



Introduction to the Internet and World Wide Web

Chapter Objectives

In this chapter, you will learn how to . . .

- Describe the evolution of the Internet and the Web
- Explain the need for web standards
- Describe universal design
- Identify benefits of accessible web design
- Identify reliable resources of information on the Web
- Identify ethical use of the Web
- Describe the purpose of web browsers and web servers
- Identify networking protocols
- Define URIs and domain names
- Describe HTML, XHTML, and HTML5
- Describe popular trends in the use of the Web

The Internet and the Web are parts of our daily lives. How did they begin? What networking protocols and programming languages work behind the scenes to display a web page? This chapter provides an introduction to some of these topics and is a foundation for the information that web developers need to know. You'll be introduced to Hypertext Markup Language (HTML), the language used to create web pages.

1.1 The Internet and the Web

The Internet

The **Internet**, the interconnected network of computer networks that spans the globe, seems to be everywhere today. It has become part of our lives. You can't watch television or listen to the radio without being urged to visit a website. Even newspapers and magazines have their place on the Internet.

Birth of the Internet

The Internet began as a network to connect computers at research facilities and universities. Messages in this network would travel to their destination by multiple routes, or paths. This configuration allowed the network to function even if parts of it were broken or destroyed. In such an event, the message would be rerouted through a functioning portion of the network while traveling to its destination. This network was developed by the Advanced Research Projects Agency (ARPA)—and the ARPAnet was born. Four computers (located at UCLA, Stanford Research Institute, University of California Santa Barbara, and the University of Utah) were connected by the end of 1969.

Growth of the Internet

As time went on, other networks, such as the National Science Foundation's NSFnet, were created and connected with the ARPAnet. Use of this interconnected network, or Internet, was originally limited to government, research, and educational purposes. The number of individuals accessing the Internet continues to grow each year. According to Internet World Stats (<http://www.internetworldstats.com/emarketing.htm>), the percentage of the global population that used the Internet was 0.4% in 1995, 5.8% in 2000, 15.7% in 2005, 28.8% in 2010, and 45% in 2015. Visit <http://www.internetworldstats.com> to explore more statistics about the usage and growth of the Internet.

The lifting of the restriction on commercial use of the Internet in 1991 set the stage for future electronic commerce: Businesses were now welcome on the Internet. However, the Internet was still text based and not easy to use. The next set of developments solved this issue.

Birth of the Web

While working at CERN, a research facility in Switzerland, **Tim Berners-Lee** envisioned a means of communication for scientists by which they could easily “hyperlink” to another research paper or article and immediately view it. Berners-Lee created the World Wide Web to fulfill this need. In 1991, Berners-Lee posted the code for the Web in a newsgroup and made it freely available. This version of the World Wide Web used **Hypertext Transfer Protocol (HTTP)** to communicate between the client computer and the web server, used **Hypertext Markup Language (HTML)** to format the documents, and was text based.

The First Graphical Browser

In 1993, Mosaic, the first graphical web browser became available. Marc Andreessen and graduate students working at the National Center for Supercomputing Applications (NCSA) at the University of Illinois Urbana–Champaign developed Mosaic. Some individuals in this



group later created another well-known web browser—Netscape Navigator—which is an ancestor of today’s Mozilla Firefox browser.

Convergence of Technologies

By the early 1990s, personal computers with easy-to-use graphical operating systems (such as Microsoft’s Windows, IBM’s OS/2, and Apple’s Macintosh OS) were increasingly available and affordable. Online service providers such as CompuServe, AOL, and Prodigy offered low-cost connections to the Internet. Figure 1.1 depicts this convergence of available computer hardware, easy-to-use operating systems, low-cost Internet connectivity, the HTTP protocol and HTML language, and a graphical browser that made information on the Internet much easier to access. The **World Wide Web**—the graphical user interface to information stored on computers running web servers connected to the Internet—had arrived!

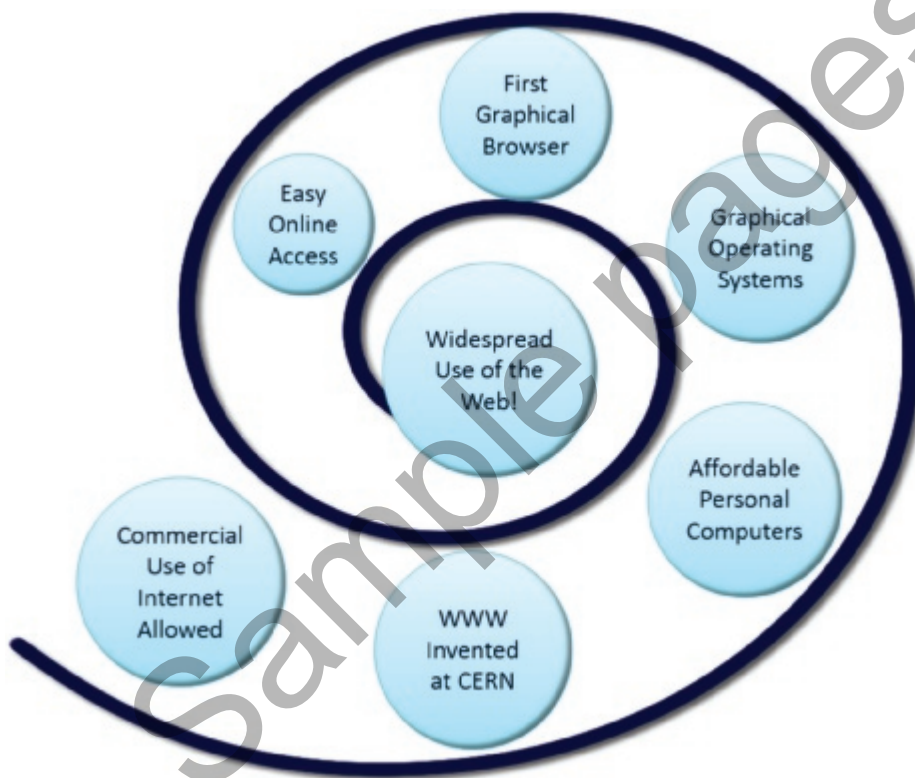


Figure 1.1 Convergence of Technologies. Used by permission of World Wide Web Consortium (W3C)

Who Runs the Internet?

