Technology Guides

TECHNOLOGY GUIDE

5

The Internet and the Web

T5.1 WHAT IS THE INTERNET?[1]

The **Internet ("the Net")** is a network that connects hundreds of thousands of internal organizational computer networks worldwide. Examples of internal organizational computer networks are a university computer system, the computer system of a corporation such as IBM or McDonald's, a hospital computer system, or a system used by a small business across the street from you. Participating computer systems, called **nodes,** include PCs, local area networks, database(s), and mainframes. A node may include several networks of an organization, possibly connected by a wide area network. The Internet connects to hundreds of thousands of computer networks in more than 200 countries so that people can access data in other organizations, and can communicate and collaborate around the globe, quickly and inexpensively. Thus, the Internet has become a necessity in the conduct of modern business.

Brief History

The Internet grew out of an experimental project of the Advanced Research Project Agency (ARPA) of the U.S. Department of Defense. The project was initiated in 1969 as ARPAnet to test the feasibility of a wide area computer network over which researchers, educators, military personnel, and government agencies could share data, exchange messages, and transfer files.

From four nodes at its initiation, the Internet has grown to millions of nodes today. The major growth occurred after commercial organizations were allowed to join ARPAnet, renamed the Internet in 1993. Each node may connect many individual users, sometimes tens of thousands of them. There are over 500 million Internet users today.

The computers and organizational nodes on the Internet can be of different types and makes. These are connected to each other by data communication lines of different speeds. The main network that links the nodes is referred to as the **backbone,** a fiber-optic network currently operated mainly by telecommunications companies such as MCI.

No central agency manages the Internet. The cost of its operation is shared among hundreds of thousands of nodes, and therefore the cost for any one organization is small. Organizations must pay a small fee if they wish to register their name, and they need to have their own hardware and software for the operation of their internal networks. The organizations are obliged to move any data or information that enters their organizational network, regardless of its source, to its internal destination, at no charge to the senders. The senders, of course, pay the telephone bills for using the backbone or regular telephone lines.

Next Generation Internet Services

Very-high-speed Backbone Network Service (vBNS) is a high-speed network designed to support the academic Internet2 (see below) and the government-sponsored Next-Generation Internet (NGI) initiative. The vBNS was first implemented as an OC-3 (155-Mbps) backbone but has been upgraded to OC-12 (622 Mbps). The goal is to increase the vBNS backbone to OC-48 (2.4 Gbps).

INTERNET2. **Internet2** is a collaborative effort by over 200 U.S. universities, working with over 70 companies and more than 40 affiliated organizations and government, to develop advanced Internet technologies and applications vital to

---

[1]See Internet 101 at *Internet-101.com, Internet101.org,* and *stimulus.com.*

support the research and education goals of higher education. Internet2 was announced in October 1996 with 34 universities, in October 1997 the University Corporation for Advanced Internet Development (UCAID) was formed to administer the Internet2 project. The primary goals of Internet2 are to create a leading-edge network capability for the national research community; enable revolutionary Internet applications; and ensure the rapid transfer of new network services and applications to the broader Internet community.

It focuses on advanced network applications, new network capabilities, middleware, and high-performance networks.

Internet2 will use part of vBNS as a backbone for providing high-speed connectivity to many of its member universities. The communications infrastructure designed to support Internet2 consists of 10 Mbs (nonshared) to the desktop, a 500 Mbs campus wide backbone, and at least 155 Mbs connectivity to GigaPoP that is a proposed national high-speed network. It supports Quality-of-Service (QoS) technology, which allows two hosts to establish a connection with a guaranteed bandwidth that is necessary for live video applications. It is expected to be 100 to 1000 times faster than the current Internet infrastructure. For details, see *internet2.edu.*

NEXT-GENERATION INTERNET.   The **Next-Generation Internet (NGI)** initiative is a multiagency (with the NSF, NIE, etc.), U.S. federal government research and development program that is developing advanced networking technologies and revolutionary applications that require advanced networking. The broad goals of the NGI initiative are to research and develop advanced end-to-end networking technologies, focusing primarily on reliability, robustness, security, quality of service guarantees for multicasting and video, and bandwidth allocation. In other words, the NGI initiative aims to create an Internet that is fast, always on, everywhere, natural, intelligent, easy, and trusted. To this end, NGI is building two test beds, one at speeds at least 100 times faster than today's Internet, and the other 1000 times faster. The university-led Internet2 and the federally-led Next Generation Internet (NGI) are parallel and complementary initiatives. For details, see *ngi.gov.*

INTERNET GOVERNANCE.   There is no CEO or board of directors of the Internet. However, the following organizations govern the various operations of the Internet:

- *Internet Society:* Enables global cooperation regarding the Internet.
- *World Wide Web Consortium (W3C):* Develops standards for the Web and promotes new Web technologies.
- *Internet Engineering Task Force (IETF):* Develops technical Internet standards and protocols.
- *Internet Architecture Board (IAB):* Oversees the Internet standards process.
- *Internet Engineering Steering Group (IESG).* Approves technical Internet standards.

T5.2   BASIC CHARACTERISTICS AND CAPABILITIES OF THE INTERNET

Accessing the Internet

There are several ways to access the Internet. From your place of work or your university, you can access an Internet-connected file server on your organization's LAN. A campus or company backbone connects all the various LANs and

servers in the organization to the Internet. You can also log onto the Internet from your home or on the road; wireless connections are also possible. Using a modem you can connect to your campus or company network from any place that offers a dial tone. Connection to the Internet is also available through commercial providers such as America Online (AOL), for which you pay a fee. Many telephone providers, such as AT&T and MCI, also sell Internet access, as do computer companies such as Microsoft. Such companies are called Internet service providers (ISPs). Following are the main Internet connection methods.

CONNECTING VIA LAN SERVER.   This approach requires the user's computer to have specialized software called a *communications stack,* which provides a set of communications protocols that perform the complete functions of the seven layers of the OSI communications model. LAN servers are typically connected to the Internet at 56 Kbps or faster. This type of connection is expensive, but the cost can be spread over multiple LAN users.

CONNECTING VIA SERIAL LINE INTERNET PROTOCOL/POINT-TO-POINT PROTOCOL (SLIP/PPP).   This approach requires that users have a modem and specialized software that allows them to dial into a SLIP/PPP server through a service provider at a cost of approximately $30 per month or less. This type of connection is advantageous, for example, for employees working at home who need to access the Internet or their own company's intranet.

CONNECTING VIA AN ONLINE SERVICE.   This approach requires a modem, standard communications software, and an online information service account with an Internet service provider. The costs include the online service fee, a per-hour connect charge, and, where applicable, e-mail service charges.

One of the first **online services (OLS),** CompuServe, was launched in 1979. Prodigy and AOL (America Online) followed with offerings targeted toward home computer users. Services include, for example, access to specialized news sources (business, sports, etc.), market information, maps, and weather. This focus contrasted with the Internet's original orientation, which was geared to the academic and research community, and that of online research services such as Lexis-Nexis (a division of Reed Elsevier), which targets large corporations.

Originally, **Internet service providers (ISPs)** were established to provide connectivity, not content. Many ISPs now offering dial-in Internet access to consumers initially were set up to provide dedicated Internet connections to educational and commercial organizations. Others (such as NetCom) began by supplying dial-in access to Internet-connected time-sharing systems. UUNet started by providing dial-up connections for routing of e-mail and Usenet news (discussion groups devoted to specialized areas of interest) between non-Internet-connected sites.

