CSE3002

Internet and Web Programming

Module I
INTRODUCTION

- Internet Overview- Networks - Web Protocols – Web Organization and Addressing - Web Browsers and Web Servers - Security and Vulnerability-Web System Architecture – URL - Domain Name – Client-side and server-side scripting.

Internet

- A network of networks, bringing together people, information, hardware and software around the World.

- The origin of Internet devised from the concept of **Advanced Research Project Agency Network (ARPANET).**
- **ARPANET** was developed by United States Department of Defense.
- Basic purpose of ARPANET was to provide communication among the various bodies of government.
- Initially, there were only four nodes, formally called **Hosts.**
- In 1972, the **ARPANET** spread over the globe with 23 nodes located at different countries and thus became known as **Internet.**

Early 80’s

- BITNET – Because Its Time Network
- CSNET – Computer Science Network

In 1986,

- NSFNET - National Science Foundation Network

To promote advanced research and education networking in the United States.

- It is the largest network in the world that connects hundreds of thousands of individual networks all over the world.
- To access the Internet, an existing network need to pay a registration fee and agree to certain standards based on the TCP/IP (Transmission Control Protocol/Internet Protocol) .

The uses of the Internet

- Send e-mail messages.
- Send (upload) or receive (down load) files between computers.
- Participate in discussion groups, such as mailing lists and newsgroups.
- Surfing the web.

Addressing

- Each computer on the internet does have a unique identification number, called an IP (Internet Protocol) address.
- The IP addressing system currently in use on the Internet uses a four-part number.
- Each part of the address is a number ranging from 0 to 255, and each part is separated from the previous part by period,
- For example, 106.29.242.17

IP Addressing

- The combination of the four IP address parts provides 4.2 billion possible addresses (256 x 256 x 256 x 256).
- This number seemed adequate until 1998.
- Members of various Internet task forces are working to develop an alternate addressing system that will accommodate the projected growth.
- However, all of their working solutions require extensive hardware and software changes throughout the Internet.

- New generation of the Internet Protocol was eventually named Internet Protocol Version 6 (IPv6) in 1995. The address size was increased from 32 to 128 bits, thus providing up to 2^{128} (approximately 3.403×10^{38}) addresses.
- Eg: 2001:db8:0:1234:0:567:8:1 (IPv6).

- The Internet Assigned Numbers Authority (IANA) manages the IP address space allocations globally and delegates five regional Internet registries (RIRs) to allocate IP address blocks to local Internet registries (Internet service providers) and other entities.

Domain Name Addressing

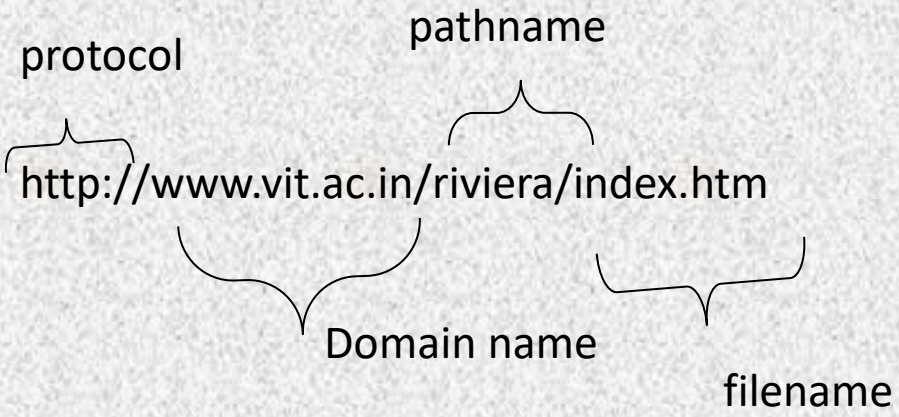
- Most web browsers do not use the IP address to locate Web sites and individual pages.
- They use domain name addressing.
- A **domain name** is a unique name associated with a specific IP address by a program that runs on an Internet host computer.
- This program, which coordinates the IP addresses and domain names for all computers attached to it, is called **DNS (Domain Name System) software**.
- The host computer that runs this software is called a **domain name server.**
- Domain names are formed by the rules and procedures of the **Domain Name System** (DNS)

