CHAPTER 2

TECHNOLOGY INFRASTRUCTURE: THE INTERNET AND THE WORLD WIDE WEB

LEARNING OBJECTIVES

In this chapter, you will learn about:

- The origin, growth, and current structure of the Internet
- How packet-switched networks are combined to form the Internet
- How Internet protocols and Internet addressing work
- The history and use of markup languages on the Web, including SGML, HTML, and XML
- How HTML tags and links work on the World Wide Web
- The differences among internets, intranets, and extranets
- Options for connecting to the Internet, including cost and bandwidth factors
- Internet2 and the Semantic Web

INTRODUCTION

Many business executives made the statement "the Internet changes everything" during the late 1990s. One of the first people to say those words publicly was John Chambers, CEO of **Cisco Systems**, in a speech at a computer industry trade show in 1996. For his company, the Internet did indeed change everything. Cisco, founded in 1984, grew rapidly to become one of the largest and most profitable companies in the world by 2000.

Cisco designs, manufactures, and sells computer networking devices. In this chapter, you will learn about these devices and how they make up the Internet. Cisco's earnings grew as telecommunications companies purchased the company's products to build the infrastructure of the Internet. Other companies also wanted to connect their business operations to the Internet; they became lucrative customers for Cisco, too. In its fiscal year ended July 2000, Cisco had sales of $19 billion and net income of $3 billion. Cisco was one of the true winners in the first wave of electronic commerce.

Because Cisco grew so rapidly during the first wave, it developed a strategy in which it built equipment before it received orders from customers. Cisco did not want to run short of equipment that customers were eager to get. In 2000, many of Cisco's most important telecommunications customers suddenly stopped expanding. Some of them even went out of business. Demand for Cisco's networking devices plummeted. The equipment that Cisco had built with the expectation of ever-increasing orders sat in its warehouses, unwanted. In 2001, Cisco posted a loss of $1 billion on sales of $22 billion. Investors drove Cisco's stock, which had been trading at $80 per share, to $14 per share.

Chambers immediately undertook a series of belt-tightening steps. Between 2001 and 2004, Cisco wrote off $2 billion in inventory that it could not sell and laid off more than 10,000 of its 35,000 employees. The company also increased the efficiency of its operations by eliminating 20 percent of its 50 different product lines and reducing its number of suppliers by 60 percent. More than 75 percent of all corporate networks use Cisco equipment, and Chambers knew that market position would be important when business picked up again. All he needed to do was to keep the company operating through the difficult times the industry was facing.

Despite the sudden drop in demand for Cisco's and other telecommunications companies' products, the number of people using the Internet continued to increase rapidly. Businesses were finding new uses for the Internet infrastructure that the telecommunications companies had built, especially now that those financially troubled telecommunications companies were selling access to their infrastructure at very low prices. Five years after reporting the $1 billion loss, Cisco was benefiting from the second wave of electronic commerce: its sales had recovered and were growing again, although at a slower rate. Cisco had weathered the storm and was ready once again to make money by supplying technologies to a world in which the Internet had changed everything.

THE INTERNET AND THE WORLD WIDE WEB

A **computer network** is any technology that allows people to connect computers to each other. Computer networks and the Internet, which connects computer networks around the world to one another, form the basic technology structure that underlies all electronic commerce. This chapter introduces you to the hardware and software technologies that make electronic commerce possible. First, you will learn how the Internet and the World Wide Web work. Then, you will learn about other technologies that support the Internet, the Web, and electronic commerce. In this chapter, you will be introduced to several complex networking technologies. If you are interested in learning more about how computer networks operate, you can consult one of the computer networking books cited in the For Further Study and Research section at the end of this chapter, or you can take courses in data communications and networking.

Millions of people use the Internet every day, but only a small percentage of them really understand how it works. The **Internet** is a large system of interconnected computer networks that spans the globe. Using the Internet, you can communicate with other people throughout the world by means of electronic mail; read online versions of newspapers, magazines, academic journals, and books; join discussion groups on almost any conceivable topic; participate in games and simulations; and obtain free computer software.

In recent years, the Internet has allowed commercial enterprises to connect with one another and with customers. Today, all kinds of businesses provide information about their products and services on the Internet. Many of these businesses use the Internet to market and sell their products and services.

The part of the Internet known as the **World Wide Web**, or, more simply, the **Web**, is a subset of the computers on the Internet that are connected to one another in a specific way that makes them and their contents easily accessible to each other. The most important thing about the Web is that it includes an easy-to-use standard interface. This interface makes it possible for people who are not computer experts to use the Web to access a variety of Internet resources.

Origins of the Internet

In the early 1960s, the U.S. Department of Defense became concerned about the possible effects of nuclear attack on its computing facilities. The Defense Department realized that the weapons of the future would require powerful computers for coordination and control. The powerful computers of that time were all large mainframe computers, so the Defense Department began examining ways to connect these computers to each other and also connect them to weapons installations distributed all over the world. The Defense Department agency charged with this task hired many of the best communications technology researchers and, for many years, funded research at leading universities and institutes to explore the task of creating a worldwide network that could remain operational, even if parts of the network were destroyed by enemy military action or sabotage. These researchers worked to devise ways to build networks that could operate independently—that is, networks that did not require a central computer to control network operations.

Early computer networks used leased telephone company lines for their connections. Telephone company systems of that time established a single connection between sender and receiver for each telephone call, and that connection carried all data along a single path. When a company wanted to connect computers it owned at two different locations, the company placed a telephone call to establish the connection, and then connected one computer to each end of that single connection.

The Defense Department was concerned about the inherent risk of this single-channel method for connecting computers, and its researchers developed a different method of sending information through multiple channels. In this method, files and messages are broken into packets that are labeled electronically with codes for their origins, sequences, and destinations. You will learn more about how packet networks operate later in this chapter.

In 1969, Defense Department researchers in the Advanced Research Projects Agency (ARPA) used this network model to connect four computers—one each at the University of California at Los Angeles, SRI International, the University of California at Santa Barbara, and the University of Utah—into a network called the ARPANET. The ARPANET was the earliest of the networks that eventually combined to become what we now call the Internet. Throughout the 1970s and 1980s, many researchers in the academic community connected to the ARPANET and contributed to the technological developments that increased its speed and efficiency. At the same time, researchers at other universities were creating their own networks using similar technologies.

New Uses for the Internet

Although the goals of the Defense Department network were to control weapons systems and transfer research files, other uses for this vast network began to appear in the early 1970s. E-mail was born in 1972 when Ray Tomlinson, a researcher who used the network, wrote a program that could send and receive messages over the network. This new method of communicating became widely used very quickly. The number of network users in the military and education research communities continued to grow. Many of these new participants used the networking technology to transfer files and access computers remotely. You will learn about these file transfer tools in Chapter 8.

The first e-mail mailing lists also appeared on these networks. A **mailing list** is an e-mail address that forwards any message it receives to any user who has subscribed to the list. In 1979, a group of students and programmers at Duke University and the University of North Carolina started **Usenet**, an abbreviation for **User's News Network**. Usenet allows anyone who connects to the network to read and post articles on a variety of subjects. Usenet survives on the Internet today, with more than 1000 different topic areas that are called **newsgroups**. Other researchers even created game-playing software for use on these interconnected networks.

Although the people using these networks were developing many creative applications, use of the networks was limited to those members of the research and academic communities who could access them. Between 1979 and 1989, these network applications were improved and tested by an increasing number of users. The Defense Department's networking software became more widely used in academic and research institutions as these organizations recognized the benefits of having a common communications network. As the number of people in different organizations using these networks increased, security problems were recognized. These problems have continued to become more important, and you will learn more about network security issues in Chapter 10. The explosion of personal computer use during the 1980s also helped more people become comfortable with computers. In the late 1980s, these independent academic and research networks merged into what we now call the Internet.

Commercial Use of the Internet

As personal computers became more powerful, affordable, and available during the 1980s, companies increasingly used them to construct their own internal networks. Although these networks included e-mail software that employees could use to send messages to each other, businesses wanted their employees to be able to communicate with people outside their corporate networks. The Defense Department network and most of the academic networks that had teamed up with it were receiving funding from the **National Science Foundation (NSF)**. The NSF prohibited commercial network traffic on its networks, so businesses turned to commercial e-mail service providers to handle their e-mail needs. Larger firms built their own networks that used leased telephone lines to connect field offices to corporate headquarters.

In 1989, the NSF permitted two commercial e-mail services, MCI Mail and CompuServe, to establish limited connections to the Internet for the sole purpose of exchanging e-mail transmissions with users of the Internet. These connections allowed commercial enterprises to send e-mail directly to Internet addresses, and allowed members of the research and education communities on the Internet to send e-mail directly to MCI Mail and CompuServe addresses. The NSF justified this limited commercial use of the Internet as a service that would primarily benefit the Internet's noncommercial users. As the 1990s began, people from all walks of life—not just scientists or academic researchers—started thinking of these networks as the global resource that we now know as the Internet. Although this network of networks had grown from four Defense Department computers in 1969 to more than 300,000 computers on many interconnected networks by 1990, the greatest growth of the Internet was yet to come.

Growth of the Internet

In 1991, the NSF further eased its restrictions on commercial Internet activity and began implementing plans to privatize the Internet. The privatization of the Internet was substantially completed in 1995, when the NSF turned over the operation of the main Internet connections to a group of privately owned companies. The new structure of the Internet was based on four **network access points (NAPs)** located in San Francisco, New York, Chicago, and Washington, D.C., each operated by a separate telecommunications company. As the Internet grew, more companies opened more NAPs in more locations. These companies, known as **network access providers**, sell Internet access rights directly to larger customers and indirectly to smaller firms and individuals through other companies, called **Internet service providers (ISPs)**.

The Internet was a phenomenon that had truly sneaked up on an unsuspecting world. The researchers who had been so involved in the creation and growth of the Internet just accepted it as part of their working environment. However, people outside the research community were largely unaware of the potential offered by a large interconnected set of computer networks. Figure 2-1 shows the consistent and dramatic growth in the number of **Internet hosts**, which are computers directly connected to the Internet.

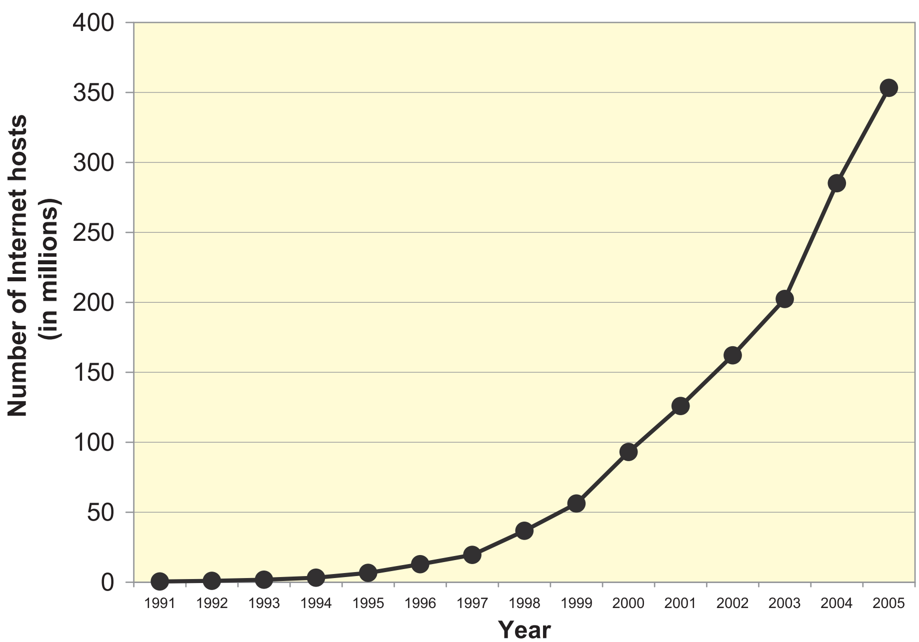

FIGURE 2-1 Growth of the Internet

In 30 years, the Internet has grown to become one of the most amazing technological and social accomplishments of the last century. Millions of people, most of whom are not computer researchers or experts, now use this complex, interconnected network of computers. These computers run thousands of different software packages. The computers are located in almost every country of the world. Every year, billions of dollars change hands over the Internet in exchange for all kinds of products and services. All of this activity occurs with no central coordination point or control, which is especially ironic given that the Internet began as a way for the military to maintain control while under attack.

The opening of the Internet to business activity helped dramatically increase its growth; however, there was another development that worked hand in hand with the commercialization of the Internet to spur its growth. That development was the World Wide Web.

Emergence of the World Wide Web

The Web is software that runs on computers that are connected to the Internet. The network traffic generated by Web software is the largest single category of traffic on the Internet today, outpacing e-mail, file transfers, and other data transmission traffic. But the Web is more a way of thinking about and organizing information storage and retrieval than it is a specific technology. As such, its history goes back many years. Two important innovations that became key elements of the Web are hypertext and graphical user interfaces.

The Development of Hypertext

In 1945, **Vannevar Bush**, who was director of the U.S. Office of Scientific Research and Development, wrote an article in *The Atlantic Monthly* about ways that scientists could apply the skills they learned during World War II to peacetime activities. The article included a number of visionary ideas about future uses of technology to organize and facilitate efficient access to information. Bush speculated that engineers would eventually build a machine that he called the Memex, a memory extension device that would store all of a person's books, records, letters, and research results on microfilm. Bush's Memex would include mechanical aids, such as microfilm readers and indexes, that would help users quickly and flexibly consult their collected knowledge.

In the 1960s, Ted Nelson described a similar system in which text on one page links to text on other pages. Nelson called his page-linking system **hypertext**. Douglas Engelbart, who also invented the computer mouse, created the first experimental hypertext system on one of the large computers of the 1960s. In 1987, Nelson published *Literary Machines*, a book in which he outlined project Xanadu, a global system for online hypertext publishing and commerce. Nelson used the term *hypertext* to describe a page-linking system that would interconnect related pages of information, regardless of where in the world they were stored.

In 1989, Tim Berners-Lee was trying to improve the laboratory research document-handling procedures for his employer, CERN: European Laboratory for Particle Physics. CERN had been connected to the Internet for two years, but its scientists wanted to find better ways to circulate their scientific papers and data among the high-energy physics research community throughout the world. Berners-Lee proposed a hypertext development project intended to provide this data-sharing functionality.

Over the next two years, Berners-Lee developed the code for a hypertext server program and made it available on the Internet. A **hypertext server** is a computer that stores files written in the hypertext markup language and lets other computers connect to it and read these files. Hypertext servers used on the Web today are usually called **Web servers**.

Hypertext Markup Language (HTML), which Berners-Lee developed from his original hypertext server program, is a language that includes a set of codes (or tags) attached to text. These codes describe the relationships among text elements. For example, HTML includes tags that indicate which text is part of a header element, which text is part of a paragraph element, and which text is part of a numbered list element. One important type of tag is the hypertext link tag. A **hypertext link**, or **hyperlink**, points to another location in the same or another HTML document.

Graphical Interfaces for Hypertext

Several different types of software are available to read HTML documents, but most people use a Web browser such as Netscape Navigator or Microsoft Internet Explorer. A **Web browser** is a software interface that lets users read (or browse) HTML documents and move from one HTML document to another through text formatted with hypertext link tags in each file. If the HTML documents are on computers connected to the Internet, you can use a Web browser to move from an HTML document on one computer to an HTML document on any other computer on the Internet.

An HTML document differs from a word-processing document in that it does not specify how a particular text element will appear. For example, you might use word-processing software to create a document heading by setting the heading text font to Arial, its font size to 14 points, and its position to centered. The document displays and prints these exact settings whenever you open the document in that word processor. In contrast, an HTML document simply includes a heading tag with the heading text. Many different browser programs can read an HTML document. Each program recognizes the heading tag and displays the text in whatever manner each program normally displays headings. Different Web browser programs might each display the text differently, but all of them display the text with the characteristics of a heading.

A Web browser presents an HTML document in an easy-to-read format in the browser's graphical user interface. A **graphical user interface (GUI)** is a way of presenting program control functions and program output to users. It uses pictures, icons, and other graphical elements instead of displaying just text. Almost all personal computers today use a GUI such as Microsoft Windows or the Macintosh user interface.