You may be surprised that there is no single person “in charge” of the global interconnected network of computer networks known as the Internet. Instead, Internet infrastructure standards are overseen by groups such as the **Internet Engineering Task Force (IETF)** and the **Internet Architecture Board (IAB)**. The IETF is the principal body engaged in the development of new Internet protocol standard specifications. It is an open international community of network designers, operators, vendors, and researchers concerned with the evolution of Internet architecture and the smooth operation of the Internet. The actual technical work of the IETF is completed in its working groups. These working groups are organized into areas by topic, such as security and routing.

The IAB is a committee of the IETF and provides guidance and broad direction to the IETF. As a function of this purpose, the IAB is responsible for the publication of the **Request for Comments (RFC)** document series. An RFC is a formal document from the IETF that is drafted by a committee and subsequently reviewed by interested parties. RFCs are available for online review at <http://www.ietf.org/rfc.html>. Some RFCs are informational in nature, while others are meant to become Internet standards. In the latter case, the final version of the RFC becomes a new standard. Future changes to the standard must be made through subsequent RFCs.

The **Internet Corporation for Assigned Numbers and Names (ICANN)**, <http://www.icann.org>, was created in 1998 and is a nonprofit organization. Its main function is to coordinate the assignment of Internet domain names, IP address numbers, protocol parameters, and protocol port numbers. Prior to 1998, the **Internet Assigned Numbers Authority (IANA)** coordinated these functions. IANA still performs certain functions under the guidance of ICANN and maintains a website at <http://www.iana.org>.

Intranets and Extranets

Recall that the Internet is an interconnected network of computer networks that is globally available. When an organization needs the communication capabilities of the Internet, but doesn't want its information to be available to everyone, either an intranet or extranet is appropriate.

An **intranet** is a private network that is contained within an organization or business. Its purpose is to share organizational information and resources among coworkers. When an intranet connects to the outside Internet, usually a gateway or firewall protects the intranet from unauthorized access.

An **extranet** is a private network that securely shares part of an organization's information or operations with external partners such as suppliers, vendors, and customers. Extranets can be used to exchange data, share information exclusively with business partners, and collaborate with other organizations. Privacy and security are important issues in extranet use. Digital certificates, encryption of messages, and virtual private networks (VPNs) are some technologies used to provide privacy and security for an extranet. Digital certificates and encryption used in e-commerce are discussed in Chapter 12.

1.2 Web Standards and Accessibility

Just as with the Internet, no single person or group runs the World Wide Web. However, the **World Wide Web Consortium** (<http://www.w3.org>), referred to as the **W3C**, takes a proactive role in developing recommendations and prototype technologies related to the Web. Topics that the W3C addresses include web architecture, standards for web design, and accessibility. In an effort to standardize web technologies, the W3C produces specifications called recommendations.

W3C Recommendations

The W3C Recommendations are created in working groups with input from many major corporations involved in building web technologies. These recommendations are not rules; they are guidelines. Major software companies that build web browsers, such as Microsoft, do not always follow the W3C Recommendations. This makes life challenging for web developers because not all browsers will display a web page in exactly the same way. The good news

is that there is a convergence toward the W3C Recommendations in new versions of major browsers. You'll follow W3C Recommendations as you code web pages in this book. Following the W3C Recommendations is the first step toward creating a website that is accessible.

Web Standards and Accessibility

The **Web Accessibility Initiative (WAI)** (<http://www.w3.org/WAI>), is a major area of work by the W3C. Since the Web has become an integral part of daily life, there is a need for all individuals to be able to access it.

The Web can present barriers to individuals with visual, auditory, physical, and neurological disabilities. An **accessible** website provides accommodations that help individuals overcome these barriers. The WAI has developed recommendations for web content developers, web authoring tool developers, web browser developers, and developers of other user agents to facilitate use of the Web by those with special needs. See the WAI's **Web Content Accessibility Guidelines (WCAG)** at <http://www.w3.org/WAI/WCAG20/glance/WCAG2-at-a-Glance.pdf> for a quick overview.

Accessibility and the Law

The **Americans with Disabilities Act (ADA)** of 1990 is a federal civil rights law that prohibits discrimination against people with disabilities. The ADA requires that business, federal, and state services are accessible to individuals with disabilities. A 1996 Department of Justice ruling (http://www.justice.gov/crt/foia/readingroom/frequent_requests/ada_coreletter/cltr204.txt) indicated that ADA accessibility requirements apply to Internet resources.

Section 508 of the Federal Rehabilitation Act was amended in 1998 to require that U.S. government agencies give individuals with disabilities access to information technology that is comparable to the access available to others. This law requires developers creating information technology (including web pages) for use by the federal government to provide for accessibility. The **Federal IT Accessibility Initiative** (<http://www.section508.gov>) provides accessibility requirement resources for information technology developers. As the Web and Internet technologies developed, it became necessary to review the original Section 508 requirements. New proposed Section 508 requirements were aligned to WCAG 2.0 guidelines and released for comment in 2015. This textbook focuses on WCAG 2.0 guidelines to provide for accessibility.

In recent years, state governments have also begun to encourage and promote web accessibility. The Illinois Information Technology Accessibility Act (IITAA) guidelines (see <http://www.dhs.state.il.us/IITAA/IITAAWebImplementationGuidelines.html>) are an example of this trend.

