
Lecture 1

Contents: World Wide Web; Web server and Web client; Internet; Hypertext, hypermedia, hyperlink; Working principle of World Wide Web; Web server and clients; Domain naming; Internet service provider; Intranet; World Wide Web versus Internet; Exercises

1.1 World Wide Web

The World Wide Web (WWW) is also known as the Web. It is a system of interconnected public webpages accessible through the Internet. The Web is not the same as the Internet. The Web is one of many applications built on top of the Internet. The Web is an information system enabling documents.

Documents and downloadable media are accessed from web servers using networks with the help of programs such as web browsers. Both servers and resources on WWW are identified and located using character strings called uniform resource locators (URLs). The Web is a convenient way to disseminate information and making it available all the times. It is a powerful tool for mass communication, e-business and e-commerce.

1.2 Web server and Web client

A web server is computer software and associated hardware that accepts requests via HTTP or its secure variant HTTPS. A user agent, commonly a web browser or web crawler, initiates communication by making a request for a web page or other resource using HTTP. The server responds with the content of that resource or an error message. A web server can also accept and store resources sent from the user agent if it is configured to do so.

The hardware used to run a web server varies according to the volume of requests that it needs to handle. At the low end, embedded systems such as a router can run a small web server at its configuration interface. *Nginx* is an open-source, high-performance web server, released in 2004. It has become one of the most used web servers, together with *Apache*. Nginx is an ideal solution for managing high-traffic websites. Many high-visibility websites and applications, such as Netflix or Pinterest, use the Nginx web server.

A resource sent from a web server can be a pre-existing file available to the web server, or it can be generated dynamically at the time of the request by another program that communicates with the server. The former can be served faster and can be easily cached for repeated requests. But the latter supports a broader range of applications.

The client side (user side) of the Web. A Web client typically refers to the Web browser in the user's machine or mobile device. It may also refer to extensions and helper applications that enhance the browser to support special services from the site.

1.3 Internet

Two or more connected computers with the capability of exchanging information is called a network. A network that usually takes the form of a small office network is called as local area network (LAN). A network that covers a city or some other large physical area can be called metropolitan area network (MAN) and, still a larger area network is known as a wide area network (WAN).

Advantages of building a network include the following:

- Communication such as e-mail facility
- Information sharing such as database accessing
- Resource sharing such as printing using networked printers

One can combine two or more networks resulting a larger network, called an *internetwork*. Consider that there are two LANs, placed in the first floor and second floor of a building. These two networks are needed to be joined together. One can join them physically by drilling a hole in the floor, and then setting up some network equipment and wiring the two networks. The resulting network is an internetwork. Such small network of networks is also called intranet. Intranet is a private network. A diagram of an internetwork is shown in Figure 1.1.

But Internet is a global network and nobody owns it. Intranet uses the Internet communication standards and tools to provide information to the restricted users.

The Internet is a global system of interconnected computer networks that uses TCP/IP protocol to communicate between networks and devices. It is a network of networks consisting of private, public, academic, business, and government networks of local to global scope, linked by a broad array of electronic, wireless, and optical networking technologies. It carries a vast range of information resources and services. Some important services offered by the Internet are given as follows: Electronic email, Video conference, Instant messaging, E-Commerce, Web services, Telnet, File transfer, Internet relay chat, VoIP, Newsgroup, Electronic fax, World wide web.

A router is a networking device that connects two or more packet-switched networks or subnetworks. It manages traffic between these networks by forwarding data packets to their intended IP addresses, and allows multiple devices to use the same Internet connection.

1.4 Hypertext, hypermedia, hyperlink

Hypertext is text which contains links to other texts. Texts are read linearly. But hypertext is not constrained to be linear.

HyperMedia is a term used for hypertext which is not constrained to be text. It can include graphics, video and sound.

Hyperlink is a link that allows a user to navigate between different computer resources like webpages. A hyperlink could be absolute, or relative or inline link. A hyperlink can be one of the following types:

Text hyperlink uses a word or phrase to take visitors to another page, file or document.

Image hyperlink uses an image to take visitors to another page, file or document.

Bookmark hyperlink uses text or an image to take visitors to another part of a web page.

E-mail hyperlink allows visitors to send an e-mail message to the displayed e-mail address.