The World Wide Web

Berners-Lee called his system of hyperlinked HTML documents the World Wide Web. The Web caught on quickly in the scientific research community, but few people outside that community had software that could read the HTML documents. In 1993, a group of students led by Marc Andreessen at the University of Illinois wrote Mosaic, the first GUI program that could read HTML and use HTML hyperlinks to navigate from page to page on computers anywhere on the Internet. Mosaic was the first Web browser that became widely available for personal computers, and some Web surfers still use it today.

Programmers quickly realized that a functional system of pages connected by hypertext links would provide many new Internet users with an easy way to access information on the Internet. Businesses recognized the profit-making potential offered by a worldwide network of easy-to-use computers. In 1994, Andreessen and other members of the University of Illinois Mosaic team joined with James Clark of **Silicon Graphics** to found Netscape

Communications (which is now owned by **Time Warner**). Its first product, the Netscape Navigator Web browser program based on Mosaic, was an instant success. Netscape became one of the fastest-growing software companies ever. **Microsoft** created its Internet Explorer Web browser and entered the market soon after Netscape's success became apparent. A number of other Web browsers exist, but Internet Explorer dominates the market today.

The number of Web sites has grown even more rapidly than the Internet itself. The number of Web sites is currently estimated at more than 75 million, and individual Web pages number more than 20 billion because each Web site might include hundreds or even thousands of individual Web pages. Therefore, nobody really knows how many Web pages exist. For example, researchers at **BrightPlanet** estimate that the number of Web sites could be more than 500 million. Figure 2-2 shows how the growth rate of the Web increased dramatically between 1997 and 2000. After a brief consolidation period during 2001–2002, the Web is once again showing rapid growth.

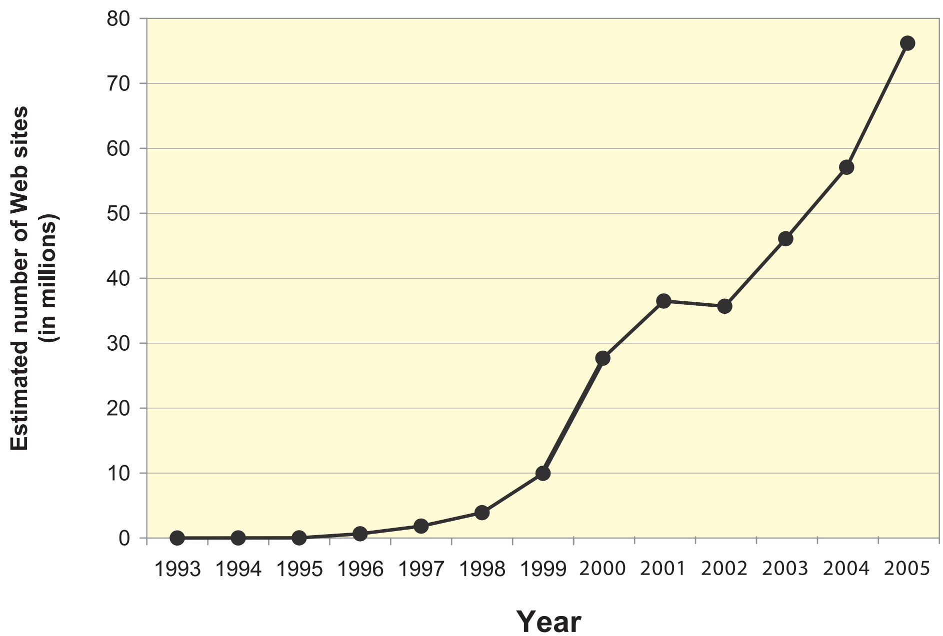

FIGURE 2-2 Growth of the World Wide Web

As more people gain access to the Web, commercial interest in using the Web to conduct business will continue to increase, and the variety of nonbusiness uses will become even greater. In the rest of this chapter, you will learn how Internet and Web technologies work to enable electronic commerce.

PACKET-SWITCHED NETWORKS

A network of computers that are located close together—for example, in the same building—is called a **local area network**, or a **LAN**. Networks of computers that are connected over greater distances are called **wide area networks**, or **WANs**.

The early models (dating back to the 1950s) for WANs were the circuits of the local and long-distance telephone companies of the time, because the first early WANs used leased telephone company lines for their connections. A telephone call establishes a single connection path between the caller and receiver. Once that connection is established, data travels along that single path. Telephone company equipment (originally mechanical, now electronic) selects specific telephone lines to connect to one another by closing switches. These switches work like the switches you use to turn lights on and off in your home, except that they open and close much faster, and are controlled by mechanical or electronic devices instead of human hands.

The combination of telephone lines and the closed switches that connect them to each other is called a **circuit**. This circuit forms a single electrical path between caller and receiver. This single path of connected circuits switched into each other is maintained for the entire length of the call. This type of centrally controlled, single-connection model is known as **circuit switching**.

Although circuit switching works well for telephone calls, it does not work as well for sending data across a large WAN or an interconnected network like the Internet. The Internet was designed to be resistant to failure. In a circuit-switched network, a failure in any one of the connected circuits causes the connection to be interrupted and data to be lost. Instead, the Internet uses packet switching to move data between two points. On a **packet-switched** network, files and e-mail messages are broken down into small pieces, called **packets**, that are labeled electronically with their origins, sequences, and destination addresses. Packets travel from computer to computer along the interconnected networks until they reach their destinations. Each packet can take a different path through the interconnected networks, and the packets may arrive out of order. The destination computer collects the packets and reassembles the original file or e-mail message from the pieces in each packet.

Routing Packets

As an individual packet travels from one network to another, the computers through which the packet travels determine the best route for getting the packet to its destination. The computers that decide how best to forward each packet are called **routing computers**, **router computers**, **routers**, **gateway computers** (because they act as the gateway from a LAN or WAN to the Internet), or **border routers** (because they are located at the border between the organization and the Internet). The programs on router computers that determine the best path on which to send each packet contain rules called **routing algorithms**. The programs apply their routing algorithms to information they have stored in **routing tables** or **configuration tables**. This information includes lists of connections that lead to particular groups of other routers, rules that specify which connections to use first, and rules for handling instances of heavy packet traffic and network congestion.

Individual LANs and WANs can use a variety of different rules and standards for creating packets within their networks. The network devices that move packets from one part of a network to another are called hubs, switches, and bridges. Routers are used to connect networks to other networks. As technologies have improved, many of the distinctions between these different types of network devices have become blurred. You can take a data communications and networking class to learn more about these network devices and how they work.

When packets leave a network to travel on the Internet, they must be translated into a standard format. Routers usually perform this translation function. As you can see, routers are an important part of the infrastructure of the Internet. When a company or organization becomes part of the Internet, it must connect at least one router to the other routers (owned by other companies or organizations) that make up the Internet. Figure 2-3 is a diagram of a small portion of the Internet that shows its router-based architecture. The figure shows only the routers that connect each organization's WANs and LANs to the Internet, not the other routers that are inside the WANs and LANs or that connect them to each other within the organization.

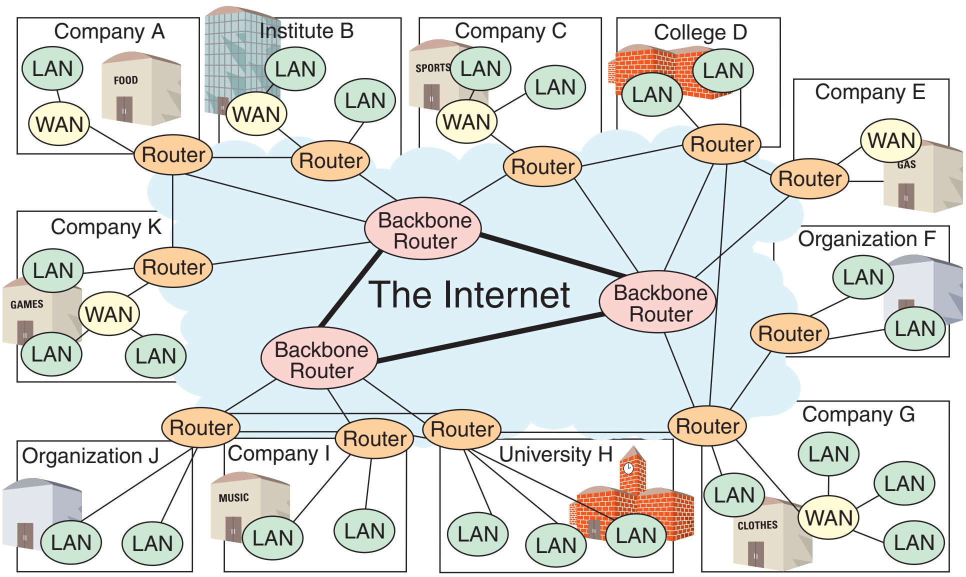

FIGURE 2-3 Router-based architecture of the Internet

The Internet also has routers that handle packet traffic along the Internet's main connecting points. These routers and the telecommunications lines connecting them are collectively referred to as the **Internet backbone**. These routers, sometimes called **backbone routers**, are very large computers that can each handle more than 50 million packets per second! You can see in Figure 2-3 that a router connected to the Internet always has more than one path to which it can direct a packet. By building in multiple packet paths, the designers of the Internet created a degree of redundancy in the system that allows it to keep moving packets, even if one or more of the routers or connecting lines fails.

INTERNET PROTOCOLS

As you learned earlier in this chapter, the first packet-switched network, the ARPANET, connected only a few universities and research centers. This experimental network grew during the next few years and used the **Network Control Protocol (NCP)**. A **protocol** is a collection of rules for formatting, ordering, and error-checking data sent across a network. For example, protocols determine how the sending device indicates that it has finished sending a message, and how the receiving device indicates that it has received (or not received) the message. A protocol also includes rules about what is allowed in a transmission and how it is formatted. Computers that communicate with each other must use the same protocol for data transmission. In the early days of computing, each computer manufacturer created its own protocol, so computers made by different manufacturers could not be connected to each other. This practice was called **proprietary architecture** or **closed architecture**. The **open architecture** philosophy developed for the evolving ARPANET, which later became the core of the Internet, included the use of a common protocol for all computers connected to the Internet and four key rules for message handling:

- Independent networks should not require any internal changes to be connected to the network.
- Packets that do not arrive at their destinations must be retransmitted from their source network.
- Router computers act as receive-and-forward devices; they do not retain information about the packets that they handle.
- No global control exists over the network.

The open architecture approach has contributed to the success of the Internet because computers manufactured by different companies (Apple, Dell, Hewlett-Packard, Sun, etc.) can be interconnected. The ARPANET and its successor, the Internet, use routers to isolate each LAN or WAN from the other networks to which they are connected. Each LAN or WAN can use its own set of protocols for packet traffic within the LAN or WAN, but must use a router (or similar device) to move packets onto the Internet in its standard format (or protocol). Following these simple rules makes the connections between the interconnected networks operate effectively.

TCP/IP

The Internet uses two main protocols: the **Transmission Control Protocol (TCP)** and the **Internet Protocol (IP)**. Developed by Internet pioneers Vinton Cerf and Robert Kahn, these protocols are the rules that govern how data moves through the Internet and how network connections are established and terminated. The acronym **TCP/IP** is commonly used to refer to the two protocols.

The TCP controls the disassembly of a message or a file into packets before it is transmitted over the Internet, and it controls the reassembly of those packets into their original formats when they reach their destinations. The IP specifies the addressing details for each packet, labeling each with the packet's origination and destination addresses. Soon after the new TCP/IP protocol set was developed, it replaced the NCP that ARPANET originally used.

In addition to its Internet function, TCP/IP is used today in many LANs. The TCP/IP protocol is provided in most personal computer operating systems commonly used today, including Linux, Macintosh, Microsoft Windows, and UNIX.

IP Addressing

The version of IP that has been in use for the past 20 years on the Internet is **Internet Protocol version 4**, abbreviated **IPv4**. It uses a 32-bit number to identify the computers connected to the Internet. This address is called an **IP address**. Computers do all of their internal calculations using a **base 2** (or **binary**) number system in which each digit is either a 0 or a 1, corresponding to a condition of either off or on. IPv4 uses a 32-bit binary number that allows more than 4 billion different addresses ($2^{32} = 4{,}294{,}967{,}296$).

When a router breaks a message into packets before sending it onto the Internet, the router marks each packet with both the source IP address and the destination IP address of the message. To make them easier to read, IP numbers (addresses) appear as four numbers separated by periods. This notation system is called **dotted decimal** notation. An IPv4 address is a 32-bit number, so each of the four numbers is an 8-bit number ($4 \times 8 = 32$). In most computer applications, an 8-bit number is called a **byte**; however, in networking applications, an 8-bit number is often called an **octet**. In binary, an octet can have values from 00000000 to 11111111; the decimal equivalents of these binary numbers are 0 and 255, respectively.

Because each of the four parts of a dotted decimal number can range from 0 to 255, IP addresses range from 0.0.0.0 (written in binary as 32 zeros) to 255.255.255.255 (written in binary as 32 ones). Although some people find dotted decimal notation to be confusing at first, most do agree that writing, reading, and remembering a computer's address as 216.115.108.245 is easier than 11011000011100110110110011110101, or its full decimal equivalent, which is 3,631,433,189.

Today, IP addresses are assigned by three not-for-profit organizations: the **American Registry for Internet Numbers (ARIN)**, the **Reséaux IP Européens (RIPE)**, and the **Asia-Pacific Network Information Center (APNIC)**. These registries assign and manage IP addresses for various parts of the world: ARIN for North America, South America, the Caribbean, and sub-Saharan Africa; RIPE for Europe, the Middle East, and the rest of Africa; and APNIC for countries in the Asia-Pacific area. These organizations took over IP address management tasks from the Internet Assigned Numbers Authority (IANA), which performed them under contract with the U.S. government when the Internet was an experimental research project.

You can use the **ARIN Whois** page at the ARIN Web site to search the IP addresses owned by organizations in North America. You can enter an organization name into the search box on the page, then click the Submit Query button, and the Whois server returns a list of the IP addresses owned by that organization. For example, performing a search on the word *Carnegie* displays the IP address blocks owned by Carnegie Bank, Carnegie Mellon University, and a number of other organizations whose names begin with Carnegie. You can also enter an IP address and find out who owns that IP address. If you enter "3.0.0.0" (without the quotation marks), you will find that General Electric owns the entire block of IP addresses from 3.0.0.0 to 3.255.255.255. General Electric can use these addresses, which number approximately 16.7 million, for its own computers or it can lease them to other companies or individuals to whom it provides Internet access services.

In the early days of the Internet, the 4 billion addresses provided by the IPv4 rules certainly seemed to be more addresses than an experimental research network would ever need. However, about 2 billion of those addresses today are either in use or unavailable for use because of the way blocks of addresses were assigned to organizations. The new kinds of devices on the Internet's many networks, such as wireless personal digital assistants and cell phones that can access the Web, promise to keep demand high for IP addresses.

Network engineers have devised a number of stopgap techniques to stretch the supply of IP addresses. One of the most popular techniques is **subnetting**, which is the use of reserved private IP addresses within LANs and WANs to provide additional address space. **Private IP addresses** are a series of IP numbers that are not permitted on packets that travel on the Internet. In subnetting, a computer called a **Network Address Translation (NAT) device** converts those private IP addresses into normal IP addresses when it forwards packets from those computers to the Internet.

The **Internet Engineering Task Force (IETF)** worked on several new protocols that could solve the limited addressing capacity of IPv4, and in 1997, approved **Internet Protocol version 6 (IPv6)** as the protocol that will replace IPv4. The new IP is being implemented gradually because the two protocols are not directly compatible. The process of switching over to IPv6 will take at least another 10 years; however, network engineers have devised ways to run both protocols together on interconnected networks. The major advantage of IPv6 is that it uses a 128-bit number for addresses instead of the 32-bit number used in IPv4. The number of available addresses in IPv6 (2^{128}) is 34 followed by 37 zeros—billions of times larger than the address space of IPv4. The new IP also changes the format of the packet itself. Improvements in networking technologies over the past 20 years have made many of the fields in the IPv4 packet unnecessary. IPv6 eliminates those fields and adds fields for security and other optional information.