Universal Design for the Web

The Center for Universal Design defines universal design as “the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design.” Examples of universal design are all around us. The cutouts on curbs that make it possible for people in wheelchairs to access the street also benefit a person pushing a stroller or riding a Segway Personal Transporter (Figure 1.2). Doors that



Focus on
Accessibility



Focus on
Accessibility



Figure 1.2 A smooth ride is a benefit of universal design. Definition from *The Principles of Universal Design* by The Center for Universal Design. Published by NC State University, ©1997

open automatically for people with mobility challenges also benefit people carrying packages. A ramp is useful for a person in a wheelchair, a person dragging a rolling backpack or carry-on bag, and so on.

Awareness of universal design by web developers has been steadily increasing. Forward-thinking web developers design with accessibility in mind because it is the right thing to do. Providing access for visitors with visual, auditory, and other challenges should be an integral part of web design rather than an afterthought.

A person with visual difficulties may not be able to use graphical navigation buttons and may use a screen reader device to provide an audible description of the web page. By making a few simple changes, such as providing text descriptions for the images and perhaps providing a text navigation area at the bottom of the page, web developers can make the page accessible. Often, providing for accessibility increases the usability of the website for all visitors.

Accessible websites, with alternative text for images, headings used in an organized manner, and captions or transcriptions for multimedia features, are more easily used not only by visitors with disabilities, but also by visitors using a browser on a mobile device such as a phone or tablet. Finally, accessible websites may be more thoroughly indexed by search engines, which can be helpful in bringing new visitors to a site. As this text introduces web development and design techniques, corresponding web accessibility and usability issues are discussed.



Focus on
Accessibility

1.3 Information on the Web

These days anyone can publish just about anything on the Web. In this section we'll explore how you can tell if the information you've found is reliable and how you can use that information.

Reliability and Information on the Web

There are many websites—but which ones are reliable sources of information? When visiting websites to find information, it is important not to take everything at face value (Figure 1.3).

Questions to ask about web resources are listed as follows;

- **Is the organization credible?**

Anyone can post anything on the Web! Choose your information sources wisely. First, evaluate the credibility of the website itself. Does it have its own domain name, such as <http://mywebsite.com>, or is it a free website consisting of just a folder of files hosted on a free web hosting site (such as weebly.com, awardspace.com, or 000webhost.com)? The URL of a free website usually includes part of the free web host's domain name. Information obtained from a website that has its own domain name will usually (but not always) be more reliable than information obtained from a free website.

Evaluate the type of domain name: Is it for a nonprofit organization (.org), a business (.com or .biz), or an educational institution (.edu)? Businesses may provide information in a biased manner, so be careful. Nonprofit organizations and schools will sometimes treat a subject more objectively.

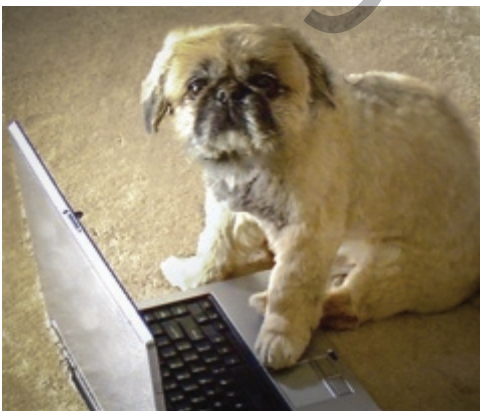


Figure 1.3 Who really updated that web page you are viewing?

- **How recent is the information?**

Another item to look at is the date the web page was created or last updated. Although some information is timeless, very often a web page that has not been updated for several years is outdated and may not be the best source of information.

- **Are there links to additional resources?**

Hyperlinks indicate websites with supporting or additional information that can be helpful to you in your research as you explore a topic. Look for these types of hyperlinks to aid your studies.

- **Is it Wikipedia?**

Wikipedia (<http://wikipedia.org>) is a good place to begin research, but don't accept what you read there for fact, and avoid using Wikipedia as a resource for academic assignments. Why? Well, except for a few protected topics, anyone can update Wikipedia with anything! Usually it all gets sorted out eventually—but be aware that the information you read may not be valid.

Feel free to use Wikipedia to begin exploring a topic, but then scroll down to the bottom of the Wikipedia web page and look for “References”—and then explore those websites and others that you may find. As you gather information on these sites, also consider the other criteria: credibility, domain name, timeliness, and links to additional resources.

Ethical Use of Information on the Web

This wonderful technology called the World Wide Web provides us with information, graphics, and music—all virtually free (after you pay your Internet service provider, of course). Let's consider the following issues relating to the ethical use of this information:

- Is it acceptable to copy someone's graphic to use on your own website?
- Is it acceptable to copy someone's website design to use on your own site or on a client's site?
- Is it acceptable to copy an essay that appears on a web page and use it, or parts of it, as your own writing?
- Is it acceptable to insult someone on your website or link to that person's site in a derogatory manner?

The answer to all these questions is no. Using someone's graphic without permission is the same as stealing it. In fact, if you link to it, you are actually using up some of the site's bandwidth and may be costing the owner money. Instead, ask the owner of the website for permission to use the graphic. If permission is granted, store the graphic on your own website and be sure to indicate the source of the graphic when you display it on your web page. The key is to request permission before using someone else's resources. Copying the website design of another person or company is also a form of stealing. Any text or graphic on a website is automatically copyrighted in the United States, regardless of whether a copyright symbol appears on the site or not. Insulting a person or company on your website or linking to the person's or company's website in a derogatory manner could be considered a form of defamation.