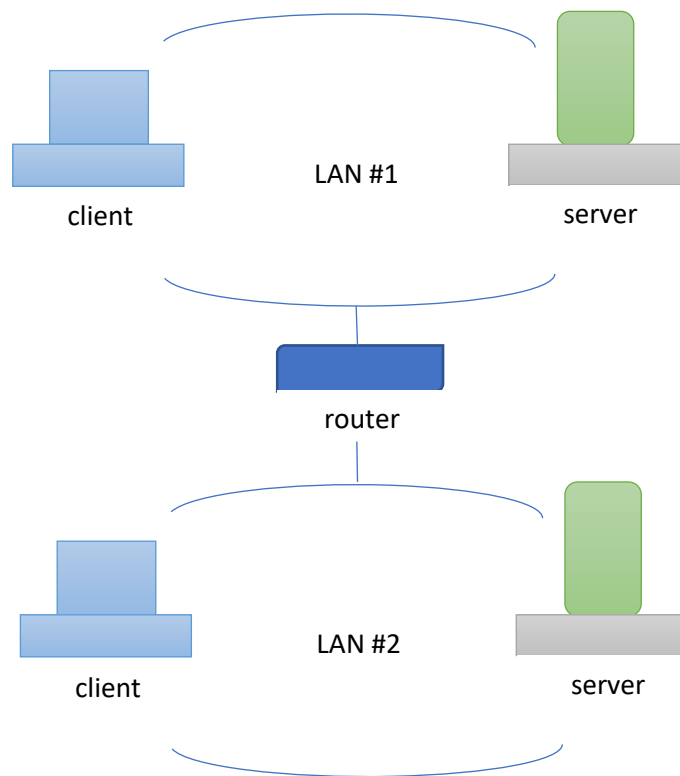


Figure 1.1: An internetwork made by joining two LANs

1.5 Working principle of World Wide Web

Hypertext Transfer Protocol (HTTP) is an application layer protocol for distributed, collaborative, hypermedia information systems. It is the foundation of data communication for the WWW, where hypertext documents include hyperlinks to other resources that the user can easily access by a mouse clicking or tapping the screen in a web browser.

HTTP is a fast and efficient protocol that controls different communications taking place between the Web browser client and the server. It uses the Transmission Control Protocol (TCP) to transport all of its control and data messages from one computer to another.

Web pages are typically grouped as a Web site. A website starts by displaying home page, also known as the index page of the website. Users can navigate into other pages by mouse clicking on hyperlinks such as menus, texts, buttons and images. These hyperlinks reference to other pages or information within a page. When one clicks a hyperlink, it jumps to another part of the same page, or a new page in the same Web site, or to another Website. A click might lead to an execution of a program, or display a picture, or download a file. All these hyperlinks are inserted using Hyper Text Markup Language (HTML) which works in harmony with HTTP.

In order to visit a website, user types a web address (URL) in the address bar of the browser. Web browser first gets the IP address of URL from a Domain Name System (DNS) server, and then connects with the target server. The server responds to the client and transfers document to the Web browser. The Web browser then translates and displays the HTML information.

1.6 Communication between Web client and Web server

Lecture 2

Contents: Web browser; Search engine; World Wide Web consortium; Domain naming system; Web development; Web hosting services; Protocol and standard; Malware; Virus and worm; ISDN; Dial-up and leased connection; Exercises

2.1 Web browsers

A web browser is an application program used for accessing websites. When a user requests a web page from a particular website, the browser gets the files from the web server. It then displays the page on the screen. Browsers are associated with devices such as desktop, laptop, tablet and smartphone. Some examples of popular browser are Chrome, Opera, Firefox, Brave, Vivaldi, Edge, Safari, DuckDuckGo, Fulldive VR Browser, and UR Browser.

2.2 Search engines

A search engine is a software designed to search information on the World Wide Web (WWW). It searches WWW in a systematic way for information as specified in search query. Search results are generally presented in a line of results, known as search engine result pages. When a user enters a query into a search engine, the engine scans its index of web pages to find those that are relevant to the user's query. The results are ranked by relevancy and displayed to the user. Each of the output results may be a mix of links to web pages, images, videos, articles, research papers, and other files. Some popular search engines and their years of appearance are given below:

WebCrawler (1993), Lycos (1994), Yahoo! Search (1995), Yandex (1997), Google (1997), MSN Search (1998), Ask.com (2006), Bing (2009), Ecosia (2009), Brave Search (2021)

2.3 World Wide Web consortium

The World Wide Web is also known as the Web, in short. The World Wide Web Consortium (W3C) is the main international standards organization for the Web. It was founded in 1994 and led by Tim Berners-Lee. The consortium is made up of member organizations that maintain staff. They work together in the development of standards for the Web. W3C also engages in education and outreach, develops software and serves as an open forum for discussion about the Web.

2.4 Domain naming system

It is a hierarchical client/server distributed database management system. In the domain name system (DNS), the clients are called resolvers and the servers are called name servers. It associates various information with domain names assigned to each of the associated entities. Another important function is to translate domain names to the numerical IP addresses needed for locating and identifying computer services and devices with the underlying network protocols. DNS is an essential component of the functionality of the Internet. Internet service provider may do this conversion or connect to a specific DNS server that does.