Thanks to the growth of Internet usage (as well as deregulation of the telecommunications industry), the number of ISPs has grown rapidly. Gartner Group estimates that there were approximately 13,000 ISPs worldwide at the end of 2001. Leading ISPs, which also now provide online content as well, include America Online (*aoltimewarner.com*), MSN (*msn.com*), United Online (*unitedonline.net*), Earthlink (*earthlink.com*), and Prodigy (*prodigy.com*).

NSP, NAP, ROUTERS, AND PACKETS.   The Internet is composed of thousands of interconnected networks, managed by multiple Internet service providers (ISPs) who control the routing of packets and their performance. Routing of packets between such networks is accomplished through the Border Gateway Protocol, which tells a receiving network how to reach customer addresses connected to the sending network. Inside this Internet connection, there are two features: *network service providers (NSPs)* and *network access points (NAPs).*

NSPs, commonly known as tier-1 carriers, are responsible for the backbone's traffic connected by a high-speed connection (called *peering circuits*). Most ISPs operate regional networks and rely on the backbones of the NSPs to provide the required long-distance connectivity. NSPs can serve their ISP customers either through dedicated access circuits or through interconnections at public NAPs.

NAPs, which were created by the U.S. National Science Foundation (NSF), are the access points at which service providers can meet and exchange packets traffic. The NAP operator provides switching equipment to interconnect service providers present at the NAP. For more information on NSPs and NAPs, see *mitretek.org/pubs/sigma/summer2003/1.pdf.*

THE TV AS A CHANNEL TO THE INTERNET.   For people who do not own a computer, television can also provide access to the Internet. Using a television set, a special connection device, and a telephone connection, viewers can surf the World Wide Web in their living rooms. The concept was pioneered by Web TV Networks, which manufactures an add-on device for the television, but several other companies are also developing similar devices. At the present time the quality of the Web image on a television screen is poor compared to what you see on a computer screen. It may take a few years before this situation improves.

An alternative is a combined PC and TV. This is a PC with a TV feature. It takes advantage of the convergence of telecommunications, television, and computer to deliver video and other multimedia content over the Internet at low cost. Technologies are now available to enable PCs to decode TV signals, and to receive full-motion video over the Internet. LCD TV monitors also enable access to the Internet through a $100 set-top box.

OTHER ACCESSES.   There have been several attempts to make access to the Net cheaper, faster, and/or easier. Special Internet terminals called "Internet Lite" or *network computers (NCs)* (see Technology Guide 1) are used by some companies. Also, **Internet kiosks** are terminals placed in public places like libraries and coffee houses, and even in convenience stores in some countries, for use by those who do not have computers of their own. (For examples, see Payphone Cyberbooth at *atcominfo.com*.) Accessing the Internet from cell phones and pagers is growing very rapidly.

BROADBAND AND DSL.   **Broadband** refers to telecommunication that provides multiple channels of data over a single communications medium, typically using some form of frequency or wave division multiplexing. There are several broadband technologies. These include Global System for Mobile communication (GSM), High-Speed Circuit-Switched Data (HSCSD), POTS, frame relay, GPRS (General Packet Radio Services), Integrated Services Digital Network (ISDN), xDSL, AppleTalk, Enhanced Data GSM Environment (EDGE), Satellite, and Fiber-Distributed Data Interface (FDDI).

There are several methods by which users can connect to the Internet—namely, through a dial-up modem, through a local area network, through a cable modem, or through a digital subscriber line (DSL) connection. **DSL** is a very high speed connection that uses the same wires as a regular telephone line. At the customer's location there is a DSL transceiver, which may also provide other services. At the DSL service provider's location, there is a DSL Access Multiplexer (DSLAM) to receive customer connections.

There are advantages and disadvantages of using DSL connection method. Advantages are:

- You can leave your Internet connection open and still use the phone line for voice calls.
- The speed is much higher than a regular dial-up modem (1.5 or even 10 Mbps versus 56 Kbps)
- DSL does not necessarily require new wiring; it can use the phone line you already have.
- The company that offers DSL will usually provide the modem as part of the installation.
- You can use a wireless system with the DSL.

Disadvantages and limitations of DSL are:

- A DSL connection works better when you are closer to the provider's central office. Otherwise, the connection may have interruptions, or it will slow down.
- The connection is faster for receiving data (i.e., downstream, usually at up to 10 Mbps) than it is for sending data (i.e., upstream, usually at 640 Kbps) over the Internet.
- The service is not available everywhere.

For further details, see *cis.state.mi.us/mpsc/comm/broadband/mediums/tech_def.htm.*

The TCP/IP Protocol Procedures and rules for transferring data across the Internet are called telecommunications **protocols.** The original participants of the Internet, mostly universities, used *TCP/IP* (described in Technology Guide 4), which is now the **Internet Protocol (IP).** The information that passes through the Internet is divided into small portions, called **packets,** whose creation and transmission are governed by TCP/IP to provide for more consistent delivery and control.

One member of the TCP/IP family of protocols is **Telnet,** a service that allows users to connect to computers other than their own and interactively search for files, text, software, and so forth. This enables researchers who commute to conduct their research without interruption. You can also use Telnet to connect to your organization's computer system when you are traveling, so you can check your e-mail, send messages, or conduct other activities.

STANDARDS AND PROTOCOLS. *Standards* are a set of rules or specifications for the design or operation of a computing device. There are *proprietary standards,* which are those developed and promulgated by companies in the hope of assuring or increasing their market share, and *open standards,* which are published and available for use by anyone. Either type may become a *de facto* standard, a

set of rules or specifications that comes into such widespread use in the marketplace that it becomes normative, or a *de jure* standard, a standard given the endorsement of an official standards body such as the International Organization for Standardization (ISO).

Protocols are standard sets of rules that govern network communications functions by describing both the format that a message must take and the way in which messages are exchanged between computers.

Quite often the word *standard* is used to refer to so-called industry standards, i.e., relatively stable principles and practices applied by leading manufacturers. This is a very vague concept. Between official standards and "industry standards," there is a wide range of "standards" that have been agreed on in some forum. For more information, see *collectionscanada.ca/9/1/p1-253-e.html.*

IPv4 AND IPv6.   *IPv4* is version 4 of the Internet Protocol (IP). It was the first version of the Internet Protocol to be widely deployed, and it forms the basis for most of the current Internet. Nearly 20 years old, IPv4 has been remarkably resilient in spite of its age, but it is beginning to have problems. Most significant of these problems is the growing shortage of IPv4 addresses, which are needed by all new machines added to the Internet.

IPv6 is version 6 of the Internet Protocol. It is the next-generation protocol designed by the Internet Engineering Task Force (IETF) to replace the current version (IPv4). IPv6 fixes a number of problems in IPv4, such as the limited number of available IPv4 addresses (128-bit instead of 32-bit). It also adds many improvements to IPv4 in areas such as routing (hierarchical routing, mobility) and network autoconfiguration to increase the security during the data transfer (IPSec: authentication and encryption). IPv6 is expected to gradually replace IPv4, with the two coexisting for a number of years during a transition period. For details, see *ipv6.org/.*

Internet Resources   Some information sources on the Internet are free. For example, using Telnet, you can access libraries and conduct research. Alternatively, the information may be owned by a commercial online service. As described above, an *online service* sells access to large (usually nationwide) databases. Such services can be used to add external data to a corporate information system in a timely manner and at a reasonable cost. Several thousand services are currently available. Representative services are Dow Jones Information Service, Mead Data Central, and Compustat.