IPv6 has a shorthand notation system for expressing addresses, similar to the IPv4 dotted decimal notation system. However, because the IPv6 address space is much larger, its notation system is more complex. The IPv6 notation uses eight groups of 16 bits ($8 \times 16 = 128$). Each group is expressed as four hexadecimal digits and the groups are separated by colons; thus, the notation system is called **colon hexadecimal** or **colon hex**. A **hexadecimal (base 16)** numbering system uses 16 digits (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, a, b, c, d, e, and f). An example of an IPv6 address expressed in this notation is: CD18:0000:0000:AF23:0000:FF9E:61B2:884D. To save space, the zeros can be omitted, which reduces this address to: CD18:::AF23::FF9E:61B2:884D.

Domain Names

The founders of the Internet were concerned that users might find the dotted decimal notation difficult to remember. To make the numbering system easier to use, they created an alternative addressing method that uses words. In this system, an address such as www.thomson.com is called a domain name. **Domain names** are sets of words that are assigned to specific IP addresses. Domain names can contain two or more word groups separated by periods. The rightmost part of a domain name is the most general. Each part of the domain name becomes more specific as you move to the left.

For example, the domain name www.sandiego.edu contains three parts separated by periods. Beginning at the right, the name "edu" indicates that the computer belongs to a four-year educational institution. The institution, University of San Diego, is identified by the name "sandiego." The "www" indicates that the computer is running software that makes it a part of the World Wide Web. Most, but not all, Web addresses follow this "www" naming convention. For an example of an exception, the group of computers that operate the **Yahoo! Games** service is named games.yahoo.com.

The rightmost part of a domain name is called a **top-level domain** (or **TLD**). For many years, these domains have included a group of general domains—such as .edu, .com, and .org—and a set of country domains. Since 1998, the **Internet Corporation for Assigned Names and Numbers (ICANN)** has had responsibility for managing domain names and coordinating them with the IP address registrars. ICANN is also responsible for setting standards for the router computers that make up the Internet. In 2000, ICANN added seven new TLDs. Four of these TLDs (.biz, .info, .name, and .pro) are general domains, the other three (.aero, .coop, and .museum) are sponsored domains. A **sponsored top-level domain** (**sTLD**) is a TLD for which an organization other than ICANN is responsible. The sponsor of a specific sTLD must be a recognized institution that has expertise regarding and is familiar with the community that uses the sTLD. For example, the .aero sTLD is sponsored by SITA, an air transport industry association that has expertise in and is familiar with airlines, airports, and the aerospace industry. Although these new domain names were chosen after much deliberation and consideration of more than 100 possible new names, many people were highly critical of the selections (see, for example, the **ICANNWatch** Web site).

In 2002, ICANN came under additional fire for acting in ways that many people thought violated the democratic principles on which the organization was founded. In 2005, ICANN approved a new sTLD for adult content sites (.xxx) as a part of its normal deliberations on 10 proposed new sTLDs. The U.S. Commerce Department, responding to pressure from conservative political groups, ordered ICANN to delay implementing the .xxx domain. Many observers believe that this intervention has seriously damaged the independence of ICANN. You can learn more about these issues on the Web sites of the **Internet Governance Project** and the **Convergence Center at Syracuse University**. Increases in the number of TLDs can make it more difficult for companies to protect their corporate and product brand names. You will learn more about these issues in Chapter 7. Figure 2-4 presents a list of the general TLDs, including the 2000 additions, and some of the more frequently used country TLDs.

Original general TLDs		Country TLDs		General TLDs added in 2000	
TLD	**Use**	**TLD**	**Country**	**TLD**	**Use**
.com	Commercial	.au	Australia	.biz	Businesses
.edu	Four-year educational institution	.ca	Canada	.info	General use
.gov	U.S. federal government	.de	Germany	.name	Individual people
.mil	U.S. military	.fi	Finland	.pro	Professionals (accountants, lawyers, physicians)
.net	General use	.fr	France		
.org	Not-for-profit organization	.jp	Japan		
		.se	Sweden		
		.uk	United Kingdom		

FIGURE 2-4 Top-level domain names

Web Page Request and Delivery Protocols

The Web is software that runs on computers that are connected to each other through the Internet. **Web client computers** run software called **Web client software** or **Web browser software**. Web client software sends requests for Web page files to other computers, which are called Web servers. A Web server computer runs software called **Web server software**. Web server software receives requests from many different Web clients and responds by sending files back to those Web client computers. Each Web client computer's Web client software renders those files into a Web page. Thus, the purpose of a Web server is to respond to requests for Web pages from Web clients. This combination of client computers running Web client software and server computers running Web server software is called a **client/server architecture**.

The set of rules for delivering Web page files over the Internet is in a protocol called the **Hypertext Transfer Protocol (HTTP)**, which was developed by Tim Berners-Lee in 1991. When a user types a domain name (for example, www.yahoo.com) into a Web browser's address bar, the browser sends an HTTP-formatted message to a Web server computer at Yahoo! that stores Web page files. The Web server computer at Yahoo! then responds by sending a set of files (one for the Web page and one for each graphic object, sound, or video clip included on the page) back to the client computer. These files are sent within a message that is HTTP formatted.

To initiate a Web page request using a Web browser, the user types the name of the protocol, followed by the characters "//:" before the domain name. Thus, a user would type http://www.yahoo.com to go to the Yahoo! Web site. Most Web browsers today automatically insert the http:// if the user does not include it. The combination of the protocol name and the domain name is called a **Uniform Resource Locator (URL)** because it lets the user locate a resource (the Web page) on another computer (the Web server).

Electronic Mail Protocols

Electronic mail, or **e-mail**, that is sent across the Internet must also be formatted according to a common set of rules. Most organizations use a client/server structure to handle e-mail. The organization has a computer called an **e-mail server** that is devoted to handling e-mail. The software on that computer stores and forwards e-mail messages. People in the organization might use a variety of programs, called **e-mail client software**, to read and send e-mail. These programs include **Microsoft Outlook**, **Mozilla Thunderbird**, **Netscape Messenger**, **Pegasus Mail**, **Qualcomm Eudora**, and many others. The e-mail client software communicates with the e-mail server software on the e-mail server computer to send and receive e-mail messages.

Many people also use e-mail on their computers at home. In most cases, the e-mail servers that handle their messages are operated by the companies that provide their connections to the Internet. An increasing number of people use e-mail services that are offered by Web sites such as **Yahoo! Mail** or **Hotmail**. In these cases, the e-mail servers and the e-mail clients are operated by the owners of the Web sites. The individual users only see the e-mail client software (and not the e-mail server software) in their Web browsers when they log on to the Web mail service.

With so many different e-mail client and server software choices, standardization and rules are very important. If e-mail messages did not follow standard rules, an e-mail message created by a person using one e-mail client program could not be read by a person using a different e-mail client program. As you have already learned in this chapter, rules for computer data transmission are called protocols.

SMTP and POP are two common protocols used for sending and retrieving e-mail. **Simple Mail Transfer Protocol (SMTP)** specifies the format of a mail message and describes how mail is to be administered on the e-mail server and transmitted on the Internet. An e-mail client program running on a user's computer can request mail from the organization's e-mail server using the **Post Office Protocol (POP)**. A POP message can tell the e-mail server to send mail to the user's computer and delete it from the e-mail server; send mail to the user's computer and not delete it; or simply ask whether new mail has arrived. The POP provides support for **Multipurpose Internet Mail Extensions (MIME)**, which is a set of rules for handling binary files, such as word-processing documents, spreadsheets, photos, or sound clips, that are attached to e-mail messages.

IMAP, the **Interactive Mail Access Protocol**, is a newer e-mail protocol that performs the same basic functions as POP, but includes additional features. For example, IMAP can instruct the e-mail server to send only selected e-mail messages to the client instead of all messages. IMAP also allows the user to view only the header and the e-mail sender's name before deciding to download the entire message. POP allows users to search for and manipulate only those e-mail messages that they have downloaded to their computers. IMAP lets users create and manipulate mail folders (also called mailboxes), delete messages, and search for certain parts of a message while the e-mail is still on the e-mail server; that is, the user does not need to download the e-mail before working with it. The tools that IMAP provides are important to the increasing number of people who access their e-mail from different computers at different times. IMAP lets users manipulate and store their e-mail on the e-mail server and access it from any number of computers. POP allows users to access new e-mail messages from only one PC after they download the old messages to another PC. The main drawback to IMAP is that users' e-mail messages are stored on the e-mail server. As the number of users increases, the size of the e-mail server's disk drives must also increase. In general, server computers use faster (and thus, more expensive) disk drives than desktop computers. Therefore, it is more expensive to provide disk storage space for large quantities of e-mail on a server computer than to provide that same disk space on users' desktop computers. You can learn more about IMAP at the University of Washington's **IMAP Connection** Web site.

Unsolicited Commercial E-Mail (UCE, Spam)

Spam, also known as **unsolicited commercial e-mail (UCE)** or **bulk mail**, is electronic junk mail and can include solicitations, advertisements, or e-mail chain letters. The origin of the term spam is generally believed to have come from a song performed by British comedy troupe Monty Python about Hormel's canned meat product, SPAM. In the song, an increasing number of people join in repeating the songs chorus: "Spam spam spam spam, spam spam spam spam, lovely spam, wonderful spam ..." Just as in the song, e-mail spam is a tiresome repetition of meaningless text that eventually drowns out any other attempt at communication.

Besides wasting people's time and their computer disk space, spam can consume large amounts of Internet capacity. If one person sends a useless e-mail to a million people, that unsolicited mail consumes Internet resources for a few moments that would otherwise be available to other users. Spam has always been an annoyance, but in recent years companies are increasingly finding it to be a major problem. In addition to consuming bandwidth on company networks and space on e-mail servers, spam distracts employees who are trying to do their jobs and requires them to spend time deleting the unwanted messages. A considerable number of spam messages include content that is offensive to the recipient. Some companies worry that their employees might sue them, arguing that offensive spam they receive while working contributes to harassment by creating a hostile work environment.

MARKUP LANGUAGES AND THE WEB

Web pages can include many elements, such as graphics, photographs, sound clips, and even small programs that run in the Web browser. Each of these elements is stored on the Web server as a separate file. The most important parts of a Web page, however, are the structure of the page and the text that makes up the main part of the page. The page structure and text are stored in a text file that is formatted, or marked up, using a text markup language. A **text markup language** specifies a set of tags that are inserted into the text. These **markup tags**, also called **tags**, provide formatting instructions that Web client software can understand. The Web client software uses those instructions as it renders the text and page elements contained in the other files into the Web page that appears on the screen of the client computer.

The markup language most commonly used on the Web is HTML, which is a subset of a much older and far more complex text markup language called **Standard Generalized Markup Language (SGML)**. Figure 2-5 shows how HTML, XML, and XHTML have descended from the original SGML specification. SGML was used for many years by the publishing industry to create documents that needed to be printed in various formats and that were revised frequently. In addition to its role as a markup language, SGML is a **meta language**, which is a language that can be used to define other languages. Another markup language that was derived from SGML for use on the Web is **Extensible Markup Language (XML)**, which is increasingly used to mark up information that companies share with each other over the Internet. XML is also a meta language because users can create their own markup elements that extend the usefulness of XML (which is why it is called an "extensible" language).

The **World Wide Web Consortium (W3C)**, a not-for-profit group that maintains standards for the Web, presented its first draft form of XML in 1996; the W3C issued its first formal version recommendation in 1998. Thus, it is a much newer markup language than HTML. In 2000, the W3C released the first version of a recommendation for a new markup language called **Extensible Hypertext Markup Language (XHTML)**, which is a reformulation of HTML version 4.0 as an XML application. The Online Companion includes a link to the **W3C XHTML Version 1.0 Specification**.

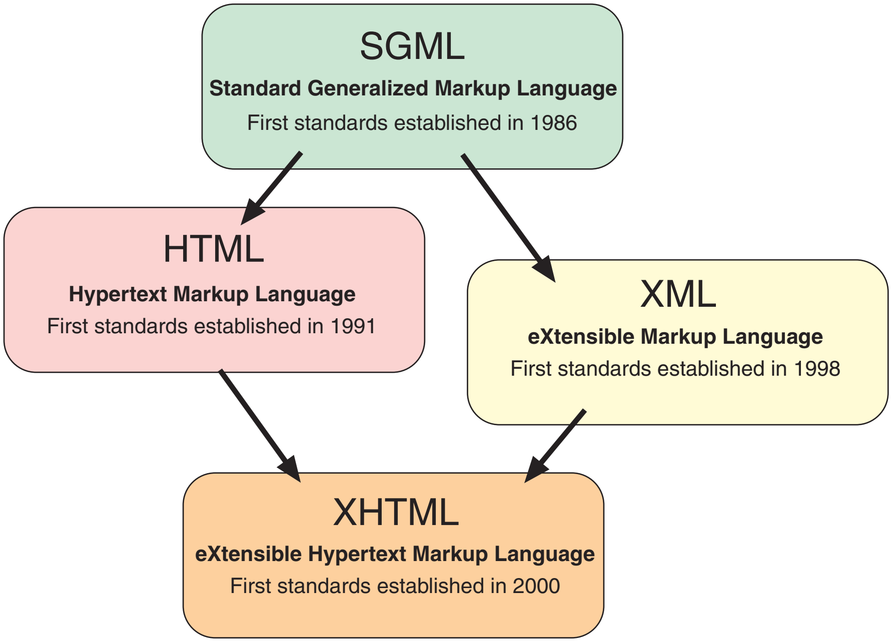

FIGURE 2-5 Development of markup languages

Standard Generalized Markup Language

Since the 1960s, publishers have used markup languages to create documents that can be formatted once, stored electronically, and then printed many times in various layouts that each interpret the formatting differently. U.S. Department of Defense contractors also used early markup languages to create manuals and parts lists for weapons systems. These documents contained many information elements that were often reprinted in different versions and formats. Using electronic document storage and programs that could interpret the formats to produce different layouts saved a tremendous amount of retyping time and cost.

A **Generalized Markup Language (GML)** emerged from these early efforts to create standard formatting styles for electronic documents. In 1986, after many elements of the standard had been in use for years, the **International Organization for Standardization (ISO)** adopted a version of GML called Standard Generalized Markup Language (SGML). SGML offers a system of marking up documents that is independent of any software application. Many organizations, such as the Association of American Publishers, Hewlett-Packard, and Kodak, use SGML because they have complex document management requirements.

SGML is nonproprietary and platform independent and offers user-defined tags. However, it is not well suited to certain tasks, such as the rapid development of Web pages. SGML is costly to set up and maintain, requires the use of expensive software tools, and is hard to learn. Creating document-type definitions in SGML can be expensive and time consuming.

Hypertext Markup Language

HTML includes tags that define the format and style of text elements in an electronic document. HTML also has tags that can create relationships among text elements within one document or among several documents. The text elements that are related to each other are called **hypertext elements**.

HTML is easier to learn and use than SGML. HTML is the prevalent markup language used to create documents on the Web today. The early versions of HTML let Web page designers create text-based electronic documents with headings, title bar titles, bullets, lines, and ordered lists. As the use of HTML and the Web itself grew, HTML creator Berners-Lee turned over the job of maintaining versions of HTML to the W3C. Later versions of HTML included tags for tables, frames, and other features that helped Web designers create more complex page layouts. The W3C maintains detailed information about HTML versions and related topics on its **W3C HTML Page**.

The process for approval of new HTML features takes a long time, so Web browser software developers created some features, called **HTML extensions**, that would only work in their browsers. At various times during the history of HTML, both Microsoft and Netscape enabled their Web browsers to use these HTML extension tags before those tags were approved by the W3C. In some cases, these tags were enabled in one browser and not the other. In other cases, the tags used were never approved by the W3C or were approved in a different form than the one implemented in the Web browser software. Web page designers who wanted to use the latest available tags were often frustrated by this state of affairs. Many of these Web designers had to create separate sets of Web pages for the different types of browsers, which was inefficient and expensive. Most of these tag difference issues were resolved when the W3C issued the specification for HTML version 4.0 in 1997, although enough of them remain to cause regular problems for Web designers.