When a domain name is typed in the address bar of a Web browser, a query is sent to the primary DNS server defined in your Web browser's configuration dialog box. The DNS server converts the

domain name to an IP address and returns this address to the system from which the query was originated. Subsequently, the IP address is used in all communications.

DNS is a TCP/IP protocol used on different platforms. The domain name space is divided into three different sections: generic domains, country domains, and inverse domain. Generic domains include *biz* for businesses and firms, *com* for commercial organizations, *gov* for government organizations, and so on. Generic domains show the behaviour for registering domains. The format of country domains is the same as generic domains. It uses two characters to abbreviate a country. For example, *in* for India, *us* for United States of America, *ru* for Russia, and so on. Inverse domain is used for mapping an address to a name for verification of an IP address. For determining whether the client is on the authorized list, it sends a query to the DNS server and ask for mapping an address to the name.

2.5 Web development

A web application is an application software that is accessed using a web browser. Web applications are delivered on the World Wide Web or intranet web to users with an active network connection. Web development is the work of developing web applications. Web applications can range from static web pages to complex web applications, electronic businesses, and social network services. We provide here some ideas of front-end and back-end development.

2.5.1 Front-end development

Front-end development involves developing visual and interactive elements of a website. Three important languages viz., HTML, CSS and JavaScript are required. HTML provides the structure of the webpage, CSS is required for styling and layout, and JavaScript achieves the dynamic behaviour and interactivity.

Front-end developers are responsible for creating the user interface of a website, looking website good and making it easy to use. It is needed to focus on design principles and user experience. Front-end developers will be working closely with designers, back-end developers, and project managers to make sure the final product meets the client's requirements and provides the best possible experience for the end-users.

2.5.1 Back-end development

Backend web development deals with the server-side of a web application. This involves tasks such as creating and managing the server-side logic, connecting the application to a database, creating server-side APIs, handling user authentication and authorization, and processing and responding to user requests. It requires a programming language such as Python, Java, Ruby, PHP, Node.js, and .NET language.

Backend developers are responsible for the development of server-side components of web applications. They need to work with databases, handle requests, create server-side APIs that can be used by frontend developers to retrieve and manipulate data, ensure the scalability of the systems so that the backend can handle a high volume of traffic and by integrating external services such as payment gateways, message queues, and cloud services.

2.6 Web hosting services

A web hosting service allows a website to be placed on a Web server so that Web clients can access it using Internet service. It offers the facilities required for them to create and maintain a site and

makes it accessible on the Web. Companies that provide web hosting services are called *web hosts*. Web hosting needs the following things:

- One or more servers to act as the host(s) for the websites
- Physical space, electricity, Internet connectivity and cooling service of the servers
- Domain name system to identify the name(s) for the sites and their hosting server(s)
- A web server running on the Web host
- Each website on server(s) needs the following things:
 - Space to keep website
 - Website-specific configuration
 - Database for some websites
 - Software for the clients to access, modify and configure the website
 - Email connectivity

2.7 Protocol and standard

Computer networks are used to communicate entities in different systems. Entities need to follow a set of rules that govern data communications, called protocol. A protocol defines what is communicated, how it is communicated, and when it is communicated. The following elements are used to define a protocol.

- Syntax: It refers to the structure or format of the data for communication.
- Semantics: It refers to the meaning of each group of bits in the message.
- Timing: It has two components: timing of sending data and data rate.

A standard is decided by an agreed-upon rules. Standards are of two types: agreed by fact or convention, and agreed by law or regulation. The first type is known as *de facto* standard, and the second one is known as *de jure* standard.

2.8 Malware

Malware is short form of malicious software. Malware refers to all code and applications written to harm devices or the people who use them. Viruses and worms are two distinct types of malware, but both can spread rapidly. Some ways of spreading malware are given below.

Email: An email can include infected attachments and malicious links containing viruses or worms.

Security vulnerabilities: A worm can scan a network, find all devices with the vulnerability, and exploit them to gain access. Developers sometimes accidentally leave vulnerabilities in their code, which hackers exploit to insert malware into your system.

P2P file-sharing: If you download music, movies, or apps from unauthorized sources, you might get more than what you are supposed to download.

Malvertising: Many websites often contain infected advertisements in the form of banners or pop-ups. If you click on them, you can unwittingly download a virus or worm. Even if you don't click, some of these advertisements can execute a drive-by-download that infects your device.

2.9 Virus and worm

First, we note their differences.

* A virus lays dormant in your system until you execute it or take some other required action. But, worms do not rely on you to trigger them.