Several magazines deal exclusively or mainly with the Internet, such as *Internet Business Advantage* (*zdjournals.com*), *Internet World* (*iw.com*), *NetGuide* (*netguide .com.au*), *IEEE Internet*, and *Wired*. There are also many online magazines such as *TechWeb* (*techweb.com*). Finally, there are several associations and societies through which members participate in activities related to the Internet, such as the *Internet Society* (*isoc.org*) and the *Electronic Frontier Foundation* (*eff.org*).

ADDRESSES ON THE INTERNET.   Each computer on the Internet has an assigned address, called the **IP (Internet Protocol) address,** that uniquely identifies and distinguishes it from all other computers. The IP numbers have four parts, separated by dots. For example, the IP address of one computer might be 135.62.128.91.

DOMAIN NAMES. Most computers also have names, which are easier for people to remember than IP addresses. These names are derived from a naming system called the **domain name system (DNS).** Network Solutions Inc. (NSI) (now part of Verisign, Inc.) once had the exclusive authority to register addresses using .com,.net, or .org domain names. The company's contract ended in October 1998, as the U.S. government moved to turn management of the Web's address system over to the broader private sector. Currently, in November 2004, 376 companies, called registrars, are accredited to register domain names from the Internet Corporation for Assigned Names and Numbers (ICANN) (*internic.net*).

All domain names are administered by the Internet Corporation for Assigned Names and Numbers (ICANN). There are two types of top-level domains (TLDs): generic (gTLDs) and country code (ccTLDs). *Generic domains* were created for use by anyone on the Internet; *country code domains* were created for use by individual countries (*iana.org/cctld/cctld-whois.htm*). There is a special top-level domain (*.arpa*) for Internet infrastructure.

Domain names consist of multiple parts, separated by dots, and are translated from right to left. For example, consider the name *software.ibm.com*. The rightmost part of an Internet name, in this case "com," is its top-level specification, or the **zone.** The abbreviation "com" indicates that this is a commercial site. The 15 top-level specifications are listed below.

ZONE	USED FOR	ZONE	USED FOR
com	commercial sites	aero	air-transport industry
edu	educational sites	biz	businesses
mil	military sites	coop	cooperatives
gov	government sites	museum	museums
net	networking organizations	name	registration by individuals
org	organizations	pro	accountants, lawyers, physicians (professionals)
store	businesses offering goods for purchase	int	registering organizations
info	information service providers		

Finishing our example, "ibm" is the name of the company (IBM), and "software" is the name of the particular machine (computer) within the company to which the message is being sent. The rightmost two letters in a domain name, if present, represent the country of the Web site. For example, "us" stands for the United States, "de" for Germany, "it" for Italy, and "ru" for Russia. In the United States, the "us" is omitted. Many companies in other countries use ".com" without the country name also.

As of July 17, 2004, there were 39,363,493 domain names. Distribution of top-level domains was as follows:

TLD type	.com	.net	.org	.biz	.info	.edu
Total no.	29,321,771	4,791,592	3,029,358	1,001,991	1,211,384	7397

Source: zooknic.com/Domains/counts.html.

Most domain names have their own commercial value. This commercial value has led to the practice of *cybersquatting*—buying a potentially coveted domain name and hoping someone wants that name enough to pay for it. The practice of cybersquatting grew out of NSI's early policy of registering domain names on a first-come, first-served basis. This policy resulted in many companies or individuals registering a domain name associated with an established firm before the established firm did. The policy resulted in disputed names and legal actions. In response, Congress passed the Anti-Cybersquatting Consumer Protection Act in November 2000.

DNSs. A set of servers called **domain name servers (DNSs)** maps the human-readable names to the IP addresses. These servers are simple databases that map names to IP addresses, and they are distributed all over the Internet. Most individual companies, ISPs, and universities maintain small domain name servers to map host names to IP addresses. Inside the DNS system, there is a big database that has billions of IP addresses currently in use and the DNS requests made every day.

URLs. A **uniform resource locator (URL)** indicates the location (or address) of a Web site you want to visit. The address consists of several parts. For example, Yahoo's address—*http://www.yahoo.com*—tells us that to access the site, we use a **hypertext transfer protocol (http),** which is a procedure for retrieving hypermedia and hypertext documents. The next part of the URL, "www," indicates that you are making a request of a Web server. "Yahoo" is the name of the company, and ".com" a first-level domain name that refers to a commercial organization. As explained earlier, the URLs are classified into generic domains (e.g., .com, .edu, .net, and .gov). Addresses can be even more detailed, including the specific name of a document at a site.

To help minimize errors, most current browers allow you to type URLs without the "http://" and "www" portions of the URL. For example, you can type the text "yahoo.com" instead of "http://www.yahoo.com."

Wide Area Information Servers (WAIS)

Wide area information servers (WAIS) is an Internet directory designed to help end users find and retrieve information over the networks by providing efficient search methods. It is a distributed information retrieval system supported by Apple Computer, Thinking Machine, and Dow Jones. Users can ask for general topics, searching for key words and phrases. The search returns a list of documents, ranked according to the frequency of occurrence of the keyword(s) used in the search. The client can retrieve text or multimedia documents stored on the server. WAIS offers simple natural language input, indexed searching for fast retrieval, and a "relevance feedback" mechanism that allows the results of initial searches to influence future searches.

Mobile Internet

The **mobile Internet** refers to the use of wireless communications technologies to access network-based information and applications from mobile devices. The mobile Internet is also called the **wireless Web.**

Mobile Internet applications have several interesting applications. Users can access the mobile Internet anywhere, and the mobile Internet device can know where in particular it is located and use that knowledge to perform services that take advantage of spatial or geographic information. These applications are

called *location-based services*. The mobile Internet also provides services that are based on the type of location the mobile Internet device is in and that allow the user to act in ways that make sense only in that location. These applications are called *presence-based services*. Both types of services are examples of *location-based commerce* (*l-commerce*).

Internet Challenges

Challenges facing the Internet in the next few years include making the Internet more suitable for e-commerce transactions, incorporating rapidly evolving technologies, standards, and regulatory frameworks, responding to the growing need for additional bandwidth, and addressing privacy concerns. Another challenge for the Internet is providing the infrastructure for very large scale projects. We consider some of these challenges in the following paragraphs.

NEW TECHNOLOGIES. Vendors are adopting new technologies more rapidly than many users and customers can implement them. For example, the two most popular Web browsers are Netscape's Navigator and Microsoft's Internet Explorer. Many of the most innovative sites on the Web use Java applets, interactive three-dimensional graphics, and video and audio clips. To access these Web sites and take advantage of their innovative content and full functionality, users must have recent versions of Navigator and Explorer.

INTERNET REGULATION. Technical organizations, such as the Internet Engineering Task Force, the World Wide Web Consortium, and others, have played an important role in the evolution of the Internet and the Web. These organizations are not formally charged in any legal or operational sense with responsibility for the Internet. However, they perform the important task of defining the standards that govern the Internet's functionality. Hardware and software vendors also have been instrumental in submitting specifications for consideration to standard bodies and in creating de facto standards of their own.

Recent attempts by governments in the United States and elsewhere to regulate the content of Internet-connected computers have generated concerns about privacy, security, and the legal liability of service providers. Some content providers have addressed these issues with filters, ratings, and restricted access. However, it is difficult to regulate content across international borders. Regulation of services such as gambling also has been debated.

