CACS101: Computer Fundamentals and Applications

# Unit 6 – Internet and WWW

SUSAN SUNUWAR

# Internet

- The Internet is a global system of interconnected computer networks that use the standard Internet protocol suite (often called TCP/IP, although not all applications use TCP) to serve billions of users worldwide.
- It is a network of networks that consists of millions of private, public, academic, business, and government networks, of local to global scope, that are linked by a broad array of electronic, wireless and optical networking technologies.

# Internet

- The Internet carries an extensive range of information resources and services, such as the inter-linked hypertext documents of the World Wide Web (WWW) and the infrastructure to support email.
- Internet is a short form of the technical term internetwork, the result of interconnecting computer networks with special gateways or routers.

# Internet

- The Internet is also often referred to as the Net. The terms Internet and World Wide Web are often used in everyday speech without much distinction. However, the Internet and the World Wide Web are not one and the same.
- The Internet establishes a global data communications system between computers. In contrast, the Web is one of the services communicated via the Internet. It is a collection of interconnected documents and other resources, linked by hyperlinks and URLs.

# Its Applications

- E-Commerce (auction, buying, selling products etc.)
- Research (on-line journals, magazines, information etc.)
- Education (e-learning courses, virtual classroom, distance learning)
- E-Governance (online filing of application (Income Tax), on-line application forms etc.)
- On-line ticket booking (airplane tickets, rail tickets, cinema hall tickets etc.)
- On-line payments (credit card payments etc.)

# Its Applications

- Video conferencing
- Exchange of views, music, files, mails, folders, data, information etc.
- Outsourcing jobs (workflow software)
- Social networking (sites like facebook, linkedin, twitter)
- E-Telephony (sites like skype)

# Connecting to the Internet

- Requirements for connecting computer to the internet are as follows:
  1. TCP/IP enabled computer
  2. Internet service provider (ISP): an internet service provider provides you with a connection to the internet and the software you will need to navigate.
  3. Connection line: telephone line, coaxial cable or fiber connection
  4. Modem: a modem converts a digital signal received from a computer into an analogue signal that can be sent along ordinary telephone lines, and back to digital at the other end.
  5. Web browser: a web browser is software used to view and download Web pages and various types of files such as text, graphics and video. Examples are Microsoft Internet Explorer or Mozilla Firefox, Google Chrome.

# Client Server Technology

- The client-server model is a computing model that acts as distributed application which partitions tasks or workloads between the providers of a resource or service, called servers, and service requesters, called clients.
- Often clients and servers communicate over a computer network on separate hardware, but both client and server may reside in the same system.
- A server machine is a host that is running one or more server programs which share their resources with clients. A client does not share any of its resources but requests a server's content or service function.
- Clients therefore initiate communication sessions with servers which await incoming requests.

# Internet as client/server technology

- Internet is massive network of networks and world wide web is a system of interlinked hypertext documents accessed via the internet. To complete the flow of accessing information over the web there is need of client/server architecture.
- Client/server describes the flow of information between two computer programs in which one computer program (client) makes a service request to another computer program (the server), which provide service requested by the client. Clients rely on servers for required resources.
- It is network architecture in which each computer or process on the network is either a client or a server. Hence, we could describe internet as client/server technology.

# Internet as client/server technology

- The server component provides a function or service to one or many clients, which initiate requests for such services.
- Servers are classified by the services they provide. A web server serves web pages, and a file server serves computer files. A shared resource may be any of the server computer's software and electronic components, from programs and data to processors and storage devices.

# Electronic Mail

- Electronic mail is a method of creating, transmitting, or storing primarily text-based human communications with digital communications systems.
- An e-mail system allows computer users on a network to send text, graphics, sounds, and animated images to other users.
- Features of e-mail are:
  - E-mail can be sent to one person or more than one person at the same time.
  - It is location independent, i.e., user can access it anywhere.
  - It can be sent at any time of the day.
  - It is faster than postal mail as it is transmitted electronically.
  - A copy of e-mail message that the sender has sent is available on the sender's computer.

# Electronic Mail

- **Advantages**
  - Faster and cheaper than traditional mailing system
- **Disadvantages**
  - Server failure leads to whole network failure
  - Increases expenses, as dedicated hardware and special software adds to the cost.

