh fashion. Whenever the destination IP address in a packet originating on the LAN does not point to a machine on one's own local network, as is the case in our example of Delhi sending to New York, the packet is forwarded to the router. This is done by a protocol called Transmission Control Protocol (TCP) mat runs on every LAN machine. We team about TCP HI Unit 9. The router forwards the same to another appropriate router on me Internet The packet travels via a number of routers on the Internet until it finally reaches the concerned router in New York that semis the packet to me appropriate machine on the LAN.

As you may observe, on me internet we have a number of possible routes that can be taken to reach the packet to New York. A typical route in our case is Delhi-Mumbai-Arnsterdam-London-New York. At every stage, the concerned router makes a routing decision as to which of the output ink must be chosen to forward the packet. The routing decision is taken according to the routing algorithm (software program) used by the router. Routing algorithm are discussed in the next section.

The routers shown at Delhi and New York institutions may also have multiple output links connecting to different routers on the Internet. It means that the concerned institutions have more than one Internet link. For example, if the packet were destined to Japan, the route chosen could be different if multiple paths were available.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

24) "Routers to data networks are Bee exchanges in telephone networks" Discuss.

....

....

....

....

8.14 ROUTING ALGORITHMS

As you may be aware, an algorithm is a step-by-step procedure to execute a task especially in a computer. Software programs implement algorithms to perform various tasks. Routing algorithms are procedures to make routing decisions. Routers execute routing algorithms to make routing decisions. Routing algorithms may be placed under two broad categories:

- Adaptive or dynamic algorithms
- Fixed or static algorithms

Dynamic algorithms adapt themselves to changing traffic conditions and network availability. For example, traffic may suddenly increase in a particular segment. As a result, tong queues of packets may build up stowing down the delivery to the destination. An adaptive algorithm may find an alternative route that may be longer but faster. Static algorithms use fixed routes for relatively long time durations. There could be more than one fixed route defined for a given destination in order of priority. Static algorithms do not monitor traffic conditions or the time to deliver. They are relatively easier to implement when compared to dynamic algorithms. They also need less processing power, i.e. CPU time.

Routing algorithms are designed to satisfy certain performance parameters. Some of the important parameters are:

1) Minimum delay for delivery
2) Minimum number of hops to reach the destination
3) Robustness
4) Stability
5) Fairness

Minimum delay may be local or global. By local we mean that the packet does not stay in the router for long. The output queue small and the packet leaves the router quickly. By global we mean tile delay in reaching tire destination. A packet may leave a router quickly but may get stuck later, tit such a case, it is better to route the packet by an alternative route even though the local delay may be longer.

A packet traversing a router is said to have done a hop. In other words, a packet hops from router to router on its way to the destination. As you are aware, every router examines the destination address in the header portion of every incoming packet. There is computational time overhead associate with this activity. Hence, it is desirable to have minimum number of routers or hops on the way to the destination.

Robustness means the ability to reach tile destination even when part of the network fails. A router should not forward a packet to a dead router on the way. In that case, the packet would never reach the destination. A robust routing algorithm would ensure that a packet is delivered to the destination at all costs. In other words there is guaranteed delivery.

Stability refers to the ability to deliver the packet as quickly as possible without the packet wandering here and mere. Sometimes loops may be formed in a network that packets may go round and round without moving forward towards the destination. A loop in the network is an unstable condition. Routing algorithms must ensure that no loops are formed. And if formed, they must be detected quickly and remedial action taken.

Fairness means delivery in a reasonable time for all types of packets. Sometimes, networks may receive high priority packets. In such case, other packets are delayed and the high priority ones are forwarded first. But such an action should not result in low priority or normal packets being delayed indefinitely. This criterion is called fairness.

There are many routing algorithms that are designed to implement one or more of the performance parameters discussed above. Some of the important algorithms are:

- Shortest path routing
- Flooding
- Hierarchical routing
- Broadcast routing

A shortest path may be determined based on one or more of the following factors:

- Link length
- Minimum local delay

- Minimum global delay
- Minimum cost
- Minimum number of hops.

Shortest path algorithm is one of the most popularly used ones. Flooding is a robust algorithm and is often used in military applications where delivery is critical. It is also very simple to implement. In flooding, an incoming packet is forwarded on all outgoing links except the one on which it arrived. The idea is that the packet will definitely be delivered to the destination via one of all the possible available routes. Hence, the algorithm is robust. Flooding generates a vast number of duplicate packets and can choke me network unless controlled. It can also cause loops easily. One of the reasons why a packet is not forwarded on the incoming link is to avoid looping. A variation of flooding is called controlled flooding. Here, a router remembers at the packets that it has forwarded. If the same packet returns to it, it is discarded straight away or after forwarding one or two more times. This is the control exercised.

Hierarchical routing maps the network in a hierarchy and forwards the packets via the appropriate hierarchical route. Broadcast routing is like flooding where the packet is sent even on the incoming route.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

25) Distinguish between dynamic and static routing.

26) What is robustness in routing? Which algorithm is designed to meet robustness requirement?

....

....

....

....

8.15 SUMMARY

This unit deals with four basic aspects of data networks, viz. topology, media access control (MAC) protocols, address resolution and routing. Data networks are designed in different geometrical shapes. The geometrical shape of the network is called the topology of the network. There are a variety of topologies in use. This unit discusses the main topologies. These include star, bus ring, mesh, tree and hybrid topologies. The way data travels in a network need not necessarily correspond to the physical topology of the network. Data travel often defines its own topology and this is called the logical topology of the network. Logical topology is generally defined by way of set of rules called protocols. Two important protocols, viz. Ethernet and token passing ring have been discussed in detail. The merit of implementing these protocols using star physical topology has been explained. Certain bask: devices that are required to build topologies and interconnecting them have been described. These include hubs, switches, repeaters, bridges, and routers.

All data packets that traverse a network carry the source and destination addresses.

These addresses use different formats at different levels. When moving from one level to another, addresses have to be translated from one format to another. This function called address resolution is required at different levels in networks. Three important address formats and the techniques for their resolution are discussed in detail. User level name addresses are resolved to IP addresses by domain name servers (DNS). IP addresses are resolved to hardware interface unit addresses by address resolution protocols (ARP). After addresses are resolved, encapsulation is required before data transmission.

Finally, the functioning of routers and the features of some of the important routing protocols have been discussed. A router is one of the basic building blocks of Internet. Internet cannot function without routers. Local routers connect to LANs on the one side and the Internet on the other. Internet routers connect to other routers in the Internet. The connectivity is such that there are at least two routes to reach a destination. Routers execute routing algorithms to make routing decisions.

Routing algorithms may be dynamic or static. Dynamic algorithms adapt themselves to changing traffic conditions and network availability. Static algorithms use fixed routes for relatively long time durations. Static algorithms are relatively easy to implement. Routing algorithms are designed to satisfy certain performance parameters. They include minimum delay, minimum number of hope, robustness, stability and fairness. Routing protocols that satisfy one or more of the performance have been discussed briefly. They include shortest path routing, flooding, controlled flooding and hierarchical routing algorithms. Shortest path algorithm is the most widely used one. It takes into factors like minimum local and global delay, minimum number of hops to the destination etc.

8.16 ANSWERS TO SELF-CHECK EXERCISES

1) Packet size is 2^{11} bytes. File size is 1 MB, i.e. 2^{20} bytes. Therefore the number of packets to be transmitted is $2^{20}/2^{11} = 2^{9}$, i.e. 512 packets.

2) The logical paths through which packets can travel from C1 to C7 are:

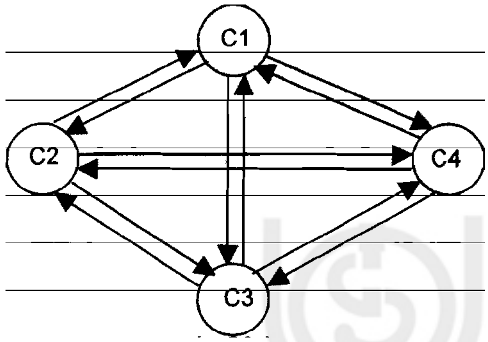
   - C1-C3-C5-C7
   - C1-C3-C5-C6-C7
   - C1-C2-C4-C5-C7
   - C1-C2-C4-C5-C6-C7

   There are a total of 4 logical paths constituting the logical topology of the network. The packet should not travel from C5 to C4, as it would result the packet going back to C1 and thus looping forever.

3) Yes. The packets of the same file may travel via different logical paths, as routing decision is taken for every packet independently. The problem that may arise is that the packets may arrive out of sequence at the destination. For example, consider a file having 10 packets. Packet 4 may be routed via longer route and Packet 5 via a shorter route. In such case, it is possible that Packet 5 arrives at the destination before Packet 4. The destination will have to take care of proper sequencing of the packets. For this purpose, whenever two or more packets of the same file are transmitted, the packets are tagged with a sequence number at the source so that the destination can sequence them properly. Packet numbering at the source and sequencing at the destination are taken care of high level protocols.

4) Half-duplex links carry information in only one direction. To support full-duplex communication, i.e. communication in both directions, we need two half-duplex links between every pair of computers. For a fully connected network with N computers, we need a total of N(N - 1)/2 full-duplex links or 2 x N(N -1)/2 half-duplex links. For 10 computers we need:

   2 x 10(10 -1)/2 = 2 x (10 x 9)/2 = 90 half-duplex links.



5) Fibre optic links are one-way (half-duplex) communication links. For two-way (full-duplex) communication, we need two optical fibre links between ever source and destination pair. The fully connected topology for four computers is shown in the Figure below. There are 12 links in the network. The formula

   N(N -1) links applies, i.e. 4 x 3 = 12.

6) The advantages of star topology are:

   - Ease of implementation
   - Ease of maintenance
   - Ease of administration: connection, disconnection etc.
   - It is a robust topology. If a link or a computer fails, the rest of the network is not affected
   - Flexibility for implementing different logical topologies, i.e. protocols including Ethernet protocol, token ring protocol and a switch.

7) The different logical topologies (or protocols) that can be implemented by a hub are Ethernet and token ring.

8) 20Base3 Ethernet system means an operating speed of 20 Mbps (mega bits per second) and the maximum distance covered is 300 meters.

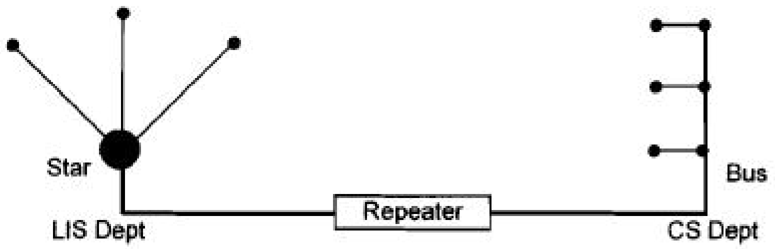
9) A station that seizes the token is expected to reintroduce the token on the ring after completing the data transmission. But it fails to so. In such a case, there will be no token on the ring. The token is said to have been lost. No other station can now transmit data and the ring is as good as 'dead'. Such a situation can be handled in the following way. When a station has data to send and finds no token or traffic on the ring for quite sometime, say 5 seconds, can introduce a new token on the ring. The new token can now be seized and new data transmissions can start.

10) Leaf nodes have no branches. Hence, the branching factor of leaf node is zero.

11) The number of nodes in a binary tree of depth 5 is (23 - 1) = 31. The number of point-to-point links is one less than the number of nodes, i.e. 30.

12) In tree topology, communication between adjacent nodes takes place by following a strict hierarchy. Hence, it is also called a hierarchical topology.

13) Both LIS and CS departments have Ethernet LAN, i.e. the same logical topology but different physical topologies. Since the departments are geographically far apart, signals will be attenuated. Hence, a repeater may be placed between the two buildings to boost the signal level as shown below.



14) If the CS department has a ring LAN, then the two protocols or the logical topologies are different. We need an intelligent device to transform packets from one format to another. Hence, we use a bridge instead of a repeater to interconnect the two departments. The bridge implements Ethernet protocol for the LIS department and the token ring protocol for the CS department. It performs necessary format conversion.

15) Comparison of repeaters and bridges:

Features	Repeater	Bridge
Signal amplification	Yes	Yes
Intelligent Device	No	Yes
Address Recognition	No	Yes
Packet Reformatting	No	Yes
Same Multiple Interfaces	Yes	Yes
Different Multiple Interfaces	No	Yes
Multiple Protocols	No	Yes

16) Carrier is a high frequency signal over which the data signals are superimposed.

17) AM broadcast frequencies lie in the range of 550 kHz to 1500 kHz. The AM Rajdhani channel in Delhi has a broadcast frequency of 666 kHz. Student is required to find out the broadcast frequency of an AM station nearby his/her city/town and record the answer.

18) There are two reasons as to why a collision may occur during retransmission attempt in Ethernet:

- The random wait time generated by two or more colliding stations might have been the same. For example, if two colliding stations generate random wait time as 0.1 sec by chance, then a collision will occur during retransmission attempt.

- A new station may join and start transmission at the same time when an old station is making a retransmission attempt.

19) A token is a specific bit pattern, say 7 ones and one zero (11111110). This pattern is unique such that it is not allowed to appear in the data.

20) Consider a token ring with 10 stations serially numbered. Let Station 2 seize the token and transmit data to Station 6. If Station 2 reintroduces the token, then the next station that would get the opportunity to transmit is Station 3. But if Station 6 reintroduces the token, then the next station that would get the opportunity to transmit is Station 7. Thus, the next station that gets opportunity to transmit changes in the two cases.

21) In data networks, destination addresses have different formats at different levels. This is required for easy implementation of a complex system. Address resolution is required to change addresses from one format to another.

22) Domain Name Server (DNS) resolves name string address provided by the user into a numerical IP address.

23) In all LANs, there is provision to broadcast information. On bus LANs using ARP the sending computer broadcasts the destination IP address of the packet. All other stations (NIC) on the LAN read this broadcast. Whichever NIC is attached to the computer that has mis IP address responds in reply. The sending computer now knows the NIC address to which the user information should be forwarded. Thus an IP address is resolved to NIC address.

24) The primary function of a router is to direct the user packets encapsulated with IP addresses in the direction of the destination. In this sense, the routers are much like telephone exchanges for the data networks. Telephone exchanges route the phone calls to the appropriate destination. Similarly, routers forward the data packets towards the destination. The telephone exchanges examine the number dialled to determine the destination. Routers examine the destination IP address in the incoming packets to decide the destination route.

25) Dynamic algorithms adapt themselves to changing traffic conditions and network availability. For example, traffic may suddenly increase in a particular segment. As a result, long queues of packets may build up slowing down the delivery to the destination. An adaptive algorithm may find an alternative route that may be longer but faster. Static algorithms use fixed routes for relatively long time durations. They do not monitor traffic conditions or the time to deliver. Static algorithms are relatively easier to implement when compared to dynamic algorithms. They also need less processing power, i.e. CPU time.

26) Robustness means the ability to reach the destination even when part of the network fails. A router should not forward a packet to a dead router on the way. In that case, the packet would never reach the destination. A robust routing algorithm would ensure that a packet is delivered to the destination at all costs. In other words there is guaranteed delivery. Flooding and controlled flooding are the routing algorithms designed to implement robustness.

8.17 KEYWORDS

Address Resolution : Given an address in one format, the process of obtaining the equivalent in another format

Algorithm : A step-by-step procedure for performing a task in a computer program

ARP : Address Resolution Protocol used to resolve IP addresses to NIC addresses in Ethernet LAN

Bridge	:	An intelligent device capable of interconnecting two dissimilar networks
Bus	:	A laid out open cable to which LAN computers are connected
Carrier Sensing	:	The process of checking whether a transmission is in progress on the bus of Ethernet LAN
Coaxial cable	:	A cable with an inner and an outer conductor placed coaxially and separated by insulating material. The overall structure is covered by a sheath
Collision Detection	:	The process of finding out if two or more stations are transmitting at the same time on one common medium
DNS	:	Domain Name Server used to resolve name string addresses into IP addresses.
Ethernet	:	A protocol used for exchange of information in bus or hub LAN
Flooding	:	A routing protocol that is robust
Hub	:	A device used in star configuration implementing one of the LAN protocols
Hybrid Topology	:	A network topology comprising two or more dissimilar network segments
LAN	:	Local Area Network
Media Access Control	:	A technique or protocol for controlling access to a common medium by multiple computers
Mesh	:	A network topology where computers are interconnected without any particular geometric shape in mind.
Multiple Access	:	Refers to where multiple computers access a common medium
NIC	:	Network Interface Card used in Ethernet LAN
Packet	:	A segment of information with specified length and structure
Protocol	:	A set of rules that govern the exchange of packets between computers
Repeater	:	A passive device that amplifies the signal level
Ring	:	LAN topology where computers are connected in the form of a ring
RIU	:	Ring Interface Unit used in ring LAN
Robustness	:	The ability of a routing algorithm to reach the destination even when part of the network fails

Router	:	A fundamental networking device used extensively in Internet to route packets towards their destination
Stability in Routing	:	The ability of routing algorithm to deliver packets as quickly as possible without the packets wandering here and there
Star	:	LAN topology configured with a central unit called hub
Switch	:	Networking device that switches incoming packets to output lines leading to their destinations
Token	:	A specific bit pattern use in token ring LAN
Token Ring	:	A LAN protocol for multiple access in the common ring
Tree	:	Network topology configured in the form of an inverted tree with the root at the top and the branches spanning downwards.

8.18 REFERENCES AND FURTHER READING

Mansfield, Kenneth C and Antonakos, James L. *An Introduction to Computer Networking*. New Delhi: Prentice Hall of India, 2002. Print

Tanenbaum. A. S. *Computer Networks*. 4^{th} Ed. New Delhi: Prentice Hall of India, 2002. Print

Viswanathan. Thiagarajan. *Telecommunications Switching Systems and Networks*. New Delhi: Prentice Hall of India, 2008. Print

www.en.wikipedia.org/wiki

UNIT 9 COMMUNICATION PROTOCOLS AND NETWORK ADDRESSING

Structure

9.0 OBJECTIVES

After going through this Unit, you will be able to understand and appreciate:

- What are protocols;
- Difference between computing and communication protocols;
- Need for communication protocols;

- Difference between connection-oriented and connectionless protocols;
- Basic packet transfer protocols like IP, TCP and UDP;
- Most widely used network computing architecture: Client-Server;
- File transfer and remote login application protocols like FTP and Telnet;
- Convergent cell switching in ATM networks;
- Fast routing technique using labels: MPLS;
- Numbering schemes used for landline and mobile phones;
- How network computers are addressed world over;
- Details of IPv4 addressing scheme;
- IPv6 features briefly; and
- Web access protocols for both wired and wireless networks.

9.1 INTRODUCTION

Quest for new knowledge is the central theme of human existence. All of us, whether we realise or not, are in the process of acquiring new knowledge all the time. When we ask a question, we are seeking knowledge. When we answer a query, we give information to the person posing the question. When a person assimilates the given information, we say that the person has acquired knowledge. Knowledge is spread via information that is communicated from one person to another in some form: oral, writing etc. Thus, knowledge, information and information communication are three entities that are closely inter-related.

It is often said that we are in the information age. In the last about six decades, information in the world has been growing at an exponential rate, i.e. doubling every 10 years. Information Communication Technology (ICT) has grown leaps and bound in the last 30 - 40 years. Instant transfer of information from one part of the world to any other part is a reality today. Underlying this development is the convergence of computer and communication technologies. This convergence process started in late 1960s and has led to the development of worldwide computer network that is now known popularly as **Internet.** A large number of home and office local area networks (LANs) and innumerable personal computers all over the world have been interconnected to form Internet. Hence, it aptly said that Internet is a **network of networks.** Information travels in the form of data packets on Internet and hence it is also called a **data network.** Data packets are of fixed length, say 2048 bytes, i.e. 2" bytes. Long messages are broken into as many packets as required before transmission. Because of packet-based transmission, the Internet also carries the nomenclature **Packet Data Network** (PDN). Since Internet is an open public network, another related nomenclature that is used sometimes is **Packet Switched Public Data network** (PSPDN). Internet is not limited to its presence only on the land but is also in ships at the seas and in planes in the air.

United Nations today has 192 countries of the world as its members. Almost all these countries have Internet connection in place. About 200,000 LANs are connected to the Internet. Over 1.5 billion people, i.e. a quarter of the world population has access to Internet. With the evolution of Internet, our life-style is changing. A number of our day-to-day activities are being carried out on the Internet. Clearly, the society is evolving towards a networked community with electronic information as the central commodity.

One might term the society of the 21st century as the **Networked Electronic Information Society** (NEIS). It is a society in which activities are centred on networks and the main commodity on the networks is electronic information in digital form.

It is important to realise that with alt its massive presence, Internet is still evolving. Today's Internet services are predominantly text and data oriented with only sprinkles of graphics, still pictures and slow motion video. Experience shows that Internet is slow for many network applications. Internet is basically designed for data transport. Real time services like voice and video transmissions experience serious quality problems. The key to the solution of current Internet problems lies in building **Global Information Infrastructure** (GII) that would have adequate capacity and efficiency to support full-scale services including high quality audio and motion video and high-resolution graphics envisioned for NEIS.

Information exchange between computers that are connected to a massive worldwide network cannot happen without standard procedures and sets of rules that govern such an exchange. A comprehensive collection of such standard sets of rules and procedures are called **communication protocols.** Furthermore, every entity on the Internet needs to be identified uniquely. This is done by assigning a **network address** to each entity. Communication protocols and network addressing are the subject matter of this unit.

9.2 WHAT ARE PROTOCOLS?

Let us start understanding protocols. The word protocol has different connotation under different circumstances. In governments, protocol means a strict official procedure in state affairs and diplomatic occasions. For example, in India the President is the Head of the State and there are protocols that govern his/her participation in state functions. These protocols specify how and where the President will be seated, who would accompany him/her, how would the dignitaries be introduced etc. In other words, they specify the accepted code of behaviour in particular situations. They may cover aspects about appearance (dress code), ways of greeting, conversation, and eating manners. All these rules help people successfully communicate and work together.

In inter-governmental dealings, the word protocol is used to denote the original draft of a diplomatic document containing especially terms of a treaty agreed to in a conference and signed by the parties concerned. You might have heard of Kyoto Protocol, a document that spells out the terms to be adhered to by the signatories for controlling and reducing carbon emissions in the world.

In science, a formal record of scientific experimental observations is often called a protocol. Procedures for carrying out scientific experiments or a record of the course of any medical treatment are also known as protocols.

The word protocol is used extensively in computers and communications as well. In computers, protocols deal with interaction between processes, exchange of messages etc. A process, as you may know, is a program in execution. In communication, protocols deal with signalling, switching, routing, forwarding, error control, monitoring, and recovery procedures in the exchange and transmission of information across entities in a network. A fundamental difference between the two is as follows. While computing protocols define rules for communication among processes within a computer, the communication protocols define rules for communication among computers. Both of them, however, deal with exchange of information. Protocols are generally software programs that implement the rules for communication. Some protocol functions are implemented in hardware, particularly those dealing with the movement of bits and bytes. In Section

9.3, we briefly discuss computing protocols and study communication protocols in greater detail in later sections.

9.3 COMPUTING PROTOCOLS

Computing and communication protocols together define sets of rules and procedures that govern all the information management functions. With electronic information being the central commodity in NEIS, information management functions become the core of technological capability in networks. There are seven functions of electronic information management that are important:

1) Generation
2) Acquisition
3) Storage
4) Retrieval
5) Processing
6) Transmission
7) Distribution

Computing and communication protocols govern all these functions. In general, information is generated by human thought process, human acts and by happenings in nature. Human intellectual activity is creative and intuitive and hence may not be amenable to protocols. Whether technology generates information is a debatable point. When data is processed in a computer, the output is considered as information. In that sense, it may be said that computers generate information. But the basic data comes from nature or human activity. However, machine generation of information can be governed by protocols.

Among the other functions, storage, retrieval and processing fall in the realm of computing protocols. The remaining functions, viz. acquisition, transmission and distribution fall in the class of communication protocols. Transmission and distribution functions may be collectively called as information *dissemination.* Transmission refers to bulk transfer between two main points. Distribution refers to transfer to end points like user computers or terminals.

Computing protocols are relatively a recent development. As you are aware, information processing, storage and retrieval are functions performed by application processes. You are familiar with applications like word processing, spread sheet, power point presentation and data base management. Computing protocols deal with information storage, retrieval and exchange among these applications. For example, how do we import information from word files into spreadsheets or vice versa? Or how do we import information from word files to power point presentation slides and vice versa? Computing protocols are being evolved to make such imports fairly easy. Some of the well-known computing protocol functions include message passing, process synchronisation and process switching, simple object access and object communication and data portability.

The idea of computing protocols is to encourage what are known as open systems design. Open systems follow industry standards and are capable of running on variety of platforms. For example, open office is an innovation in computing protocols. Many Java products use open computing protocols. Microsoft has recently announced a number of open computing protocols and has made them available in public domain. Open computing protocols offer greater opportunity and choice for freelance developers as they conform to industry standards. In contrast, closed protocols are proprietary in nature and are vendor specific.

Interoperability in computer systems is the main goal of computing protocols. By interoperability we mean the ability of different applications to interwork with each other using common data. User does not have to reformat and copy data from one application to another. Interoperability principles include:

- Ensuring open connections
- Promoting data portability
- Enhancing support for industry standards
- Driving open approach across competitors.

Although the open approach is currently limited to application packages from the same vendor, increasingly computing protocols are addressing issues for interoperability across different vendor products and platforms. Interoperability concept is also applicable to networked computers.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Differentiate between computing and communication protocols.

....

....

....

....

9.4 COMMUNICATION PROTOCOLS: GENERAL CONCEPTS

Communication protocols deal with all aspects of communication functions that are required for information exchange among computers in a network or across networks. They are designed especially in the context of Internet. We have already discussed in Unit 8 the protocols that are used for information exchange in LANs. On the Internet, communication related functions include:

- Breaking up messages into packets
- Packet sequencing and reassembly
- Synchronisation or handshaking for information exchange
- Signalling: start and end of messages
- Switching: routing or forwarding of messages towards their respective destinations
- Connectionless and connection oriented transfers

- Message encapsulation and de-capsulation
- Format conversion
- Error detection and correction
- Setting up and termination of sessions
- Recover from unexpected loss of connection.

We discuss each one of the above functions in order in the following paragraphs.

Packet formation: As mentioned earlier, in data networks information is transferred in the form of packets. Packets are of fixed length with an upper bound on the size. For example, the maximum size of a packet in Ethernet is 1500 bytes. User messages longer than the maximum permissible packet size need to be broken into multiple packets.

Packet sequencing: As explained in Unit 8, routers take forwarding decisions independently for each packet. Even static routing algorithms may have more than one route defined for the same destination. Depending upon the path taken, the packets may arrive out of sequence at the destination. If the packets belong to the same message, they cannot be delivered to the user program unless they are properly sequenced. The concerned communication protocol needs to perform this function of sequencing the packets of the same message in proper order. Breaking up a message into multiple packets at the source and reassembling them at the destination are complementary functions performed by communication protocols.

Synchronisation: A packet transmission cannot start unless the receiving station is ready. The sending and the receiving stations exchange handshake signals and synchronise their transfer process. Handshake signals are like sending a query 'Are you ready?' and receiving a response like 'OK, go ahead'. The synchronisation process includes agreeing on transfer speeds and the required buffer sizes. When the transfer is in progress, the receiving station may want a pause for reasons like buffer full. Handshake signals are exchanged to enforce 'pause' and 'resume' actions.

Signalling: Once synchronisation is achieved, the actual transmission starts. At this stage, the sending station must signal to the receiving station the start of the packet. This is usually done by sending 'start of text (STX)' bit pattern. Similarly, the end of the packet is indicated by 'end of text (ETX)' bit pattern.

Routing: We have discussed this function in detail in Unit 8.

Transfer modes: There are two fundamental ways in which information transfer takes place in our life: connectionless and connection oriented. These transfers are analogous to postal communication and telephone communication respectively. In postal system, we write a letter, post the same and expect it to reach the addressee. The postal system delivers on the best-of-efforts basis. While the letters are delivered most of the time, some get lost somewhere. In telephone communication a connection is first established between the parties concerned and then the communication takes place.

Encapsulation: Consider a packet traversing six routers before reaching the destination. Let the source station and the first four routers be on the same network. The last two routers and the destination station belong to another incompatible network. The fourth router will now have to encapsulate the user packet to make it compatible to the destination network. Encapsulation is like putting one envelope (user packet) into another and writing the addresses differently on the outer envelop. At the destination, the outer envelope (encapsulated packet) is discarded and the original information obtained. This is termed as de-capsulation.

Format conversion: Sometimes when moving packets between incompatible networks, pack formats may have to be changed. An example is moving packets between Ethernet and Token ring LANs, which calls for format conversion.

Error handling: Errors occur in data transmission. These have to be detected and corrective action taken. Error detecting codes are used to detect errors. There are two basic techniques available for error recovery. One is when an error is detected in a packet, it is discarded and the sending station is requested to retransmit the packet. This technique is called automatic repeat request (ARQ). Handshake mechanism is used to request retransmission of the packet. The other is to use forward error correction codes (FEC) that are capable of both detecting and correcting errors at the receiving end.

Sessions: A variety of tasks are performed on the networks by establishing sessions between a server and a client computer. Online search of databases, remote job entry, remote login to a time sharing system and fie transfer between two systems are examples of different types of sessions. Different sessions have different requirements. For example, a dialogue may be two-way simultaneous or one-way alternating. A large file transfer session may call for establishing roll back points to recover from connection failures. Session protocols perform functions required to establish, successfully execute and terminate properly different types of sessions.

Packet loss: It is not unusual to experience unexpected loss of connections in networks. You might have had this experience while accessing Internet. Some Internet browsers including Microsoft's Internet Explorer have provision to resume a session that was terminated unexpectedly, say due to a power failure. Many communication protocols have features to recover from unexpected connection failures. This is particularly so in sessions related protocols.

In this section, we have studied the general features that are required in communication protocols. In the next section, we look at the details of some of the commonly used communication protocols.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

2) We use signalling as a matter of fact in our daily life. Give any four examples of such signalling.

3) Is SMS a connectionless or connection oriented service?

4) When you are typing on a computer terminal, you make a mistake. Then you correct it. Which one of the techniques, ARQ or FEC you are using? Give reasons.

5) Many word processing packages have auto correct features. Which one of the techniques, ARQ or FEC is used there? Give reasons.

....

....

....

....

9.5 COMMON COMMUNICATION PROTOCOLS

The field of ICT is replete with protocols. Hundreds of protocols have been defined for various purposes. Many are very specialised, some are rarely used and some are defunct. You are already familiar with computing and communication protocols. There are other classes of protocols such as *data (bits & bytes) transmission protocols, routing protocols, access protocols, services protocols* and *applications protocols.* As a user of networks, you need to be concerned with only about a dozen protocols. This is much like a language dictionary having over 100,000 words and the average vocabulary of a person being about 4000 words.

Extensively used communication protocols include:

- Internet Protocol (IP)
- User Datagram Protocol (UDP)
- Transmission Control Protocol (TCP)
- File Transfer Protocol (FTP)
- Remote Login Protocol (Telnet)
- Internet Control Message Protocol (ICMP)
- Dynamic Host Configuration Protocol (DHCP)
- Post Office Protocol 3 (POP3)
- Simple Mail Transfer Protocol (SMTP)
- Internet Message Access Protocol (IMAP)
- Cell Switching Protocols (ATM)
- Muti Protocol Label Switching (MPLS)
- HyperText Transfer Protocol (HTTP)
- Wireless Application Protocol (WAP)
- Lightweight Transport Protocol (LTP)
- General Packet Radio Service (GPRS)
- Simple Network Management Protocol (SNMP)

Of the above, the first three protocols, viz. IP, UDP and TCP are basic protocols used by a variety of Internet services and applications. We discuss them in Section 9.6. FTP and Telnet are most extensively used service or application level Internet protocols. A large number of applications on the Internet use what is known as Client-Server architecture. FTP and Telnet and web browsers also use this architecture. We present this architecture in Section 9.7. We discuss FTP and Telnet in Section 9.8.

Routers use ICMP to report any abnormal event on the network. An example of an abnormal event that a router may discover is the outage of the network in some segment. Such an event may be reported to all other routers on the network as welt as the to the network management centre. ICMP is also used to monitor the functioning of Internet. ICMP, however, is not discussed in this course. DHCP is used for managing IP address allocation in local networks. This is an advanced protocol meant for network administrators and as such we do not discuss the same. POP3, SMTP and IMAP are all e-mail related protocols. They are not discussed in this unit.

At the data transmission level, viz. transfer of bits and bytes a new convergent switching technique has emerged in the 1990s. This technique is known as cell switching and the associated transfer mode is called Asynchronous Transfer Mode (ATM). There is a set of protocols associated with ATM. We present an introduction to ATM in Section 9.9.

Multi Protocol Label Switching (MPLS) is router-based technique for routing IP packets fast. An introduction is given to MPLS in Section 9.10.

HTTP is the widely used web access protocol designed to work with desktop and laptop computers. WAP and LTP are wireless access protocols designed to work with small portable devices like mobile phones. These protocols are discussed in Section 9.14.

A related protocol is GPRS that is used to send packets over slow-speed wireless links. GPRS is not discussed here. SNMP is discussed in Unit 11 that deals with network management.

9.6 BASIC COMMUNICATION PROTOCOLS: IP, UDP, TCP

9.6.1 Internet Protocol (IP)

Internet protocol (IP) is fundamental to the operation of Internet. All services on the Internet use IP for sending or receiving packets. No computer can be connected to the Internet without the IP running on it. Hence all computer operating systems like Windows provide IP bundled with them. IP software is usually memory resident. IP specifies exactly how a packet must be formed and how an Internet router should deal with the packet.

Packet and packet switching are generic terms used in a variety of contexts in ICT. For example, a network not conforming to Internet standards may use packet switching and define its own packet structure. In order to distinguish from other packets and to uniquely identify IP packets, the term *IP datagram* or simply *datagram* is used. We use the term datagram to mean an IP packet in this course module.

Although IP is a communication protocol defining datagram formats and transfer details, it serves an important purpose that goes almost unnoticed. Once on the Internet, a user can create and send datagrams to any computer on the Internet irrespective of where the destination computer is located. The user is unmindful of the presence of different component networks and a host of routers that interconnect them. Thus, IP makes a network of networks appear as one giant seamless data network.

An important aspect of IP is that it delivers datagrams on the best-of-efforts basis. The delivery is not guaranteed. The central idea in IP design has been internetworking and fast transfer of datagrams and not aspects like recorded delivery. Such aspects are taken care of higher-level protocol like TCP that we discuss later.

In general, a packet has the structure shown in Fig. 9.1. We use the term payload to denote the sum total of data handed over to IP for transmission. The payload may be pure user data or user data encapsulated with other information by any of the communication protocols.

The datagram header has a mandatory fixed length part and an optional variable length part as shown in Fig. 9.1(b). The fixed length is 20 bytes and the variable length can be up to 40 bytes making the maximum size of the header as 60 bytes. The different fields of the fixed part are illustrated in Fig. 9.1 (c) where each row is 32-bit or 4 bytes long. The source and destination addresses are 32-bit each corresponding to IPv4 address format.

IP addresses have two versions: Version 4 and Version 6 abbreviated as IPv4 and IPv6 respectively. IPv4 uses 32-bit address and IPv6 128-bit. IPv4 has been in use for a very long time, over 30 years, and most of the computers on the Internet have IPv4 addresses as of now. IPv6 has been introduced recently. Over the years, IPv6 is expected to replace IPv4 addresses. IP addresses are discussed in detail in Sections 9.13. The 'version' field in the header specifies the version to which the header belongs. Version information in each datagram permits the coexistence of different versions and smooth transition from one version to another.

Header	Payload

(a) A Generalised Packet format

Mandatory Fixed length	Optional Variable length

(b) Datagram header Darts

Version | Header	Service	Datagram length		
Datagram identifier				Fragment identifier
Time to	Upper layer	Header error control		
Source Address				
Destination Address				
Optional Fields up to 10 32-bit words				

(c) Mandatory Fields of IPv4 datagram header

Fig. 9.1: IPv4 datagram formats

The maximum size of IP datagrams can be up to 64 k bytes including the header and the text part. But rarely such a big size is used. Different networks are allowed to set their own limit for the maximum size of the datagram well below the theoretical limit of 64 k. This maximum size set by a network is called the *maximum transfer unit* (MTU) of that network. This provision further complicates processing of datagrams. If a datagram is delivered to a network with a size greater than the MTU of the network, then the datagram needs to be fragmented for transportation within that network and reassembled at the exit point of that network. In such a case, we need a provision to identify the datagram and its fragments. IP header fields, datagram identifier* and 'fragment identifier' in Fig. 9.1(c) are provided for this purpose. While reassembling the fragments, IP must know the original protocol from which the fragments came. This is specified in the field 'Upper layer protocol'. The one-bit 'M' field when set to 'V implies 'More fragments to come'. This bit is set to '1' in all but the last fragment. The last fragment will have this bit set to '0'. There may be certain applications where fragmentation may not be acceptable. The one-bit 'D' field, if set to '1' would mean Do not fragment'. In such cases, the route will be so chosen that no fragmentation occur.

You may recall that sometimes packets may wander indefinitely without getting delivered to the destination due to routing errors. The field Time to live' is used to exercise control over such malfunctioning. The field 'service type' addresses issues like priority etc. Other fields in the header are self-explanatory.

9.6.2 User Datagram Protocol (UDP)

UDP provides connectionless service at the user level. It uses IP for this purpose. In that sense, UDP is a higher-level protocol when compared to IP. Here, a user submits his/her entire message to UDP with a request for transfer to the specified destination. User message is a payload for UDP. In turn, UDP encapsulates this with its own header and passes the same to IP as payload. User datagrams are different from IP datagrams. User datagrams do not conform to IP standard. They are just chunks of information of any size. UDP encapsulates the user datagram with its own header to form UDP datagram. UDP may split a user datagram into multiple UDP datagrams conforming to IP standards.

UDP datagram is shown in Fig 9.2. Now let us see as to why UDP adds its own header to the user data. Many application processes or users on a computer may use UDP simultaneously. Hence, UDP needs to maintain an identity of individual process and its corresponding destination process. This information is kept in its header in the form of source and destination port numbers so that the datagram may be delivered to the correct destination process along with the source identification. In Fig. 9.2 each row is 4 bytes long. With two rows the UDP header is 8-bytes long. The port fields in the header identify the source and destination processes or applications.

Source Port	Destination Port
UDP Length	UDP Error
UDP Data or payload	

.

Fig. 9.2: UDP datagram structure

Each port field is 16-bit long. The destination port value is used to deliver the user datagram to the correct application. The destination application may use the source port value for sending a response to the source application. The port address feature is the one that distinguishes UDP from IP. Otherwise, the functional capability of UDP is the same as that of IP. As in the case of IP, UDP messages may be lost, duplicated and delivered out of sequence.

The value of the UDP length field specifies the total length of the datagram including data part and the header. The use of UDP error control field is optional. This field is used only for the header portion of the PDU, i.e. the error control is done only for the header. The UDP does not perform error control at the datagram level. This must be taken care of at the application level. The payload supplied by the user or an application program follows the header. The entire UDP datagram with its header and user data becomes the payload for IP.

UDP, being a connectionless service functions on the best-of-efforts basis. There is no delivery acknowledgement in UDP. There is no guarantee of delivery. But it is used extensively like the postal system. The protocol is simple, efficient and fast. There are a large number of applications where occasional non-delivery is acceptable. If the underlying network is reliable, UDP is very effective.

9.6.3 Transmission Control Protocol (TCP)

TCP is a connection-oriented service. It uses IP and is at a higher level. In fact, UDP and TCP are at the same level. TCP is guaranteed delivery service. TCP provides reliable and error free communication. It achieves this in four ways:

1) Detects errors in datagrams and uses ARQ technique for error recovery

2) Recognises duplicate datagrams and discards all but one

3) Detect lost datagrams and retransmit the same

4) Sequences datagrams received out of sequence

You are already familiar with error detection and the use of automatic repeat request. As you aware that some routing algorithms send out multiple copies of datagrams to achieve robustness. TCP checks for duplicate datagrams and accepts only the error free copy received first. Detection of lost datagrams is done using acknowledgement and timer mechanisms. Receipt of every datagram is acknowledged by the destination. At the time of sending a datagram the source initiates a timer with a value within which the acknowledgement must be received. If the timer expires and no acknowledgement has been received, the source concludes that the datagram is lost and despatches another copy. By adopting the above said four mechanisms TCP is able to provide reliable and error free transmission.

TCP being a connection-oriented service establishes a connection between two communicating programs before data transfer begins. The service progresses in four phases as in the following:

- Source requests TCP for a connection by giving the destination identity. TCP contacts the destination
- Destination responds with a positive acknowledgement
- Data transfer takes place
- Connection terminated

This is very much like what happens in a telephone conversation. Much as the way we use both telephone and postal systems extensively in our daily life, both TCP and UDP are used extensively on the Internet. Since TCP and IP are closely interlinked, vendors bundle both the software routines as part of the operating system. This is why you always hear of TCP/IP together.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

6) How many fields are there in the mandatory portion of IP header?

7) Why is fragmentation required while transferring IP datagrams?

8) What action is expected of a router if the field ‘D’ is set to ‘1’ in the IP header?

9) What purpose the port fields in UDP header serve?

10) How does TCP detect lost datagrams?

....

....

....

....

9.7 CLIENT-SERVER ARCHITECTURE

As mentioned earlier, FTP, Telnet and Web browsers are based on client-server architecture. Client-Server architecture is the most widely used form of computation on data networks. It has evolved from interactive computing model that was prevalent in the 1960s and 1970s. In interactive computing, a user interacts with a mainframe computer via a terminal that may be dumb or smart. The interaction model follows a master-slave approach. The mainframe computer acts as the master and the terminal as the slave. The slave terminal is under the complete control of the master computer.

With the advent of personal computers and data networks, the master-slave model of interaction has given way to **peer-to-peer interaction model.** Peer-to-peer interaction permits arbitrary communication among computers on the network. No distinction is made among the computers. A PC may contact another PC or a large mainframe as easily. Similarly, a mainframe computer can contact another mainframe or a PC. Distributed computing has become the norm. Distributed computing means any form of computation between two or more computers communicating over network.

Computers called **servers** that provide different types of services are on the networks. The services are accessible to other computers that are treated as **clients** of the service-providing computers. This model of interaction is known as the **client-server architecture.** A computer on the network may act both as a server and a client. When it provides service, it is a server and when it accesses the services of another computer, it is a client. We may thus say that the client-server architecture is a form of distributed computing with peer-to-peer interaction.

The client-server configuration is depicted in Fig 9.3. There are two machines and a network in the configuration: a server machine, a client machine and a data network.

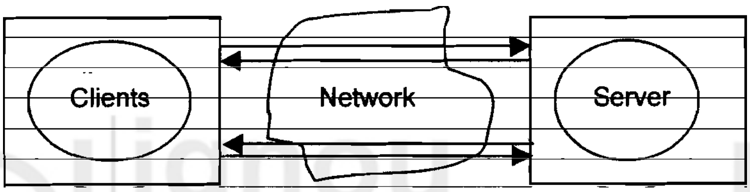


Fig. 9.3: Client -Server configuration

The server and the client interact via the data network. The server provides a set of information and computational services that are availed by remote clients. As shown in Fig. 9.3, usually many clients access one server simultaneously. It must be noted that the server and client machines do not actually interact. It is the server program and the client program that interact although we normally speak of server and client interaction. Support for multiple clients is possible only because of program-to-program interaction. A server creates as many processes of the same program as there are clients logged on to it. Use of multiprogramming and time-sharing features of the server operating system makes this possible. The server machine is one, but the instances of a server application program are many. This is how many users access one web site simultaneously. Client-server interaction may take place at one of the following three levels:

1) Human - Server Program

2) Human - Human

3) Client Program - Server Program

The first case is the most popular one with human client and a machine server. A typical example is that of user accessing information from a server, say searching a database. An example of the second case is student - teacher chat session. In online learning, student is tutored by a teacher. The student is the client and the teacher is the server. Timed periodic file transfer or email transfer between two or more machines are examples of the third case.

In all client-server interactions, it the client that always initiates a session. The server is ready and waiting without doing anything. When a client request comes, the server program responds. This is like a shopkeeper ready to sell with his shop open but the actual transaction takes place only when a customer arrives. The server service must be available on 24 x 7 (24 hours a day, 7 days a week).

Server systems are generally more powerful than client systems. They fall in one of the following categories:

- PC servers
- Workstation servers
- Mainframe servers

PC servers typically use standard 32-bit microprocessors. They have large RAM and hard disk capacity. They are ruggedised for continuous uninterrupted running with backup power systems and cooling systems where required. PC servers must have an operating system that can handle multiple users, as many client PCs may connect to the server at a time. Such operating systems are known as *network operating system (NOS).* Some of the popular NOS are MS Windows NT, Windows 2003, Novell Netware and Linux. All the servers are designed to support simultaneous access from many clients.

Workstation servers use high power or custom designed microprocessors. They are generally 64-bit or 128-bit microprocessor based systems. Workstation servers run under Unix like operating system that has a rich set of tools for supporting a wide variety of applications. Unix is a more reliable and secure operating system when compared to Windows. Linux is a recent addition to the world of operating systems and is considered a suitable substitute for both Unix and Windows. Linux is available in the open software domain. Some predict that in future, both PCs and workstations may run Linux instead of Windows or Unix. However, experience so far has not shown this to be true.

Mainframe computer based servers are even more reliable and powerful than Unix workstation servers. Mainframe based servers are often called enterprise servers to convey the fact that they are more powerful than PC servers or workstation servers.

Client systems are of two types:

- Desktop personal computers
- Mobile stations

The most popularly used desktop systems are Intel microprocessor based computers running Microsoft's Windows operating system. Such systems are sometimes called

Wintel systems signifying Windows operating systems and Intel microprocessor. The other class of desktop personal computer is Apple Macintosh. Mobile stations may be smart cellular phones (like Blackberry or iPOD), notebook computers and personal digital assistants or tablet PCs etc.

One of the powerful features of client-server architecture is its **scalability.** An application may start on a low-end PC and move in steps to a large PC, workstation and mainframe as the number of users rises. Interestingly, the upgradatton may happen without the user ever being aware of it.

In client-server architecture all applications have two program parts: a server program part and a client program part. The server program part is responsible for providing the specified services and the client part enables access to the services. Hence, anyone developing applications that would run on a network needed to develop both server and client parts of the software. The client software needs to be distributed to all client machines that may be spread all over the world. For example, you cannot access a server site that stores PDF (portable document format) files without an Acrobat Reader that is the client software for accessing PDF databases. In the early days, new server applications used custom-designed client software. Soon, it was obvious that roiling out client software to thousands of users all around the world is rather time consuming and expensive. With the arrival of World Wide Web (WWW), most of the network applications are designed to be **web-enabled** so that browsers now available with most of the PCs can access the applications without having to have special client software. In other words, the client software is embedded in the browsers. Two of the well-known browsers are Internet Explorer from Microsoft and Netscape Navigator from Netscape Communications.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

11) What are the differences between interactive computing and client-server computing models?

12) Can we say that Internet uses peer-to-peer communication? Why?

13) What mechanisms are used to support multiple clients on the same server?

....

....

....

....

9.8 APPLICATION LEVEL COMMUNICATION PROTOCOLS: FTP, TELNET

9.8.1 File Transfer Protocol (FTP)

FTP is used to transfer files from one computer to another on the Internet. FTP works in an interactive mode. A repertoire of FTP commands is available for interaction. A user invokes FTP client application on his/her computer as the first step. The user then enters the identity of the remote computer from which files are to be transferred using 'open' command of FTP. The FTP client then invokes TCP to establish connection with the remote computer. Once the connection is established, FTP server is activated at the remote computer.

At the next step, the user is authorised to access the remote computer by inputting a valid user name and password. A valid user account is required for this purpose. When the authorisation is successful, the user may examine and select a file on the remote computer by using the list command 'Is'. He/She then uses FTP 'gef command to transfer the file to his/her own computer. FTP client allows a user to transfer a file to the remote computer from the local computer as well. For this purpose, user invokes the 'send' command of FTP. The FTP client application is closed by 'bye' command.

FTP recognises only two types of files: text and binary. Any non-text file is treated as binary file. Examples include audio, computer programs, spread sheets and graphics data. Text files have to be strictly according to one of the standard character encoding schemes like ASCII or EBCDIC. If in doubt about the nature of the fife, it is best to specify binary format. Binary format will transfer a text file as well successfully. However, transfer of text files is more efficient and faster. Where known, it is a good idea to specify 'text' as the file type. However, If an incorrect type is specified, the resulting file may be malformed.

There are server systems on the Internet, which make available files to general public. Examples include servers providing government circulars or legal judgements. Such public files can be accessed without the user having an account on the server. FTP client makes this possible by providing an account called anonymous' with password as 'guest'.

Since FTP application runs on client-server model, the FTP server must run under multiprogramming and time-sharing operating system to enable multiple clients to access the server simultaneously.

9.8.2 Remote Login (Telnet)

Telnet allows an Internet user to log into a remote time-sharing computer and access and execute programs on the remote machine. For this purpose, the user invokes a Telnet client on his/her machine and specifies the identity of the remote machine. Telnet client makes a connection to the remote computer using TCP. Once the connection is established, the remote computer (Telnet server program) takes over the user's display and issues a login command. The user follows the regular login procedure by giving his/her account name and password. From then on the user computer behaves exactly like a terminal on the remote system. When the user logs out, the remote computer breaks the Internet connection and the Telnet client on the local machine exits automatically.

Remote login is a general access feature. The generality makes it a powerful tool on the Internet. It enables the programs on the remote computer accessible without having to make any changes to the programs themselves. The installation of the Telnet server on the time-sharing system is ail that is required. The telnet client and server together make the user computer appear as a standard terminal on the remote system. Hence, no changes are required as far as the application on the remote system is concerned. In view of this generality, different arbitrary brand of computers can be connected to the remote system. In effect, any computer on the Internet can become a Telnet client to any Telnet server on the Internet. Unlike FTP or e-mail, Telnet allows the user to interact dynamically with the remote system. Due to this, Telnet service is very popular.

Telnet sessions may run into occasional problems. The application program on the remote computer may malfunction or freeze. The local computer then hangs. We need a mechanism to come out of this situation. Remember that during a Telnet session, two programs are running: one the program on the remote computer and the other the Telnet client on the local machine. Telnet makes a provision to switch between these two programs. Once a Telnet session is established, every keystroke by the user is passed on to the remote computer. A special combination keystroke, like *Ctrl* +], is reserved to revert to the local program. The Telnet client examines every keystroke of the user before passing on the same to the remote machine. If the special combination key is pressed, it stops communication with the remote machine and allows communication with the local client program. The user can then terminate connection with the remote computer, close the Telnet client and resume local operations.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

14) Let the IGNOU LIS course modules be available on a fictitious FTP server called "cm.lis@Jgnou.ac.in". Write down the FTP commands and responses to download this unit to your computer.

....

....

....

....

9.9 SWITCHING LEVEL CONVERGENCE PROTOCOL: ATM

There are three major forms of switching techniques used in telecommunication and networks:

1) Circuit Switching

2) Packet Switching

3) Cell Switching

Circuit switching is the oldest technique used in telephone networks and has been in existence for over 120 years. Packet switching is about 50 years old used in data networks like the Internet. Cell switching is the most recent one evolved during mid 1990s used in new telecommunication infrastructure.

Before we proceed to discuss these techniques, definitions of two terms are in order: *channel* and *circuit.* A *channel* is defined as an information pipe with some specified characteristics like bandwidth, capacity, level of attenuation and noise immunity. A channel is a one-way link. A *circuit* is a two-way link and comprises two channels that enable two-way information flow between two entities. The two channels of a circuit need not have the same characteristics. If they do, then the circuit is said to be *symmetric.* Otherwise, the circuit is said to be *asymmetric.* Some authors tend to use the term *channel* to mean a physical medium. This is incorrect. A physical medium like optical fibre may carry several thousand information channels in a multiplexed mode.

Circuit switching is connection oriented. A circuit comprising two channels for two-way communication is established between the two communicating entities before the information transfer begins. The circuit is established using dedicated physical resources. The physical resource may be copper wires, optical fibres, radio or satellite links or a combination of these media. The circuit remains dedicated for the communicating pair until it is released, it is unavailable to any other communication need while dedicated to the communicating pair.

The main advantage of circuit switching is that once the circuit is established there is a direct connection between the communicating entities and the network is transparent to them. The information flows smoothly over the circuit from one end to another. The information is delivered in proper sequence and there is no possibility of out of sequence delivery. There is no delay caused by network elements like routers. The main disadvantage of circuit switching is that the scarce network resources remain dedicated for the entire duration of the information transfer phase and are heavily under utilised. Be it a telephonic conversation or computer interaction, mere are pauses during the session and the dedicated resources remain idle during the pauses.

In packet switching, messages are split as packets at the source and delivered to the network. *The* network transports the packets to the destination. Packet switched networks adopt two different approaches for transporting packets from the source to the destination:

- Datagram approach
- Virtual circuit approach

You are already familiar with datagram transport and the associated problems of out-of-sequence arrival at the destination and datagram losses. Virtual circuit approach was conceived to take advantage of in-sequence delivery of circuit switching while better utilizing the physical resources. The virtual circuit approach draws upon the idea of circuit estabtishment as in circuit switching. Instead of establishing a dedicated circuit, it establishes a fixed route from the source to destination. Since the packets follow the fixed route, they are delivered in order to the destination. Need for re-sequencing does not arise.

Virtual circuit makes routing more efficient and reduce the header overhead. As soon as a virtual circuit (fixed route) is established between a communicating pair, the same is given a unique identifier called *virtual circuit number* (VCN). The VCN defines the source and destination addresses, the message and the route. Hence, VCN together with the packet number uniquely identifies the packet. VCN plus packet number is much smaller in size when compared to the elaborate identification described earlier. Thus the header size and the transmission overhead are reduced significantly. Routing is also made simpler as the VCN is used to index a table to find out the outgoing link. There is no analysis of destination address and route determination.

Although virtual circuit concept is a major step forward, it still suffers from possible loss of packets and non-smooth flow of information. Since routers are involved, queues may build up and packets may be discarded. Dynamic variation in queue lengths may result in packets being delivered with different delay times thus interrupting smooth flow and affecting real time services.

Cell switching is the most recent switching technique evolved during 1990s. The main objective of cell switching has been to minimise the problems experienced in virtual circuit switching. This is done in two ways:

- To redefine the packet as a cell that is very small in size
- To leap forward in the speeds of virtual circuit switching.

Cell switching is designed to cause minimal network delay ensuring at the same time efficient utilisation of network resources. You are aware of MTU and the associated problem of possible segmentation and reassembly. This problem is completely avoided in cell switching. The entire infrastructure uses a standard cell size of 53 bytes. The cell has 48 bytes of payload and 5 bytes of header. Now let us understand the merits of cell switching. Ceil switching is built on a very reliable and ultra fast network infrastructure. The reliable technology almost rules out cell tosses. Even if a cell or two is lost very rarely, the effect is unnoticeable in real time services like voice and video. The small size of the cell makes the loss imperceptible to hearing or viewing. In data services of course, recovery is required.

Cell switching uses virtual circuit principle. The cells are guaranteed to be delivered in sequence. Virtual circuit reduces switching overheads significantly and makes switching extremely fast. For this reason, cell switching is sometimes called **fast packet switching.**

The networks that use cell switching are called **Asynchronous Transfer Mode (ATM)** networks. The reason for this is that the cells of a particular message are not switched in a fixed time frame say every millisecond. They are switched as they arrive. The arrival is a mixed bag of cells from different messages or services. They are switched in the order in which they arrive. Consecutive cells do not necessarily belong to the same message or service. In other words, the cells of a message or service are not continuous or synchronous in time. Hence, the term asynchronous transfer is used. Asynchronous transfer ensures effective utilisation of the network resources. The resources are not dedicated to one service.

In contrast, in the conventional circuit switched networks the information transfer is continuous and synchronous. In synchronous transfers, information pieces arrive in fixed time frame, say one byte every microsecond. In ATM, the cell arrival is not time synchronous. The time gap between the arrivals of two consecutive cells of the same message is not fixed but a variable one. However, the variability is very small because of the high-speed switching of ATM. For all practical purposes, the services perceive synchronous arrival. ATM is a technique that marks the convergence of both circuit and packet switching. Hence, ATM protocols are often referred to as convergence protocols.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

15) Why is cell switching superior to circuit and packet switching?

....

....

....

....

9.10 MULTI PROTOCOL LABEL SWITCHING: MPLS

As you know, virtual circuit makes routing more efficient and reduce the header overhead by using VCN. The VCN uniquely defines the source and the destination, the message and the route. Use of VCN reduces the header size and the transmission overhead.

Routing is made simpler as the VCN is used to index a table to find out the outgoing link. Multi Protocol Label Switching (MPLS) is an attempt to bring VCN concept to IP packets. Here, an IP packet is assigned a label that uniquely identifies the destination. In fact, the IP packet is encapsulated with the label header. The label is then used to index into a table to find out the outgoing link to be used for forwarding the packet. There is no examination of the destination address every time. This greatly simplifies routing overhead and makes the IP packets move faster through the network. This is particularly useful where large volume data transfers are involved as in the case FTP service.