INTERNET EXPANSION. The Internet was not designed to provide a mass-market interchange of high-density information. As a result, the massive growth of Internet traffic has strained some elements of the network. The strains manifest themselves as slowdowns in retrieval time, unreliable transmission of streamed data, and denial of service by overloaded servers. The Internet's design, with many potential transmission paths, is in theory highly resistant to outages caused by failed links. In practice, however, the Internet often is affected by software problems.

A wide range of factors contributes to congestion or slowdowns. These problems include improperly configured networks, overloaded servers, rapidly changing Internet usage patterns, and too much traffic for available bandwidth. Approaches to solve these problems include installing high-speed transmission media to accommodate large amounts of data; bigger, faster routers and more sophisticated load balancing and management software to handle peak traffic

periods; local caching (storing) of frequently requested Web pages to improve response times; and more reliable tiers of service for those willing to pay for them.

INTERNET PRIVACY. Web sites collect information with and without consumers' knowledge. One way to collect information at Web sites is through registration (as on Amazon.com). Visitors to the site enter data about themselves, and obviously know that such information is available for future use by the company that collects the data.

The most common way Web sites collect information, though, is through "clickstream" data—that is, information about where people go within a Web site and the content they see. Clickstream data are most commonly collected by cookies. A **cookie** is a small data file placed on users' hard drives when they first visit a site. This software can be used to exchange information automatically between a server and a browser without a user seeing what is being transmitted.

Cookies are useful in tracking users' actions and preferences. When a user goes back to a site, the site's computer server can read the usage data from the cookies. This background information can then be used to customize the Web content that is given to the user. That information is stored in a database and can be used to target ads or content, based on the preferences tracked. Netscape and Internet Explorer browsers support cookie technology.

The Federal Trade Commission randomly checks Web sites to see if site operators are posting privacy notices that explain how personal information—such as e-mail addresses, shopping habits, and consumer financial data—is being used and whether that information is being protected from unwarranted intrusion. Privacy on the Internet at this point is not a sure thing.

There are bills in the U.S. Congress related to Web privacy, ranging from laws to regulate spamming (unsolicited e-mail) to legislation restricting disclosure of subscriber information by online services. Three possibilities exist:

- The government should let groups develop voluntary privacy standards but not take any action now unless real problems arise.
- The government should recommend privacy standards for the Internet but not pass laws at this time.
- The government should pass laws now for how personal information can be collected and used on the Internet.

Privacy is such an emerging important issue in the world of the Internet that it should be continuously considered in designing and using information systems.

INTERNET INVESTMENT. One final issue to consider is investment in Internet infrastructure. In the past, investments to improve the shortcomings of the Internet infrastructure were insufficient. As a result, equipment costs have outstripped revenue for the Web hosting, ASP, and outsourced computing companies as they have tried to catch up with the infrastructure needs. As entrepreneurs and financial backers develop products that will make data centers cheaper to equip and easier to manage, data centers will redirect innovation toward new technologies for truly distributed computing. According to Tier I Research, Web hosting is forecast to become a $28.5 billion business by 2005—a sevenfold increase. Companies like HP and IBM, in tandem with large telecoms, will offer computing-on-tap, so corporations will pay for computing

power much the same way they purchase electrical power—that is, they will pay for what they use. In other words, future deployment of hardware and software could become a service-centric computing model.

T5.3   BROWSING AND THE WORLD WIDE WEB

When you visit a library or a bookstore, you may browse through the books and magazines, reading a page in one and looking at a picture in another. Similarly, you can browse through the vast resources of the Internet. Initially, the amount of information on the Net was relatively small, so by going through catalogs and indices, it was possible to browse and find things quickly. But when the amount of information became very large, it was necessary to make the browsing more efficient. In 1991, a new way to organize information was introduced—the World Wide Web (WWW, or "the Web"). The **World Wide Web** (or "the Web")

The **Web** is a vast collection of interconnected pages of information that are stored on computers around the world that are connected to the Internet. The Web is a system with universally accepted standards for storing, retrieving, formatting, and displaying information (text, pictures, video, etc.) via a client/server architecture. The browser that runs on your PC is the client. When you search for information and services, Internet (or Web) servers process your request. The Web handles all types of digital information, including text, hypermedia, graphics, and sound. It uses graphical user interfaces, so it is very easy to use. The technology underlying the World Wide Web was created by Timothy Berners-Lee, who in 1989 proposed a global network of hypertext documents that would allow physics researchers to work together. This work was done at the European Laboratory for Particle Physics (known by its French acronym CERN) in Geneva, Switzerland.

Tools for Browsing the Web

Documents accessible on the Web contain **hyperlinks (links)** to other documents. Such links are used to connect documents, either internal or external. The links are basically an implementation of *hypertext*. When you click on a link, the browser will automatically display the associated document, called a **Web page,** on the topic you are exploring, regardless of where it is located on the Web. The **home page** is the starting point for your search. From a single Web page located on a computer attached to the Internet you can browse the Web by clicking on any interesting links you see. In most cases, the home page will lead users to other pages. All the pages of a particular company or individual are known as a **Web site.** Each Web site is a computer network, such as the one in your university, that has a connection to the Internet. Most Web pages provide a way to contact the organization or the individual. The person in charge of an organization's Web site is its **Webmaster.**

The most widely used browers are Netscape Navigator and Microsoft Explorer. These use GUI interfaces and require multimedia hardware and software. Similar capabilities, but without graphics, are provided by Lynx, a text-based browser developed by the University of Kansas.

NETSCAPE'S BROWSER SUITE.   Netscape's browser suite includes Netscape Navigator, Netscape Mail, Netscape Instant Messenger, Netscape Composer, and Netscape Address Book. The suite provides functions for running Web applets, audio playback, streaming media, Web content, and Net2Phone for

free PC-to-phone calls anywhere within the United States. Netscape Communicator is a comprehensive set of components that integrates e-mail, Web-based word processing, and chat to allow users to easily communicate, share, and access information.

MICROSOFT INTERNET EXPLORER.   Faced with the tremendous lead in the browser marketplace that Netscape established, Microsoft embarked on a strategy to gain market share and penetrate the installed base. It gave Internet Explorer away for free and bundled it with the Windows operating systems. This approach was successful commercially but also resulted in scrutiny by the U.S. government.

In October 1997 the U.S. Justice Department filed a petition in federal court to prevent Microsoft from requiring personal computer manufacturers to bundle Microsoft's Internet browser software with Microsoft's Windows operating systems. At issue was whether or not Microsoft tried to monopolize the Internet browser software business by refusing to let PC makers license the Windows operating systems unless they also shipped their PCs with Internet Explorer. Microsoft maintained that Internet Explorer was an enhancement of Windows, not a separate product, and that the company therefore was not violating its antitrust settlement. According to Microsoft, Internet Explorer's tight integration with Windows offers users the advantage of "one-stop computing."

In early November 2001, the Justice Department reached a settlement with Microsoft in the antitrust case. Under the settlement, a panel of three independent monitors will work onsite at Microsoft to oversee its conduct and review its accounts. Crucially, Microsoft must provide rival software firms with information to allow them to develop competing "middleware" products—in other words, software programs that interact with the computer operating system. Microsoft was prevented from retaliating against computer manufacturers and software rivals who brought out competing products, and had to deal with licensing partners on uniform terms.

METABROWSERS.   Web surfing has typically been a one-page-at-a-time experience. Then Yahoo offered My Yahoo, which enabled personalization of news, sports, financial data, entertainment, and other topics. **Metabrowsers** automate the tasks of creating metadata, letting users concentrate on the task of cataloguing Web resources (*metabrowser.spirit.net.au/manual/Manual.1.10.doc*). Octopus is a leading example. Other metabrowsing services include CallTheShots (c*alltheshots.com*), Quickbrowse (*quickbrowse.com*), Katiesoft (*katiesoft.com*), and Yodlee (*yodlee.com*).