Uniform Resource Locators

- The IP address and the domain name each identify a particular computer on the Internet.
- However, they do not indicate where a Web page's HTML document resides on that computer.
- To identify a Web pages exact location, Web browsers rely on Uniform Resource Locator (URL).
- URL is a four-part addressing scheme that tells the Web browser:

➢ What transfer protocol to use for transporting the file
➢ The domain name of the computer on which the file resides
➢ The pathname of the folder or directory on the computer on which the file resides
➢ The name of the file

Structure of a Uniform Resource Locators



http => Hypertext Transfer Protocol

What is Web?

- The **Web (World Wide Web)** consists of information organized into Web pages containing text and graphic images.
- It contains hypertext links, or highlighted keywords and images that lead to related information.
- A collection of linked Web pages that has a common theme or focus is called a **Web site**.
- The main page that all of the pages on a particular Web site are organized around and link back to is called the site's **home page**.

Web Protocols – HTTP – HTTPS- FTP- TCP/IP

- **Protocols** specify interactions between the communicating entities.
- A network protocol defines rules and conventions for communication between network devices.
- Network protocols include mechanisms for devices to identify and make connections with each other, as well as formatting rules that specify how data is packaged into messages sent and received.

- Modern protocols for computer networking all generally use packet switching techniques to send and receive messages in the form of packets - messages subdivided into pieces that are collected and re-assembled at their destination.
- Hundreds of different computer network protocols have been developed each designed for specific purposes and environments.

- Modern operating systems contain built-in software services that implement support for some network protocols.
- Applications like Web browsers contain software libraries that support the high level protocols necessary for that application to function.
- For some lower level TCP/IP and routing protocols, support is implemented in directly hardware (silicon chipsets) for improved performance.

- Each packet transmitted and received over a network contains binary data (ones and zeros that encode the contents of each message).
- Most protocols add a small *header* at the beginning of each packet to store information about the message's sender and its intended destination.
- Some protocols also add a *footer* at the end.
- Each network protocol has the ability to identify messages of its own kind and process the headers and footers as part of moving data among devices.

Web browser

- A web browser is a software application for retrieving, presenting and traversing information resources on the World Wide Web.
- An information resource is identified by a Uniform Resource Identifier (URI/URL) that may be a web page, image, video or other piece of content.

- The first web browser was invented in 1990 by Sir Tim Berners-Lee.
- Berners-Lee is the director of the World Wide Web Consortium and also the founder of the World Wide Web Foundation. His browser was called WorldWideWeb and later renamed Nexus.

- Nexus – 1990
- Erwise – web browser with a graphical user interface – 1992
- Mosaic – first popular browser – 1993
- Netscape Navigator – 1994
- Internet Explorer – 1995
- Opera – 1996
- Mozilla Foundation – 1998
- Apple’s Safari – 2003
- Chrome – 2008
- Microsoft Edge – 2016

Web server

- The primary function of a web server is to store, process and deliver web pages to clients.
- The communication between client and server takes place using the Hypertext Transfer Protocol (HTTP).
- Pages delivered are most frequently HTML documents, which may include images, style sheets and scripts in addition to text content.

- Web browser initiates communication by making a request for a specific resource using HTTP and the server responds with the content of that resource or an error message if unable to do so.

Internet Service Provider (ISP)

- A commercial organization with permanent connection to the Internet that sells temporary connections to subscribers.
- **Internet Service Provider (ISP)** is a company offering access to internet. They offer various services:
  - Internet Access
  - Domain name registration
  - Dial-up access
  - Leased line access

Connection Types

There exist several ways to connect to the internet. Following are these connection types available:

- Dial-up Connection
- Integrated Services Digital Network
- Digital Subscriber Line
- Cable TV Internet connections
- Satellite Internet connections
- Wireless Internet Connections