MPLS is a router-based solution to improve the router efficiency. This is not a protocol that runs on any user machine. User machines run only the conventional communication protocols like TCP and FTP. We need MPLS-capable routers to implement MPLS. Only MPLS-capable routers can assign labels and handle MPLS packets. There are two ways in which the labels are assigned to IP packets: data-driven and control-driven assignments.

In data-driven assignment, when a packet enters a MPLS-capable router, it contacts the next MPLS-capable router and asks for a label for the destination address. The next MPLS-capable router in turn connects to the next one and the process is continued until the destination router is reached. Thus a fixed route is formed for all packets to the same destination. The first router now encapsulates the IP packet with the label supplied by the next router and forwards the packet. From then on, the label is used for routing. The name mufti protocol is used to signify the fact that the MPLS-capable routers can forward IP packets from a variety o f protocols like TCP and FTP.

In control-driven assignment, a destination router creates labels for all its host computers and passes them to its neighbours. The neighbours in turn create labels and contact other neighbours. The process is continued until all the routers acquire the path. Thereafter, the label is used for routing.

9.11 TELEPHONE AND MOBILE NUMBERING

Every entity in any network needs to be uniquely identified. Otherwise, the entity cannot be accessed. In telephone and mobile networks the entity is a phone instrument and it is number that uniquely identifies the entity. In Internet, the entity is a computer and it is IP address that uniquely identifies the entity. Although it is called an address, IP address is also a number. At the user level, the addresses are specified by string of characters on the Internet, (e.g. ignou.com). In a sense, similar character addressing is also available in telephone networks by way of directories where one looks up the number corresponding to a name. The addressing or numbering scheme follows a structure. We discuss the telephone and mobile addressing schemes in this section and the IP addresses in Section 9.13.

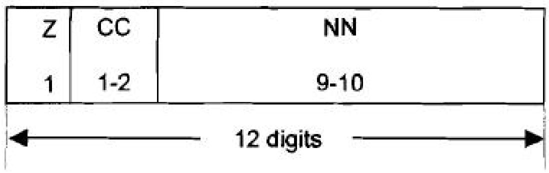
9.11.1 Landline Telephone Numbering

Telephone numbering worldwide follows an international standard set by International Telecommunications Union (ITU). The details are specified in the standards E.160 - E.163 of ITU. In ITU parlance, the numbering scheme is called **numbering plan.** As per the plan, the world is divide into 9 zones with each zone being identified by a zone code as indicated in Table 9.1. The zone names in Table 9.1 are representative. For exact delineation, one is advised to refer to the standards. Europe is given two codes, as there are many countries there.

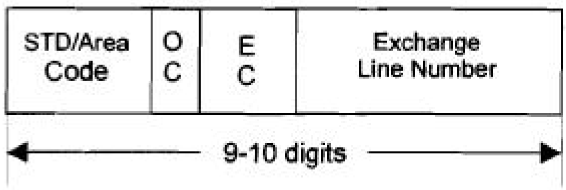
Table 9.1: World zones for telephone numbering

Zone	Code	Zone	Code
North America	1	Australia	6
Africa	2	Russia	7
Europe 1	3	Far East	8
Europe 2	4	South Asia	9
South America	5	-	-

The structure of the number is illustrated in Fig. 9.4. The maximum size of the number is 12 digits. The first digit is the zone number. The remaining 11 digits are divided between



(a) International telephone number



(b) National telephone number

EC = Exchange code   OC = Operator code

Fig. 9.4: Telephone Number Structure

country code (CC) and the national number (NN). The country code is one or two digits. With the zone code added, effectively the country code is 3-digit long. In common usage, zone code is not mentioned separately. It is included as part of country code. For example, the country code for India is mentioned as '91'. But to be precise, one should say that the zone code for India is '9' and the country code is '1'. Together they make '91'. The country code is kept to be of variable length. The general principle adopted is that the countries with large population are assigned short codes of two digits (1 zone + 1 country). The countries with smaller population are assigned longer codes of three digits (1 zone + 2 country). For example, in zone 9, Maldives has a 3-digit country code '960'. With just over 200 countries in the world this coding would work for millenniums to come, in fact, forever. Countries with large population like India get 10 digits for national number and smaller countries tike Maldives nine digits. Ten digits provide for a maximum of 10 billion connections. For India with a population of 1.2 billion, this is adequate for times to come. You may appreciate that the telephone numbering plan has been designed with farsightedness.

National telephone number has four parts in it as shown in Fig. 9.4(b). Subscriber Trunk dialling (STD) code may be further subdivided as one digit region code within the country and one or more digits of sub area codes within the region. In India, eight regions have been identified for telephone numbering. These are numbered 1 - 8 as shown in Table 9.2. The region descriptions are indicative and actual area covered is as per Department of Telecommunications guidelines.

Table 9.2: Indian regions for telephone numbering

Region	C	Region	C
Delhi, NCR, Haryana, Punjab	**1**	U.P. & Bihar	**5**
Mumbai, Maharashtra,	**2**	Orissa	**6**
Kolkata & North East	**3**	Central	**7**
Tamil Nadu & Kerala	**4**	Karnataka &	**8**

The sub area codes are kept to be of variable length. The general principle adopted is that the sub areas with large population are assigned short codes of two digits (1 region + 1 sub area). The sub areas with smaller population are assigned longer codes of three or more digits. For example, Delhi has a code '11' and Noida '120'. Similarly, Mumbai has a code '22' whereas Bopal, a town near Ahemadabad has the code '2707'.

Operator code (OC) is used when there is more than one service provider in an area. Until the early 90s, India had only the state operator, the Department of Telecommunications, providing telecom services in the country. But now telecom is opened up to private operators. We generally have more than one operator in major cities and towns. The operator code is used to identify the different service providers.

Every service provider has more than one telephone exchange in a city. Exchange code (EC) identifies the telephone exchange to which the subscriber is connected. Usually two or three digits are provided for EC. If the number of exchanges exceeds 99, we need three digits. This is the case in cities like Delhi and Mumbai. In smaller cities and towns, only two digits may be used.

The last part of the national telephone number is the line number assigned to the subscriber in the telephone exchange to which he/she is connected. Exchanges are usually designed to support 1000 or 10,000 subscribers. Accordingly, the line number may have 3 or 4 digits.

9.11.2 Mobile Phone Numbering

Technically speaking, there is no reason as to why mobile phone numbering could not follow the same numbering plan as the landline phone numbering. After all, the mobile is another telephone instrument except that it works on radio technology instead of landline (electrical or optical cable) technology. In fact, initially mobile phones in the United States used the same numbering scheme as the landlines. But, commercial considerations have led to a different scheme of numbering for mobile phones. In the beginning, the cost of mobile technology was relatively higher when compared to the landline technology. Mobile service providers needed to charge the customers higher. You may be aware that in the initial days of mobile communications, the incoming calls to mobile phones used to attract incoming call charges and the outgoing calls used to cost about six times the landline charges. The charge differential being so high, users needed to know whether they are calling a mobile phone or a landline phone. Further, roaming feature of mobile phones and the associated charging policies made the distinct identification of mobiles phones necessary. Hence, the need arose to distinguish a mobile phone from a landline phone. Thus was bom a different numbering scheme for mobile phones.

In general the series of numbers starting with '9' was reserved all over the world for future use while the landline numbering plan was evolved. When mobile technology came up and a need arose for distinguishing mobile phone numbering, it was decided to use the '9' series for mobile numbering. As you have learnt, India has 10-digit national number with its country code being '91'. The 10-digit national number starting with '9' was allotted to mobile phones. The '9' series provides for one billion numbers and it was considered adequate to meet the needs of mobile users in India, particularly because the cost being high not much penetration was expected. But, the history proved otherwise.

India has over 500 million mobile users in the country as of June 2010. Close to 750 million mobile numbers have already been allotted. If the rate of growth in mobile users continues at the present rate, we would run out of mobile number space soon. Hence, Telecommunications Regulatory Authority of India (TRAI) has opened up unused '8' series numbers for mobile users. An interesting fact is that India has only about 450 landline users and the growth rate here is not very significant. Considering this, it is likely over a period of time that number space for landline users may be reduced and the space thus freed may be allotted to mobile users. However, it is important to note that the potential for high-speed applications is much higher in the case of landline communications. This is because the radio bandwidth is limited whereas the landline bandwidth is unlimited.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

16) What is the maximum number of telephone numbers (both landline and mobile together) that can be assigned using the 10-digit national number?

17) Identify the different components of the international telephone number 911129534336.

....

....

....

....

9.12 NUMBER PORTABILITY

Number portability is a feature that allows a telephone user to retain his telephone number permanently. It is like the PAN (Permanent Account Number) allotted by the Income Tax Department or the Social Security Number assigned to individuals in the United States. Such numbers remain permanent for the lifetime of the individual irrespective of where the individual lives or works with. The number portability feature implements a similar concept. Imagine that you are given a telephone number once in your lifetime and that number remains valid for your complete lifespan. Would that not be very interesting? Number portability attempts to do just that. However, we have a long way to go in this regard.

Number portability needs to be considered in three situations from the users' point of view:

- Change of location
- Change of operator or service provider

- Change of service from landline to mobile or vice versa

Accordingly, three kinds of number portability are discussed from the telecom network point of view:

- Location portability
- Operator portability
- Service portability

Location portability implies that if a user moves his residence or place of work from one locality to another in the same city or moves from one city to another, his/her telephone number does not change. Both intra-city and inter-city movements have to be taken care of.

Operator portability implies that if a user moves from one operator to another, say from Airtel to Vodafone, his/her telephone number does not change. You may recall that the national number has a field (OC) that identifies the service provider. When number portability is introduced, OC may lose significance and may just be used by the old operator to redirect the call to the new operator.

Service portability implies that a user may move from one form of service to another and yet retain the original number. At present, three forms of service are available: landline telephone, mobile phone and voice-over-IP. Service portability must ensure mobility of the user among all these three services.

So far our discussions were limited to one portability requirement at a time. Portability requirements may occur in combinations as well. The following combinations may arise:

- Location + operator portability. A user may shift from one city to another and may want to change the operator also at the same time.
- Location + service portability. A user may shift from one city to another and may want to change from one form of service to another at the same time.
- Operator + service portability. A user may want to change the operator and the service as well.
- Location + Operator + Service portability. All aspects being changed at the same time.

Major changes may be required in the telecommunication equipments for implementing number portability. Hence, although number portability is being talked about for many years now, its implementation is not wide spread.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

18) A mobile user moves from the city of Hyderabad to Mangudi, a village in Tamil Nadu, which does not have mobile network coverage but has landline connectivity. What portability aspects would come into picture in this case?

....

....

....

....

9.13 IP ADDRESSING: IPV4, IPV6

As explained earlier, every entity in a network needs to be uniquely identified. Otherwise, the entity cannot be accessed. In telephone and mobile networks the entity is a phone instrument and it is number that uniquely identifies the entity. In Internet, the entity is a computer and it is IP address that uniquely identifies the entity. Although it is called an address, IP address is also a number. An IP datagram cannot be delivered to the destination unless the destination is uniquely and unambiguously identified and IP address does exactly that. Computers connected to the Internet are called **hosts** and hence the term host address is used extensively.

There are two versions of IP addresses under active use. The most widely used one is defined in Version 4 of Internet Protocol abbreviated as **IPv4.** IPv4 has been in successful use for about 30 years and it uses a 32-bit address. With 32 bits we can uniquely address 2^{32} hosts, i.e. approximately 4 billion (4×10^9) hosts. Remember that the world has a population of over 6 billion. If everyone were to have a computer in this world, we would not have enough IP addresses to assign to each one of these. Further, the IP address is a structured one having different address formats. Structuring has the effect of reducing the effective address space to a much smaller number than 4 billion. With the rapid growth of Internet over the last 30 years we are now on the brink of running out of addresses for new machines. It is in this context, a new version of IP, Version 6 abbreviated as **IPv6** has been recently standardised and is being introduced on internet. Addresses in IPv6 are 128 bits long. With 128 bits we can have approximately 256×10^{36} unique addresses. Such an address space is unlikely to run out in the foreseeable future and must serve the mankind at least for a few millenniums. The two addresses, IPv4 and IPv6 are interoperable and would coexist for many decades to come. IP addresses are assigned and managed by a non-profit corporation called *Internet Corporation for Assigned Names and Numbers* (ICANN) to ensure uniqueness in naming and numbering hosts. In this Unit, we discuss IPv4 in detail and IPv6 briefly. Readers interested in more details of IPv6 may refer to Further Reading material listed at the end of this unit.

IPv4 address is structured reflecting the objectives of Internet. You may recall that Internet is a network of networks. Hence, at the level of ICANN, the main interest concerns networks rather than hosts. ICANN, through its agents around the world assigns addresses to networks that contain many hosts. As you are aware, each network is connected to the Internet via a **router** or a **layer-3 switch.** The router has an Internet network address and is capable of forwarding datagrams towards destination networks. On its own network, it distributes the datagrams to the respective hosts. The host addresses are assigned by the respective network owners and maintained on the router.

For the purpose of assigning network addresses by ICANN, the networks are classified under three categories: large, medium and small signifying the number of hosts on the network. Corresponding to three network categories, there are three address classes: **Class A, Class B** and **Class C** respectively. The general structure of IPv4 address has three fields as shown in Fig. 9.5. Class A provides for large, Class B for medium and Class C for small networks. The 'Class' field is of variable length of 1 - 3 bits. A one-bit class field with a value '0' specifies Class A addresses. The 2-bft field with value '10' and the 3-bit field with value '110' specify Class B and C addresses respectively. Often, the class field and the network number field together are called **network address** and the host field as **host address.** We also use the same convention in this unit.

Class	Network No.	Host

Fig. 9.5: IPv4 Address structure

In Class A address, 7-bit pattern following the first bit specifies the network number. Seven bits provide for $2^7 = 128$ bit patterns. Two of the 7-bit patterns are reserved for special purposes. They are all zeros '0000000' and all ones '1111111' patterns. The all zeros pattern implies the local network in which the host itself is located. The all ones pattern is used for loop back testing of protocols and applications. That leaves us with 126 Class A networks world over. The remaining 24 bits are used for host addresses supporting up to 16 million hosts on each network. With 16 million hosts on a single network, Class A represents the largest possible network on the Internet using IPv4 addresses.

In Class B address, 14-bit pattern following the first two bits specifies the network number. Leaving out the special patterns mentioned above, this means that up to $(2^{14} - 2) = 16,382$ medium sized networks may exist on the Internet. Much as in the case of network addresses, special patterns are reserved for similar purposes in host addresses as well. Taking this into consideration, each medium sized network may have up to $(2^{16} - 2) = 64$ k hosts.

In Class C address, 21-bit pattern following the first three bits specifies the network number. Leaving out the special patterns mentioned above, this means that up to $(2^{21} - 2)$ H" 2 million small sized networks may exist on the Internet and each such network may have up to $(2^8 - 2) = 254$ hosts. Class C networks are ideally suited for small organisations. They are used extensively. Some small organisations may have more than 254 hosts, say ranging from 300 to 1000, but may not have as many as 16 k hosts to warrant a Class B address. This, in fact is the case with organisations like universities, research laboratories and large corporate houses. In such cases, the organisation is allotted as many Class C addresses as needed. For example, three Class C addresses can support up to 762 (3×254) hosts. This approach of using multiple Class C addresses helps in conserving Class B addresses. If Class B addresses were to be assigned to such organisations, the address space would remain heavily under utilised.

In addition to the above three primary classes of addresses, there are two special categories of addresses, Class D & E. Class D address is used for multicasting and Class E is reserved for future use. Multicasting is the distribution of datagrams to many hosts that have the same address group. Note that broadcasting is the distribution of datagrams to all the hosts. Hence, multicasting can be called as 'limited broadcasting'.

For the sake of convenience and clarity, the 32-bit IPv4 addresses are presented in a dotted decimal notation. The addresses are viewed as four bytes (32 bits) and the decimal value of each byte is written with periods separating them. As you know, an 8-bit pattern can have values ranging from 0 —255. An example address in decimal notation is 183.41.235.7. The equivalent binary address is 10110111 00101001 11101011 00000111. With experience, it is felt that a hexadecimal representation would have served the purpose of clarity much better. As you know, hexadecimal representation uses a base of 16 using symbols 0 through 9 and A, B, C, D, E and F. The hexadecimal representation of this address is B7.29.EB.07.

Consider the case of an organisation that has 9,000 hosts. We would need to allocate a Class B address for this organisation to avoid allocating too many Class C addresses. Recall that one Class B address can support up to 16,382 computers. In this case, 7,382 addresses are wasted because the same network address cannot be used for another organisation. This is yet another example of how address space remains unutilised. The net result is the loss of address space. It was realised, though rather late, that a large segment of address space remains unused in IP class based address structure. While address space remained unused with the existing users, addresses for new users were becoming unavailable. A class definition with incremental number of hosts would have been far better. By the time this realisation came, the damage had been done. In order to contain further damage, Internet management introduced a classless addressing mechanism known as **Classless Inter Domain Routing** (CIDR) in late 1990s. The basic idea behind CIDR is to allocate the remaining IP addresses in blocks of contiguous addresses without any class consideration. With CIDR, an organisation can seek provision for hosts in powers of 2 such as 256, 512,1024, 2048 and so on. While network address length is fixed in Class A, B, C networks as 8, 16, 24 bits respectively, in CIDR the network address length may lie in the range of 8 - 31 bits theoretically. In practice, however, the range is 12 - 24. Note that the network address length of 8, 16, and 24 in CIDR automatically correspond to Class A, B and C networks respectively.

While CIDR makes more efficient utilisation of IP address space, it needed major changes in the routers all over the Internet, as the algorithm for routing has to undergo a sea change to handle both classed and classless addresses. Routing was relatively simpler and the router configuration was small with classed addressing. With CIDR the situation has changed. Routers have to maintain a much larger database including information about the length of the network address field. This is kept in a 32-bit field called **subnet mask** by setting as many higher order bits of the subnet mask to T as the length of the network address. For example, if the network address length is 20 bits, the subnet mask in binary notation is 11111111 11111111 11110000 00000000 and in decimal notation is 255.255.240.0. In decimal notation 255 means all the 8 bits of the byte, are set to '1' and 240 implies that four higher order bits of the byte are set to '1'. Therefore, twenty higher order bits of the subnet mask are set to '1' in this case.

IPv4 addressing is a classical example of how ad hoc and non-visionary decisions at the global level can turn out to be messy. It is worth noting that such a mess has never occurred in other fields like telephone and ISDN numbering. One of the major drawbacks of Internet is that many such shortsighted decisions exist and more and more ad hoc solutions are being found. This is clearly due to the lack of rigorous standardisation process such as the ones followed in ISO and ITU.

The introduction of CIDR and the concept of subnet mask led to a slight modification to the decimal notation for the IP address. Since the network address is now of variable length, its length is indicated by a number at the end of decimal notation after placing a 7. Example of new notation is 183.241.060.000/23. Here the network address is of length 23 bits. The subnet mask is 255.255.254.0

Let us now briefly discuss IPv6 addresses. As mentioned earlier, IPv6 addresses are 128 bits long. This is a very large address space, i.e. $2^{128} = 10^{38}$, i.e. approximately 10^{28} times the world population. This number is large enough to probably address almost every little thing on the earth. Therefore, it is impossible that this number space will ever get exhausted, certainly not for many millenniums. It is expected that in future items like refrigerators, air conditioners, cars, buses, ships, airplanes, and even bicycles will all be assigned IPv6 addresses so that they can be controlled and guided from the Internet.

IPv6 addresses are structured along the IPv4 addresses. The address space being very large, over 20 classes of addresses have been proposed. Five significant changes have been introduced in IPv6:

1) Large address space provided by 128-bit addresses

2) Flexible multiple header format. There is one base header that is mandatory and is of fixed size of 40 bytes. More extension headers can be introduced optionally.

3) Improved control options

4) Permits pre-allocation of resources

5) Provision for protocol extension.

There are only five fields in the fixed header portion other than the source and the destination addresses. These are *version, flow label, payload length, next header* and *hop limit.* You may recall that Ipv4 header has 11 fields in its header excluding the source and destination address.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

19) What are the lengths of network addresses in Class A, B, and C IP addresses?

20) How many hosts can be supported in one Class C address? Explain.

21) Write the following binary IP address in decimal and hexadecimal notations: 00011100 10101000 1100110011 00001100. What class of IP address is this? What is the network number (decimal) in this address?

22) Write the subnet mask in binary for /22 Classless address suffix?

23) How many hosts can be supported in classless address with suffix /27?

....

....

....

....

9.14 WEB COMMUNICATION PROTOCOLS: HTTP, WAP, LTP

As you are aware, World Wide Web (WWW) or simply Web is very popular on the Internet. These days, a large number of business houses, government agencies and many individuals have their own web sites. The number of web sites is growing day by day. Many business houses are currently upgrading their web sites to facilitate electronic commerce. In this context, it is necessary for you to learn about the communication protocol used for web access. Before we discuss the protocol let us briefly review the basics of web.

The concept of Web started in 1989 in France. In 1994, a consortium called World Wide Web Consortium (W3C) was established. W3C now has many countries as its members and is responsible for development and standards for Web and its access. From the users' point of view, web is a collection of documents scattered all over the world that are accessible over the Internet. The documents are often referred to as web pages. These documents are **hypertexts** as they contain embedded links to other documents. Embedded links are in the form of Uniform Resource Locator (URL) that contains three parts: a resource name, the identification of the server in which the resource is located and the protocol that can be used to access the resource. In effect, URL is unique identifier for a specific resource on the Internet anywhere in the world. An embedded URL is called **hyperlink.**

Web pages are of two types: static and dynamic. Static web pages are designed using a language called *Hypertext Markup Language* (HTML) that allows a developer to place text, graphics, sound, video and hyperlinks in a web page Being a mark up language, HTML defines how documents are to be formatted. In the process, it mixes the contents and format information. This poses serious problems while editing the pages. To overcome this deficiency, two new languages called *extensible Mark up Language* (XML) and *extensible Style Language* (XSL) have been developed. XML sets standards for structuring the contents and XSL for formatting the pages. Thus, the content and formatting are separated. These languages are being used increasingly these days.

As you are aware, web access is based on the client-server architecture that was discussed in Section 9.7. The web browser that acts as the client sends a web page request using URL and the server responds by returning the requested document. It is often necessary for the server to keep track of the user preferences for presenting information. The server does this by storing what are called **cookies** on the clients system. Cookies are short strings of data that a server sends along with a web page and uses the same later for meeting user preferences. The user, however, has the option of blocking cookies being stored on his/her system.

The protocol used for communication between the web browser client and the server is called *Hypertext Transfer Protocol* (HTTP). For secure applications, *Secure Hypertext Transfer Protocol* (HTTPS) is used. HTTPS is discussed in Unit 12 on Network Security. We discuss HTTP below.

HTTP is used universally to access web services all over the Internet. It specifies how a client may send requests to servers and how the servers may respond. The requests are sent in the form of ASCII (American Standard Code for Information Interchange) strings and the responses are received in the form of Multipurpose Internet Mail Extension (MIME) format. HTTP establishes a TCP connection on Port 80 of the server and uses the same for sending requests and receiving responses. More than one request may be sent without waiting for the responses. This is called pipelining of requests.

The first word of the ASCII string is one of the reserved words that specify the operation requested. For example, the reserved word GET signifies a read request for a web page and PUT for storing a page. PUT operation calls for authentication of the user. The authentication information usually follows the PUT request. The information that follows the operational request is called **request header** and need to be specified in a particular format.

In HTTP parlance, these operational requests are called **methods.** Other methods include POST, DELETE, and TRACE etc. POST is used append information to an existing page. It is used in the case of notice and news boards. POST method requires authentication so that only authorised users can post notice or news. DELETE is a request for removing the web page and obviously requires authentication. TRACE is a request for echoing the message that is being sent. This is used for diagnostic purposes.

The response from the server begins with a 3-digit status word that informs the client whether the request was successfully processed or not. It also indicates the reason in the case of unsuccessful processing. Typical failure messages include 'no content found', 'page not found', 'page removed' and 'forbidden page' etc.

With the advent of digital wireless access to Internet, considerable interest was generated in making small portable devices like mobile phones access web using wireless finks. Two access protocols **Wireless Application Protocol (WAP)** and **Lightweight Transport Protocol (LTP)** were developed for this purpose. These protocols were optimally designed to work with low bandwidth wireless links and wireless devices with a slow CPU, small amount of memory and a small screen. Obviously, such restrictions do not apply to desktop or laptop PCs. The device and link capability dictated the design of wireless protocols. In designing HTTP, these restrictions were not there. Over the time, the handsets have been made very powerful and wireless communication links have also become faster. The latest example in this category is iPhone4 from Apple Corporation. Accordingly, the later versions of WAP and LTP are also more sophisticated.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

24) How do the design considerations differ for wireless web access protocols when compared with HTTP?

....

....

....

....

9.15 SUMMARY

This unit has dealt with two distinct but closely related aspects: Internet communication protocols and network addressing. Protocols are generally software programs that implement the rules and procedures for communication. Some protocol functions are implemented in hardware too. There are protocols for computing purposes as well. Computing protocols are relatively a recent development. They define rules for information exchange among processes within a computer. Communication protocols define rules for exchange of information among computers. They deal with all aspects of communication functions that are required for information exchange among computers in a network or across networks.

An overview discussion of the communication protocols in Section 9.4 brings out their general functionalities like breaking messages into packets, packet sequencing and reassembly, message encapsulation and de-capsulation, error detection and correction and loss recovery. A list of commonly used communication protocols is given in Section 9.5. Then, the basic or fundamental protocols without which Internet cannot function are discussed in detail. They include IP, TCP and UDP. Internet Protocol (IP) is responsible for transporting packets from source to destination. Transmission Control Protocol (TCP) provides assured quality services that ensure errorless and lossless data transmission. User Datagram Protocol (UDP) is a low overhead, fast and simple protocol that delivers user messages on best-of-efforts basis.

The unit then covers Client-Server architecture that is fundamental to running remote applications on the Internet. It is the most widely used form of computation model on data networks. It has evolved from interactive computing model of yesteryears. Thereafter, two application level protocols that use client-server model for communication are discussed. File Transfer Protocol (FTP) is used for transferring files from one computer to another on the Internet. FTP works in an interactive mode using a repertoire of commands. Remote login protocol (Telnet) allows an Internet user to log into a remote time-sharing computer and access and execute programs on the remote machine.

The unit then focuses on two switching level communication protocols ATM and MPLS. Asynchronous Transfer Mode (ATM) is the new communication protocol used in basic telecommunication infrastructure. It uses the principle of cell switching that combines the advantages of both circuit and packet switching techniques. ATM is extremely reliable and fast. Routers use Multi Protocol Label Switching (MPLS) to speed up the process of routing packets across networks.

The unit then turns its attention to addressing entities uniquely on the networks. First, the numbering plans for telephone and mobile networks are discussed. Both international numbering and national numbering in India are elaborated. Another important issue, viz. number portability is then discussed. Number portability needs to be considered at three levels: Location portability, Operator portability and Service portability.

Addressing in data networks is then discussed. Version 4 of IP address (IPv4) is discussed in detail bringing out its limitations and merits. Developments in IPv6 are then briefly presented.

Finally, the unit discusses the web communication protocols. The universally used HyperText Transfer Protocol (HTTP) is described in detail. Brief features of wireless web protocols Wireless Application Protocol (WAP) and Lightweight Transfer Protocol (LTP) are then presented.

9.16 ANSWERS TO SELF-CHECK EXERCISES

1) Computing protocols define rules for communication among processes within a computer. Communication protocols define rules for communication among computers connected to the same or different networks.

   Computing protocols are concerned with storage, retrieval and processing functions of information management. Communication protocols are concerned with acquisition, transmission and distribution functions of information management.

2) Examples of signalling from our daily life:

   - A bus conductor's whistle to stop and start the bus
   - Flagging of a sport event like running race
   - Indicator lights in cars
   - Caller tunes in mobile phones.

3) Small Messaging Service is a connectionless service. One prepares a message and sends it across expecting it to be delivered. The service is provided on the best-of-efforts basis.

4) ARQ is the technique used here. You observe (detect) an error, erase (discard) it and input the right character (retransmit).

5) FEC is the technique here. The package detects the error and corrects it automatically. There is no retransmission by the user. This is forward correction at the receiving end.

6) Thirteen including source and destination addresses.

7) Internet is a network of networks. Each component network has its own maximum transfer unit (MTU) defined. If a datagram is delivered to a network with a size greater than the MTU of the network, then the datagram needs to be fragmented for transportation within that network and reassembled at the exit point of that network.

8) There are certain applications that cannot run with fragmentation. For such applications, the 1-bit 'D' field is set to '1'. This would mean 'Do not fragment'. If a router finds this bit set to '1', it must route the datagram via such networks that have MTU equal to or greater than the datagram size. Then no fragmentation will occur.

9) The port fields in UDP header identify the source and destination processes or applications. Using the information in these fields, UDP is able to deliver the datagram to the correct destination application. The source port identifies the source application that is sending the datagram, tf the destination so desires, it can send a reply datagram to the source application by using the source port address.

10) In TCP, detection of lost datagrams is done using acknowledgement and timer mechanisms. Receipt of every datagram is acknowledged by the destination. At the time of sending a datagram the source initiates a timer with a value within which the acknowledgement must be received. If the timer expires and no acknowledgement has been received, the source concludes that the datagram is lost and dispatches another copy.

11) In interactive computing, a user interacts with a mainframe computer via a terminal that may be dumb or smart. The interaction model follows a master-slave approach. The mainframe computer acts as the master and the terminal as the slave. The slave terminal is under the complete control of the master computer.

12) In client-server architecture, a computer on the network may act both as a server and a client. When it provides service, it is a server and when it accesses the services of another computer, it is a client. During interaction, the client is not under the control of the server. Client can do its own computing. Both the client and the server act independently and hence share the status of being peers. We may thus say that the client-server architecture is a form of distributed computing with peer-to-peer interaction.

13) Yes. Internet uses peer-to-peer communication. Any computer can contact any other computer. No computer is under the control of another. All computers are considered autonomous and can function independently. Hence, we say Internet uses peer-to-peer communication.

14) Machines do not communicate in client-server interaction. The interaction is between the server program and client program running on the server machine and the client machine respectively. As many instances of server program are activated as there are clients accessing the service. One instance is dedicated to one client. This is made possible by using the multi-programming and time-sharing features of the server operating system.

ftp>open

(to) cm.lis@ignou.ac.in

Connected to cm.lis@ignou.ac.in

LIS Course Module Services at IGNOU

cm.lis@ignou.ac.in FTP server ready

Name: yourusername

Password:…………..

Login OK

ftp> Is

PORT command successful.

Opening ASCII mode data connection for file list

Block 1: Basics of ICT

Block 2: Middleware Technologies

Block 3: Network Fundamentals

Block 4: Internet Tools and Services

Transfer complete

Xxx bytes received in xx seconds

ftp> Is Block 3: Network Fundamentals

PORT command successful.

Opening ASCII mode data connection for file list

Unit 8: Topology

Unit 9: Communication Protocols and Network Addressing

Unit 10: Protocol Architecture

Unit 11: Network Applications and Management

Unit 12: Network Security

Transfer complete

Xxx bytes received in xx seconds

ftp> get

(remote file): Block 3: Network Fundamentals/ Unit 8: Topology

(local-file): mydocuments/topology

PORT command successful.

Opening ASCII mode data connection for file list

Transfer complete

Local: mydocuments/topology

Xxx bytes received in xx seconds

ftp>bye Goodbye.

15) In packet switching, there are problems like out-of-sequence arrival at the destination and datagram losses. There is also the problem of segmentation and reassembly due to maximum transfer unit (MTU) limitation. Routing overheads are high in packet switching. In circuit switching, physical resources remain dedicated leading to their inefficient use. Cell switching overcomes these problems.

    Cell switching is designed to cause minimal network delay ensuring at the same time efficient utilisation of network resources. The physical resources do not remain dedicated.

    Cell switching is built on a very reliable and ultra fast network infrastructure. The reliable technology almost rules out cell losses. Even if a cell or two is lost very rarely, the effect is unnoticeable in real time services like voice and video. The small size of the cell makes the loss imperceptible to hearing or viewing. In data services of course, recovery is required.

    Cell switching uses virtual circuit principle. The cells are guaranteed to be delivered in sequence. Virtual circuit reduces switching overheads significantly and makes switching extremely fast. It is for these reasons that cell switching is superior to both packet and circuit switching.

16) 10-digit national number can support 10^{10} i.e. 10 billion numbers.

17) The 12-digit international telephone number given is '911129534336'. Here '91' stands for country code for India. The country code may further be subdivided as zone code '9' and country code within the zone as '1'. Next '11' stands for area code for Delhi. The area code may be further subdivided as region code as '1' and sub area code within the region as '1'. The following digit '2' stands for operator code, which in this case is MTNL. Digits '953' is the exchange code and '4336' is the exchange line number for the subscriber. The combination '29534336' is called the subscriber number.

18) Since the user is moving location, location portability is required. Since the user is a mobile subscriber and there is no mobile coverage available in the new place, service portability is required. It is assumed the operator is the same in the new place. Otherwise, operator portability is also required.

19) The lengths of network addresses in Class A, B, and C IP addresses are 8, 16, 24 bits respectively.

20) Class C address has 8 bits for host address. With 8 bits $2^8 = 256$ hosts can be supported. But the host addresses of all zeros and all ones are reserved for special purposes. Hence, a maximum of 254 hosts can be supported in Class C address.

21) The given binary address is '00011100 10101000 11101001 00001101'. Its decimal equivalent is 28.168.233.13. Hexadecimal equivalent is 1C.A8.E9.0D. This is Class C address as the most significant digit is zero. Seven bits following the first digit gives the network number which in this case is '0011100' and is 28 in decimal.

22) The subnet mask in binary for /22 Classless address suffix '11111111 11111111 11111100 00000000'

23) Classless address suffix is /27. This means 5 bits are available for host address. Leaving out the special reserved patterns of all zeros and all ones, we can have $(2^5 - 2) = 30$ hosts.

24) HTTP is designed for desktop and laptop and other high-end computers like servers. Here, there are no limitations of memory, computing power and screen size. It also assumes the availability of high-speed data links of at least 64 kbps. The emphasis on HTTP design is flexibility and powerful features. On the other hand, wireless protocols have to work with small portable devices like mobile phones. Here, the screen size is small, the available memory is very low and the CPU is not powerful. Data link speeds may be as low as 1.2 kbps or even less. Hence, the emphasis on WAP and LTP is high efficiency with essential minimal features only.

9.17 KEYWORDS

Client-Server	:	A computing and communication model used extensively in Internet
Connectionless	:	A service or protocol that commences information transfer without establishing a connection with the destination
Connection-oriented	:	A service or protocol that establishes a connection between the source and destination before information transfer commences
DHCP	:	Dynamic Host Configuration Protocol
Encapsulation	:	The process of covering a packet with another layer of header with a different format
Error control	:	The process of detecting and correcting errors
FTP	:	File Transfer Protocol
HTTP	:	HyperText Transfer Protocol
ICT	:	Information Communication Technology
ICMP	:	Internet Control Message Protocol
IMAP	:	Internet Message Access Protocol
Interoperability	:	Ability of different applications to interwork with each other using common data
IP	:	Internet Protocol
IPv4	:	IP Version 4 using 32-bit addresses
IPv6	:	IP Version 6 using 128-bit addresses
LTP	:	Lightweight Transport Protocol
NEIS	:	Networked Electronic Information Society
Open Protocols	:	Protocols that follow industry standards and are capable of running on a variety of platforms
POP3	:	Post Office Protocol Version 3
Protocol	:	A set of rules and procedures for information exchange between computers and applications

SMTP	:	Simple Mail Transfer Protocol
SNMP	:	Simple Network Management Protocol
TCP	:	Transmission Control Protocol
Telnet	:	Remote Login Protocol
UDP	:	User Datagram Protocol
WAP	:	Wireless Application Protocol.

9.18 REFERENCES AND FURTHER READING

Black, U. *Computer Networks: Protocols, Standards and Interfaces*. 2nd Edition. New Delhi: Prentice Hall of India, 1999. Print

Homer, Douglas E. *The Internet*. New Delhi: Prentice Hall of India, 2000. Print

Homer, Douglas E. *Internetworking with TCP/IP*, Volume I. 3rd Edition. New Delhi: Prentice Hall of India, 2001. Print

Lin, Yi-Bing. *Wireless and Mobile Network Architectures*. Singapore: John Wiley & Sons (Asia), 2001. Print

Mansfield, Kenneth C and Antonakos, James L. *An Introduction to Computer Networking*. New Delhi: Prentice Hall of India, 2002. Print

Panko, R. R. *Business Data Networks and Telecommunications*. 4th Ed. New Delhi: Prentice Hall of India, 2002. Print

Stallings, W. *Data and Computer Communications*. 5th Ed. New Delhi: Prentice Hall of India, 2000. Print

Tanenbaum, A. S. *Computer Networks*. 4th Ed. New Delhi: Prentice Hall of India, 2002. Print

Viswanathan, Thiagarajan. *Telecommunications Switching Systems and Networks*,. New Delhi: Prentice Hall of India, 2010. Print

UNIT 10 PROTOCOL ARCHITECTURE

Structure

10.0 OBJECTIVES

After going through this Unit, you will be able to understand and appreciate:

- What are protocol architectures and protocol stacks;
- What layered communication is;
- What are open systems;
- Open System Interconnection (OSI) reference model;
- Components of Internet protocol architecture;
- How Internet protocols are layered;
- Bluetooth technology;
- Bluetooth protocol stack;
- Integrated Services Digital Network (ISDN);
- Broadband ISDN (BISDN);
- Protocol model of BISDN;
- ATM protocol stack;

- Synchronous Optical Network (SONET) bit rates hierarchy;
- How mobile networks function; and
- Mobile network protocol architecture.

10.1 INTRODUCTION

Computer networks are complex systems. You already may have gathered an idea about the complexities involved while reading Units 8 and 9. Interconnecting computers with wires or wireless links is only a small part of the task of establishing a computer network. Even here, one has a variety of options. Different types of wires and cables, wireless systems of different speeds and a host of topologies to choose from are some of the considerations at this stage. Then, there are innumerable protocols to work with. Routers, hubs, repeaters and switches and many other devices need to be appropriately chosen and used. There are security considerations and finally applications to run. No doubt, we are dealing with a very complex system while designing and establishing a computer network.

10.2 PROTOCOL ARCHITECTURE AND PROTOCOL STACK

The subject matter of this unit is protocol architecture. Much as the way general architecture helps us in planning, designing and building complex structures, protocol architecture helps us in planning and designing complex network solutions.

Design and construction of any complex system follow three fundamental principles that are pursued in three different stages:

- Top-Down design
- Stepwise refinement
- Bottom-Up build

Top-Down design is the most widely used approach in the design of any complex system. The first step in this approach is the conceptual design. At this stage, one needs to take a holistic view of the complete system including its scalability and future growth. One is not concerned with individual parts or components at this stage. It is like working out an architectural or artistic overview of a major office or housing complex. One makes an attempt to visualise the system in its entirety identifying the major subsystems. The interconnection and interaction of these major subsystems are broadly spelt out. Many alternative conceptual designs are usually worked out, merits and demerits of each design are discussed and debated and one design is chosen finally.

The next stage in planning, design and execution of a major project is stepwise refinement. The chosen conceptual design is now refined or exploded to review further details. The major subsystems are broken down into smaller subsystems. The smaller subsystems are further refined to identify the components that go to make the subsystems. At every step, details of interfaces and interconnection among the subsystems are meticulously planned and precisely spelt out. The stepwise refinement process is thus continued until full details of the design are worked out.

The third and the last stage is the execution of the project, i.e. building the complex system. This is done following the bottom-up principle. As you know, actual building of a structure follows a brick-by-brick approach. Bottom-Up build signifies this aspect.

The proper components are selected first. Sometimes, even the components need to be designed and tested separately. Using the components selected or designed, smallest subsystems are assembled and tested. Many small subsystems are then interfaced to form larger subsystems and these are tested. Finally, the larger subsystems are interconnected to make up the full system. The full system then goes through trial runs before it is declared operational.

The three stages in realising a complex system as listed above are not watertight compartments. There is always certain amount of back and forth movement between the stages. For example, during stepwise refinement some constraint may come to light that might call for minor changes in the conceptual design. Similarly, during the construction process some incompatibility may come to light that might call for some changes in the details worked out during stepwise refinement process. The objective in the overall exercise must be to minimise such back and forth movements.

As you may know, architecture is the art or science of planning and building complex structures. It deals with the manner in which the elements of the complex structures are arranged and organised. Similarly, protocol architecture is the way of planning and building complex network systems. It deals with the manner in which the different protocols are arranged, interlinked and made to interact to achieve complex network functions.

Protocol architecture is defined as a comprehensive organised collection of a set of protocols at different levels with well-defined input/output interfaces. The protocols are designed to function in a co-ordinated manner to perform a variety of network functions to support different network applications. Levels define a hierarchical structure. Protocols in adjacent levels interact. The architecture is a comprehensive view of the entire functionality available. The levels at which different protocols exist are also defined.

Protocol stack is a chosen subset of protocols from the architecture to perform specific functions to support a specific application. While protocol architecture is the general structure reflecting the capability for performing a variety of functions, protocol stack is a specific structure to perform certain specified functions. Many authors use the terms *protocol architecture* and *protocol stack* interchangeably. We, however, make a distinction between the two terms in this course material as defined above.

Designing and supporting a network application follows the three-step process outlined above. With the aid of the protocol architecture and other subsystems required, a conceptual design for an application is finalised. At this stage, various protocol options are evaluated and interfaces to application-specific subsystems are considered. The conceptual design is then refined to arrive at the protocol stack that is required to support the application. The application is then implemented using bottom-up build approach. First, some gross and simple functions are tested and then detailed functions are integrated into the application.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Differentiate between protocol architecture protocol stacks.

....

....

....

....

10.3 LAYERED ARCHITECTURE

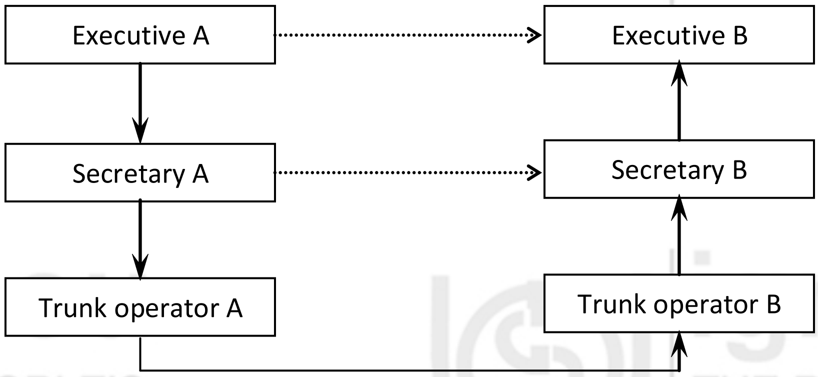


Fig. 10.1: Three-layer structure for trunk call connection

Layering is a natural choice for communication architectures. This is well illustrated by an example.

Consider the activities that are involved when executives A and B of two companies in different cities want to converse over a trunk telephone connection. For the sake of illustration, let us assume that there is no subscriber truck dialling (STD) facility between the two cities and trunk operators put through the trunk calls. Let executive A be the calling party. As a first step, he requests his secretary to connect him to executive B. His secretary in turn calls up the trunk operator A and communicates the calling number, called number, nature of the call, the name of the particular person called etc. Then, the trunk operator A calls up the trunk operator B in the other city and communicates the details. Remote trunk operator B now calls up the secretary of executive B, who in turn confirms with executive B that he would like to receive the call and requests the operator to put through the call. This process is depicted in Fig. 10.1.

A few interesting observations are in order:

1) A three-layer structure is used in this communication process.

2) The conversation between two adjacent layers is strictly business like.

3) There is generally a little private and informal conversation between the two trunk operators and between the two secretaries on account of their familiarity with each other. In other words, persons at the same level or layer exchange information in their own private way.

4) A layer obtains services from its immediate lower layer and provides services to its immediate upper layer. In this sense, a layer acts both as a user as well as a service provider.

5) A layer can communicate only with its adjacent layers.

6) There are fairly well defined functions to be performed by each layer.

7) It is immaterial as to how the functions of each layer are implemented. For example, the secretary may ask his assistant to book the call and as far as the executive is concerned, it is immaterial who books the call.

10.4 PRINCIPLES OF LAYERING

In fact, the above observations regarding a simple telephone conversation are stated as some of the important layering principles in arriving at a standard layered architecture known as *open system interconnection* (OSI) reference model or architecture evolved by International Standards Organisation (ISO). We state the ISO-OSI principles in the following section and discuss the ISO-OSI architecture in Section 10.4.

Principles of layering were first enunciated in the context of evolution of ISO-OSI reference model. Most important principles of layering are listed briefly in the following:

1) Layering is the most natural choice for data or voice communication network architecture.
2) Layers may be created to handle functions that are manifestly different in the process performed or technology involved.
3) Similar functions are to be placed in the same layer.
4) Boundaries are to be created at points such that the number of interactions across the boundaries are minimised.
5) Functions in a layer are to be localised and made autonomous to the extent possible so that the layer may be redesigned without affecting the interfaces with adjacent layers.
6) Redesign layers to improve performance by taking advantage of new advances in hardware and software technology.
7) A layer may be divided into sub-layers if the local function can be subdivided into independent modules
8) Layers are numbered bottom up. The bottom most layer is numbered 1.
9) A layer offers services to the upper layer immediately above.
10) A layer takes services from the lower layer immediately below.
11) Entities in the same layer but not in the same computer system are called **peer entities**.
12) Peer entities communicate using what are known as **peer protocols**.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

2) Particular protocol architecture has 10 layers. To which layer, Layer 5 provides services? From which layer, Layer 8 obtains services?
3) Does Layer 10 provide services? If so, to whom?
4) Does Layer 1 take services? If so, from whom?

....

....

....

....

10.5 ISO-OSI REFERENCE MODEL

As you are aware, data communication among computers involves a number of functions such as physical transmission of bits, error control, and routing and session establishment. In order to implement these functions efficiently, vendors of computer systems evolved their own architectures. Examples of vendor specific architectures are System Network Architecture (SNA) of IBM and Digital Network Architecture (DNA) of Digital Equipment Corporation (DEC). Such architectures permit interconnection of computers from the same vendor but not from different vendors. Systems or networks, which are not open to other vendor systems for networking, are known as 'closed' systems or networks. In order that heterogeneous computer systems from different vendors may be interconnected as a network, an architecture that is used as standard by all the vendors is required. The heterogeneity covers the following aspects:

- Systems of different vendors
- Systems under different management
- Systems of different complexities
- Systems of different technologies

ARPANET, the network project supported by Advanced Research Projects Agency of the Department of Defence, United States, is one of the pioneering efforts in connecting heterogeneous systems. The efforts put in and the experiences gained in the project have significantly contributed to the emergence of a set of world standards for computer communication. These standards, now well known as ISO-Open System Interconnection (ISO-OSI) standards, are widely accepted. The standards are based on a reference architecture which is described in the ISO standard IS 7498. International telecommunication Union (ITU) has adopted this standard under its own number X.200. The ISO-OSI architecture is considered 'open', as any vendor's system conforming to these standards is capable of organising information transfer with any other vendor's system that also conforms to the same standards.

OSI reference model proposes a general layered concept, with provision for adding or deleting layers as demanded by factors like service complexity, technology options, etc. ISO has recommended and standardised a 7-layer architecture shown in Fig. 10.2 taking into account various functions involved in data communication. On top of the seventh layer is the user who runs applications on the network. Hardware and software modules that implement the different functions of a layer are called **entities**. As mentioned earlier, the corresponding entities in the same layer but in different systems are called **peer entities**. The peer entities communicate using **peer protocols**.

Figure 10.2 shows two end systems that communicate with each other via two intermediate nodes. All the seven layers are active in the end systems. Only the first three layer functions come into action in the intermediate nodes. Entities in the first three layers always communicate with peer entities in the adjacent systems. The communication proceeds on a link-by-link basis from source to destination. Hence, layers 1-3 are called **link-to-link** layers. In contrast, the communication in layers 4-7 occurs between peer entities in the end systems. Hence, layers 4-7 are called **end-to-end** layers. The physical layer is concerned with transmission of bit streams either in synchronous or asynchronous mode. The data link layer handles errors and organises reliable transport of layer-2 data on a link-by-link basis. The network layer is concerned with the processing of destination addresses, routing of data units and internetworking. Routing algorithms are executed in this layer. Since the intermediate nodes perform the functions of the first three layers, they are also referred to as layer-3 switches. Layer-3 switches perform the routing function as well. The routing functions are implemented in hardware and hence layer-3 switches are must faster than conventional routers. They are also less expensive. As a result, layer-3 switches are replacing many routers particularly in campuses.

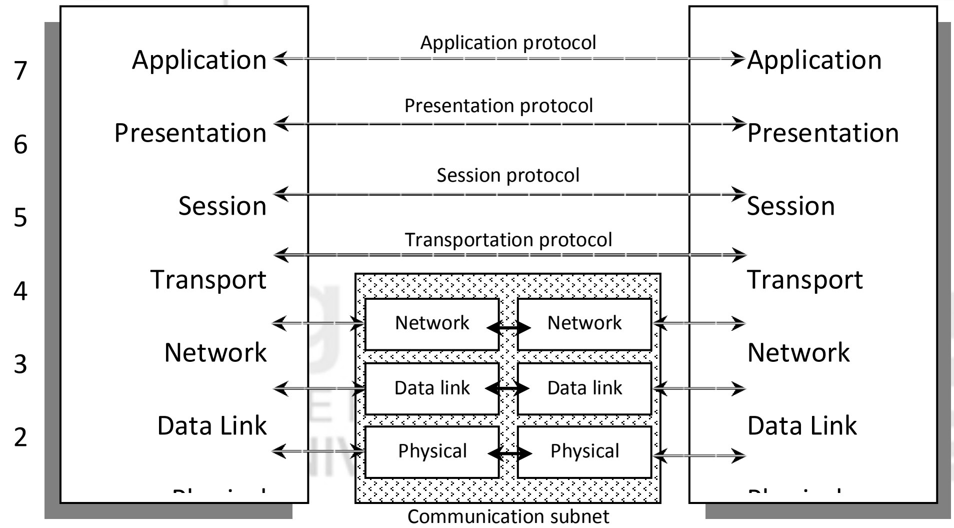


Fig. 10.2: ISO-OSI Reference Model

The transport layer is the first end-to-end layer and is concerned with reliable transport of full user messages between the two end systems. It is an interface layer between the user applications and the underlying network. This layer performs the function of splitting or segmenting the user messages into data units of appropriate size as required by the network layer. At the receiving end, this layer assembles the data units in proper order to reconstruct the full user messages before delivering the same to the upper layer. To ensure reliable transport of user messages, this layer retransmits data units that are lost in transmission or received with errors at the destination. The sessions layer permits two users or a user and an application to establish a session for purposes such as chat, interactive computing or information retrieval. The presentation layer deals with the syntax and semantics of information exchanged. It takes care of differences in data representation in the two end systems that may belong to two different vendors. The top layer is the application layer that enables a user to run different applications like electronic mail, accessing a web page etc. A number of different protocols have been developed that run at different layers of the OSI model.

The work of ISO on open systems followed ARPANET efforts and was based on ideas from ARPANET and the industry. ARPANET being a defence project, its results were considered confidential initially. Hence, the networking community looked up to ISO for open system standards. However, the elaborate procedure of ISO standardisation took time to finalise the architecture. In the meantime, the U.S Department of Defence decided to adopt and make public the ARPANET standards. Networking community started using ARPANET technology and soon it became the *de facto* standard in the public domain. ARPANET technology was adopted by Internet and is used widely today. We discuss Internet architecture in the next section. Although ISO-OSI architecture arrived late, there were many aspects in it that were comprehensive and structured. ARPANET protocols were refined later based on OSI recommendations. ISO-OSI is treated as a **reference model** even today and the new protocols and architectures are evaluated with respect to OSI recommendations. All new major developments in communications area come out with their reference model, protocol architecture or protocol stack. We discuss the important ones in Sections 10.7 - 10.11.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

5) How many layers are there in the end systems and in the intermediate nodes of OSI reference model?

6) Distinguish between link-to-link layers and end-to-end layers in OSI reference model.

....

....

....

....

10.6 INTERNET PROTOCOL ARCHITECTURE: TCP/IP ARCHITECTURE

Internet protocol architecture is popularly known as TCP/IP architecture. As you are aware, IP and TCP are two fundamental protocols used widely in the Internet. Since Internet grew in the initial stages based on these two protocols, its architecture came to be known as TCP/IP architecture. The architecture is presented in Fig. 10.3. The number layers in Internet protocol architecture is a debated topic. Some authors claim that there are only four layers and some talk about five layers. While there is complete agreement about the top three layers, many authors club layer 1 and 2 into one layer and call it by different names. The reason for this is that Internet standardisation process has mainly concentrated on the top three layers and allowed the bottom two layers to follow standards of different technologies. We present and consider 5-layer Internet protocol architecture in this course. In Fig. 10.3, we have shown the nomenclature for different layers along with some protocols in each layer. The protocol repertoire is rather large for each layer except for layer 4. As mentioned in Unit 8, the complete set of protocols run into several hundreds. We now briefly discuss each of the layers of the Internet protocol architecture in the following paragraphs.

Layer 5	Application	FTP, Telnet, SMTP, HTTP
Layer 4	Transmission Control	TCP, UDP
Layer 3	Internet	IP, ICMP, IGP, EGP, BGP
Layer 2	Network Access	Ethernet, ARP, 802.11, 802.16
Layer 1	Physical	X.21, 802.3, 802.5, SONET

Fig. 10.3: Internet Protocol Architecture

The physical layer deals with reliable movement of bits and bytes over point-to-point communication links. The technology and the techniques used depend on the network in use. X.21 is the protocol used in wide area telecommunication data networks. 802.3

and802.5 define the techniques for Ethernet and token ring respectively. SONET is discussed in Section 10.10. There are other protocols that are used for radio networks.

The network access layer deals with media access control protocols. You are familiar with Ethernet and Address Resolution Protocol. 802.11 and 802.16 deal with media access for wireless LAN and broadband wireless LAN respectively.

Internet layer runs IP and other routing protocols like Interior Gateway Protocol (IGP), Exterior Gateway Protocol (EGP) and Border Gateway Protocol (BGP). IGP deals with routing in local systems. EGP and BGP deal with routing between two autonomous computers connected to the Internet. These protocols are used by routers.

Transmission control layer has the most widely used protocols that you are already familiar with. You are also familiar with the protocols in the application layer.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

7) Name any two protocols that form part of Layers 2 – 5 of the Internet protocol architecture.

....

....

....

....

10.7 BLUETOOTH PROTOCOL STACK

As you are aware, Bluetooth is a short-range wireless convergent technology. It is a low power radio technology covering a small range of distances up to 10 metres. The purpose of the technology is to make bluetooth enabled devices that are in the vicinity of a master device to communicate in a wireless mode. The technology is very simple to use. All that one needs to do is to bring a bluetooth device close to bluetooth-enabled computer and a communication can start. There is no cable or modem or driver to be installed. The simplicity of its use is very attractive to users. Bluetooth functions in master-slave configuration. Usually a bluetooth-enabled computer acts as a master. Slave devices are dumb, basically doing whatever the master tells them to do. For example, a mobile phone may connect to a laptop computer and send and receive electronic mail if the two devices have bluetooth interface incorporated in them. Bluetooth communication calls for no additional infrastructure like mobile communication that calls for a radio network to be in place. Bluetooth devices can communicate directly without any network support.

Layer 4	Application
Layer 3	Link manager
Layer 2	Baseband
Layer 1	Physical radio

Fig. 10.4: Bluetooth Protocol Stack

Bluetooth protocol stack is shown in Fig. 10.4. Being a very different technology with very different objective, Bluetooth evolved its own protocol stack quite independent of OSI or Internet protocol architecture. Bluetooth stack has four layers. The physical layer deals with radio transmission and modulation. Its main objective has been to arrive at low-cost device so that the Bluetooth technology would have wide acceptance.

10.8 ISDN REFERENCE MODEL

The baseband layer deals with media access and is somewhat analogous to MAC protocols. It specifies how master controls slaves by defining time slots and allocating them to the slave dumb devices. It also decides the order in which the dumb devices communicate. The link manager establishes logical channels, authenticate devices, perform power management and ensure quality of service.

Integrated Services Digital Network (ISDN) has been perhaps the most important development to emerge in the field of telecommunications in the 1980s, and it will probably continue to dominate the developments in decades to come. Unlike many other developments, ISDN is a well-conceived and planned area of development in the field of telecommunications. Based on a study with the objective of exploring the use of digital technology, International Telecommunication Union (ITU) adopted and issued a definition of ISDN in 1972 as:

Integrated Services Digital Network: An integrated digital network in which the same digital switches and digital paths are used to establish different services, for example, telephony and data.

This definition was further refined and in 1984, a comprehensive and generic definition of ISDN was adopted as:

> An ISDN is a network, in general evolving from telephony IDN, which provides end-to-end digital connectivity to support a wide range of services, including voice and non-voice services, to which users have access by a limited set of standard multipurpose user-network interfaces.