Issues like these, related to intellectual property, copyright, and freedom of speech, are regularly discussed and decided in courts of law. Good Web etiquette requires that you ask permission before using others' work, give credit for what you use (“fair use” in the U.S.



**Focus on
Ethics**

copyright law), and exercise your freedom of speech in a manner that is not harmful to others. The **World Intellectual Property Organization (WIPO)**, <http://wipo.int>, is dedicated to protecting intellectual property rights internationally.

What if you'd like to retain ownership, but make it easy for others to use or adapt your work? **Creative Commons**, <http://creativecommons.org>, is a nonprofit organization that provides free services that allow authors and artists to register a type of a copyright license called a Creative Commons license. There are several licenses to choose from, depending on the rights you wish to grant. The Creative Commons license informs others as to exactly what they can and cannot do with your creative work. See <http://meyerweb.com/eric/tools/color-blend> to view a web page licensed under a Creative Commons Attribution-ShareAlike 1.0 License.



Checkpoint 1.1

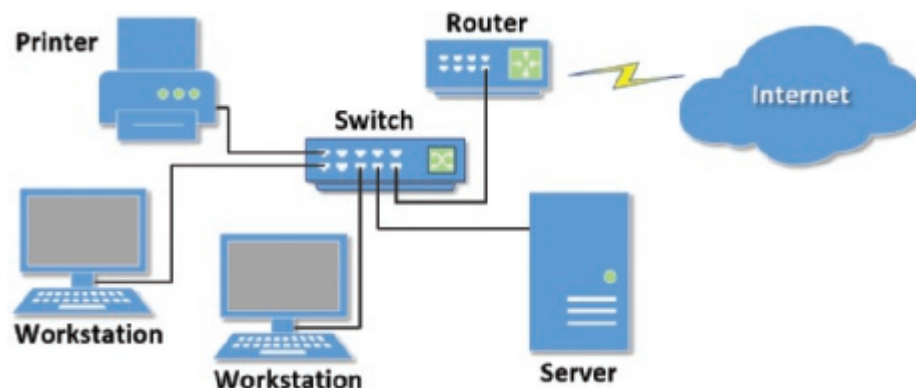
1. Describe the difference between the Internet and the Web.
2. Explain three events that contributed to the commercialization and exponential growth of the Internet.
3. Is the concept of universal design important to web developers? Explain your answer.

1.4 Network Overview

A **network** consists of two or more computers connected for the purpose of communicating and sharing resources. Common components of a network are shown in Figure 1.4 and include the following:

- Server computer(s)
- Client workstation computer(s)
- Shared devices such as printers
- Networking devices (routers, hubs, and switches) and the media that connect them

Figure 1.4
Common components of a network



The **clients** are the computer workstations used by individuals, such as a personal computer (PC) on a desk. The **server** receives requests from client computers for resources such as files. Computers used as servers are usually kept in a protected, secure area and are accessed only by network administrators. Networking devices such as hubs and switches provide network connections for computers, and routers direct information from one network to another. The **media** connecting the clients, servers, peripherals, and networking devices may consist of copper cables, fiber optic cables, or wireless technologies. Networks vary in scale. A **local area network (LAN)** is usually confined to a single building or group of connected buildings. Your school computer lab may use a LAN. If you work in an office, you probably use a computer connected to a LAN. A **wide area network (WAN)** is geographically dispersed and usually uses some form of public or commercial communications network. For example, an organization with offices on both the East and West Coasts of the United States probably uses a WAN to provide a link between the LANs at each of the offices.

A **backbone** is a high-capacity communication link that carries data gathered from smaller links that interconnect with it. On the Internet, a backbone is a set of paths that local or regional networks connect to for long-distance interconnection. The Internet is a group of interconnected networks with very high-speed connectivity provided by the Internet backbones.

1.5 The Client/Server Model

The term **client/server** dates from the 1980s and refers to personal computers joined by a network. “Client/server” can also describe a relationship between two computer programs—the client and the server. The client requests some type of service (such as a file or database access) from the server. The server fulfills the request and transmits the results to the client over a network. While both the client and the server programs can reside on the same computer, typically they run on different computers (Figure 1.5). It is common for a server to handle requests from multiple clients.

The Internet is a great example of client/server architecture at work. Consider the following scenario: An individual is at a computer using a web browser client to access the Internet. The individual uses the web browser to visit a website, let’s say <http://www.yahoo.com>. The server is the web server program running on the computer with an IP address that corresponds to yahoo.com. It is contacted, locates the web page and related resources that were requested, and responds by sending them to the individual.

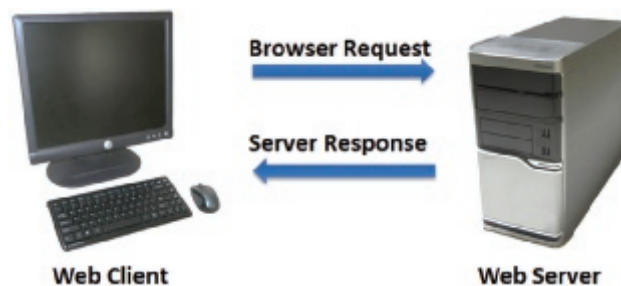


Figure 1.5 Web client and web server

Here's how to distinguish between web clients and web servers:

Web Client

- Connected to the Internet when needed
- Usually runs web browser (client) software such as Internet Explorer or Firefox
- Uses HTTP
- Requests web pages from a server
- Receives web pages and files from a server

Web Server

- Continually connected to the Internet
- Runs web server software (such as Apache or Internet Information Server)
- Uses HTTP
- Receives a request for the web page
- Responds to the request and transmits the status code, web page, and associated files

When clients and servers exchange files, they often need to indicate the type of file that is being transferred; this is done through the use of a MIME type. **Multipurpose Internet Mail Extensions (MIME)** are rules that allow multimedia documents to be exchanged among many different computer systems. MIME was initially intended to extend the original Internet e-mail protocol, but it is also used by HTTP. MIME provides for the exchange of seven different media types on the Internet: audio, video, image, application, message, multipart, and text. MIME also uses subtypes to further describe the data. The MIME type of a web page is text/html. MIME types of GIF and JPEG images are image/gif and image/jpeg, respectively.

