COURSE TECHNOLOGY
CENGAGE Learning™

# E-Business
# Tenth Edition

*Chapter 3*
*Web Server and E-Mail Technologies*

# Learning Objectives

In this chapter, you will learn:

- Web server basics
- Software for Web servers
- E-mail management and spam control issues
- Internet and Web site utility programs
- Web server hardware

# Web Server Basics

- Chapter topics
  - Basic technologies to build online business Web sites
    - Server software and hardware
    - Utility function software
- Client/server architectures
  - Used in LANs, WANs, and the Web
  - Client requests server services
- Servers
  - Have more memory and larger, faster disk drives

# Web Server Basics (cont'd.)

- Web browser software
  - Uses Web browser software (Web client software)
  - Make computers work as Web clients
  - Web browser also called Web client software
  - **Platform neutrality**
    - Ability of a network to connect devices that use different operating systems
    - Critical in rapid spread and widespread Web acceptance

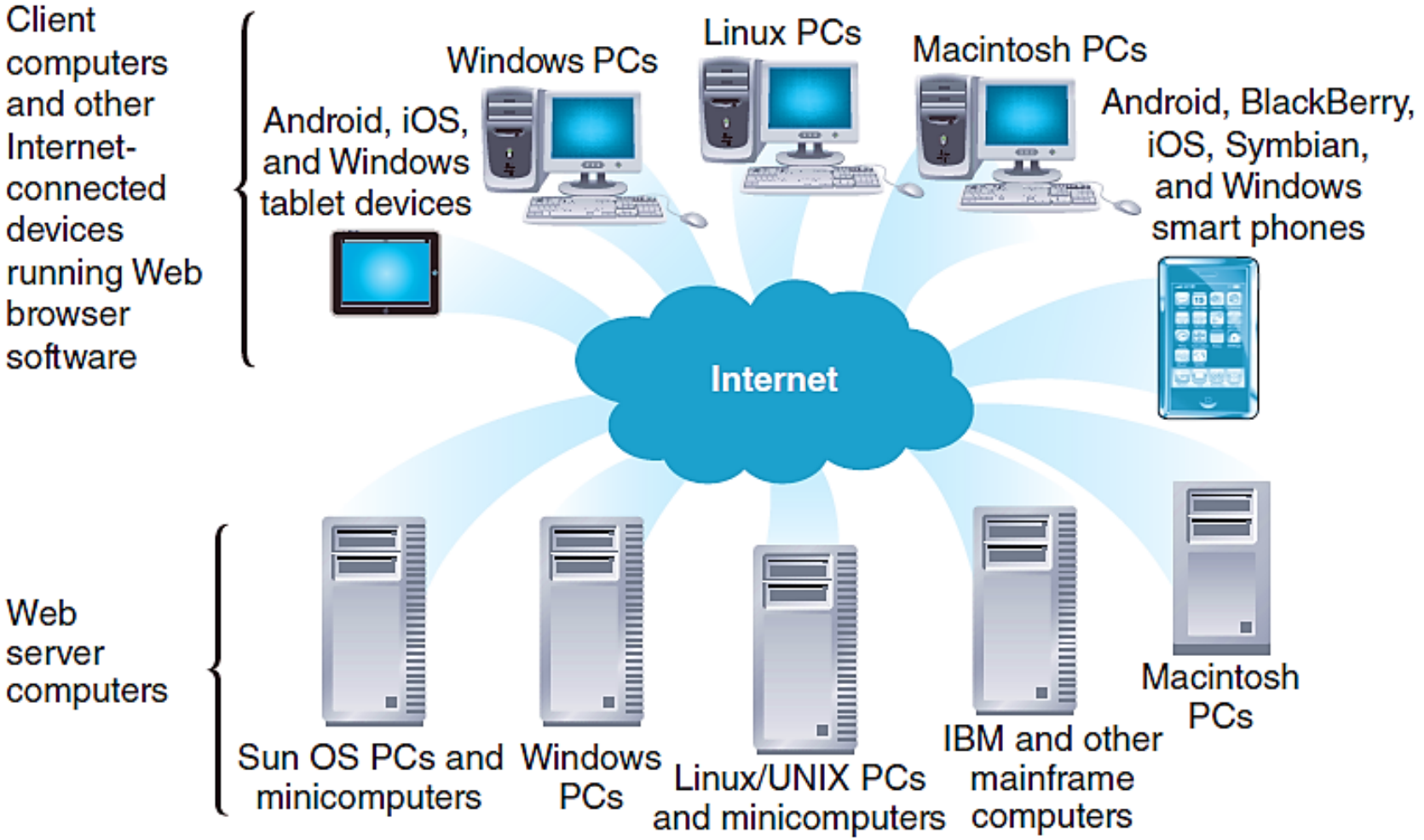





**FIGURE 3-1** Platform neutrality of the Web

# Web Server Basics (cont'd.)

- Web server
  - Main job: respond to Web client requests
  - Main elements:
    - Hardware, operating system software, and Web server software
- Web site goals followed by site development estimations
  - Number of visitors
  - Number of pages viewed during an average visit
  - How large the pages will be
  - Maximum number of simultaneous visitors