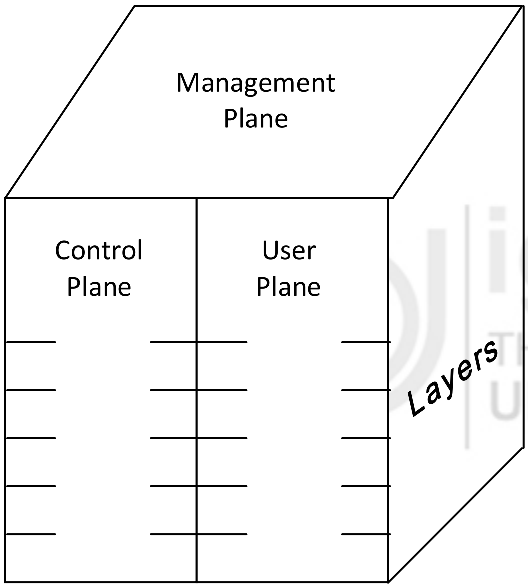


Fig. 10.5: BISDN Reference Model

In the above definition, IDN stands for Integrated Digital Network. Note that IDN did not envisage service integration. ISDN proposed a basic interface of what is known as 2B + D to users. 2B means two digital voice channels and D stands for data channel. It is used for signalling and data transfer. ISDN was designed to work up to speeds of 2 Mbps. Soon the need for higher speeds were felt. Broadband ISDN (BISDN) is defined as a network that caters to speeds higher than 2048 kbps, i.e. 2 Mbps.

Although started off with speeds just higher than 2 Mbps, after many years of deliberations, the basic speed of BISDN has been finalised as 155 Mbps, i.e. about 75 times the maximum speed of ISDN. This is because of the ambitious aim of BISDN to offer studio quality video and imaging services that demand very high bandwidth. The idea is to be able to distribute a wide variety of cultural, entertainment and educational materials to home and offices virtually on demand. The maxim is that "You ask for it and you get it". All these meant a quantum jump in technology, signalling and control, management and user services. This is reflected in the BISDN reference model depicted in Fig. 10.5. The model is designed along the lines of OSI reference model. Three distinct major components are visualised: User services, signalling and control, and network management. Each one of these components merits its own protocol architecture. Hence the reference model is a broad outline of these components having three different planes. Each plane has up to seven layers. Since a wide spectrum of user services were envisaged ranging from traditional voice service to video on demand service, a set of protocol layers is envisaged for the user plane. For example, conventional telephony is accessible in layer 3 and needs no protocol at higher layers. On the other hand, video services that call for maximum of bandwidth resources require protocols for presentation and session establishment. Protocols for these are at layers 6 and 5 respectively. Depending on the service accessed by user, different sets of protocols in different layers are invoked.

For signalling and control a protocol architecture called Signalling System 7 (SS7) is evolved. A set of telecommunications network management protocols are designed for management plane.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

8) Identify the important features of ISDN from 1984 ITU definition.

....

....

....

....

10.9 ATM PROTOCOL STACK

With the conception of BISDN at 155 Mbps came ATM. Cell switching and asynchronous transfer mode are chosen as the basic transport mechanism for BISDN. This development resulted in the definition of ATM protocol architecture. The ATM protocol architecture is depicted in Fig. 10.6. The architecture has three main layers. You may recall that OSI layering principles recommend formation of sub-layers if distinct functions are involved in a layer. This principle is used in forming ATM reference model and sub-layers are defined in Layer 1 and Layer 3. Both of them have two sub-layers each.

Layer	Sub-Layer	Functions
ATM Adaptation Layer	Convergence sub-layer	Convergence
	Segmentation and reassembly sub-layer	Segmentation and reassembly
ATM Layer	NIL	Flow control Cell header formation and analysis VC links management
Physical Layer	Transmission convergence sub-layer	Error control Cell management
	Physical medium	Bit timing

Fig. 10.6: ATM Protocol Architecture

10.10 SONET HIERARCHY

At the lowest level, bit transmission at rates of 155 Mbps and above take place. In addition, physical layer performs header error control functions and cell management that includes multiplexing and demultiplexing of cells. The second layer, ATM layer deals with flow control, cell header formation and management of virtual circuit links. ATM adaptation layer converges a variety of real time and non-real time services of different speeds to ATM standards. In addition it deals with segmentation and reassembly.

Synchronous Optical Network (SONET) is the optical transmission network for ATM. Coaxial cables are also used for transmission. SONET uses light as carrier often referred to as Optical Carrier (OC). Coaxial cables use electrical signals. SONET constitutes one of the underlying transport networks for ATM. SONET does not have a protocol hierarchy or architecture. It has what is called transmission hierarchy defining speeds at which segments of the network may carry information. The SONET/SDH hierarchy is shown in Table 10.1. Transmission in SONET uses synchronous technology. Hence, its transmission hierarchy is also called as Synchronous Digital Hierarchy (SDH). ITU uses the nomenclature SDH. The minimum speed of ATM is defined to be 155.52 Mbps corresponding to STM-1 of SDH. STM stands for Synchronous Transmission Multiplex.

Table 10.1: SONET-OC/ITU-SDH Operating Speeds

SONET-OC	ITU-SDH	Speeds (Mbps)
OC-1	-	51.84
OC-3	STM-1	155.52
OC-9	STM-3	466.56
OC-12	STM-4	622.08
OC-18	STM-6	933.12
OC-24	STM-8	1244.16
OC-36	STM-12	1866.24
OC-48	STM-16	2488.32

In transmission, there is a basic data rate that gets multiplexed for higher speeds. Thus, OC-3 means three OC-1 channels are multiplexed. But for ATM the basic speed is 155.52 Mbps defined as STM-1. ATM starts with 155.52 Mbps and its higher speeds are multiple of the basic rate. STM-4 implies that four 155.52 Mbps streams are multiplexed to obtain a rate of 622.08 Mbps. This corresponds to 12 SONET basic rate channels. ITU has also defined a similar hierarchy for coaxial cable based hierarchy knows Synchronous Transport Signal (STS). STS-1 starts at the same level as OC-1 at 51.84 Mbps.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

9) What is the basic bit rate in ATM? What is its nomenclature in SDH? Which level of SONET corresponds to this rate?

10) What are the bit-rates corresponding to STM-4 and OC-16?

....

....

....

....

10.11 MOBILE NETWORK PROTOCOL ARCHITECTURE

As you are aware, mobile communication is major development that has taken place in the last two decades. It is a complex communication system. It offers a wide variety of services: voice, text and multimedia messaging, and a host of value added services. It is important for you to understand the protocol architecture of this system.

First let us understand how the system functions. A simple network configuration diagram is shown in Fig. 10.7. Let us digress a little from the main technical discussion to focus on a societal issue. Do you know that your mobile phone is continuously transmitting electromagnetic signals? Have you heard of advice not to keep your mobile phone close to your heart? This is because it is suspected that continuous transmission of electromagnetic signals near the heart may be harmful to it. There is no health study to prove this conclusively. If you do follow the advice, it is only a precautionary measure. On the other hand, radiation from mobile towers seems to be reaching levels that are harmful to health. Now back to technical discussion.

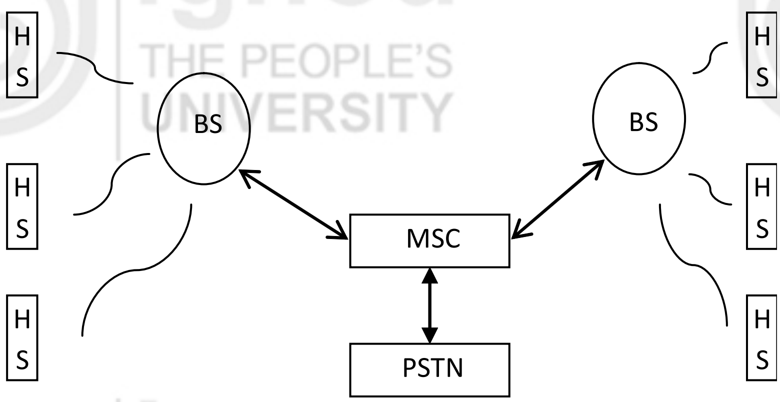


Fig. 10.7: Simple mobile network configuration

Continuous radiation from user handsets (HS) keeps them in touch with the nearest base station (BS). The base stations have two major subsystems. One, the tower and the antennas mounted on them. Collectively, they are called transceiver subsystem. The other subsystem is the control electronics housed at the basement of the tower, generally referred to as the base station control subsystem. The radio link between the base station and the handsets is modelled along the ISDN interface with two voice channels and one data channel. It is through the data channel that a handset is in continuous contact with the base station. It is also the data channel that is used for signalling and data services like SMS. You would have noticed that your mobile has facility to receive two phone calls simultaneously.

The radio coverage area of a base station is called a cell. The nomenclature 'cell phone' comes from this definition of cell. The cells are arranged in the form of a beehive with hexagonal area coverage. Two adjacent cells have different frequencies of operation. This is necessary to distinguish between the radio coverage of two adjacent base stations and to avoid interference between signals emanating from adjacent base stations. When a cell phone moves from one cell to another, a handoff procedure is followed to ensure that the transition is smooth without any break in the conversation. You may note that the frequency of operation of the handset changes when it moves from one cell to another. The handsets are designed to operate over the complete range of frequencies used by the base stations. This is what allows your handset function even when you are roaming.

The handsets are remotely monitored. Status information about handsets is stored in a database called mobility database (MDB). When you switch off a handset, it sends a signalling message to the BS, which in turn passes on the information to MDB via Mobile Switching Centre (MSC).

All base stations are connected via landline or dedicated microwave link to MSC. It is through MSC a connection between two mobile phones is established. MSC searches the MDB to ascertain the location and status of the called mobile and then establishes the connection via the concerned BS. MSC is also connected to the landline network, Public Switched Telephone Network (PSTN) for providing connectivity between mobiles and landline phones.

Having seen how a mobile network functions, we may now focus on the protocol architecture of mobile networks. The architecture is shown in Fig. 10.8. There are four layers in the architecture. The physical layer deals with four types of transmissions: radio transmission between BS and HS, microwave or landline transmission between BS and MSC and landline transmission between MSC and PSTN. Similarly, the link management layer also deals with four types of links. The third layer, network layer handles routing. Routing is done in four contexts in mobile networks:

- Mobile to Mobile routing with the same operator
- Mobile to Mobile routing with different operators
- Mobile to Landline routing with the same operator
- Mobile to Landline routing with different operators

Layer 4	Application	SMS, MMS, Music etc.
Layer 3	Network	Routing
Layer 2	Link manager	Radio, Microwave, Landline
Layer 1	Physical	Radio, Microwave, Landline

Fig. 10.8: Mobile Protocol Architecture

Application layer deals with applications like text and multi media messaging and value added services like music on demand, scores of cricket matches etc.

As you may be aware, mobile communications have undergone three generations of development: 1G, 2G and 3G. The fourth generation has been announced and is coming up soon. The above protocol architecture is applicable to 2G systems. 3G and 4G systems are broadband systems offering a wide variety of user services and may be expected to have architecture along the lines of BISDN.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

11) What are the different communication links that the link manager of mobile networks needs to deal with?

12) What are the different contexts of routing that the network layer needs to deal with?

....

....

....

....

10.12 SUMMARY

This unit is concerned with protocol architectures. Protocol architectures are layered structures. Layering is a natural choice for any communication process. International Standards Organisation (ISO) has evolved 7-layer reference architecture for data communication. ISO lays down well-defined principles for layering. The layers are numbered from the bottom starting with 1. A layer hides all the layers below it from the layer just above it. Layering principles promote autonomous functions in each layer and also the formation of sub-layers for independent local functions. Most protocol architectures are built around OSI model. In any architecture, each layer houses many protocols that are useful in a variety of contexts. An application may use none or only one or at best two protocols from each layer to run the application. The selection of protocols for a specific application and the associated layered structure is called a protocol stack.

Internet protocol architecture is popularly known as TCP/IP architecture. TCP/IP architecture has adopted 4/5-layer structure. In this unit, we have considered 5-layer architecture. Bluetooth, which is a specific application, has four layers in its protocol stack. Broadband integrated services digital network (BISDN) is seen as the future telecommunication infrastructure. Considering very high speeds and plethora of user services in future networks, BISDN has proposed a 3-dimensional reference model with protocol stacks for user, network control and network management. Asynchronous transfer mode (ATM) network is considered to be the basic infrastructure for BISDN. In view of this, ATM protocol stack is presented. ATM has a 3-layer protocol stack with sub-layers in two of the layers. The optical network SONET provides the underlying bit transmission infrastructure for ATM. The transmission hierarchy associated with SONET is presented. Finally, the unit concludes with a discussion on mobile network protocol architecture.

10.13 ANSWERS TO SELF-CHECK EXERCISES

1) Protocol architecture is defined as a comprehensive organised collection of a set of protocols at different levels with well-defined input/output interfaces. The protocols are designed to function in a co-ordinated manner to perform a variety of network functions to support different network applications. Levels define a hierarchical structure. Protocols in adjacent levels interact. The architecture is a comprehensive view of the entire functionality available. The levels at which different protocols exist are also defined.

   Protocol stack is a chosen subset of protocols from the architecture to perform specific functions to support a specific application. While protocol architecture is the general structure reflecting the capability for performing a variety of functions, protocol stack is a specific structure to perform certain specified functions.

2) Layer 5 provides services to Layer 6. Layer 8 obtains services from Layer 7.

3) Layer 10 being the top layer provides no service to any other layer. But it provides services to the user. User interfaces with the top layer of any protocol architecture.

4) Layer 1 being the bottom most layer it takes service from no one.

5) There are 7 layers in the end systems and 3 layers in the intermediate nodes.

6) The first three layers (Layers 1 – 3) of OSI reference model are called link-to-link layers. Lin-to-link layers exist in all the end systems as well as in the intermediate nodes. Entities in the end-to-end layers always communicate with peer entities in the adjacent systems. The communication proceeds on a link-by-link basis from source to destination.

   The top four layers (Layers 4 – 7) of OSI reference model are called end-to-end layers. End-to-end layers are present only in the end systems and not in the intermediate nodes. The communication in end-to-end layers occurs between peer entities in the end systems.

7) Protocols of the Internet protocol architecture:

   Layer 2: Token ring, Ethernet, 802.11, 802.16 (any two)

   Layer 3: IGP, EGP, BGP, IP (any two)

   Layer 4: TCP, UDP

   Layer 5: FTP, Telnet, SMTP, MIME, SNMP (any two)

8) The important features of ISDN from 1984 ITU definition are:

   - Provides end-to-end digital connectivity

- Supports a wide range of services, including voice and non-voice services
- Users have access to the services by using a limited set of standard multipurpose user-network interfaces.

9) The basic bit rate in ATM is 155.52 Mbps. Its nomenclature in SDH is STM-1. The level of SONET corresponding to this rate is OC-3.

10) The bit-rates corresponding to STM-4 and OC-16 are 622.08 Mbps and 829.44 Mbps respectively.

11) The different communication links that the link manager of mobile networks needs to deal with are:
   - Radio link between mobile handset and the base station
   - Microwave link between the base station and the MSC.
   - Landline link between the base station and MSC
   - Landline link between MSC and landline network PSTN.

12) The different contexts of routing that the network layer of the mobile protocol architecture needs to deal with are:
   - Mobile to Mobile routing with the same operator
   - Mobile to Mobile routing with different operators
   - Mobile to Landline routing with the same operator
   - Mobile to Landline routing with different operators

10.14 KEYWORDS

2B + D : User interface in ISDN with 2 voice channels and one data channel.

ATM : Asynchronous Transfer Mode. Cell transport network used in BISDN.

BGP : Border Gateway Protocol. Used for routing information between autonomous systems connected to the Internet.

BISDN : Broadband ISDN. Conceived to be the future telecommunication infrastructure.

Bit rates hierarchy : Defines a basic rate and a set of multiplexed rates for data transmission.

Bluetooth : Short-range wireless technology.

Cell (mobile network) : Radio coverage area by a base station.

EGP : Exterior Gateway Protocol. Used for routing information between autonomous systems connected to the Internet.

IDN : Integrated Digital Network without service integration.

IGP	:	Interior Gateway Protocol. Used for routing information between systems connected to a local network.
ISDN	:	Integrated Services Digital Network that supports service integration.
ISO	:	International Standards Organisation.
Layered Architecture	:	A comprehensive set of protocols arranged in layers.
MDB	:	Mobility database. Used for holding status and location information about mobile handsets.
MSC	:	Mobile Switching Centre that establishes connection between calling and called numbers in mobile networks.
OC	:	Optical Carrier. Unit of data rate in SONET.
Open systems	:	Systems that use international standards so that the can interconnect with other heterogeneous systems
OSI	:	Open System Interconnection. A reference protocol architecture model proposed by ISO.
Protocol Architecture	:	A comprehensive organised collection of a set of protocols at different levels with well-defined input/output interfaces.
Protocol stack	:	A chosen subset of protocols to perform specific functions to support a specific application.
PSTN	:	Public Switched Telephone Network. Commonly used landline telephone network.
Reference Model	:	Protocol architecture model that is used as a reference to compare other architectures.
SDH	:	Synchronous Digital Hierarchy. Defines digital data rate hierarchy in optical transmission.
SONET	:	Synchronous Optical Network. One of the bit transport mechanism used in ATM.
SS7	:	Signalling System 7. Signalling protocol stack used in telecommunication networks
STS	:	Synchronous Transport Signal. A hierarchy of data rates supported using cables.

10.15 REFERENCES AND FURTHER READING

Bryce, James Y. *Using ISDN*. 2nd Edition. New Delhi: Prentice Hall of India, 1998. Print

Homer, Douglas E. *Internetworking with TCP/IP*, Volume I. 3rd Edition. New Delhi: Prentice Hall of India, 2001. Print

Lin, Yi-Bing. *Wireless and Mobile Network Architectures*. Singapore: John Wiley & Sons (Asia), 2001. Print

Mansfield, Kenneth C and Antonakos, James L. *An Introduction to Computer Networking*. New Delhi: Prentice Hall of India, 2002. Print

Stallings, William. *ISDN: An Introduction*. New York: Macmillan Publishing Company, 1989. Print

Stallings, William. *ISDN and Broadband ISDN with Frame Relay and ATM*. 4th Ed. Singapore: Pearson Education Asia, 2001. Print

Tanenbaum, A. S. *Computer Networks*. 4th Edition. New Delhi: Prentice Hall of India, 2002. Print

Verma, Pramod K. *ISDN Systems: Architecture, Technology, and Applications*. New Jersey: Prentice-Hall Inc, 1990. Print

Viswanathan, Thiagarajan. *Telecommunications Switching Systems and Networks*. New Delhi: Prentice Hall of India, 2010. Print

UNIT 11 NETWORK APPLICATIONS AND MANAGEMENT

Structure

11.0 OBJECTIVES

After going through this Unit, you will be able to understand and appreciate:

- The need for global and national information infrastructures: GII, NII;
- The difference between network services and user applications;
- Fundamental forms of digital information;
- Interactive and distributive applications;
- Differences between broadcast, multicast and unicast transmissions;
- How text messaging (SMS) is implemented in mobile networks;
- The evolution of SMS language;
- How multimedia messages are sent in mobile networks;
- E-mail system and its features;
- What interactive television is;
- Interactive Music (IM) or Music-on-demand application;

- How real time service delivery is achieved;
- Performance issues in real time service delivery; and
- Network management and SNMP.

11.1 INTRODUCTION

As you know, the world is evolving towards a *Networked Electronic Information Society* (NEIS). Networked society means one in which a large proportion of the world population is interconnected or networked by some form of telecommunication system and the people carry out their day-to-day activities using the network predominantly. Day-to-day activities may involve tasks such as banking, ticket booking for travel or entertainment programmes, product ordering, financial transactions, exchange of mails, retrieving of information from a database, downloading of music files, simple telephone conversation etc. Electronic information is central to all these tasks.

Today's network infrastructure of NEIS is Internet. Internet is only a minuscule of a network that is envisioned for NEIS. Today's Internet services are predominantly text and data oriented with only sprinkles of graphics, still pictures and slow motion video. Only about one-sixth of the world population is connected to the Internet. Even with this level of service and connectivity, Internet is having serious problems of address space and bandwidth capacity. Experience shows that Internet is slow for many network applications and the quality of services is far from acceptable level in many cases. Internet is designed for data transport, and real time services like voice and video transmissions have serious quality problems. Internet is predominantly built over voice grade telecommunication infrastructure, its protocols have heavy overheads, and there are too many ad hoc solutions for problems encountered during operations. All these compound to almost insurmountable difficulties in bringing up Internet to any meaningful level of performance for NEIS.

The vision of NEIS calls for transportation of high quality audio including high fidelity music and high quality motion video apart from high-resolution graphics. With the present level of development and trend, support of such services on the present Internet is almost impossible. The key to the evolution of NEIS lies in building **Global Information Infrastructure (GII)** that would have adequate capacity and efficiency to support full-scale services envisioned for NEIS. What are these full-scale services and applications that would run on GII? How would such a massive infrastructure be managed? These are the subject matter of this unit. The way to GII is via **National Information Infrastructure (NII)** that needs to be set up by the each of the nations. In India, Telecommunications Regulatory Authority of India (TRAI) has proposed the setting up of a broadband fibre optic network spanning the entire country in the timeframe of 2011 – 2015 at an estimated cost of Rs. 30,000 crores.

Before we proceed further, two pertinent points are to be noted. First we need to make a distinction between services and applications. Services are provided by networks whereas applications are designed and executed by the users. Often, these terms are used interchangeably. But that is not correct. We illustrate the distinction by an example. Facilities offered by an e-mail server fall under network services whereas actual despatch of e-mail falls under user applications. A user requires network services to be able to run his/her applications.

Second is about the fundamental nature and forms of information. Advances in computer and communication technologies have brought about the representation, recording and communication of information in electronic form. Electronic information may be in analog or digital in representation. However, the terminology 'digital information' implies that the representation is entirely in digital form. At present, there is a perceptible trend towards the use of digital technology in both communication and computer fields. Everything electronic is moving towards digital technology. One may say that there is a digital revolution that is currently sweeping the world. As a result, electronic information is also going digital. These days, even sound and video are being recorded using digital technology. Many of you may be aware that many cinema theatres have modernised their projection system and use digital (Dolby) sound systems. Digitally recorded audio and video CDs are in common use today.

11.2 SERVICE AND APPLICATION TYPES

Finally, human beings are the ones who perceive information. Of the five senses, we use only vision and hearing for perceiving information. In vision, we make a distinction between language oriented text information and pictorial information. Accordingly, we have text and video as components of electronic information. One may now say that there are three fundamental components of electronic information: audio, text and video. All the three components are referred to as data, e.g. text data and video data. These three components together constitute **multimedia** information. We now move on to studying the different types of services and applications in the context of NEIS.

GII would bring digital data sockets to home that would support speeds of the order of 155 Mbps. Compare this with the present day so called broadband connection at 512 kbps. GII socket speeds would be about 450 times faster than the maximum speed that is available today. This is because of the ambitious aim of NEIS to offer studio quality video and imaging services that demand very high bandwidth. Imagine a communication facility at home that allows the members of the household to view a movie of their choice at any time of the day and for any chosen duration without any interrupting advertisements. The movie may be viewed partly, a bookmark created and resumed sometime later. Such a service, known as **video on demand** is one of most exacting services to be offered in NEIS. It is for this purpose that such high speeds are required in NEIS. In fact, the proposed precise socket speed is 155.84 Mbps corresponding to OC-3 standard of SONET or STM-1 of Synchronous Digital Hierarchy (SDH). You studied about SONET and SDH and the speed hierarchies in Unit 10.

Interestingly, all the network services or user applications that one can ever visualise for NEIS or otherwise can be placed under two broad categories:

- Interactive services or applications
- Distributive services or applications

A user needs interactive services to be able to run interactive applications and distributive services to run distributive applications. Interactive services are those in which there is two-way exchange of information. Such exchange of information may occur between two end users or a user and an application running on a server or between two application programs. Conferencing application where more than two end users participate is also interactive. Similarly, more than two application programs may interact with each other. Distributive services are those in which the information transfer is primarily one way, i.e. from a server to a user. Both the interactive and distributive services can be further classified as shown in Fig. 11.1. Interactive services include conversational services, messaging services and retrieval services. Distributive services include broadcast, multicast and unicast services as well as cyclic services.

Conversational services provide real time circuits (comprising two channels) for full-duplex communication among the conversing parties. The capacity and characteristics of the circuits are based on the requirements or demands of the user. Audio conferencing, video conferencing and collaborative research meetings are some example applications that run using conversational services. Please note that all these applications are interactive.

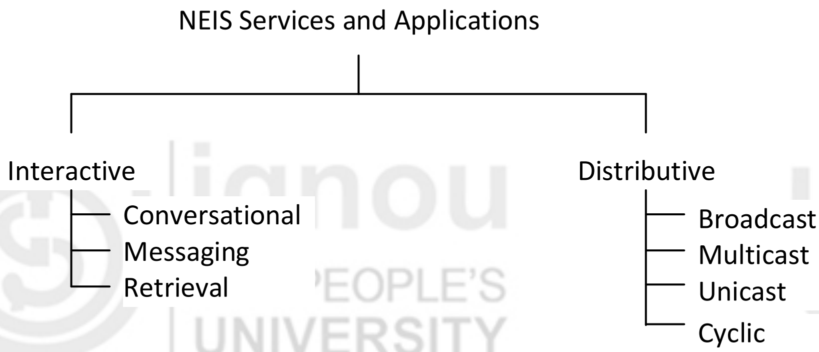


Fig.11.1: Categories of Services and Applications in NEIS

Messaging services are non-real time. They offer interpersonal communication among users using store-and-forward technique, mailbox or message handling functions. Since these services are not real time, they place less demand on the network resources. Typical messaging services include electronic text messaging, online chat, multimedia messaging, electronic mail, audio mail and video mail. As you know, electronic mail replaces the mailing of a letter. Similarly, audio and video mails replace the mailing of audiocassettes or videocassettes. Video mail is expected to become one of the most prevalent forms of messaging in NEIS. This is particularly so in the context of quality digital video cameras and monitors becoming available at affordable prices.

As students of Library and Information Science, you are familiar with retrieval services. Using retrieval services, user is able to access information stored in web sites or information centres that are, in general, available for public use. The service is interactive because the information is made available on user demand with specific queries. Based on the user query, a search process is initiated and the requested information is retrieved. Based on the output, the user can modify his/her query and obtain more specific information suiting his/her needs. The user thus has control over the information being provided to him/her.

Distributive services can be of one-time delivery type or cyclical delivery type. One time delivery includes broadcast, multicast and unicast. Broadcast services provide a continuous flow of information that is distributed from a central source to all the authorised receivers/users of the network. Every user has access to this information but has no control over it. Users simply tap into the flow of information.

Multicast is limited broadcast. The information from the central source is not distributed to all users but to a specified group of users only. The users are identified by a **group address**. Many groups may exist and users may be part of one or more groups. They may also be part of no group at all. Multicasting is directed towards groups. The information is delivered to all users of the group.

Unicast is also a form of distributive service. Here the information from the central source is delivered only to one user. Video-on-demand and music-on-demand are example applications that use unicast service. Here, there is a direct real time connection between the central server and the user. Thus, point-to-point connection characterises unicast application.

In addition to the above, there are two other related, but less used, forms of distributive services: **cluster cast** and **any cast**. Both these are also group based user services, but function differently. In cluster cast, the central source distributes the information to the nearest (usually) user who in turn distributes the information to all the other users in the group. There could also be other criterion for selecting the first recipient. In any cast, the information is distributed to only one of the users in the group. The information is not further distributed. The criterion for selecting a user could be based on who is less occupied currently. A typical application that uses any cast is query direction in a call centre. A user query is directed to the call centre executive who is currently free or to the one who becomes free next.

Cyclic information services distribute a set of information entities repeatedly to users in any one of the cast categories. The information is usually updated in every cycle. Example of applications include distribution of day temperatures, say every one minute or stock prices of selected stocks, say every 30 seconds.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) A mall has a video surveillance system installed. How would you categorise this application, interactive or distributive? Give reasons.

2) Give an example of text messaging system that is used commonly these days.

3) Differentiate between broadcast and multicast.

4) Differentiate between multicast and cluster cast.

5) Differentiate between unicast and any cast.

....

....

....

....

11.3 ELECTRONIC TEXT MESSAGING

Electronic text messaging primarily refers to Small Messaging Service (SMS) offered by mobile communication systems or Instant Messaging Service (IMS) offered by online chat rooms on the Internet. In the 1990s electronic mail used to be referred to as text messaging. But this is not appropriate in the present context for two reasons. One, text messaging implies short and quick communication in the present day context. Second, today's electronic mail is capable of carrying multimedia information. Text messaging is like conventional telegram. Electronic mail is more like conventional letter. Facilities in text messaging are limited and restrictive. Facilities offered by e-mail servers are fairly extensive. We study SMS in this section. E-mail service is covered in Section 11.5.

Most of you are familiar with SMS. Some of you might have used IMS and done online chats as well. Both SMS and IMS are part of messaging services and are interactive. In

SMS, the interaction may or may not be in real time. In IMS, it usually happens in real time. However, IMS generally supports a feature that allows a user to leave a text message for the other party even though he/she is not online.

SMS is a store-and-forward packet transfer service. It uses one of the control channels of the mobile network to send and receive short messages. The speed of the control channel used for SMS is fairly low. It generally operates at 110 bits per second (bps) or about 11 characters per second. Because of the limited screen, memory and processor capacity of the mobile handset, the length of short messages (SM) is limited to 140 bytes or characters. With 11 characters per second speed, it may take about 15 seconds for a message to be transferred. Longer messages are sent by concatenating multiple 140-byte messages. A continuity flag is used in the message header for this purpose. There is a SMS service centre (SMS-SC) in each mobile network that receives, stores and forwards the short messages.

Two types of short message services are supported on mobile networks:

- Cell broadcast SMS
- Point-to-point SMS

In cell broadcast SMS, a base station transmits short messages to all the mobiles in its coverage area. It is through this service that you receive a large number of advertisements on your mobile. If multiple base stations are chosen for transmitting the same message, then the number of mobiles that would receive the message goes up tremendously. This is how advertisements are distributed to a large number of users.

The base stations can be programmed to provide cyclic services or do multicast. In the case of cyclic services, information can be sent to users periodically. As an example, temperature in the coverage area of the base station can be sent to users, say every half hour or so. Using multicast, paid services can be implemented, i.e. information is sent only to users in a specified group, who have paid for the service. Using both cyclic and multicast, paid cyclic services can be implemented. Examples include providing cricket or badminton score updates.

Mobile networks support three categories of point-to-point SMS:

- Message originating from a mobile and terminating on another mobile
- Message originating from mobile and terminating on non-mobile device
- Message originating from non-mobile device and terminating on a mobile

Non-mobile devices include fax machines and personal computers. Receipt and delivery mechanisms differ in each of the above categories. Operating procedures also differ. Mobile networks use three special subsystems to implement SMS:

- SM service centre (SM-SC)
- SMS inter working Mobile Switching Centre (SMS-IWMSC)
- SMS gateway Mobile Switching Centre (SMS-GWMSC)

Mobile originated messages are first delivered to SMS-IWMSC, which in turn, passes the same on to SM-SC. Non-mobile devices receive and deliver messages from/to the SM-SC. SM-SC also connects to the SMS-GWMSC for delivering messages to mobiles. In effect, SMS-IWMSC acts as the input interface for mobile originated messages, SM-SC as input/output interface for non-mobile devices and SMS-GWMSC as the delivery interface for mobile terminating messages. As mentioned earlier, SMS is a store-and-forward service. Short messages cannot be sent directly from the sender to the recipient. They have to pass through the SM-SC.

Mobile-terminating short messages can be targeted to one of three destinations in a mobile:

- User specific
- Mobile equipment (ME) specific
- Subscriber Identity Module (SIM) specific

User messages are displayed to the users. ME-specific message may be used for activating or deactivating a function on the ME remotely. The SIM card processes SIM-specific messages. Using SIM-specific messages, special functions can be triggered by the network operator.

Limitation of 140-character messages in SMS has led to the evolution of what one might call as 'SMS Language' (SMSL). SMSL uses abbreviations to allow maximum use of the limited space as well as to compose messages quickly. Two techniques are used in framing abbreviations:

- Use of similar sounding letters or numbers to replace words or syllables
- Omission of letters from words, especially vowels.

The first technique is phonetic based one. Let us see some examples. The letter 'C' sounds the same as the words 'see' or 'sea'. The letter 'U' sounds the as the word 'you'. So the message "C U later" means "See you later" saves four characters in total length of 13 characters. This is a saving of about 25%. Similarly, numbers like 2, 4, and 8 can be used to substitute words or syllables that sound similar. For example, GR8 may mean 'great' and 'I wa8 4 u" may mean 'I wait for you'. Examples of the second technique are 'pls' for 'please' and 'msg' for 'message'. Another interesting abbreviation is 'ILU' for 'I love you'. Of course 'ILU' has been in use in Hindi movies for decades now.

Another feature of SMS is the use of what are called 'emoticons' that are representations of facial expressions formed using keyboard characters. Examples are :-) for happiness and :-(for unhappiness. Emoticons are quick and amusing way of conveying the emotions of the sender. Draw the above two emoticons on paper by rotating them clockwise by 90° and observe the emotions that they convey. Many SMS abbreviations are finding the status of accepted conventions for most users of the language. Example includes B2B that means 'Business to Business".

To simplify message generation, mobile handsets execute programs that use *predictive text input* algorithms. These algorithms predict the full word being typed based on the first 3 or 4 characters input. Some sophisticated algorithms even predict the next possible word.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

6) Distinguish between cell broadcast SMS and point-to-point SMS.

7) A car garage wants to send promotional offers to its registered customers on their mobile. Suggest a suitable SMS solution for the same.

8) What is the role of SM-SC?

9) A mobile operator has the capability to remotely activate WAP on mobiles. Which type of SMS is required in this case?

10) Recast the following sentence in SMS language 'Happy to see you tomorrow'

....

....

....

....

....

....

....

....

11.4 MULTIMEDIA MESSAGING

As you know, multimedia means a combination of video, sound and text. Multimedia messaging (MMS) usually means multimedia information being exchanged on mobile networks. In principle, transmission of multimedia information demands much higher bandwidth than text transmission (SMS). Maximum bandwidth is required when multimedia information comprises motion video with associated sound and overlaid text. MMS currently supported widely in mobile networks comprises only still picture. Any text in it is also a scanned picture. Such MMS requires relatively low bandwidth. But even this bandwidth is large when compared to the bandwidth required for SMS. As mentioned earlier, SMS is sent using control channels of mobile networks. Control channels have very limited bandwidth. MMS cannot be sent through these channels. Hence, MMS is sent via traffic channels of mobile networks. Traffic channels that carry voice have larger bandwidth and support higher data rates.

2G (second generation) or 2.5G mobile networks are still prevalent in most parts of the world. These networks do not have adequate bandwidth to transmit motion video etc. This is the reason why today's MMS is limited to still pictures. 3G mobile networks have larger bandwidth and hence can support the transmission of motion video. This is how these networks offer video calling and TV reception.

Digital image or still picture is stored and transmitted using a microscopic process. The picture is formed as a matrix of dots called *pixels* or *pels.* The word pixel or pel is a short form for picture element. The density of dots could vary from 75 dots per inch (dpi) to 2400 dpi both horizontally and vertically. The horizontal and vertical dot densities together are called the *resolution of the picture*. Larger the resolution, the better is the clarity of the picture. Larger the resolution, the higher is the bandwidth required for transmission. In 2G mobile systems, the picture resolution varies from 75 × 75 dpi to 300 × 300 dpi. The clarity is low at these resolutions. The pixels may be in colour or in black and white (B&W). Accordingly, the still picture is in colour or in B&W. In B&W, the pixels represent grey levels leading to the appearance of different shades. The pixel values are stored as binary numbers. Usually, grey values in B&W pictures are stored using 4-bit numbers. This allows 16 grey levels to be represented from black to white. Colour values require much larger binary string. Colour pixels use 8, 12, 16, or 24-bit representation. As a result, colour pictures require larger bandwidth for transmission.

There are three commonly used formats for storing and transferring digital images:

- Tagged Image File Format (TIFF)
- Graphics Image Format (GIF)
- Joint Picture Expert Group (JPEG) format.

TIFF has been developed as the common format for image scanners and DTP software. GIF has been developed for use on the Internet. GIF uses 8-bit representation for the pixels and hence can represent only 256 colours or grey levels. In this sense, it has limited resolution but the file sizes are small and can be transported easily across Internet. JPEG format is an image coding standard that has been optimised for continuous tone products such as photographs. It supports 16 million colours. In MMS, GIF format is used widely.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

11) A 1" × 2" picture is stored in GIF using 75 × 75 dpi resolution. How many bytes are required to store this image?

....

....

....

....

11.5 ELECTRONIC MAIL

Electronic mail, popularly known as e-mail, is the most widely used form of communication on the Internet and other computer networks for over two decades now. It has grown exponentially to the point where its volume per day is far in excess of the conventional paper mail.

A comparison of fax and e-mail is illustrative of the power of e-mail. Fax communication is terminal-to-terminal, whereas electronic mail communication is user-to-user. Being terminal-to-terminal communication, fax works on a circuit switched mode. E-mail is a store and forward (S&F) service and uses packet switching. In fax, messages destined to a number of users in the same office are sent to one terminal from where it is distributed by an operator or a messenger. The message is open and can be seen by the operator or the messenger. There is no privacy. On the other hand, electronic mail is delivered to the mailboxes of individuals. Privacy is ensured as the mail is delivered to individual's mailbox, which can be opened only by the intended recipient.

Being a person-to-person communication system, electronic mail turns out to be a cheaper alternative to telephone conversation and eliminates the time spent in establishing phone calls. For a telephone conversation to materialise, both the calling and the called party must be present simultaneously. It is not unusual that Mr X calls Y and Y is not present. Some time later, when Mr Y is in, he returns the call only to find Mr X is not at his desk. In fact, some studies indicate that as much as 70 per cent of the business phone calls during business hours do not succeed in the first attempt due to non-availability of the called party. Electronic mail permits communication between two parties without the parties actually being present simultaneously.

Another important advantage of electronic mail is its ability to reduce the consumption of paper in the office. Internal memos and reports can be exchanged electronically without using paper. Being a computer based messaging system, files prepared using office automation packages like word processor, database manager and spreadsheet package can be easily exchanged as electronic mail. This facility has the potential of improving office efficiency considerably.

E-mail systems have two major subsystems:

- User Agent (UA)
- Message Transfer Agent (MTA)

User agent (UA) usually runs on the user machine and the MTA on a mail server. When e-mail services offered by the Internet Service Providers (ISP), the UA functions are available on the web site of the ISP. This is sometimes referred to as web mail. Alternatively, the user may download the messages on to his/her machine and use a local UA for managing mails. Microsoft's Outlook express and MS outlook are examples of UA that run on user machines.

UA performs functions relating to the preparation, submission, and receipt of messages. It also assists the user in other message functions such as filing, replying, retrieving and forwarding. Message transfer agent (MTA) is concerned with transfer of messages across the network. It obtains messages from the source UA and delivers the same to the destination UA. On receiving a message, the MTA performs either a delivery function or a routing function. If the destination UA is in the same local network as the MTA, then the MTA performs a delivery function; otherwise it performs a routing function. The message is sent to another MTA that is en route to the destination. Another important function of the MTA is the reporting of the status of a message. The status may be 'delivered', 'rejected', 'lost', or 'destination unknown' etc.

Submission of a message by UA to MTA is via a protocol called Simple Mail Transfer Protocol (SMTP). Delivery of messages from MTA to the UA uses one of two protocols: Post Office Protocol version 3 (POP3) or Internet Message Access Protocol (IMAP).

We now list and discuss some of the important facilities offered by the UA of current mail systems:

- Mailbox management
- Address book
- Group mailing or mailing list
- Mail acknowledgement
- Mail encryption
- Digital signature
- File attachment

Present day mail services offer powerful mailbox management features. Standard components of mailboxes include:

- Inbox
- Outbox
- Sent Items
- Deleted Items
- Drafts
- SPAM

All these components can be further subdivided as folders much as the way files are organised in offices. For example, inbox may contain a folder for each contact. Items in these components can be arranged in a variety of ways: date wise: received or sent date, name wise, priority wise, subject wise etc. Such arrangements may be in ascending or descending order.

Outbox contains mails that are ready for despatch but not yet sent. The folder 'sent items' contains the messages that have actually been despatched. The folder 'Deleted items' is similar to a waste paper basket that has not been emptied. This folder can be emptied by selecting it and executing a 'delete' command. SPAM folder contains messages that are identified as SPAM, i.e. unwanted. SPAM mail is discussed in Unit 12 on Network Security.

Every e-mail user needs to have one or more e-mail addresses or ids. The address is of the form *username@mailserver.company.com*. It contains a user name, the mail server in which he/she is registered, the name of the company that owns the e-mail server and provides mail services and the top domain name. If the company is unique like *Yahoo*, then separate mail server name may not exist, e.g. rahul@yahoo.com. It is difficult to remember e-mail addresses in the above form. Address book helps us to overcome this difficulty. The easy to remember names of the users and their e-mail addresses are stored in the address book along with a lot more details like phone and mobile numbers, office and home addresses etc. While sending e-mail, user is selected by name and the e-mail address is automatically filled in by the UA from the address book.

Address books have provision to define groups or mailing lists. A group name is defined and individuals are added to this group. Messages may be sent to a group name. In this case, the message is despatched to all those who are part of the selected group. This feature is useful for distributing circulars or in a coordinated multi member project.

A user may request an acknowledgement for delivering a message. It is like registered mail with acknowledgement due. An acknowledgement may be sent as soon as the recipient opens and sees the mail. This may happen automatically or with the consent of the user.

E-mails can be encrypted for maintaining secrecy. Digital signatures may be affixed to indicate authenticity of the message. Encryption and digital signature are discussed in Unit 12 on Network Security.

E-mail systems allow files to be attached as part of e-mail. There is usually a limit on the number files that may be attached and the size of the individual files. There is usually no restriction on the type of file. As a result, files containing multimedia information can be sent with the mail.

Much as the conventional mail, e-mail also has the concept of **envelope** and **contents**. The envelope carries the destination address and other delivery information like priority. The source address is not carried as part of the envelope but is placed as part of the contents. The envelope also has provision to mark carbon copy (Cc) and blind carbon copy (Bcc) of the mail to others. Blind carbon copy hides from the recipient the fact that copy has been marked to others. The contents are delivered to the user entirely. The contents portion has two parts: **header** and **body**. The header part contains the source address and a subject line. The body contains the sender's message.

Although multimedia files can be attached with e-mail, often need was felt for embedding multimedia information in the body of the message itself. This need has led to the development of a protocol called Multi-purpose Internet Mail Extension (MIME). Using MIME message format, multimedia information such as video images and sound bites can be embedded in the message body.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

12) List any four advantages of electronic mail.

13) What are the functions of UA and MTA?

14) Match the most appropriate in the following:

A.	Carbon copy	F.	Message submission
B.	User names	G.	POP 3
C.	SMTP	H.	Envelope
D.	Mail downloading	I.	Mailbox
E.	Sent items	J.	Address book

15) What is group mailing? How it is done?

16) Where is the sender's address placed in e-mail?

....

....

....

....

....

....

....

....

11.6 INTERACTIVE TELEVISION (ITV)

Interactive television (ITV) is a network application for the user. First let us understand what interactive TV means. Today, television is a broadcast service. Different channels transmit pre-scheduled programmes. The user can select a channel of his/her choice and view the programme that is being telecast currently. The user will have to put up with interrupting and often annoying advertisements. The user has no control over what programme he/she can see and when. ITV offers flexibility and control in viewing programmes. Conventional TV transmission may be analog or digital whereas ITV is digital only. Conventional TV is available via sky, satellite or cable communication whereas ITV is available via network. Hence, pre-requisite for ITV is a broadband network communication facility at user premises capable of handling high quality multimedia, i.e. voice, video and data. This network facility allows two-way communication between ITV service provider system and the user TV system. In ITV, one-way video broadcasting system is turned into a two-way multimedia communication system. We may say that ITV is a television set with a return path to the ITV service provider. The return path makes it interactive. Information flows not only from the ITV service provider to the viewer but also from the viewer to the ITV service provider and/or his computer.

When fully developed, ITV would allow users to view a programme of their choice at any time of the day and for any chosen duration without any interrupting advertisements. The programme may be viewed partly, a bookmark created and resumed sometime later. Thus every individual user gets full flexibility of what to view and when. This is what is fully interactive television. Unlike broadcast service where the viewer can only watch one of the many programmes currently being broadcast, ITV gives the viewer an individual choice of content that is exclusive to that viewer.

ITV is not available commercially as of now and is in the evolution stage. ITV is evolving in three stages:

- Broadcast of TV programmes on the Internet
- Multicast of TV programmes on the Videonet
- Unicast of TV programmes on the Videonet

The first requirement of ITV is to make TV programmes available on the network. True ITV is possible only when NII and GII are in place. In the absence of these, attempts are being made to deliver video programmes on the Internet. Each ITV service provider broadcasts one programme at a time as per previously announced schedule. A user may log on to a video site and watch the programme that is being telecast at that time. This is much like conventional television except that the programme is available via Internet. This arrangement calls for Internet access from your TV set. This is provided by many ISPs today. Strictly speaking, there is no real interaction in this case except for logging in. However, this is the first step towards the evolution of ITV. Video is brought to your TV or computer using a technique called **video streaming**. Internet protocols have been evolved to support video streaming. We will learn more about streaming technique in Section 11.8.

In the second stage, the development concentrates on providing the necessary basic infrastructure for ITV. A video network called **Videonet** is being evolved. Videonet is a high-speed network capable of transmitting and distributing high quality video information. It consists of **video servers**, a **backbone network** and a set of **video switches** or **video distribution servers** located in what are called **distribution centres**. ATM with SONET is a suitable candidate for the backbone network. Video servers have massive storage facility and store a large number of programmes to choose from. Depending upon user requests, they despatch many programmes to the video distribution centres via the backbone network. To start with, video distribution centres will contain video switches. Video distribution servers will replace them in the future. Video switches distribute the programmes locally to the end users in multicast mode. Only one copy is received from the video server and distributed to many users locally.

This is like using group address. All the members of the group receive the same programme at the same time.

At this level of evolution, the user interaction improves in two ways: First, the user has a choice of logging on to one of many video distribution centres and second, he/she has a large number of programmes to choose from. Obviously, this is not yet real interaction as no individual can exercise any control on the programme that is being telecast. However, distribution centres may offer some control for the group as a whole. Group multicast is commercially viable because many people have interest in seeing popular programmes at any point of time.

The third stage of evolution aims at full interaction features. Full interaction implies every individual would have the ability to control and view programmes. This can be achieved only by unicast. The distribution centres will send individual video streams to each user who gets full control and flexibility in viewing. Such a service termed as **video on demand**.

How does a user interact with ITV server? Facility for interaction is another area of evolution in ITV. This is also evolving in three stages:

- Programmes are broadcast via Videonet and the user interaction path is provided via telephone network.
- Programmes are broadcast via Videonet and the user interaction path is provided via set-top box.
- Video on demand programmes unicast via Videonet and the user interaction path is via set-top box.

The set-top box is so called because it is usually placed on top of the TV set. It is more than a TV tuner. It has a computer with a phone, coaxial cable or satellite link to the ITV service provider and the Internet. There is a phone modem or a cable modem or a network card, which connects the set-top box to a public data network. Recently, television sets are being manufactured with built-in set-top features so that a separate box sitting on top of the TV is not required. This approach also points to what is likely to be the future home viewing set-up.

The potential use of ITV is enormous. To name a few, ITV may be used for marketing, advertising, child counselling, public relations, education and even politics. For the user, there is the promise of choice, fun, convenience and empowerment. Sitting at home, one will be able to get literally any product or service delivered at the touch of a button. This is known as *t-commerce*, television commerce. Users may click on advertisements to know more about the product. Viewers may choose camera angles while watching their favourite sports event. This indeed is exciting. Users may pause and resume programmes that they are viewing, provided the programmes are not live and are being delivered from video distribution servers located in video distribution centres. With live programmes that are not available on video servers, the users may record the programmes for later viewing.

ITV does not come free. Users will have to pay for ITV services. Two modes of payments are in vogue. First, a user pays a fixed subscription. Second is the *pay-per-view* model. In the latter, the user pays only for the programme s/he views. The two models are analogous respectively to the post-paid and pre-paid schemes that are used in mobile networks. Obviously, broadcast services are the cheapest and the unicast services are the most expensive. Multicast services fall in between.

There is one serious aspect of ITV, which is now spreading to Internet. All ITV systems have a feature called *click stream analysis.* This feature creates a complete record of the clicks that a user performs on his/ her set-top box. This record is later analysed to build profiles of users. In a positive sense, the purpose is to provide the user information that s/he is interested in a focussed manner. The negative aspect is that the service provider is actually treading into the private life of individuals and may use the information collected to blackmail users.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

17) Discuss ITV features.

18) What are the basic infrastructure components of ITV?

....

....

....

....

....

11.7 INTERACTIVE MUSIC

Interactive Music (IM) is an interactive network application. It is similar to ITV. It is also called **music-on-demand**. The user has flexibility and full control in listening to music on the network like Internet. The control includes functions like *pause, resume, stop, fast forward* and *fast backward*. Pause and resume are similar to creating bookmarks and continuing from thereon later. Music-on-demand services are evolving in two directions:

- Delivering music on the Internet in broadcast mode. This further evolves to **Internet radio**.
- Establishing audio servers and audio distribution switches to deliver audio programmes in unicast mode.

Unlike in the case of ITV, delivery of chosen music on multicast mode is not very attractive from a commercial point of view. It is unlikely that many people would choose the same song or music programme at the same time except when some major musical programmes take place, which are far and few. For example, many may hear a programme of Michael Jackson or a music festival but may not choose the same music on a day-to-day basis. In the case of ITV, popular TV programmes appear almost on daily basis. For example, there are over 500 movies are produced in a year. In addition, there are innumerable other TV programmes. A large number of persons have interest in viewing the same video programme at the same time. Hence, it makes sense to provide multicast for IRV. As in the case of ITV, charging for music on demand can be based on subscription or *pay-per-listen* mode. Pay-per-listen generally has an initial payment as well.

There are some differences between music-on-demand and Internet radio. In music-on-demand, the music bit is selected by the user whereas in Internet radio, the station plays out the programme. There is no user interaction for selection. User also has no control like *pause, resume* etc. Some radio stations play a second channel that is delayed by about 10 minutes to allow the users to take a break. After the break, the user may switch over to the delayed channel and listen to the programme from where s/he left. Another important difference is that music-on-demand is a unicast service whereas Internet radio is a multicast service. Many persons listen to Internet radio at a time. Hence, the same audio stream has to be sent to different destinations.

Audio is brought to your music system or computer using a technique called **audio streaming**. Internet protocols have been evolved to support audio streaming. We will learn more about streaming technique in Section 11.8.

11.8 APPLICATION DELIVERY

There are two fundamental ways in which music or video programmes are delivered on the network:

- Streaming technique
- Use of distribution servers

Both techniques are used for delivering audio and video. Streaming is more suitable for music. Distribution server approach is better suited for delivering video programmes. We discuss these two approaches in the following. We discuss *streaming* in the context of audio delivery and distribution server approach in the context of video delivery.

Streaming Audio

Streaming audio application is designed under client-server architecture. There is a media server on the server side in addition to the usual web site. There is a media player on the client side in addition to the usual web client, i.e. web browser. The media server and the media player interact via the convergent *real time protocol* (RTP). Initially, the user interacts with the server via the HTTP browser protocol for selecting the music bit. Thereafter, the media server and the media player come into picture to play the music in real time. The streaming audio configuration is depicted in Fig. 11.2.

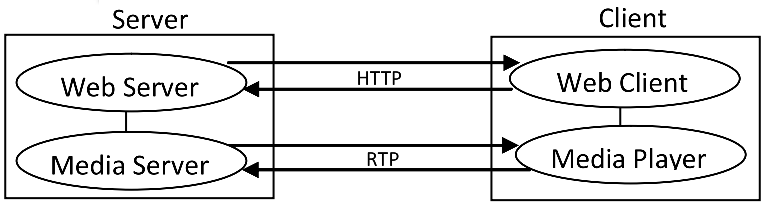


Fig. 11.2: Configuration for music delivery on Internet

In streaming audio, the digitised music is formed as small packets containing about 5 – 8 ms of music. Each digitised sample is 8 bits in size and is generated every 125 ìs. Each sample is not transmitted as it is generated. About 40 - 64 samples are grouped together as a packet and sent. Since packet transmission is on store and forward basis, different packets may take different time to reach the destination. Some packets may even be lost on the way. Delay variation and packet loss cause jitters in the music affecting the quality. Streaming audio takes care of these problems by providing a music buffer at client's end.

At the start of a listening session, the buffer is filled with music samples before the music is played out to the user. The buffer holds about 15-20 seconds of music. The user has to wait only for a short while until the buffer is filled to the required level before he can start listening to music. The music packets continue to arrive from the server while the media player is playing out the music and emptying the buffer contents. This is a streaming operation, i.e. on the one side the buffer is being emptied and on the other it is being filled. Hence the scheme is named as *streaming audio*. The provision of the buffer ensures that continuous music is available even in the presence of variable packet delays. The streaming operation is similar for video distribution as well and usually called **video streaming**.

Video distribution

Video distribution is done via **Videonet**. Videonet is a high-speed network capable of transmitting and distributing high quality video information. It consists of three parts: **video transmission centres (VTC)**, **backbone network** and a set of **video distribution centres (VDC)**.

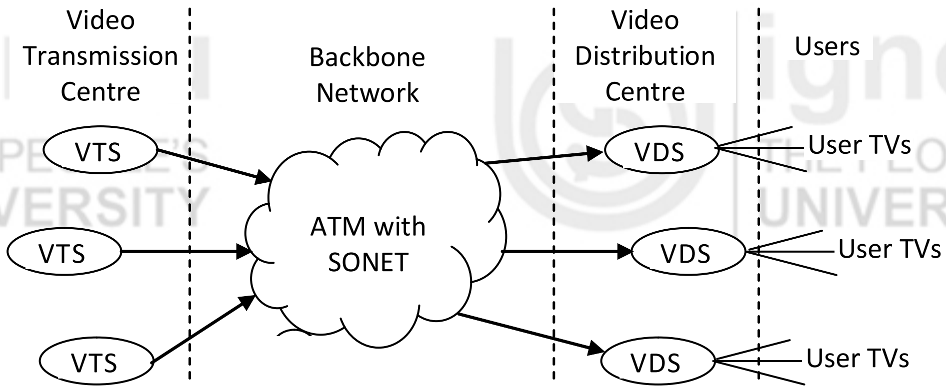


Fig. 11.3: Videonet Configuration

Video transmission centres house **video transmission servers (VTS)** or simply **video servers**. Video distribution centres house **video distribution servers (VDS)**. Video switches are used in place of distribution servers in case of video streaming. The configuration of Videonet is shown in Fig. 11.3 using VDS. Backbone network is broadband network like ATM with SONET. Video transmission servers have massive storage facility and store a large number of programmes to choose from. To get an idea of the quantum of storage required, let us consider storing movies. On an average, a movie requires 4 GB of storage. To store 10,000 movies, we would need storage of 40,000 GB or 40 terabytes (TB). With the present day technology, such capacities can only be obtained by using what are called 'disk farms'. Disk farms are a collection of disk arrays interconnected by a very high-speed network structure offering very large storage capacities.

Depending upon user requests, VTS despatch many programmes to the video distribution centres via the backbone network. Video distribution servers store the programmes locally and distribute the same to the end users in unicast or multicast mode. VDS is similar to VTS except that they have much smaller storage capacity, say 80 GB to store about 20 movies. The fundamental difference between video switches and VDS is that video switches do not have storage facility and stream the video to the users from the transmission servers in real time. VDS store the most recently requested video programmes by the users. Popular programmes remain locally in VDS, since many users view them. If a programme is available in VDS, the access is instantaneous. If it is to be downloaded from VTS, there is considerable delay. When a user makes a request for a programme, s/he is informed of the availability of the same locally or otherwise. VDS acts like cache memory for VTS.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

19) Why is streaming more suitable for music than the use of distribution servers?

20) Where is the buffer located in audio streaming? What purpose does it serve?

21) What are disk farms? Why do we need them?

....

....

....

....

11.9 PERFORMANCE ISSUES

As mentioned earlier, Internet is packet switched network designed for data transport. Packet switching is ideally suited for data transfer. Attempting to deliver real time services like music and video on a packet switched network gives rise to performance issues on account of the following:

- Time taken for packets to travel from source to destination varies from packet to packet.
- Packets may arrive out of sequence.
- Packets may be lost.

Performance issues arriving out of the first two can be taken care of by providing buffers at the receiving end and suitably managing them. The performance issue arising out of the third can be minimised by some special techniques. Let us see the issues and the solution thereof in the following paragraphs.