A web server determines the MIME type of a file before the file is transmitted to the web browser. The MIME type is sent along with the document. The web browser uses the MIME type to determine how to display the document.

How does information get transferred from the web server to the web browser? Clients (such as web browsers) and servers (such as a web server) exchange information through the use of communication protocols such as HTTP, TCP, and IP, which are introduced in the next section.

1.6 Internet Protocols

Protocols are rules that describe how clients and servers communicate with each other over a network. There is no single protocol that makes the Internet and Web work; a number of protocols with specific functions are needed.

File Transfer Protocol (FTP)

File Transfer Protocol (FTP) is a set of rules that allow files to be exchanged between computers on the Internet. Unlike HTTP, which is used by web browsers to request web pages and their associated files in order to display a web page, FTP is used simply to move files

from one computer to another. Web developers commonly use FTP to transfer web page files from their computers to web servers. FTP is also commonly used to download programs and files from other servers to individual computers.

E-mail Protocols

Most of us take e-mail for granted, but there are two servers involved in its smooth functioning: an incoming mail server and an outgoing mail server. When you send e-mail to others, **Simple Mail Transfer Protocol (SMTP)** is used. When you receive e-mail, **Post Office Protocol** (POP; currently **POP3**) and **Internet Message Access Protocol (IMAP)** can be used.

Hypertext Transfer Protocol (HTTP)

HTTP is a set of rules for exchanging files such as text, images, audio, video, and other multimedia on the Web. Web browsers and web servers usually use this protocol. When the user of a web browser requests a file by typing a website address or clicking on a hyperlink, the browser builds an HTTP request and sends it to the server. The web server in the destination machine receives the request, does any necessary processing, and responds with the requested file and any associated files (such as image files, media files, and other related files).

Transmission Control Protocol/Internet Protocol (TCP/IP)

Transmission Control Protocol/Internet Protocol (TCP/IP) has been adopted as the official communication protocol of the Internet. TCP and IP have different functions that work together to ensure reliable communication over the Internet.

TCP

The purpose of **TCP** is to ensure the integrity of network communication. TCP starts by breaking files and messages into individual units called **packets**. These packets (see Figure 1.6) contain information such as the destination, source, sequence number, and checksum values used to verify the integrity of the data.



Figure 1.6 TCP packet

TCP is used together with IP to transmit files efficiently over the Internet. IP takes over after TCP creates the packets, using IP addressing to send each packet over the Internet via the best path at the particular time. When the destination address is reached, TCP verifies the integrity of each packet by using the checksum, requests a resend if a packet is damaged, and reassembles the file or message from the multiple packets.

IP

Working in harmony with TCP, **IP** is a set of rules that controls how data is sent between computers on the Internet. IP routes a packet to the correct destination address. Once sent, the packet gets successively forwarded to the next closest router (a hardware device designed to move network traffic) until it reaches its destination.

Each device connected to the Internet has a unique numeric **IP address**. These addresses consist of a set of four groups of numbers, called octets. The current version of IP, **Internet Protocol Version 4 (IPv4)**, uses 32-bit (binary digit) addressing. This results in a decimal number in the format of xxx.xxx.xxx.xxx, where each xxx is a value from 0 to 255. Theoretically, this system allows for at most 4 billion possible IP addresses (although many potential addresses are reserved for special uses). However, even this many addresses will not be enough to meet the needs of all of the devices expected to be connected to the Internet in upcoming years.

IP Version 6 (IPv6) will be the next standard IP protocol and will provide a huge increase in the number of possible addresses and many technological advances. IPv6 was designed as an evolutionary set of improvements to the current IPv4 and is backwardly compatible with it. Service providers and Internet users can update to IPv6 independently without having to coordinate with each other. IPv6 provides for more Internet addresses because the IP address is lengthened from 32 bits to 128 bits. This means that there are potentially 2^{128} unique IP addresses possible, or 340,282,366,920,938,463,463,347,607,431,768,211,456 addresses. (Now there will be enough IP addresses for everyone's PC, notebook, cell phone, tablet, toaster, and so on!)

The IP address of a device may correspond to a domain name. The **Domain Name System (DNS)** associates these IP addresses with the text-based URLs and domain names you type into a web browser address box. For example, at the time this book was written, one of Google's IP addresses was 173.194.116.72. You can enter this number in the address text box in a web browser (as shown in Figure 1.7), press Enter, and the Google home page will display. Of course, it's much easier to type "google.com," which is why domain names such as google.com were created in the first place! Since long strings of numbers are difficult for humans to remember, the Domain Name System was introduced as a way to associate text-based names with numeric IP addresses.

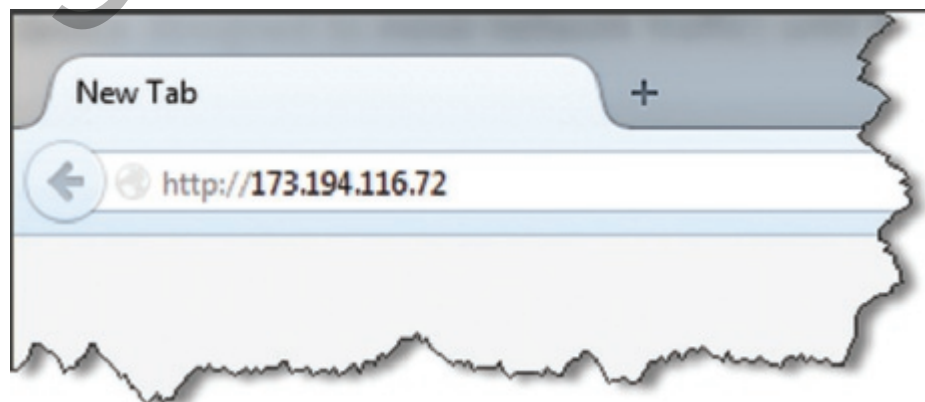


Figure 1.7 Entering an IP address in a web browser



FAQ What is HTTP/2?