# Dynamic Content Generation

- **Dynamic page**
  - Web page content shaped by program
- **Static page**
  - Unchanging page retrieved from Web server file(s)
- Web sites using collection of HTML pages
  - Changed by editing HTML (cumbersome)
    - Specific query-customized pages not allowed
- **Dynamic content**
  - Nonstatic information constructed in response to Web client's request
  - Gives user an interactive experience

# Dynamic Content Generation (cont'd.)

- Approaches for creating dynamic content
  - **Client-side scripting**
    - Software operates on the Web client (browser)
    - Software changes Web page display in response to a user's actions
    - Software examples: JavaScript or Adobe Flash
  - **Server-side scripting**
    - Program runs on a Web server
    - Program creates Web page in response to request for specific information from a Web client

# Dynamic Content Generation (cont'd.)

- Dynamic page generation technologies
  - Server-side scripts mixed with HTML-tagged text
  - Examples:
    - Microsoft **Active Server Pages** (**ASP**): **ASP.NET**
    - Sun Microsystems **JavaServer Pages** (**JSP**): **Java servlets**
    - Open-source Apache Software Foundation **Hypertext Preprocessor (PHP)**
    - Adobe **ColdFusion**
  - Server-side languages generally use:
    - **Common Gateway Interface** (**CGI**)

# Dynamic Content Generation (cont'd.)

- Dynamic page generation tools
  - **AJAX** (asynchronous JavaScript and XML)
    - Creates interactive Web sites looking like applications
    - Example: Google Maps
  - **Ruby on Rails**
    - Creates dynamic Web pages with interface looking like application
  - **Python**
    - Scripting language

# Multiple Meanings of “Server”

- Server
  - Computer providing files, making programs available to other computers connected to it through a network
  - **Server software**
    - Makes files and programs available
    - May be part of the operating system (OS)
    - Server OS software may be referred to as server software (confusing)
  - May connect through a router to the Internet
    - Run Web server software

# Multiple Meanings of “Server” (cont’d.)

- Web server
  - Computer connected to the Internet
  - Runs Web server software
    - Makes server’s files available to other computers
- **E-mail server**: handles incoming and outgoing e-mail
- **Database server**
  - Runs database management software
- **Transaction server**
  - Runs accounting and inventory management software

# Various Meanings of “Server” (cont’d.)

- “Server” describes several types of computer hardware, software
  - Context needed to determine the intended meaning

# Web Client/Server Architectures

- Web browser requests files from Web server
  - Transportation medium: the Internet
  - Request formatted by browser using HTTP
  - Request sent to server computer
  - Server receives request
    - Retrieves file containing requested Web page
    - Formats using HTTP
    - Sends back to client over the Internet
  - Client Web browser software
    - Displays page on client machine

# Web Client/Server Architectures (cont'd.)

- Repeating process
  - Client requests, server responds; client displays result
  - Possible result:
    - Dozens or even hundreds of separate server responses
  - Graphics and other objects
    - May be slow to appear in client's Web browser window
- **Two-tier client/server architecture**
  - One client and one server computer
    - Create and read messages

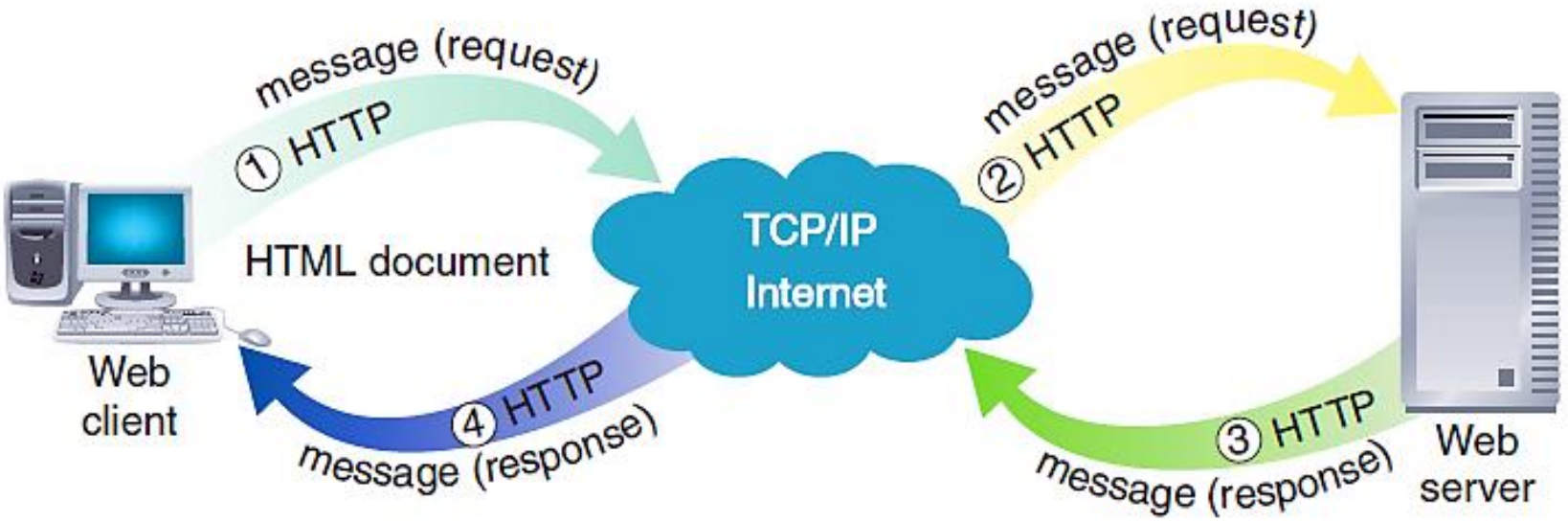



**FIGURE 3-2** Message flows in a two-tier client/server network

# Web Client/Server Architectures (cont'd.)

- **Request message**
  - Web client message sent to request file(s) from a Web server
  - Three major parts
    - **Request line**: contains command, target resource name, protocol name, version number
    - Optional **request headers**: file type information client accepts
    - Optional **entity body**: passes bulk information to server