In packet switched networks, the packets are moved from node to node until the respective destination nodes are reached. At every intermediate node, packets may experience queuing delay. Depending on the queue lengths the wait times for the packets vary. Hence, even with the same route different packets may take different times to reach the destination. In real time services, such delay variation causes jitter. The sound reproduced at the receiving end is not natural and the video flickers. If the packets are delivered out of sequence, the resulting audio or video is distorted. If the packets are lost, the signal is broken. Parts of the audio and video are missed.

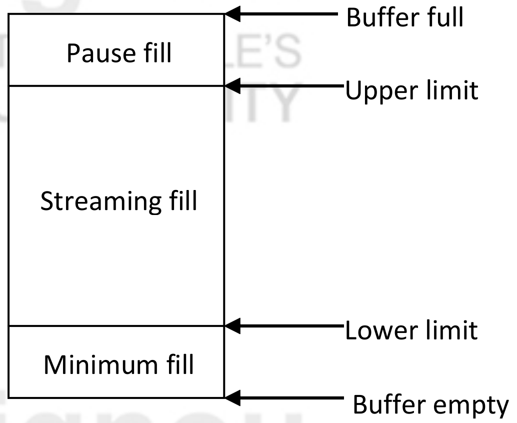


Fig. 11.4: Buffer management for streaming

Let us now see how buffering at the receiving end helps to overcome some of these problems. A typical buffer organisation is shown in Fig. 11.4. Buffer is part of media player at the user end. Refer to Fig. 11.2. When the user makes a request for a music or video bit to be played, the buffer is empty. The media server and media client interact and information streaming is started. The buffer is being filled. When the fill reaches the lower limit level, i.e. the minimum fill portion of the buffer is full; the media player starts playing the information to the user. The minimum fill portion of the buffer holds about 15-20 seconds of information. The user has to wait for this short period before s/he can start listening to music or viewing the video. While the media player is playing out, the media server continues to stream and fill the buffer. When the fill reaches the upper limit level, the media player asks the media server to pause the flow momentarily. 'Pause' is a flow control command. It controls the flow of information. There is a small time gap before the pause command takes effect. Information would continue to stream during this period and is received in the 'pause fill' portion of the buffer. The flow then stops but the media player continues to play the information to the user. The buffer is being emptied. When the information level reaches the lower limit level, the media player issues the flow control command 'Resume' to the media server. The information flow then resumes. There is a small time gap before the resume command takes effect. In the meantime, the media player continues to play from the minimum fill portion of the buffer to the user. On the one side the buffer is being emptied and on the other it is being filled. Delay variation occurring at the input side of the buffer does not affect the play out to the user. Out sequence packets can be sequenced on the input side without disturbing the flow on the output side. Thus user gets continuous and smooth information.

All flow control commands like 'pause' and 'resume' are implemented using a protocol known as *real time control protocol* (RTCP), which is used in conjunction with RTP. There are two other protocols associated with streaming services. They are protocols used for establishing a real time session before streaming operation starts and managing the stream flow once streaming starts. The protocols are *session initiation protocol* (SIP) and *real time streaming protocol* (RTSP) respectively.

Loss of packets results in break in signal to the user. The user will observe breaks in the information. In real time services, lost packets cannot be recovered in time to fill the gap. In order to minimise the effect of such packet loss, some media severs send alternate samples in packets, say odd numbered samples in one packet and even numbered ones in another. The two packets together provide the full information. If one packet is lost, the alternate samples are lost and the effect may not be perceivable by the user. If full packets of sequential samples are lost, the effect will be perceivable to the user.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

22) What purpose does the lower limit in buffer serve?

23) The size of buffer required for video streaming is much larger than that required for audio streaming. Why?

....

....

....

....

11.10 WHY NETWORK MANAGEMENT?

The complexity of networks is ever increasing. More and more users are converting their single device end equipment into multiple workstation local area networks (LAN). These LANs connect with each other via wide area networks (WAN) resulting in the well-known Internet infrastructure. At the applications level, the Internet infrastructure has given birth to what are known as **Intranets** and **Extranets**. An Intranet is a secure web-based private network that supports the business requirements of a corporate body across different geographical locations. An Extranet is an extension of the Intranet that connects a company's Intranet to the networks of its business partners and selected customers and suppliers. Both Intranets and Extranets operate on the same principles as Internet except that they are more secure closed networks amongst an identified set of users. In contrast, Internet is a public network.

Clearly, the management of such a complex network is not possible by manual means and calls for a set of automated and well-defined management tools and applications. To ensure interoperability of network management systems the international bodies like ISO, ITU and Internet Society have undertaken extensive studies and standardisation in the area of network management.

Historically, ISO is the first international body to give attention to network management aspects. As early as 1979, ISO initiated standardisation activities for network management as part of its standardisation process for ISO-OSI reference model. Its initial concern was to deal with management aspects of data networks. Later the work generalised to cover all types of telecommunication networks. ISO's network management standards are broadly covered under the subject heading **Common Management Information Protocol** (CMIP). ITU started working on management aspects in the 1980s. Its recommendations are covered under the subject **Telecommunications Management Network** (TMN).

The development of network management products based on ISO/ITU standards took time due to the efforts directed towards a comprehensive definition of features in these standards. In the meantime, the unprecedented growth in Internet led to a pressing and urgent need for network management tools for Internet. The Internet Society through its Internet Architecture Board (IAB) and its subordinate group Internet Engineering Task Force (IETF) decided to introduce a simplified version of network management derived from ISO/ITU proposals. In 1989, a standard known as **Simple Network Management protocol** (SNMP) was adopted for TCP/IP based Internets. The implementation of SNMP turned out to be a remarkable success demonstrating that the much needed network management tools actually solved real-life problems. The diffusion of SNMP, soon after its introduction, far exceeded the general expectations resulting in the definition of SNMP version 2 in 1993 and SNMP version 3 in 1998.

11.11 SIMPLE NETWORK MANAGEMENT PROTOCOL (SNMP)

CMIP, TMN and SNMP concentrated mainly on WAN management tools. In the early 1990s, it was realised that local networks (LAN) too require considerable network management support for a variety of purposes that we elaborate in the next section. For meeting such needs, a supplement to SNMP known as Remote Monitoring of Network (RMON) was issued in 1993 and an upgraded version RMON2 was issued in 1995.

Network management functions are placed under five broad areas as defined by ISO and adopted by ITU and Internet Society. These areas are:

- Fault Management
- Configuration and Name Management
- Accounting Management
- Performance Management
- Security Management

The order in which these areas are listed above leads to an easy to remember acronym formed by taking the first letter of each area: F-CAPS. Simple Network Management Protocol (SNMP) and Remote Monitoring (RMON) together are designed to support the above management functions. SNMP has three major components:

- The protocol itself
- Structure of Management Information (SMI)
- Management Information Base (MIB)

The protocol defines the format of SNMP messages and the rules on how the messages are exchanged. SMI specifies rules to name and define individual objects that need to be managed. It is also used to define the type of management information to be collected. MIB, as the name implies, is a database of all information objects with their attributes or variables and their values. MIB is maintained by each device and contains information about the managed objects in that device. Any device like router, bridge or switch is usually called 'Managed Device' if it implements SNMP.

SNMP is envisaged for use by environments defined by ISO, ITU and IAB (Internet). SMI allows definition of data objects for each of these agencies. For Internet SMI defines eight categories under which information can be maintained in MIB:

System: Information about the hardware, operating system and operations of the device. Example variable: System up time.

Interfaces: Information about individual network interfaces in a device. Example variable: Number of interfaces.

Address Translation: Information about address mappings: Example variable: IP address and the corresponding NIC address.

Internet Protocol: Statistics about IP. Example variable: No of datagrams received.

Internet Control Message Protocol: Statistics about ICMP. Example variable: Number of ICMP messages received.

Transmission Control Protocol: Statistics about TCP. Example variable: Number of TCP segments received.

User Datagram Protocol: Statistics about UDP. Example variable: Number of UDP datagrams received.

Exterior Gateway Protocol: Statistics about EGP. Example variable: Number of EGP messages received.

SMI and MIB are independent of the network management protocol used. Following SMI rules, vendors can define their own MIB variables in any of the categories. This is very useful for testing new products or enhancements to existing products in real network environment.

So far we have discussed about how management information is defined and organised. This information needs to be accessed and acted upon by the Network Management System (NMS). Access to management information is provided by the protocol. SNMP supports five commands for this purpose:

Get request: Fetch the value of a specified variable. This command is usually issued by NMS.

Get-next request: Fetch the value of the next variable. Name of the variable is not specified.

Get response: Reply by a device to a fetch command from NMS. The requested value is sent in this reply.

Set request: Issued by NMS to store a value in a specified variable.

Trap: Specify an event, on the occurrence of which the device sends a response to the NMS.

As is seen from the foregoing discussions, SNMP provides a means to specify and organise management information variables and to access the same by a set of commands.

As mentioned earlier, RMON was defined to support network management functions in LANs. They include the following:

- Monitor traffic type and network usage
- Identify network problems quickly
- Plan for future growth of the network
- Protocol decoding
- Protocol usage

RMON2 allows collection of statistics regarding protocol usage in a variety of ways such as:

- Network segment wise
- Network address wise
- Traffic wise between a pair of network addresses
- Application wise

Protocol analyser and LAN explorer are two devices that are used extensively for monitoring LAN segments under SNMP and RMON environment.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

24) What are the main components of SNMP?

25) What is RMON? What is its main purpose?

....

....

....

....

11.12 SUMMARY

This unit covered main user applications that run on Internet and network management. The applications discussed include non-real time ones like text and multimedia messaging and real time ones like interactive television and music on demand. To start with, a distinction is made between application layer network services and user applications. A user requires network services to be able to run his/her applications. User applications fall under two broad categories: interactive and distributive applications. Interactive applications may be conversation, messaging or retrieval oriented. Distributive services may be broadcast, multicast or unicast. Cyclic services are also distributive. In addition, two other forms of distributive services viz. cluster cast and any cast are also discussed.

Under electronic text messaging, the unit covers Small Messaging Service (SMS) in detail. SMS is of two types: Cell broadcast SMS and point-to-point SMS. In cell broadcast SMS, a base station transmits short messages to all the mobiles in its coverage area. Point-to-point SMS is between two mobile users or one mobile user and a landline user. Every mobile network has Small Message Service Centre (SM-SC) through which all SMS messages pass. SM-SC serves as a clearinghouse for short messages. Mobile-terminating short messages can be user specific, mobile equipment specific or SIM specific.

The unit then covers Multimedia Messaging Service (MMS). Transmission of multimedia information demands much higher bandwidth. MMS in 2G mobile networks supports only still picture transmission. 3G mobile networks have larger bandwidth and hence can support the transmission of motion video. Digital image is formed as a matrix of dots called *pixels.* The horizontal and vertical dot densities together are called the *resolution of the image*. Larger the resolution, the better is the clarity of the picture and the bandwidth required is higher. Three formats are commonly used for storing and transferring digital images: TIFF, GIF and JPEG.

Electronic mail application is then discussed. E-mail systems have two major subsystems: User Agent (UA) and Message Transfer Agent (MTA). UA performs functions relating to the preparation, submission, and receipt of messages. It also assists the user in other message functions such as filing, replying, retrieving and forwarding. Message transfer agent (MTA) is concerned with transfer of messages across the network. Some of the important facilities offered by the UA include mailbox management, address book, group mailing or mailing list, mail acknowledgement and mail encryption. Mailbox management covers management of inbox, outbox, sent items etc.

The discussions in the unit then turn to real time applications. Both Interactive Television (ITV) and Interactive Music (IM) offer flexibility and control in viewing and listening to programmes. ITV is also known as Video-on-demand and IM as music-on-demand. Both are two-way multimedia communication systems. The return path makes the systems interactive. Information flows not only from the service provider to the viewer but also from the viewer to the service provider. The systems allow users to view or listen to a programme of their choice at any time of the day and for any chosen duration. The programme may be viewed/heard partly, a bookmark created and resumed sometime later. Every individual user gets full flexibility of what to view/hear and when. Pay-per view or pay-per-listen are charging schemes in ITV and IM applications respectively.

Delivery of real time programmes on packet switched networks is discussed next. Streaming audio or streaming video technique is used on packet networks. Distribution servers are used on Videonet, which is a high-speed network capable of transmitting and distributing high quality video information. Performance issues relating to application delivery are then discussed. In packet networks, delay in packet delivery, out of sequence reception and packet loss give rise to performance issues.

Finally, network management aspects and the associated protocol for Internet are discussed. The ever-increasing complexity of networks calls for a set of automated and well-defined management tools and applications. ISO, ITU and Internet Society have undertaken extensive studies and standardisation in the area of network management. The respective standards are Common Management

11.13 ANSWERS TO SELF-CHECK EXERCISES

Information Protocol (CMIP) from ISO, Telecommunications Management Network (TMN) from ITU and Simple Network Management protocol (SNMP) from Internet society. Local networks (LAN) too require considerable network management support. A supplement to SNMP known as Remote Monitoring of Network (RMON) was issued for this purpose. SNMP has three major components: the protocol, Structure of Management Information (SMI) and Management Information Base (MIB).

1) Video surveillance system is an interactive application. The reason is that there are human beings monitoring the video images generated by the cameras. They can control the system. They can choose which camera to monitor or even pan the camera. Hence it is interactive.

2) An example of commonly used text messaging system is SMS in mobiles.

3) Multicast is a form of limited broadcast. In broadcast, information is distributed to all users, whereas in multicast, it is distributed to a group of users who form a subset of all users.

4) In multicast, the central source distributes information to all users in a group directly. In cluster cast, the central source distributes information to any one user in the group who in turn distributes the information to other users in the group.

5) In unicast, the central source distributes information to a specific user using point-to-point connection. In any cast, information is sent to any one of the users in a specified group.

6) In cell broadcast SMS, a base station transmits short messages to all the mobiles in its coverage area. Point-to-point SMS is between two users.

7) Cell multicast may be used as the solution as explained in the following. Since the promotional material is to go only to registered users, general broadcast cannot be used. We need to use group address for the registered users and multicast the messages to them. Registered customers may be spread out in different geographical areas. Therefore, all the concerned cells in different geographical areas must be used in multicast mode to achieve the functionality desired.

8) SMS-SC receives, stores and forwards the short messages. It acts as the input/output interface for non-mobile devices.

9) Wireless access protocol (WAP) is a mobile functionality. Hence, we require ME-specific (mobile equipment specific) SMS. In some cases, a SIM-specific message may also be required as the SIM card might perform some WAP related functions.

10) SMS language may be as “J 2 C U 2morow”

11) The picture size is 1" × 2". The resolution is 75 × 75 dpi. Therefore, the total number of pixels is (1 × 75) × (2 × 75) = 11,250. The picture is stored in GIF format. GIF uses 8 bits, i.e. one byte to represent pixel values. Therefore, the number of bytes required to store this image are 11,250 bytes.

12) Advantages of e-mail are listed below. Any four from the following list may be given as answer:

- Electronic mail communication is user-to-user and not terminal-to-terminal communication as in the case of fax.
- The mail is delivered to individual's mailbox and can be opened only by the intended recipient. Therefore, privacy is ensured.
- E-mail service uses store and forward (S&F) and packet switching form of communication. Hence, it uses network resources more effectively. It is cheaper than circuit switched services like fax.
- Local distribution of faxes call for a person whereas the distribution is automatic in e-mail.
- Electronic mail is a cheaper alternative to telephone conversation and eliminates the time spent in establishing phone calls.
- Electronic mail permits personal communication between two parties without the parties actually being present simultaneously on the network.
- Electronic mail reduces the consumption of paper in offices. Internal memos and reports can be exchanged electronically without using paper.
- Being a computer based messaging system, files prepared using office automation packages like word processor, database manager and spreadsheet package can be easily exchanged with electronic mail. This facility has the potential of improving office efficiency considerably.

13) User Agent (UA) performs functions relating to the preparation, submission, and receipt of messages. It also assists the user in other message functions such as filing, replying, retrieving and forwarding.

Message transfer agent (MTA) is concerned with transfer of messages across the network. It obtains messages from the source UA and delivers the same to the destination UA. On accepting a message, the MTA performs either a delivery function or a routing function. If the destination UA is in the same system as the MTA or is attached to the MTA directly, then the MTA performs a delivery function; otherwise it performs a routing function. Another important function of the MTA is the reporting of the status of a message such as 'delivered', 'rejected', 'lost', or 'destination unknown'.

14) Most appropriate matching are:

10.0 A – H

10.1 B – J

10.2 C – F

10.3 D – G

10.4 E – I

15) Electronic mail subscribers can be placed in groups. Mail can be sent to groups and it gets automatically delivered to all the members of the group. This is called group mailing. A group is also referred to as mailing list.

    Address books have provision to define groups or mailing lists. A group name is selected and individuals are added to this group. When a message is sent to a group it is delivered to all those who are part of the selected group.

16) The sender's address is placed in the header part of the contents portion.

17) ITV offers flexibility and control in viewing programmes. ITV is digital. ITV is available via a broadband network that allows two-way communication between ITV service provider system and the user TV system. ITV allows users to view a programme of their choice at any time of the day and for any chosen duration without any interrupting advertisements. The programme may be viewed partly, a bookmark created and resumed sometime later. Thus every individual user gets full flexibility of what to view and when. Unlike broadcast service where the viewer can only watch one of the many programmes currently being broadcast, ITV gives the viewer an individual choice of content that is exclusive to that viewer.

18) The basic infrastructure of ITV is Videonet. The infrastructural components of Videonet are video servers, a backbone network and a set of video switches or video distribution servers located in what are called video distribution centres.

19) In streaming, information is directly transferred to the user from the server. Distribution servers store information locally and distribute the same to the users. This is efficient if many people are likely to choose the same programme over a short period of time. It is unlikely that many people would choose the same music on a day-to-day basis whereas popular video programmes are viewed by a large number of people. Hence storing music locally is unnecessary overhead and not commercially remunerative. Therefore, streaming is better for music.

20) In audio streaming, the buffer is located at user end. The provision of the buffer ensures that continuous and smooth music is available to the user even in the presence variable packet delays.

21) Disk farms are a collection of disk arrays interconnected by a very high-speed network structure. We need them to build storage capacities in terabytes range for storing large quantity of information, particularly video.

22) Lower limit level of the buffer has two purposes. First is at the starting time when the buffer is empty. When the information fill reaches the lower limit level i.e. the minimum fill portion of the buffer is full, the media player starts playing the information to the user. Second is when the information is being played out and the server is in the pause state. In this state, the buffer is being emptied. When the information fill depletes to the lower limit level, the media player issues the flow control command 'Resume' to the media server. The information flow then resumes. There is a small time gap before the resume command takes effect. In the meantime, the media player continues to play from the minimum fill portion of the buffer to the user.

23) Video occupies larger bandwidth. The rate of the digitised video is much larger than that of the rate for audio. More bits are received per second. To store about 10 seconds of video, we require more storage space than that is required for 10 seconds of audio. Hence, the size of buffer required for video streaming is much larger than that required for audio streaming.

24) The main components of SNMP are the protocol, structure of management information (SMI) and management information base (MIB).

25) RMON stands for Remote Monitoring of Network. It is the tool for monitoring management information parameters in LANs.

11.14 KEYWORDS

Any cast	:	A form of information distribution: To Any one of the entities in a group.
Broadcast	:	A form of information distribution: To all entities in a network.
Cell broadcast SMS	:	A form of short message (SM) distribution: To all the mobiles in cell's coverage area.
Cluster cast	:	A form of information distribution: To the nearest (usually) entity which in turn distributes the information to all other entities in a group.
CMIP	:	Common Management Information Protocol evolved by ISO.
GII	:	Global Information Infrastructure.
Group mailing	:	A feature for sending e-mail to all users in a group.
Interactive Music	:	An interactive network application where the user has flexibility and full control in listening to music.
Interactive Television	:	An interactive network application where the user has flexibility and full control in viewing video programmes.
Media player	:	Software program at the user end capable of receiving and playing multimedia programmes to the user.
Media server	:	Software program at the server end capable of sending multimedia programmes to the media player.
MMS	:	Multimedia Messaging Service.
MTA	:	Message Transfer Agent: A component of e-mail systems.
Multicast	:	A form of information distribution: To all entities in a group.
Multimedia	:	Information containing audio, text and/ or video.
Music-on-demand	:	Another name for Interactive Music.
NEIS	:	Networked Electronic Information Society.
NII	:	National Information Infrastructure.

Pay-per-listen	:	A form of payment in Interactive Music.
Pay-per-view	:	A form of payment in Interactive Television.
Point-to-point SMS	:	Short Message exchanged between two users via SM-SC.
RMON	:	Remote Monitoring of Network
RTP	:	Real Time Protocol
SMI	:	Structure of Management Information: A component of SNMP.
SM-SC	:	Small Message Service Centre.
SNMP	:	Simple Network Management Protocol.
Streaming	:	Technique used for sending audio or video in real time.
TMN	:	Telecommunications Management Network: A network management standard evolved by ITU.
UA	:	User Agent: A component of e-mail systems.
Unicast	:	A form of information distribution: To only one user.
Videonet	:	A high-speed network capable of transmitting and distributing high quality video information.

11.15 REFERENCES AND FURTHER READING

Homer, Douglas E. *Internetworking with TCP/IP*, Volume I. 3rd Edition. New Delhi: Prentice Hall of India, 2001. Print

Lin, Yi-Bing. *Wireless and Mobile Network Architectures*. Singapore: John Wiley & Sons (Asia), 2001. Print

Mansfield, Kenneth C and Antonakos, James L. *An Introduction to Computer Networking*. New Delhi: Prentice Hall of India, 2002. Print

Simoneau, Paul. *SNMP: Network Management*. New Delhi: Tata McGraw-Hill, 1999. Print

Stallings, William. *ISDN: An Introduction*. New York: Macmillan Publishing Company, 1989. Print

Stallings, William. *ISDN and Broadband ISDN with Frame Relay and ATM*. 4th Ed. Singapore: Pearson Education Asia, 2001. Print

Tanenbaum, A. S. *Computer Networks*. 4th Edition. New Delhi: Prentice Hall of India, 2002. Print

Verma, Pramod K. *ISDN Systems: Architecture, Technology, and Applications*. New Jersey: Prentice-Hall Inc, 1990. Print

Viswanathan, Thiagarajan. *Telecommunications Switching Systems and Networks*. New Delhi: Prentice Hall of India, 2010. Print

UNIT 12 NETWORK SECURITY

Structure

12.0 OBJECTIVES

After going through this Unit, you will be able to understand and appreciate:

- Different aspects of and importance of information security;
- Internal and external threats to organisations;
- Computer system security and security over networks;
- Information properties prone to attacks;
- Information security attackers and their attacks;
- Authentication, authorisation and audit features;
- Password design and password policies;
- Password authentication process;
- Firewall features and the use of proxy servers;
- How secure transactions are done on web sites;
- Internet security features such as anti-virus, anti-spyware etc.;

- What are spam mails, phishing and cookies;
- Encryption and digital signature techniques;
- Why digital certificates are required; and
- How to send secure e-mail.

12.1 INTRODUCTION

In the pre-computer era, information in an organisation used to be secured physical and administrative measures. For example, a conventional library has manual checks at the entry/exit point. In other office areas, secure cabinets and restricted physical access were used for information security. With the advent of computerisation and networking, information is being stored and transmitted electronically. As a result, information security has shifted largely to electronic domain. Some aspects like access to server room are still controlled by physical means. In this unit, we are concerned only with securing electronic information and not with the physical aspects of security. Electronic locks are a part of physical security and are not discussed here. We are purely concerned with securing information that is either stored or transmitted electronically.

The electronic information security measures fall under two different classes: **computer security** and **network security**. The measures used for securing information stored inside a computer fall under the class computer security. The measures used while transmitting information over a network fall under the class network security. The business nature of an organisation and its interface with the external world would determine the extent of measures required in each class. In this unit, we shall be studying about both the classes of information security.

Threat to information resources may come from within an organisation or from external sources. Accordingly, we have **internal threats** and **external threats.** For example, internally, an employee of the organisation may try to tamper with the financial data to gain pecuniary benefits. Externally, a competitor organisation may try to steal the strategic plans of a company. Security measures are required in both cases. The extent of measures is usually based on the threat perception in each case. Threat is evaluated by studying the nature of the business of an organisation. The external threat is high in the case of organisations that provide public information services as in the case of a library or are involved in sensitive or controversial activities. On the other hand, an organisation that has an Internet connection for the benefit of its own employees may not face serious external threats. In this unit, we shall be studying about security measures to deal with both internal and external threats.

12.2 WHY INFORMATION SECURITY?

The answer to the question 'Why information security?' is as simple or complex as the answer to the question 'Why security in the normal life of society?' It is because there are some persons in the society who attempt to derive benefits from required to secure attempts to destroy others' resources in a surreptitious (covert) manner. Such attempts are to be prevented and thwarted. Security is meant to do this. Information is a valuable commodity and some persons attempt to derive benefits by exploiting others' information resources. Attempts may also be made to destroy other's information resources to cause harm to the owner. Hence there is the need to secure information resources. In the context of information security, one may associate four properties to any piece of information. These properties are:

- Availability
- Confidentiality or Secrecy
- Integrity
- Authenticity.

If information is not secured, one or more of these properties may be compromised. For example, if a communication link is broken or a hard disk stolen, then information becomes **unavailable**. Information must be available to bona fide users during specified times. Genuine accesses should not be denied. An attempt to make information unavailable is formally referred to as **denial of access** attempt. If an unauthorised person gets access to some information, then the **confidentiality** associated with that resource gets lost. This amounts to illegal copying of information even though it is not physically erased. If an unauthorised person modifies the contents in an information resource, then the **integrity** of information is lost. In such a case, the internal consistency of information is lost. For example, the contents of a philosophical website dealing with the axiom 'Unity in Diversity' is modified to include biodiversity and eco systems. In such a case the consistency of information is lost. Finally, if an unauthorised person is able to insert some counterfeit information objects into the system, then the **authenticity** of information is lost. This is equivalent to committing fraud. For example, an impostor may post a circular in the name of the Government of a country. Thus, information security is essential to preserve all the four properties.

People who tend to violate information security provisions are generally known as **intruders**. They are usually put under two categories: **attackers** and **hackers** of information resources. Attackers are more harmful than hackers. Attackers are further classified as **passive** or **active** attackers. Similarly, hackers are also classified as **white hat** or **black hat** hackers.

The main aim of hackers is to gain unauthorised access to information resources. White hat hackers attempt to expose security loopholes in a system and usually do not cause serious harm. They tend to annoy the owners of the system. Black hat hackers exploit security loopholes for personal gains or to harm someone. A typical example is an employee who tampers pay records to get monetary benefits for self or to cause monetary loss to a colleague. Another example is to change the contents of a website to make fun of the site. A reported case includes that of a hacker who modified the contents of a medical website dealing with abortions to preach religious teachings against abortions.

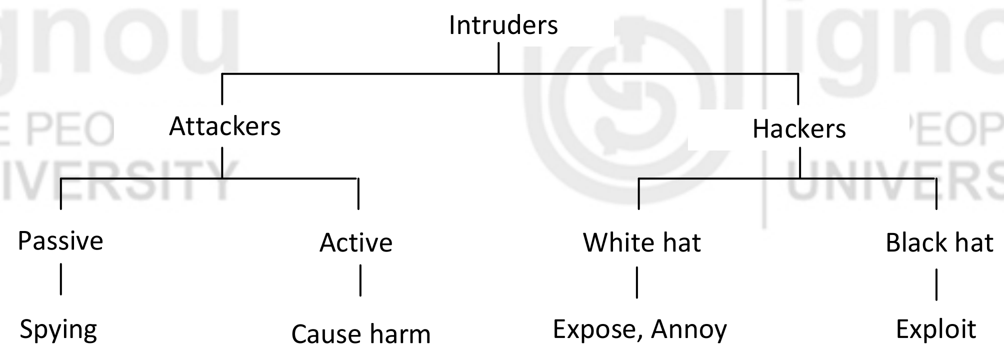


Fig. 12.1: Intruders of Information Systems

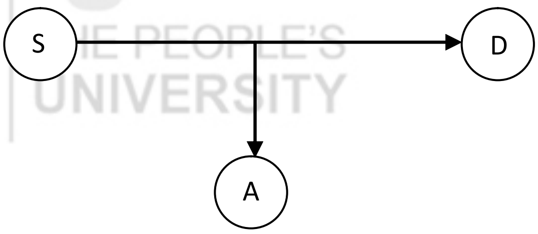
Attackers tend to break the security system and cause serious harm to the owners of the system. Their scale of operation is usually large and often borders espionage. Attacks on military sites and sensitive technology sites fall in this category. We will learn more about the type of attacks in the next section. Figure 12.1 summarises the intruder classification. The dividing line between black hat hackers and attackers is somewhat blurred. It is usually the scale of operation that distinguishes between the two. While a black hat hacker causes limited damage, an attacker could cause massive damage to the owner of the information resource.

12.3 TYPES OF ATTACKS

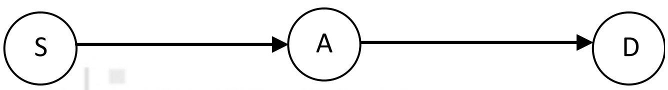
Passive attacks involve eavesdropping, listening to the traffic, monitoring the transmission parameters such as time and frequency of transmission, message length, and destination address. The contents of the message are not accessed. In the case of active attacks, the message and the transmission are tampered with in a variety of ways. Usually passive attacks precede active attacks. In the context of both passive and active attacks, the potential security threats may be listed as follows:

- Unauthorised access
- Eavesdropping
- Masquerade
- Modify
- Delete
- Delay
- Replay
- Reroute
- Misroute
- Flood network

Before we elaborate on the above security threats, let us see how attackers organise their attacks. Attackers use basically two techniques to organise their attacks and break the security system. These techniques are illustrated in Fig. 12.2.



(a) Eavesdropping



(b) Interspersed attacker

Fig. 12.2: Security attack modes

The normal flow of information is between a source (S) and destination (D). A passive attacker (A) just listens to the channel as shown in Fig. 12.2(a) and does not disturb the normal flow of information. An active attacker positions self in between the source and destination as shown in Fig. 12.2(b) and causes a variety of harmful effects. An incoming message may be blocked or deleted causing an interruption in communication between the source and destination. The incoming message may be modified and transmitted to the destination. The attacker may copy the credentials of the source and masquerade (pose oneself) as the sender and originate new messages on behalf of the sender. The message may also be delayed if time is critical in delivery. Replaying the message may cause the same action to be taken more than once, e.g. double payment. Rerouting or misrouting can be done when the attacker has access to alternative communication channels. Rerouting is usually done to send the message via the attacker's message analysis centre. The message may be copied there, analysed and even modified. In networks, routes are often chosen dynamically depending upon the traffic conditions. Hence, the destination may not be able to detect the deliberate rerouting at all. Misrouting is done to misguide the recipient and possibly cause a wrong action to be taken. For example, an army command to move the troops intended for western sector may be misrouted to eastern sector causing troop movement in the wrong sector. Rerouting and misrouting are serious forms of attacks.

An unprotected information system that is exposed to threats faces many **risks**, some of which are listed in the following:

- Information theft
- Unauthorised use or misuse of system resources
- Theft of services, i.e., steal and offer the same service at a lower cost
- Denial of service, i.e., genuine users is denied of services. This may happen if the attacker floods and chokes the network with unwanted information flow.

In view of the above risks, an information system must be protected. But, protection does not come free. There is a cost associated with implementing security measures. Ideally, one would like to have the best security measures. But the cost of such a solution may be prohibitive. Hence, the cost of security measures must be evaluated against the probable loss that may occur in the event of an attack. If the loss due to an attack is not very high, simple security measures will do. But if the loss is perceived to be significant, then stringent security measures are required. For example, the loss due to copying of a public information resource is not as serious as stealing the same resource. In this case, security measure against theft is more important than one for preventing copying.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Identify each of the following incidents as caused by hacking or by attack. Also identify if the incident is caused by a white-hat or black-hat hacker or by a passive or active attacker. In the case attacks, identify the type of attack. Also identify in each case, which one of the four information properties is compromised.

   a) A confidential governmental report is published in the media.

   b) An employee's pay record is tampered to give him more than his genuine salary.

   c) A literary website page is replaced with pornographic material.

   d) An instruction to a stockbroker for investment is delivered late to lose market advantage.

   e) An instruction to include four new users to a sensitive computer system is delivered with two additional bogus users.

   f) An online money transfer transaction gets executed twice.

g) A mutual fund receives instructions to change investment portfolio from a customer who actually never sent one.

2) Why do attackers reroute messages?

....

....

....

....

....

....

....

....

12.4 AAA SECURITY

AAA security is predominantly concerned with access to information resources residing on a server connected to a network or on a standalone computer. The access may come from a user or a computer program running on another computer or server. AAA stands for:

- Authentication
- Authorisation
- Auditing

Authentication is the process of ensuring that users who access the system are genuine. While genuine users are given access to the system, bogus users denied permission. Authorisation is the process by which the access to the information resources by a genuine user is restricted as per the access policy of the organisation. For example, a user may be permitted only to read the contents of a file but may not be allowed to modify the contents. Auditing is the process by which the use of the system is monitored. Audit raises alerts when a user makes an attempt to access resources in a way that s/he is not permitted to. Audit detects any unusual activity on the system and raises an alarm. Thus audit plays an important role in detecting attacks in the early stages.

AAA security may be implemented at many levels. The levels are:

- User
- Message
- Client machine
- Server
- Process
- Session

The most widely implemented AAA security is at the user level. Implementation at other levels is undertaken for specialised applications. For example, client machine authentication is implemented in stock market trading application. Server level authentication is necessary while accessing web sites that deal with financial transactions. In some applications, the authentication process needs to authenticate both the communicating entities, e.g. client machine and the server. In critical applications, the entire communication session is authenticated, i.e. every transaction or message exchange is authenticated. This is to ensure that no intruder comes on to the network after the initial authentication. In this unit, we confine our discussions only to the user level AAA security.

The most widely implemented user level AAA security is based on passwords. Password authentication involves three aspects:

- Password design
- Password policies
- Authentication process

Password must be designed to be complex but at the same time easy to remember for the individual user. Password should always be designed by the user and memorised by him/her. The designed password must be such that an intruder would not be able to guess or break the password. A complex password is designed by following five guidelines as given below:

1) Mix and use both upper and lower case letters
2) Use numerals as part of the password
3) Use special characters
4) Passwords must be at least eight characters long. The longer the password, it is more difficult to break it.
5) Password should not contain a string from the user name (ID).

The complexity of a password is usually rated as *weak, medium* and *strong*. For example, *boyandgirl* is a weak password. The password *boy12girl* is rated as medium and *Boy&12Girl* is a strong one. It may be noted that *Boy* and *Girl* are not usually parts of the user name. We have chosen strings that are not part of user name, mixed both upper and lower case letters, and included numerals and a special character in designing this password. It is also easy to remember this password. But it is difficult for a hacker or attacker to guess and break the password.

After having discussed the password design guidelines, we now look at the password policies. A typical set of policies is given below:

The password must have a minimum of 8 characters and a maximum of 32 characters.

1) The password must be changed every 30 days.
2) The system shall alert the user of the need to change the password four days prior to the expiry date. The alert shall appear every time a user logs on.
3) If the user fails to change the password by the due date, the system shall force the user to change the password the next time s/he attempts to log on. Unless the password is changed, the user shall not be allowed to log on.
4) If a user fails to enter his/her correct password within three attempts, his/her account shall be locked out. Only contacting the system administrator can then open the account.

The above policies constitute a typical set. The actual set may vary from organisation to organisation. For example, one organisation may enforce password change every 15 days whereas another may specify 45 days.

We now look at password authentication process. The main concentration of password authentication process is to hide the password from intruders. The passwords are stored in the system in an encrypted form. The encryption is often made irreversible so that the original password cannot be deciphered from the encrypted one. When the password is entered for authentication, it is encrypted and compared with the stored one. If a match occurs, the authentication becomes valid. When the password is transmitted on the network, it is encrypted, but this time using a reversible encryption process. The receiver system decrypts the received string to obtain the password. In this case, if the intruder catches the encrypted string, s/he would be able to crack the password or use the string to pose as the genuine user. To avoid this, sophisticated authentication algorithms are used and they produce a different encrypted string every time. Even if an intruder catches an encrypted string, it would be useless for the next log on.

We now consider the authorisation process. Once a user is authenticated, his/her access to the system is controlled by means of an **access control list (ACL).** ACL is the mechanism for limiting access to certain items of information based on user's identity. The ACL specifies the resources that a user can access and the permitted access for each resource. The resources include servers, directories, specific objects, folders and files. Access permissions range from full access at one extreme to total denial at the other extreme. A typical set of access permissions is given in the following:

- Full access
- Total denial
- Read only
- Read and Print
- Read and Copy
- Read & Write (Modify)
- Append
- Create
- Delete
- Change ownership
- Deny any

While read & print allows a user to get a printed copy of the resource, read & copy allows the user to save the resource as a file in the local user system. Append allows a user to add own material to the resource but prohibits him/her from modifying the existing contents. Change ownership permission is only given to system administrators and not to users. As opposed to total denial, deny any permits the system administrator to prohibit a particular access. For example, a user may be denied delete operation explicitly.

Auditing process tracks the activities of users and records specific events that appear to pose security threats. The events to be recorded and reported to the system administrator are specified in audit policies.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) What is the importance of auditing in AAA security?

4) How passwords are rated? Give one example of for each category of rating. Explain the reasons for the rating you have given for each password.

5) Distinguish between 'total denial' and 'deny any' access permissions.

6) Which access permission gives the user the power to modify the contents of an information resource?

....

....

....

....

....

....

....

....

12.5 FIREWALLS AND PROXY SERVERS

The literal meaning of firewall is a wall or a partition that is designed to inhibit or prevent the spread of fire. In the context of computer networks, firewall is a device that is designed to protect an organisation's local network from external attacks while permitting the local network users to safely access the external resources. It is a wall that prevents *electronic fires* (external security attacks) from spreading to the local network. It acts as a gatekeeper for the local network. No one can enter, (i.e., no access can be made to) the local network without the permission of the gatekeeper.

Firewalls are designed to operate at packet level, i.e., at the network layer level or transport layer level in the OSI reference model. On the other hand, **proxy servers** are computer systems that perform firewall functions at the application level i.e., at the sessions layer or applications layer level. Hence, proxy servers are also called as **application level firewall**.

Both firewalls and a proxy servers act as single entry/exit point where security checks and audit inspections are imposed. This is illustrated in Fig. 12.3. They are part of the internal local network but can connect to the external resources safely. They are designed to be immune to attacks or any penetration from outside or even inside. They use highly secure operating systems. In this sense they are called **trusted systems**.

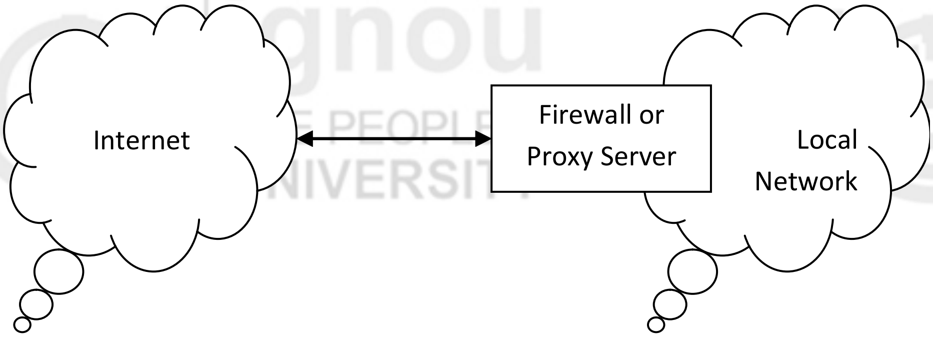


Fig. 12.3: Firewall or Proxy Server in the network

As mentioned earlier, firewalls are designed to operate at packet level. They use what is known as *packet filtering technology*. There are two versions of this technology:

- Static Packet Filtering technology
- Dynamic Packet Filtering technology

Static packet filtering (SPF) firewalls check the header information of incoming and outgoing service request packets and determine whether a service is to be established or not or whether a packet is to be forwarded or discarded. Forwarding or discarding decision is taken based on a set of security rules defined by the organisation. In general, SPF firewalls examine the following information for regulating the traffic:

- Destination address
- Source address
- Destination service port number
- Source service port umber
- Flag settings, where relevant

On the Internet, services are associated with certain port numbers. For example, web access takes place on port 80 and email receipt on port number 110. Over one thousand port numbers are reserved for specific services. Hence, by examining port numbers, a firewall can identify the service that is being requested. From the port number values and the corresponding security rules, the firewall can decide to block or allow a particular service.

Flag is a set of bits with each bit or a group of bits conveying a specific meaning. For example, one bit in the flag represents an acknowledgement for a request. If this bit is set, it means that the packet is a response to a request that has already originated under an established service. In this case, the firewall would pass the packet without any further checking. Once a bona fide service is established, there is a direct communication between the external and internal systems. In general, after the establishment of a service, SPF firewalls do not examine the subsequent packets of that service. This is to reduce the processing load and time overhead of examining the packets. Please note that every packet that is examined by the firewall contributes to delay in delivery of the packet.

Dynamic packet filtering (DPF) firewalls are more advanced than the SPF firewalls. In addition to establishing initial bona fide service connections, DPF firewalls continuously monitor the state of the service to ensure that no security lapses occur during the course of service communication. Obviously, the DPF firewalls add to delivery delay. To reduce this delay, they are designed to be more powerful than SPF firewalls. Consequently, they are more expensive.

Proxy servers are application level firewalls. They are also sometimes referred to as application level gateways. Unlike packet level firewalls, proxy servers do not allow any direct communication between internal and external systems. Instead they act like language interpreters between two speakers who speak different languages. The speakers carry on the conversation but never speak to each other directly. They only speak to the interpreter who translates what is communicated in one language into the other. Proxy servers function exactly this way. An interpreter tries to do a precise translation of what is being said unmindful of harmful or desirable effects the conversation may have. But, the proxy servers also examine the incoming and outgoing information to ensure that the security rules of the organisation are adhered to. In that sense, the proxy servers perform certain additional function when compared to an interpreter. A typical proxy server configuration is shown in Fig. 12.4.

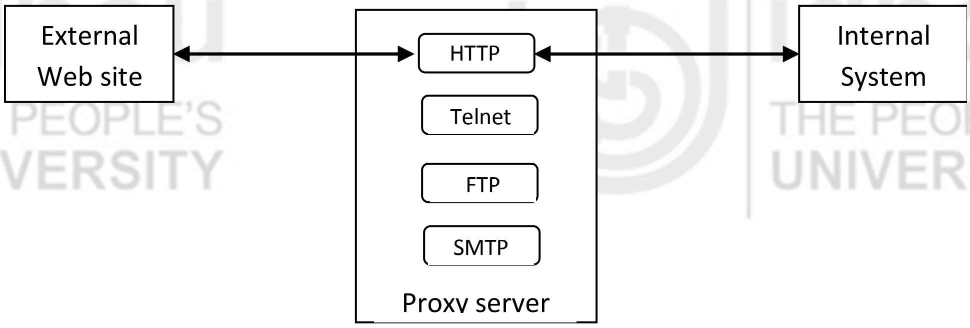


Fig. 12.4: Typical Proxy Server configuration

The proxy server shown in Fig. 12.4 supports four applications, viz. web access (HTTP), remote log in (Telnet), file transfer (FTP) and mail transfer (SMTP). A web access service connection is illustrated between an internal and an external system using hypertext transfer protocol (HTTP). Note that there is no direct connection between the internal system and the web site. The connection is via HTTP application in the proxy server. When the internal system wishes to access the web server, it sends a request to the proxy server. The proxy server, after ensuring compliance to the security rules formulates its own request to the web server. The web server replies to the proxy server and it does not even know the existence of an internal system. The proxy server examines the reply to ensure conformity to the security rules and then passes the reply on to the internal system. Thus all internal systems are protected from external systems.

Apart from the packet filtering firewalls and proxy servers, there is another type of firewall known as circuit type firewall. This is not used widely. A popular brand of DPF firewall uses the terminology S*tateful Inspection* to denote dynamic packet filtering. Finally, some packet level firewalls have other security features like anti-virus and anti-spam built in.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

7) What is the main function of a firewall?

8) How do firewalls and proxy servers perform their respective functions?

9) Distinguish between static and dynamic packet filtering.

10) What are the fundamental differences between packet level firewalls and proxy servers?

....

....

....

....

....

....

....

....

12.6 WEB SECURITY

These days, a large number of business houses, government agencies and many individuals have their own web sites. The number of web sites is growing day by day. Many business houses are currently upgrading their web sites to facilitate electronic commerce. At the same time, attacks on Internet and web sites are becoming serious. Almost all the four security properties of associated with information resources can be compromised by security attacks on the web sites. If an attacker gains read/write access to a web site, the integrity of information is lost. If an attacker is able to steal client data or business data from the web site, confidentiality is lost. Client data can be used to impersonate a client leading to the failure of authentication process. Hence, it has become imperative to design web sites that are secure and are immune to security threats.

Web security may be implemented at three levels:

- Network layer level
- Transport layer level, i.e. transmission control (TCP)
- Application layer level

At the network level, the emphasis is to make Internet Protocol (IP) secure. The secure Internet protocol is known as **IPSec**. When implemented at the transmission control level, the solution places emphasis on secure transport functions. Two most widely known solutions at this level are **Secure Socket Layer** (SSL) and **Transport Layer Security** (TLS). Application level security solutions are specific to each individual application. The solutions are implemented as part of the applications. Two well-known application level solutions are **Secure Electronic Transaction** (SET) and **e-mail security**. The SET is used widely in electronic commerce and financial applications like on-line banking, electronic stock trading and purchasing merchandise using credit cards. In this section, we will study about SSL and SET. E-mail security is dealt with later in another section. Equivalent of IPSEC in wireless networks is **Wi-Fi Protected Access** (WPA). Both IPSec and WPA are complex and their discussion is beyond the scope at BLIS level.

A company called Netscape Communications Corporation originated Secure Socket Layer (SSL) concept. The SSL solution has evolved over the years and is widely used presently. SSL is being adopted as Internet standard under the nomenclature Transport Layer Security (TLS). The discussions about SSL equally apply to TLS.

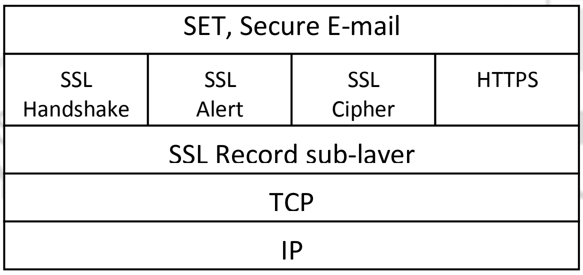


Fig. 12.5: SSL Protocol Stack

Socket is a connection end-point. SSL creates dynamically one connection end-point for each of the communicating entities. These sockets are then used for secure communication. SSL is designed to function as two sub-layers above the TCP layer in TCP/IP protocol stack. This is shown in Fig 12.5.

Immediately above the TCP layer is the SSL Record Protocol layer that provides basic security services to other higher level SSL and HTTPS functions. As you know, HTTP is the protocol used for web interaction. When HTTP is secured using SSL, the protocol is known as HTTPS, i.e. Secure HTTP. SSL handshake protocol is used for authenticating both the server and the client. SSL alert protocol generates SSL related security alerts. SSL cipher protocol deals with message encryption. We study encryption in a later section of this unit.

Secure Electronic Transaction (SET) is an application level service designed to protect credit card transactions on the Internet. SET specifies encryption and security protocols and standards to enable the users to utilise the existing credit card payment infrastructure on an open network like the Internet. The initiative for SET came from the major credit cad provider institutions like MasterCard and Visa. SET specifications support three functionalities:

- Provide a secure communication channel among the parties involved in a transaction, like the user, the merchant and the card issuer.
- Offer a trusted environment by issuing digital certificates.
- Ensure privacy among the transacting entities.

Digital certificates are explained in a later section of this unit.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

11) Why do we need web security?

12) What are the different levels at which web security is implemented?

13) What is a socket? How is it used by SSL?

14) How credit card transactions are secured on the Internet?

....

....

....

....

12.7 MALICIOUS SOFTWARE

A malicious computer program is one that harms the normal functioning of a computer. Malicious computer programs enter the computer systems via Internet, and removable mass storage devices such as floppies, pen drives and CDs. The rate of growth of malicious programs has been on the rise in the last ten years. It has also been observed that the cost of recovering from a malicious program attack is rising rather significantly. Downtime time for critical systems increase and services become unavailable for prolonged periods. In view of this, considerable attention is being given at present to prevent malicious program attacks. As you know, *prevention is better than cure*. Among the prevention measures are the use of malicious software removal tools and anti software like anti-virus.

A variety of malicious computer programs exist, virus being one of them. Malicious programs tend to exploit the vulnerabilities in system software, in utility programs such as editors and compilers and in application software. They fall under two broad categories:

- Programs that can exist on their own and run on their own
- Pieces of code that need a host program to run

Under the first category are programs called **bacteria** and **worms.** Under the second category are programs called as **trapdoors, logic bombs, Trojan horses** and **viruses.** Of these, bacteria, worms and viruses are capable of replicating themselves. Of the six different malicious programs mentioned above, we briefly discuss five of them in this section. Viruses are discussed in greater detail in the next section.

Bacteria are programs that generally cause **denial of service** attack. They do not affect the existing files or programs. Since they are independent stand alone programs, they replicate themselves and start growing exponentially, viz. one copy to two, two to four, four to eight and so on. Eventually, these replicated programs (say, 100s in numbers) take up all the resources of a computer system, viz. processor capacity, memory and disk space. The result is that no resources are available for normal programs and the regular services come to standstill.

Worms are malicious programs that spread from system to system via network connections. To travel from one system to another, they use one of the following methods:

- A copy of the worm program is sent as email from one system to another. It then executes itself on the new system. Then it further propagates to other systems in a similar fashion.
- A worm may execute itself on another computer via the remote execution feature of the operating system. For this purpose, the worm looks up tables to see which other computers are accessible and which of those allow remote execution. Then it selects a suitable system and moves to that system.
- A worm may log on to a remote system as a user and then use commands to copy itself on to the remote system. Then it travels to another system in a similar manner. This uses **Telnet** feature of the network operating system.
- A worm may also use file transfer facility (**FTP**) to copy its source code from one system to another.

Once active within a system, a worm can perform disruptive or destructive action within the system like bacteria or virus.

Often, programmers build in special entry points in programs at the time of development. Such entry points help them debug and diagnose programs. These entry points usually skip many standard procedures to enter the program. Some of the steps skipped may be security steps. Although many special entry points are closed before releasing the program in the field, some points are left open for field level diagnosis and debugging. In the context of information security, such open entry points are called **trap doors.** A person who is knowledgeable about the existence of such entry points can exploit the feature to enter the program bypassing security provisions and wreck the operation of the system. The existence of trap doors is verified by performing what is known as **penetration testing.**

Logic bombs are codes that get activated when certain conditions are met. Such conditions include particular date and time, the presence or absence of certain files, running of a particular application or the appearance or disappearance of particular user ID. A well-known example of logic bomb is the one that explodes on Friday, the thirteenth. If the date 13 happens to be a Friday, this bomb explodes. Another interesting example is that of a contractor who built in a logic bomb to explode if his payment was not released by a particular date.

A **Trojan horse** is a seemingly useful program that contains hidden code, which when invoked performs unwanted harmful functions. Attackers who are otherwise unable to get access to systems that they want to attack use Trojan horses. They would supply seemingly beneficial program (often free of cost) to a bona fide user. When the user executes the program, the malicious code becomes active. To avoid suspicion, the code actually causes harm after sometime (say, a day or two) so that the user does not suspect the supplier of the program.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

15) What are the commonly known malicious programs? Which of these need support to exist and execute?

16) How do bacteria cause denial of service attack?

17) What are the different network utilities through which worms spread?

18) Discuss the useful and harmful effects of trap doors.

....

....

....

....

....

....

....

....

....

12.8 VIRUSES

A computer virus is a malicious program code designed to cause harm to the normal functioning of computer systems. A virus is not an independent program. It needs a vehicle to ride and cause damage. The vehicle could be an executable program file or a data file like word document or a picture file. Much like the biological virus, a computer virus also spreads from program to program, data file to data file and from one system to another.

How does a virus get into a computer in the first place? People obtain programs, utilities and data from a variety of sources. These are imported into the system in a variety of ways. They are brought in through floppy disks, pen drives, CDs, downloaded from Internet or copied from another system using some network. It is this process of bringing in external program and data that launches a virus into a system. Obviously, day-to-day functioning calls for external interaction that is unavoidable. The only thing that one can do is to guard one's system from being infected by a virus. One needs to ensure that only virus-free programs and data are brought into the system. This is done by using what are known as **anti-virus** programs that keep a continuous watch on what enters the system. Anti-virus programs scan all incoming code and data to ensure virus-free import. They also scan every removable storage device whenever they are mounted on the system. Viruses, if found are neutralised or healed. If this is not possible, the import process is rejected. The user is given an appropriate message when scanning is done. The results of scanning are also displayed to the user. Despite all this, virus attacks take place. This is because the writers of virus code continuously innovate new techniques to befool the existing anti-virus programs. They produce new type of virus that is not previously known to the anti-virus programs. Such viruses enter the system. Then the anti-virus program is updated to deal with the new virus. This is a continuous process and that is the reason why it is essential to keep anti-virus programs updated on a daily basis and scan the system periodically. However, if one is using the system for surfing only trusted sites, downloading data or programs from certified sources only and using a mail system that has virus scanning built in, then the system can function safely without anti-virus programs. But the usage has to be restrictive and care should be exercised in choosing sites. Usually, the operating systems like Windows provide their own software firewalls that offer a good amount of protection.

There are two broad categories of viruses:

- Those attach themselves to executable program files
- Those attach themselves to data files.

The latter category of viruses is also known as **macro viruses**. Macro is a piece of code that is placed in data files. Macros execute themselves when the data files are open. If a virus is placed as a macro in a data file, it gets executed as soon as the data file is open. Virus then copies itself to other data files and thus spreads. There are many viruses that belong to the first category. The important ones are as follows:

- Parasitic virus
- Memory resident virus
- Boot sector virus
- Stealth virus
- Polymorphic virus

Parasitic viruses are the most common. They attach themselves to executable files. When an infected program file is executed, the virus becomes active and replicates itself by infecting other executable files. Memory resident viruses lodge themselves in main memory and infect every program that executes. Boot sector virus find place in boot record of the system and spreads when the system is booted. A stealth virus is explicitly designed to hide itself from anti-virus programs. Polymorphic viruses mutate (change) with every infection making it difficult to locate replicated copies even if the original has been detected.

During its lifetime a typical virus goes through the following four stages:

- Dormant phase
- Propagation phase
- Triggering phase
- Execution phase

When a virus enters a computer system, it usually remains idle (dormant). It then replicates itself (propagates) whenever the infected file is executed or opened. A virus gets activated (triggered) to perform its intended function based on some event, much like the logic bomb. Finally, the virus performs its destructive function. It is not essential that every virus goes through all the above four stages. For example, a virus may execute itself as soon as it enters the system.

Antivirus programs usually go through three phases of operation:

- Detection
- Identification
- Removal

When an infection takes place, the anti-virus program scans the system files to detect the presence of the virus. Once detected, the program identifies the specific type of virus and determines what is called the **signature** of the virus. The system is then scanned for the occurrence of the signature and the virus is removed wherever it occurs. Since polymorphic viruses change the signature at every replication, they are difficult to remove at one go. Multiple scanning may be required.

Antivirus technology has advanced significantly over the years. Starting with simple scanners, they now incorporate advanced features like heuristic scanners etc. The antivirus programs are considered to have evolved over four generations. The current most sophisticated programs are called fourth generation anti virus programs.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

19) How does a virus enter a computer system?

20) What are the functions of anti-virus programs?

21) What is a macro? What are macro viruses? Why are they most common?

22) What are the stages through which a virus moves during its lifetime?

....

....

....

....

12.9 SPYWARE, SPAM, PHISHING AND COOKIES

There is yet another class of malicious programs that do not cause harm directly to the functioning of the computer system but are of nuisance value or exploit personal data in other ways. Spyware is a malicious program that launches itself in the start up menu of a computer and gets executed every time the system is booted. This program tracks the activity of the users of the system and passes the collected information over the Internet to its parent site for analysis. In this sense, the program behaves like a spy. The rouge parent site may misuse the information collected to cause harm to the user. Some spyware programs can actually interfere with the functioning of the computer. They may change the Internet browser settings resulting in slow connection speeds or redirect access to non-genuine websites.

Spam is unsolicited advertising via e-mail. A spammer sends and advertisement to tens of thousands of e-mail addresses and mailing lists. A recent study shows that 95% of Internet traffic comprises spam e-mails. It is not completely possible to stop spam. Spam guards can be used identify spam mails and ensure that they do not carry virus or other malicious programs. The source address of the spammer can be input to the spam guard, which then will block all the mails from that source. However spammer will generate new source addresses and keep sending spam. As a result spam adds to the nuisance value.

Phishing is a type of deception to steal valuable personal data, such as credit card number, bank account number, e-mail ID, passwords etc. Attackers pose like genuine sites and seek secret information. For example, an attacker site may pose as your bank and ask for your bank account details citing some convincing reason. If provided, the attackers would then use the information to the customer's disadvantage. Phishing filters are used to guard against stealing of valuable information. Phishing filters monitor the web sites accessed to ensure that sites are genuine.