Lecture 3

Contents: Introduction; Editors for writing HTML program; Installing notepad++ editor; Writing the first program; Executing a HTML program; Basic concepts / elements; Origin and current version of HTML; Exercises

3.1 Introduction

HTML stands for *hypertext markup language*, used widely as a language to write web pages. Hypertext documents are interconnected by hyperlinks, which are typically activated by a mouse click, or keypress set, or screen touch. The term *hypertext* is also sometimes used to describe tables, images, and other presentational content formats with integrated hyperlinks. Hypertext is one of the key underlying concepts of the World Wide Web, where Web pages are often written in HTML.

3.2 Editors for writing HTML program

There are many editors by which an HTML program could be developed. A few names are given as follows:

- Atom
- Notepad ++
- Sublime Text
- Visual Studio Code
- Adobe Dreamweaver CC
- Froala
- CoffeeCup

We shall take notepad++ as an editor for coding HTML programs. notepad also serves as an editor.

notepad++ is a free editor and it supports several languages. Running in the MS Windows environment, its use is governed by GNU General Public License.

Based on the powerful editing component *Scintilla*, notepad++ is written in C++ and uses pure Win32 API and STL which ensures a higher execution speed and smaller program size.

NOTE: GNU stands for GNU's Not Unix; API has full form application programming interface; STL is a file format native to the stereolithography CAD software created by 3D Systems. STL has several acronyms such as *standard triangle language* and *standard tessellation language*.

3.3 Installing notepad++ editor

It is a free software. Visit website <https://notepad-plus-plus.org/downloads/> to download notepad++ software. Otherwise, you can give a search using “download notepad++” using Google search. Follow the instructions to install the software and choose option to create a shortcut at the desktop.

3.4 Writing the first program

Every HTML program has extension `.html`. One can execute an HTML program in different ways.

▷ Firstly, one can just double click the icon of the file containing the HTML program.

▷ One could write or copy the entire path of the file containing HTML program in the address bar of a browser and then press the “Enter” key of the keyboard. Then the program gets executed.

HTML programs are tag-based. Here is an example of program file `first.html`.

```
<html>
<head>
<title>First Page</title>
</head>
<body>
<h1>This is an H1 heading</h1>
<p>My day starts with a morning walk of 45 minutes. </p>
</body>
</html>
```

Figure 3.1: An HTML program that displays an H1 heading and a paragraph

The output is displayed on the screen. After the execution of the program given in Fig. 3.1, the title “First Page” is displayed in the tab, placed above the screen. Note that program title can be different from the file name. The file name is given here as “first.html”. The output of the program is displayed on the screen, and highlighted below:

This is an H1 heading

My day starts with a morning walk of 45 minutes.

It has two lines: first line is an H1 heading, the string put within `<h1>` and `</h1>` tags, and second line is a paragraph, put within `<p>` and `</p>` tags. There are six different heading tags: H1, H2, H3, H4, H5, H6. Students are encouraged to replace H1 by one of the remaining headings and check the output.

3.5 Text formatting tags

Let us work with different text processing tags such as bold, italic, underlined, strong, super script and sub script. The program in Fig. 3.2 illustrates these tags. The output of the program is given below.

This is a line of bold text.

But, this line is kept in italic.

This is an underlined text.

But, this line looks strong.

Write 100 to the power 2: 100²

and x subscript 3: x₃

Note that we have used tag `
` to display the next output at the next line, i.e., a line-break given in the text.

3.6 Basic concepts / elements

▷ An HTML program may start with the statement `<!DOCTYPE html>`. This declaration represents the document type as `html` program, and helps browsers to display web pages correctly.

▷ How do you view HTML source code? Right-click on an HTML page and select "View page source" in Chrome or in Edge, or similar in other browsers.

▷ An HTML element is everything from the start tag to the end tag. Refer to Fig. 3.1. Here, `<title>First Page</title>` is an element. The following is an example of nested element.

```
<head>
```

```
<title>First Page</title>
```

```
</head>
```

▷ An element containing another element is known as nested element. HTML elements with no content are called empty elements. In Fig 3.2, `
` is an empty element.

```
<html>
<head>
<title>Text formatting tags</title>
</head>
<body>
<b>This is a line of bold text.</b><br>
<i>But, this line is kept in italic.</i><br>
<u>This is an underlined text.</u><br>
<strong>But, this line looks strong.</strong><br>
Write 100 to the power 2: 100<sup>2</sup><br>
and x subscript 3: x<sub>3</sub>
</body>
</html>
```

Figure 3.2: Different text processing tags

▷ HTML is not case sensitive. This means that `<body>`, `<Body>` and `<BODY>` are the same tag.