# Web Client/Server Architectures (cont'd.)

- Server receiving request message executes command included in message
  - Retrieves Web page file from disk
  - Creates **response message**: sent back to client
    - Identical in structure to request message (slightly different function)
    - **Response header line**: server HTTP version, response status, status information explanation
    - **Response header field**: information describing server's attributes
    - **Entity body**: returns HTML page requested

# Web Client/Server Architectures (cont'd.)

- **Three-tier architecture**
  - Extends two-tier architecture
    - Allows additional processing before server responds to client's request
  - Often includes databases and related software applications
    - Supplies information to the Web server
  - Web server uses software applications' output when responding to client requests

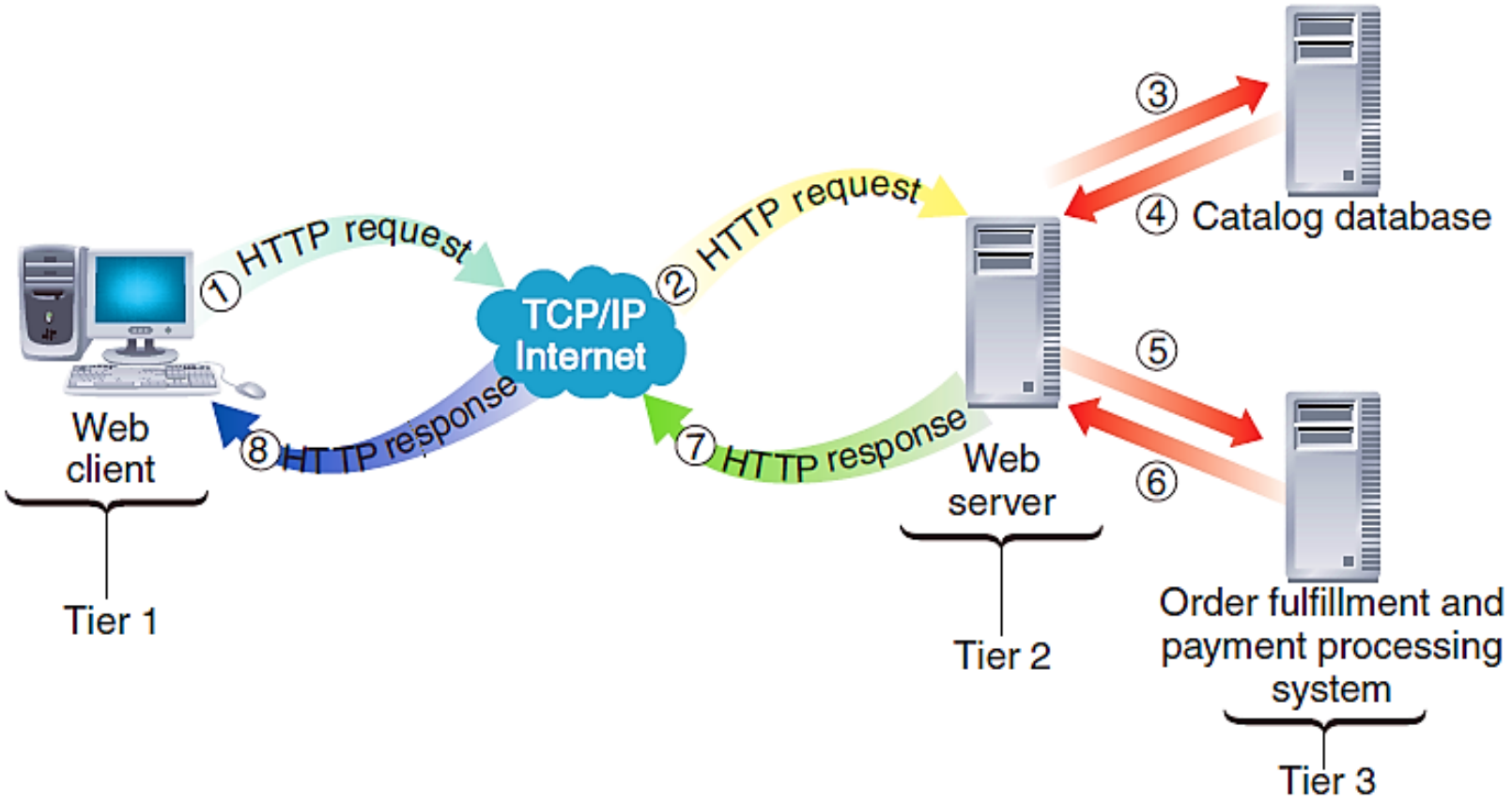





**FIGURE 3-3** Message flows in a three-tier client/server network

# Web Client/Server Architectures (cont'd.)

- **n-tier architectures**
  - More than three tiers
  - Example: catalog-style Web site search, update, display functions
    - Track customer purchases stored in shopping carts, look up sales tax rates, keep track of customer preferences, query inventory databases, keep company catalog current