# How E-mail works

- It works on client-server model.
- Both, the sender and recipient of e-mail are e-mail clients.
- E-mail server is a combination of processes running on a server with a large storage capacity.
- The e-mail client interacts with the e-mail server to send or receive e-mail. Most email servers provide email services by running two separate processes on the same machine – Post Office Protocol 3 (POP3) and Simple Mail Transfer Protocol (SMTP). Some e-mail servers also run another process on the machine – Internet Message Access Protocol (IMAP)

# Video Conferencing

- Video conferencing refers to live, visual connection between two or more people in which two or more sets of hardware and software interact while simultaneously transmitting and receiving video and audio signals from two or more geographic locations.
- It is a live connection between people in separate locations for the purpose of communication, involving audio and often text as well as video.
- It may involve sharing documents, various presentation materials, whiteboards.
- Reason for implementing video conferencing is to save travel time and costs.

# Internet Service Providers

- ISP, also called Internet access provider or IAP is a company that offers its customers access to the Internet.
- ISP connects to its customers using a data transmission technology appropriate for delivering Internet Protocol datagrams, such as dial-up, DSL or cable modem.

# Domain Name and Domain Name Server

- Domain names are named addresses provided to devices attached to the network. It is a way to identify and locate computer connected to the internet. Although IP addresses uniquely designate a host connected to the internet, they are not convenient to use as they are difficult to remember. No two organizations have the same domain name.
- A domain name usually has tow parts separated by periods called dot. E.g., google.com. The last portion of the domain name is the top-level domain name which describes the type of the organization.
- Each domain name corresponds to an IP address. The DNS server is responsible for translation of domain names to IP address.
- Domain names are case insensitive so com, COM means same.

# Advantages of Domain Names

- Domain names tend to be **memorable**.
- The domain name is unique.

# Internet Address

- An Internet Protocol (IP) address is a numerical identification (logical address) that is assigned to devices participating in a computer network utilizing the Internet Protocol for communication between its nodes. It is numbered address provided to computers and other devices attached to the internet to identify them uniquely.
- IP addresses are unique.
- All machines connected to the Internet agree to use the same scheme for establishing an address.

# IPV4

- In IPV4 an address consists of 32 bits which limits the address space to 4294967296 ($2^{32}$) possible unique addresses.
- IPV4 addresses are canonically represented in dot-decimal notation, which consists of four decimal numbers, each ranging from 0 to 255, separated by dots, e.g., 172.16.254.1

# IPV6

- An Internet Protocol Version 6 address (IPV6 address) is a numerical label that is used to identify a network interface of a computer or other network node participating in an IPV6 enabled computer network.
- IPV6 implements a new addressing system that allows for far more addresses to be assigned than with IPV4.
- It does not implement interoperability features with IPV4. Exchanging traffic between the two networks requires special translator gateways, but this is not generally required as most computer operating systems and software implement both protocols for transparent access to both networks, often known as dual stack support.
- IPV6 addresses have a size of 128 bits.
- e.g., 2001:0db8:85a3:0042:0000:8a2e:0370:7334

# Internet Protocols

- A protocol is the set of standard rules for data representation, signaling, authentication and error detection required to send information over a communication channel.
- Protocols for digital computer network communication have features intended to ensure reliable interchange of data over an imperfect communication channel.

# Transmission Control Protocol / Internet Protocol (TCP/IP)

- TCP/IP is the protocol used on the Internet.
- The IP does the following:
  - Envelops and addresses the data.
  - Enables the network to read the envelope and forward the data to its destination.
  - Defines how much data can fit in a single envelope (a packet)
- The TCP does the following:
  - Breaks data into packets that the network can handle efficiently
  - Verifies whether all the packets have arrived at their destination
  - “Reassembles” the data
- The IP protocol deals only with packets whereas TCP enables two hosts to establish a connection and exchange streams of data. TCP guarantees delivery of data and guarantees that packets will be delivered in the same order in which they were sent.

# Hypertext Transfer Protocol (HTTP)

- It is a TCP/IP based communication protocol which is used to deliver virtually all files and other data, collectively called resources, on the World Wide Web. These resources could be HTML files, image files, query results, or anything else.
- A browser works as an HTTP client because it sends requests to an HTTP server which is called web server. The web server then sends responses back to the client. The standard port for HTTP servers to listen on is 80 but it can be changed to any other port like 8080.
- HTTP is connectionless.
- HTTP is media independent.
- HTTP is stateless.