Cookies are packets of text sent by a web server to a web client. The client returns them unchanged during subsequent access to the web server. The web server uses cookies to keep track of the earlier accesses. Cookies are not programs and are usually harmless. They are neither viruses nor spyware. However, it is better to delete the cookies once a web access is closed. This can be done by appropriate tool available as part of the Internet browser.

There are two types of cookies:

- Temporary
- Persistent

Temporary cookies are deleted by the web site when the site access is closed. Persistent cookies stay in the system for a long duration. Some of them may have a specific lifetime (say, 15 days) after which they are automatically deleted. Others may stay permanently unless specifically deleted by the user. The persistent cookies are used to arrange the content of the accessed web site according to the user preference. Once the user indicates his/her preference, the required information is stored in the persistent cookie and is used by the web site to arrange the contents when it is opened next time. Cookies are sometimes called as web cookies, tracking cookies or HTTP cookies.

12.10 ENCRYPTION

Encryption is the art of hiding the meaning contained in a message. Decryption is the reverse process of encryption to extract the hidden meaning. The general field of study of encryption and decryption and related aspects is known as **cryptography**. Cryptography is an age-old technique for communicating secretly between two parties. Military personnel and persons in love have been using this technique for ages. There are two fundamental ways of encryption:

- **Substitution** where each character or a group of characters in a message is substituted with another.
- **Transposition** where the characters are jumbled but not changed.

When each character is substituted by another, it is called **monoalphabetic** substitution. When a group of characters is substituted by another group, it is called **polyalphabetic** substitution. We illustrate the substitution and the transposition techniques by taking a simple example. Consider the word 'COMPUTER'. A substitution encryption for this word is FRPSXWHU. Here, each letter has been substituted by the letter that occurs two letters after the current one in English alphabet, i.e. the third letter. The encrypted text, i.e. FRPSXWHU in this case, is often referred to as **cipher text**. Julius Cesar, the Roman emperor, first used monoalphabetic substitution encryption. Cesar's cipher used a shift of three letters, i.e. A is substituted by D, B by E, C by F and so on. A transposed encryption for the word COMPUTER is 'UTERCOMP', where the characters are shifted in a round robin fashion by four positions. The characters are the same as in the original word, but their positions have been changed, i.e. they are transposed.

Obviously, the simple substitution and transposition ciphers as illustrated above are easy to be broken by an attacker. By simple trial one can easily determine the shift value in these cases. The substitution and transposition can be made complex by using keys. For example, let us use the numeric key 12340123 to transpose the word COMPUTER. The key is as long as the text to be transposed. Each digit in the numeric key specifies the number of positions by which the corresponding character must be shifted in a round robin fashion. The transposition operation leads to the cipher ECROUMTP. The letters are jumbled and their order of occurrence is changed. But the letters themselves have not changed. Now it is more difficult to break the cipher. If only the sender and the receiver know the numeric key, then secret information can be exchanged between the two. The attacker will not be able to break the coded message unless he/she can figure out the secret key used for encryption and decryption.

In this example, the key length is the same as the text length, i.e. 8 characters. In practice, one cannot have very long keys. Hence, the message is broken into small blocks and the blocks are then encrypted by using the same key. Sophisticated encryption algorithms have been designed by using these two fundamental techniques and keys. Among them are Data Encryption Standard (DES), Triple DES (3DES) and International Data Encryption Algorithm (IDEA). A general cryptography scheme is shown in Fig. 12.6. User message is encrypted using the encryption key and the encryption algorithm to produce cipher text. The cipher text is decrypted using the decryption key and decryption algorithm to obtain the original user text.

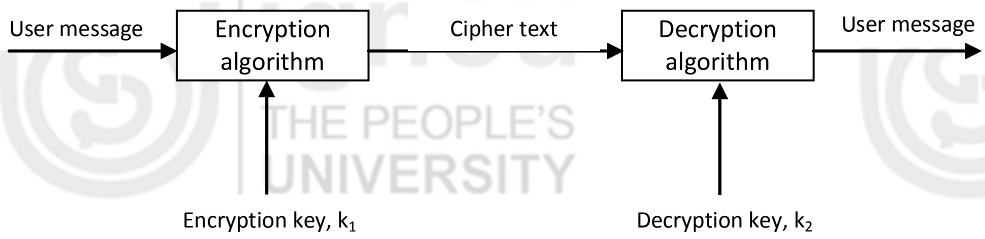


Fig. 12.6: Cryptography Scheme

We can think of encryption as the process of putting the user message in safe box and locking it with the key. The safe box is then transported to the receiver. Because the box is locked (cipher text), it is safe during the passage. Decryption is the process of opening the safe box with the key and retrieving the message.

Depending on the properties of keys used, two cryptographic schemes are in use:

- Symmetric cryptography where key k_1 = key k_2
- Asymmetric cryptography where key $k_1 \neq$ key k_2 but the two form a special pair.

Symmetric cryptography is like the one where a safe box has two identical keys. The same key can be used for both encryption and decryption. One key each is available with both the communicating parties. Both of them keep the key secret. Whenever one wants to send a secret message, he/she encrypts the message using the key. The receiver decrypts the message with the same copy of the key. The main problem with symmetric cryptography lies in the exchange of key between the communicating parties. Consider the case where the two parties are in different countries. The key should not be exchanged over the network because an attacker may copy the key. In high-risk situation, voice exchange over telephone line may not also be safe. The telephone line may be tapped. Postal communication may have the same constraint. Such serious security concerns apply only to military and other sensitive applications. Exchange of keys over telephone would work for most situations. However, many key exchange algorithms have been designed to ensure safe key exchange over the network itself. Because the keys are private to the two communicating parties, this scheme is also called **private key cryptography**.

Asymmetric cryptography can be likened to a safe box with two special keys that form a pair. The keys are such that if the safe is locked with one key it can be opened only with the other key. Both keys can be used to lock and open the safe except that the locking and the opening keys have to be different but from the same pair. Consider a key pair (k_1, k_2). If the safe is locked with k_1, it can only be opened with k_2 and no other key, not even k_1. Similarly, if the safe is locked with k_2, it can only be opened with k_1 and no other key, not even k_2. That is the special property of the keys. Every user on the network has this key pair generated. The user keeps key k_1 secret and makes k_2 public. If anyone wants to send a secret message, then the message can be encrypted using k_2 and sent to the user. The user decrypts the message using the key k_1. Since no key other than k_1 can decrypt the message, only the recipient with key k_1 can decrypt the message, which means the message has been passed on to the receiver secretly. Since one key is made public, this asymmetric scheme is also called **public key cryptography**.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

23) The string given below is encrypted using mono alphabetic substitution. Spaces between words have been deleted and the capitalisation has been hidden to make deciphering difficult. The encrypted string is:

    ymnxnxgqnxhtzwxjtknlstz

    Determine the original unencrypted message including spaces and capital letters.

24) Why is the symmetric cryptography called private key cryptography?

25) What is public in asymmetric cryptography?

26) In asymmetric cryptography, a message is encrypted by key k_2. How can this be decrypted?

....

....

....

....

12.11 DIGITAL SIGNATURE

Why do we use signatures? A signature authenticates the contents of a document. A signed document is an authentic source of information. Digital signatures are a means of authenticating electronic documents. We saw in the previous section as to how the asymmetric (public key) cryptography can be used to convey secret information. In this section, we will show that the scheme has inherent capability to authenticate documents as well. We will then describe how digital signatures are formed.

Consider an asymmetric cryptography key pair of some user *A*, (Ak_1, Ak_2). The key Ak_1 is kept secret and the key Ak_2 is made public. Anyone who wants to convey secret information to user *A* encrypts the message with Ak_2 and transmits the same to user *A*. This message can be decrypted only by using key Ak_1 that is kept secret by the user *A*. (Recall the key properties of asymmetric scheme). Even if an intruder intercepts the message while it is being communicated, he/she cannot read the contents, as he/she does not know the secret key Ak_1. Thus, information is conveyed safely and secretly to user A.

Now, consider the case when user A wants to communicate an authenticated document to someone on the network. He/she then encrypts the document using key Ak_1 and transmits the same on the network. Now, by key pair properties, this document can be decrypted only by key Ak_2. The key Ak_2 is publicly known to everybody and hence the recipient can decrypt the message. There are two things that you must note now. One, this message is not secret and an intruder who intercepts the message can read the contents using key Ak_2 that is public. Two, the fact that this message can be decrypted only by Ak_2 implies that it has been encrypted by Ak_1 that is known only to user *A*. Hence the message has been originated by user *A*. In other words, this is an authentic message from user *A*.

In the examples considered above, either a message secret or authenticated but not both. Often we need to send an authenticated message secretly. How is this possible? In fact, the asymmetric scheme is capable of supporting both secrecy and authentication simultaneously. This is achieved by carrying out two levels of encryption and decryption. We shall now see how this happens? In addition to user *A*, let us consider user *B* with asymmetric key pair (Bk_1, Bk_2). Let us suppose that user *A* wants to send an authenticated message secretly to user *B*. The two-level encryption and decryption process is depicted in Fig. 12.7. The user A performs two levels of encryption in the following order:

- First, encrypt the original message *M* with Ak_1. Let the encrypted message be *X*.
- Second, encrypt *X* with Bk_2 to get *Y* and transmit the same to user *B*.

The first encryption ensures authentication and the second secrecy. The original message goes through the following transformation: *M* ’! *X* ’! *Y*. At the receiving end user *B* performs two levels of decryption in the following order:

- First, decrypt *Y* with Bk_1 to get *X*.
- Second, decrypt *X* with Ak_2 to get the original message *M*.

The received message *Y* goes through the following transformation: *Y* ’! *X* ’! *M*.

Because the transmitted message *Y* is encrypted with the public key of user *B*, only user *B* can open it with his secret key. Hence secrecy is achieved. After first decryption, *B* gets message *X*. Since this is encrypted with the secret key of *A* this can be opened only using the public key of *A*. The fact that the message is decrypted with AK_2 implies that the sender is *A*, i.e. the message is authentic from user *A*.

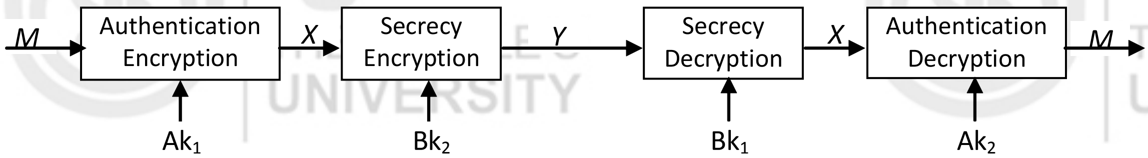


Fig. 12.7: Two-level asymmetric cryptography

Thus, the public key cryptography scheme is capable of transporting authenticated documents secretly. In other words, both authentication and secrecy are achievable in asymmetric cryptography. The problem of key exchange also does not exist. There is no secret key that needs to be exchanged as in the case of symmetric scheme.

Encrypting a full document is expensive in terms of processing time and communication overheads. In practice, there are numerous instances where the contents of a document are not secret but they need to be authentic. Examples include government circulars, court orders and department notices. All of them are public documents. However, their content must be authentic. This is where digital signature comes handy. Use of digital signatures reduces the encryption overheads. Here, the main contents of the document are in plain text and the signature is encrypted using asymmetric scheme.

Consider the example of a government circular to be issued on the network. The official who issues the circular appends his/her name encrypted using his/her own secret key k_1. Any recipient can verify the signature by decrypting it with the public key k_2 of the official concerned. This procedure does not ensure that the contents are safe. An intruder may intercept the circular, change the plain text contents but retain the encrypted signature. To safeguard from such an eventuality, the official concerned generates what is called a

message digest and encrypts the same along with his/her name. Message digest is short string of characters or numerals that unique to the message. Even if one letter is changed in the message, the digest value will change. Now recipient can generate the same message digest and match it with what is sent by the official in encrypted form. If a match occurs, the message is safe. Otherwise, the message has been tampered. The procedure for generating the message digest is public. Anybody can generate the digest and verify with what is sent.

Digital signatures can be used in private communication too. In other words, every user of the network may have one's own digital signature and send authenticated documents across. There is one problem here. What happens if user *X* poses as user *Y*? To safeguard from this problem, **digital certificates** are used. These certificates are issued by well known certified authorities (CA). These certificates are digitally signed by the CA and confirm that the sender is the genuine person and not a fake. Digital certificates are required only in sensitive applications where security threats are high.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

27) User *A* encrypts a message with key Bk_2 and sends the same on the net. Who can read this message and how? What does this operation ensure?

28) User A encrypts a message with key Ak1 and sends the same on the net. Who can read this message and how? What does this operation ensure?

....

....

....

....

12.12 E-MAIL SECURITY

Electronic mail is one of the most widely used applications on the network. The power and speed of electronic mail has prompted users and businesses to seek ways of sending confidential information over the network. Thus was born the need for secure electronic mail. There are two most popular schemes used for sending secure e-mail:

- Pretty Good Privacy (PGP)
- Secure Multi-purpose Internet Mail Extension (S/MIME)

PGP was largely developed by a single person, Phil Zimmermann in 1991, who made it available freely on the Internet. It is open source software that can be downloaded from the net. PGP provides the following features:

- Digital signature
- Message encryption for secrecy
- Compression
- Compatibility to e-mail systems
- Segmentation.

We have already studied digital signature and encryption. PGP offers facility to compress (zip) messages to economise on storage and transmission. Most e-mail systems put restriction on the number and size of the files that can accompany e-mail. Segmentation and reassembly at the receiving end are used to handle large messages that exceed the size permitted by e-mail systems.

Most of the electronic mail systems use Simple Mail Transfer protocol (SMTP) for transporting mail from one system to another. SMTP is an Internet approved protocol. There are some limitations with SMTP. To overcome these limitations, Multipurpose Internet Mail Extension (MIME) was developed. Security features have been added to MIME and S/MIME has been developed. S/MIME also has all the features of PGP and is more sophisticated than PGP in some ways. Prior to S/MIME, Internet task force developed e-mail security software known as **Privacy Enhanced Mail** (PEM). PGP and PEM are very similar. But, PEM never became popular. S/MIME was developed as a sophisticated replacement for PEM.

12.13 SUMMARY

This unit has dealt with information security issues in a networked environment. Information, being a valuable commodity, needs to be secured from hackers and attackers. If the information security is violated, the availability, confidentiality, integrity and authenticity of information may be compromised.

Access to information resources is secured by authentication, authorisation and auditing (AAA security) at different levels. User authentication is via passwords. They need to be designed to be complex but easy to remember. The complexity must be strong enough so that attackers cannot break the passwords easily. An organisation must enumerate password policies to ensure better security. The networked environment comprises LAN and Internet. Firewalls and proxy servers protect the organisation's LAN from attacks originating from the Internet.

World wide web (WWW) is used extensively in a networked environment. A number of applications involving financial transactions are being supported on the web increasingly. Such applications require security. Secure socket layer (SSL) and transport layer security (TLS) are two standards that support web security. Secure electronic transaction (SET) is an application-level standard that ensures secure credit card transactions on open networks like Internet. Wireless networks are becoming popular these days. Security in wireless networks is an important issue. Unsecured wireless networks are vulnerable for misuse. There are reported cases of unsecured wireless networks being hacked and used for terror activities. It is essential to install Wi-Fi Protected Access (WPA) feature with wireless networks.

In a networked environment, malicious software like virus, bacteria and worms enter user systems and cause damage to information resources. Some of the malicious programs exist on their own and others ride on executable user files or data files. The ones that ride on data files use macros and are called macro viruses. Macro viruses are more common than viruses that attach themselves to executable files. There is another class of malicious programs that do not directly damage information resources but are of nuisance value or attack indirectly by collective sensitive and confidential information. They include spyware, spam mails, Phishing attacks and cookies.

Encryption is the most important security tool used extensively in all security applications. Substitution and transposition are two fundamental encryption techniques. They are made very powerful by using keys. Encryption can be implemented in symmetric or asymmetric form. Symmetric cryptography is also known as private key cryptography and the asymmetric scheme as public key cryptography. While symmetric scheme ensures secrecy, asymmetric scheme can ensure both secrecy and authenticity. Hence, it is used more extensively on the Internet. Digital signatures affixed using asymmetric cryptography principle allow public documents to be sent in plain text while ensuring integrity of information that is authenticated. To facilitate secure digital signature use in private domain, trusted certificate authorities (CA) issue digital certificates.

Electronic mail is the most extensively used service on the network. In sensitive applications, information needs to be exchanged securely via e-mail. For this purpose, pretty good privacy (PGP) and secure multipurpose Internet mail extension (S/MIME) protocols have been developed. Both of them support encryption, digital signature and compression. They are compatible with most mail systems.

12.14 ANSWERS TO SELF-CHECK EXERCISES

1) The answers to different parts of this question are:

   a) White hat hacker. Confidentiality.

   b) Black hat hacker. Integrity.

   c) White hat hacker. Integrity and authenticity.

   d) Attacker. Delayed delivery. Availability.

   e) Attacker. Modify. Integrity and authenticity.

   f) Attacker. Replay. Authenticity.

   g) Attacker. Masquerade. Authenticity.

2) Attackers reroute the messages to analyse and cause an active attack. Usually, rerouting is done via the attacker's own server, where the message is analysed in detail. An attack strategy is then worked out. The attack is then carried out as per the strategy. The strategy may be to delay, modify or substitute the message. The onward transmission of the message may be blocked resulting in apparent breakdown of communication between source and destination.

3) Auditing brings out potential attacks before they happen. For example, attacker trying to crack a user's password may be caught if he does not succeed in, say, 3 attempts. Audit generates alerts about all unusual activities. These alerts can help prevent security attacks. Audit also alerts when an employee attempts to carry out activities that he/she is not expected to perform. These are raised as security alerts.

4) Passwords are rated as: weak, medium and strong. Examples: IGNOUBLIS is a weak password. This is weak because it uses two well known strings without numerals or special characters. I1g2n3o4U5 is a medium password. This is medium because it uses both letters and numerals. I$g5N6o&U* is a strong password. This is strong because it uses both upper case and lower case letters, numerals and special characters. It is easy to remember as the upper and lower cases alternate and special characters and numerals occur in a sequence on the keyboard.

5) 'Total denial' to an information resource means that the concerned user cannot access the resource at all. 'Deny any' means a particular access. For example, 'deny copy' prohibits a user from copying the contents of a resource but may allow him to read the information.

6) 'Write" permission would allow the user to modify the contents.

7) The main function of a firewall is to protect the internal systems of an organisation from external security attacks.

8) Firewalls and proxy servers act as a single point of entry to and exit from the organisation for information flow. They are part of the internal network and provide safe access to external resources. Refer to Fig. 12.3.

9) Static filtering examines the packets whenever a new connection is requested. In general, after the establishment of a service, SPF firewalls do not examine the subsequent packets of that service. Dynamic packet filtering (DPF) firewalls are more advanced than the SPF firewalls. In addition to establishing initial bona fide service connections, DPF firewalls continuously monitor the state of the service to ensure that no security lapses occur during the course of service communication.

10) Packet level firewalls permit direct connection between the internal user and the external resources. Proxy servers do not allow internal users to interact with external resources. Firewalls operate at packet level whereas the proxy servers at the application level.

11) These days, a large number of business houses, government agencies and many individuals have their own web sites. The number of web sites is growing day by day. Many business houses are currently upgrading their web sites to facilitate electronic commerce. At the same time, attacks on Internet and web sites are becoming serious. Almost all the four security properties of associated with information resources can be compromised by security attacks on the web sites. If an attacker gains read/write access to a web site, the integrity of information is lost. If an attacker is able to steal client data or business data from the web site, confidentiality is lost. Client data can be used to impersonate a client leading to the failure of authentication process. Hence, it has become imperative to design web sites that are secure and are immune to security threats.

12) Web security may be implemented at three levels:

- Network layer level
- Transport layer level
- Application layer level

13) Socket is a connection end-point. SSL creates dynamically one connection end-point for each of the communicating entities. These sockets are then used for secure communication.

14) Credit card transactions on the Internet are secured by the use of an application level service called Secure Electronic Transaction (SET) service. SET is designed to protect credit card transactions on the Internet. SET specifies encryption and security protocols and standards to enable the users to utilise the existing credit card payment infrastructure on an open network like the Internet. SET specifications support three functionalities:

- Provide a secure communication channel among the parties involved in a transaction, like the user, the merchant and the card issuer.
- Offer a trusted environment by issuing digital certificates.
- Ensure privacy among the transacting entities.

15) There are six commonly known malicious programs: bacteria, worms, trapdoors, logic bombs, Trojan horses and viruses. Bacteria and worms are self-supporting. Trapdoors, logic bombs, Trojan horses and viruses need some executable or data file to attach themselves.

16) Since bacteria are independent stand alone programs, they replicate themselves and start growing exponentially, viz. one copy to two, two to four, four to eight and so on. Eventually, these replicated programs (say, 100s in numbers) take up all the resources of a computer system, viz. processor capacity, memory and disk space. The result is that no resources are available for normal programs and the regular services come to standstill. Hence genuine users are denied of service, which is a denial of service attack.

17) There are four network utilities through which worms spread to other systems: e-mail, remote login (Telnet), remote execution and file transfer (FTP) facility.

18) Trap doors are built-in special entry points in programs. Such entry points help them debug and diagnose programs. This is useful. A person who is knowledgeable about the existence of such entry points can exploit the feature to enter the program bypassing security provisions and wreck the operation of the system. This is harmful.

19) People obtain programs, utilities and data from a variety of sources. These are imported into the system in a variety of ways. They are brought in through floppy disks, pen drives, CDs, downloaded from Internet or copied from another system using some network. It is this process of bringing in external program and data that launches a virus into a system.

20) Anti-virus programs keep a continuous watch on what enters the system. They scan all incoming code and data to ensure virus-free import. They also scan every removable storage device whenever they are mounted on the system. Viruses, if found are neutralised or healed. If this is not possible, the import process is rejected. The user is given an appropriate message when scanning is done. The results of scanning are also displayed to the user.

21) Macro is a piece of code that is placed in data files. Macros execute themselves when the data files are open. Viruses that attach themselves to a data file as a macro are called macro viruses. Macro viruses are more common because people import a lot more data files into the system than program files. Macro viruses come in as part of data files.

22) During its lifetime a typical virus goes through the following four stages:

- Dormant phase
- Propagation phase
- Triggering phase
- Execution phase

23) The decrypted string is: **This is BLIS course of IGNOU**.

24) In symmetric cryptography, the encryption and decryption keys are identical and are private and secret to thee sender and receiver. Hence, the nomenclature.

25) In asymmetric cryptography, the key k_2 of the key pair (k_1, k_2) is made public.

26) A message encrypted by key k_2 can be decrypted by using key k_1.

27) Since the message is encrypted using Bk_2, i.e. the public key of user *B*, it can only be decrypted using the secret key of user *B*, i.e. Bk_1. Only user *B* can read this message. Hence this process ensures secret message being passed on to user *B*.

28) Since the message is encrypted using Ak_1, i.e. the secret key of user *A*, it can only be decrypted using the public key of user *A*, i.e. Ak_2. The key Ak_2 is public and is known to everyone on the network. Hence anybody can read this message. This process ensures that the message is an authentic one from user *A*.

12.15 KEYWORDS

Access control list : A list used to control the access to information resources such as files by any user

Attackers : Those who cause harm to information resources like modify contents, deny access, steal data, replace the resource etc.

Authenticity : The guarantee that information is actually from the source it is claimed to be.

Availability : Availability of information resource to genuine users.

Cipher Text : The text obtained after encryption.

Confidentiality : Privacy of information.

Cookie : A piece of text message sent by a web site to user machine to keep track of browsing history by the user.

Cryptography : The branch of study covering encryption, decryption and other related aspects.

Decryption : The process of recovering hidden information from an encrypted text. It is the reverse process of encryption.

Digital certificate : A certificate issued by Certifying Authority assuring that the user of a public key is genuine.

Digital signature : An electronic signature to a document that ensures integrity of the contents and affirms that the source is authentic.

Encryption : The process hiding useful information from those oother than the intended recipient.

Firewall : A hardware device or software on a computer, which protects the internal network of an organisation from external attackers.

Hackers : Persons who violate security provisions to expose loopholes or for personal gains.

Integrity : A property of information resource that assures the contents are not corrupted.

Intruders	:	Those who enter the network with mala fide intentions.
Macro	:	A piece executable code attached to a data document like word document.
Macro virus	:	A virus that attaches itself to a data resource.
Malicious programs	:	Programs that enter user computer systems and cause harm to information resources or spy or cause nuisance.
Masquerade	:	Pose as a genuine user by deceitful means.
Pretty Good Privacy	:	PGP. It is a security service for e-mail.
Phishing	:	The process of collecting sensitive and confidential user information by deceitful means.
Proxy servers	:	Systems that hide internal users from external network.
S/MIME	:	Secure Multipurpose Internet mail Extension used for sending secure e-mail.
SET	:	Secure Electronic Transaction. Used for securing credit card transactions on open network like Internet.
Spam	:	An unsolicited e-mail often containing advertisements.
Spyware	:	A malicious program that spies on user activity.

12.16 REFERENCES AND FURTHER READING

SSL : Secure Socket Layer. SSL is a suite of services, standards and protocols for securing access to web sites.

Brenton, Chris. *Mastering Network Security*. New Delhi: BPB Publications, 1999. Print

Stallings, William. *Network Security Essentials.* Delhi: Pearson Education, 2003. Print

Tanenbaum, A. S. *Computer Networks*. 4th Ed. New Delhi: Prentice Hall of India, 2002. Print

Viswanathan, Thiagarajan. *Telecommunications Switching Systems and Networks*. New Delhi: Prentice Hall of India, 2008. Print

UNIT 13 E-MAIL AND E-MESSAGING

Structure

13.0 OBJECTIVES

After reading this Unit, you will be able to:

- provide a detailed account about Email and Email service Providers;
- explain in detail various Protocols used in Email service; and
- discuss about Web 2.0 tools in Email.

13.1 INTRODUCTION

Electronic Mail is one of the most prominent uses of networked communication technology. Better known as email, this is one of the most widely used forms of communication today. Electronic mail (email) has many advantages over other forms of communication: it is easy to use, free of charge, fast, and delivers information in a digital format. With suitable encoding methods, email can be used to send any kind of computer file, including pictures, sounds, programs, and movies.

Email started in 1965 as a way for multiple users of a time-sharing mainframe computer to communicate among themselves. ARPANET had a significant role in popularising email. The foundation for today's global Internet email service was created in the early ARPANET.

13.2 EMAIL

Email is the most popular and the quickest method of transferring message over network. There are several service providers over WWW, offering free email services. However, organisation can offer own mailing service for the stakeholders of organisation. The service is so popular and effective that many of the organisations operate only on email for internal and external communication.

13.2.1 Defining Email

Electronic mail, or email, is the transmission of text-based messages among networked computers. Email is one of the earliest and most basic messaging resources on the Internet and in many ways it still acts as the lowest common denominator for computer communications.

Features

- It is faster and more secure than conventional mail.
- It requires less physical effort to edit and send a letter of communication.
- Once the hardware, software, and Internet connection are in place, email on the Internet is free, even if messages is to be sent to the other side of the world.
- Unlike communication by telephone, email does not require the attention of both parties at the same time.

13.2.2 Needs of Email

Email has become an important and integral part for the people who are living away or who has someone far away. For different persons the reasons are different towards using email. In general the reasons are:

1) An email ensures faster/easier delivery of messages as long as email address is correct.
2) It provides time-stamped proof of an interaction. Also, many email services (such as Gmail) collate the conversation on the same subject into single threads.
3) It is more secure and inexpensive compared to other modes of communication.
4) It is easy to archive for future recall. Most of the email services provide search facility through emails.
5) An email can be edited and rephrased as much as it is desired before sending to the recipient(s).
6) It is easy to send the same piece of information to several people simultaneously such as circulation of memos, agendas, and minutes, or disseminate educational material.

13.2.3 Email Address

An email address is a unique address, which identifies a location to send and receive email. The email address contains ***username***, followed by an @ symbol, and then domain name i.e. username@domainname

e.g. abc@yahoo.co.in

An email address starts with a **user name** (*abc* in this case) that refers to the recipient's mailbox. Then, **sign** @ followed by the **host name** (*yahoo.co.in* in this case), also known as **domain name**. Normally, the Domain name has three parts (two part in case of United States) separated by two period (.) symbols. Reading from the left the domain name the first part is *yahoo*, is the name of a machine, which is a **mail server** or the computer where the recipient has an electronic mailbox. The first part of the **domain name** ends with period followed by rest of the part known as top-level domain (TLD). The TLD may have two parts the first part represents the type of organisation and the second part represents country code (according to the name of the country). In the given example, ***co.in*** is the top-level domain, where ***.co*** qualifies that ***yahoo*** is a company. This part of the domain name indicates the type of organisation (***.com*** represents commercial organisation whereas ***.gov*** refers to a governmental setup). The last part represents the country where ***yahoo*** has registered this machine or hosted. It is two characters long country code. In the given example, it is ***.in*** which represents that ***yahoo*** has registered/hosted this machine in India.

13.3 TYPES OF EMAIL SERVICES

There are different types of email services based on the type of provider organisation and the terms and conditions.

13.3.1 Free Web-based Email Services

These are the commonly used email accounts accessible through web browser (such Internet Explorer, Firefox etc.). They generally use HTTP Protocol for accessing mail. Example: Yahoo Mail, Hotmail, Gmail offer free web based email services. These services are run on financial profit from advertisements. Advertisers pay service providers to expose email account holders to advertisement.

The service providers give users a document to accept terms and conditions for use of service, which, one must before signing for the service. These services have a provision for spam filtration and virus scanning. Almost all the major services provide secure login using HTTPS (Hypertext Transfer Protocol Secure). HTTPS can be identified in the URL line while signing into the account (https://gmail.com or https://mail.yahoo.com).

13.3.2 Priced Web-based Email Services

There are also priced web based email service providers. Many of the free web based email service providers such as Yahoo! And Hotmail do offer premium account on payment basis. Advantage of utilising such services include more secured transaction of communication; personalised email address; better spam filteration as well as increased storage space. Priced email services are mostly useful in business related message exchanges.

13.3.3 Private Email Services

Most of the institutions/organisations have their own dedicated mail server and offer mail account for free to their staff and other members of the institution. These accounts exist only till one is member of that institution. For example: abc@ugc.ac.in; abc@satyam.net etc.

13.4 TYPES OF EMAIL ACCOUNT

Based on access to email provided, there are two types of Email Account

13.4.1 POP/IMAP Account

A POP account is based on Post Office Protocol and supports "offline" Email management. In a POP account, when a user connects to a mail server through a mail client, the client retrieves all the messages from the server, stores them locally and marks them as new/unread messages. Subsequently, the downloaded messages get deleted from the server and the connection is closed. Most of the ISPs (Internet Service Providers) offer POP mail.

An IMAP account is based on Internet Message Access Protocol. It allows their users to work with their messages in both online and offline modes. In this, the email client retrieves the message headers from the server and can store local copies of the messages in a local (temporary) cache. All the messages are left on the server until the user deletes them. This mechanism allows multiple email clients to access a single mailbox and is often used for corporate / business emails (e.g. sales@company-domain.com).

13.4.2 Email Forwarder

This type of email accounts will forward any incoming mail to another email address. Normally, all the service providers give email forward service.

13.4.3 Mailing List

A Mailing list consists of its subscribers/members' email addresses. Any email sent to a mailing list account will be distributed to all the subscribers of the mailing list. For example, lis-forum@ncsi.iisc.ernet.in

13.4.4 Auto Responder

Within an email account, an autoresponder can be set with a readymade reply to any incoming email such as *message for successful receipt of mail*; *vacation email*, etc.

13.4.5 Email Bouncer

An email bouncer enables to send a 'fake' bounce message to the sender. Thus, the spammer is led to believe that the email account is inactive or unrecognised by the server, and in many cases will remove the email address from its mailing list.

13.4.6 Email Blackhole

To avoid spam mails from certain addresses, a blackhole for those addresses is created so as to avoid/discard any messages coming from those addresses.

Self Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Discuss needs of email.
2) What is email forward?

....

....

....

....

13.5 STRUCTURE AND FEATURES OF EMAIL

Internet email messages consist of two major sections: Header and Body

13.5.1 Header

The header of an email is structured into various fields such as summary, sender, receiver, and other information about the email. The header of an email can be easily distinguished from the body of the email.

Various fields included within header are:

1) **From**: Contains the email address, and optionally name, of the sender of the message.
2) **To:** The email address(es), and optionally name(s), of the receiver(s) of the message.
3) **Subject:** A brief summary of the contents of the message.
4) **Date:** The local time and date when the message was originally sent.
5) **CC:** Stands for Carbon Copy. It contains email address(es) of those who will receive a copy of the message in addition to receiver(s) mentioned in To field.
6) **Message-ID:** It shows the number assigned to the message by the mail program at the host machine.
7) A series of **Received:** These are the lines, showing details of the systems through which the email has passed (useful for troubleshooting if mail bounces back).
8) A **Reply-to:** gives the preferred address for replies (usually but not always the same as the sender's address).

Note: Blank spaces are not allowed in an email address. Also, an email address is not case-sensitive.

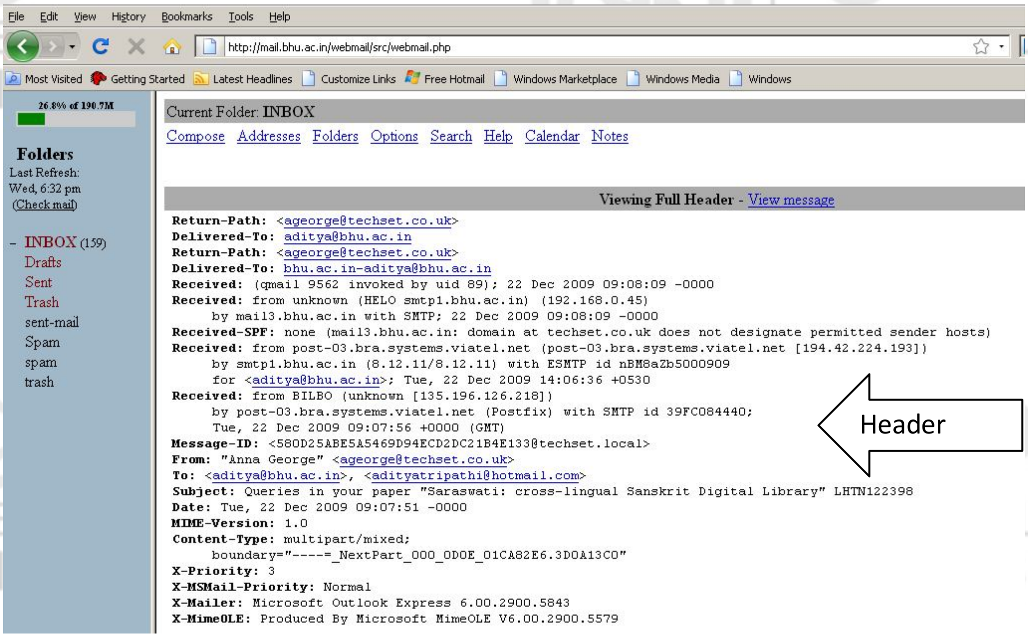


Fig 13.1: Header Section of an email

13.5.2 Body

The Body of an email contains the message itself mostly in text form. An account may be configured to automatically assign a signature (of the user) at the end. Signature is the text appearing at the end of the body by default in each message. Normally, it is the name of the sender and other contact details.

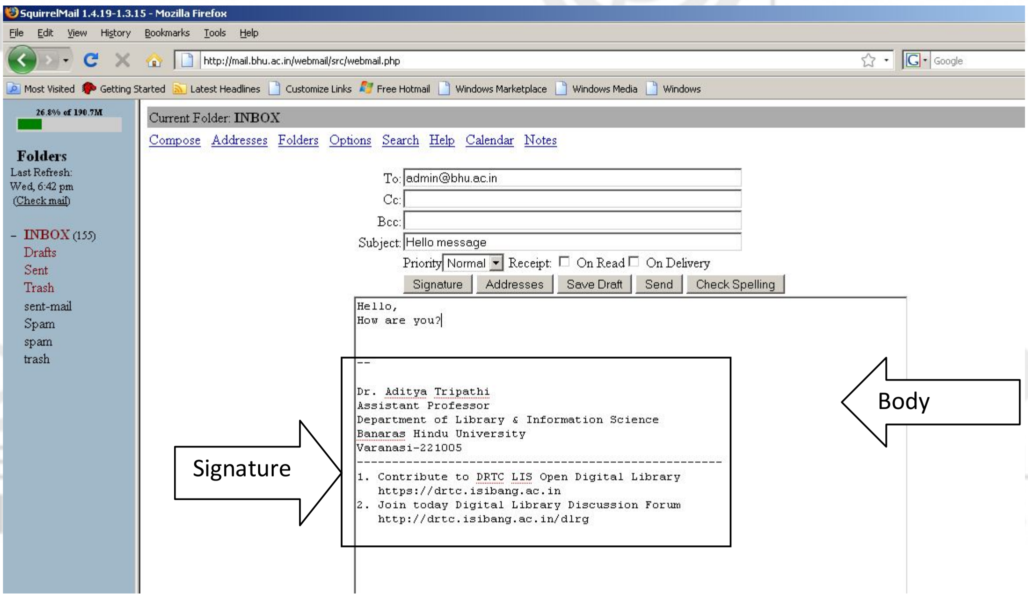


Fig13.2: Body Section of an email

13.5.3 Features

1) Email is based on push technology, i.e. Email is delivered to the recipient so they don't have to work to get it — they just need to open their Inbox to access the email.

2) Most of the email clients do offer to create MIME (Multipurpose Internet Mail Extensions) and HTML emails with colorful fonts, graphics and links.

3) The email account can be set to remind an upcoming event.

4) The accounts do offer facility for spell check while writing and email.

5) Any kind of document including multimedia objects can be sent through an email as attachments. However, the maximum size of the object, which can be attached or received, is fixed and set by the service providers. For example, MSN Hotmail can handle an object with 10 MB as attachment.

6) An email collates threads of communication on a single subject. Therefore, it sometimes acts as time stamped proof of communication. Gmail of Google offers this service.

7) A message can be saved/ printed along with all communication details.

8) A list of contacts (along with other details such as phone number, fax, etc.) can be created within an account with an ease to recall nickname associated with each email address. Hence, the user only has to enter one word for an email, instead of the full address.

9) Web-based email programs usually have a virus scan function that scans attachments before they are sent along with the main email.

10) One common feature of all email programs is the use of folders. These folders include an inbox, drafts folder, sent items folder and deleted messages folder. The users can also create other folders to better sort their email. Folders help categorise email according to its subject or importance. Filters are also included in email programs. These allow a user to define certain words or phrases that the program will look for in a message. The programs will then delete the message, forward it to a specified address, or put it in a particular folder.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) What are the parts of an email?

....

....

....

....

13.6 FUNCTIONING OF EMAIL SYSTEMS

In this section we will know how an email system works.

13.6.1 Protocols

Protocol can be defined as a set of rules to perform a specific task. The mail server and its client exchange information with each other using a variety of protocols. A protocol is a standardised mechanism used at each end of a communication channel, to achieve proper transmission of information. The most common of these email protocols is listed below:

IMAP Protocol

IMAP (Internet Message Access Protocol) is used by the client, which is used to read the email like Firebird, Outlook express, Apple Mail etc. (also called as Access Client). In this, email is received and held on Internet server for email account holder is read through a client. IMAP4 is the latest version in use. Only specific email messages requested by the user get downloaded from the server. It works well even for slower connection to Internet as it requires only a small data transfer. Also, one can create and manipulate folders or mailboxes on the server, make a keyword search through email, delete messages etc.

POP3 Protocol

The **POP (Post Office Protocol 3)** protocol provides a simple, standardised way for users to access their mailboxes and download messages on their computers using email clients.

This protocol permits to download all email messages from the mail server to a local computer. One can also choose to leave copies of their emails on the server as well. The advantage is that once the messages are downloaded mail can be read without an internet connection. Most current email client applications support POP3.

SMTP Protocol

The **SMTP (Simple Mail Transfer Protocol)** protocol is used by the Mail Transfer Agent (MTA) to deliver an email to the recipient's mail server. The SMTP protocol can only be used to send emails, not to receive them. The vast majority of mail servers use SMTP.

HTTP Protocol

The **HTTP protocol** is not a protocol dedicated for email communications, but it can be used for accessing mailbox. This protocol can be used to compose or retrieve emails from an email account. Hotmail, Yahoo are few examples which are using HTTP as an email protocol.

13.6.2 Delivery Agent

Mail Transfer Agent

Mail Transfer Agent is a piece of software which transfers messages or mails from one host or machine to other. It is often referred as mail server. MTA is the program which forwards message from one machine to other and finally delivers to the destination. An MTA receives a mail and puts a Received stamp in the header of the mail and checks for the recipient on the host. If the recipient is not present on the host MTA transfers the message to next host and thus passing through various hosts message is delivered to the recipient. The message is forwarded using SMTP. Example of MTA is qmail, sendmail, postfix etc.

Mail Delivery Agent

A mail delivery agent or Message Delivery Agent (MDA) is computer software used by the Mail Transfer Agent to deliver email to a particular user's mailbox. When the mail server receives a mail or in other words MTA (Mail Transfer Agent) using SMTP than the mail is placed in the targeted user's mailbox. This work is done by MDA locally. Some of the mail delivery agents are: binmail, delivermail; maildrop, postfix-maildrop etc.

13.6.3 Access Client

The term *email access client* refers to an agent acting as a client towards an email server to access an email account. It is kind of application software. Some of the examples of Access Client are Outlook Express, Outlook or Thunderbird.

Note: Thunderbird is a free open source, cross-platform email client that has quickly gained huge popularity, with a reason. It is extremely easy to use, fairly powerful and robust.

13.6.4 Setting up Account

Several websites provide free email services. Following is an example for creating an email account on Yahoo! Mail.

1) Go to the website **https://mail.yahoo.com/** (Fig. 13.3). This provides a login page for the users who already have a mail account with Yahoo. For new users of Yahoo, they need to utilise **Sign Up**.

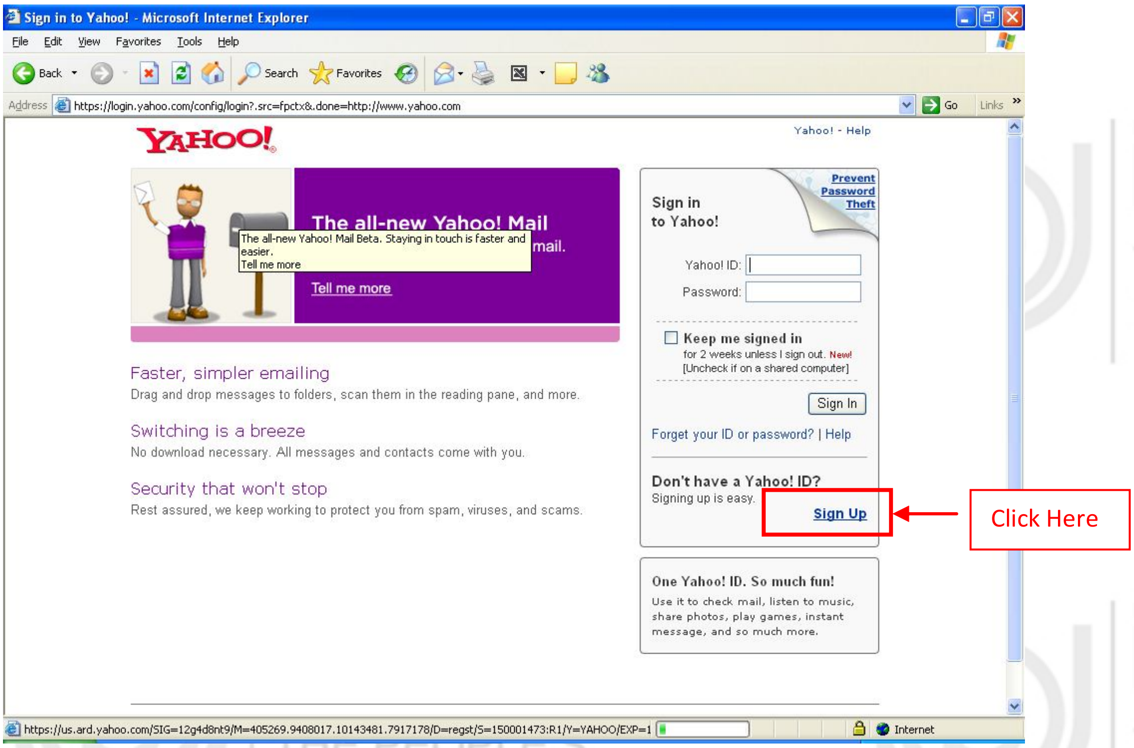


Fig 13.3: Yahoo Mail Homepage

2) Click on **Sign Up**. This will open a Registration Form (Fig. 13.4).

3) Enter the registration details and choose an ID and password for your Yahoo account. Availability of a Yahoo ID can be checked through the button Check. If an ID is unavailable, this will enlist other options for Yahoo ID.

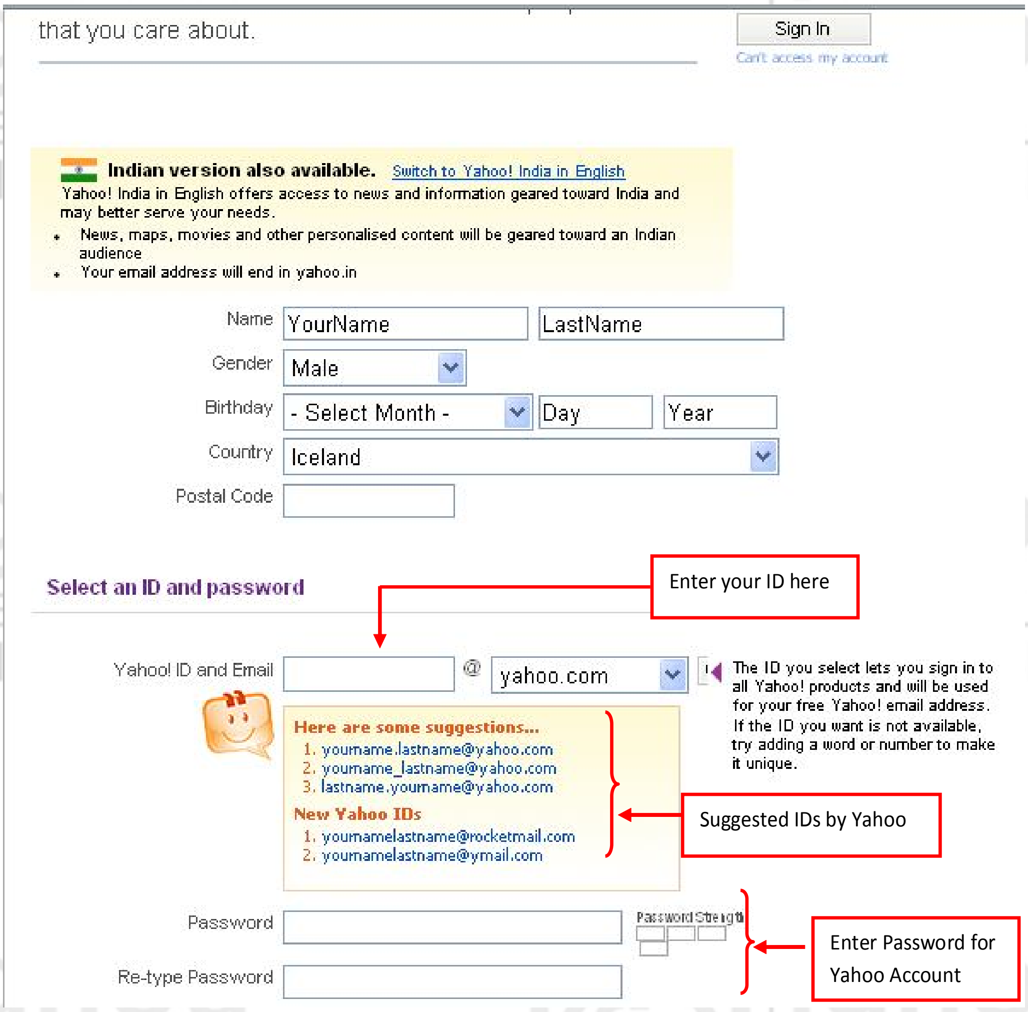


Fig. 13.4: Yahoo Mail registration Form (I)

4) Enter the verification code and click the button "Create My Account" (Figure 13.5).

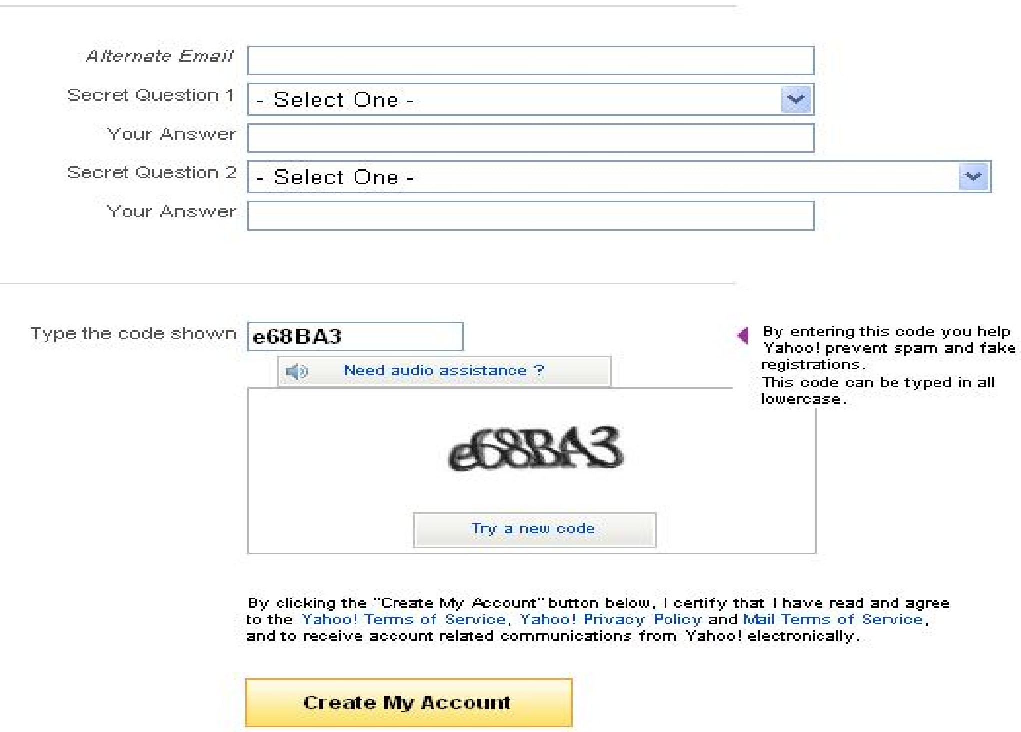


Fig. 13.5: Yahoo Mail registration Form (II)

5) This will open a new page confirming the creation of new yahoo account along with registration details Fig. 13.6. Click **Continue** to access the newly created account.

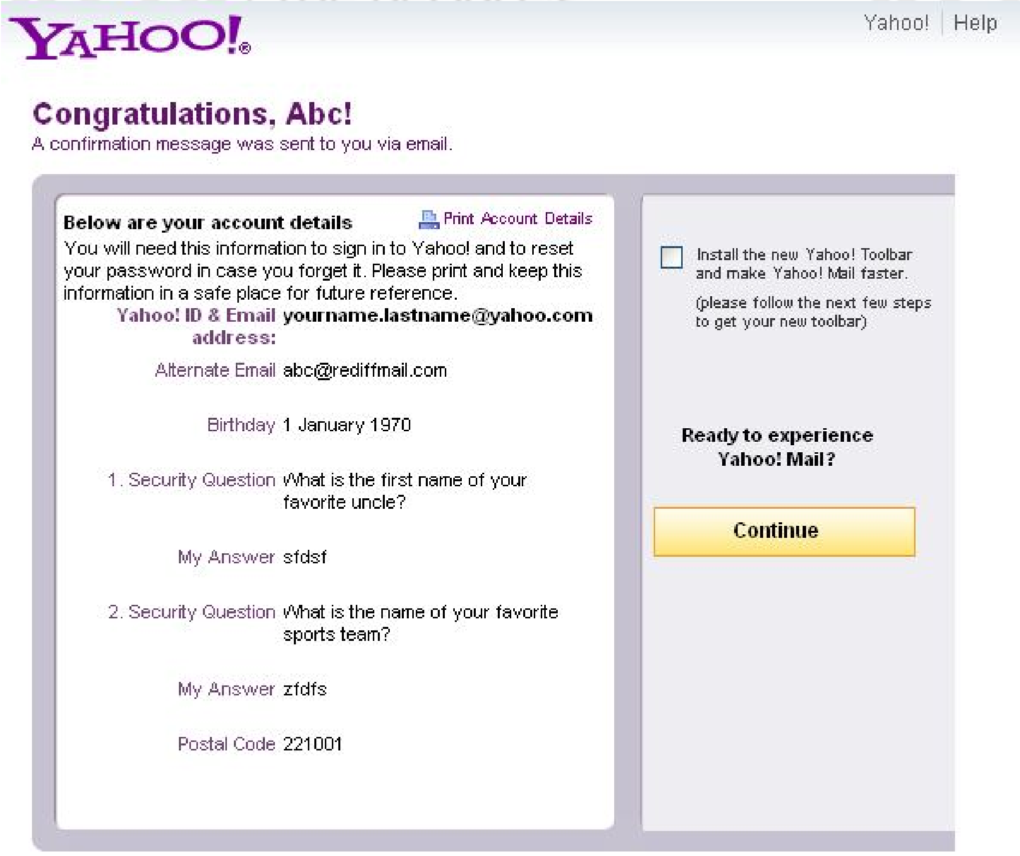


Fig. 13.6: Confirmation Message for new account

NOTE: To access the account for the next time, the Yahoo ID and password can be used on Sign-in page to access the mail account.

Most of the Web based Mail services offer four standard email folders: **Inbox**, **Sent**, **Drafts** and **Trash**, each of them are described below:

- **Inbox:**Lists all emails received from other email accounts. Highlights newly received/unread mails.
- **Sent:** A copies of messages sent are put into the **Sent** folder provided that the mail account is set to save all the sent messages.
- **Drafts:** A place for storing unfinished messages. If writing of a message is not yet finished and needs to be stopped in between, clicking the **Save** button puts the message into the **Drafts** folder, which can be accessed later on.
- **Trash:** stores email that is deleted from other folders. The messages are not truly deleted until they are deleted from this folder.

A message can be opened by clicking its **Subject**. Once the mail gets opened, at the top, options are available to delete, print or forward a message.

Note: Messages in Trash folder that are more than few days old will automatically be deleted.

13.6.5 Folder Management

Once a Yahoo account is established, one can create folders other than the already existing permanent folders (such as Inbox, Sent, Drafts, Trash and Spam). There is panel at the extreme right side of the panel called Folders (Figure 13.7).

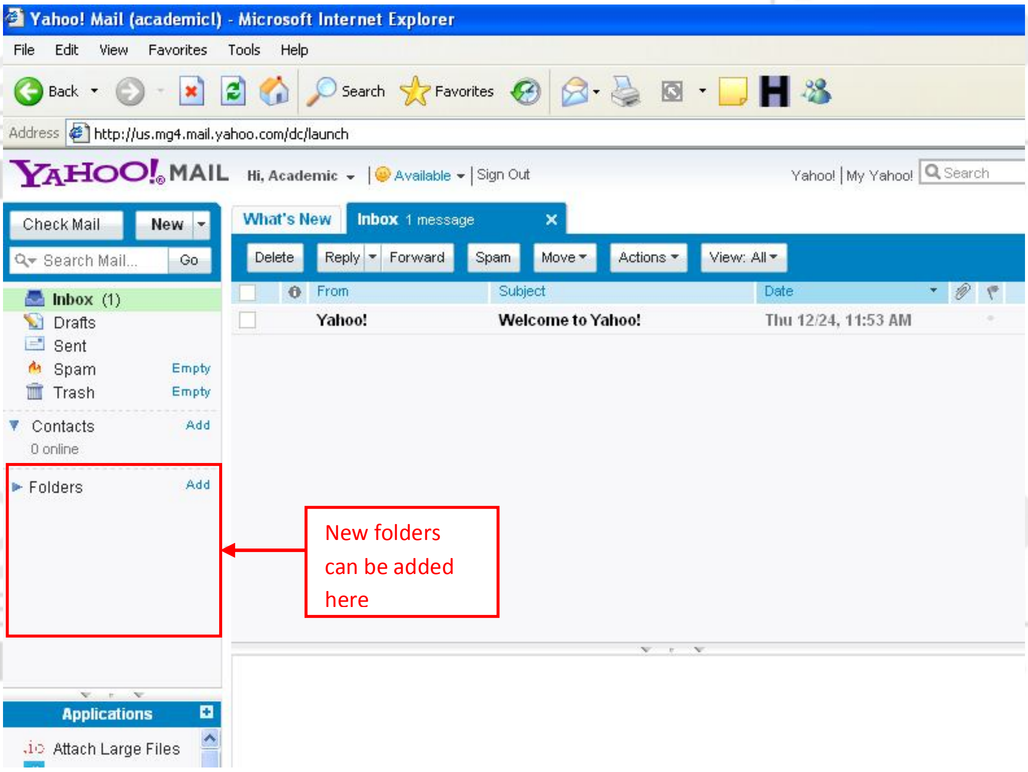


Fig. 13.7: The Folders Panel

To Create a New Folder

1) Click on **Add** (listed in front of the option **Folders**). This will create a new untitled folder below the panel Folders.