# Software for Web Servers

- Web server software may:
  - Run on one or several computer operating systems
- Section topics
  - Learn about operating system software used on most Web servers
  - Learn about Web server software itself
  - Learn about other programs
    - Running on Web servers or other computers as part of electronic commerce operations

# Operating Systems for Web Servers

- Operating system tasks
  - Running programs, allocating computer resources, providing input and output services
  - Larger system responsibilities
    - Tracking multiple users, ensuring no interference
- Web server operating systems software
  - Microsoft Windows Server products
  - Linux
  - UNIX-based operating systems
    - Example: FreeBSD

# Operating Systems for Web Servers (cont'd.)

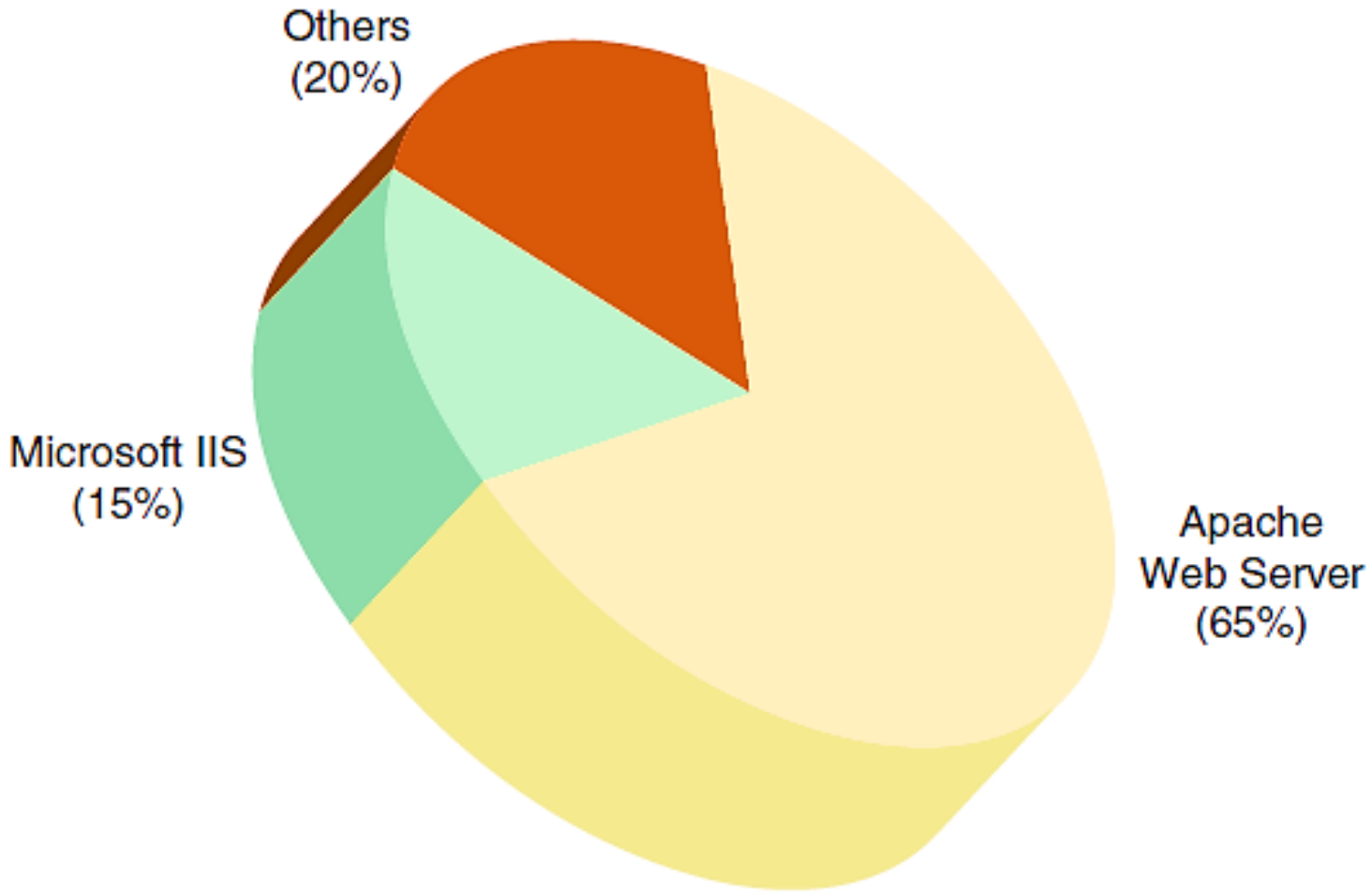
- Microsoft server products
  - Considered simple to learn and use
  - Raise security concerns
- Linux
  - Open-source
  - Fast, efficient, and easy to install
  - Can be downloaded free from the Web
  - Most companies buy it through a commercial distributor
    - Includes additional utilities, support

# Operating Systems for Web Servers (cont'd.)

- Linux (cont'd.)
  - Commercial Linux examples: Mandriva, Red Hat, SCO Group, SuSE Linux Enterprise
- UNIX-based operating system
  - Solaris