HTML Tags

An HTML document contains document text and elements. The tags in an HTML document are interpreted by the Web browser and used by it to format the display of the text enclosed by the tags. In HTML, the tags are enclosed in angle brackets (<>). Most HTML tags have an **opening tag** and a **closing tag** that format the text between them. The closing tag is preceded by a slash within the angle brackets. The general form of an HTML element is:

```
<tagname properties> Displayed information affected by tag </tagname>
```

Two good examples of HTML tag pairs are the boldface character-formatting tags and the italic character-formatting tags. For example, a Web browser reading the following line of text:

```
<B>A Review of the Book <I>HTML Is Fun!</I></B>
```

would recognize the <B> and </B> tags as instructions to display the entire line of text in bold and the <I> and </I> tags as instructions to display the text enclosed by those tags in italics. The Web browser would display the text as:

A Review of the Book *HTML Is Fun!*

Some Web browsers allow the user to customize the interpretations of the tags, so that different Web browsers might display the tagged text differently. For example, one Web browser might display text enclosed by bold tags in a blue color instead of displaying the text as bold. Tags can be written in either lowercase or uppercase letters; the tag <b> has the exact same meaning as the tag <B>. Although most tags are two-sided (they use both an opening and a closing tag), some are not. Tags that only require opening tags are known as one-sided tags. The tag that creates a line break (</br>) is a common one-sided tag. Some tags, such as the paragraph tag (<p> ... </p>), are two-sided tags for which the closing tag is optional. Designers often omit the optional closing tags, although many Web designers argue that this practice is poor markup style.

In a two-sided tag set, the closing tag position is very important. For example, if you were to omit the closing bold tag in the preceding example, any text that followed the line would be bolded. Sometimes an opening tag contains one or more property modifiers that further refine how the tag operates. A tag's property may modify a text display, or it may designate where to find a graphic element. Figure 2-6 shows some sample text marked up with HTML tags and Figure 2-7 shows this text as it appears in a Web browser. The tags in these two figures are among the most common HTML tags in use today on the Web.

Other frequently used HTML tags (not shown in the figures) let Web designers include graphics on Web pages and format text in the form of tables. The text and HTML tags that form a Web page can be viewed when the page is open in a Web browser by using the menu commands View, Source (in Internet Explorer) or View, Page Source (in Netscape Navigator). A number of good Web sources (such as the **W3C Getting Started with HTML** page) and textbooks are available that describe HTML tags and their uses, and you may wish to consult them for an in-depth look at HTML.

```
<html>

  <head>

     <title>HTML Tag Examples</title>

  </head>

  <body>

     <h1>This text is set in Heading One tags</h1>
     <h2>This text is set in Heading Two tags</h2>
     <h3>This text is set in Heading Three tags</h3>


     <p>
     This text is set within Paragraph tags. It will appear
     as one paragraph; the text will wrap at the end of each
     line that is rendered in the Web browser no matter where
     the typed text ends. The text inside Paragraph tags is
     rendered without regard to extra spaces typed in the text,
     such as these:              Character formatting can also
     be applied within Paragraph tags. For example, <b>this
     text is set in bold</b>, <font face="Arial">this text is
     set in Arial</font>, and <i>this text is set in italics</i>.
     </p>

     <pre>
The Preformatted tag instructs the Web browser to render the text
      exactly    the    way     it    is    typed,
as in this example.
     </pre>

     <p>
     HTML includes tags that instruct the Web browser to place
     text in bulleted or numbered lists:
     </p>

     <ul>
       <li>Bulleted list item one</li>
       <li>Bulleted list item two</li>
       <li>Bulleted list item three</li>
     </ul>

     <ol>
       <li>Numbered list item one</li>
       <li>Numbered list item two</li>
       <li>Numbered list item three</li>
     </ol>

     <p>
     HTML also includes tags for left, center, and right justification:
     </p>


     <div align="left">Text aligned left</div>
     <div align="center">Text aligned center</div>
     <div align="right">Text aligned right</div>

     <p>
     The most important tag in HTML is the Anchor Hypertext
     Reference tag, which is the tag that provides a link to
     another Web page (or another location in the same Web page).
     For example, the underlined text
     <a href="http://www.course.com/">Course Technology</a>
     is a hyperlink to the Web page of the publisher of this book.
     </p>

   </body>

</html>
```

FIGURE 2-6 Text marked up with HTML tags

This text is set in Heading One tags

This text is set in Heading Two tags

This text is set in Heading Three tags

This text is set within Paragraph tags. It will appear as one paragraph; the text will wrap at the end of each line that is rendered in the Web browser no matter where the typed text ends. The text inside Paragraph tags is rendered without regard to extra spaces typed in the text, such as these: Character formatting can also be applied within Paragraph tags. For example, **this text is set in bold**, this text is set in Arial, and *this text is set in italics*.

```
The Preformatted tag instructs the Web browser to render the text
      exactly    the    way     it    is    typed,
as in this example.
```

HTML includes tags that instruct the Web browser to place text in bulleted or numbered lists:

- Bulleted list item one
- Bulleted list item two
- Bulleted list item three

1. Numbered list item one
2. Numbered list item two
3. Numbered list item three

HTML also includes tags for left, center, and right justification:

Text aligned left

Text aligned center

Text aligned right

The most important tag in HTML is the Anchor Hypertext Reference tag, which is the tag that provides a link to another Web page (or another location in the same Web page). For example, the underlined text Course Technology is a hyperlink to the Web page of the publisher of this book.

FIGURE 2-7 Text marked up with HTML tags as it appears in a Web browser

HTML Links

The Web organizes interlinked pages of information residing on sites around the world. Hyperlinks on Web pages form a "web" of those pages. A user can traverse the interwoven pages by clicking hyperlinked text on one page to move to another page in the web of pages. Users can read Web pages in serial order or in whatever order they prefer by following hyperlinks. Figure 2-8 illustrates the differences between reading a paper catalog in a linear way and reading a hypertext catalog in a nonlinear way.

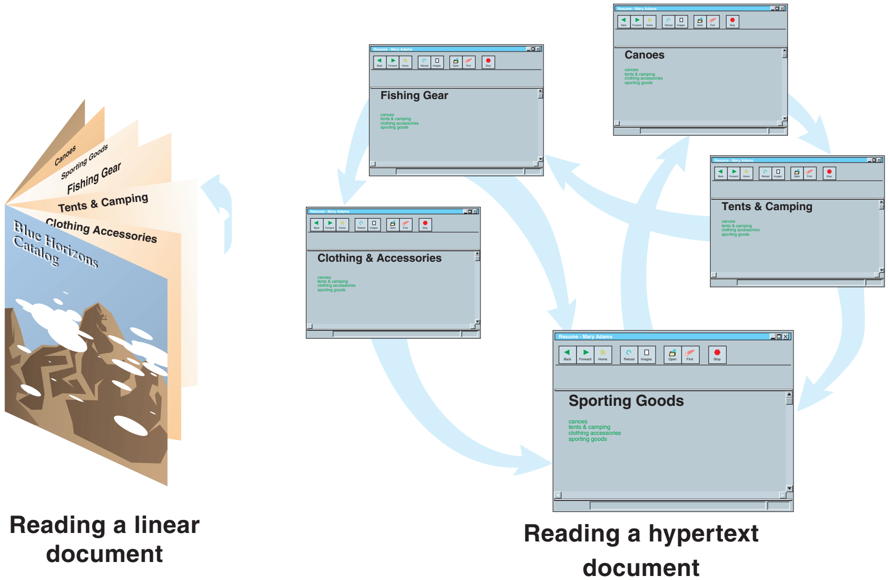

FIGURE 2-8 Linear vs. nonlinear paths through documents

An electronic commerce Web site can use links to direct customers to pages on the company's Web server. The way links lead customers through pages can affect the usefulness of the site and can play a major role in shaping customers' impressions of the company. Two commonly used link structures are linear and hierarchical. A **linear hyperlink structure** resembles conventional paper documents in that the reader begins on the first page and clicks the Next button to move to the next page in a serial fashion. This structure works well when customers fill out forms prior to a purchase or other agreement. In this case, the customer reads and responds to page one, and then moves on to the next page. This process continues until the entire form is completed. The only Web page navigation choices the user typically has are Back and Continue.

Another link arrangement is called a hierarchical structure. In a **hierarchical hyperlink structure**, the Web user opens an introductory page called a **home page** or **start page**. This page contains one or more links to other pages, and those pages, in turn, link to other pages. This hierarchical arrangement resembles an inverted tree in which the root is at the top and the branches are below it. Hierarchical structures are good for leading customers from general topics or products to specific product models and quantities. A company's home page might contain links to help, company history, company officers, order processing, frequently asked questions, and product catalogs. Many sites that use a hierarchical structure include a page on the Web site that contains a map or listing of the Web pages in their hierarchical order. This page is called a **site map**. Figure 2-9 illustrates the linear and hierarchical structures. Of course, pages combining linear and hierarchical structures are also possible.

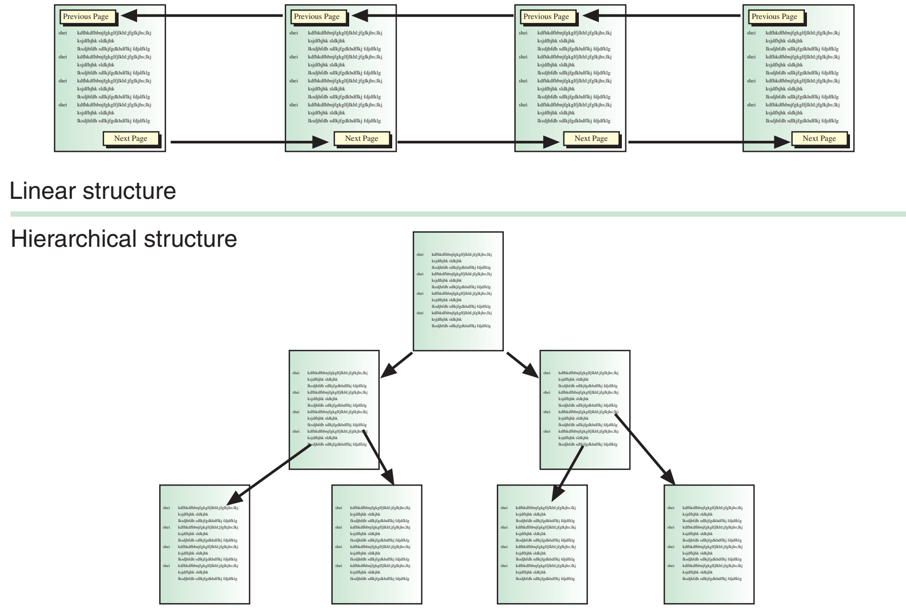

FIGURE 2-9 Two alternative hyperlink structures

In HTML, hyperlinks are created using the HTML **anchor tag**. Whether you are linking to text within the same document or to a document on a distant computer, the anchor tag has the same basic form:

```
<A HREF="address">Visible link text</A>
```

Anchor tags have opening and closing tags. The opening tag has a hypertext reference (HREF) property, which specifies the remote or local document's address. Clicking the text following the opening link transfers control to the HREF address, wherever that happens to be. A person creating an electronic résumé on the Web might want to make a university's name and address under the Education heading a hyperlink instead of plain text. Anyone viewing the résumé can click the link, which leads the reader to the university's home page. The following example shows the HTML code to create a hyperlink to another Web server:

```
<A HREF="http://www.sandiego.edu">University of San Diego</A>
```

Similarly, the résumé could include a local link to another part of the same document with the following marked up text:

```
<A HREF="#references">References are found here</A>
```

In both of these examples, the text between the anchors appears on the Web page as a hyperlink. Most browsers display the link in blue and underline it. In most browser software, the action of moving the mouse pointer over a hyperlink causes the mouse pointer to change from an arrow to a pointing hand.

Scripting Languages and Style Sheets

Versions of HTML released by the W3C after 1997 include an HTML tag called the object tag and include support for cascading style sheets. Web designers can use the object tag to embed scripting language code on HTML pages. The most common scripting languages used on Web pages are JavaScript, JScript, Perl, and VBScript. Scripts written in these languages and embedded on Web pages can execute programs on computers that display those pages. You can learn more about embedding script languages in HTML documents (also called **client-side scripting**) by taking a course in Web programming or reading a book such as Kathleen Kalata's *Internet Programming* (a full reference for this book appears in the For Further Study and Research section at the end of this chapter).

Cascading style sheets (CSS) are sets of instructions that give Web developers more control over the format of displayed pages. Similar to document styles in word-processing programs, CSS let designers define formatting styles that can be applied to multiple Web pages. The set of instructions, called a **style sheet**, is usually stored in a separate file and is referenced using the HTML style tag; however, it can be included as part of a Web page's HTML file. The term *cascading* means that designers can apply many style sheets to the same Web page, one on top of the other. For example, a three-stage cascade might include one style sheet with formatting instructions for text within heading 1 tags, a second style sheet with formatting instructions for text within heading 2 tags, and a third style sheet with formatting instructions for text within paragraph tags. A designer who later decides to change the formatting of heading 2 text can just replace the second style sheet with a different one.

Extensible Markup Language (XML)

As the Web grew, HTML continued to provide a useful tool for Web designers who wanted to create attractive layouts of text and graphics on their pages. However, as companies began to conduct electronic commerce on the Web, the need to present large amounts of data on Web pages also became important. Companies created Web sites that contained lists of inventory items, sales invoices, purchase orders, and other business data. The need to keep these lists updated was also important and posed a new challenge for many Web designers. The tool that had helped these Web designers create useful Web pages, HTML, was not such a good tool for presenting or maintaining information lists.

In the late 1990s, companies began turning to XML to help them maintain Web pages that contained large amounts of data. XML uses paired start and stop tags in much the same way as database software defines a record structure. For example, a company that sells products on the Web might have Web pages that contain descriptions and photos of the products it sells. The Web pages are marked up with HTML tags, but the product information elements themselves, such as prices, identification numbers, and quantities on hand, are marked up with XML tags. The XML document is embedded within the HTML document.

XML includes data management capabilities that HTML cannot provide. To better understand the strengths of XML and weaknesses of HTML in data management tasks, consider the simple example of a Web page that includes a list of countries and some basic information about each country. A Web designer might decide to use HTML tags to show each information item the same way for each country. Each information item would use a different tag. Assume that the Web designer in this case decided to use the HTML heading tags to present the data. Figure 2-10 shows the data and the HTML heading tags for four countries (this is only an example; the actual list would include more than 150 countries). The first item in the list provides the definitions for each tag. Figure 2-11 shows this HTML document as it appears in a Web browser.

```
<html>

   <head>

     <title>Countries</title>

   </head>

   <body>

     <h1>Countries</h1>

     <h2>CountryName</h2>
     <h3>CapitalCity</h3>
     <h4>AreaInSquareKilometers</h4>
     <h5>OfficialLanguage</h5>
     <h6>VotingAge</h6>

     <h2>Argentina</h2>
     <h3>Buenos Aires</h3>
     <h4>2,766,890</h4>
     <h5>Spanish</h5>
     <h6>18</h6>

     <h2>Austria</h2>
     <h3>Vienna</h3>
     <h4>83,858</h4>
     <h5>German</h5>
     <h6>19</h6>

     <h2>Barbados</h2>
     <h3>Bridgetown</h3>
     <h4>430</h4>
     <h5>English</h5>
     <h6>18</h6>

     <h2>Belarus</h2>
     <h3>Minsk</h3>
     <h4>207,600</h4>
     <h5>Byelorussian</h5>
     <h6>18</h6>

   </body>

</html>
```

FIGURE 2-10 Country list data marked up with HTML tags

FIGURE 2-11 Country list data as it appears in a Web browser

These figures reveal some of the shortcomings of using HTML to present a list of items when the meaning of each item in the list is important. The Web designer in this case used HTML heading tags. HTML has only six levels of heading tags; thus, if the individual items had any more information elements than shown in this example (such as population and continent), this approach would not work at all. The Web designer could use various combinations of text attributes such as size, font, color, bold, or italics to distinguish among items, but none of these tags would convey the meaning of the individual data elements. The only information about the meaning of each country's listing appears in the first list item, which includes the definitions for each element. In the late 1990s, Web professionals began to consider XML as a list-formatting alternative to HTML that would more effectively communicate the meaning of data.