HTTP/2 will be the first major update to HTTP, which was first developed in 1999. As websites have become more image and media intensive, the number of requests needed to display a web page and its related files have increased. A major benefit of HTTP/2 will be quicker loading of web pages by processing multiple concurrent HTTP requests. Visit the following resources for more information about HTTP/2:

- <https://tools.ietf.org/html/draft-ietf-httpbis-http2-17>
- <http://readwrite.com/2015/02/18/http-update-http2-what-you-need-to-know>
- <https://http2.github.io>

1.7 Uniform Resource Identifiers and Domain Names

URIs and URLs

A **Uniform Resource Identifier (URI)** identifies a resource on the Internet. A **Uniform Resource Locator (URL)** is a type of URI which represents the network location of a resource such as a web page, a graphic file, or an MP3 file. The URL consists of the protocol, the domain name, and the hierarchical location of the file on the web server.

The URL `http://www.pearsonglobaleditions.com/felke-morris/chapter1/index.html`, shown in Figure 1.8, denotes the use of HTTP protocol and the web server named `www` at the domain name of `pearsonglobaleditions.com`. In this case, the root file (usually `index.html` or `index.htm`) of the chapter 1 directory will be displayed.



Figure 1.8 URL Describing a file within a folder

Domain Names

A **domain name** locates an organization or other entity on the Internet. The purpose of the Domain Name System (DNS) is to divide the Internet into logical groups and understandable names by identifying the exact address and type of the organization. The DNS associates the text-based domain names with the unique numeric IP address assigned to a device.

Let's consider the domain name `www.google.com`. The portion `google.com` is the domain name that is registered to Google. The `.com` is the top-level domain name. The `google` is considered a second-level domain name. The `www` is the name of the web sever (sometimes called the **host**) at the `google.com` domain.

A **subdomain** can be configured to house a separate website located at the same domain. For example, Google’s Gmail can be accessed by using the subdomain “gmail” in the domain name (gmail.google.com). Google Maps can be accessed at maps.google.com, and Google News Search is available at news.google.com. See <http://www.labnol.org/internet/popular-google-subdomains/5888/> for a list of the top 40 Google subdomains. The combination of a host/subdomain, second-level domain, and top-level domain name (such as www.google.com or mail.google.com) is called a **fully qualified domain name (FQDN)**.

Top-Level Domain Names

A **top-level domain (TLD)** identifies the rightmost part of the domain name, starting with the final period. A TLD is either a **generic top-level domain (gTLD)**, such as .com for commercial, or a **country-code top-level domain**, such as .fr for France. The **Internet Assigned Numbers Authority (IANA)** website has a complete list of TLDs (<http://www.iana.org/domains/root/db>).

Generic Top-Level Domain Names (gTLDs)

The Internet Corporation for Assigned Names and Numbers (ICANN) administers gTLDs (<http://www.icann.org>). Table 1.1 shows several common gTLDs and their intended use.

Table 1.1 Top-level domains

Generic TLD	Used By
.aero	Air-transport industry
.asia	Pan-Asia and Asia Pacific community
.biz	Businesses
.cat	Catalan linguistic and cultural community
.com	Commercial entities
.coop	Cooperative
.edu	Restricted to accredited degree-granting institutions of higher education
.gov	Restricted to government use
.info	Unrestricted use
.int	International organization (rarely used)
.jobs	Human resources management community
.mil	Restricted to military use
.mobi	Corresponds to a .com website—the .mobi site is designed for easy access by mobile devices
.museum	Museums
.name	Individuals
.net	Entities associated with network support of the Internet, usually Internet service providers or telecommunication companies
.org	Nonprofit entities
.post	Universal Postal Union, an agency of the United Nations
.pro	Professionals such as accountants, physicians, and lawyers
.tel	Contact information for individuals and businesses
.travel	Travel industry

The .com, .org, and .net TLD designations are currently used on the honor system, which means that, for example an individual who owns a shoe store (not related to networking) can register shoes.net.

Expect the number and variety of gTLDs to increase. ICANN accepted proposals for almost 2000 new gTLDs in 2012. A wide variety of new gTLDs were proposed including place names (.quebec, .vegas, and .moscow), retail terms (.blackfriday), financial terms (.cash, .trade, and .loans), technology terms (.systems, .technology, and .app), and whimsical, fun terms (.ninja, .buzz, and .cool). Some of the first new gTLDs to become available included .bike, .guru, .holdings, .clothing, .singles, .ventures, and .plumbing. ICANN has set a schedule to periodically launch new gTLDs. Visit <http://newgtlds.icann.org/en/program-status/delegated-strings> for a list of new gTLDs.

Country-Code Top-Level Domain Names

Two-character country codes have also been assigned as TLD names. These codes were originally intended to be meaningful by designating the geographical location of the individual or organization that registered the name. Table 1.2 lists some popular country codes used on the Web.