# Web Server Software

- Commonly used Web server programs
  - Apache HTTP Server, Microsoft Internet Information Server (IIS)
- Other Web server programs
  - Oracle iPlanet, nginx, lighttpd
- Netcraft January 2012 Web survey indicates:
  - Web server software market share stabilized in recent years




Source: Netcraft Web Surveys, http://www.netcraft.com

**FIGURE 3-4** Percent of active Web sites that use major Web server software products

E-Business, Tenth Edition

# Web Server Software (cont'd.)

- Apache HTTP Server
  - 1994: Rob McCool developed Apache
  - Extension had original core system with patches
    - Known as "a patchy" server ("Apache")
  - Reasons Apache dominated Web since 1996
    - Free and performs efficiently, active contributors to online forum
  - Runs on many operating systems and supporting hardware
    - FreeBSD-UNIX, HP-UX, Linux, Microsoft Windows, SCO-UNIX, and Solaris

# Web Server Software (cont'd.)

- Microsoft Internet Information Server
  - Bundled with Microsoft Windows Server OS
  - Used on many corporate intranets
  - Used by small and large sites
  - Run only on Windows server operating systems (by design)
  - Supports ASP, ActiveX Data Objects, SQL database queries
  - Produces dynamic Web pages by:
    - Including HTML pages, ActiveX components, scripts

# Finding Web Server Software Information

- Netcraft Web site
  - "What's that site running?" link
    - Leads to search function page
    - Provides operating system, Web server software specific site now running
    - Provides past site information

# Electronic Mail (E-Mail)

- Electronic commerce important technologies
  - Web
    - Provides interactions between Web servers and clients
  - E-mail:
    - Used to gather information, execute transactions, perform other electronic commerce related tasks
    - Originated in 1970s on ARPANET
    - Most popular form of business communication

# E-Mail Benefits

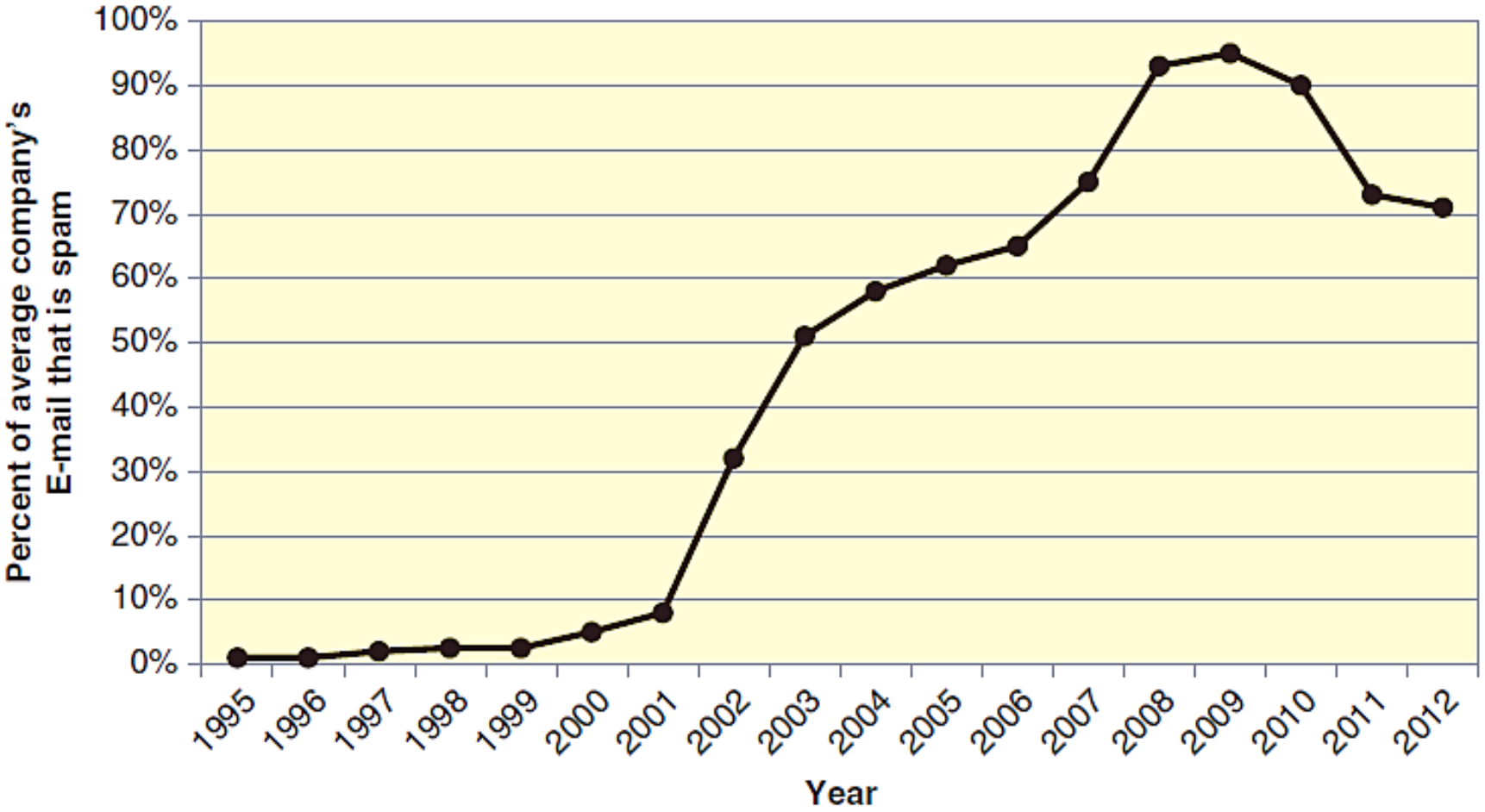
- Reason people originally attracted to the Internet
- Conveys messages in seconds
  - Contains simple ASCII text or character formatting
- Useful feature
  - **Attachments**: often most important message part
- E-mail uses
  - Confirm receipt of customer orders, confirm shipment of items ordered, send information about a purchase to buyer, announce specials and sales, keep in touch with customers