XML differs from HTML in two important respects. First, XML is not a markup language with defined tags. It is a framework within which individuals, companies, and other organizations can create their own sets of tags. Second, XML tags do not specify how text appears on a Web page; the tags convey the meaning (the semantics) of the information included within them. To understand this distinction between appearance and semantics, consider the list of countries example again. In XML, tags can be created for each fact that define the meaning of the fact. Figure 2-12 shows the countries data marked up with XML tags. Some browsers, such as Internet Explorer, can render XML files directly without additional instructions. Figure 2-13 (on the next page) shows the country list XML file as it would appear in an Internet Explorer browser window.

declaration

root element

```
<?xml version="1.0"?>

<CountriesList>

  <Country Name = "Argentina">
    <CapitalCity>Buenos Aires</CapitalCity>
    <AreaInSquareKilometers>2,766,890</AreaInSquareKilometers>
    <OfficialLanguage>Spanish</OfficialLanguage>
    <VotingAge>18</VotingAge>
  </Country>

  <Country Name = "Austria">
    <CapitalCity>Vienna</CapitalCity>
    <AreaInSquareKilometers>83,858</AreaInSquareKilometers>
    <OfficialLanguage>German</OfficialLanguage>
    <VotingAge>19</VotingAge>
  </Country>

  <Country Name = "Barbados">
    <CapitalCity>Bridgetown</CapitalCity>
    <AreaInSquareKilometers>430</AreaInSquareKilometers>
    <OfficialLanguage>English</OfficialLanguage>
    <VotingAge>18</VotingAge>
  </Country>

  <Country Name = "Belarus">
    <CapitalCity>Minsk</CapitalCity>
    <AreaInSquareKilometers>207,600</AreaInSquareKilometers>
    <OfficialLanguage>Byelorussian</OfficialLanguage>
    <VotingAge>18</VotingAge>
  </Country>

</CountriesList>
```

FIGURE 2-12 Country list data marked up with XML tags

```
<?xml version="1.0" ?>
<CountriesList>
- <Country Name="Argentina">
    <CapitalCity>Buenos Aires</CapitalCity>
    <AreaInSquareKilometers>2,766,890</AreaInSquareKilometers>
    <OfficialLanguage>Spanish</OfficialLanguage>
    <VotingAge>18</VotingAge>
  </Country>
- <Country Name="Austria">
    <CapitalCity>Vienna</CapitalCity>
    <AreaInSquareKilometers>83,858</AreaInSquareKilometers>
    <OfficialLanguage>German</OfficialLanguage>
    <VotingAge>19</VotingAge>
  </Country>
- <Country Name="Barbados">
    <CapitalCity>Bridgetown</CapitalCity>
    <AreaInSquareKilometers>430</AreaInSquareKilometers>
    <OfficialLanguage>English</OfficialLanguage>
    <VotingAge>18</VotingAge>
  </Country>
- <Country Name="Belarus">
    <CapitalCity>Minsk</CapitalCity>
    <AreaInSquareKilometers>207,600</AreaInSquareKilometers>
    <OfficialLanguage>Byelorussian</OfficialLanguage>
    <VotingAge>18</VotingAge>
  </Country>
</CountriesList>
```

FIGURE 2-13 Country list data marked up with XML tags as it would appear in Internet Explorer

The first line in the XML file shown in Figures 2-12 and 2-13 is the declaration, which indicates that the file uses version 1.0 of XML. XML markup tags are similar in appearance to SGML markup tags, thus the declaration can help avoid confusion in organizations that use both. The second line and the last line are the root element tags. The root element of an XML file contains all of the other elements in that file and is usually assigned a name that describes the purpose or meaning of the file.

The other elements are called child elements; for example, Country is a child element of CountriesList. Each of the other attributes is, in turn, a child element of the Country element. The names of these child elements were created specifically for use in this file. If programmers in another organization were to create a file with country information, they might use different names for these elements (for example, “Capital” instead of “CapitalCity”), which would make it difficult for the two organizations to share information. Thus, the greatest strength of XML, that it allows users to define their own tags, is also its greatest weakness.

To overcome that weakness, many companies have agreed to follow common standards for XML tags. These standards, in the form of data type definitions (DTDs) or XML schemas, are available for a number of industries, including the **ebXML** initiative for electronic commerce standards, the **eXtensible Business Reporting Language (XBRL)** for accounting and financial information standards, **LegalXML** for information in the legal profession, and **MathML** for mathematical and scientific information. A number of industry groups have formed to create standard XML tag definitions that can be used by all companies in that industry. **RosettaNet** is an example of such an industry group. In 2001, the W3C released a set of rules for XML document interoperability that many researchers believe will help resolve incompatibilities between different sets of XML tag definitions. A set of XML tag definitions is sometimes called an **XML vocabulary**. Hundreds of publicly defined XML vocabularies are currently circulating, many of which are registered with the **XML Registry**. You can learn more about XML by reading the **W3C XML Pages**.

Although it is possible to display XML files in some Web browsers, XML files are not intended to be displayed in a Web browser. XML files are intended to be translated using another file that contains formatting instructions or to be read by a program. Formatting instructions are often written in the **Extensible Stylesheet Language (XSL)**, and the programs that read or transform XML files are usually written in the Java programming language. These programs, sometimes called **XML parsers**, can format an XML file so it can appear on the screen of a computer, a wireless PDA, a mobile phone, or some other device. A diagram showing one way that a Web server might process an HTTP request for an XML page appears in Figure 2-14.

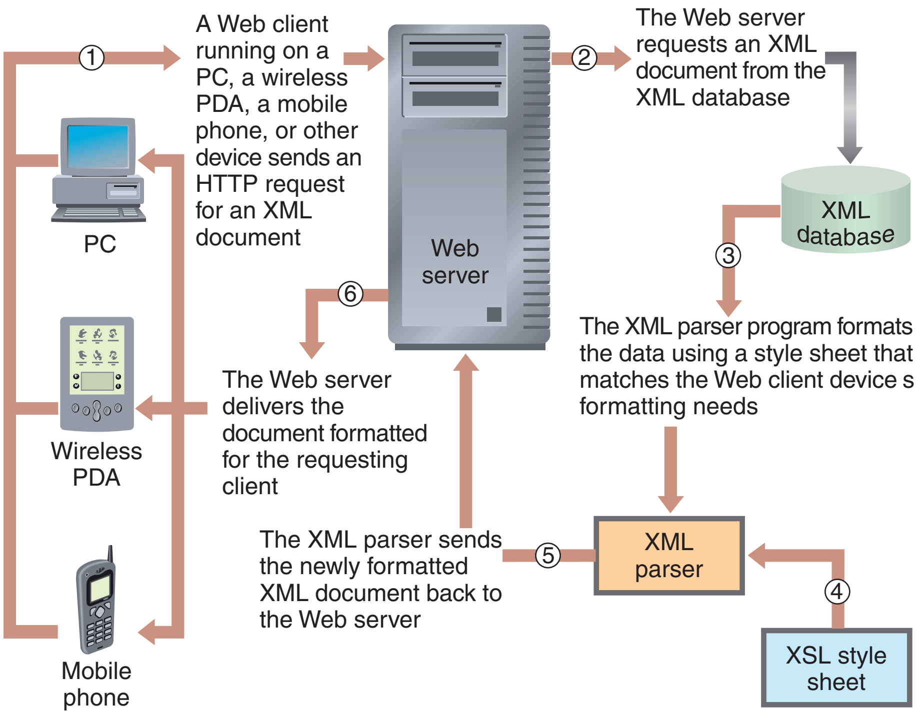

FIGURE 2-14 Processing a request for an XML page

HTML and XML Editors

Web designers can create HTML documents in any general-purpose text editor or word processor. However, one of the special-purpose HTML editors can help Web designers create Web pages much more easily. There are many freeware, shareware, and commercial

HTML editors available for download on the Internet, including CoffeeCup, HomeSite, and CuteHTML. The Additional Resources section of the Online Companion for this chapter includes links to Web sites that offer downloads of these **HTML Editing Programs**.

HTML editors are also included as part of more sophisticated Web site design and creation programs that are sometimes called Web page builder software. With these programs, Web designers can create and manage complete Web sites, including features for database access, graphics, and fill-in forms. These programs display the Web page as it will appear in a Web browser in one window and display the HTML-tagged text in another window. The designer can edit in either window and changes are reflected in the other window. For example, the designer can drag and drop objects such as graphics onto the Web browser view page and the program automatically generates the HTML tags to position the graphics.

Web site design and creation software also provides maintenance tools that allow the designer to create a Web site on a PC and then upload the entire site (HTML documents, graphics files, and so on) to a Web server computer. When the site needs to be edited later, the designer can edit the copy of the site on the PC and instruct the program to synchronize those changes on the copy of the site that resides on the Web server. Examples of Web site design and creation programs include **Microsoft FrontPage** and **Macromedia Dreamweaver**.

XML files, like HTML files, can be created in any text editor. However, programs designed to make the task of designing and managing XML files easier are also available. These programs include Epic Editor, TurboXML, XMetal, and XML Spy. These programs provide tag validation and XML creation capabilities in addition to making the job of marking up text with XML tags more efficient. You can find links to these programs' Web sites in the Additional Resources section of the Online Companion under the heading **XML Editing Programs**.

INTRANETS AND EXTRANETS

Not all TCP/IP networks connect to the Internet. Many companies build internets (small "i"), or interconnected networks, that do not extend beyond their organizational boundaries. An **intranet** is an interconnected network (or internet), usually one that uses the TCP/IP protocol set, and does not extend beyond the organization that created it. An **extranet** is an intranet that has been extended to include specific entities outside the boundaries of the organization, such as business partners, customers, or suppliers. Although fax, telephone, e-mail, and overnight express carriers have been the main communications tools for business for many years, extranets can replace many of them at a lower cost.

Intranets

Intranets are an excellent low-cost way to distribute internal corporate information. Based on the client-server model, intranet requests for files, documents, or schematic drawings work the same way they do on the Internet. An intranet uses Web browsers and Internet-based protocols, including TCP/IP, FTP, Telnet, HTML, and HTTP. Because intranets are compatible with the Internet, information from intranets can be shared among departments that use different technologies as well as among external consumers. Intranets are often the most efficient way to distribute internal corporate information, because producing and distributing paper is usually slower and more expensive than using Web-based communications.

Companies can also use intranets to reduce software maintenance and update costs for their employees' computer workstations. Computing staff can place software updates and patches on the intranet, and then provide a script to update employee workstations automatically the next time they log on.

Extranets

Extranets are networks that connect companies with suppliers, business partners, or other authorized users. Each participant in the extranet has access to the databases, files, or other information stored on computers connected to the extranet. An extranet can be set up through the Internet, or it can use a separate network.

Some extranets start out as intranets that eventually provide access of intranet data to select Internet users. For example, for many years, FedEx let customers track their packages by calling a FedEx toll-free number, and then giving the operator a tracking number. In the early 1990s, FedEx began giving package-tracking software to any customer who wanted it. Once it was installed on the customer's computer, the software dialed the FedEx computer using a modem, queried the status of the customer's package, and displayed the results on the customer's computer with no operator required. In the mid-1990s, FedEx eliminated client-machine software and made package tracking available on its Web site. Instead of having thousands of programs running on customers' computers, FedEx has its customers use their Web browsers to run one program running on its Web site. This Web-based system is called **FedEx Ship Manager**, and it gives customers Web access (from any browser on any computer on the Web) to package tracking, air bill creation, shipment logging, and FedEx supply shipments. Critical information, such as a package's location, is stored and made available to customers through the FedEx Ship Manager section of the FedEx extranet.

Public and Private Networks

A **public network** is any computer network or telecommunications network that is available to the public. The Internet is one example of a public network. Although a company can operate its extranet using a public network, very few do because of the high level of security risks. The Internet, as you will learn in later chapters, does not provide a high degree of security in its basic structure.

A **private network** is a private, leased-line connection between two companies that physically connects their intranets to one another. A **leased line** is a permanent telephone connection between two points. Unlike the normal telephone connection you create when you dial a telephone number, a leased line is always active. The advantage of a leased line is security. Only the two parties that lease the line to create the private network have access to the connection. The largest drawback to a private network is cost. Leased lines are expensive. Every pair of companies wanting a private network between them requires a separate line connecting them. For instance, if a company wants to set up an extranet connection over a private network with seven other companies, the company must pay the cost of seven leased lines, one for each company. If the extranet expands to

20 other companies, the extranet-sponsoring company must rent another 13 leased lines. As each new company is added, costs increase by the same amount and soon become prohibitive. Vendors refer to this as a **scaling problem**; that is, increasing the number of leased lines in private networks is difficult, costly, and time consuming. As the number of companies that need to join the extranet increases, other networking options become appealing.

Virtual Private Network (VPN)

A **virtual private network (VPN)** is an extranet that uses public networks and their protocols to send sensitive data to partners, customers, suppliers, and employees using a system called IP tunneling or encapsulation. **IP tunneling** effectively creates a private passageway through the public Internet that provides secure transmission from one computer to another. The virtual passageway is created by VPN software that encrypts the packet content and then places the encrypted packets inside another packet in a process called **encapsulation**. The outer packet is called an **IP wrapper**. The Web server sends the encapsulated packets to their destinations over the Internet. The computer that receives the packet unwraps it and decrypts the message using VPN software that is the same as, or compatible with, the VPN software used to encrypt and encapsulate the packet at the sending end. The "virtual" part of VPN means that the connection seems to be a permanent, internal network connection, but the connection is actually temporary. Each transaction between two intranets using a VPN is created, carries out its work over the Internet, and is then terminated.

VPN software must be installed on the computers at both ends of the transmission. A VPN provides security shells, with the most sensitive data under the tightest control. The VPN is like a separate, covered commuter lane on a highway (the Internet) in which the passengers are protected from being seen by the vehicles traveling in the other lanes. Company employees in remote locations can send sensitive information to company computers using the VPN private tunnels established on the Internet.

Unlike private networks using leased lines, VPNs establish short-term logical connections in real time that are broken once the communication session ends. Establishing VPNs does not require leased lines. The only infrastructure required outside each company's intranet is the Internet. Companies such as **Aventail**, **Cisco**, **SonicWall**, and **V-ONE** are making VPNs simpler to install and maintain.

Extranets are sometimes confused with VPNs. Although a VPN is an extranet, not every extranet is a VPN. Figure 2-15 shows a diagram of a VPN. VPNs usually work as part of a firewall (the firewall-VPN combinations are shown as brick walls in the figure). A firewall is a program or hardware device that protects information inside an organization's network from attacks that originate outside the network. You will learn more about VPNs, firewalls, and other network security devices in Chapter 10.

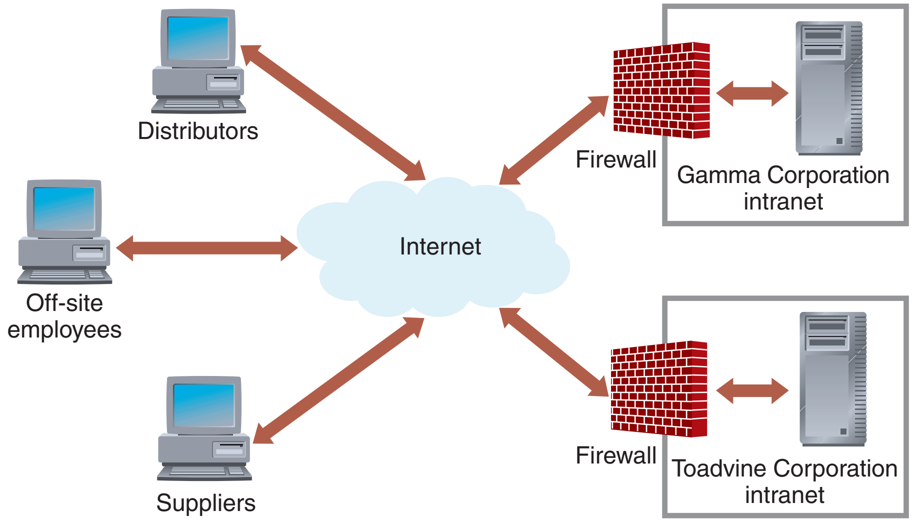

FIGURE 2-15 VPN architecture example

INTERNET CONNECTION OPTIONS

The Internet is a set of interconnected networks. A corporation or individual cannot become part of the Internet without a telephone connection or a connection to a LAN or intranet. Larger firms that provide Internet access to other businesses, called **Internet access providers (IAPs)** or ISPs, usually offer several connection options. This section briefly describes current connection choices and presents their advantages and disadvantages.