# File Transfer Protocol (FTP)

- File Transfer Protocol (FTP) is a network protocol used to transfer data from one computer to another through a network such as the Internet. It is used for exchanging and manipulating files over a TCP computer network.
- An FTP client may connect to an FTP server to manipulate files on that server.
- FTP works on the client/server principal.
- FTP sites are typically used for uploading and downloading files to a central server computer, for the sake of file distribution.
- A special FTP software is used to connect to download and upload files to an FTP site.
- Files can be transferred and stored on computers called FTP servers. Files on FTP servers are compressed to decrease the size and make file transfer time shorter.

# Simple Mail Transfer Protocol (SMTP)

- It is a protocol for sending e-mail messages between servers. Most of the e-mail systems that send mail over the Internet use SMTP to send messages from one server to another.

# Post Office Protocol (POP)

- It is designed to allow single user computers to retrieve electronic mail from a POP server.
- The POP server might be a computer with the permanent Internet connection whereas its clients might only connect to it occasionally, e.g., by modem.

# Terminal Network (Telnet)

- Telnet is a protocol that enables one computer to connect to another computer and such control is referred to as remote login.
- The user's computer which initiates the connection is referred to as the local computer and the machine being connected to, which accepts the connection, is referred to as the remote, host computer.
- Once connected, the user has full control over the remote host during the telnet session.

# Gopher

- It is a protocol designed to search, retrieve and display documents from remote sites on the Internet.
- It is also possible to initiate online-connections with other systems via Gopher.
- Information accessible via Gopher is stored in Gopher servers. These servers do not just contain files, directories and searchable databases but also reference to other servers.
- Gopher was created as a piece of software to utilize some of the services that were becoming available on the Internet.

# Wide Area Information Service (WAIS)

- It is an internet search tool based on Z39.50 standard.
- This standard describes a protocol, for computer-to-computer information retrieval.
- A WAIS client program enables the user's computer to contact a WAIS server, submit a search query and receive a response to that query.
- WAIS has the capability of simultaneously searching in more than one database.

# Intranet

- An intranet is a private computer network that uses Internet technologies to securely share any part of an organization's information or operational systems with its employees.
- An intranet is built from the same concepts and technologies used for the Internet, such as client-server computing and TCP/IP. Internet protocols may be found in an Intranet, such as HTTP (web services), SMTP (e-mail), and FTP (file transfer).
- Intranet can be understood as a private version of the Internet confined to an organization.
- When Internet access is provided it is usually through a network gateway with a firewall, shielding the intranet from unauthorized external access. This gateway often also implements user authentication, encryption of messages, and often virtual private network (VPN) connectivity for off-site employees to access company information, computing resources and internal communications.

# Internet vs Intranet vs Extranet

| Basis of Comparison | Internet | Intranet | Extranet |
|---|---|---|---|
| Network Accessibility | Public | Private | Private |
| Availability | Global system. | Specific to an organization. | To share information with suppliers and vendors it makes the use of public network. |
| Coverage | All over the world. | Restricted area upto an organization. | Restricted area up to an organization and some of its stakeholders or so. |
| Content accessibility | It is accessible to everyone connected. | It is accessible only to the members of organization. | Accessible only to the members of organization and external members with logins. |

# Internet vs Intranet vs Extranet

| Basis of Comparison | Internet | Intranet | Extranet |
|---|---|---|---|
| No. of computers connected | It is largest in number of connected devices. | The minimal number of devices are connected. | The connected devices are comparable with Intranet. |
| Owner | No one. | Single organization. | Single/ Multiple organization. |
| Network purpose | Its purpose is to share information throughout the world. | Its purpose is to share information throughout the organization. | Its purpose is to share information between members and external, members. |
| Users | General public. | Employees of the organization. | Employees of the organization which are connected. |

# Advantages of Intranet

- **Workforce productivity**: Intranets can also help users to locate and view information faster and use applications relevant to their roles and responsibilities. Users can access data held in any database the organization wants to make available, anytime increasing employees' ability to perform their jobs faster, more accurately, and helps to improve the services provided to the users.
- **Time**: With intranets, organizations can make more information available to employees.
- **Communication**: Intranets can serve as powerful tools for communication within an organization. By providing information on the intranet, staffs can keep up-to-date with the strategic focus of the organization.