# E-Mail Drawbacks

- Time spent answering e-mail
  - Managers: five minutes per e-mail
  - Average person: two hours a day
  - Creating resentment
- **Computer virus (virus)**
  - Program attaching itself to another program
    - Causes damage when host program activated
  - Attachment can contain viruses
  - Cost for e-mail convenience
    - Virus protection software; dealing with security threats

# Spam

- Magnitude of spam problem
  - In 2009: 24-hour period showed 220 billion spam e-mail messages sent
  - Researchers believe spam growth has leveled off
    - Appears to be declining slightly




Source: Symantec *Intelligence Reports*, *Spam and Phishing Reports*, and *Spam Reports*; www.symantec.com

**FIGURE 3-5** Spam as a proportion of all business e-mail

# Spam (cont'd.)

- Antispam efforts and software products
  - E-mail server software
    - Limit amount of spam getting to employees
  - Client-based spam-filtering programs
  - Set filters available within client e-mail client software
  - Most effective
    - Eliminate spam before downloaded to user

# Solutions to the Spam Problem

- Methods to limit spam and its effects
  - Passing new laws
  - Technical changes in Internet mail-handling systems
  - Use existing laws and current technologies
    - Requires cooperation from large numbers of organizations and businesses
  - Use tactics available for individual e-mail users

# Solutions to the Spam Problem (cont'd.)

- Individual user antispam tactics
  - Limit spammers access to e-mail address
  - Use complex e-mail address
    - xq7yy23@mycompany.com
  - Control e-mail address exposure
    - Spammer software robots search for e-mail addresses
    - Discussion boards, chat rooms, other online sources
  - Use multiple e-mail addresses
    - Switch to another if spammers use one
  - Use filtering techniques
    - Based on contents

# Solutions to the Spam Problem (cont'd.)

- Basic content filtering
  - Content-filtering techniques differ in terms of:
    - Content elements examined
    - Spam indications
    - How strictly message classification rules applied
  - Basic content filters examine e-mail headers
  - Filtering task software location
    - **Client-level filtering**: individual users' computers
    - **Server-level filtering**: mail server computers

# Solutions to the Spam Problem (cont'd.)

- Basic content filtering (cont'd.)
  - **Black list spam filter**
    - Looks for known spammers in incoming messages' From addresses
    - Requires list to be continually updated
  - **White list spam filter**
    - Looks for good sender From addresses in incoming messages
    - High **false positives** rate
  - Used in client-level or server-level filters
    - Can also use approaches together with other content-filtering approaches

# Solutions to the Spam Problem (cont'd.)

- Challenge-response content filtering
  - Compares all incoming messages to a white list
    - If sender not on white list, automated e-mail response sent (challenge)
    - Challenge asks sender to reply to e-mail (response)
    - Reply must contain response to a challenge presented in the e-mail
  - Designed so human can respond easily
  - More information
    - Carnegie Mellon University CAPTCHA Project site

# Solutions to the Spam Problem (cont'd.)




**FIGURE 8-6** Example of a challenge that uses distorted letters and numbers

- Drawbacks
  - Potential abuse
  - Doubles amount of useless e-mail messages sent

# Solutions to the Spam Problem (cont'd.)

- Advanced content filtering
  - More effective than basic content filters
  - Looks for spam indicators in entire e-mail message
    - Indicator identified: message's spam "score" raised
  - Indicator types
    - Words, word pairs, certain HTML codes, information about where word occurs
  - Problems
    - Spammers stop including defined indicators

# Solutions to the Spam Problem (cont'd.)

- **Bayesian revision** statistical technique
  - Additional knowledge used to revise earlier probability estimates
- **Naïve Bayesian filter**
  - Software begins by not classifying messages
  - User reviews messages
  - Message type indicated to software: spam (not spam)
  - Software gradually learns message element

# Solutions to the Spam Problem (cont'd.)

- Naïve Bayesian filters' effectiveness
  - Very effective client-level filters
  - Major drawback: users must update filters regularly

# Solutions to the Spam Problem (cont'd.)

- Legal solutions
  - January 2004: U.S. CAN-SPAM law went into effect
    - Spam decreased first three months
  - After no threat of broad federal prosecution:
    - Spam rates increased
  - CAN-SPAM regulates:
    - All e-mail messages for advertising or promoting commercial product or service
    - Messages promoting Web site content