Connectivity Overview

ISPs offer several ways to connect to the Internet. The most common connection options are voice-grade telephone line, various types of broadband connections, leased line, and wireless. One of the major distinguishing factors between various ISPs and their connection options is the bandwidth they offer. **Bandwidth** is the amount of data that can travel through a communication line per unit of time. The higher the bandwidth, the faster data files travel and the faster Web pages appear on your screen. Each connection option offers different bandwidths, and each ISP offers varying bandwidths for each connection option. Traffic on the Internet and at your local service provider greatly affects **net bandwidth**, which is the actual speed that information travels. When few people are competing for service from an ISP, net bandwidth approaches the carrier's upper limit. On the other hand, users experience slowdowns during high-traffic periods.

Bandwidth can differ for data traveling to or from the ISP depending on the user's connection type. **Symmetric connections** provide the same bandwidth in both directions. **Asymmetric connections** provide different bandwidths for each direction. **Upstream bandwidth**, also called **upload bandwidth**, is a measure of the amount of information that can travel from the user to the Internet in a given amount of time. **Downstream bandwidth**, also called **download** or **downlink bandwidth**, is a measure of the amount of information that can travel from the Internet to a user in a given amount of time (for example, when a user receives a Web page from a Web server).

Voice-Grade Telephone Connections

The most common way to connect to an ISP is through a modem connected to your local telephone service provider. **POTS**, or **plain old telephone service**, uses existing telephone lines and an analog modem to provide a bandwidth of between 28 and 56 Kbps. Some telephone companies offer a higher grade of service called **Digital Subscriber Line (DSL)** protocol. DSL connection methods do not use a modem. They use a piece of networking equipment that is similar to a network switch, but most people call this piece of equipment (incorrectly) a "DSL modem." **Integrated Services Digital Network (ISDN)** was the first technology developed to use the DSL protocol suite and has been available in parts of the United States since 1984. ISDN is more expensive than regular telephone service and offers bandwidths of between 128 Kbps and 256 Kbps.

Broadband Connections

Connections that operate at speeds of greater than about 200 Kbps are called **broadband** services. One of the newest technologies that uses the DSL protocol to provide service in the broadband range is **asymmetric digital subscriber line (ADSL**, usually abbreviated **DSL)**. It provides transmission bandwidths from 100 to 640 Kbps upstream and from 1.5 to 9 Mbps (million bits per second) downstream. For businesses, a **high-speed DSL (HDSL)** connection service can provide more than 768 Kbps of symmetric bandwidth.

Cable modems—connected to the same broadband coaxial cable that serves a television—typically provide transmission speeds between 300 Kbps and 1 Mbps from the client to the server. The downstream transmission rate can be as high as 10 Mbps. In the United States alone, more than 100 million homes and organizations have broadband cable service available, and more than 70 million homes subscribe to cable television. The latest estimates indicate that there are more than 9 million cable modem subscribers in the United States and another 3 million households that have broadband DSL or satellite connections.

DSL is a private line with no competing traffic. Unlike DSL, cable modem connection bandwidths vary with the number of other subscribers competing for the shared resource. Transmission speeds can decrease dramatically in heavily subscribed neighborhoods at prime times—in neighborhoods where many people are using cable modems simultaneously.

Connection options based on cable or telephone line connections are wonderful for urban and suburban Web users, but those living in rural areas often have very limited telephone service and no cable access at all. The telephone lines used to cover the vast distances between rural customers are usually **voice-grade lines**, which cost less than telephone lines designed to carry data, are made of lower grade copper, and were never intended to carry data. These lines can carry only limited bandwidth—usually less than 14 Kbps. Telephone companies have wired most urban and suburban areas with **data-grade lines** (made more carefully and of higher grade copper than voice-grade lines) because the short length of the lines in these areas makes it less expensive to install than in rural areas where connection distances are much longer. It is also likely that urban and suburban lines will someday be leased to companies willing to pay the higher fees charged for data-grade lines.

LEARNING FROM FAILURES

NORTHPOINT COMMUNICATIONS

In 1997, Michael Malaga was a successful telecommunications executive with an idea. He wanted to sell broadband Internet access to small businesses in urban areas. DSL technology was just gaining acceptance, and leased telephone lines were available from telephone companies. He wanted to avoid residential customers because they would soon have inexpensive cable modem access to meet their broadband needs. He also wanted to avoid suburban and rural businesses to keep the telephone line leasing costs low (lease charges are higher for longer distances). He and five friends started NorthPoint Communications with $500,000 of their combined savings and raised another $11 million within a few months. After six months, the company had raised more money from investors and had acquired 1500 customers, but it was posting a net loss of $30 million. On the strength of its number of customers, the company began the task of raising the $100 million that Malaga estimated it would need to create the network infrastructure.

Independent DSL providers such as NorthPoint were pressed by customers to install service rapidly, but had to rely on local telephone companies to ensure that their lines would support DSL. In many cases, the telephone companies had to install switches and other equipment to make DSL work on a particular line. The telephone companies often were in no rush to do this because they also sold DSL service, and speedy service would be helping a competitor. The delays led to unpredictable installation holdups and many unhappy NorthPoint customers. Customers with problems after the service was installed often were bounced from the telephone company to NorthPoint, without obtaining satisfactory or timely resolutions of their problems.

Although NorthPoint was unable to make its relationship with each customer profitable, Malaga and his team were rapidly raising money in the hot capital markets of the time. The company raised $162 million before its first stock offering in 1999, which brought in an additional $387 million. At that time, the company had 13,000 customers, which means that NorthPoint had raised more than $42,000 from outside investors for each customer. Considering that each customer would generate revenue of about $1000 peryear, the economics of the business did not look good. By the end of 1999, NorthPoint had spent $300 million of the cash it had raised to build its network infrastructure and reported an operating loss of $184 million. At this point, NorthPoint was operating in 28 cities.

During the next year, the company continued to raise additional funds, gain more customers, and lose money on each customer. In August 2000, the telephone company Verizon agreed to purchase 55 percent of the company for $800 million paid in installments. The total funding that NorthPoint had obtained by the end of 2000, including the partial payments received from Verizon, added up to $1.2 billion. By the end of the year, NorthPoint was in 109 cities and needed to spend $66 million in cash per month just to stay in business. Verizon withdrew from the purchase agreement, the stock plunged, and the layoffs began.

continued

NorthPoint filed for bankruptcy in January 2001 and sold its networking hardware to AT&T in March for $135 million. AT&T was not interested in continuing the DSL business (it just wanted the hardware), so NorthPoint's 87,000 small business customers lost their Internet service overnight. In many of the cities that NorthPoint had served, there were no competitors to pick up the service. Because the capital markets of the late 1990s were so eager to invest in anything that appeared to be connected with the Internet, NorthPoint was able to raise incredible amounts of money. However, NorthPoint sold Internet access to customers for less than it cost to provide the service. No amount of investor money could overcome that basic business mistake.

Leased-Line Connections

Large firms with large amounts of Internet traffic can connect to an ISP using higher bandwidth connections that they can lease from telecommunications carriers. These connections use a variety of technologies and are usually classified by the equivalent number of telephone lines they include. (The connection technologies they use were originally developed to carry large numbers of telephone calls.)

A telephone line designed to carry one digital signal is called DS0 (digital signal zero, the name of the signaling format used on those lines) and has a bandwidth of 56 Kbps. A **T1** line (also called a DS1) carries 24 DS0 lines and operates at 1.544 Mbps. Some telecommunications companies offer **fractional T1**, which provides service speeds of 128 Kbps and upward in 128-Kbps increments. **T3** service (also called DS3) offers 44.736 Mbps (the equivalent of 30 T1 lines or 760 DS0 lines). All of these leased telephone line connections are much more expensive than POTS, ISDN, or DSL connections.

Large organizations that need to connect hundreds or thousands of individual users to the Internet require very high bandwidth. NAPs use T1 and T3 lines. NAPs and the computers that perform routing functions on the Internet backbone also use technologies such as **frame relay** and **asynchronous transfer mode (ATM)** connections and **optical fiber** (instead of copper wire) connections with bandwidths determined by the class of fiber-optic cable used. An OC3 (optical carrier 3) connection provides 156 Mbps, an OC12 provides 622 Mbps, an OC48 provides 2.5 Gbps (gigabits, or 1 billion bits per second), and an OC192 provides 10 Gbps.

Wireless Connections

For many people in rural areas, satellite microwave transmissions have made connections to the Internet possible for the first time. In the first satellite technologies, the customer placed a receiving dish antenna on the roof or in the yard and pointed it at the satellite. The satellite sent microwave transmissions to handle Internet downloads at speeds of around 500 Kbps. Uploads were handled by a POTS modem connection. For Web browsing, this was not too bad, since most of the uploaded messages were small text messages (e-mails and Web page requests). People who wanted to send large e-mail attachments or transfer files over the Internet found the slow upload speeds unsatisfactory.

In recent years, companies such as **DirecPC**, **DIRECWAY**, and **StarBand** have begun offering satellite Internet connections that do not require a POTS modem connection for uploads. These connections use a microwave transmitter for Internet uploads. This transmitter provides upload speeds as high as 150 Kbps. Initially, the installation charges were much higher than for other residential Internet connection services because a professional installer was needed to carefully aim the transmitter's dish antenna at the satellite. Recently, the accuracy of the antennas improved, and some of these companies now offer a self-installation option that drastically reduces the initial cost. For installations in North America, the antennas must have a clear line of sight into the southwestern sky. This requirement can make these services unusable for many people living in large cities or on the wrong side of an apartment building.

No discussion of Internet connections would be complete without mention of the wireless devices that can be connected to the Internet. People today use mobile phones, wireless personal digital assistants (PDAs), tablet computers, and even laptops equipped with wireless network cards to connect to networks that, in turn, are connected to the Internet. Several wireless standards are in use today and more are being developed.

Bluetooth and Ultra Wideband (UWB)

One of the first wireless protocols, designed for personal use over short distances, is called **Bluetooth**. (The protocol was developed in Norway and is named for Harald Bluetooth, a 10th century Scandinavian king.) Bluetooth operates reliably over distances of up to 35 feet and can be a part of up to 10 networks of eight devices each. It is a low-bandwidth technology, with speeds of up to 722 Kbps. Bluetooth is useful for tasks such as wireless synchronization of laptop computers with desktop computers and wireless printing from laptops, PDAs, or mobile phones. These small Bluetooth networks are called **personal area networks (PANs)** or **piconets**.

One major advantage of Bluetooth technology is that it consumes very little power, which is an important consideration for mobile devices. Another advantage is that Bluetooth devices can discover one another and exchange information automatically. For example, a person using a laptop computer in a temporary office can print to a local Bluetooth-enabled printer without logging in to the network or installing software on either device. The printer and the laptop computer electronically recognize each other as Bluetooth devices and immediately can begin exchanging information.

Another wireless communication technology, **Ultra Wideband (UWB)**, provides wide bandwidth (up to about 480 Mbps in current versions) connections over short distances (30 to 100 feet). UWB was developed for short-range secure communications in military applications during the 1960s. Many observers believe that UWB and other similar technologies will be used in future personal area networking applications such as home media centers (for example, a PC could beam stored video files to a nearby television) and in linking mobile phones to the Internet.

Wireless Ethernet (Wi-Fi)

The most common wireless connection technology for use on LANs is called **Wi-Fi**, **wireless Ethernet**, or **802.11b** (802.11 is the number of the technology's **network specification**, which is the set of rules that equipment connected to the network must follow). A computer equipped with a Wi-Fi network card can communicate through a wireless access point connected to a LAN to become a part of that LAN. A **wireless access point (WAP)** is a device that transmits network packets between Wi-Fi-equipped computers and other devices that are within its range. The user must have authorization to connect to the LAN and might be required to perform a login procedure before the laptop can access the LAN through the WAP.

Wi-Fi has a potential bandwidth of 11 Mbps and a range of about 300 feet. In actual installations, the achieved bandwidth and range can be dramatically affected by the construction material of the objects (such as walls, floors, doors, and windows) through which the signals must pass. For example, reinforced concrete walls and certain types of tinted glass windows greatly reduce the effective range of Wi-Fi. Despite these limitations, organizations can make Wi-Fi a key element of their LAN structures by installing a number of WAPs throughout their premises. Wi-Fi devices are capable of **roaming**, that is, shifting from one WAP to another, without requiring intervention by the user. Increasingly, Wi-Fi is becoming available in public places such as airports, convention centers, hotels, and office lobbies. The users of these networks authorize a connection charge when they log in and then have access to the wireless LAN's resources, including access to the Internet.

In 2002, an improved version of Wi-Fi, called **802.11a** (the 802.11b protocol was easier to implement, thus it was introduced first) was introduced. The 802.11a protocol is capable of transmitting data at speeds up to 54 Mbps, but it is not compatible with 802.11b devices. Later in 2002, the **802.11g** protocol, which has the 54 Mbps speed of 802.11a and is compatible with 802.11b devices, was introduced. Because of its compatibility with the many 802.11b devices that were in use, 802.11g was an immediate success. The next protocol in this series, **802.11n**, is expected to offer even greater speeds (up to 320 Mbps). Most industry experts expect the 802.11n protocol to be released in 2006 or 2007.

Some organizations operate WAPs that are open to the public. These access points are called **hot spots**. Some organizations allow free access to their hot spots, others charge an access fee. A growing number of retail establishments, such as McDonald's, Panera, and Starbucks, offer hot spots. Hotels have found that installing a WAP can be cheaper and easier than running network cable, especially in older buildings. Some hotels offer wireless access free, others charge a small fee. There are several Web sites that offer hot spot directories that show hot spots by location, but these sites tend to open and close frequently, so these directories become out of date rather quickly. The best way to find hot spots (or a hot spot directory) is to use your favorite search engine.

Some communities have installed wireless networks that can be accessed from anywhere in the area. For example, the city of Grand Haven, Michigan, installed a metropolitan area Wi-Fi network. Grand Haven is a growing town on the shores of Lake Michigan. The company that built the network, Ottawa Wireless, sells network access to residents and businesses throughout the area. The company offers access not only on land, but on boats up to 15 miles out on Lake Michigan. Several small company owners use this network to conduct their online business while sailing!

Fixed-Point Wireless

In some areas, companies such as **Etherlinx** and **Getwireless.net** are beginning to offer fixed-point wireless service. One version of **fixed-point wireless** uses a system of repeaters to forward a radio signal from the ISP to customers. The **repeaters** are transmitter-receiver devices (also called **transceivers**) that receive the signal and then retransmit it toward users' roof-mounted antennas and to the next repeater, which receives the signal and passes it on to the next repeater, which can be up to 20 miles away. The users' antennas are connected to a device that converts the radio signals into Wi-Fi packets that are sent to the users' computers or wireless LANs. Another version of fixed-point wireless directly transmits Wi-Fi packets through hundreds, or even thousands, of short-range transceivers that are located close to each other. This approach is called **mesh routing**. As Wi-Fi technologies improve, the number and variety of options for wireless connections to the Internet should continue to increase.

Cellular Telephone Networks

In 2003, there were about 500 million mobile phones in the world and industry experts expect that number to grow by 50 million or more per year for at least the next several years. These phones are often called cellular (or cell) phones because they broadcast signals to (and receive signals from) antennas that are placed about 3 miles apart in a grid, and the hexagonal area that each antenna covers within this grid is called a cell.

Although cell phones were originally designed to handle voice communications, they have always been able to transmit data. However, their data transmission speeds were very low, ranging from 10 Kbps to 384 Kbps. Several changes in cell phone technology have increased those speeds in today's most capable cell phones to 2 Mbps. The devices that combine the latest technologies available today are called **third-generation (3G) cell phones**. Many cell phones have a small screen and can be used to send and receive short text messages using a protocol called **short message service (SMS)**. Some cell phones and combination phone-PDAs include tiny Web browsers. These devices can provide users with Web access, e-mail connections, and other Internet services in addition to short message service. In 2004, mobile telephone and PDA manufacturers introduced the first devices (telephones and PDAs) that could automatically switch their connections between the cellular network and a local Wi-Fi network. Many industry analysts expect these types of devices to become the standard within a few years.

Today, companies that sell mobile telephone services also sell Internet access through their cellular networks. These companies offer a variety of pricing plans and service levels, but most charge a fixed fee plus a charge for the amount of data traffic sent or received during the month.