2) Type the name for the new folder e.g. FORUM (Fig. 13.8)

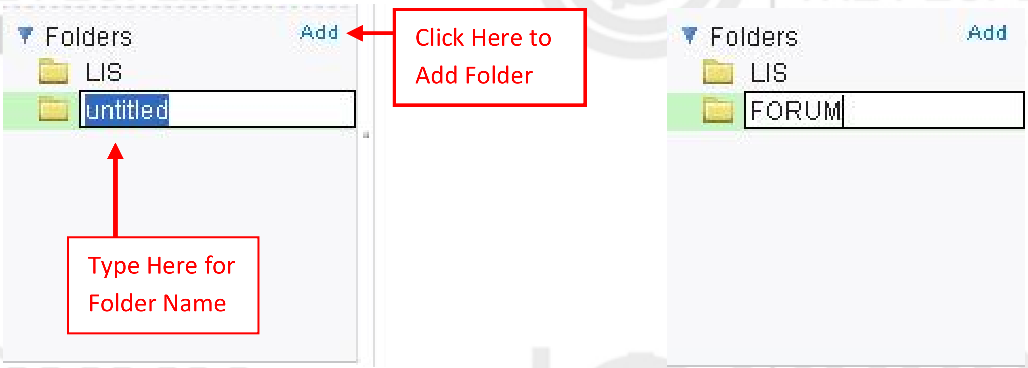


Fig. 13.8: To Create New Folder

To Rename a Folder

The steps are as follows (Fig. 13.9):

1) Select the folder to be renamed by clicking on it. Now, either press **F2** on keyboard or right click on mouse and select **Rename.**

2) Now, type the new name for the folder and press Enter from keyboard.

3) The folder will get assigned the new name.

Fig. 13.9: To Rename a Folder

To Delete/Remove a Folder:

1) Select the target folder (e.g. LIS)

2) Delete all mails from the target folder.

NOTE: A folder can be deleted only if it is empty i.e., no mails are present. Hence, before deleting a folder each mail should be deleted from the folder.

3) Now, Press **F2** on keyboard or right click on mouse and select **Remove**. This will prompt a message to confirm the action. Click OK. (Fig. 13.10)

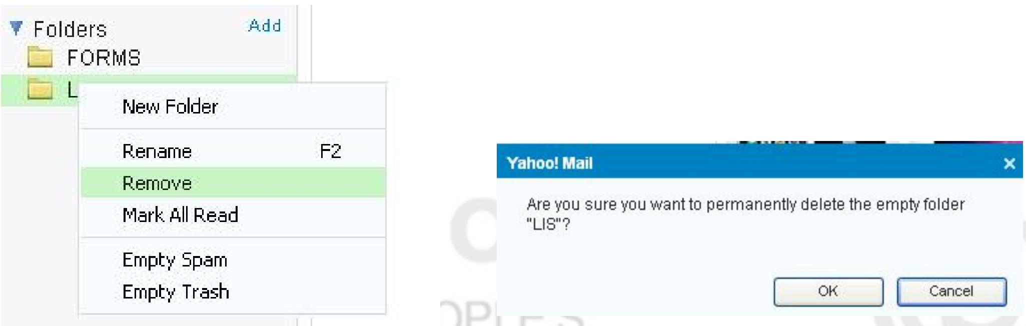


Fig. 13.10: Removing the LIS Folder

4) The folder is deleted successfully (Fig. 13.11).

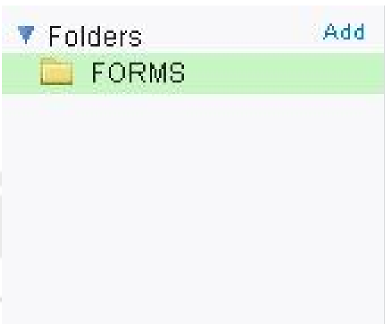


Fig. 13.11: LIS folder is deleted from the list

To Move an Email to a Folder

1) Go to the Inbox or Sent Folder.

2) Select the target email.

3) Click **Move** button of the mail task panel (Fig. 13.12). All the created folders (in this case FORMS and LIS) as well as Spam and Trash folders will be listed.

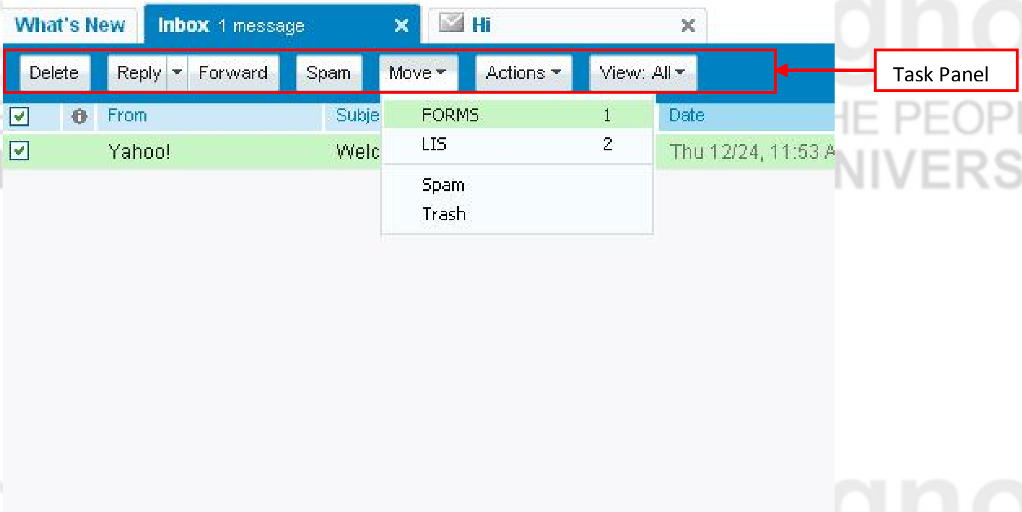


Fig. 13.12: The Task Panel of Inbox

4) Select the folder where the mail has to be moved (e.g. FORMS folder).

5) The mail is successfully moved to the target folder (Fig. 13.13)

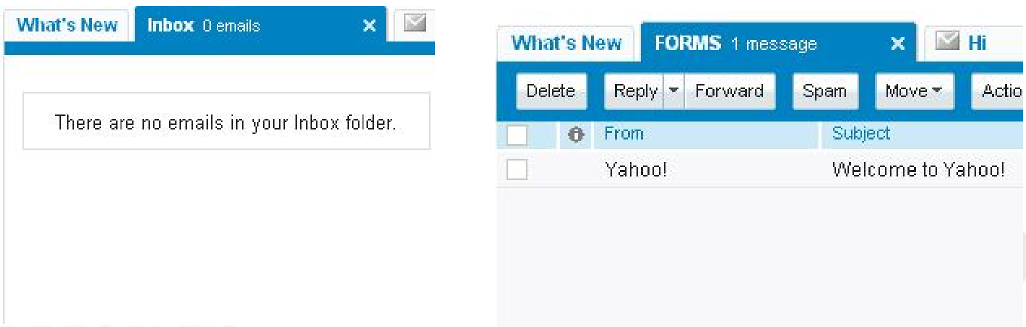


Fig. 13.13: The Empty Inbox and FORMS Folder with Message

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) What is Mail Transfer Agent?

....

....

....

....

13.7 MESSAGING

Messaging is a method of communication between two people or organisation. Messaging could be done using the power of Internet or through cell phones. Messaging software connects two people and facilitates them to communicate through text or voice or both of them.

Two types of Messaging

a) **Asynchronous Messaging**: The term "asynchronous messaging" means a method of communication between programs in which a program places a message on a message queue and leaves. It really does not bother how the message will be delivered. It is the delivery agent or the kind of infrastructure ensures delivery of message, even if the recipient is offline. For example, delivery of emails.

b) **Synchronous messaging**: In this kind of communication sender and receiver both have to be in connectivity while transferring the message. For example, telephonic conversation. Hence, a program places a message in a message queue and then waits for a reply to its message before resuming further.

13.7.1 Instant Messaging

Instant messaging (IM) is a form of real-time communication between two or more people based on typed text or using audio or video. The message is conveyed via devices connected over a network. Most IM programs provide these features:

- **Instant messages/Chat:** sending and receiving text/notes with an online friend.
- **Chat Rooms:** a common platform where two or more than two people can communicate.
- **Files/Web links/Videos/Images:** can be shared over network.
- **Talk:** Instead of a phone, Internet can be used to actually talk with friends. e.g Google+ Hangouts
- **Mobile capabilities:** Instant messages can be sent to mobile/cell phones.

13.7.2 Unified Messaging

Unified Messaging (or UM) is combination of different media into one channel. A user can access information into different media using a single device. Normally, Unified Messaging is common in mobile communication where, voice, text and fax can be accessed using one mailbox. It provides power to reach people almost anywhere, at any time and the flexibility to allow people to control when they can be reached.

13.8 ISSUES WITH MESSAGING

IM is increasing in popularity in both professional and personal applications. However, as with most of the things are Internet based, the increasing use of instant messaging has led to an associated increase in the number of security risks.

13.8.1 Spamming

Often in the email accounts one receives lot of unwanted mails which includes threats, promotional mails and so on. These messages are called Spam. There are two kinds of spams, mail lists and individual mails. Often users subscribe mail lists where lots of promotional materials are posted in bulk. These materials are of generic nature. There is another type of spamming where individual users are targeted for example, individual messages from unknown for illegal business or receipt of some unknown prize. Almost all the service providers have mechanism for filtering such messages based on the content or the subject line of the messages.

13.8.2 Privacy

Many Email service providers have their own privacy policies. Most of the free web based email services have financial dependency on advertisers and in return they share user information such as their email address information with advertisers. Also, details of the communication can be shared with government agencies on demand. Thus, one must be aware of the privacy policies of an email service provider before going for their services.

13.8.3 Security

Email /Instant messaging account are vulnerable to hijacking or spoofing. Spoofing is a phenomenon where an attacker hijacks another user's email/IM account and impersonates as user with others. Things apart the information exchanged over network can also be read over network by a third party. Hence, it is important to ensure information exchange.

There are several ways which have been devised to overcome all these problems which are enlisted below:

1) One of the best ways to secure the information being transmitted along an IM network is to encrypt it.

2) Keeping message logs for tracing any kind of mischief.

3) If file transfer via the instant messaging network is not required, then an instant messaging system that does not allow for files to be transferred should be utilised.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

5) What is a Spam?

....

....

....

....

13.9 WIDGETS AND UTILITIES

Widgets are small programming code that users can add to their webpage, personalised homepage, web browser, desktop, blog or social network. Mostly, these codes are embedded within an image file, hence, can be evoked through a mouse click or through a keyboard command by a computer or Internet user. A widget is used to enhance the look and feel of a website i.e., to make it look more fanciful. Widgets are considered as an offering by Web 2.0 to internet users.

There are two kinds of widgets

a) **Web Widget**

They are web-based, which means they can only be part of a website. It cannot be put on desktop unless customised to be used as desktop widget. The best example of a web widget is a YouTube video being displayed on a website other than YouTube. Similarly, *AddThis* (http://www.addthis.com) is a simple free web widget which can be used to bookmark a particular piece of information through various social networking websites such as Delicious (http://delicious.com/), Stumbleupon (http://www.stumbleupon.com), Google Bookmarks.

b) **Desktop Widget**

These widgets need not be part of a website and could be integrated on the desktop of a computer. These application accesses the Internet to provide the latest information. For example, a widget "Desktop Weather" displays the temperature and forecast for a given (configured) geographical area through the website **Weather.com.**

Fig. 13.14 : Web Widget

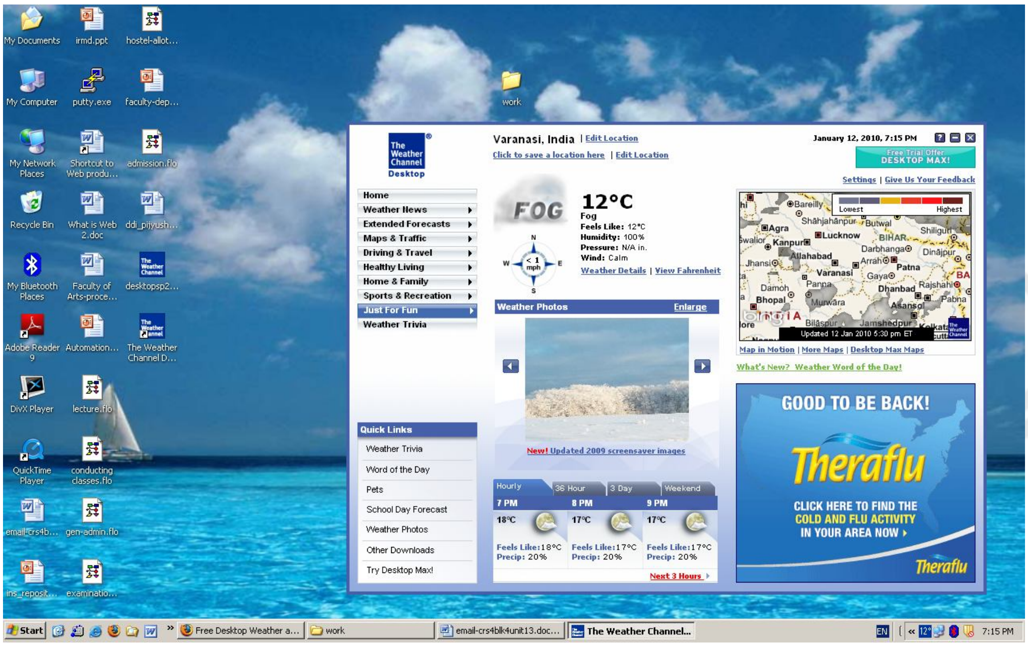


Fig. 13.15: Desktop Widget

Widgets can be delivered as:

- **Graphics:** Used to create banners, unusual fonts and animated images or text.
- **Gadgets:** hit counters, weather forecasts, calendars and maps
- **Entertainment:** Streaming of TV listings, daily quotes, interesting facts or games
- **Social:** provides link to social bookmarking or networking pages such as Facebook or Twitter e.g. Addthis.com
- **Audio Visual:** e.g. Streaming of YouTube videos directly to a web page

Utilities:

Utilities are useful for fixing minor problems or mis-configurations and handle day-to-day chores associated with computers. They are designed to help manage and tune the computer hardware, operating system or application software by performing a single task or a small range of tasks. Hence, the objective is to keep a system running at peak performance. It is also known as service *program*, service routine, tool, or *utility* routine.

Some of the examples of mail utilities are:

- **Cryptographic** utilities encrypt streams and files;
- **Mail Send Utility** to send multiple mails to a desired mail list from command line;
- **Email Finder Pro** n/a Fast and simple email address extraction utility;
- **JPEE Email Utility Lite 5.3.2** is known as a tool to merge and extract email, parsing data. JPEE is available as a free, easy to use, highly configurable, email communications software implementation;
- **Google Email Uploader** is a useful and smart desktop utility for Windows with the ability to upload email and contacts from other desktop email programs (like Microsoft Outlook) into your Google Apps mailbox. The Email Uploader preserves information such as sent dates and sender/recipient data, as well as the folder structure used by the other email program.

13.10 SUMMARY

Email is the most used service on Internet. It has become the fastest medium for transferring the message as well as data. The importance and authenticity of email is now being recognised officially and many organisation use email for day-to-day communication and do broadcast there circulars and orders by email. The service is offered by several service providers free of cost however, one can set his own mail server for communication. In this chapter we have learnt about the email systems and different types of email addresses. In due course we have also learnt about Instant Messaging services. There are some issues like individual privacy and security which has been discussed in detail in order to give insight of the problems.

13.11 ANSWERS TO SELF CHECK EXERCISES

1) Email is the fastest method of communication over Internet. Email stands for electronic mail. An email is delivered within no time irrespective to the distance. It carries text, audio, video and graphics as an attachment. Using email one can

transfer any piece of information or data to a number of people in a single stroke. It is one of the most secure medium for one to one as well as one to many communication.

2) Forward is a facility normally provided with all the email services. If one wants to change his email address to a new one and wants to receive mails from the previous accounts in the newer one he/she can put a forward on the previous accounts which will forward all emails received on the previous accounts to the new account.

3) An email has two parts:

   - Header
   - Body

   Header contains information regarding whom the mail is sent and email address of sender with brief subject about the content of message and the date mail is sent. The header also shows email addresses of all those who have received a copy of the mail.

   Body contains the message from the sender.

4) Mail Transfer Agent (MTA) is a piece of software, which transfers messages or mails from one host or machine to other. MTA forwards message through different machines to the destination. A mail could be transferred through several machines or MTAs and each machine stamps over email.

5) Spam is unwanted emails received in the inbox. Most of the such mails are propaganda materials. Apart from this these mails may contain harmful programs or virus.

13.12 KEYWORDS

Attachments : Files, which can be sent along with email.

Blacklist : This is a list of email addresses, which are blocked for sending mails. In other words, mail server refuses to accept mails from blacklisted email addresses.

Blind Carbon Copy (BCC) : Copy of an email message sent to recipients but his email address does not appear in the message.

Bounce : Returned messages that do not reach to destinations.

Carbon Copy (CC) : Carbon copy is a copy of an email sent to addresses other than the main recipient.

DNS : Stands for Domain Name Server, which translates domain names into IP addresses.

Download Message : Allows saving a copy of email message on local computer

Email harvesting : A program to scan Web pages for the purpose of collecting email addresses.

Email headers	:	Section of an email message contains the sender and recipient's email addresses and routing information.
Filters	:	Filters automatically filter incoming emails into designated folders. Filtering is done based on different fields of an email like, recipient address, subject, domain name and so on.
Forwards	:	Feature of an service which automatically forwards incoming emails to another email account.
MIME (Multi-Purpose Internet Mail Extensions)	:	An extension of the email standard that allows users to exchange multimedia files.
Plain text	:	Text in an email message that contains no formatting elements (i.e., color, bold, italics, etc.), pictures, or HTML.
Post Office Protocol (POP3)	:	A standard that enables emails to be retrieved from a remote mailbox. It allows to collect emails from an account to an email client program like, Outlook.
Proxy	:	A computer system or router that breaks the connection between sender and receiver, giving anonymity.
Reminders	:	Added feature of email systems that allows setting up reminders on a certain date.
Search Facility	:	A feature within the email system to search an email based on keywords/phrase as well as on certain criteria.
Signature	:	Section in the email message which goes by default with the sent email normally carrying name, address, website, address, phone number, etc. It appears at the end of the message.
SMTP	:	A protocol standard that enables emails to be transferred from sender machine to destination machine.
SMTP(Simple Mail Transfer Protocol)	:	Stands for Simple Mail Transfer Protocol, used to send email on the Internet.
Snail mail	:	Postal service delivers letters at home in paper format via postman.
Soft bounces	:	Bounce of a mail with a message due to temporary network congestion.
Spam	:	Unsolicited, unwelcome email sent in a large volumes.

Spell Check : Feature with email service to correct any spelling errors.

Spoofing : A method of spamming where the spammer uses false email address which does not exist.

Web mail : Web browser based email client like squirrel mail.

White list : A list authorised email addresses from which email messages can be delivered regardless of spam filters.

13.13 REFERENCES AND FURTHER READING

1. Email Protocols: IMAP, POP3, SMTP and HTTP. <http://www.emailaddressmanager.com/tips/protocol.html>
2. Using Folders. <http://www.earthlink.net/webmail/help/earthlink/en_US/using_folders.html>
3. What are utility programs. <http://depts.alverno.edu/cil/mod1/software/utility.html>
4. Widgets. <http://www.mediafuturist.com/files/widgets_trendmarker_feb08_forweb_200803251313542.pdf>
5. Electronic mail. http://tldp.org/LDP/nag/node186.html
6. Unified messaging. <http://www.iec.org/online/tutorials/unified_mess/topic02.html>
7. Email. (n.d.). Retrieved from Wikipedia. <http://en.wikipedia.org/wiki/Email>
8. Multipurpose Internet Mail Extensions. (n.d.). Retrieved from Wikipedia. <http://en.wikipedia.org/wiki/Multipurpose_Internet_Mail_Extensions>
9. Simple Mail Transfer Protocol. (n.d.). Retrieved from Wikipedia: <http://en.wikipedia.org/wiki/Simple_Mail_Transfer_Protocol>
10. Learn The Net: Email. (n.d.). Retrieved from Learn The Net: <http://www.learnthenet.com/english/section/email.html>
11. The history of email. (n.d.). Retrieved from Net History: <http://www.nethistory.info/History%20of%20the%20Internet/email.html>
12. Levine, J. R., Young, M. L. & Baroudi, C. *The Internet for Dummies*. 11th ed. Hoboken: Wiley Publishing, 2007. Print
13. Gralla, P. *How the Internet works*. 4th ed. Indianapolis: Que Publishing,1998. Print
14. Thomas, Brian J. *The Internet for scientists and engineers: online tools and resources*. 3rd ed. Oxford: Oxford University Press, 1997. Print
15. Lindsell-Roberts, S. *135 tips on email and instant messages: plus blogs, chatrooms, and texting.* Boston: Houghton Mifflin Company, 2008. Print
16. Wood, D. *Programming Internet Email*. Sebastopol: O'Reilly & Associates, 1999. Print

UNIT 14 WORLD WIDE WEB

Structure

14.0 OBJECTIVES

After reading this Unit, you will be able to:

- discuss the growth and development of WWW;
- explain the underlying technology behind WWW;
- describe Web 2.0 technology; and
- understand services of Web 2.0 technology and its impact.

14.1 INTRODUCTION

Internet has changed the life of people. While traveling in the train, airplane or any other mode of transportation one can keep oneself busy and connected to the world. It has influenced education, commerce, governance and entertainment, which have tremendous impact on day-to-day life of an individual. Internet was there even before twenty years but it is WWW (World Wide Web), which has brought radical change in the use of Internet. World Wide Web consists of interlinked hypertext documents. The transfer medium is known as Hypermedia, which carries the multimedia objects such as images, as well as audio and video files over network in addition to text. There is no doubt about the fact that the actual influence of Internet on society could be seen only after the introduction of WWW. It was a revolutionary break-through in the emerging technological environment.

14.2 WORLD WIDE WEB

The credit of developing World Wide Web (WWW) goes to Sir Tim Berners-Lee in 1989, and later on Robert Cailliau in 1990 at CERN laboratory in Switzerland where it was demonstrated over distributed hypermedia servers (ý1). The servers store hypertext documents, which can be accessed via a client (i.e., a web browser). The language used for creation of document is known as HyperText MarkUp Language (HTML). These web documents or hypertext documents are linked to each other using a specific pointer system known as Uniform Resource Locator (URL). These pointers are like *handles* and have the capability to fetch the document stored and scattered across various web servers. The servers where these hypertext objects are stored host the hypermedia documents, honouring the request to serve the documents to a client. Internet is collection of such servers also known as Web servers placed in different parts of world. Hence, WWW is a service over Internet. It is a collection of documents available over host servers worldwide.

A special kind of application software is used to access hypertext documents know as Web Browser. There are many web browsers available like, Mozilla Firefox, Internet Explorer, Opera, and so on. WWW supports hosting, dissemination and playing of multimedia documents that includes audio video.

In 1989, Sir Tim Berners Lee wrote a proposal "Information Management: A Proposal" for conceptualising WWW at CERN Lab. In 1991 the first web browser was released. The web browser used to work on hypertext with Graphical User Interface on platform called as NeXTStep. The web browser was a WYSIWYG (What you see is what you get) type with facility to support hyperlinks (Fig. 14.1).

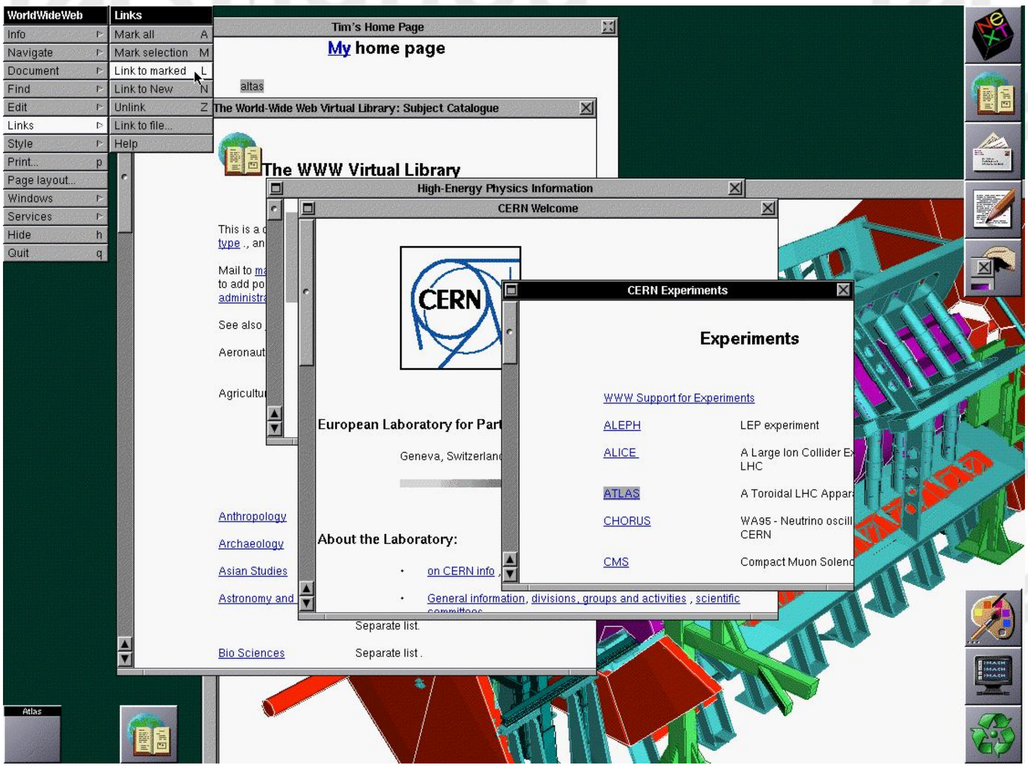


Fig. 14.1: Early Web Browser

14.3 CONCEPTUAL FRAMEWORK OF WWW

Web is one of the most widely used components of Internet. The Web allows access to information dispersed all across the world on different servers. The information is available in diversified form such as text, graphics, animation, photos, audio and video. The Web physically consists of:

a) a personal computer or mobile device;

b) web browser software;

c) connection to an Internet service provider;

d) computers called servers to host digital data and routers and switches that direct the flow of information.

14.3.1 Communication Architecture

The basic web structure of WWW is two tiered popularly known as Client Server Model (Fig. 14.2). Those machines that provide services (like Web servers or FTP servers) to other machines are **servers**; and the machines utilising those services are known as **clients**. This kind of architecture depends upon the following:

a) **Standard Representation of Information on Web:** Markup Languages such as HTML, XML are some of the standards available for the content representation over the Web.

b) **Transfer Protocols:** These are the different protocols for transferring information between computers on the Internet. HyperText Transfer Protocol is the underlying protocol used by web. HTTP is a synchronous request-reply protocol that requires direct, online connections. HTTP defines how messages are formatted and transmitted, and what actions Web servers and browsers should take in response to various commands. For example, when you enter a URL in your browser, this actually sends an HTTP command to the Web server directing it to fetch and transmit the requested Web page. HTTP is called a *stateless* protocol because each command is executed independently, without any knowledge of the commands that came before it.

c) **Addressing Protocols:** Web utilises protocols to identify a web object based on names and addresses. These are known as Uniform Resource Identifiers (URIs). Uniform Resource Locator (URL) is an example of URI. It is the global address of web documents and objects. The first part of the address is called as protocol identifier and it indicates what protocol to use, the second part is called as resource name and it specifies the IP address or the domain name where the resource is located. The protocol identifier and the resource name are separated by a colon and two forward slashes. An example of addressing and addressing protocol is as follows:

http://www.bbc.co.uk/news/world-latin-america-11617094

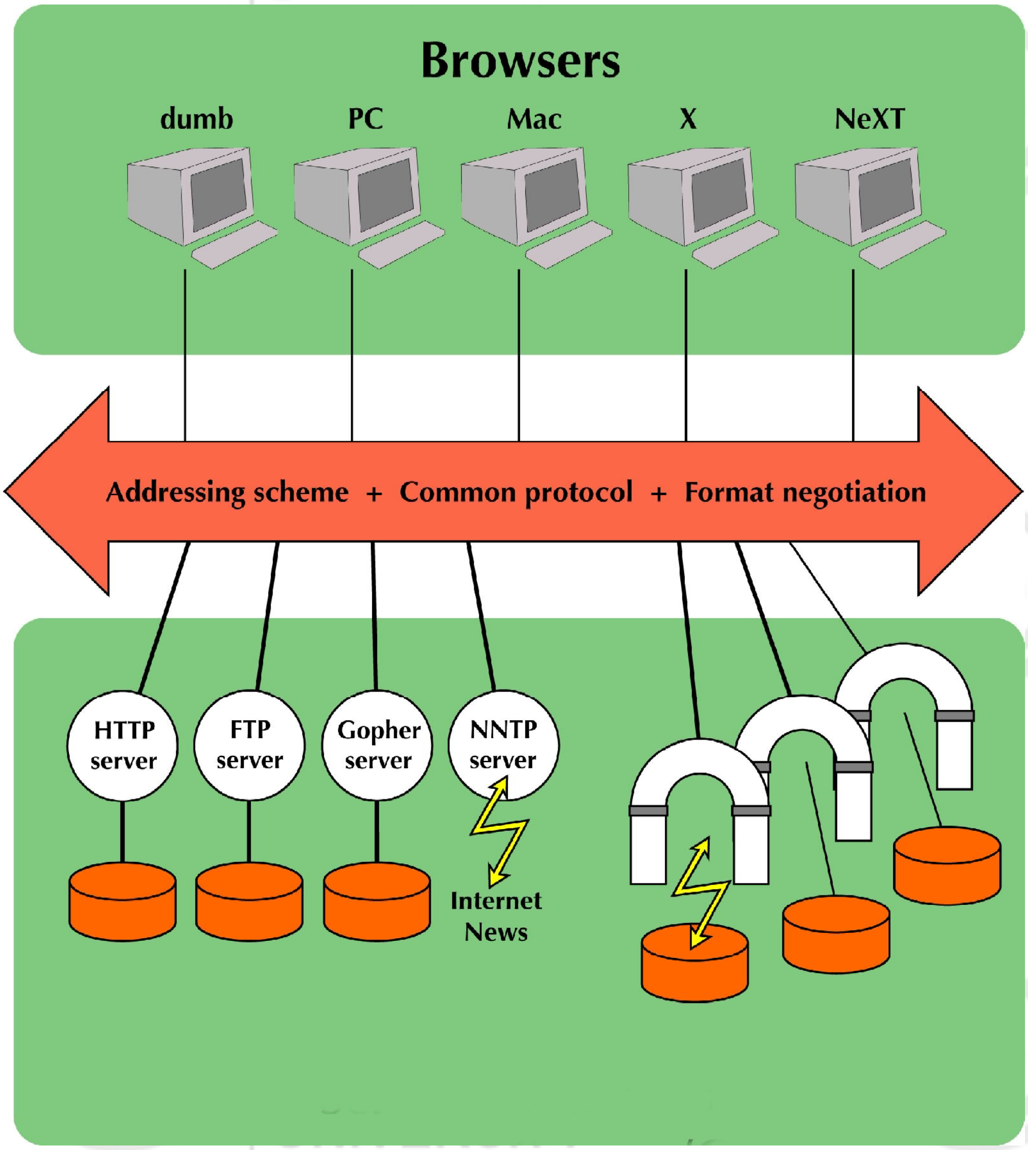


Fig. 14.2: Early Architecture of WWW given by Tim Berners Lee and Robert Cailliau

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) What is a Web Browser?

2) What do you mean by addressing protocol?

....

....

....

....

14.3.2 Protocols

Tim Berners-Lee implemented the HTTP protocol in 1991 at CERN, the European Center for High-Energy Physics in Geneva, Switzerland. HTTP stands at the very core of the World Wide Web. According to the HTTP 1.0 specification (26):

"The Hypertext Transfer Protocol (HTTP) is an application-level protocol with the lightness and speed necessary for distributed, collaborative, hypermedia information systems. It is a generic, stateless, object-oriented protocol, which can be used for many tasks, such as name servers and distributed object management systems, through extension of its request methods (commands). A feature of HTTP is the typing and negotiation of data representation, allowing systems to be built independently of the data being transferred".

- HTTP provides a comprehensive addressing scheme with the concept of URI as a location URL. HTTP based hyperlink address is rendered in following format:

  http://host:port-number/path/file.html

  http://drtc.isibang.ac.in:80/

- HTTP protocol is based on a request from a client and response from a server. A server hosts web pages and on request, delivers these web pages to client. The default port from any HTTP request is 80, however one can change the port number to any other port number as desired. However, one port number can be used only to run one program. This kind of request/response format of communication is known as Client – Server Architecture. Every request is treated as a new request and there is no dedicated connection between client and server. This is known as stateless connection.

- HTTP supports multimedia objects also known as Internet Media Types (MIME format). The header of information informs the client what kind of data will follow.

14.4 MARKUP LANGUAGES

It is an encoding system to annotate the text of a web document. The system consists of notations known as tags. They control the structure, formatting as well as the relationship between different parts of a document. A marked-up document thus contains two types of text: text to be displayed and markup language on how to display it.

Example of markup language is HyperText Markup Language (HTML), one of the document formatting languages of the World Wide Web.

A marked-up document will typically look like this:

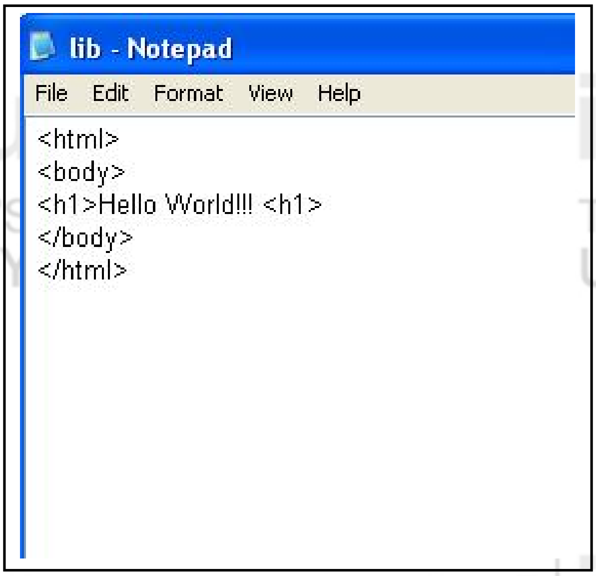


Fig. 14.3: lib.html (view through the notepad application)

The output of this document will be:

Fig.14. 4: lib.html (output view through Mozilla Firefox)

14.4.1 Definition and Need

A markup language is a methodology to annotate text for its representation and processing over World Wide Web. IBM researcher Charles Goldfarb is considered as "father" of markup languages. A markup language uses codes, which are also known as tags to describe the layout and formatting of the document. Also, it can also describe the types of information conveyed by a given text.

For example:

<html>

<Address> Flat No. 20, Rajajipuram, Lucknow </Address>

</html>

Thus, the piece of content included within <Address> tag is an address or a location. Over web, search engines processes these codes or tags and interpret the information accordingly. Thus, markup languages are important to define the meaning and context of the text on web. It also facilitates standard representation of text by different browsers.

14.4.2 Types of Markup Languages

Standardised Generalised Markup Language (SGML)

The markup languages that carry the instruction for text processing are known as 'Procedural Markup'. The idea of markup was to format a particular kind of document. But later on it was felt that markup languages could be used for system-to-system information interchange also. This was first realised by Charles Goldfarb, Ed Mosher and Ray Lorie when they were working with legal documents. They designed the first markup language known as GML (Generalised Markup Language) based on the following observation:

- The document processing programs needed to support a common document format.
- The common format needed to be specific to their domain - for example legal documents.

- To achieve a high degree of reliability, the document format would have to follow specific rules.

For example, take an example of a memorandum:

To: Bishwanath Dutta

CC: Bibhuti Bhushan Sahoo

From: Aditya Tripathi

Date: 27.01.2003

Subject: Appointment order

We are extremely happy to inform you that you are selected as the Coordinator of Knowledge Management Team.

There are six fields in this document:

- Who the document is intended for (the **To**: field).
- Who has been sent a copy of document (the **CC**: field).
- Who sent the document (the **From**: field).
- The date of document written (the **Date:** field).
- The subject of document (the **Subject:** field).
- The document body.

The structure of this document is fixed and one is bound to write it in the same structure. Hence, porting the information across systems will not be a problem as the structure of document is always same. The definition of the structure of document is known as 'DTD (Document Type Definition)'.

Once GML was designed, Goldfarb fine-tuned his work and proposed the SGML (Standardised Generalised Markup Language) which was further approved by ISO (International Organisation for Standardisation) in 1986. Hence, SGML is for defining the format in a text document. Readable by both humans and computer programs, SGML is usable in a wide range of applications such as print publishing, CD-ROMs, and database systems. SGML was not a language itself but it was a meta language to develop other markup languages. HTML (Hypertext Markup Language) is a derivative of SGML. HTML is more like a formatting language. Thus it is difficult to pull out what kind of data is stored inside a HTML document. Once this difficulty was understood, the need for domain specific tags was felt, for information interchange, which is not possible with HTML. Hence the XML was developed. It is always said that XML is more near to SGML when compared to HTML.

HyperText Markup Language (HTML)

HTML stands for HyperText Markup Language. It is a language, which is used to develop web pages. It is a collection of several tags to describe visuals of a webpage. The goal of HTML is to provide a display format to the given set of data so that it can be read on a web browser. HTML was originally designed by Sir Tim Berners-Lee in 1991 at CERN Lab. HTML is an offshoot of SGML.

HTML is the building block of a website. It allows multimedia objects to be embedded in the webpage including audio, video, text and graphics. The language consists of tags. A tag is an element (known as HTML element), which has certain properties. These properties are applied on the data embedded in between tags. It is an individual component of an HTML document. Hence, HTML documents are collection of tags. These tags may simply contain data or can co-exist with other tags establishing parent-child relation. A tag has certain attributes, which are applied, on the contained data or on the child tags (or elements).

The number of tags used in HTML is fixed hence the language uses closed vocabulary. The structure of a web page is as follows:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<HTML>
  <HEAD>
    <TITLE>My first HTML document</TITLE>
  </HEAD>
  <BODY>
     <P>Hello world!
  </BODY>
</HTML>
```

Structure of HTML Document:

An HTML document has two parts,

- Head, and
- Body

'HEAD', contains elements (tags) for TITLE of the document. The 'TITLE' element stores information about the title of the document.

<TITLE>Website of Indira Gandhi National Open University</TITLE>

There is another element used in 'HEAD' section i.e., 'META' element. 'META' element stores information about the document such as author, copyright, location, relation, keywords and so on.

<META name="Author" content="Sneha Tripathi">

In the META tag first attribute or the property is defined under the NAME attribute and its value is given under CONTENT. In the above example Author is an attribute and value of author is 'Sneha Tripathi'.

These attributes are the attributes of the document, which is being described. Sometimes an attribute may use closed vocabulary or a scheme. In such cases META element also specifies the SCHEME used.

<META scheme="ISBN" name="identifier" content="0-8230-2355-9">

The second part of an HTML document is BODY element. The body of a document contains the document's content. The BODY element contains all the tags or elements, which are used to display the data over a web browser. It includes a variety of tags such as

<H1>…</H1>

<H2>…</H2>

<H3>…</H3>

<H4>…</H5>

<TABLE>…</TABLE>

<P></P>

<B></B>

<I></I>

Each element inside BODY tag can have various attributes, which is defined in HTML Standard Specification. The current version of HTML specification is 4.01. The next version of HTML, which is due, is HTML 5.0.

Extensible Markup Language (XML)

According to the abstract from the XML Specification version 1:

"The extensible Markup Language (XML) is a subset of SGML that is completely described in this document. Its goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML. XML has been designed for ease of implementation and for interoperability with both SGML and HTML."

- XML stands for eXtensible Markup Language.
- XML is a markup language much like HTML.
- XML was designed to describe data.
- XML tags are not predefined in XML unlike HTML where the tags are pre-defined.
- XML uses a DTD (Document Type Definition) to describe the data.
- XML with a DTD is designed to be self-descriptive.

Following are the goals kept in mind while developing the specification for XML:

i) XML shall be straightforwardly usable over the Internet.

ii) XML shall be compatible with SGML.

iii) It shall be easy to write programs to process XML files.

iv) The processors could read the XML document easily.

v) XML document should be human-legible and reasonably clear.

vi) The XML design should be prepared quickly.

vii) The design of XML should be formal and concise.

viii) XML document shall be easy to create.

ix) Terseness in XML is of minimum importance.

XML is different from HTML in the following ways:

i) XML was designed to carry data.

ii) XML is not a replacement for HTML.

iii) XML was designed to describe data and to focus on what data is.

iv) HTML was designed to display data and to focus on how data looks.

v) HTML is about displaying information. XML is about describing information.

What can be done with XML?

i) XML does not DO Anything

XML was not designed to DO anything. Maybe it is a little hard to understand, XML was not developed to DO anything. XML is created as a way to structure, store and send information.