Here is how the Octopus (*octopus.com*) provides portal personalization by giving consumers control over what they view: Users assemble customized pages, called "views" on Octopus, choosing to start with a blank view or to use one of the sample views arranged in a menu. To add elements, users grab items from the menu and drag them onto their pages. The views are dynamically linked to the Web sites providing the content. Users who click on any page element—for example, a news headline or a stock chart—are connected directly to that site. Users can store their views privately on Octopus, or they can "publish" them so other users can take a look. For users, the service is free. Octopus plans to make money by charging Web publishers a fee each time users click through to their sites.

NEW BROWSERS. There also are a couple of newcomers to the scene, Opera and NeoPlanet. Opera is renowned for being the world's fastest browser. It is much smaller than other major browsers. In its standard configuration, it almost fits on a floppy disk, yet it is all you need to surf the Web. NeoPlanet offers a "skinnable" browser. NeoPlanet *skins* are collections of graphics and sounds. It also utilizes personalization tools, allowing you to control your view of the online world. Designed to be small, full-featured, and easy to use, this free browser can look and function the way that best complements your needs.

OFFLINE BROWSERS. **Offline browsers** (or pull products) enable a user to retrieve pages automatically from Web sites at predetermined times, often during the night. WebWhacker (*bluesquirrel.com*) and WebCopier (*maximumsoft.com*) are offline browsers that allow users to define a group of sites by their URLs and then download text and images from those sites to their local storage. Web-Whacker and WebCopier let users determine how much of a Web site to retrieve—title pages only, any linked pages, or all pages.

WEB SERVERS. A **Web server,** also called an *HTTP server,* is a computer (usually dedicated) storing all Web pages and/or Web content. Users can access the Web server using a Web browser with a unique URL or IP address (see Figure T5.1). There are many types of Web servers. The best known are Microsoft's Internet Information Systems (IIS) and Linux's Apache.

The Apache HTTP Server Project is an effort to develop and maintain an open-source HTTP server for modern operating systems, including UNIX and Windows NT. The goal of this project is to provide a secure, efficient, and extensible server that provides HTTP services in sync with the current HTTP standards. In October 2004, Netcraft Web Server Survey found that more than 67 percent of the Web sites on the Internet were using Apache.

Others Web servers, used in particular niche markets, include the following:

- *Jigsaw* is W3C's leading-edge Web server platform, providing a sample HTTP 1.1 implementation and a variety of other features on top of an advanced architecture implemented in Java.
- *WN* is a Web server that runs on a wide variety of UNIX platforms and is freely available at no cost for any use under the terms of the GNU General Public License.
- *Small HTTP Server,* version 3.05, created by Max Feoktistov, claims to be the smallest Web server while having numerous functionalities. For details, see *home.lanck.net/mf/srv/index.htm.*
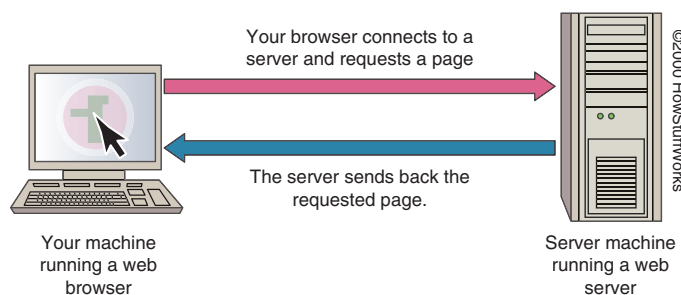

FIGURE T5.1 Relationship of a Web browser to a Web server.
Source: computer .howstuffworks.com/web-server.htm/printable

Your browser connects to a server and requests a page

The server sends back the requested page.

©2000 HowStuffWorks

Your machine running a web browser

Server machine running a web server

Creating Web Documents

To write a Web document for the Internet or an intranet, various software languages can be used. The most common is **HyperText Markup Language (HTML),** which formats documents and incorporates dynamic hypertext links to other documents stored on the same or different computers. HTML was derived from the more complex **Standard Generalized Markup Language (SGML),** a text-based language for describing the content and structure of digital documents. HTML is a simpler subset of SGML and incorporates tables, applets, text flow around images, superscripts, and subscripts. Using these hypertext links (which are typically blue, bold, and underlined), the user points at a highlighted word, clicks on it, and is transported to another document. Users are able to navigate around the Web freely with no restrictions, following their own logic, needs, or interests. Others include Dynamic HTML (DHTML), eXtensible Markup Language (XML), and Virtual Reality Modeling Language (VRML) (see Technology Guide 2).

The introduction of new tools is making the creation of Web documents easier. Applications such as FrontPage from Microsoft and DreamWeaver from Macromedia make Web page development as easy as creating a document with word-processing software.

Another useful tool for creating Web documents is **Java,** developed by Sun Microsystems (as described in Technology Guide 2). Applications programmed in Java essentially become plug-and-play objects that operate on any operating system. Java applets are small applications designed to expand the capabilities of the Web browser. They can be programmed rapidly and are reusable and portable. Using the applets, it is easy to develop special applications in finance, manufacturing, marketing, and education. Some of these include features such as animation. Java applications are easily distributed on the networks.

Another technology that has evolved from Java is called **Jini.** Jini is a network protocol that lets a device connected to a network announce to other devices on the network that it has joined up and make known what services it provides. For example, if a new printer is added to a network, via the Jini protocol, it can tell the PCs on the network that it is there as well as what services it provides (e.g., color printing). For details, see *wwws.sua.com/software/jini/specs.*

Search Engines

Browsing on the Web is a fascinating activity; you can click on the links and discover a wealth of information new to you. But a random search may be inefficient when you are looking for specific information. The solution then is to use a search engine.

A **search engine** is a tool that makes your browser more effective. It enables you to locate information by using *key words* in the same way that you would search online library resources (e.g., ABI Inform). The search engine matches the key words with abstracts and presents the results to you. The abstracts contain links to the source documents so that you can view them in full.

Some of the major Internet-based search engines include Google, Yahoo, Lycos, HotBot, WebCrawler, Alta Vista, and Excite. Each engine uses a slightly different approach, syntax, and presentation mode, though they all have similar capabilities. Search engines can be reached directly by entering their address in the location box of most browsers. Yahoo's address, for example, is *yahoo.com*. When you enter the Yahoo site, you are presented with many resources that can be reached by the hyperlinks found there. For example, by requesting "Option Trading," you can receive information about buying and selling financial options. With sites such as Yahoo and Excite constantly beefing

themselves up into the online equivalent of mega-malls, online resources are astounding, and are expanding very frequently.

Some new search engines do nothing but search. Google is one such search engine. It uses sophisticated next-generation technology to quickly produce the most relevant search results with every query.

Search engines for large numbers of Web pages, such as those that attempt to cover the entire Internet, do so by maintaining databases that model the Web's structures. Through a combination of information-trolling robots that collect information automatically about Web pages and developer registration, search engines select a large number of Web sites to be indexed. Their databases are then populated with information about the contents of each page deemed useful. Many Web sites have search engines embedded within them. Also, some engines search not only Web pages, but also gopher sites, FTP resources, or USENET news groups. Search engines select pages for inclusion in their databases in two primary ways:

- **Web crawlers** traverse the Web automatically, collecting index data on one of two search principles: *depth first,* which follows only the links that are deemed relevant to a topic, or *breadth first*, which collects the entire network of links from a given starting point regardless of the page contents. Web crawlers are variously called *spiders, ants, robots, bots*, and *agents*.
- **Registration** is allowed by most search sites. Web developers can register their sites or pages by submitting a form. This process enables developers to ensure that their sites eventually will be included in the search index.