**FIGURE 3-7** U.S. Federal Trade Commission Spam information site home page

# Solutions to the Spam Problem (cont'd.)

- Technical solutions
  - Slowing down acknowledgment messages
    - Originating computer will slow (must continue to scan for acknowledgment)
    - Will not send more messages until acknowledgment received
    - Requires defending company to develop way to identify computers sending spam

# Solutions to the Spam Problem (cont'd.)

- **Teergrubing** ("tar pit"): launching a return attack
  - Sending e-mail messages back to computer originating suspected spam
- Teergrubing objective
  - Ensure computer sending spam is trapped
  - Drag down ability to send spam
  - Concern: counterattack might violate laws
- Ultimate spam problem
  - New e-mail protocols providing absolute verification of e-mail message source

# Web Site Utility Programs

- TCP/IP supports utility programs (tools)
  - Run on Web server or client computers
- Earliest Internet utility program
  - E-mail
    - Most important utility
    - Key element in electronic commerce strategies

# Finger and Ping Utilities

- **Finger** program
  - Runs on UNIX operating systems
  - Provides information about other network users
  - Many organizations disable Finger command
    - Privacy and security
  - Built into some e-mail programs
- **Ping**: Packet Internet Groper
  - Tests connectivity between two Internet-connected computers
  - Provides performance data about connection
  - Available as freeware and shareware

# Tracert and Other Route-Tracing Programs

- **Tracert** (TRACE RouTe)
  - Sends data packets to every computer on path
    - Between one computer and another computer
  - Clocks packets' round-trip times
  - Provides indication of time message needs to travel from one computer to another and back
  - Ensures remote computer online
  - Pinpoints data traffic congestion
  - Calculates and displays:
    - Number of hops between computers
    - Time to traverse entire one-way path

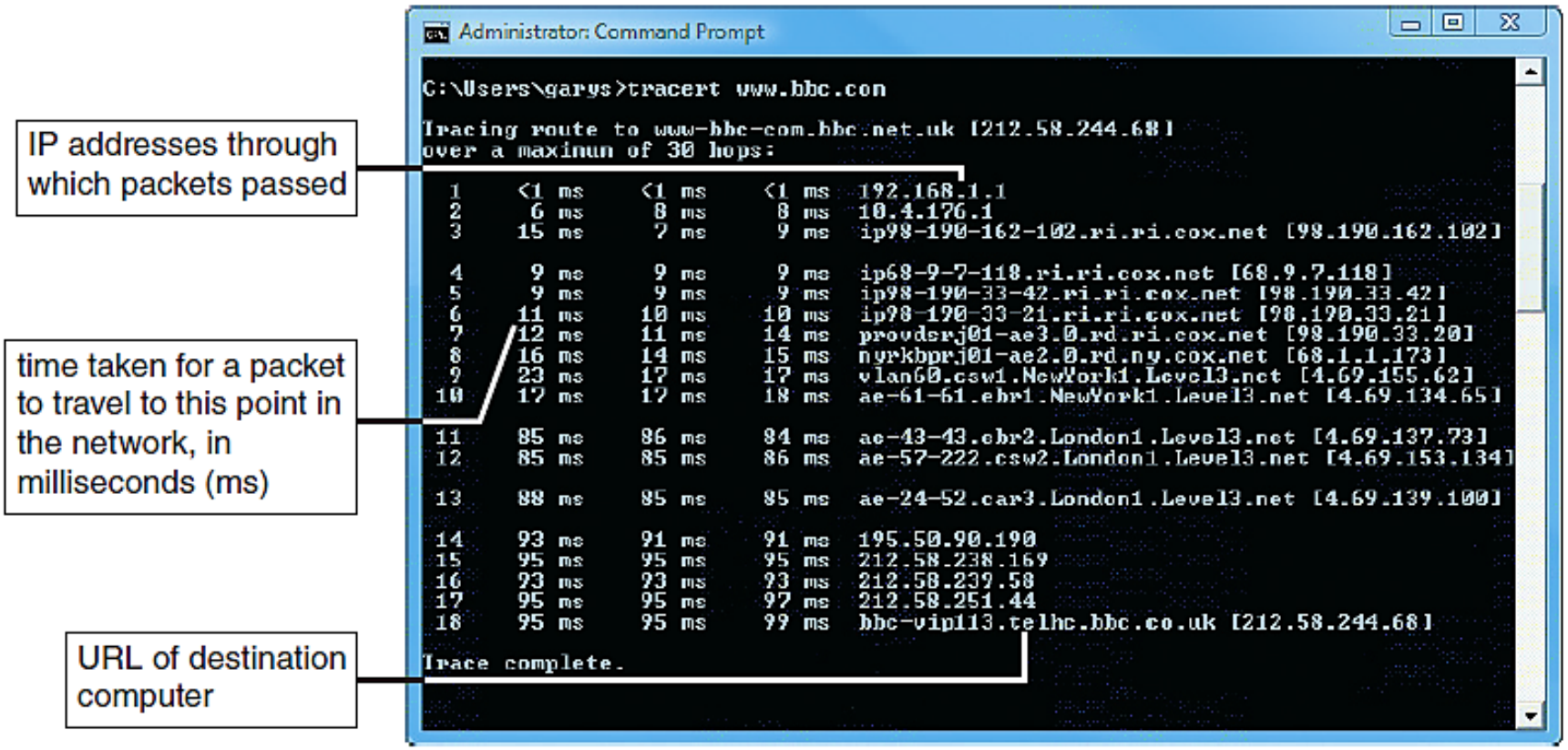



**FIGURE 3-8** Tracing a path between two computers on the Internet

# Telnet and FTP Utilities

- **Telnet** program
  - Provides remote login capability
  - Useful if no Web interface
  - Availability
    - Free Internet downloads, Microsoft Telnet.exe
  - Provides remote troubleshooting
  - **Telnet protocol**: set of rules used by Telnet program
  - Web browser Telnet client
    - “telnet://” followed by remote host domain name
  - Telnet use decreasing