```
<?xml version="1.0" encoding="UTF-8" ?>
<book>
   <title>Application of expert systems in libraries and information centres</
title>
    <author>
     <f_name>Anne</f_name>
     <l_name>Morris</l_name>
    </author>
   <edition>1st Edition</edition>
   <place>London</place>
   <publisher>Bowker-Saur</publisher>
   <physical_desc>241 p.</physical_desc>
</book>
```

Fig. 14.5: An XML Document in a Web Browser

The example shows the structure of a document, which describes a book, titled 'Application of expert systems in libraries and information centres'. The book has a title, author, edition, publication related information, etc. Author Name is further divided into first name (f_name) and last name (l_name). Inside these tags the actual data is stored. If one browses the document in the browser, data will appear embedded in the tag without having any kind of formatting.

ii) Define Your Own Tags

XML provides the facility to create domain specific tag set which facilitates the information interchange within a specific domain. For example, NewsML is developed for information interchange among the news agencies like Reuter and others.

iii) XML is Not a Replacement for HTML

It is important to understand that XML is not a replacement for HTML. It is useful for describing the data.

iv) XML can be used to Exchange Data

With XML, data can be exchanged between incompatible systems. In the real world, computer systems and databases contain data in incompatible formats. One of the most time –consuming challenges for developers has been to exchange data between such systems over the Internet. Converting the data to XML can greatly reduce this complexity and create data that can be read by many different types of applications.

v) XML can be Used to Share Data

With XML, plain text files can be used to share data. Since XML data is stored in plain text format, hence it provides a software as well as hardware-independent way of sharing data.

This makes it much easier to create data that different applications can work with. It also makes it easier to expand or upgrade a system to new operating systems, servers, applications, and new browsers.

vi) XML can Make Data More Useful

With XML, data is available to more users. Since XML is independent of hardware, software and application, one can make their data available to more than only standard HTML browsers.

Other clients and applications can access XML files as data sources, like they are accessing databases. The data can be made available to all kinds of "reading machines" (agents). For example, a data set can be used to see a webpage in a web browser of a computer or it can be used see the display in mobile phone.

vii) XML is used to Create New Languages

XML is the mother of WAP (Wireless Application Protocol) and WML (Wireless Markup Language). The Wireless Markup Language (WML), used to markup Internet applications for handheld devices like mobile phones, is written in XML.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) Discuss HTTP protocol for information communication over WWW.

....

....

....

....

14.5 WEB 2.0

14.5.1 Definition and Need

The term Web 2.0 is given by Dale Dougherty, Head of Maker Media division of O'Reilly in 2003. Since then it has become a popular concept. Web 2.0 refers to the second generation of the Web, which enables people with no specialised technical knowledge to create their own websites, to self-publish, create and upload audio and video files, share photos and information and complete a variety of other tasks.

While there is no set definition of Web 2.0, it generally refers to the use of the web more as a social platform where users can participate by generating their own content alongside the content provided by the websites.

"Web 1.0 was all about connecting people. It was an interactive space, and I thinkWeb 2.0 is, of course, a piece of jargon, nobody even knows what it means" (12). The above two definitions and excerpt from the talk with Tim Berners Lee show that there is no hard and fast definition of Web 2.0 however it is agreed that it is the second generation of Web where users can participate in generation of web content with much of the knowledge technology. The spectrum of web based content and services cover online shopping, email, chatting, discussion forums, blogs, wikis, social networks, YouTube, Twitter and so on. The products are more personalised now compared to more generic as it used to be.

Another aspect of Web 2.0 is the generation of product and services which are specialised and more users' centric. The trend is to develop tailor made products on demand like online e-learning modules; information digests systems, literature review etc. that are more user oriented services. The Web 2.0 has empowered users to be more interactive with the existing services over World Wide Web.

14.5.2 Need and Features

Following points are instrumental towards the development of Web 2.0 technologies:

- Users' Participation;
- User centric Services;
- Decentralisation and Interoperability;
- Hiding Technological Complexity;
- Modularity.

Users' Participation

With legacy web, the communication was more from source to users where as reverse was not possible. Web 2.0 provides means and tools which empowers users not just to communicate back to the source but also generate content for the website. A user can express their feelings and view about the content. They can agree or disagree with the source. This participation leads in collaboration and development of innovative thoughts.

User Centric Services

The trend of services and its delivery is transformed in Web 2.0 environment. Earlier it was technology, which used to ride the market. But with the advent of new technologies, web services are more customer/user oriented. These services require enough flexibility in their modus-operandi to meet the ever changing needs of users. The user may keep changing their priorities and services options and the services have to stand themselves to each call of their users.

Federated and Interoperability

The applications in Web 2.0 environment are distributed over different nodes. And each node is responsible for its own services. However, collaboratively all the nodes can generate one service through a single platform. The service nodes generate services in a standard format, which can be amalgamated to a single service. This makes systems interoperable over a federated environment.

Hiding Technological Complexity

Web 2.0 platform aid their users to create services or contents without knowing much of the technology. Thus, it is more user-centric. Web 2.0, in principle, is really not about the technological complexity though technology plays a major part from backend. The technology is developed in such a way that users need not bother about the hassles of technology rather they should concentrate on content of the services. The technology is kept hidden from the user.

Modularity

Modules are components of any system. A system functions in coordination of different modules to offer a service or product. Modular approach towards a system provides flexibility for adding or removing any feature out of the system. Legacy Web was about providing information in a robust but inflexible way. In Web 2.0, the modularity provides facility to add or remove components offering flexibility to a great extent.

Features of Web 2.0 Applications

Tim O'Reilly suggests that the true test of a Web 2.0 service relies on amalgamation of some or all of the following features: (18)

- services, not packaged software, with cost-effective scalability;
- control over unique, hard-to-recreate data sources that get richer as more people use them;
- trusting users as co-developers;
- harnessing collective intelligence;
- leveraging the long tail through customer self-service;
- software above the level of a single device; and
- Light-weight user interfaces, development models, and business models.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

4) Write features of Web 2.0 applications.

....

....

....

....

14.5.3 Web 2.0 Applications

Collaborative Web

One of the important features of Web 2.0 is Collaborative services. In a collaborative approach content is generated by a number of persons working at different locations. Online Wikipedia is a good example of collaborative content development.

- **Wikipedia**

Wikipedia (23) is a community-based encyclopedia where anonymous volunteers from Internet contribute articles. It is free to use and one can write as well as change the

Wikipedia articles. The project was started in 2001 and has become one of the most consulted website over Internet.

It is a live service, which is being continuously updated. Newer topics are being added every moment. However the control over the content is kept by editing process to stop any kind of vandalism and misinformation.

The topics in the Wikipedia is arranged in classified manner with following broad subjects:

- Humanities
- Social Sciences
- Natural Sciences
- Formal Sciences
- Professions and Applied Sciences

Each subject is further divided into narrower subjects. Articles are the leaf nodes, which are interlinked as and when, referred within the text.

Each article starts with a small introduction followed by a content of topics covered in the article. At the end, list of references, suggested readings (See also), External links (links outside Wikipedia website) and Further Readings are given to facilitate the reading.

Example: http://en.wikipedia.org (Fig. 14.6)

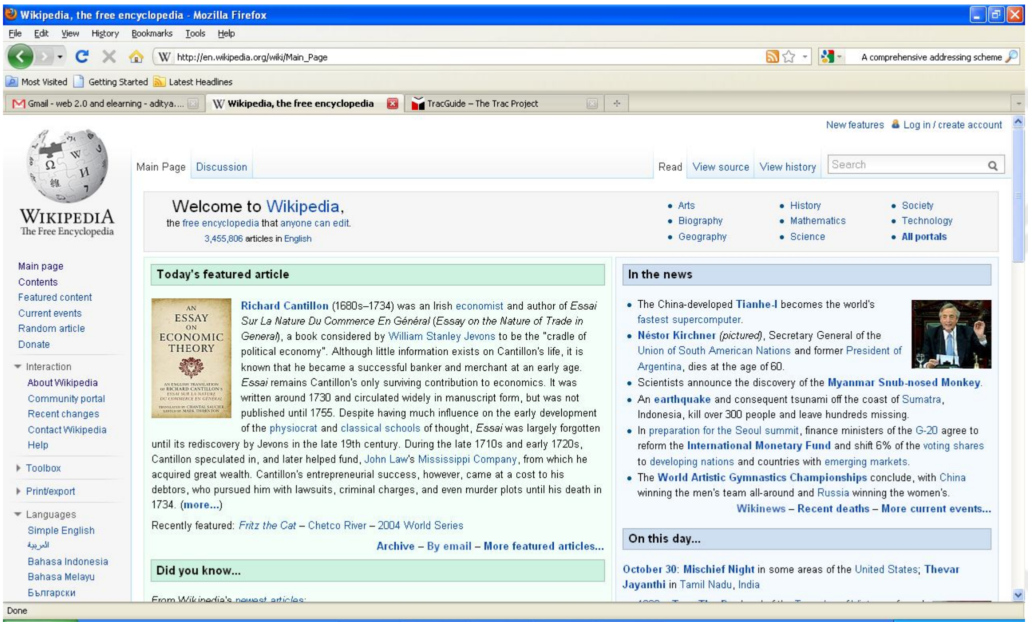


Fig. 14.6: Wikipedia

- **Blogging**

Phenomenon of blogging has picked up immensely among new as well as older generation. According to a statistics, a new blog is created on average every second; 54,000 posts are created every hour, translating into roughly 1.3 million new posts per day. The area or the usage where blog is created is known as blogosphere. A blog promotes freedom of speech, interaction with audiences, and a tool for aggressive marketing.

Blogger offers service to host individual blogs online. Initially, it was started as a small company but now owned by GOOGLE. It facilitates people to express their thoughts online.

Another important blogging website is Wordpress.com (Fig. 14.7) which allows hosting of individual blogs and maintains it.

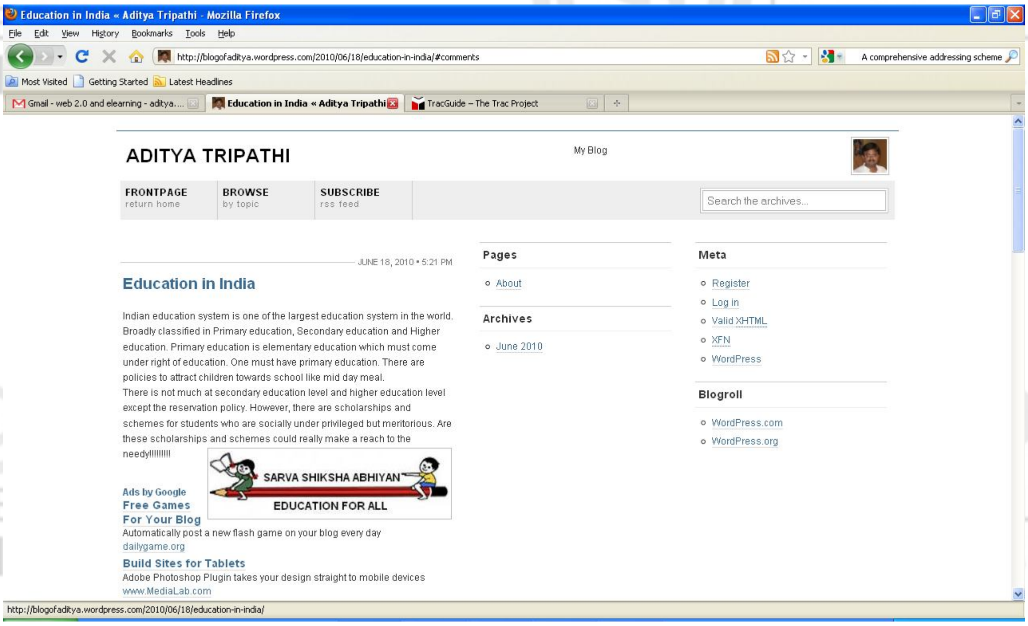


Fig. 14.7: An Example of a Blog at Wordpress.com

These blogging websites provide an administrative panel to administer one's own blogging site. Normally, two kinds of subscription are offered by these blogging sites, free and fee based subscription. The later one is used by commercial organisations, celebrities and so on whereas free ones are available for all. The topics in a blog can be classified and comments can be moderated by the blog administrator.

Example of Blogs: http://www.blogger.com/
http://wordpress.com/

- **Project Management System**

Managing a project requires collaborating different activities together with definite time schedules. Particularly, during the project different versions of products are developed which has to be tracked so that they do not create confusion.

A project management system is ideal for managing project of software development. However, it can manage projects apart from software development. These systems support charting down the work-plan and share it with all the members. These systems use version control for a particular set of file or software and highlight the differences among the different version. These systems have strong user control mechanism, which does not allow a non-authenticated user to look into the areas, which he/she is not concerned with. These systems highlight the landmarks and project the time frame for the stakeholder participating in the project.

Trac is an open source lightweight project management tool, which can be implemented as a web based API. Trac supports to monitor and resolve individual bugs, issues, feature requests, and ideas. It has a ticketing system, which is numbering system for any of the issues mentioned during the project. Trac has supervision module with built-in documentation server, which can be used to keep of documentation in a form of a Wiki.

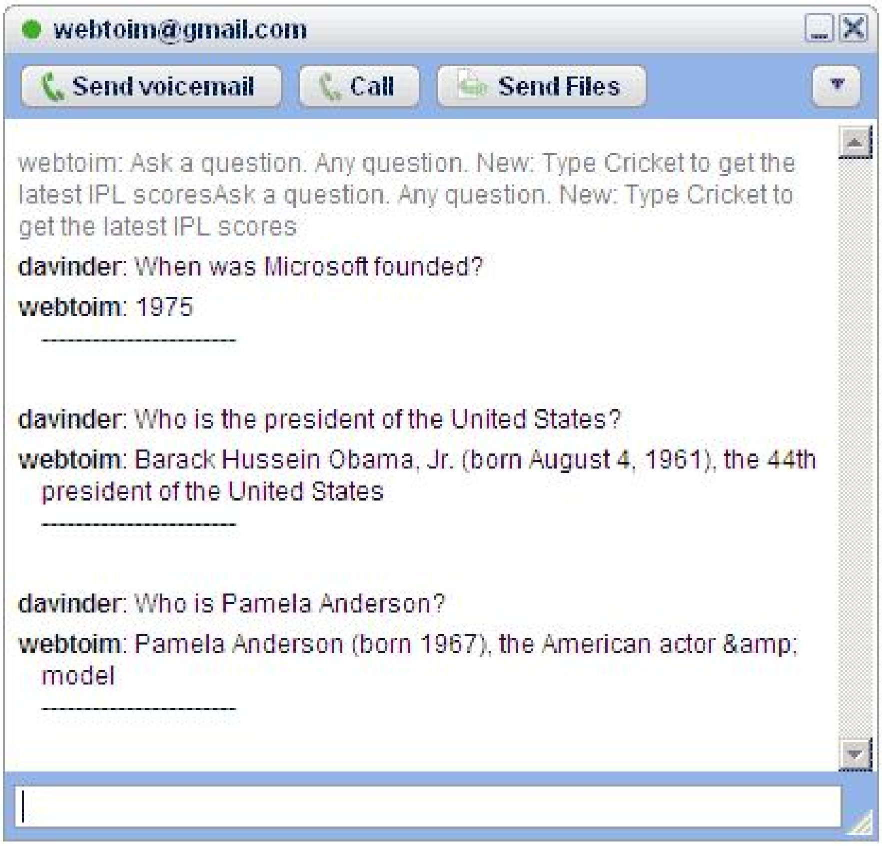
Bugzilla is another Defect tracking system, used to track the bugs of software. It has a feature to notify through email about the changes in the versions of software system.

Interactive Web

Interactive tools of Web 2.0 are used for collaboration of users to share their ideas and work. It covers all the tools, which are used for sharing works and conducting online discussions and meetings.

- **Online Chat**

There are many free applications, which can be used for online chatting using text, audio and video. Important chatting applications are, Skype, GTalk (or GoogleTalk), Yahoo Chat. These applications are free to use over computer-to-computer chatting. One can hold conferencing as well as one to one chat. These applications can be also used to transfer the files over Internet.



- **Document Sharing Tool**

These are tools which can be used to share documents, images and audio visuals over Internet.

Document Sharing: GoogleDoc

It is free tool for sharing documents over Internet. It supports online upload, editing and sharing of documents, spreadsheet, presentation, drawing and forms. It supports the following functionalities:

- Real-time collaborative highlighting in documents
- In-cell dropdown and cell validation in spread-sheets
- Shows all formula
- Spell checker
- Page sizes

- o Auto-linking the text in the documents
- o A new curve tool in drawings
- o Convert files in the document list
- o In built dictionary – one can add terms in the dictionary also
- o Create Tables and draggable rows and columns to resize
- o Document translation
- o Searching of text in the document
- o Mobile support
- o Format painter in spreadsheets
- o In built Optical Character Recognition (OCR) for PDF (Portable Document Format)
- o A keyboard shortcut pop-up and more in drawings
- o Including drawings in the text
- o Sorting of text
- o Create forms

GoolgleDocs (Fig. 14.8) has almost all the functionalities to support document texting and sharing with a group of people.

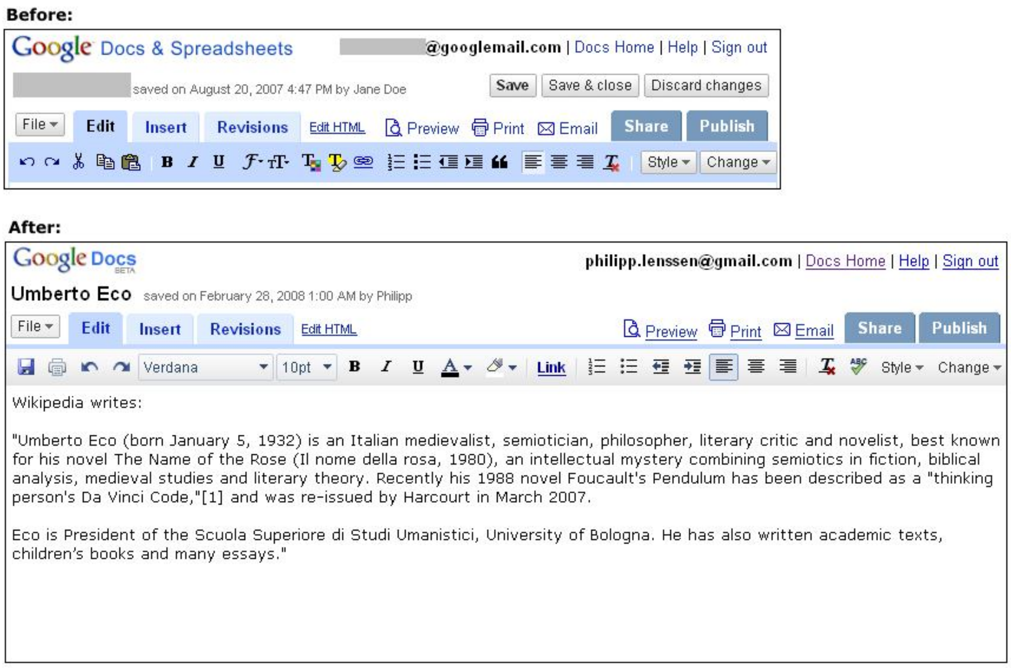


Fig. 14.8: GoogleDocs

Image Sharing: Picasa

Image sharing is an important feature of Web 2.0. People can share their albums online, which can be viewed online by others. One can define the level of access to their photographs. Local album on computer can be synchronised with online album using a client software and any addition in the local album will be uploaded to the website. Picasa (Fig. 14.9) is an online image sharing tool by Google. Another example of image sharing is Flicker.

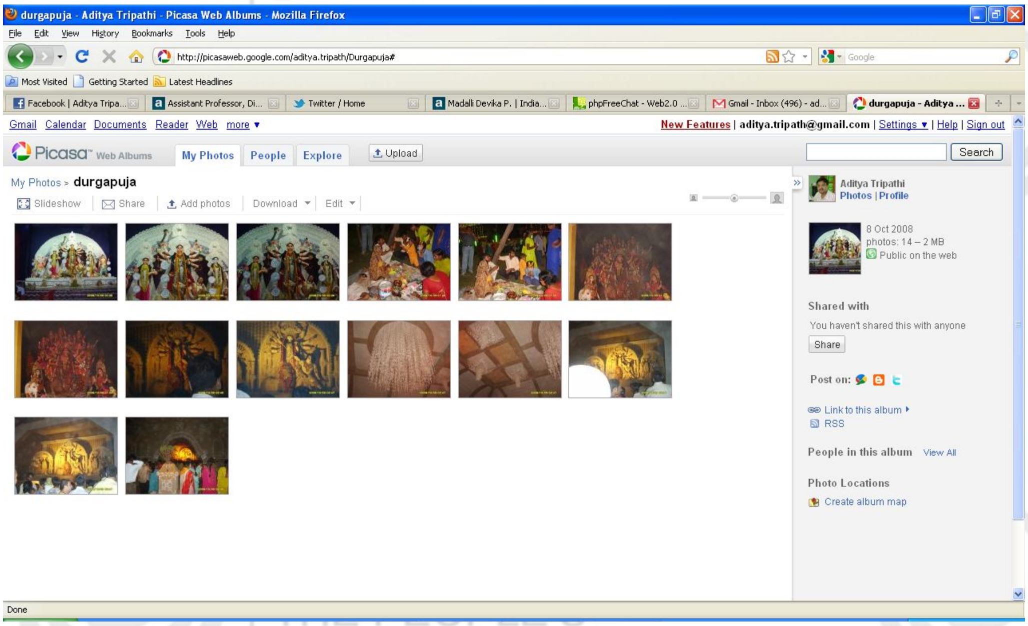


Fig. 14.9: Picasa

Movie Sharing: YouTube

It a website to share video clips (Fig. 14.10). Presently, it is owned by Google and acting as subsidiary firm to Google. It was developed in 2005 by some of the employees of PayPal, a website for making online transaction. This technology uses Adobe Flash Video technology to display a wide variety of user-generated video content, including movie clips, TV clips, and music videos, as well as amateur content such as video blogging and short original videos. Individuals have uploaded most of the content on YouTube. Many of the media corporate also do offer their content through YouTube. Anyone can watch the video, as they are available in open domain free of cost.

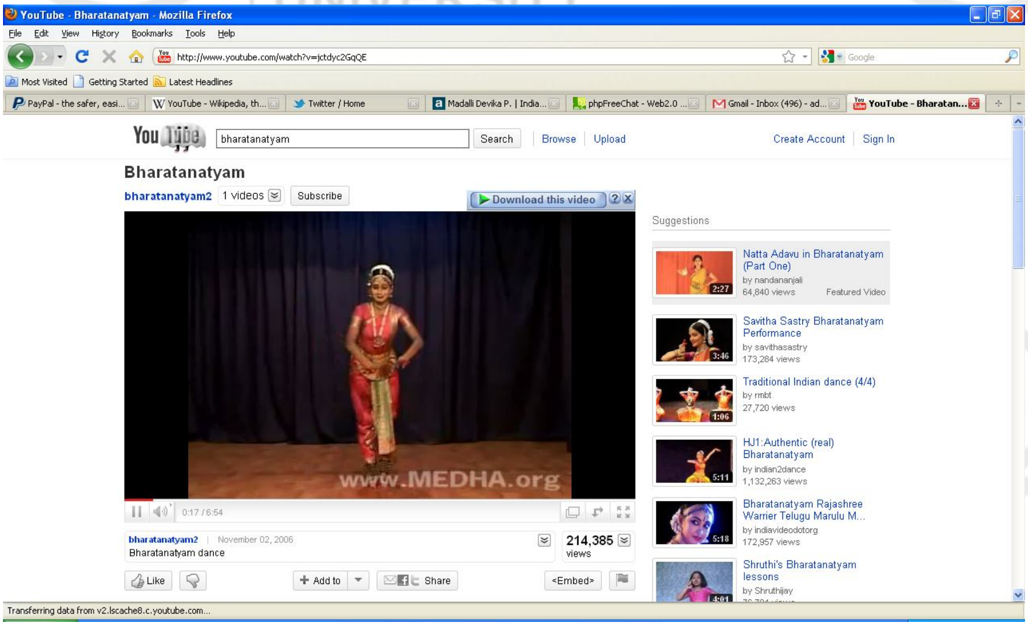


Fig.14.10: YouTube

Social Networking

Social networking is a remarkable phenomenon of Web 2.0 implementation. People can connect online with each other. They can search for old friends or form a group of people with same kind of interest. Such sites are not only used by common man to connect to mass or by celebrities but they have become a tool for advertising firms and products. These social networking sites contain several applications like, image sharing, video sharing, online chatting, emailing, commenting or sending public as well as private messages to an individual. These sites are available over mobile applications in mobile phones so that one can keep connected to people even when he/she is travelling.

- **Facebook**

With all the above mentioned features, Facebook (Fig. 14.11) presently hosts more than 500 million active users and on any given day 50% of the registered users use Facebook. It has become such an important phenomenon for the people using Internet that on average each user has 130 friends or people to which they are connected with. One can share any object with others, which includes pages, groups, events and community pages. It has become a largest sharing platform where around 550 thousand objects are being shared globally. The popularity of Facebook is so much that it is being translated by user community voluntarily using its translation application. Facebook is widely available as an application in mobile phones of various brands. There are more than 200 mobile operators in 60 countries working to deploy and promote Facebook.

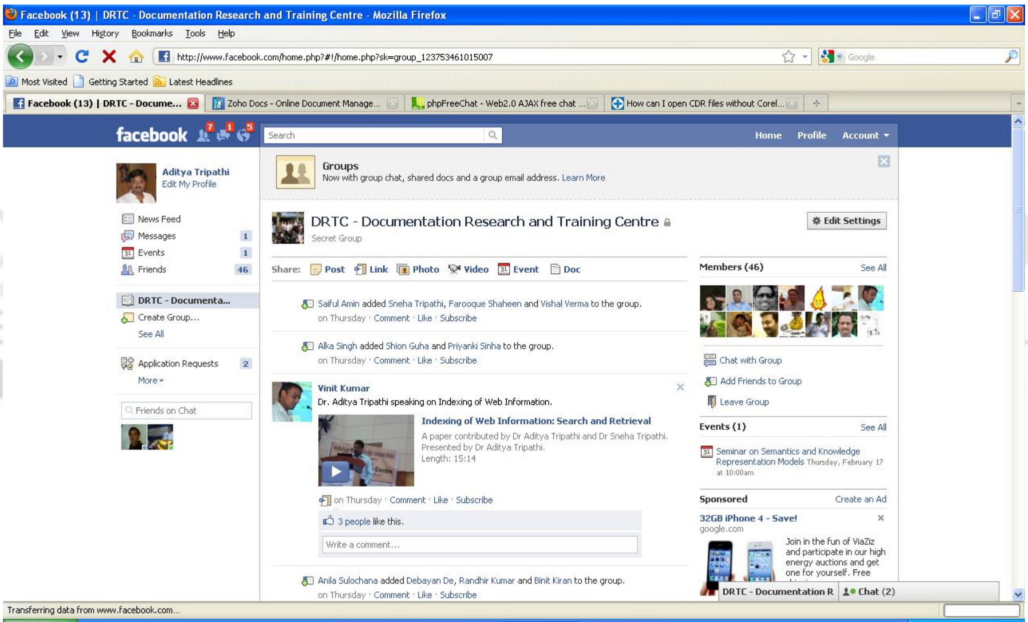


Fig. 14.11: A Facebook Account

Twitter and Orkut are some of the other examples of social networking website. Twitter is very popular among the celebrities. People can follow each other over Twitter.

- **Academia.edu**

This is a social networking site specially designed for the people with academic interests. They can share their ideas with others. It facilitates sharing research interests by hosting published works of their users. He/she can share research papers and articles. It also sends notification to its users if others are searching them. It collects jobs and other academic opportunities for the user as per his/her interests. It is very popular socialising site among the academic and research community.

Information Mashup

These are website which uses the content and services from different service provider about different aspects of subject. For example, if someone is looking for Delhi. A

Mashup website would provide information about the map of Delhi, Weather Forecast, News from Delhi, Photographs of important places, mode of transport, administrative set up and so on. The number of the services, which can be included in Information Mashup, is never ending. Mashup is important to make already existing data more useful, pertinently for personal and professional use. These are hybrid services and often look as an advanced version of web portals, which used to host static webpages.

- **Indianrailinfo**

Indianrailinfo (Fig. 14.12) is a mashup of different kind of information about Indian railways. This website collects information from different websites of Indian railways. It is one stop portal for information about arrival/departure of trains from a station, PNR status, seat availability and so on. IndianrailInfo is also connected through Google, Yahoo and Facebook and it can extract user's information through these websites.

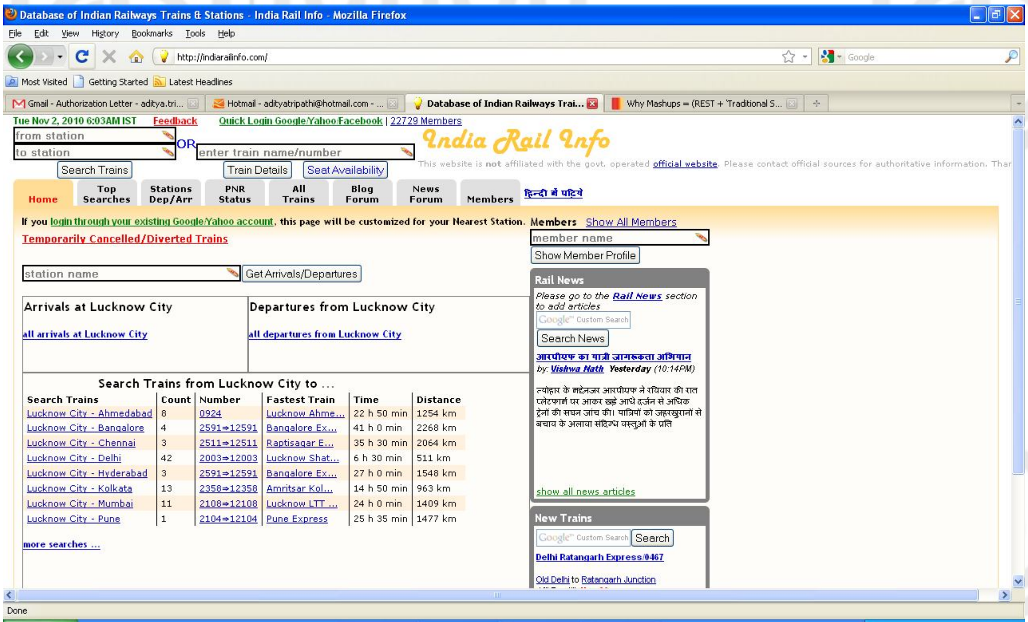


Fig. 14.12: Indianrailinfo

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

5) Name different types of Web 2.0 applications.

....

....

....

....

14.6 IMPACT OF WEB 2.0 TOOLS OVER WWW AND SEMANTIC WEB

Though there is no clear demarcation between Web 1.0 and Web 2.0 but the product and services offered under the umbrella of later is quite different from the previous. The services are more focused and interoperable in nature. One application or web service could exploit the data from other. These services are wise enough to correlate the data of users need and services offered by service providers. Semantic web is about developing such kind of services or agents, which are more meaningful from the users' point of view. Implementation of web ontology can bring in more contexts in the pursuit of information and web-based services from the user's point of view.

14.7 SUMMARY

In the present chapter we learnt how the present Web has evolved from a simple idea of Sir Tim Berners Lee. It has become a platform of many complex services like, blogs, mashups, and social networking websites. The existing technology is still in evolutionary stage and more is expected to come in the form of Semantic Web. In the present chapter we could create understanding about,

- The framework on which Web functions
- The languages which are used to develop web applications
- Various Web 2.0 based services.

14.8 ANSWERS TO SELF CHECK EXERCISES

1) Web Browser is an application software to view webpages over Internet. It is used to retrieve, view and transfer information over WWW.

2) Addressing protocols is a standard to identify a web object by its names and addresses. It is known as Uniform Resource Identifiers (URIs). Uniform Resource Locator (URL) is a kind of URI which is used for locating web objects or webpages over WWW. It is a global addressing system, which is translated to IP address in order to locate the server containing information. The first part of the address is known as protocol identifier and it indicates what protocol to use; the second part is called as resource name and it specifies the IP address through the domain name where the resource is located. The protocol identifier and the resource name are separated by a colon and two forward slashes.

3) HTTP provides mechanism to address a website. It is a networking protocol for distributed, collaborative, hypermedia information systems. HTTP is used for data communication over World Wide Web. It supports retrieval, which includes search, front-end update, and annotation. HTTP protocol is based on a request from a client and response from a server. A server hosts webpages and on request delivers it to the client. This kind of request/response format of communication is known as Client – Server Architecture. Every request is treated as a new request and there is no dedicated connection between client and server. This is known as stateless connection. HTTP includes support to SMTP, NNTP, FTP, Gopher, and WAIS protocols. The last two are no more in use for real world web communication. Hence, HTTP allows basic hypermedia access to resources available from diverse applications.

4) Following are the key features of Web 2.0 applications,

   - Users' Participation
   - User centric services
   - Decentralisation and Interoperability
   - Hiding technological complexity
   - Modularity

5) Different types of Web 2.0 applications are:

- Collaborative Web Tools – Wikipedia, Drupal, Joomla, Blog, Trac, Bugzilla etc.
- Interactive Web Tools – Gtalk, Yahoo Chat, GoogleDoc, Picasa, Flickr
- Social Networking Tools- Orkut, Facebook, Twitter etc.
- Information Mashup – Google Map, Indiarailinfo, etc.

14.9 KEYWORDS

Blog : A website where entries are displayed in a reverse chronological order. A Blog is used to communicate to the audience and reverse. A blog may use text, images, and links to other blogs, web pages, and other media related.

eLearning : Stands of Electronic learning. E-learning is a method to impart distance education through electronic media which includes, online and offline media.

Flickr : It is a photo sharing website and web services where members can share photographs. The photographs uploaded on the website can be tagged for efficient retrieval.

Folksonomy : Folksonomy is community based labeling or tagging system performed by the Internet community (users). It is an open-ended labeling system for web objects including web pages, photographs, software and so on. This tagging system is intended to make a body of information increasingly easier to search, discover, and navigate over time.

Gopher : A distributed document search and retrieval network protocol designed for the Internet, which is now obsolete.

Intelligent Agents : A piece of software, which has intelligence to draw the inference from the fact.

Learning Management System (LMS) : Piece of software that enables the management and delivery of learning content and resources to students. Normally, these software works in online environment.

Listserv : An email-based mailing list.

Podcasting : Multimedia file distributed over the Internet using syndication feeds (RSS), for playback on mobile devices and personal computers.

RSS : Stands for Really Simple Syndicate. A standard mechanism of offering content, which can be read by web browsers as bookmark. It is a simple XML-based system that allows users to subscribe to their favourite websites.

Semantic Web : An idea of more meaningful web where resource discovery will be assisted by intelligent agents.

YouTube : A Web 2.0 based service to share and view video clips.

Web 2.0 : New generation of Web where users can have better control and voice towards using web services and products. The services are more user centric and interactive.

Wiki : An effective tool for collaborative authoring.

14.10 REFERENCES AND FURTHER READING

1. Arms, William. *Digital Libraries*, M.I.T. Press, 1999. <http://www.cs.cornell.edu/wya/DigLib/MS1999/Chapter2.html>

2. Bugzilla. < http://www.bugzilla.org>

3. E-mail. (n.d.). Available at *Wikipedia*: <http://en.wikipedia.org/wiki/E-mail>

4. Enhancing Web 2.0 literacy: Using blogging as a management tool. *Strategic Direction*. 24.11 (2008) pp. 12-13. DOI: 10.1108/02580540810914977

5. Extensible Markup Language (XML) 1.0 Available *at World Wide Web Consortium* (1998) < http://www.w3.org/TR/1998/REC-xml-19980210.html#sec-intro>

6. Flynn, Peter (ed.). *The XML FAQ: Frequently-Asked Questions about the Extensible Markup Language*. (2010) < http://xml.silmaril.ie/basics/markup/ >

7. Glossary. Available at *European Union's Publisher Website* <http://publications.europa.eu/vademecum/vademecum/9313fdfe-c49e-119e-45c6a6441e63e066_en.html>

8. Gralla, P.. *How the Internet works*. 4th ed. Indianapolis: Que Publishing, 1998. Print

9. How the web began. (n.d.) Available at CERN website. <http://public.web.cern.ch/public/en/about/webstory-en.html>

10. HTTP Properties (2006) <http://www.wdvl.com/Internet/Protocols/HTTP/Properties.html>

11. Johnson, Tony. *Spinning the World-Wide Web*. (1994) <http://www-sld.slac.stanford.edu/sldwww/beamline/article.html#concepts>

12. Laningham, Scott (ed.) IBM developerworks Interviews: Tim Berners-Lee. (2006). <www.ibm.com/developerworks/podcast/dwi/cm-int082206txt.html>

13. Learn The Net: E-mail. (n.d.). <http://www.learnthenet.com/english/section/email.html>

14. Levine, J. R., Young, M. L. & Baroudi, C. *The Internet for Dummies*. 11^{th} ed. Hoboken: Wiley Publishing, 2007. Print

15. Lindsell-Roberts, S. *135 tips on email and instant messages: plus blogs, chatrooms, and texting*. Boston: Houghton Mifflin Company, 2008. Print

16. Multipurpose Internet Mail Extensions. (n.d.). Available at *Wikipedia.* <http://en.wikipedia.org/wiki/Multipurpose_Internet_Mail_Extensions>

17. Nations, Daniel. *What is Web 2.0? How Web 2.0 Is Defining Society.* (2010). <http://webtrends.about.com/od/web20/a/what-is-web20.htm>

18. O'Reilly, T. *What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software*. (2005). <www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1>

19. Simple Mail Transfer Protocol. (n.d.). Available at *Wikipedia*. <http://en.wikipedia.org/wiki/Simple_Mail_Transfer_Protocol>

20. The history of email. (n.d.). Available at *Net History*. <http://www.nethistory.info/History%20of%20the%20Internet/email.html>

21. Thomas, Brian J. *The Internet for scientists and engineers: online tools and resources*. 3^{rd} ed. Oxford: Oxford University Press, 1997. Print

22. Trac. < http://trac.edgewall.org/ >

23. Wikipedia< http://www.wikipedia.org/>

24. Wood, D. *Programming Internet Email*. Sebastopol: O'Reilly & Associates, 1999. Print

25. Young, Linda. A social media glossary. (2008) < http://www.capilanou.ca/help/login-page/active-cms/glossary.html>

26. Yves, Lafon. *HTTP - Hypertext Transfer Protocol Available at World Wide Web Consortium.* (2010) < http://www.w3.org/Protocols/Overview.html>

UNIT 15 SEARCH ENGINES

Structure

15.0 OBJECTIVES

After going through this Unit, you will be able to:

- describe different search tools available to locate information on web;
- discuss search engines as an effective information retrieval tool;
- know different search strategies to increase the performance of search tools; and
- explain search engine architecture.

15.1 INTRODUCTION

Internet has created revolutionary changes in this era of Information Technology. For many, it is one stop platform to find or locate any information they are interested in. Traditionally, librarians had the job to assist their users to locate the information they needed. But, now the scenario has changed a lot. Internet has in offer a variety of search tools such as search engines, search directories to locate the information on web.

A search on web is a simple process and can be conducted by simply issuing a query to the search tool. The search tool in return will look for the information in its web based information databases and retrieves those, which are relevant to the query. Searching is an iterative process i.e. one needs to keep working on their query unless the exact information is located.

The very first tool used for searching on the Internet was Archie. The name stands for "archive" without the "v." It was created in 1990 by Alan Emtage, a student at McGill University in Montreal. Veronica (*V*ery *E*asy *R*odent-*O*riented *N*et-wide *I*ndex to *C*omputerised *A*rchives) and Jughead (*J*onsy's *U*niversal *G*opher *H*ierarchy *E*xcavation *A*nd *D*isplay) were two other popular search programs.

There are three basic types of search tools that most people use to find what they are looking for on the Web: Search Engines, Subject Directories and Meta Search Tools. Search Engines are more generic and much larger than Subject Directories. Meta Search Tools get their results from several search engines. The following sections will provide an elaboration on these search tools.

15.2 SEARCH ENGINES

Search engine is a tool for locating information from a collection. Search engines uses information about the information (such as metadata, catalogue) stored in the database to locate information. Sometimes they perform full text search within the document from first character to last character.

The search is done on pattern matching algorithm whether it is a database or full text.

15.3 TYPES OF SEARCH TOOLS

15.3.1 Search Directory

Search directories are classified collections of documents. They are good for searching with a context. These directories are good for browsing. In subject directories, documents are pre classified by a person. Librarians' Internet Index; Google Directory; Yahoo!; dmoz are some of the examples of subject directories.

There are two basic types of directories:

- Academic and Professional Directories: These are often created and maintained by subject experts to support the needs of researchers. INFOMINE, from the University of California, is a good example of an academic directory.

- Commercial Directories: These cater to the needs of general public. Directories of Yahoo! and Google are examples of commercial directories.

Fig. 15.1: Librarians' Internet Index (www.lii.org)

15.3.2 Search Engines

World Wide Web is a network of several information databases. In recent years, an exponential growth in these databases has made it difficult to locate a particular piece of information. Internet offers a powerful tool known as search engine to manage, filter and retrieve the information for their users.

Search engines are automated tools for searching information from a collection using metadata stored in the database of search engine. In other words, it is an information retrieval system and assists in locating information on web.

Google and Yahoo! are most popular search engines.

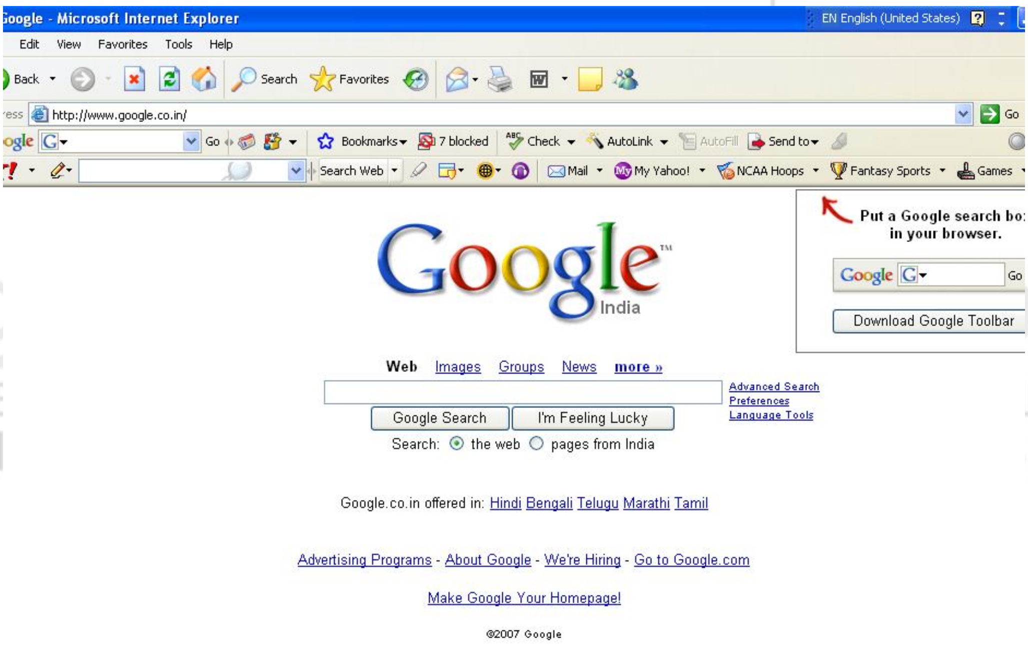


Fig. 15.2: Search Interface of Google

15.3.3 Meta-search Engines

Meta Search engines are online tools (search engines) which performs simultaneous search on more than one search engine at a time. These search engines aggregates the results into a single list and displays them according to their source. e.g. Dogpile is a metasearch engine and gets its results from Google, Yahoo, MSN Search, Ask , About, MIVA, LookSmart, and more.

Example: Dogpile, WebCrawler, Browsys

Fig. 15.3: Search Interface of "Dogpile", a Metasearch Engine

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) Write a short note on Search Tools?

2) Name any four search engines and two metasearch engines?

....

....

....

....

15.4 FEATURES OF SEARCH TOOLS

The searching tools follows pattern matching algorithm. There are many types of searches can be done using search tools,

15.4.1 Keyword Search

When searching is done using a keyword it is known as Keyword searching. Keyword may occur at any place in the document or in the metadata field. This kind of search has higher recall value.

15.4.2 Boolean Search

Logical AND, OR and NOT are known as Boolean operators. When Boolean operators are used for searching it is known as Boolean search. The operators are used for combining more than one word with certain conditions. These kind of searching also known as Combinatorial search.

➢ **AND**

This operator will retrieve all the documents which contains all the keywords occurring at both ends of the AND operator.

Syntax: <Search Term **A**> AND <Search Term **B**>

Example: Library ***AND*** Information

Output:

1) The above query will retrieve only those documents which contains both the terms Library and Documentation

2) The precision in search is more. The number of documents retrieved will be less hence less is the recall value.

➢ **OR**

This operator will retrieve all the documents which contains all the keywords occurring at both ends of the OR operator.

Syntax: <Search Term **A**> OR <Search Term **B**>

Example: Library ***OR*** Information

Output:

1) The above query will retrieve all documents which contains both the terms Library and Documentation

2) The recall in search is more. The number of documents retrieved will be more but the precision in retrieved documents will be less.

➢ **NOT or AND NOT**

These operators increase the precision of the search result. The query can be made more specific by using these operators. Using the capitalised AND NOT operator preceding a search term eliminates documents that contain that term.

Syntax

<Words to be searched> AND NOT <Words not to be searched>

Example:

If user is looking for information on Drivers and do not want documents that include information relating to the Screw Drivers the query could be "Driver" AND NOT Screw.

15.4.3 Proximity Search

This is another kind of Combinatorial search where the proximity of two words is checked. The term proximity means 'nearness of words'. Proximity is given in terms of number of words by which two words should be separated. There are two kinds of proximities,

1) Near Proximity

2) Exact Proximity

1) **Near Proximity**

Near Proximity brings range of search results where the number of proximity is from adjacent to the mentioned number proximity. For example, for two keywords COLLEGE and LIBRARIANS, if the proximity of 3 is applied between them it will bring the results as follows,

COLLEGE LIBRARIANS

COLLEGE FOR LIBRARIANS

COLLEGE OF LIBRARIANS

COLLEGE WITHOUT LIBRARIANS

COLLEGE WITH THE LIBRARIANS

COLLEGE OF INDIAN LIBRARIANS

COLLEGE OF THE LIBRARIANS

The Near Proximity would bring the results where the search terms would be separated by no word to n-1 word (where n is number of proximity). In other words, near proximity brings all the proximities which are lesser than the mentioned number.

2) **Exact Proximity**

Exact proximity brings the results with exact number of proximity mentioned. It does not bring the results which have lesser number of proximity. For example, if the number of proximity is set to 3 between two keywords i.e. COLLEGE and LIBRARIANS. The retrieved result would be,

COLLEGE WITH THE LIBRARIANS

COLLEGE OF INDIAN LIBRARIANS

COLLEGE OF THE LIBRARIANS

15.4.4 Truncation Search

Truncation means concatenation of words. In other words, if the root string of the words is searched it brings all the derivatives derived out of the given root string. Truncation is of three types based on truncation techniques:

1) **Left Truncation**

When the root string is concatenated from the left side, it is known as left truncation. For example, if the left truncation is implemented for the root string ISM, it will bring all the words which ends with the string ISM, like

BRAHAMINISM

COMMUNISM

SUPHISM

2) **Right Truncation**

When the root string is concatenated from the right side it is known as right truncation. For example, right truncation is used with the root string CLASS, it will bring all the words which starts with the root string CLASS, like

CLASS

CLASSIFICATION

CLASSIFICATIONIST

CLASSIFIER

15.4.5 Case Sensitive Search

One of the major features of search tools is their support to search words based on their case. In other words, search tools can differentiate between Upper and lower cases. For example, DUKE and duke will bring different search results based on the case. In an ordinary/plain search, search tool performs searching irrespective of their cases. However, if case sensitive search is invoked, search tool brings exact search string based on the case of search string.

15.4.6 Limiting Search

There are certain conditions based on which the searches can be narrowed down, for example, by Date, by Domain, by media type, by Document Directory Depth, by Page Depth and so on. This kind of condition reduces the number of search results and increases the relevancy of final output.

15.4.7 Fields Search

Field Search is a kind of **limiting search** to a particular field of the database. Searching can be done within a given context. For example, searching within Title or searching within Author or searching within both the fields. This kind of searching is known as Field search.

15.4.8 File Types Search

When the searching is restricted to a particular file type like, MS-Word, PDF, PPT etc. it is known as file type search.

15.4.9 Stop Words

While searching documents in a collection or over Web, some frequently occurring words like prepositions, conjunction etc. should be avoided. In order to avoid such redundant words from the search results, the tools contain a file called stop word file. This file lists all the words which are to be avoided from being indexed. This saves space of storage and reduces time of search.

15.4.10 Ranking

Search tools present the search results in some order. Normally, when system is small the presented results are arranged in alphabetical order. But when results run in several pages, it becomes important to present the most relevant document on the top followed by less relevant one. Thus, it is important to rank the retrieved documents based on their relevancy to their users. Search engines have an automated mechanism to rank the retrieved results according to the relevancy of each retrieved search result.

For example, PageRank of Google is an algorithm for measuring weightage of results based on link analysis.

15.4.11 Family Filters

Family filters are used to reduce, if not remove, the objectionable matter to appear on search results. Search engines do provide functionality for setting family filters as safe search. Family filters are used by Google (as Safe Search), Altavista, Yahoo and so on. Apart from the search engines there are tools (e.g. Naomi, which is a freeware) which can be loaded on computers to stop display of obscene matters.

15.4.12 Fuzzy Search

Fussy search is one of the major features of today's information retrieval system. It brings out results based on approximations. In other words, these are error correction algorithms. For example, if a keyword is miss-spelt search algorithm used in searching attempts to render the search result according to correct spelling. Such algorithms are known as Soundex and Metaphone algorithms. Levenshtein distance algorithm is one such kind of algorithm used by Lucene search engine, an open source search tool.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) Write short note on Boolean Search?

4) What is a Proximity Search?

5) What is "PageRank"?

....

....

....

....

15.5 ARCHITECTURE OF SEARCH TOOLS

WWW is a huge source of information and search engines are tools or agents for locating information. Everything and anything can be located over Internet using search engines. Search Engines are tools which provide a kind of interface for users to search the web. A Search Engine basically has three components:

- Web Crawler
- Metadata storage
- Search Agent

15.5.1 Web Crawler

Web crawler is also known as robot or spider. It is a program which goes to each and every site over Internet and indexes the content of the webpage. The content includes metadata information and the text from the webpage. Text of the page can be indexed as a whole or only few lines or bytes of data are stored. This index is stored within search engines database with corresponding URL (Uniform Resource Locator).

Hence, Web crawler is a program used by search engine in order to extract data from the web pages so that pages can be searched using the search engine's interface. Following are the names of web crawlers used by popular search engines:

Search engine	**Robot**
Google	Googlebot
Yahoo	Slurp
MSN	MSNbot

15.5.2 Metadata Storage

Metadata is data about data. Web Crawlers extract metadata like title, author, filename, file type, file size, links and so on from search engines. The metadata information is extracted from meta-tag of webpage, file name, file extension, etc. The collected metadata is stored in the repository of search engine, which is a database, in the form of index. Normally, many of the system follow a kind of keyword indexing. But the Keyword indexing is good for recall not for precision. In such cases the context for search is lost. However, the use of metadata is good for preserving the context of search term. The metadata is stored in the form of metadata index inside the search engine database.

Fig. 15.4: Architecture of Search Engine

The most commonly used metadata schema is Dublin Core Metadata Initiative (DCMI) over Internet. The standard is developed and maintained by DCMI and DCMI Task Groups. There are 15 elements given in Dublin Core. Apart from these 15 elements there are other metadata set vocabularies which should be used with 15 elements.

Title	Documentation Research & Training Centre – Home page
Creator	Aditya Tripathi
Publisher	Documentation Research & Training Centre
Identifier	http://drtc.isibang.ac.in
Subject	Library and Information Science
Format	txt/html
Language	English
Rights	Indian Statistical Institute

Fig. 15.5: Dublin Core Metadata Record

15.5.3 Search Agent

Search Agent is a set of search programs which receives query from the user interface. The received key word is passed through several algorithms. Some commonly used algorithms are as follows:

- Boolean operator
- Data clustering algorithms
- Error correction algorithms.

Once the query or search term is processed, the agent performs search within the repository of the search engine and retrieves the search result and send to user interface.

Boolean Operators

The Boolean operators are AND, OR and NOT. These operators are used to generate combinatorial search. AND and NOT operators increase precision where as OR increases recall of search results. The shaded area represents retrieved records in the following example (Fig. 15.6). Almost all the search engines provide facility of using Boolean operators. These operators can also be used to combine keywords present in different fields.

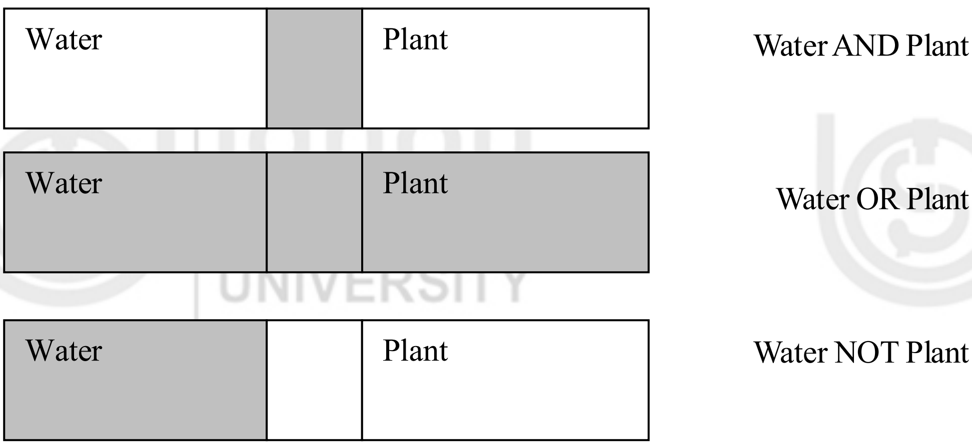


Fig. 15.6: Boolean Operators

Clustering Algorithms

Cluster is group of like entities. Search engines use clustering technique to classify like concepts for more meaningful retrieval. Clustering is a technique by which similar kind of data are grouped based on certain characteristics. This technique is very useful when there is large set of data for effective and efficient retrieval. For example, the faculty member of a university can be clustered according to their area of specialisation.

The clustering algorithm attempts to identify groups in a given set of data or population based on likeness of certain characteristics or traits. Thus it creates a picture of a big group and then inside several sub groups. The algorithms attempts to identify core entity which is also known as Centroid. Hence, centriod is the centre of the concept or the core concept. The other concepts are placed around the core concept. The likeness decreases with the increasing distance of the concept from the core concept or the centroid. Hence, to determine cluster membership and sise, the algorithm evaluates the distance between concept and the core. Such techniques are highly dependent on use of statistics for generating clusters.

Error correction Algorithms

While typing the keywords often the searcher commits mistake. Search engines are equipped with the algorithms which correct the spelling and yield meaningful results. Google's "Did you mean" uses these kind of algorithms. In case of incorrect spelling or if a word is not found in its database, Google suggests alternate spellings for the keyword.

Soundex and metaphone algorithms are examples of such kind of algorithms. Both algorithms are based on the pronunciation of a word.

In **soundex algorithm** a numeric code is assigned to each character used in a word and when search is performed, words with similar codes are also brought out in search result. Metaphone also works on same algorithm but unlike soundex which encodes a word on letter-by-letter basis, it encodes groups of letters i.e. a word.

Metaphone algorithm embodies more accurately the rules of pronunciation in language. Such algorithms are well established for English as a language. Both algorithms return all words that exactly match the desired word as well as all similar sounding words (homophones).

15.5.4 User Interface

User Interface is the part which interacts with the user. It is like a switch board for the user for invoking the system to perform search for needed information. There are two parts of User Interface:

- Search Interface
- Result Interface

Search Interface

This is the end from where users enter his/her search terms. It is one of the major components which initiate the communication between users and the system. The Search Interface performs following tasks:

1) *Capturing user input/query*

It is also known as front-end. The interface captures keywords given by users. It passes it to the search agent. The look and feel of the front-end should be easy to operate.

2) *Search refinement*

Search interface should have facility to refine the search. The refinement facility should be given within the displayed search results. Hence, user interface should provide facilities for modifying search statement. Sometimes user interface gives facility for browsing pre-classified categories.

3) *Advance Search statement*

User interface should have another interface for advance level search. The advance level search includes, use of Boolean operators, image search, file-type search, language search, date-wise search and so on.

Result Interface

Display of search results is another important aspect of searching. It should be in user friendly format and customisable by the user.

Ranking of search result is an important feature. The search engines use algorithms for ranking. Google uses an algorithm called PageRank. In this, link analysis is done for ranking of retrieved documents. Some search engines rank their results based on the frequency of occurrence of search terms. Statistical techniques are widely used for this purpose.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

6) What is a “Web Crawler”?

7) What is Clustering?

8) What is Soundex Algorithm?

....

....

....

....

15.6 CHALLENGES

Searching on Internet has gone a long way. In the era of Web 2.0, search engines are well equipped with several algorithms like soundex, metaphones, page rank, etc. for improved retrieval of search results. One area which is still required to be improved is context based searching. Though clustering techniques are applied for setting the context but these methods are not fools proof. The implementation of Semantic web is ahead which promises search agents with specialised search strategies and features. These agents would not only perform the search, they will also guide the users in taking decision.