Table 1.2 Country codes

Country Code TLD	Country
.au	Australia
.de	Germany
.es	Spain
.eu	European Union (a group of countries rather than a single country)
.in	India
.jp	Japan
.ly	Libya
.nl	The Netherlands
.us	United States
.ws	Samoa

The IANA website at <http://www.iana.org/domains/root/db> has a complete list of country-code TLDs. Domain names with country codes are often used for municipalities, schools, and community colleges in the United States. For example, the domain name www.harper.cc.il.us denotes, from right to left, the United States, Illinois, community college, Harper, and the web server named “www” as the site for Harper College in Illinois.

Although country-code TLD names were intended to designate geographical location, it is fairly easy to obtain a domain name with a country-code TLD that is not local to the registrant. Examples of non-geographical use of country-code TLDs include domain names such as mediaqueri.es, webteacher.ws, and bit.ly. Visit <http://register.com>, <http://godaddy.com>, and many other domain name registration companies for examples of readily available country-code TLDs.

Domain Name System (DNS)

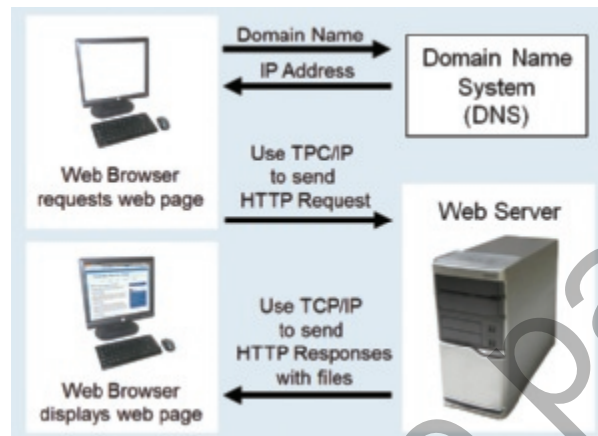
The DNS associates domain names with IP addresses. The following happens each time a new URL is typed into a web browser:

1. The DNS is accessed.
2. The corresponding IP address is obtained and returned to the web browser.

3. The web browser sends an HTTP request to the destination computer with the corresponding IP address.
4. The HTTP request is received by the web server.
5. The necessary files are located and sent by HTTP responses to the web browser.
6. The web browser renders and displays the web page and associated files.

We all get impatient sometimes when we need to view a web page. The next time you wonder why it is taking so long to display a web page, think about all of the processing that goes on behind the scenes before the browser receives the files needed to display the web page (Figure 1.9).

Figure 1.9
Accessing a
web page



1.8 Markup Languages

Markup languages consist of sets of directions that tell the browser software (and other user agents such as mobile phones) how to display and manage a web document. These directions are usually called tags and perform functions such as displaying graphics, formatting text, and referencing hyperlinks.

Standard Generalized Markup Language (SGML)

SGML is a standard for specifying a markup language or tag set. SGML in itself is not a document language, but a description of how to specify one and create a document type definition (DTD). When Tim Berners-Lee created HTML, he used SGML to create the specification.

Hypertext Markup Language (HTML)

HTML is the set of markup symbols or codes placed in a file intended for display on a web browser. The web browser renders the code in the HTML file and displays the web page document and associated files. The W3C (<http://www.w3.org>) sets the standards for HTML.

Extensible Markup Language (XML)

XML was developed by the W3C as a flexible method to create common information formats and share the format and the information on the Web. It is a text-based syntax designed to describe, deliver, and exchange structured information. It is not intended

to replace HTML, but to extend the power of HTML by separating data from presentation. Using XML, developers can create whatever tags they need to describe their information.

Extensible Hypertext Markup Language (XHTML)

XHTML was developed by the W3C to reformulate HTML 4.01 as an application of XML. It combines the formatting strengths of HTML 4.01 and the data structure and extensibility strengths of XML. XHTML has been used on the Web for over a decade and you'll find many web pages coded with this markup language. At one point the W3C was working on a new version of XHTML, called XHTML 2.0. However, the W3C stopped development of XHTML 2.0 because it was not backward compatible with HTML4. Instead, the W3C decided to move forward with HTML5.

HTML5—the Newest Version of HTML

HTML5 is intended to be the successor to HTML4 and will replace XHTML. HTML5 incorporates features of both HTML and XHTML, adds new elements, provides new features such as form edits and native video, and is intended to be backward compatible. The W3C approved HTML5 for Candidate Recommendation status in late 2012. HTML5 reached final Recommendation status in late 2014. The W3C continues its development of HTML and is adding more new elements, attributes, and features in HTML5.1 which is currently in draft form.



Checkpoint 1.2

1. Describe the components of the client/server model as applied to the Internet.
2. Identify two protocols used on the Internet to convey information that use the Internet, but do not use the Web.
3. Explain the similarities and differences between a URL and a domain name.

1.9 Popular Uses of the Web

E-Commerce

Continued growth is expected for **e-commerce**, which is the buying and selling of goods and services on the Internet. Forrester Research projects that by 2018, over \$414 billion will be spent on online retail sales (<https://www.forrester.com/US+eCommerce+Forecast+2013+To+2018/fulltext/-/E-RES115513>). With over 3 billion people online worldwide, as reported by Internet World Stats (<http://www.internetworldstats.com/emarketing.htm>), that's quite a few potential shoppers! As mobile web access becomes more commonplace, e-commerce will be regularly done not only from stationary computers, but also from portable devices—tablets, netbooks, smartphones, and technology we haven't even imagined yet.

Mobile Access

Accessing the Web with devices other than the standard desktop, notebook, and netbook computers is a growing trend. Bloomberg Businessweek reported that electronics manufacturers expect a huge demand for **tablet** devices (such as the Apple iPad, Amazon Kindle Fire, and Microsoft Surface) and anticipated \$49 billion in sales by 2015 (<http://buswk.co/fK2Q9e>). A study (<http://stateofthedia.org/2012/mobile-devices-and-news-consumption-some-good-signs-for-journalism/>) by the PEW Research Center determined that many people own more than one mobile Internet-enabled device: over half of people who own a laptop computer also own a smartphone, 23% of owners of laptops also own a tablet, 13% own all three devices. Web designers must consider how their pages will display and function not only on personal computers, but also on smartphones, tablets, and other mobile devices.