When a user enters a search query, the engine searches its database for relevant Web pages. It assembles a list of pages sorted by relevance or other, user-specified weighting factors. Some sites also remove redundant references to pages from the list. Search results are returned as a list of relevant pages that then can be retrieved via hyperlinks. Different engines can produce results that vary widely, ranging from not finding critical pages (*poor recall*) to finding hundreds of thousands of documents with few that are relevant (*poor precision ratio*). In Fall 2004, Google (*google.com*) launched a search engine for the academic world. Its URL address is *scholar.google.com.* It contains the search results of academic research papers and journal documents.

Google is the largest search engine and the first search engine to index more than 1 billion pages. Google's method of searching the Web is called PageRank. The more links there are to a page, the higher it moves in Google's ranking. PageRank improves both recall and precision ratio.

To get even better results, **metasearch engines** (such as Spider, Savvy Search, Search, ixquick, Dogpile, ithaki, metaEUREKA, ProFusion, Mamma, Pandia, MetaXChem, Turbo10, Kartoo, Widow, *webtaxi.com,* Mulimeta, Metacrawler, All-in-One, and Web Compass) integrate the findings of the various search engines to answer queries posted by the users.

Gopher and Veronica

Not every file or document on the Internet is accessible on the Web. A predecessor to the Web, **Gopher,** is a "burrowing tool" that provides access to a wide range of *textual information* available on the Internet. It is a delivery system for retrieving documents from the Internet, based on a set of menus, submenus, and sub-submenus (down to many levels). Although the Web is rapidly eclipsing it, there are still thousands of Gopher servers worldwide providing access to

hundreds of thousands of documents. Gopher can be reached directly via your browser by typing *www.gopher.com*.

Veronica (Very Easy Rodent-Oriented Netwide Index to Computer Archives) provides the capability of searching for text that appears in gopher menus. When the user enters a key word, Veronica will search through thousands of gopher sites to find titles containing that keyword. It places these files on a temporary menu on the local server, so users can browse through them. For details, see *www.ou.edu/research/electron/internet.*

Downloading Software and Files

Tens of thousands of software packages can be downloaded. Many of them are free either as public domain (*cheapware*), or copyrighted but given away (*veronica.htm*). Others are called *shareware*, meaning you can use them on a trial basis and are expected to purchase them if you like them (see *shareware.com*).

Internet users can also retrieve documents including graphics, voice, and other media for use on their PCs. Working with the Web or Gopher, it is usually fairly easy to do so. Many of the documents are transferable from the Internet using a member of the TCP/IP family called **file transfer protocol (FTP).** Using special FTP software, you can easily download files stored in an *FTP site*, a server that archives such files. Like commercially sold software, downloadable software (cheapware, freeware, and shareware) goes through a series of improvements or updates. Version 5.0 is released and shortly afterward updated to 5.1, then 5.2, and so forth. Unlike most commercial software, the updates for cheapware, freeware, and shareware are free to download.

Archie is a tool that allows users to search the files at FTP sites. It regularly monitors hundreds of FTP sites and updates a database (called an *Archie server*) on software, documents, and data files available for downloading. If users click on a listing from an Archie server, it will take them to another computer system where relevant files are stored. Once there, the Archie server may allow users to continue their searches for files until they locate what they need. Archie database searches use subject key words, such as "Bowl Championship Series" or "network computers." The Archie database will return lists of sites that contain files on these topics. For details, see *archie.is.co.za/ftpsearch.*

T5.4   COMMUNICATION TOOLS FOR THE INTERNET

Several communication tools are available on the Internet. The most notable are discussed below.

Electronic Mail

Electronic mail (e-mail) allows multiple-access communication delivered exclusively on a computer network. With e-mail, a person can send letters to anyone connected to the system. When a message is sent, it arrives at an individual's "mailbox." The receiver, when connected to the network, is notified that he or she has mail. The receiver can then read the mail, send a reply, edit the mail, save it, or forward it to another person.

E-mail has become an important communication tool in many organizations. Several vendors produce e-mail software. Some well-known software packages are Microsoft Outlook, LotusNotes, QuickMail, Netscape Messenger, Eudora, Groupwise, Pegasus Mail, and Xtramail. Windows-based e-mail and mail-enabled applications that are enterprisewide are increasing the attractiveness, friendliness, and capabilities of e-mail. E-mail is also becoming connected to voice mail, and

sometimes the two are integrated. Netscape and Microsoft browsers include e-mail capabilities and so do online services such as AOL.

Many Web-based e-mail providers (such as Hotmail, Coolmail, and Yahoo Mail) offer e-mail accounts for free. People can send and receive messages from any computer connected to the Internet.

Online Chatting

Chat programs allow you to send messages to people who are connected to the same channel of communication *at the same time*. It is like a global conference call system, most often in written form. Anyone can join in the online conversation. Messages are displayed on your screen as they arrive, even if you are in the middle of typing a message. You can even use voice chat, which, like voice e-mail, is free of charge, even if it is long distance.

Two major types of chat programs exist. The first is a Web-based program, called Webchat, which allows you to send messages to Net users using a Web browser and visiting a Webchat site. (The Budweiser Web site, *budweiser.com*, for example, has a popular chat program.) The second is an e-mail-based (text only) program called **Internet Relay Chat (IRC).** A business can use IRC to interact with customers, provide online experts' answers to questions, and so on.

INSTANT MESSAGING. **Instant messaging (IM)** refers to a type of communications service that enables you to create a private chat room with another individual in order to communicate in real time over the Internet. Such communication is analogous to a telephone conversation but uses text-based, not voice-based, communication. Typically, the instant messaging system alerts you whenever somebody on your private list is online. You can then initiate a chat session with that individual.

ICQ ("I Seek You") is a popular instant messaging chat program that informs users who is on-line at any time and enables users to contact them at will. With ICQ, people can chat, send messages, files, and URLs, and play games. Several similar programs exist (e.g., Instant Messenger). Microsoft's NetMeeting software also has an online chat function in addition to its videoconferencing feature.

PEER-TO-PEER. **Peer-to-peer (P2P)** is an electronic file-swapping system that allows users to share files, computing capabilities, network bandwidth, and storage. It allows groups of programs that have similar attributes to communicate with each other. P2P usually results in higher costs for server infrastructure, deployment, and administration. It is mainly for commercial use, whereas instant messaging is for chat and group meetings.

Similar to WinMX and KaZaa, *BitTorrent* is a kind of P2P downloading software. Basically, it has the unique feature of downloading software from a server (also called the "seed"), while the same user can upload files to other users. Thus, the software-downloading users form virtually a file-transfer collaboration web among them. This web can achieve the purpose of substantially mitigating the traffic congestion of the server containing the downloading software.

Newsgroups

Using a chat program requires you to communicate in real time. It may also be a waste of time, if the chatters are not focused on a topic of interest to you. The solution is to join (subscribe to) a **newsgroup** (also known as *Usenet*, or just *News*). Newsgroups are organized in a directory and are divided into categories and subcategories. There are many thousands of them, each of which is dedicated to a specific topic. Once you join such a group, you can read all the

messages sent by other members. You can respond by posting a message for all members, or e-mail to just one individual. In contrast to a chat room, where all communication is done in real time, here communication is done via messages in a different time/different place mode.