Many companies have seen great business potential for these wireless networks and the devices connected to them. They use the term **mobile commerce** or **m-commerce** to describe the kinds of resources people might want to access (and pay for) using wireless devices. You will learn more about revenue models that use wireless technologies in Chapter 3 and cost-reduction strategies that use wireless technologies in Chapter 5. In Chapter 11, you will learn how some companies are using these technologies to process online payments for goods and services. Figure 2-16 summarizes speed and cost information for the most commonly available wired and wireless options for connecting a home or business to the Internet.

Service	Upstream speed (Kbps)	Downstream speed (Kbps)	Capacity (number of simultaneous users)	One-time startup costs	Continuing monthly Costs
Residential and small business services					
Modem (POTS)	28–56	28–56	1	$0–$20	$12–$20
ISDN	128–256	128–256	1–3	$60–$300	$50–$90
ADSL	100–640	4,500–9,000	1–4	$50–$100	$40–$160
Cable modem	300–1,000	1,000–10,000	1–4	$0–$100	$40–$120
Satellite	125–150	400-500	1–3	$600–$1,200	$60–$70
Business services					
Leased digital line (DS0)	64	64	1–10	$50–$200	$40–$150
Fractional T1 leased line	128–1,544	128-1,544	5–180	$50–$800	$100–$1,000
HDSL	768-1,000	768–1,000	50–100	$300–$1,500	$400–$900
T1 leased line	1,544	1,544	100–200	$100–$2,000	$900–$1,600
T3 leased line	44,700	44,700	1,000–10,000	$1,000–$9,000	$5,000–$12,000
Large business, ISP, NAP, and Internet 2 services					
OC3 leased line	156,000	156,000	1,000–50,000	$3,000–$12,000	$9,000–$22,000
OC12 leased line	622,000	622,000	Backbone	Negotiated	$25,000–$100,000
OC48 leased line	2,500,000	2,500,000	Backbone	Negotiated	Negotiated
OC192 leased line	10,000,000	10,000,000	Backbone	Negotiated	Negotiated

FIGURE 2-16 Internet connection options

INTERNET2 AND THE SEMANTIC WEB

At the high end of the bandwidth spectrum, a group of network research scientists from nearly 200 universities and a number of major corporations joined together in 1996 to recapture the original enthusiasm of the ARPANET with an advanced research network called **Internet2**. When the National Science Foundation turned over the Internet backbone to commercial interests in 1995, many scientists felt that they had lost a large, living laboratory. **Internet2** is the replacement for that laboratory. An experimental test bed for new networking technologies that is separate from the original Internet, Internet2 has achieved bandwidths of 10 Gbps and more on parts of its network.

Internet2 is also used by universities to conduct large collaborative research projects that require several supercomputers connected at very fast speeds, or that use multiple video feeds—things that would be impossible on the Internet given its lower bandwidth limits. Internet2 promises to be the proving ground for new technologies and applications of those technologies that will eventually find their way to the Internet. One of the most important elements of Internet2 today is the **Abilene** project. Abilene is a high-bandwidth backbone built by the Internet2 organization with Indiana University and three commercial partners: Juniper Networks, Nortel Networks, and Qwest.

The Internet2 project is focused mainly on technology development. In contrast, Tim Berners-Lee has announced a project that will blend technologies and information to create a next-generation Web, which he calls the Semantic Web. The **Semantic Web** project, if successful, would result in words on Web pages being tagged (using XML) with their meanings. The Web would become a huge machine-readable database. People could use intelligent programs called **software agents** to read the XML tags to determine the meaning of the words in their contexts. For example, a software agent given the instruction to find an airline ticket with certain terms (date, cities, cost limit) would launch a search on the

Web and return with an electronic ticket that meets the criteria. Instead of a user having to visit several Web sites to gather information, compare prices and itineraries, and make a decision, the software agent would automatically do the searching, comparing, and purchasing.

The key elements that must be added to Web standards so that software agents can perform these functions include XML, a resource description framework, and an ontology. You have already seen how XML tags can describe the semantics of data elements. A **resource description framework (RDF)** is a set of standards for XML syntax. It would function as a dictionary for all XML tags used on the Web. An **ontology** is a set of standards that defines, in detail, the relationships among RDF standards and specific XML tags within a particular knowledge domain. For example, the ontology for cooking would include concepts such as ingredients, utensils, and ovens; however, it would also include rules and behavioral expectations, such as that ingredients can be mixed using utensils, that the resulting product can be eaten by people, and that ovens generate heat within a confined area. Ontologies and the RDF would provide the intelligence about the knowledge domain so that software agents could make decisions as humans would.

Most observers agree that Berners-Lee and the researchers at the W3C who are defining the elements of the Semantic Web have a great deal of complex work to do before the results are usable. You can learn more about this project by following the link in the Online Companion to the **W3C Semantic Web** pages.

Summary

In this chapter, you learned about the history of the Internet and the Web, including how these technologies emerged from research projects and grew to be the supporting infrastructure for electronic commerce today. You also learned about the protocols, programs, languages, and architectures that support the Internet and the World Wide Web. TCP/IP is the protocol suite used to create and transport information packets across the Internet. IP addresses identify computers on the Internet. Domain names such as www.amazon.com also identify computers on the Internet, but those names are translated into IP addresses by the routing computers on the Internet. HTTP is the set of rules for transferring Web pages and requests for those Web pages on the Internet. POP, SMTP, and IMAP are protocols that help manage e-mail. Unsolicited commercial e-mail (or spam) has become a major irritation for internet users.

Hypertext Markup Language, or HTML, was derived from the more generic meta language SGML. HTML defines the structure and content of Web pages using markup symbols called tags. Over time, HTML has evolved to include a large number of tags that accommodate graphics, Cascading Style Sheets, and other Web page elements. Hyperlinks are HTML tags that contain a URL. The URL can be a local or remote computer. The better HTML editors facilitate Web page construction with helpful tools and drag-and-drop capabilities.

Extensible Markup Language, or XML, is also derived from SGML. However, unlike HTML, XML uses markup tags to describe the meaning, or semantics, of the text, rather than its display characteristics. XML offers businesses hope for a common language that they will be able to use to describe products, services, and even business processes to each other in common, shared databases. XML could help companies dramatically reduce the costs of handling intercompany information flows.

Intranets are private internal networks that use the same protocols as the Internet. Employees can access the intranet and find, view, or print information just as they would Internet-based material. When companies want to collaborate with suppliers, partners, or customers, they can connect their intranets to each other and form an extranet. The three types of extranets are: public network, private network, and virtual private network. Virtual private networks, or VPNs, provide security at a low cost, whereas public network extranets have no security at all.

Internet service providers offer many different types of connections to the Internet. Basic telephone connections are the most economical and easiest to install, but they are the slowest. Broadband cable, satellite microwave transmission, and DSL services provide Internet access at relatively high speeds. Other, more expensive options provide the bandwidth that larger businesses need. A variety of wireless connection options are becoming available for businesses and homes. The wireless connection options available through cell phones show promise in creating new opportunities for revenue generation, cost reduction, and payment-processing applications.

Internet2 is an experimental network built by a consortium of research universities and businesses that provides a test bed for creating and perfecting the networking technologies of tomorrow. The W3C Semantic Web project holds out hope that in the future, many mundane user interactions with the Web will be handled by intelligent software agents.

Key Terms

802.11a, 802.11b, 802.11g, 802.11n
Anchor tag
Asymmetric connection
Asymmetric digital subscriber line (ADSL or DSL)
Asynchronous transfer mode (ATM)
Backbone router
Bandwidth
Base 2 (binary)
Bluetooth
Border router
Broadband
Bulk mail
Byte
Cascading style sheets (CSS)
Circuit
Circuit switching
Client/server architecture
Client-side scripting
Closed architecture
Closing tag
Colon hexadecimal (colon hex)
Computer network
Configuration table
Data-grade lines
Digital subscriber line (DSL)
Domain name
Dotted decimal
Download
Downstream bandwidth (downlink bandwidth)
Electronic mail (e-mail)
E-mail client software
E-mail server
Encapsulation
Extensible Hypertext Markup Language (XHTML)
Extensible Markup Language (XML)
Extensible Stylesheet Language (XSL)
Extranet
Fixed-point wireless
Fractional T1
Frame relay
Gateway computer
Generalized Markup Language (GML)
Graphical user interface
Hexadecimal (base 16)
Hierarchical hyperlink structure
High-speed DSL (HDSL)
Home page
Hot spot
HTML extensions
Hypertext
Hypertext element
Hypertext link (hyperlink)
Hypertext Markup Language (HTML)
Hypertext server
Hypertext Transfer Protocol (HTTP)
Integrated Services Digital Network (ISDN)
Interactive Mail Access Protocol (IMAP)
Internet
Internet access provider (IAP)
Internet backbone
Internet host
Internet Protocol (IP)
Internet Protocol version 4 (IPv4)
Internet Protocol version 6 (IPv6)
Internet service provider (ISP)
Internet2
Intranet
IP address
IP tunneling
IP wrapper
Leased line
Linear hyperlink structure
Local area network (LAN)
Mailing list

Markup tags (tags)
Mesh routing
Meta language
Mobile commerce (m-commerce)
Multipurpose Internet Mail Extensions (MIME)
Net bandwidth
Network access points (NAPs)
Network access provider
Network Address Translation (NAT) device
Network Control Protocol (NCP)
Network specification
Newsgroup
Octet
Ontology
Open architecture
Opening tag
Optical fiber
Packet
Packet-switched
Personal area network (PAN)
Piconet
Plain old telephone service (POTS)
Post Office Protocol (POP)
Private IP address
Private network
Proprietary architecture
Protocol
Public network
Repeater
Resource description framework
Roaming
Router
Router computer (routing computer)
Routing algorithm
Routing table
Scaling problem
Semantic Web
Short message service (SMS)
Simple Mail Transfer Protocol (SMTP)
Site map
Software agents
Spam
Sponsored top-level domain (sTLD)
Standard Generalized Markup Language (SGML)
Start page
Style sheet
Subnetting
Symmetric connection
T1
T3
TCP/IP
Text markup language
Third-generation (3G) cell phones
Top-level domain (TLD)
Transceiver
Transmission Control Protocol (TCP)
Ultra Wideband (UWB)
Uniform Resource Locator (URL)
Unsolicited commercial e-mail (UCE)
Upload bandwidth
Upstream bandwidth
Usenet (User's News Network)
Virtual private network (VPN)
Voice-grade lines
Web
Web browser
Web browser software
Web client computer
Web client software
Web server
Web server software
Wi-Fi (wireless ethernet, 802.11b, 802.11a, 802.11g, 802.11n)
Wide area network (WAN)
Wireless access point (WAP)
World Wide Web (WWW)
World Wide Web Consortium (W3C)
XML parser
XML vocabulary

Review Questions

RQ1. What were the main forces that led to the commercialization of the Internet? Summarize your answer in about 100 words.

RQ2. Describe in two paragraphs the origins of HTML. Explain how markup tags work in HTML, and describe the role of at least one person involved with HTML's development.

RQ3. In about 200 words, compare the POP e-mail protocol to the IMAP e-mail protocol. Describe situations in which you would prefer to use one protocol or the other and explain the reasons for your preference.

RQ4. In about 400 words, describe the similarities and differences between XML and HTML. Provide examples of at least two situations in which you would use XML and two situations in which you would use HTML.

RQ5. Use your favorite search engine and the links in the Online Companion (under the heading **Internet Connection Options**) to search for more information about broadband satellite connections, DSL connections, wireless connections, and cable connections. Prepare a four-column table (one column for each technology) in which you list the advantages and disadvantages of each connection method. Include at least two advantages and two disadvantages for each connection method.

Exercises

E1. You are the assistant to Julie Davidson, the sales manager of Old Reliable Life Insurance Company. Julie is interested in equipping her sales force with the technology they need to sell Old Reliable's insurance products. Most of her salespeople visit customers in their homes or offices. Today, the salespeople carry a laptop computer to show value projections and cash flow summaries for various policies. Many of them also carry a PDA for appointments and a mobile phone. Julie would like to ensure that salespeople have access to the home office server computers while they are making their sales presentations to customers. This access will let salespeople download the latest product information and obtain online assistance from office staff and in-house experts when the salespeople get a question from a customer that they are not able to answer. A correct and quick answer to a customer's question can often help close a difficult sale. Julie asks you to investigate various options for giving salespeople remote access to the home office server computers. She wants you to consider both wireless (directly to the laptop computers or through salespeoples' cell phones or PDAs) and wired options. Prepare a report for Julie in which you briefly review at least four options, writing no more than three paragraphs for each option. Then choose the best wired option and the best wireless option and write a one-page evaluation of strengths and weaknesses for each of them. Use the Online Companion links and your favorite Web search engines to do your research.

E2. Bridgewater Engineering Company (BECO), a privately held machine shop, makes industrial-quality, heavy-duty machinery for assembly lines in other factories. It sells its presses, grinders, and milling equipment using a few inside salespeople and telephones. This traditional approach worked well during the company's start-up years, but BECO is getting a lot of competition from abroad. Because you worked for the company during the summers of your college years, BECO's president, Tom Dalton, knows you and realizes that you are Web savvy. He wants to form close relationships with the steel companies and small-parts manufacturers that are BECO's suppliers so that he can tap into their ordering systems and request supplies when he needs them. Tom wants you to investigate how he can use the Internet to set up such electronic relationships. Use the Web and the links in the Online Companion to locate information about extranets and VPNs. Write a report that briefly describes how companies use extranets to link their systems with those of their suppliers, then write an evaluation of at least two companies (using information you have gathered in your Web searches) that could help develop an extranet that would work for Tom. Close the report with an overview of how BECO could use VPN technologies in this type of extranet. The three parts of your report should total about 700 words.

E3. Frieda Bannister is the IT manager for the state of Iowa's Department of Transportation (DOT). She is interested in finding ways to reduce the costs of operating the DOT's vehicle repair facilities. These facilities purchase replacement parts and repair supplies for all of the state's cars, trucks, construction machinery, and road maintenance equipment. Frieda has read about XML and thinks that it might help the DOT send orders to its many suppliers throughout the country more efficiently. Use the Online Companion links, the Web, and your library to conduct research on the use of XML in state, local, and federal government operations. Provide Frieda with a report of about 1000 words that includes sections that discuss what XML is and explain why XML shows promise for the ordering application Frieda envisions. Your report should also identify other DOT business processes or activities that might benefit from using XML. The report should also include a summary of the main disadvantages of using XML today for integrating business transactions. End the report with a brief summary of how the W3C Semantic Web project results might help the DOT operate more efficiently in the future.

E4. As you learned in this chapter, XML allows users to define their own markup tags. You also learned that this flexibility can lead to problems when IT professionals who have developed tag sets for their own organizations are asked to share information with other organizations that are using other tag sets. One way organizations can avoid this problem is to agree to follow common standards. A common standard for financial information is XBRL. Accountants and financial analysts around the world have agreed to use XBRL to format financial statements and other reports. In about 300 words, outline the advantages that companies and financial analysts can obtain by using the XBRL standard. You can research this subject in your school library or online using your favorite search engine and the links provided for this exercise in the Online Companion.

Cases

C1. Covad

In 1996, three enterprising executives decided to leave their jobs at Intel and form a company that would take advantage of an opportunity provided by the recently enacted Telecommunications Act of 1996. The law eliminated the monopoly that local telephone companies had held and allowed other companies to offer telecommunications services to businesses and individuals in what had been the local telephone companies' protected service areas. Since the goal of the company was to offer converged voice and data services, the founders named the company **Covad**.

During its first two years, the company became a solid regional company in the San Francisco Bay and Silicon Valley areas that sold Internet access to businesses and ISPs. Its ISP customers provided DSL access to smaller businesses and residential customers. But the Internet boom was in full swing, and in 1998, Covad hired U.S. West senior vice president Robert Knowling to take the company to the next level. Over the next two years, Covad raised more than $2 billion from stock and bond offerings and expanded into 98 metropolitan areas throughout the country. By the end of 1999, it had more than 200,000 customers, including AOL, MCI, and some of the fastest growing regional ISPs in the country. It was following the lead of its main competitor, NorthPoint Communications (featured in this chapter's Learning from Failures feature), and pursuing a strategy that included rapid expansion using external funds.