# Advantages of Intranet

- **Business operations and management**: Intranets are also being used as a platform for developing and deploying applications to support business operations and decisions across the internetworked enterprise.
- **Cost-effective**: Users can view information and data via web-browser rather than maintaining physical documents such as procedure manuals, internal phone list and requisition forms.
- **Promote common corporate culture**: Every user is viewing the same information within the Intranet.
- **Enhance Collaboration**: With information easily accessible by all authorized users, teamwork in enabled.

# Disadvantages of Intranet

- **Performance Limitations**: Some applications that have been well optimized for conventional systems create a heavy system workload while migrating them to an Internet platform.
- **Software compatibility problems**: It is an evolving technology that requires upgrades and could have software incompatibility problems.
- Security features can be inadequate.
- Inadequate system performance management and poor user support.
- Maintaining content can be time consuming.
- Some employees may not have PCs at their desks.

# World Wide Web(WWW) and its Evolution

- World Wide Web is a repository of information spread all over the world and linked together.
- Early web pages contained only text, but due to rapid advancements in technology, the web pages now contain images and other multimedia.

# Architecture of Web

- The WWW uses the concept of hypertext and hypermedia.
- In hypertext environment, information is stored in a set of documents that are linked using a concept of pointer(hyperlink).
- The hypertext documents contains text only.
- The hypermedia document can contain images, graphics, and sound.
- The web consists of a vast worldwide collection of documents or web pages called pages.
- Web browser is used to view the web pages.
- The browser fetches the page requested, interprets the text and formatting commands in it, and displays the page, properly formatted, on the screen.

# Uniform Resource Locator(URL)

- URL is a string of characters that identifies a particular internet resource.
- It specifies where an identified resource is available and the mechanism for retrieving it.
- It is also referred to as a Web address.
- URL represents a standardized addressing scheme for Internet resources and helps the users to locate these resources by indicating exactly where they are.
- URL's have three parts:
  - Protocol name
  - DNS name of machine on which the page is loaded.
  - Local name uniquely indicating the file name
  - e.g., http://www.abc.com/about.html

# Uniform Resource Locator(URL)

- The general form of URL is **protocol://address/path** where,
  - protocol defines the method used to access the web page, e.g., http, ftp
  - address is the internet address of the server where the web page resides. It contains the service (e.g., www) and the domain name (e.g., google.com), and
  - Path is the location of web page on the server.

# Browsers

- Web browser is the software, which is used to access the Internet and the WWW.
- It is used to access and view the web pages of the various websites available on the Internet and interact with text, images, videos, music, games and other information typically located on a web page at a web site on the World Wide Web or a local area network.
- It acts as an interface between the user and the inner workings of the internet like www.
- E.g., Mozilla Firefox, Google Chrome, Microsoft Internet Explorer, Opera, etc.

# Search Engine

- Search engine is a sophisticated piece of software, accessed through a page on a website that allows you to search the web by entering search queries into a search box.
- The search engine then attempts to match your search query with the content of web pages that is stored , or cached, and indexed on its powerful servers in advance of your search.
- The web search engine runs the search string against the database, returns a list of resources that match the criteria, and displays the results for the user.

# Web Servers

- Web Server refers to server software, or hardware dedicated to running said software that can serve contents to the World Wide Web.
- A web server processes incoming network requests over HTTP and several other related protocols. It is a program that uses HTTP to serve files that create web pages to users in response to their requests, which is sent by their computers HTTP connection.
- Every web server that connects to the Internet will have a unique IP address.
- Web server enables the hosting providers to manage multiple domains on a single server.

# Apache web server

- One of the most popular web server developed by Apache Software Foundation.
- Apache is an open-source software which supports almost all operating systems.
- It is designed to create web servers that can host one or more HTTP-based websites.
- It supports multiple programming languages, server-side scripting, an authentication mechanism and database support.
- Customization of Apache web server is easy.
- Modules can be added to the server as required.
- It is more stable.

# IIS web server

- Internet Information Services (IIS), formerly known as Internet Information Server, is a web server produced by Microsoft.
- This server has all the features just like Apache, but it is not an open source and adding personal modules is not easy.
- It works with all the windows operating system platforms.

# Proxy Server

- A proxy server verifies and forwards incoming client requests to other servers for further communication.
- A proxy server is located between a client and a server where it acts as an intermediary between the two, such as a web browser and a web server.
- Purposes of proxy server:
  - To provide internal system security
  - To speed up resource access
  - To apply access policies for tracking organizational Internet use or assessing employee progress.
  - To bypass special controls, such as parental or security controls
  - To scan for viruses and malware