Depending on your online services or Internet service provider, you may have access to an average of more than 100,000 newsgroups. However, not all newsgroups may be available at your site; they are so numerous that they create a great deal of Internet traffic, so your ISP management may block some. You may try access services such as America Online to reach more newsgroups. AOL and similar services have their own newsgroups. Newsgroups are popular for discussion and for posting business opportunities such as creating partnerships or selling items.

Many newsgroups have FAQs specific to the subject of the newsgroup. **FAQ** is an acronym for *Frequently Asked Questions*. FAQs are compilations of information that are usually the result of certain questions constantly being asked in a newsgroup. Before asking a question in a newsgroup, check out the appropriate FAQs. If you cannot find the answer to your question there, then you can post your question to the newsgroup.

Participation in chat or newsgroup forums can be anonymous (you can use a nickname). Make sure your messages will not cause embarrassment or offend anyone. This may be difficult to do since you are not acquainted with all possible readers.

Mailing Lists and Electronic Bulletin Boards

Similar to newsgroups are e-mail-based *discussion* or *mailing lists*. To add your e-mail name and address to a mailing list, you **subscribe** to it; to remove your name, you **unsubscribe** from the mailing list. Some mailing lists are called LISTSERVs, named after a popular mailing list of software products. It is important to note the difference between the name of the list you are subscribed to (and to which you will send your messages), and the name of the listserver that maintains the list. The name of the list is usually subject@somesite.edu (for example, *isworld@listserve.heanet.ie*).

ELECTRONIC BULLETIN BOARDS. **Electronic bulletin boards (EBBs)** are mailing lists on which users can leave messages for other people and receive massive amounts of information (including free software). They are similar to the Web newsgroups, but initially they resided on the noncommercial portion of the Internet and on private information providers' networks. To access an EBB, you need to be a member and sometimes pay a nominal fee. Special-interest groups, like users of specific software, display messages on these boards and exchange experiences about the software. Universities and schools use the EBB to advertise classes and special programs.

Portals

As discussed at some length in Chapter 4 in the textbook, a **portal** is a Web site designed to offer a variety of Internet services from a single convenient location. Most portals offer the following services: search engine; news, sports, and weather; reference such as yellow pages and maps; shopping malls; e-mail; and chat rooms. Popular portals include Altavista, Excite, Lycos, Microsoft Internet Start, Netscape Netcenter, Snap, and Yahoo.

Internet Telephony

Internet vendors are providing products that emulate traditional public switched telephone network (PSTN) applications. **Internet telephony** (also called **Voice**

over IP or **VoIP**) lets users talk across the Internet to any personal computer equipped to receive the call for the price of only the Internet connection.

VoIP carries voice calls over the Internet, either partially or completely bypassing the public switched telephone network. Sound quality can be poor due to latency and jitter. *Latency* is delay during the transmission process. *Jitter* occurs when large amounts of data clog networks at certain times. When there is too much latency and/or jitter, callers tend to miss about one out of every four or five words.

Internet Fax The use of the Internet for real-time fax transmissions is emerging as an application that may signal a shift of traditional analog communications from the telephone companies to the packet-switched Internet. This application is useful because faxes can be sent long distances at local telephone rates, and delivery can be guaranteed through store-and-forward mechanisms.

Faxing is one of the top forms of communication in the business world. Some 800 billion pages were faxed worldwide in 2000. With 100 million fax machines worldwide generating a telephone bill of more than $100 billion per year, companies spend a large percentage of their yearly phone bills on faxing. (See *auburn.edu/~rainerk/internet_web.html*.)

Most fax machines are group resources, shared by several users. Many of these users are unhappy with some of the inconveniences of faxing: They do not like having to print out a document, pick it up at the laser printer, walk down the hall to the fax machine, and wait for the fax to go through. Faxes are not secure or encrypted, nor are they private. At the receiving end, they often lie in the fax machine or on a desktop for all to see.

An **Internet fax service** overcomes many of these drawbacks. Fax service from an Internet service provider connects desktop computers and standard fax machines to a fax server located within the ISP's network. The same service can also connect desktop e-mail to the ISP's fax servers so that faxes can be originated as easily as sending an e-mail.

T5.5 OTHER INTERNET TOOLS

Streaming Audio and Video **Streaming** allows Internet users to see and hear data as they are transmitted from the host server instead of waiting until the entire file is downloaded. For example, RealNetworks' RealAudio allows a Web site to deliver live and on-demand audio over the Internet and can work over connections as slow as 14.4 Kbps.

Streaming audio enables the broadcast of radio programs, music, press conferences, speeches, and news programs over the Internet. In the future, streaming audio and Internet telephony use will overlap and complement one another.

Streaming video has other business applications, including training, entertainment, communications, advertising, and marketing. Streaming audio and video are being used to deliver market-sensitive news and other "live" status reports to stock traders, to brief sales people on new products, to deliver corporate news to employees, and to view TV commercials for approval.

Streaming audio and video vendors include Apple (QuickTime), Microsoft (ActiveMovie), and RealNetworks (RealVideo). A San-Diego company called PacketVideo has developed software for streaming video to mobile devices wirelessly. This company, which has six patents on video compression technology, has developed small-foot-print software that can be used in portable devices to

play MPG4-standard video streams. Although wireless networks usually suffer from a lot of packet loss, PacketVideo solves this problem by intelligently throwing out packets when the network gets bogged down.

Real-Time Audio and Video

In **real-time audio** and **real-time video,** the transmission from the source is live or only slightly delayed. Real-time audio and video applications include point-to-point conversations between two people; conferences among more than two people; collaborative "white boarding" (where two or more users can interactively create graphic images) and shared hypermedia documents; live broadcasts of news, talk shows, or sporting events; and broadcasts of music and concerts. Real-time audio and video vendors include Intel (Video Phone), Microsoft (NetMeeting), and CUseeMe Networks.

Web Services

The next step in software evolution will have software applications provided in the form of services delivered over the Internet. Web Services are unique pieces of computer code (or components) accessed through a Web site that deliver a specific type of function. Web Services allow us to transparently access rich software content from any site on the Web.

The great promise of Web Services is the ability to deliver applications to users at a much lower cost. As much as 30 percent of the price of traditional software is tied to the cost of distribution—pressing CDs, packaging them in boxes, and shipping them to retail outlets. Digital distribution eliminates the physical element. Some day soon, even medical services will be accessible over the Web.

IMPLEMENTATION OF WEB SERVICES. .NET is Microsoft's platform for XML Web Services. The .NET framework allows unrelated Web sites to communicate with each other and with programs that run on personal computers. .NET means that one click could set off a cascade of actions without requiring the user to open new programs or visit additional Web sites. C# (pronounced *C-sharp*) is Microsoft's object-oriented language that enables programmers to quickly build a wide range of applications for the .NET platform.

Push Technology

As the amount of information available on the Internet grows, new mechanisms for delivering it to consumers are being developed. Since its inception, the Web has been based on a pull model of information access. The Web user must actively seek out information by specifying a page to be "pulled down" to the desktop by typing in a URL, following a hot link, or using the search results from a Web search site. However, passively placing content on a Web site and waiting for people to come browse is not well-suited to establishing and fostering strong relationships with customers or prospects.

Therefore, an alternative push model of information delivery has emerged. In this model, information is "pushed" to the user's desktop. **Push technology** now makes it possible to automatically supply information to users by means of a process running on either the user's desktop or a network server. With millions of Web sites available for browsing, the only way to guarantee that users receive certain content is to send or "push" it to them. Push client packages typically are given away free, and the companies that publish them rely on advertising for revenue.