In 2000, Covad's largest customers, the ISPs that sold Internet access to smaller companies and residential users, stopped paying their Covad bills because their customers were disappearing. The Internet bubble had burst. Covad had expanded too fast and was in serious trouble. It had put all of its investors' money into equipment and infrastructure during its rapid growth and had no cash reserves to take it through a period of slower growth.

In 2001, the company brought in a new manager, Charlie Hoffman, to take the company through a Chapter 11 bankruptcy reorganization. Covad's bondholders received 19 cents on each dollar after the reorganization. This gave the company a much lower debt payment load and allowed it to focus on rebuilding its business. Hoffman changed the basic strategy of the company by decreasing its emphasis on sales to ISPs who would resell Internet DSL access to small businesses and residential customers. Instead, Covad began selling these access services directly to those customers. By 2004, Covad had more than 700,000 DSL and T1 customers and was getting close to making a profit.

Although the company's original plans were to sell both voice and data services, Covad grew rapidly selling only data services. In 2004, Covad began offering telephone voice services over its data lines (this service is called voice-over-IP, or VoIP). In the direct market for Internet access and VoIP services, Covad faces serious competition from cable companies, who have offered Internet access for many years and who are now also offering VoIP services. Covad's Internet access services also face competition from independent DSL ISPs (including many of Covad's former customers) and from satellite and fixed-point wireless access providers. Finally, Covad's VoIP services also face competition from local telephone companies and the large national mobile telephone service providers.

Required:

1. Prepare a 200-word analysis in which you describe why Covad's recent strategy of selling Internet access directly rather than through ISPs has been successful.
2. Use the links in the Online Companion for this chapter, your favorite search engine, and resources in your library to learn more about Covad and its current competitors in the cable industry (primarily **Cox Cable**, **Comcast**, and **Time Warner Cable**, but there are others). Prepare a report of about 600 words in which you outline and analyze the strengths and weaknesses that Covad has with respect to each of these competitors.
3. Use the links in the Online Companion for this chapter, your favorite search engine, and resources in your library to learn more about Covad and its current competitors in the local telephone company and national wireless industries. Prepare a report of about 600 words in which you outline and analyze the strengths and weaknesses that Covad has with respect to each of these competitors.

Note: Your instructor might assign you to a group to complete this case, and might ask you to prepare a formal presentation of your results to your class.

C2. Portable Fun Instruments

Yash Gupta is the founder and president of Portable Fun Instruments (PFI), a company that has had great success in the handheld game market. Its first products were dedicated handheld devices that each offered a specific game, such as backgammon, checkers, or chess. As the power of microprocessors grew, and the size and cost of those microprocessors shrank, PFI was able to build better and more complex games into its devices.

Today, PFI offers a wide variety of dedicated devices on which users can play card games, adventure games, and sports simulations, and solve various kinds of puzzles. Most of the elements in the game displays are graphics, not words. This helps PFI sell the devices in many different markets around the world without having to build separate interfaces for each language. PFI's game devices have retail prices that range between $30 and $70, but the retailers and distributors buy them from PFI for prices that range between $12 and $25.

PFI is profitable because Yash has worked hard to keep development and production costs low. Most of the programming is done in Bangalore, India, and the devices are built in production facilities located in Xixiang, China and Penang, Malaysia. Although Yash has been successful in controlling production costs, he worries about continuing to operate the company with a long-term strategy that requires PFI to build a new physical device for each sale. The large retail chains that have become PFI's main customers are always asking for discounts and reduced prices on new orders, and production costs are creeping upward even though the facilities are located in some of the lowest-cost areas in the world.

Yash wants to explore the potential PFI has for moving its games to other platforms. PFI has translated some of its games to PDA platforms, such as the Palm Operating System and the Microsoft Windows CE Operating System (sometimes called the Pocket PC system), but the results have been disappointing. Most PDA users are businesspeople who use their PDAs for appointments, address books, travel expenses, and other data management functions. These users are not avid game players, and sales of PFI's games for these platforms have also not been strong.

Some of PFI's marketing team members have been telling Yash about the success that Japan's DoCoMo has had offering a variety of entertainment products that are downloaded from the Internet and that are displayed on mobile phones. DoCoMo charges users for their use of these products, which include games, and shares those fees with the providers of the products. Yash has also heard that a U.S. mobile phone technology company, Qualcomm, is offering a similar service called Binary Runtime Environment for Wireless (BREW). Yash believes that these markets might be worth pursuing.

Yash has hired you as a consultant to investigate DoCoMo, BREW, and any similar delivery systems for selling online access to PFI's games to mobile phone users. He is also interested in learning more about the programming and markup languages used by these delivery systems. The programming teams in Bangalore are becoming familiar with XML because they are doing contract programming projects for other companies, and Yash wants to know if these skills will help them adapt PFI's games to mobile phone delivery systems.

Required:

1. Use the links in the Online Companion for this case, your favorite search engine, and resources in your library to learn more about DoCoMo, BREW, and similar content delivery systems for mobile phones. Prepare a 400-word executive summary for Yash that describes each delivery system, identifies the company or companies behind each system, and outlines the current availability of each system for content providers such as PFI.
2. Prepare a report for Yash and the PFI executive team in which you outline and analyze the strengths and weaknesses of each content delivery system. Your report should conclude with a specific recommendation regarding the suitability of each content delivery system for PFI's games. This report should be about 500 words in length.
3. Prepare a 300-word report that describes the programming and markup languages used in each delivery system. In this report, discuss whether the PFI programming team's experience with XML will help it adapt PFI's games to mobile phone content delivery systems.

Note: Your instructor might assign you to a group to complete this case, and might ask you to prepare a formal presentation of your results to your class.

For Further Study and Research

Alschuler, L. 2001. "Getting the Tags In: Vendors Grapple with XML-Authoring, Editing and Cleanup," *Seybold Report on Internet Publishing*, 5(6), February, 5–10.

Angwin, J. 2002. "ICANN Leader Seeks Big Changes in How Internet Is Governed," *The Wall Street Journal*, February 26, B6.

Babcock, C. 2001. "XML Databases Offer Greater Search Capabilities," *Interactive Week*, 8(18), May 7, 11–13.

Bannan, K. 2002. "Satellite: The Only Game Out of Town," *PC Magazine*, 21(3), February 21, 99.

Benner, J. 2002. "Getting a Lock on Broadband," *Salon.com*, June 7. (http://www.salon.com/tech/feature/2002/06/07/broadband/print.html)

Bergman, M. 2001. *The Deep Web: Surfacing Hidden Value*. Sioux Falls, SD: BrightPlanet.com. (http://brightplanet.com/technology/deepweb.asp)

Bonner, P. 2002. "The Semantic Web," *PC Magazine*, 21(13), July, IP01–IP02.

Bosak, J. and T. Bray. 1999. "How XML Will Fix the Web: Tags Categorizing Facts, Not Formats, Speed Up Transactions," *Scientific American*, 280(5), May, 89.

Boyle, M. 2002. "The Shiniest Reputations in Tarnished Times," *Fortune*, 145(5), March 4, 70–72.

Brewin, B. 2004. "Michigan City Turns on Citywide Wi-Fi," *Computerworld*, July 30. (http://www.computerworld.com/mobiletopics/mobile/wifi/story/0,10801,94928,00.html)

Campbell, T. 1998. "The First E-Mail," *Pretext Magazine*, March. (http://www.pretext.com/mar98/features/story2.htm)

Caulfield, B. 2002. "Wi-Fi Goes to Work," *Business 2.0*, 3(6), June, 122–123.

Computergram Weekly. 2003. "Cisco Sees Flat Sales But Chambers More Cheerful," May 7, 7–8.

Cope, J. 2001. "IPv6: Is It Inevitable?" *Computerworld*, 35(22), May 28, 58–59.

Costa, D. 2001. "Cable: This Technology Is the Simplest and Most Popular Option," *PC Magazine*, 20(3), February 6, 149–151.

Davis, K. and E. Burt. 2001. "Mad as Hell about DSL," *Kiplinger's Personal Finance Magazine*, 55(7), July 2001, 84–85.

Dornan, A. 2003. "Unwiring the Last Mile," Network Magazine, 18(1), January, 34–37.

Duffy, J. 2003. "RBOCs and Cable Wage Turf War," Network World, August 18, 11–14.
Dyck, T. 2002. "Going Native: XML Databases," *PC Magazine*, 21(12), June 30, 136–139.
The Economist. 2002. "ICANNOT," March 2, 59.
Eisenzopf, J. 2002. "Is VoiceXML Right for Your Customer Service Strategy?" *New Architect*, 7(3), March, 20–21.
Elgin, B. and M. Roman. 2002. "Cisco Rides Again," *Business Week*, May 20, 50.
Fensel, D., J. Hendler, H. Lieberman, and W. Wahlster. 2002. *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. Cambridge, MA: MIT Press.
Fitchard, K. 2004. "Covad's Quiet Authority," *Telephony*, 245(12), June 7, 34–39.
Fixmer, R. 2002. "Broadband Homeland," *eWeek*, March 4, 41–43.
Floyd, M. 2002. "XML Exposed," *PC Magazine*, 21(12), June 30, 132–140.
Fruhlinger, J. 2002. "Broadband and the New User Experience," *Web Techniques*, 7(2), February, 23–25.
Garcia, J. and J. Wilkins. 2001. "Cable Is Too Much Better to Lose," *The McKinsey Quarterly*, January, 185–188.
Goldfarb, C. 1981. "A Generalized Approach to Document Markup," *ACM Sigplan Notices*, (16)6, June, 68–73.
Hannon, N. and M. Willis. 2005. "Combating Everyday Data Problems with XBRL" *Strategic Finance*, 87(1), July, 57–59.
Hannon, N. and M. Willis. 2005. "Combating Everyday Data Problems with XBRL, Part 2," *Strategic Finance*, 87(2), August, 59–61.
Hardy, Q. 2001. "Cisco Kidding?" *Forbes*, 167(11), May 14, 52–53.
Hardy, Q. 2002. "The Great Wi-Fi Hope," *Forbes*, 169(6), March 18, 56–64.
Hawn, C. 2001. "Management By Stock Market: NorthPoint Rode the Web Wave," *Forbes*, 167(10), April 30, 52–53.
Henderson, T. and T. Ritchey. 2002. "Bring in the (802.11) A Team," *Network World*, 19(4), January 28, 55–56.
Henschen, D. 2005. "XBRL Offers a Faster Route to Intelligence," 8(8), August, 12.
Hochmuth, P. 2002. "What Customers Want From Cisco," *Network World*, 19(2), June 3, 1.
Huston, J. 2001. "Scaling the Internet," *Satellite Broadband: The Cutting Edge of Satellite Communications,* 2(3), March, 18–22.
Kalata, K. 2001. *Internet Programming with VBScript and JavaScript.* Boston: Course Technology.
Kaven, O. 2004. "Wired Ethernet and 802.11g Outpace the Rest," *PC Magazine*, 23(6), April 6, 104.
Kristof, N. 2005. "When Pigs Wi-Fi," *The New York Times*, August 7, 13.
LaBarba, L. 2001. "DSL Pains Reach End Users," *Telephony*, 240(15), April 9, 14–15.
Lawson, S., K. Miyake, and J. Evers. 2002. "IPv6 Enters the Real World," *InfoWorld*, 24(7), February 18, 35–36.
Liebman, L. 2001. "XML's Tower Of Babel," *InternetWeek*, April 30, 25–26.
Lowry, T. 2004. "Satellite's Hot Pursuit of Cable, *Business Week*, May 24, 46.
Malnig, A. 2005. "XBRL: Deep Drilling for Financials," *Seybold Report: Analyzing Publishing Technologies*, 5(4), May 18, 11–14.
Markoff, J. 2002. "Two Tinkerers Say They've Found a Cheap Way to Broadband," *The New York Times*, June 10, C1.
Marsan, C. 2004. "It's a New Domain-Name Game," *Network World*, March 1, 1, 14.
Martin, M. 2002. "Not Part of the Cable and DSL Boom?" *Network World*, February 11, 21.

Metz, C. 2002. "The All Mail Revue," *PC Magazine*, 21(9), May 7, 85–99.

Milstein, S. 2003. "Be Your Own Wireless Network," *The New York Times,* February 27, 4.

Mitchell, R. 2002. "Wireless at Full Throttle," *Computerworld*, 36(20), May 13, 63.

Nelson, T. 1987. *Literary Machines*. Swarthmore, PA: Nelson.

Nielsen, J. 2003. "Mobile Devices: One Generation From Useful," *Alertbox,* August 18. (http://www.useit.com/alertbox/20030818.html)

Nolle, T. 2002. "Are Cable Companies the Key to Local Access?" *Network Magazine*, April, 104.

Nolle, T. 2002. "Why Is Cisco Making Money?" *Network World*, 19(22), June 3, 51.

O'Connor, R. 2000. "Under Construction: Two Research Groups Work to Build a Better Internet," *Interactive Week*, 7(46), November 13, 44–48.

Olivia, R. 2001. "The Promise of XML," *Marketing Management*, 10(1), Spring, 46–49.

Panko, R. 2005. *Business Data Networks and Telecommunications*. Fifth Edition. Upper Saddle River, NJ: Prentice Hall.

PC Magazine. 2001. "Turn XML into HTML," 20(11), IP01–IP04.

Pimm, F. 2001. "Boeing Shows How XML Can Help Business," *Computerworld*, 35(11), March 12, 28–29.

Port, O. 2002. "The Next Web," *Business Week*, March 4, 96–102.

Pringle, D. 2005. "Wi-Fi Woes: Wireless Networks Are Great—If You Can Figure Out How to Set Them Up," *The Wall Street Journal*, July 18, R11.

Rendleman, J. 2002. "Cisco Positioned to Profit From Changing Market," *Information Week*, May 13, 32.

Richtel, M. 2004. "Where Entrepreneurs Go and the Internet Is Free," *The New York Times*, June 7. (http://www.nytimes.com/2004/06/07/technology/07wifi.html)

Robertson, D. 2001. "Tweaking Protocols," *Satellite Broadband: The Cutting Edge of Satellite Communications*, 2(2), February, 26–28.

Rupley, S. 2005. "I'm Drowning in Cables," *PC Magazine*, 24(14), August 23, 100.

Schonfeld, E. 2002. "Unwiring the Masses," *Business 2.0*, 3(6), June, 18–20.

Schrick, B. and M. Riezenman. 2002. "Wireless Broadband in a Box," *IEEE Spectrum*, June. (http://www.spectrum.ieee.org/WEBONLY/publicfeature/jun02/wire.html)

Shafer, S. 2002. "Secure Extranets Arrive," *InfoWorld*, 24(10), March 11, 27.

Simon, B. 2003. "Some Bet the Future of Broadband Belongs to Regional Bells, Not Cable," *The New York Times,* July 21. (http://www.nytimes.com/2003/07/21/technology/21BROA.html)

Spangler, T. 2002. "Crossing the Broadband Divide," *PC Magazine*, 21(3), February 21, 92–101.

Swanson, S. 2002. "Post Office In—and Out—of Wireless Early," *Information Week*, April 8, 24.

Tie, R. 2005. "XBRL: It's Unstoppable: Interview With Charles Hoffman," *Journal of Accountancy*, August, 32–35.

Thurm, S. 2002. "Cisco Profit Exceeds Expectations," *The Wall Street Journal*, May 8, A3.

Vogelstein, F. 2004. "The Cisco Kid Rides Again," *Fortune*, 150(2), July 26, 132–137.

Weber, T. 2002. "Wanted: A Peace Envoy to End Net's Bickering Over Address System," *The Wall Street Journal*, March 25, B1.

White, C. 2002. *Data Communications and Computer Networks: A Business User's Approach,* 2nd Edition. Boston: Course Technology.

Wildstrom, S. 2002. "Broadband Needs Home Improvement," *Business Week*, April 1, 19.

Witte, G. 2003. "Bringing Broadband Over the Mountain: Roadstar Puts Wireless Technology to the Test," *The Washington Post*, September 15, E1.

Wylie, I. 2005. "Bluetooth Killers," *Fast Company*, July, 30.
Zhang, M. and R. Wolff. 2004. "Crossing the Digital Divide: Cost-Effective Broadband Wireless Access for Rural and Remote Areas," *IEEE Communications Magazine*, 42(2), February, 99–105.