Blogs

The trend of keeping a web log, or blog, has been driven by individuals as a forum for personal expression. A **blog** is a journal that is available on the Web; it is a frequently updated page with a chronological list of ideas and links. Blog topics range from politics to technical information to personal diaries. Blogs can focus on one subject or range across a diverse group of topics—it's up to the person, called a blogger, who has created and maintains the blog. Bloggers usually update their blogs daily with easy-to-use software designed to allow people with little or no technical background to update and maintain a blog. Many blogs are hosted at blog communities such as <http://www.wordpress.com>. Others are hosted at individual websites, such as the blog kept by the web designer Eric Meyer at <http://meyerweb.com>. Businesses have noted the value of blogs as communication and customer relationship tools. Companies such as Microsoft (<http://blogs.msdn.com>) and IBM (<https://www.ibm.com/developerworks/mydeveloperworks/blogs>) utilize blogs in this manner.

Wikis

A **wiki** is a website that can be updated immediately at any time by visitors, using a simple form on a web page. Some wikis are intended for a small group of people, such as the members of an organization. The most powerful wiki is Wikipedia (<http://wikipedia.org>), an online encyclopedia, which can be updated by anyone at any time. Wikis are a form of social software in action—visitors sharing their collective knowledge to create a resource freely used by all. While there have been isolated incidents of practical jokes, and inaccurate information has occasionally been posted at Wikipedia, the given information and linked resources are a good starting point when exploring a topic.

Social Networking

Blogs and wikis provide web visitors with new methods for interacting with websites and other people—a use referred to as **social computing**, or **social networking**. A trendy activity these days is participating in a social networking site such as Facebook (<http://www.facebook.com>) or LinkedIn (<http://www.linkedin.com>). Research firm eMarketer predicts that in 2017, 78% of Internet users will access “a social network at least once per month” (<http://www.emarketer.com/Article/Social-Networking-Reaches-Nearly-One-Four-Around-World/1009976>). If it seems to you as if most of your friends are on Facebook, that may

be the case: As of 2015, Facebook had over 1.49 billion monthly active users (<http://newsroom.fb.com/company-info>). While LinkedIn was created with professional and business networking in mind, businesses have also found it useful to create Facebook sites to promote their products and services.

Twitter (<http://www.twitter.com>) is a social networking site for **microblogging**, or frequently communicating with a brief message (140 characters or less) called a **tweet**. Twitter users (called **twitterers**) tweet to update a network of friends and followers about their daily activities and observations. Twitter is not limited to personal use. The business world has also discovered the marketing reach that Twitter can provide. Visit <https://business.twitter.com/basics> for insights on how to use Twitter to promote your business and communicate with customers.

Cloud Computing

Document collaboration sites like Google Drive and Microsoft OneDrive, blogs, wikis, and social networking sites are all accessed via the Internet (the “cloud”) and are examples of **cloud computing**. The National Institute of Standards and Technology (NIST) defines cloud computing as the on-demand use of software and other computing resources hosted at a remote data center (including servers, storage, services, and applications) over the Internet. Expect to see more public and private use of cloud computing in the future.

RSS

Really Simple Syndication, or Rich Site Summary (RSS) is used to create newsfeeds from blog postings and other websites. The RSS feeds contain a summary of new items posted to the site. The URL to the RSS feed is usually indicated by the letters XML or RSS in white text within an orange rectangle. A **newsreader** is needed to access the information. Some browsers, such as Firefox, Safari, and Internet Explorer, can display RSS feeds. Commercial and shareware newsreader applications are also available. The newsreader polls the feed URL at intervals and displays the new headlines when requested. RSS provides web developers with a method to push new content to interested parties and (hopefully) generate return visits to the site.

Podcasts

Podcasts are audio files on the web that take the format of an audio blog, radio show, or interview. Podcasts are typically delivered by an RSS feed, but can also be made available by providing the link to a recorded MP3 file on a web page. These files can be saved to your computer or to an MP3 player (such as an iPod) for later listening.

Web 2.0

Flickr (<http://www.flickr.com>) and Pinterest (<http://pinterest.com>) are two social software sites that provide information-sharing opportunities. Flickr, a photo-sharing site, calls itself the “best way to store, search, sort, and share your photos.” Pinterest describes itself as “a place to discover ideas for all your projects and interests, handpicked by people like you.” Pinterest users “pin” photos from websites they have found useful or intriguing on “boards” that are organized by categories. Users can search for pins, follow other users, and follow boards. Pins of interest are displayed. Websites such as Wikipedia, Flickr, Twitter, and Pinterest are examples of what is called **Web 2.0**. While a consensus on the

definition of Web 2.0 has yet to be reached, think of it as the next step in the transition of the Web from isolated static websites to a platform that utilizes technology to provide rich interfaces and social networking opportunities. Visit <http://www.go2web20.net> and use the search engine to find Web 2.0 sites. You may also read Tim O'Reilly's informative Web 2.0 essay at <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html> for more information on this developing topic.

The single trend that you can expect to remain the same for the foreseeable future is the trend of constant change. Internet and web-related technologies are in a constant state of development and improvement. If constant change and the opportunity to learn something new excite you, web development is a fascinating field. The skills and knowledge you gain in this book should provide a solid foundation for your future learning.



FAQ What is the next big thing on the Web?

The Web is changing by the minute. Check the textbook's companion website at <http://www.pearsonglobaleditions.com/felke-morris> for a blog that will help you stay current about web trends.

Sample pages