One of the earliest products embodying the push model was the PointCast Network from PointCast. PointCast developed software that used the Web browser as a platform and displayed customizable news and other information

on the user's screen as a screensaver. PointCast fell from favor in many organizations because the incoming broadcasts often overloaded a recipient organization's local area network and went out of business.

Push technology is useful in the workplace, in the consumer market, and as a mechanism for software distribution. In the workplace, push technology can provide timely, prioritized distribution of information over a corporate network. For example, the software can be oriented to an organization's different departments to focus attention on important communications. In the consumer market, push technology can enhance traditional Web advertising. Users no longer need to find advertisements; instead, the users' attention can be directed to the advertisements. In addition, the quality of the presentations can be improved by tuning the software specifically to the user's platform and connection speed.

Information Filters

Information filters are automated methods of sorting relevant from irrelevant information. These filters help people access information with more precision; that is, they help people reduce information overload. As the information available over the Internet continues to grow, users increasingly need to narrow the content through which they wish to search.

The most publicized information filters are those programs that screen out adult content from Web browsers. Intended for home markets, examples of Internet screening software include SurfControl's CyberPatrol (*cyberpatrol.com*) and Net Nanny (*netnanny.com*). These programs prevent access to a list of sites deemed unacceptable by the company providing the software.

In response to public concerns and to preempt possible federal regulation, AOL, Disney, Microsoft, Netscape, and other companies are supporting the Platform for Internet Content Selection (PICS), a specification for labeling Web content. PICS embeds labels in HTML page headers to rate different Web sites by category, much like existing movie and television ratings. Microsoft's Internet Explorer browser and Netscape's Navigator browser allow parents to block categories and set a password.

A more active method of filtering information uses **intelligent agents.** The goal of agents is to create applications that automatically carry out tasks for users without their intervention, other than initial configuration and updating with new requests.

Clipping Services

The number of publications, traditional and electronic, available online continues to increase. In digital format, publications are easily amenable to efficient or automated clipping by use of a **clipping service.** For example, Excite offers News-Tracker, a free clipping service. Through it, users can track up to 20 news topics and retrieve articles from a database of more than 300 publications. Excite generates revenue by advertising. As a marketing manager, for example, you could use a clipping service to look for information on new products or markets your company is considering. As a public relations officer of a company, for example, you could find references to your company in those publications and could also be on the lookout for references to competing companies in your industry.

Personalized Web services offer the ability to generate Web content that is personalized for individual Web site visitors. Information about the user is gathered from activity during the current or previous visits to a Web site, the type of browser, or browser preference settings. For example, the ability to let a site visitor define how the Web page is displayed (e.g., with or without frames or Java) is a type of personalization because this home page is generated dynamically,

based on the user's previous setup. This feature has become a requirement of any online news service, appearing in the Web sites of publications ranging from the *Wall Street Journal* to *Wired*.

Web Authoring

Web sites have become important creative media with the added benefits of multimedia and dynamic database-driven content. Tools for **Web authoring**—for page and site design—range from ASCII text editors to full-featured, integrated development environments.

The limiting factor that underlies all layout and design tools is what the most commonly used Web browsers can display. Standard HTML, which is constantly evolving and being extended with proprietary enhancements by browser vendors, is the common denominator. Graphics files in the CompuServe Graphics Interchange Format (GIF) are common, as are graphics in the Joint Photographic Experts Group (JPEG) format. Browsers can be extended with additional capabilities through plug-in applications and software components that are able to display other types of content, ranging from text formatted more richly than HTML allows, to animated graphics, audio tracks, and video clips. Popular Web authoring tools include Adobe's GoLive, Microsoft's FrontPage, and Netscape's Navigator Gold and Composer.

In summary, Table T5.1 shows a recent evaluation by PCWorld.com of various Internet tools discussed in this appendix. The tool cited in the "Product Name" column represents PCWorld's choice of the best tool in its category.

TABLE T5.1 **Evaluation of Internet Tools**

Category	Product Name	Beneficial Features
E-mail client	Outlook Express 6	Offers strong security measures and a smooth interface; can accommodate many users with multiple e-mail accounts.
Instant messaging	AOL Instant Messenger	Has a simple interface but still gives you powerful features like multiuser chat and voice chat.
Media player	RealNetworks RealOne	Attractive, simple-to-use interface; copies music to both its own format and MP3.
Ad blocker	WebWasher 3	Filters out ads, pop-up windows, scripts, cookies, Web bugs, other Web annoyances.
Cookie manager	The Limit Software Cookie Crusher	Provides detailed control over bits of code that Web sites and advertisers use to identify you.
Newsgroup reader	Forte Free Agent 1.21	Great for offline news reading and for downloading and decoding batches of file attachments.
File sharing system	BearShare 2.3	Anonymous; easy for former Napster users to understand.
Metasearch engine	Copernic 2001 Basic	Searches as many as 80 engines in seven categories.
Web archiver	AskSam SurfSaver 2.2	Gives you great control over what's saved; creates searchable archives; works with Netscape Navigator 4.x and IE 4 or higher.
Web authoring	Evrsoft 1st Page 2000	Includes samples for creating Web effects; allows real-time previews; grows with your skills.
File transfer	RhinoSoft.com FTP Voyager 8	Offers a sophisticated interface and a transfer scheduler; allows for one-click Web site synchronization.
Firewall	Zone Labs ZoneAlarm 2.6	Blocks hackers trying to access your system; prevents applications on your PC from gaining unauthorized Internet access.

Source: Compiled from various files in the *PCWorld.com* site (accessed June–July 2003).

REFERENCES

Associated Press, "Ordinary Folk Eexcluded," *International Herald Tribune*, June 29–30, 2002.

"Broadband Air Pipe on Trial," *Western Australia E-commerce Centre*, January 14, 2003.

Comer, D. E., and R. E. Droms, *Computer Networks and Internets*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

Comer, D. E., *The Internet Book: Everything You Need to Know about Computer Networking and How the Internet Works*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

Gauntt, J., et al., "E-Business Technology Forecast," *PricewaterhouseCoopers Global*, May 1999.

Gralla, P., *How the Internet Works: Millennium Edition*, 5th ed. Indianapolis, IN: Que, 1999.

Malik, O., "Data Central," *Red Herring*, November 2001.

Miastkowski, S., "How It Works: Wireless Application Protocol," *PCWorld*, July 3, 2000.

Minoli, D., and A. Schmidt, *Internet Architectures*. New York: Wiley, 1999.

Ploskina, B., "Tool Puts Office-Phone Function on the Road," *Interactive Week*, December 11, 2000.

Pogue, D., "Web Site Eases Remote Access to Homebound PCs," *I-Tech*, June 11, 2001.

Quelch, J. A., and L. R. Klein, "The Internet and International Marketing," *Sloan Management Review*, Spring 1996.

Scott, D. J., *Introduction to the Internet*. Upper Saddle River, NJ: Prentice-Hall 1999.

Sliwa, C., "Jini: Promising Technology, but Will It Catch On?" *Computerworld*, February 8, 1999.

Spanbauer, S., "Browsing and Beyond," *PC World*, February 2002.

Spangler, T., "Internet Video Takes A Wireless Leap."

Stapleton, R. M., "Bigger Better Faster: Here Comes Internet2," *Interactive Week*, August 31, 2000.

Stott, D., "Internet2," *Technology Watch*, Issue 1, Winter 2000.