The idea of Semantic web is still evolving and it has still long way to go. But there are products available in the market which has started making mark on the Internet arena like, Wikipedia, Social networking sites, blogs and so on.

15.7 SUMMARY

This Unit highlights the features of search tools. Search engines are web based search tools to search the documents or objects over Internet. The Unit has discussed different types of search engines i.e. search engine, meta-searching engine and search directory. Search engines have two kinds of search interfaces, simple and advance. Advance search facilitates different kinds of searches such as Boolean search, proximity search, truncation search, case sensitive search, fields level search, file types search, stop words, sorting, etc.

The Unit elaborates the architecture of search engines with reference to user interface, search agents and web crawler.

15.8 ANSWERS TO SELF CHECK EXECRICES

1) To locate information on web, search tools play an important role. The very first tool used for searching was Archie.

   Internet offers various types of search tools:

   a) search engine

   b) subject directories

   c) metasearch engines

2) Search Engine:

   a) Google

   b) Yahoo!

   c) Alta-Vista

   d) Vivisimo (vivisimo.org)

   Meta-search Engine:

   e) Dogpile

   f) Browsys

3) When Boolean operators are used in a search it is known as Boolean Search. The Logical operators such as AND, OR, NOT are known as Boolean Operators.

4) Proximity means 'nearness'. Hence, when a search is based on the proximity or nearness between two words is known as Proximity Search.

5) PageRank is an algorithm used by Google for measuring weightage of results based on link analysis to rank the retrieved results.

6) Web crawler is a program used by search engine in order to extract data from the web pages so that pages can be searched using the search engine's interface. It is also known as robot or spider.

7) Clustering is a technique by which similar kind of data are grouped based on certain characteristics. This technique is very useful when there is large set of data for effective and efficient retrieval.

8) Soundex is an error-correction phonetic algorithm. If there is minor error in the spelling of the search term or there are cases of homophones, this algorithm helps in retrieving the results.

15.9 KEYWORDS

Directories	:	Lists of pages classified into useful categories, (like Yahoo or Looksmart).
Exact Match	:	If a document contains exact match to the query, then only it will get retrieved. Increases precision of result but low recall value.
Query	:	A query is the combination of the word or words used for searching.
Recall	:	Total retrieved records against a query.
Precision	:	Total relevant records retrieved against a query.
Relevance	:	The extent to which retrieved records against a query satisfies the end-user.
Search engine	:	A program that indexes web documents and facilitates user to perform search on them.
Weighting	:	Weighting is a heuristic technique designed to improve the relevance ranking algorithms.
Index term	:	It is a pre-defined term which can be used to refer to the content of a document.
Full Text Search	:	It is a methodology in which all the words which compose the text of the document are used as indexing terms.
Fussy Model	:	It is "set theoretic model" of document retrieval based on fussy theory.
Inverted file	:	An index composed of vocabulary and list of occurrences.

15.10 REFERENCES AND FURTHER READING

1. Besroukov, Nikolai. Searching Algorithms (2009). <http://www.softpanorama.org/Algorithms/searching.shtml>

2. Find Information: Advanced Web Searching. <www.learnthenet.com/english/html/77advanc.htm>

3. Franklin, Curt. How Internet Search Engines Work <http://www.howstuffworks.com/search-engine1.htm>

4. History of Search Engines: From 1945 to Google 2007. <http://www.searchenginehistory.com/>

5. Houghton, Bernard and Convey, John. *On-line information retrieval systems: an introductory manual to principles and practice*. Clive Bingley: London. Print

6. Information Skills Online (2006). <http://infoskills.mq.edu.au/scd/end.htm>

7. Meta Search Engines. <www.lib.berkeley.edu/TeachingLib/Guides/Internet/MetaSearch.html>

8. Subject Directories (2004). <http://www.learnwebskills.com/search/subject.html>

9. Sullivan, Danny. Where Are They Now? Search Engines We've Known & Loved. *Search Engine Watch*, Mar 4, 2003.< http://searchenginewatch.com/2175241>

10. Taylor, Chris. Introduction to metadata (2007). <http://www.library.uq.edu.au/iad/ctmeta4.html>

11. Walt Howe. A brief history of the Internet. <http://www.walthowe.com/navnet/history.html>

UNIT 16 INTERACTIVE AND DISTRIBUTIVE SERVICES

Structure

16.0 OBJECTIVES

After reading this Unit, you will be able to:

- describe interactive services and the benefits associated with them;
- explain prominent distributive services with examples; and
- discuss issues involved with the implementation of these services.

16.1 INTRODUCTION

Interactivity is defined as one to one communication between the two systems. This is the state of a system when the responses actually depend on the inputs received from outside of the system. If we look at the present services available on Web one can easily say that interactivity has been taken care seriously by the developers in the near past. The days of having a system stand-alone and with no interactions are gone. This is the pressure of more involvement of user collaboration causing the implementation of more interactivity. The interactivity further makes the system more user friendly, which subsequently enhances the user satisfaction.

On the other hand, these services provide great chance of collaboration and exchange of information.

16.2 DISTRIBUTED SERVICE

Distributed services are offered through distributed sites. In other words, the machines, which host the services, are placed at different geographical locations. However, these machines interact in a standard manner to offer one or multiple service under one umbrella service. In the subsequent sections we will discuss some of the important distributed services.

16.2.1 Web Directory

The Web contains trillions of terabytes of data and the information is not organised properly. There are some search engines available but sometimes getting answers to questions like, “What resources does the World Wide Web have on Algebra?” or “What kind of information is available on Knowledge Management?” get very tough. The web directory serves as an important service for providing answers to the above-mentioned questions. Web directories are nothing but a topical list of Internet resources arranged in a hierarchical way. Unlike search engines where Web is indexed by using robots and web directories are human maintained and created. Often these people are volunteers or sometimes hired. Generally, web directories are meant to be browsed by subject or topics, they can be searched by keywords too. Some of the popular web directories are listed in table 1 and main page screenshots given in the figures following the table from fig. 16.1 to 16.3.

Table 16.1: Popular Web Directories

Name	Directory URL	Remarks
BUBL Information Service	http://bubl.ac.uk/	From Librarians
Google Web Directory	http://www.google.com/dirhp	
Yahoo Directory	http://dir.yahoo.com/	World’s first web directory
DMOS Open Directory Project	http://www.dmos.org/	Public involvement

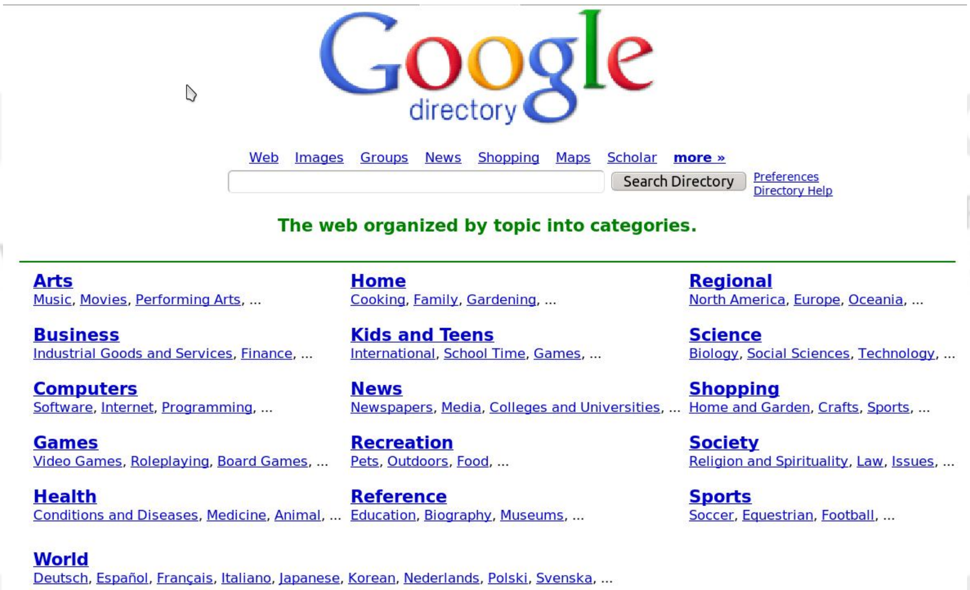


Fig. 16.1: Google Web Directory (http://www.google.com/dirhp)

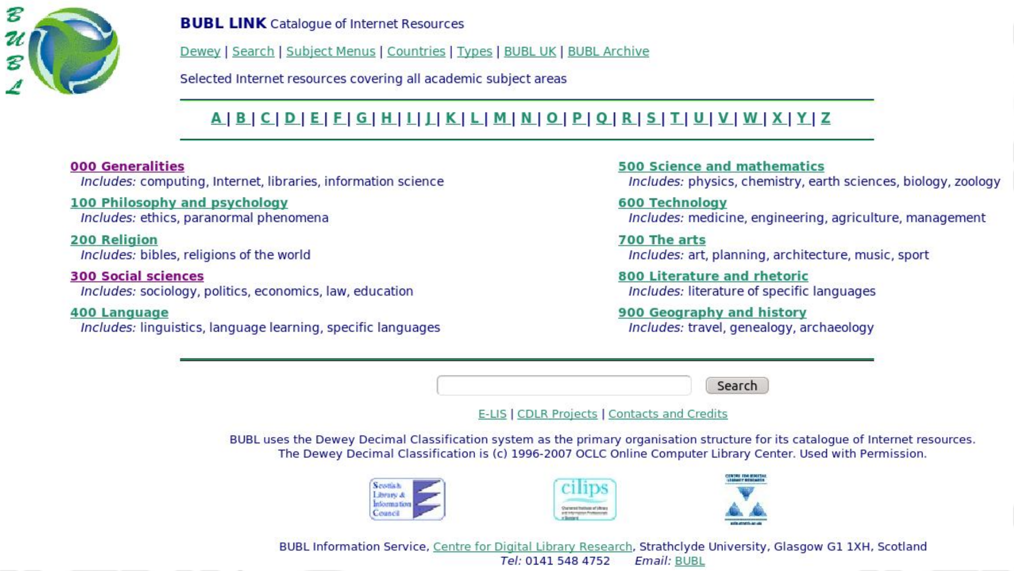


Fig. 16.2: BUBL Information Service (http://bubl.ac.uk/)

Fig. 16.3: DMOS Open Directory Project (http://www.dmos.org/)

The human element involved in creating and maintaining directories creates both advantages and disadvantages for the user. Some of the advantages are:

i) They contain fewer resources and reduces the information avalanche.

ii) Many directories rate, annotate and categorise the chosen resources.

iii) Directories increase the probability of retrieving the relevant results.

On the other hand, they have some disadvantages too:

i) Most of the directories follow their own hierarchical arrangement, which sometimes leads to arbitrary arrangement too.

ii) Because of human based maintenance the update is not so frequent and sometimes not at all happening.

iii) The subjectivity of the annotation and rating is often under question.

iv) It involves a good understanding or guess about the closest topic about the subject to be researched upon. For example if a user digs in some topic in the directory and finds no resource s/he has to guess or look for another topic.

16.2.2 Bulletin Board

Bulletin or message boards provide a facility for discussion under various topics. They allow individuals to respond to topics or threads in the group, or to begin a new topic or thread by posting a comment or question. The messages posted to a discussion board are permanently visible to everyone who has access to it. Most of the discussion boards implement staged users, the users enter the board as primitive member, in board specific terminology they are known as "newbie" and then by asking questions and posting answers to the question asked by others they get promoted to "starred" members. This is a common phenomenon in many of the Bulletin Boards.

There are millions of bulletin boards available on the Internet. Many news oriented websites, search engines, social networking websites and special interest sites such as people using a particular type of personal computer or sharing an interest or a particular hobby or political issue provide bulletin boards. Some of the major Internet service providers also provide facilities for groups to set up their own bulletin boards and other means of sharing information and communicating among each other. Many libraries provide bulletin boards within their library websites as a means of enabling their users to discuss ideas and share information. Some sites provide bulletin boards for LIS professionals and provide opportunities to share good practices, discuss hot topics or to gain support.

A wide range of software packages are available to enable the use of bulletin boards within websites one such software is PHPBB. Almost all the bulletin boards provide some of the following features:

- They provide a basic search facility, by topic, author and keyword.
- Tools to enable view bulletins in hierarchical format, popularly known as threaded and unthreaded view.
- Facilities to select and save the bulletins.
- Facility to indicate the read and unread messages.

Below are some screenshots of available web bulletin boards:

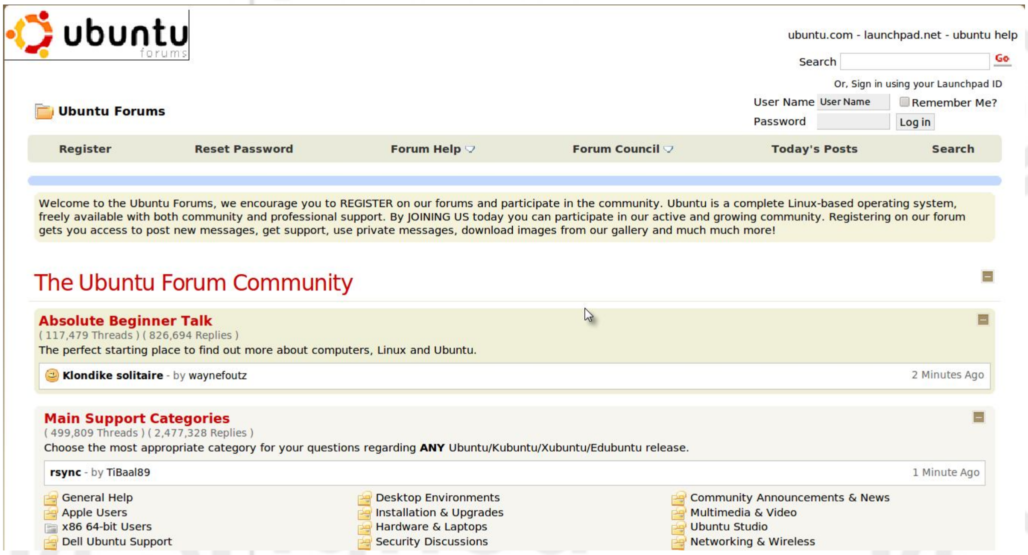


Fig.16. 4: Ubuntu Forums: Online bulletin board for Ubuntu users. (http://ubuntuforums.org/)

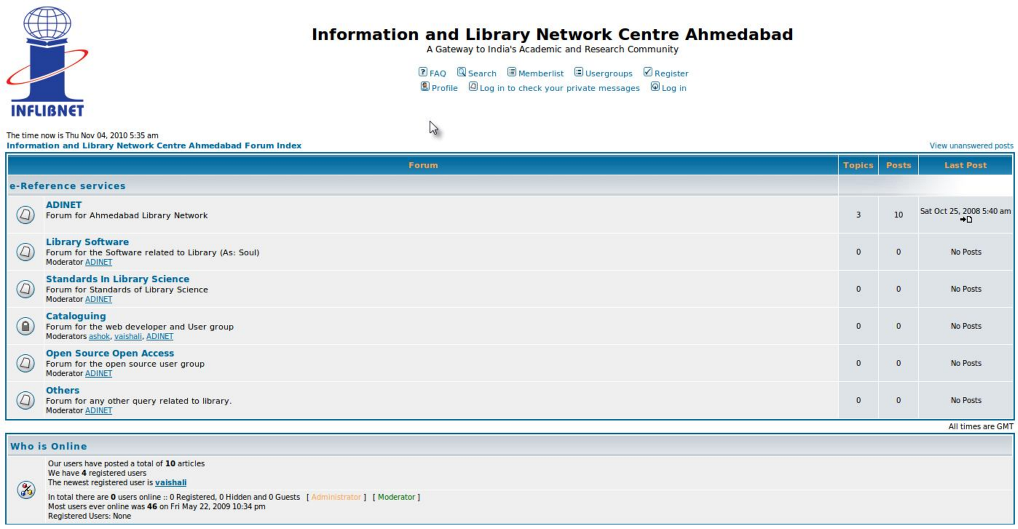


Fig. 16.5: INFLIBNET Bulletin Board (http://www.inflibnet.ac.in/forum/)

16.2.3 Mailing Lists and Discussion Lists

Mailing lists, discussion lists or listservs are services facilitate sending e-mails to a group of individuals with ease. These different names refer to the same process whereby one can send e-mail to a large group of people rather like using CC functionality provided by various mail applications. They are usually fully or partially automated through the use of software such as GNU's Mailman, Listserv, Mailbase etc. This service is hosted on hosting server that provides a reflector address on the same server capable of receiving email. The hosting service also maintains a list of all the different Mailing lists and the people who subscribe to the lists. The software processes incoming messages sent to the reflector address and depending on the content of the messages they are acted upon internally (in case of messages containing commands directed at the software itself) or are distributed to all e-mail addresses subscribed to the mailing list (See fig. 16.6). Joining a mailing list is called "subscribing" and leaving a list is called "unsubscribing".

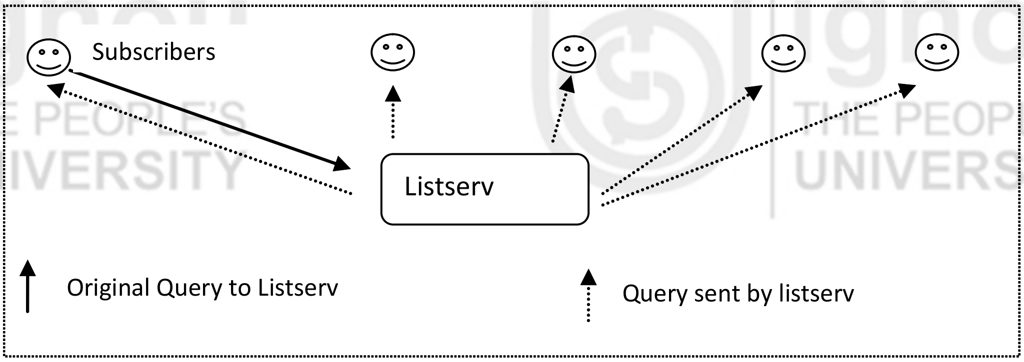


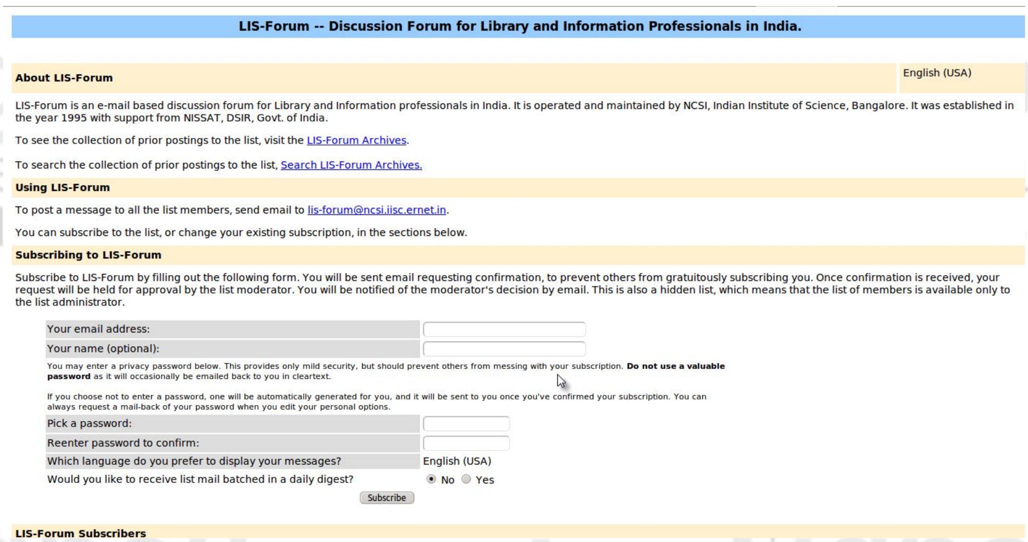
Fig. 16.6: Diagram Showing the Operation of a Mailing List/Listserv

There are thousands of mailing lists available on Internet; these are devoted to a varied range of topics and individuals. The popular open source developments and W3C standards first open a public mailing list to have ideas and to discuss on the standard in hand and after the discussion the changed are made to the standard. The best use of mailing list is for discussion, some libraries use mailing list to alert the users about new arrivals or sending the table of contents of new journal issues. In general mailing lists provide a forum to:

- Get advices on buying new systems.
- Request any factual information from the group.
- Support the group members in solving problems.
- Announce upcoming meetings and conferences.
- Inform about new vacancies.
- Discuss some prevailing issues faced by the profession.
- Information about new websites, productions and publications.

Some mailing lists have the functionality to send mails in two modes viz. digest and mail-as-arrives. The digest mode is good for heavily active lists. Digest is a consolidated mail containing string of mails sent by mailing list. This helps users' who wants to look all the mails in one stretch and does not want to be disturbed regular receipt of mail. Users' can fix the frequency of digest mail once in day or two or the way they want. Otherwise users' receive mails regularly as the mail is shot by mailing list. Some examples are:

1) **LIS Forum:** largest mailing list comprising mostly library and information science professional from India. It is operated and maintained by NCSI, Indian Institute of Science, Bangalore. (http://ncsi.iisc.ernet.in/mailman/listinfo/lis-forum)



LIS-Forum -- Discussion Forum for Library and Information Professionals in India.

About LIS-Forum English (USA)

LIS-Forum is an e-mail based discussion forum for Library and Information professionals in India. It is operated and maintained by NCSI, Indian Institute of Science, Bangalore. It was established in the year 1995 with support from NISSAT, DSIR, Govt. of India.

To see the collection of prior postings to the list, visit the LIS-Forum Archives.

To search the collection of prior postings to the list, Search LIS-Forum Archives.

Using LIS-Forum

To post a message to all the list members, send email to lis-forum@ncsi.iisc.ernet.in.

You can subscribe to the list, or change your existing subscription, in the sections below.

Subscribing to LIS-Forum

Subscribe to LIS-Forum by filling out the following form. You will be sent email requesting confirmation, to prevent others from gratuitously subscribing you. Once confirmation is received, your request will be held for approval by the list moderator. You will be notified of the moderator's decision by email. This is also a hidden list, which means that the list of members is available only to the list administrator.

Your email address:

Your name (optional):

You may enter a privacy password below. This provides only mild security, but should prevent others from messing with your subscription. **Do not use a valuable password** as it will occasionally be emailed back to you in cleartext.

If you choose not to enter a password, one will be automatically generated for you, and it will be sent to you once you've confirmed your subscription. You can always request a mail-back of your password when you edit your personal options.

Pick a password:

Reenter password to confirm:

Which language do you prefer to display your messages? English (USA)

Would you like to receive list mail batched in a daily digest? No Yes

Subscribe

LIS-Forum Subscribers

Fig. 16.7: Mailing List of NCSI, Indian Institute of Science, Bangalore. (http://ncsi.iisc.ernet.in/mailman/listinfo/lis-forum)

2) **DLRG:** Digital Library Research Group is a mailing list for helping library professional in solving software and hardware related problems. This is operated and maintained by Documentation Research and Training Centre (DRTC), Bangalore. It mainly covers topics related to digital library, library management systems and other data import and backup.

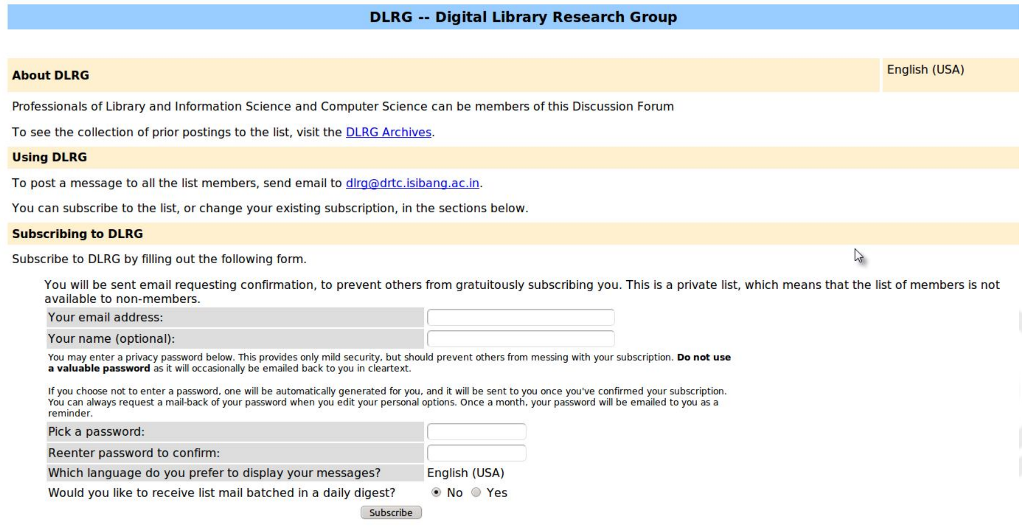


Fig. 16.8: Digital Library Research Group of DRTC, Bangalore

16.2.4 Resource Sharing

Resource sharing is a partnership where several libraries share one or more of their functions, for example, acquisitions, processing, storage and delivery of services. Each member has something useful to share, is willing to share and a plan exists to accomplish this. Major goal of resource sharing is to augment the local holdings by providing access to collections of other libraries.

One of the major approaches of resource sharing is sharing of library catalogues. For this purpose, libraries use Z39.50 protocol. In this environment, different libraries can be searched with one single interface. This interface can also be used to download the library catalogue entry so that duplication of work can be avoided among the libraries.

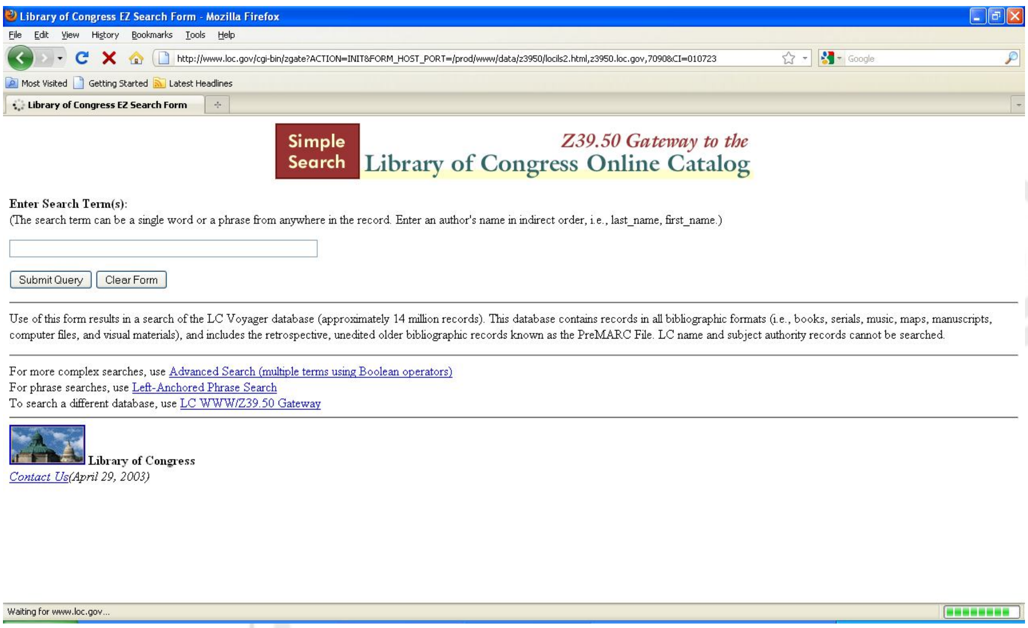


Fig. 16.9: Library of Congress Z39.50 Search

16.2.5 Online Document Repositories

Document repository systems are digital libraries containing an organisation's documents. They are commonplace for all the important documents. These systems allow controlling how documents are created, accessed, stored, and even disposed of. These may look similar to a file server but they have many differences such as. It allows creating templates for a particular class of documents for example, one can create a template for user requests wherein the user or staff may feed data wherever required and in the case of standard information the data will be prefilled. This ensures the consistency in data entry and ease in data entry.

Wherever there is sensitivity of information contained in the documents, the repository will need to be secured. For this the technology of encryption and authentication mechanisms is to be used. The online document repositories are popular in corporate world. In big organisations the employees have their username and password to access the system.

Some of the institutions and university libraries are also using document repositories to store and provide access to their research output.

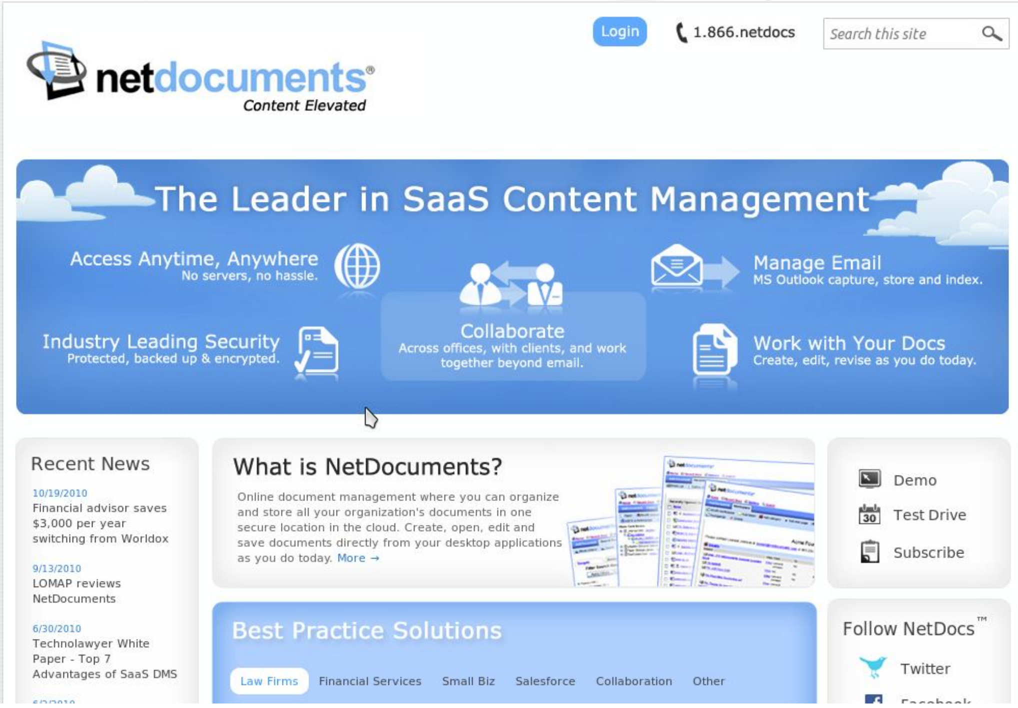


Fig. 16.10: Netdocuments (http://www.netdocuments.com/)

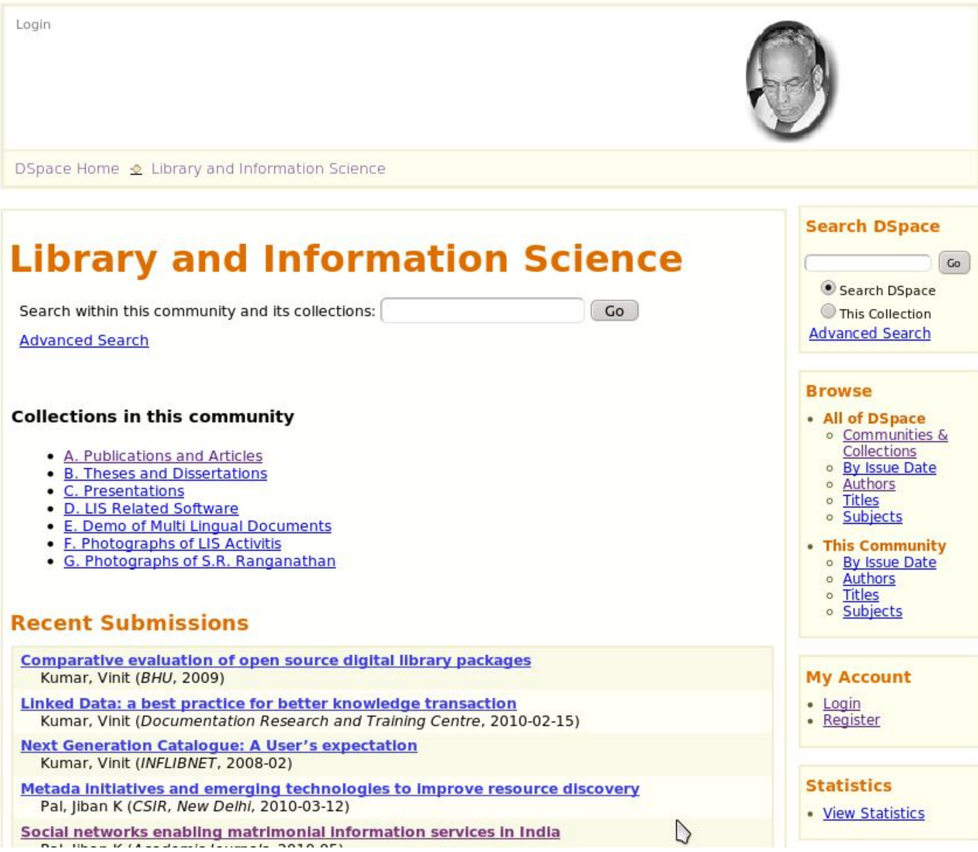


Fig. 16.11:Librarians Digital Library (LDL)(https://drtc.isibang.ac.in)

16.2.6 Web Portals

Web portals are a gateway to the information services on the Web or other sites on the Web. These act as a convenient location of sites of related interest, as it is seen in subject specific portals like infolibrarian.com or in general portals like yahoo.com. These portals provide other services such as e-mail, news, stock prices, information, databases and entertainment.

A truly effective portal must include:

- A single point of access (Single sign on)
- Unified search across all information sources
- Personalisation
- Applications integration
- Collaboration
- System security
- Openness
- Links to help files

Fig. 16.12: Yahoo Web Portal (www.yahoo.co.in)

There are subject specific web portals. These portals include documents and resources within the scope of subject. An example, in library science is infolibrarian.com.

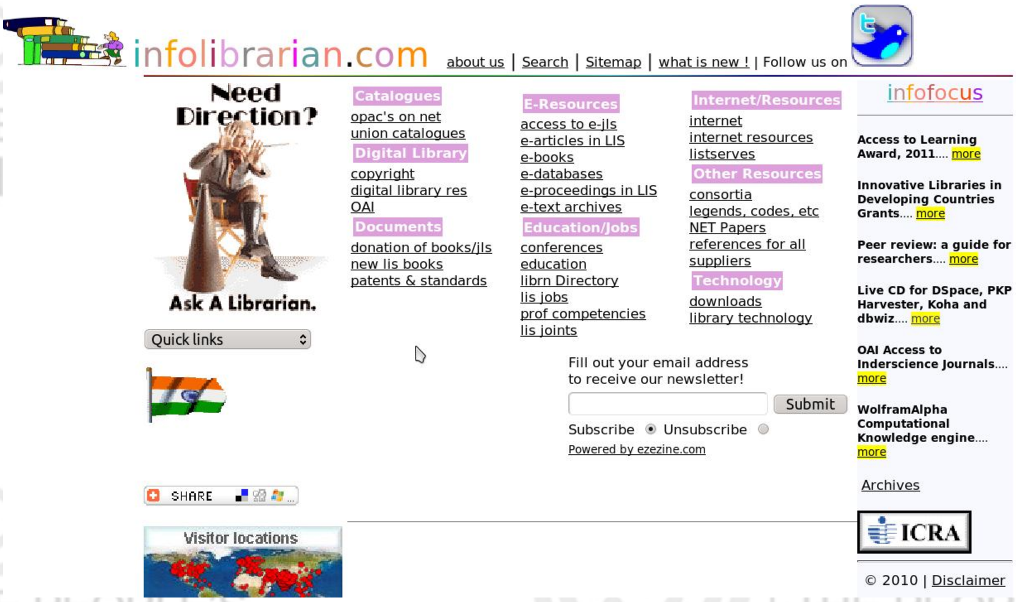


Fig. 16.13: Infolibrarian.com (http://www.infolibrarian.com/)

16.2.7 E-mail

E-mails are very common and popular method of exchanging information. E-mails are commonly and regularly used for both formal and informal communication. It is also used as a popular means of keeping up-to-date and solving small queries that arise everyday in the workplace.

International Encyclopedia of Information and Library Science defines e-mail as "a method of sending messages, data files, etc. by electronic means from one computer with network access to another". The receiving server or machine is usually equipped with a storage area, or mailbox, in which the messages are deposited. The access to mailbox is restricted by password and only the authenticated users having the password can read the mailbox. Users can read their incoming messages on-screen when they choose and, if they wish, print them out or download them on to a disk. In the beginning, the e-mails were delivered, as user to user which required both the users to be online. The present day e-mail servers provide users with the capabilities to store and forward the messages, which mean that the need to be online is not required. Further the users can login to the server and can check the messages at any time. The messages are exchanged between the users through a protocol known as Simple Mail Transfer Protocol (SMTP) with software programmes called Mail Transfer Agents (MTA) such as Exim4, Postfix and Sendmail. The MTAs provide user to retrieve the messages through standard protocols viz. Post Offices Protocol (POP) or Internet Message Access Protocol (IMAP) some corporate organisation have some proprietary standards such as Lotus Notes and MS Exchange server. The messages are written using clients which are desktop based applications such as Thunderbird, MS Outlook, Apple Mail, etc. or web based such as Webmail etc.

Some of the existing services and clients on Internet:

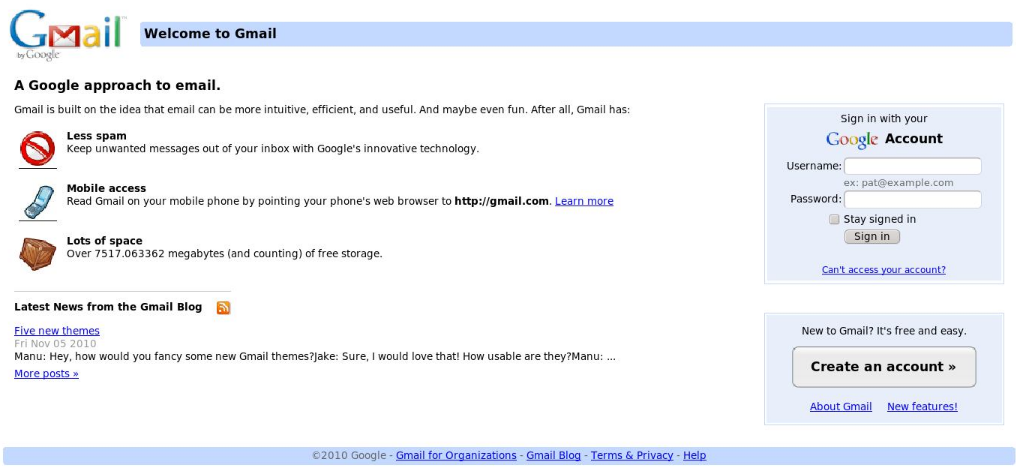


Fig. 16.14: Gmail - Google's Mail Service: (http://mail.google.com)

Fig. 16.15: Organisational Mail Service

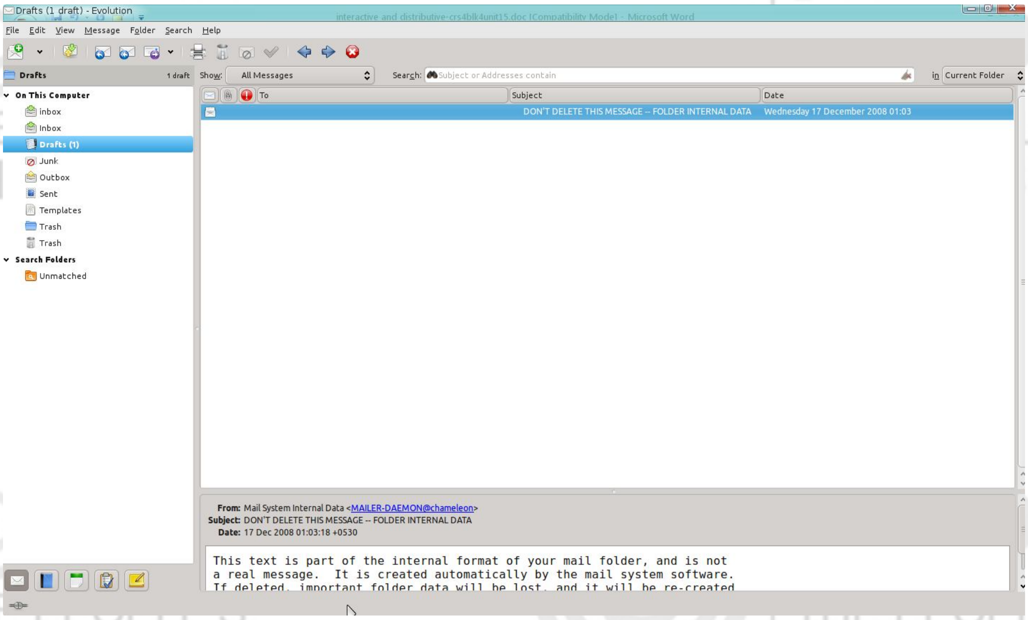


Fig.16.16: Evolution – Default Mail Client in Ubuntu Linux Distribution

The email services have some advantages and disadvantages too. The advantages are:

- The speed for delivery of messages is quick and almost instantaneous.
- They are reliable and secure efforts are going on to make them more reliable and secure.
- They are environment friendly, as there is no use of papers. Similarly the messages are stored permanently so it is easy to find the old messages.
- The use of graphics such as a picture etc. adds value to the service. Because of its one-to-many nature.
- It can be used as publicity and advertising tools as well.

Some disadvantages are also associated with email service such as:

- Because of its electronic nature and the facility to send attachments there are chances for viruses to get distributed through it.
- There is a big problem of spamming, which is nothing but sending unsolicited emails or advertisements. Checking and deleting these unsolicited emails can unnecessarily consume a lot of time of users.
- As the system is password based and contains lot of private data it gets more prone to security threats.

16.2.8 Online Storage and Searching

The electronic systems are very prone to mishandling and chances of getting corrupt due to myriad of reasons such as virus attack, natural disaster, theft, lost or broken, etc. Some online services on Internet provide the capability to store the files for easy retrieval and backup. On top of it they provide features like:

- Access files from any computer and anywhere, which reduces the need to carry them physically or sending them to own e-mail address.
- Some services allow viewing and listening media like pictures and audio files.
- Provide easy way to share the uploaded files among friends, family or the world.
- Remote access to the files, via a desktop client or website.

These services are also encrypted and properly authenticating looking at the sensitivity of the data. These services are mainly priced, but for small storage space they are free. The clients are seamlessly integrated to the operating system of the user. Some services (like, dropbox.com) dedicate a folder in the users' computer and allow dragging and dropping any file in the folder automatically synchronous to the online server. Similarly, some services provide programs to schedule the time of online syncing. Some of the online file storage services are listed in the following table.

Table 16.2: Online Storage and Searching Services

Service Name	URL	Priced/free
Dropbox	www.dropbox.com	Free till 8GB
SugarSync	https://www.sugarsync.com/	30 Day trial
OpenDrive	http://www.opendrive.com/	Free 5GB
Mozy	http://mozy.com/	Free till 2GB
Box.net	http://www.box.net/home	Free till 5GB

Apart from these exclusive storage services we have services like Google Drive (https://drive.google.com/ý) and Microsoft SkyDrive (https://skydrive.live.com/ý) wherein one can use existing e-mail accounts to use the storage facility. These services are also available on mobile devices like smart phones and tablets.

16.2.9 E-publishing

The information produced is disseminated by different model of publishing. New technologies have transformed the process of publishing and distribution of information. Electronic publishing is the process for production of typeset quality documents containing text, graphics, pictures, tables, equations etc. in digitised form. It uses new technology allowing publishers to deliver documents and other contents quickly and efficiently as well.

There are two important modes of e-publishing i.e., online and offline. Online publishing is in the form of online journals, websites, online database, e-books and so on. Offline publishing is done over some storage media like, Compact Disc, DVD and so on. Now-a-days e-books are being published on DVD and CDROM.

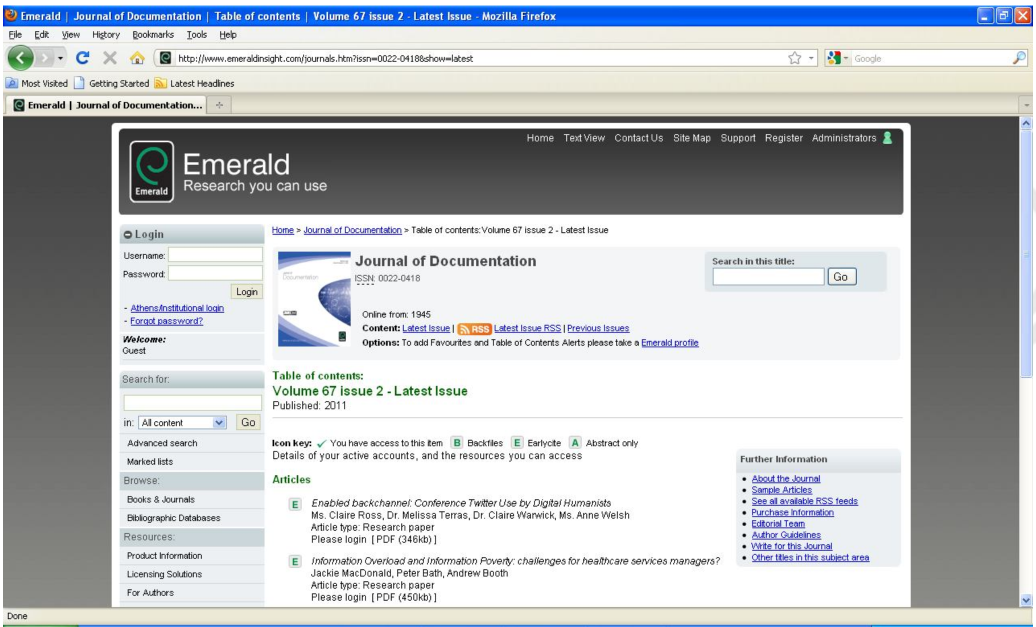


Fig. 16.17: Emerald Online Publishing

16.2.10 Webcasting

Webcasting is a service based on push technology, which means updating the systems by which users of the Internet, or Intranets, can receive news and other information through periodic and unobtrusive transmissions. The webcasting is an easy way to distribute the relevant content to the users. A webcast is a media file distributed over the Internet using streaming media technology to distribute a single content source to many simultaneous listeners/viewers. Webcasting is the Web-enabled broadcasting and integration of dynamic rich media. In simpler terms webcasting is "broadcasting" over the Internet. This involves one-to-many communication type. Webcasting informs users about material relevant or related to their interests before any specific request is made. This is done on the basis of pre-selected topics stored in a user profile. The user profile is further updated on the basis of the received feedback. The websites selected are referred to as channels; the updates received from the channels are supplied in its raw form or sometimes it is sent after processing.

With the use of very cheap and accessible technology it has become very easy to webcast anything such as:

- Meetings
- Special events: Birthday celebrations, festivals etc.
- Collaborative engineering
- Conferences with key partners
- Discussions with supply chain partners
- Sales presentations and demonstrations
- Voting, panel discussions
- Interviews
- Keynote addresses
- Training

There are many instances of webcasting available now-a-days with topics covering technology, computers, news, cooking, tutorials etc. The news webcasting is very popular these days. Webcasting is sometimes referred as "buffering media" too. Some examples of webcasting are:

1) **YouTube**

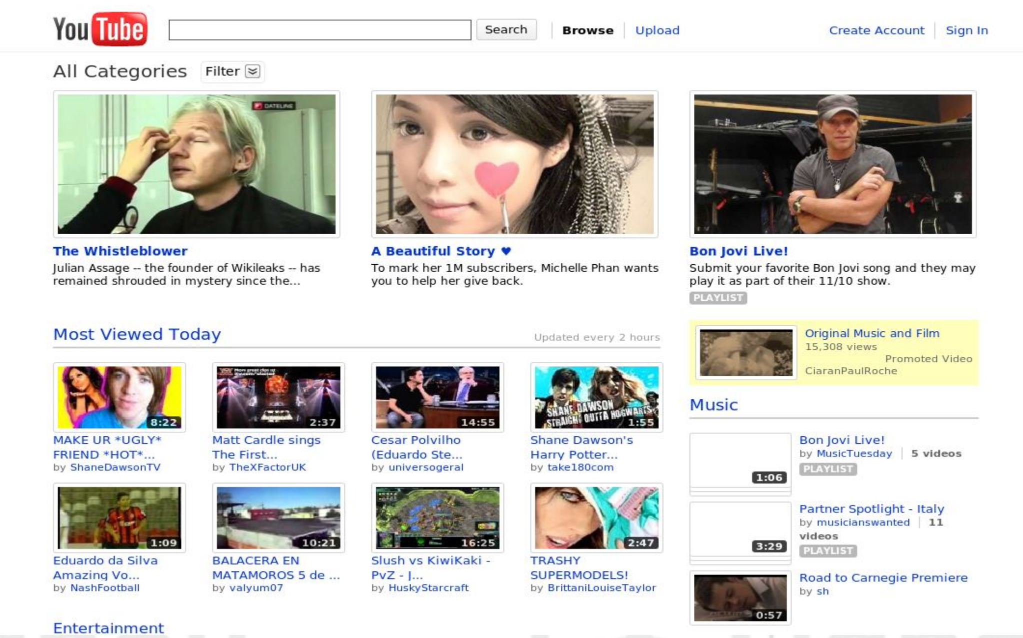


Fig. 16.18: Webcasting through YouTube(http://www.youtube.com/)

2) **BBC World News**

Fig. 16.19: News Casting(http://www.bbc.co.uk/iplayer/console/bbc_world_service)

3) **IGNOU Gyan Darshan**

Fig. 16.20: IGNOU Gyan Darshan Channel (http://www.ignouonline.ac.in/Broadcast/)

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

1) What is Web Directory?

2) Discuss advantages of Email communication.

....

....

....

....

....

....

....

....

16.3 INTERACTIVE DISTRIBUTED SERVICES

With the implementation of Web 2.0 technology the services are not only distributed but they are interactive also. They are more users' centric or customer oriented. The implementations of interactive services are manifold. Some of these are discussed in the subsequent sections.

16.3.1 Interactive Learning

Networked technologies such as Internet and World Wide Web are dramatically changing education and training as they enable people to access information and communicate with others across terrestrial boundaries, cultures and on a global scale. Interactive learning as defined by Chartered Institute of Personnel and Development emphasises the importance of connectivity and interactivity states "Learning that is delivered, enabled or mediated by electronic technology, for the explicit purpose of training in organisations. It does not include stand-alone technology-based training such as the use of CD-ROMs in isolation". Use of interactive learning involves text, graphics, audio, video and animation. Apart from this, the programmes are enhanced by providing additional support, e.g. using synchronous and asynchronous communication applications such as e-mail, discussion groups, chat rooms and video conferencing. The learning usually takes place through the use of web-based training programmes, where the learner typically follows a pre-specified learning process that includes opportunities for practice and assessment and feedback activities. It also takes place through blended approaches that involve learners experiencing a mixture of face-to-face and online learning experience.

There are online learning environments both on commercial and open source platforms. Moodle is one example of hosting online e-learning programmes.

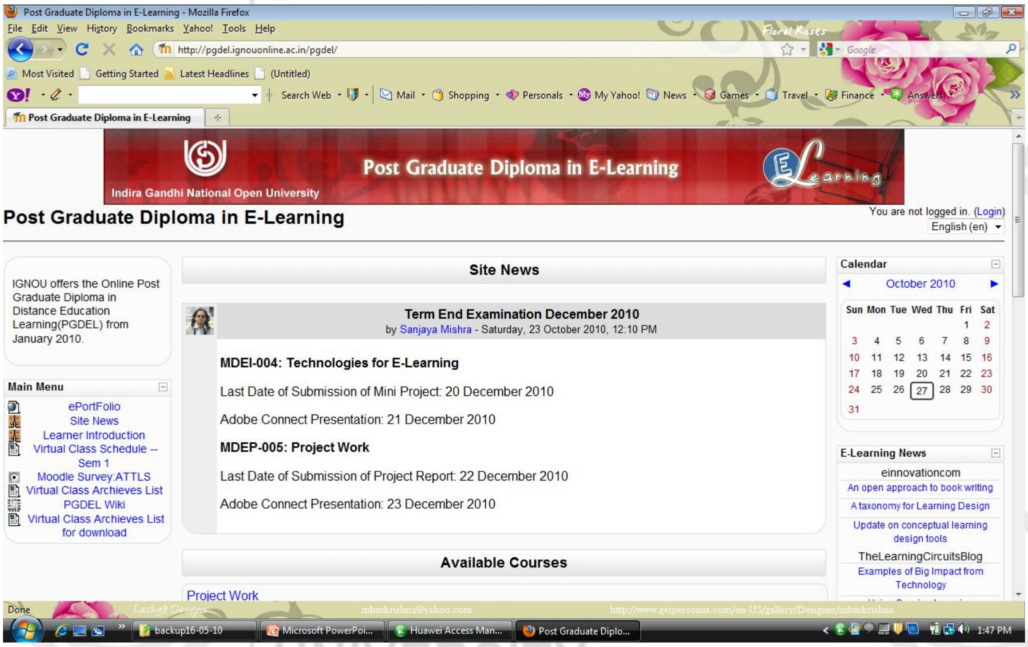


Fig. 16.21: Elearning Programme Using Moodle LMS

16.3.2 Interactive Business and Trading

One of the major implementation of existing web technology is seen over online trading and business. With the evolving standards and inter-operability among systems business houses are extending their sale counter over Internet. A user can order the products and make the payments online. An important interactive service over such online shops is use of cart facility where customer can collect the products s/he wants to buy and at the end make the payment. These carts are interactive tools where one can add or remove the product as one does in real stores.

The interactive environment of business has been extended to booking tickets online with selecting the best available option to travel. Online trading of live stocks in share market is another area which has been revolutionised. The live trading of stocks can be done. The trading window displays live prices of stocks every second. Same kind of trading services are also offered by many financial banks online where a client can buy and sell stocks and immediately money is deducted and deposited in the account as the case may be.

Fig.16.22: Online Stock Trading Screen

(Source: http://www.aaasec.com.my/stocksen/quicktour3.asp)

16.3.3 Remote Computing and File Transfer

Remote computing is a regular phenomenon in networked environment. It is used to share expensive hardware and software among a group of people. For example, if one has to perform certain set of analysis on a certain set of data. S/he can transfer the file using File Transfer Protocol (FTP) from present machine to remote machine where such analysis facility is available. Thereafter, whole processing can be done on remote machine using the resources from the remote machine. In such cases, the local machine acts as a dumb terminal.

With the increased use of Internet and implementation of new technologies, the form of remote computing has changed to Cloud Computing. In cloud computing the term cloud represents the use of Internet. The resources in cloud computing are distributed and often the user of computing facility is also not aware fully of location or the machine which is offering service. The cloud computing may use more than one resource distributed at different locations far apart.

Cloud computing is a way to reduce the cost on expensive computing exercises. It offers a sustainable environment for users to use and share their costly devices with others. One of the major implementation of cloud computing is Beowulf cluster. These are set of inexpensive computers, which are collectively used to process data as high-performance parallel computing device. These are shared machines working parallel over network.

16.3.4 Interactive Communication

Interactive communication is an important factor for the innovation and collaboration in present ideas hunting world. There are some prominent and powerful tools available to help to do so such as Microsoft Exchange (http://www.microsoft.com/exchange) server. It is a server based system and provides users with collaboration tools similar to Microsoft Outlook e-mail, calendaring application. Exchange provides shared calendars, shared address books, e-mail and other collaboration tools. The MS Exchange user typically accesses an Exchange account via a web interface or the Exchange desktop application.

There are tools, which are used for interactive real-time chat either through text or for conferencing. The participants can interact in real time to share their views and provide feedback or answer immediately. We will now discuss some of the live chatting and web conferencing tools.

Some of the tools for live chat are as follows:

1) Microsoft NetMeeting (http://www.microsoft.com/windows/netmeeting): It provides real-time chat, file transfer, interactive whiteboard and file sharing. NetMeeting also provides face-to-face video conferencing using webcam or other source and real-time audio communication.

2) Windows Messenger: Messenger allows for sending instant messages to other Windows users, audio, video, communication, file exchange, text based chat, and integration with 'remote assistance' to allow remote login to the local PC by a technical support service. Messenger is provided with most versions of Microsoft Windows.

3) ICQ (http://web.icq.com) This application provides similar functionality to Windows Messenger and may also be used via a web browser interface at the ICQ website. ICQ is a worldwide chat system, and should be used with care in any serious context.

Tele/video conferencing is the need of the day particularly on the business and trade. Facility of conducting online tele/video conferencing reduces the burden (financial and physical stress) for the business houses. This has come up as a major solution where long distance travel is required. It has become an interactive medium to communicate real time face to face. People traveling at far off place can keep in touch with their family and friends. Normally, computer-to-computer communication is free however service for calling a telephone is paid which is very nominal. This has revolutionised the global communication scenario.

The conferencing can be done one to one basis or among a group of nodes simultaneously. Some popular services are listed in Table 16.3.

Table 16.3: Conferencing Tools

Application Name	Features	URL
Pidgin	Same as Window Messenger	http://www.pidgin.im/
Skype	Similar features as NetMeeting, but more popular	http://www.skype.com/intl/en-us/home
Google+ Hangouts	Online chatting-audio, video and text chatting	http://www.google.com/hangouts/
Yahoo Messenger	Online chatting-audio, video and text chatting	http://messenger.yahoo.com

16.3.5 Interactive Search Agent and Document Delivery

Though it is in a very nascent stage but there are services, which are of interactive nature. Interactive search agents are normally domain specific. One of the implementation of domain specific search agents is seen in travel booking where a single search result yields many options. The user can choose which ever suits them and further refine the search and book for their travel.

Similarly, Google and Yahoo have launched a search service to make search more interactive and user oriented. The search interface helps users to find most suitable search term in order to form their search query. While typing the search term in the box all the related search terms are also displayed in anticipation to help user to find the search query.

The search engines provide facility to preview the searched pages within the search result page so that users can gaze the importance of web page towards their search and save their valuable time. Google has launched personalised searching alerts for any new addition by email. The service is known as Google Alert. Search engines do provide option to search within the blog, webpages, or statistics and so on.

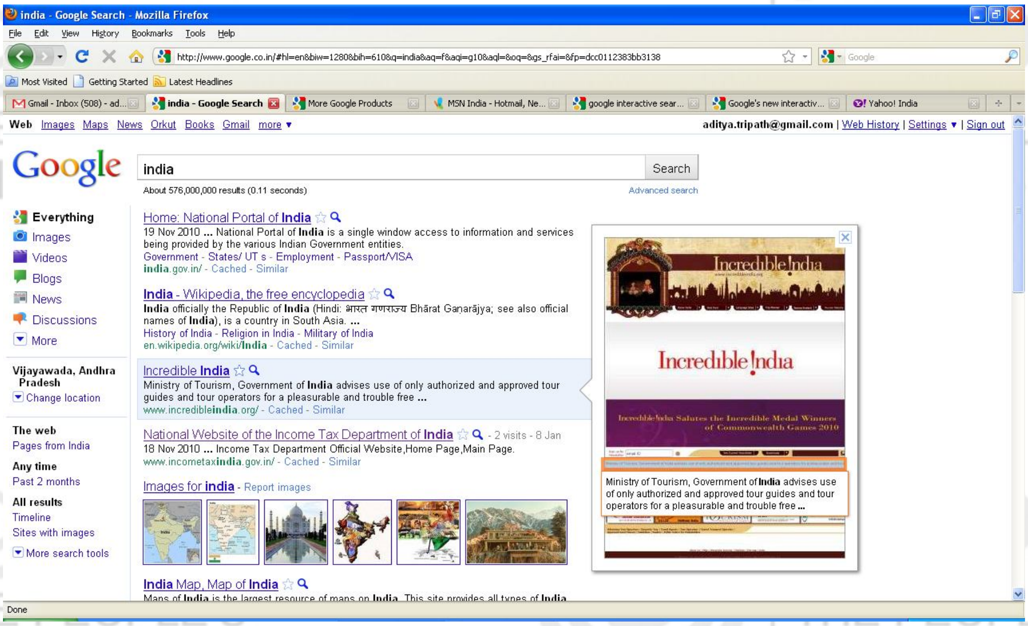


Fig. 16.23: Interactive Web Searching by Google

16.3.6 Interactive Bookmarking

People keep working online and come across several resources. These resources are book marked for any future use. The online bookmarking services provide method to bookmark resources online and use wherever required. One of the examples of such bookmarking is Google Bookmarks. The bookmarks can be categorised online and labeled.

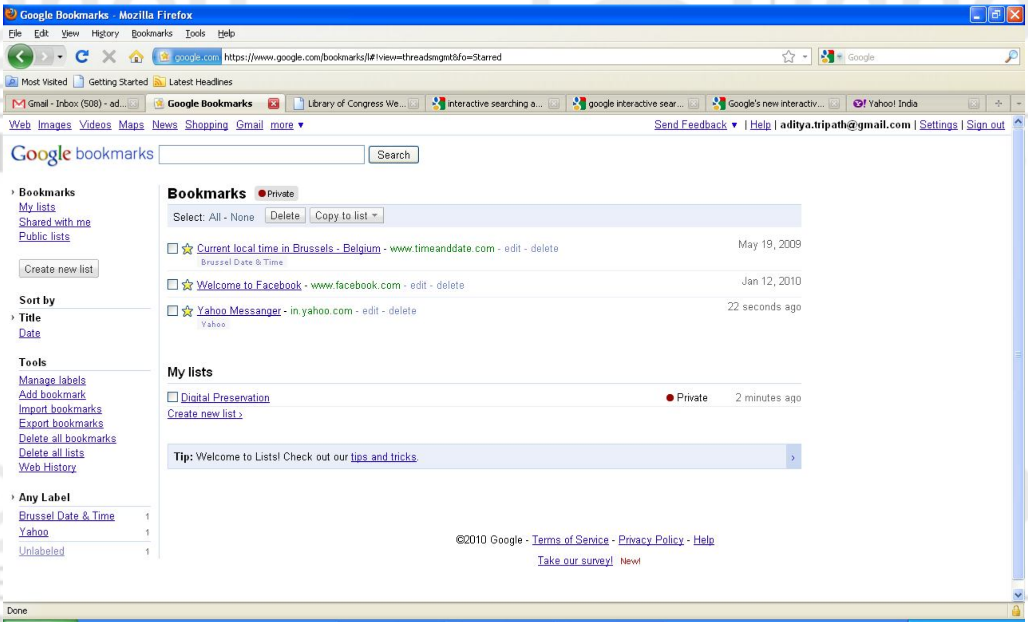


Fig. 16.24: Google Bookmark

16.3.7 Interactive Translation Service

There are several websites, which provide online translation service. However, the machine translation is not a matured technology but there are several implementations to it particularly in the web environment.

The tools like, Yahoo Babel Fish translates online given text and webpage. Google presents language tools for translation. It supports Indic language translation.

Fig. 16.25: Google Translation

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) What do you understand by Interactive Web service?

....

....

....

....

16.4 SECURITY AND PRIVACY ISSUES

Security is always an important issue in the design of any interactive and distributive service. Internet security means to prevent any unauthorised access to networked resources and services. Typical methods for ensuring security include:

1) **User authentication:** Integrating the service with a user records database (such as Active Directory) and an authentication system, such as LDAP (Lightweight Directory Access Protocol) should allow for use of the normal network username and password to access the services, restricting access to individuals without and institutional network account.

2) **IP:** The IP is a unique numerical network address on every computer. Access to the implemented services may also be limited using IP restriction, i.e., limiting the access to a range of known computers, probably located within an organisation or distributed over location.. For distributed locations the IPs of the machines should be known to the system running the IP restrictions.

3) **User roles:** The provision of having access control lists (ACLs) to define the roles of the registered users provides good level of security and reduces the chances of abuse. Some roles such 'system administrator' to manage the hardware and software at the highest level can be created. Similarly, service specific roles like 'moderator' etc. can be created for moderation of chats and discussions. Use of roles is an important security consideration, because higher level roles provide access to complex features, which if used improperly could damage the system or result in data loss (e.g. removal of user accounts).

Similarly there are privacy issues in communicating over a network, since all communications can be intercepted and recorded without participants being aware. The distributive services and social network present nowadays on web are more concerned with the privacy of the users. There are reports that the user information available in some of the services were mishandled and reached in the hands of marketing companies. These companies use the users data to propagate their products and to do business analysis etc. This further is breach of privacy of the user, as the offers offered by these companies will bug the users personal life. The Encyclopedia for Library and information Science defines privacy as "The quality of protection for aspects of the life, and information about the life, of an individual or a group, from the intervention or knowledge of others". The handling of privacy is under the eyes of media also, so utmost care should be taken. Such as:

1) **Privacy policy:** Every service asking users about their personal details should draft a privacy policy. Respect for the privacy of certain aspects of information is a counterbalance to the principle of freedom of information.

2) **Encryption:** Since all the communication in interactive and distributive technologies is transferred over a network these can be intercepted and recorded without the knowledge of the persons involved. Encryption technologies especially the public key encryption system should be used to keep the communications coded and readable only by the intended participants.

3) **Moderation policy:** When a system provides the interaction platform it may happen that some users may abuse the platform, especially in the case of discussion forums and chats. This is very important issue as the improper handling will not only refrain the fair users it will also make the environment nasty as well. To limit the abuse some moderation policy should be framed and implemented with some moderators. The steps to disallow the anonymous postings, banning of such kind of users etc. can be framed in the moderation policy. Similarly for chat, the sessions may usually be recorded providing a text record of all messages.

16.5 SUMMARY

In the present Unit we studied that Web has become a medium to deliver different types of services. The mode of many of the old services has changed to be more interactive and user oriented. Most of these services available are free and users can use them in the format they desire and as per their requirement.

We also learnt about different types of distributed services like bulletin board, forums, web portals, web casting and so on in this Unit.

The easy access to plethora of interactive web services has brought up several issues like data privacy and encryption, which has been discussed as well.

16.6 ANSWERS TO SELF-CHECK EXERCISES

1) Web directory is a topical list of Internet resources arranged in a hierarchical way. It is organised Web site listings created by human reviewers and act as a search tool created by human. It is quite different from search engine and does not display lists of web pages based on keywords; instead, it lists web sites by category and subcategory. The categorisation is usually based on the whole web site rather than one page or a set of keywords, and sites are often limited to inclusion in only a few categories.

2) Following are the advantages of E-mail

- The speed for delivery of messages is quick and almost instantaneous.
- They are reliable and secure and now efforts are on to make them more reliable and secure.
- They are environment friendly, as there is no use of papers. Similarly the messages are stored permanently so it is easy to find the old messages.
- The use of graphics such as a picture etc. adds value to the service. Because of its one-to-many nature.
- It can be used as publicity and advertising tools as well.

3) An interactive web service is online service given by a service provider or agent. In an interactive service user can provide input or feedback to the service provider. He can also modulate the mode of delivery of service or interface of service environment if he is permitted to do so.

16.7 KEYWORDS

Bookmark : A *bookmark* is a locally stored Uniform Resource Identifier (URI).

Database : A *database* is an application that stores and organises data for fast retrieval of information.

Digest : A style or format of distribution of electronic mailing lists in which multiple messages are placed together and distributed as a single unit.

E-publishing : Publishing of e-books and electronic articles over web to establish digital libraries or build organisational repositories.

Encryption : Coding of data for secured transmission.

Forum : An online group where visitors may read and post topics of common interest.

Mailing Lists : A mailing list is an electronic discussion forum that anyone can subscribe to. When someone sends an e-mail message to the mailing list, a copy of that message is broadcast to everyone who is subscribed to that mailing list.

MIME : **It stands for** ***M*****ultipurpose** ***I*****nternet** ***M*****ail** ***E*****xtensions.** *It is* a standard for formatting non-ASCII messages so that they can be communicated over the Internet.

Network : A group of two or more computer systems linked together.

Online document repositories : Document repository systems are digital libraries containing an organisation's documents.

Personalisation	:	User-specific tailoring of information.
Portal	:	These web sites serve as a single platform to access content and applications on a topic. It is also known as a "gateway," a "research guide," a "virtual library" or a "Web directory."
Query	:	A request for information from a database.
Remote Access	:	In web parlance, it is defined as access to a computer or a network from a remote distance.
Repository	:	A place where multiple databases or files are located for distribution over a network.
SMTP	:	Simple Mail Transfer Protocol is a set of standards to define the process of exchange e-mails between two servers.
Thread	:	A posting (message) on a discussion group or a mailing list and all of the responses to it builds a *thread*. To "follow a thread" is to read a series of messages sharing a common subject.
Thunderbird	:	Free, open source, cross-platform e-mail and news client based on Mozilla code.
Web Directory	:	A search tool created by editors or trained researchers who categorise or classify Web sites by subject.
Webcasting	:	A web-enabled broadcasting of information to the users.

16.8 REFERENCES AND FURTHER READING

1. Allan, Barbara. *E-learning and teaching in library and information services*. Facet publishing: London, 2002. Print

2. Burchard, Paul. *Towards Standards for an Interactive Web*. Available at <http://www.geom.uiuc.edu/hypernews/get/interactive/index.html>

3. Casey, Michael E. and Savastinuk, Laura C. Library 2. 0: A Guide to Participatory Library Service. *Information Today*: New Jersey, (2007), 174 p.

4. Catherall, Paul. *Delivering E-learning for information services in Higher Education*. Chandos Publishing: Oxford, 2005. Print

5. Chad, K and Miller, P. *Do Libraries Matter: The rise of Library 2.0*. 2005 <http://www.talis.com/applications/downloads/white_papers/DoLibrariesMatter.pdf>

6. Chapanis, Alphonse. Interactive Human Communication. *Scientific American*, 232 (1975) p.36-47. < http://web.media.mit.edu/~geek/class/ChapanisSciAm.pdf >

7. Cohen, Laura B. (ed.). *Library 2.0 initiatives in academic libraries*. Association of College & Research Libraries. (2007) 167 p.

8. Crichlow, Joel M. *Distributed Systems: Computing over Networks*. Haryana: PHI Learning, 2009. Print

9. da Silva, Alberto Manuel Rodrigues. Tutorial on web services. <http://cita2003.fing.edu.uy/articulosvf/WS1.pdf>

10. Edwards, Jim. Web Casting - The Future of Online Education. <http://www.teach-nology.com/tutorials/webcast/>

11. Holmberg, Kim, et.al. What is Library 2.0?. *Journal of Documentation*, 65.4 (2009) pp.668 - 681

12. Interactive Learning Opportunities On The Internet <http://www.techteachconcepts.com/interactive_learning.htm>

13. Internet Relay Chat (IRC), < *http://www.livinginternet.com/r/ri_early.htm*>

14. John-Okeke, Rita. Web-based legal information services and academic law libraries in Nigeria. *Library Hi Tech News*, 25. 9 (2008) pp.12 - 16

15. Khosrow-Pour, Mehdi (ed.). *Advanced topics in electronic commerce*. Idea Group, 2005. Print

16. Kidd, Terry T. and Chen, Irene. *Wired for learning: an educator's guide to web 2.0.* Information Age Publishing, 2009. Print

17. Kovacs, Diane K. "Electronic publishing in libraries: introduction", *Library Hi Tech*, 17.1(1999), pp.8 - 9

18. Librarians and the Internet, Social Media, and Web 2.0 <http://www.library20.org/>

19. Mailing Lists, <http://www.livinginternet.com/l/lp.htm>

20. Maness , Jack M. "Library 2.0 Theory: Web 2.0 and Its Implications for Libraries" *Webology.* 3.2 (2006) <http://www.webology.ir/2006/v3n2/a25.html>

21. Open Source Document Repository, <http://www.vclcomponents.com/s/0__/open_source_document_repository/>

22. Salber, D., Privacy Issues in Computer-Mediated Communication. *Science-and-Engineering-Ethics*, (1996)p. 238-240.

23. Shou, Joe. A history of Web portals and their development in libraries. *Information Technology and Libraries*. (2003) <http://www.accessmylibrary.com/article-1G1-109847209/history-web-portals-and.html>

24. Steinbrenner, Karin "PORTALS", Evolving Technology Committee. (2001) <http://www.humane.eu/fileadmin/wsan_docs/ws2004/ws2004_Fri_Veling_1_Educause.pdf>