# Telnet and FTP Utilities (cont'd.)

- **File Transfer Protocol (FTP)**
  - Part of TCP/IP rules defining formats
    - Transfer files between TCP/IP-connected computers
  - Useful services
    - Displaying remote, local computers' directories
    - Changing current client's or server's active directory
    - Creating and removing local and remote directories
  - Uses TCP and its built-in error controls:
    - To copy files accurately

# Indexing and Searching Utility Programs

- Search engines (search tools)
  - Search for requested documents on specific site or entire Web
- Indexing program
  - Provides full-text indexing
- Browser search methods
  - Compare index terms to requester's search term
  - Use complex relevance ranking rules
    - Advanced search engine software (Google)
- Web server software contains indexing software

# Data Analysis Software

- Web servers capture visitor information
  - Placed into Web **log file** (grows quickly)
- Third-party Web log file analysis programs summarize information
  - Query log file
  - Return gross summary information or accumulating details
- Popular Web log file analysis programs
  - Adobe SiteCatalyst, Urchin from Google, WebTrends

# Link-Checking Utilities

- **Dead link**
  - Displays error message rather than Web page when clicked
- **Link rot**
  - Site contains a number of links that no longer work
- **Link checker**
  - Examines each site page
    - Reports broken, incorrect URLs
  - Identifies **orphan files**
    - Web site file not linked to a page

# Link-Checking Utilities (cont'd.)

- Link checker (cont'd.)
  - Script checking and HTML validation
- Link-checking programs
  - Adobe Dreamweaver, Elsop LinkScan (separate utility)
- **Reverse link checker**
  - Checks company's link exchange program sites
  - Ensures link exchange partners fulfilling obligation
    - Include link back to company's Web site
  - Example: LinxCop

# Remote Server Administration

- **Remote server administration**
  - Web site administrator controls Web site
    - Monitor server activity
    - Manipulate server
    - Access from any Internet-connected computer
  - Provides convenience
  - Examples
    - LabTech Software
    - NetMechanic

# Web Server Hardware

- Hosting electronic commerce operations
  - Wide variety of computer brands, types, sizes used
  - Small companies
    - Run Web sites on desktop PCs
  - Most Web sites
    - Operate on computers designed for site hosting

# Server Computers

- Comparing desktop PCs to server computers
  - Servers use faster and higher-capacity hardware
- Costs
  - Low-end: $1,000-$1,500
  - More common: $2,000-$100,000
- Companies selling Web server hardware provide Web site configuration tools
- Housing Web server computers
  - Freestanding cases
  - Installed in equipment racks

# Server Computers (cont'd.)

- **Blade servers**: servers-on-a-card
  - Small: 300 installed in single 6-foot rack
- Fundamental Web server job
  - Process and respond to HTTP Web client requests
- **Virtual server** (**virtual host**)
  - Maintains more than one server on one machine
  - Different groups have separate domain names
    - All domain names refer to same physical Web server

Think Stock

**FIGURE 3-9** Rack-mounted blade servers

E-Business, Tenth Edition

# Web Servers and Green Computing

- Electrical power needs:
  - Operating server
  - Cooling the room where server is housed
- Green computing
  - Efforts to reduce the environmental impact of large computing installations
- Novel approaches using natural cooling
  - Google server facility in Finland
  - Facebook server in Lulea, Sweden
  - Other companies: Hewlett-Packard, FedEx, Harris Corporation

# Web Server Performance Evaluation

- **Benchmarking**: testing to compare hardware and software performance
- Elements affecting overall server performance
  - Hardware, operating system software, server software, connection speed, user capacity, type of Web pages delivered
  - Connection speed (T3 faster than T1)
  - Number of users server can handle
    - Important and hard to measure

# Web Server Performance Evaluation (cont'd.)

- **Throughput**: HTTP requests hardware and software process in a unit of time
- **Response time**: time server requires to process one request
- Choosing Web server hardware configurations
  - Run tests on various combinations, consider scalability, compare standard benchmarks
  - Independent testing lab: Mindcraft
- Benchmarks developed by:
  - Standard Performance Evaluation Corporation

# Web Server Hardware Architectures

- Electronic commerce Web sites use tiered architecture
  - Divides work of serving Web pages
  - May use more than one computer within each tier
- **Server farms**: large collections of servers
  - Lined up row after row
- **Centralized architecture**
  - Uses a few large and fast computers
    - Requires expensive computers
    - More sensitive to technical problems
    - Requires adequate backup plans

# Web Server Hardware Architectures (cont'd.)

- **Distributed architecture** (**decentralized architecture**)
  - Uses large number of less-powerful computers
    - Spreads risk over large number of servers
    - Uses less-expensive servers
    - Requires additional hubs or switches to connect servers to each and the Internet
    - Requires cost of load balancing

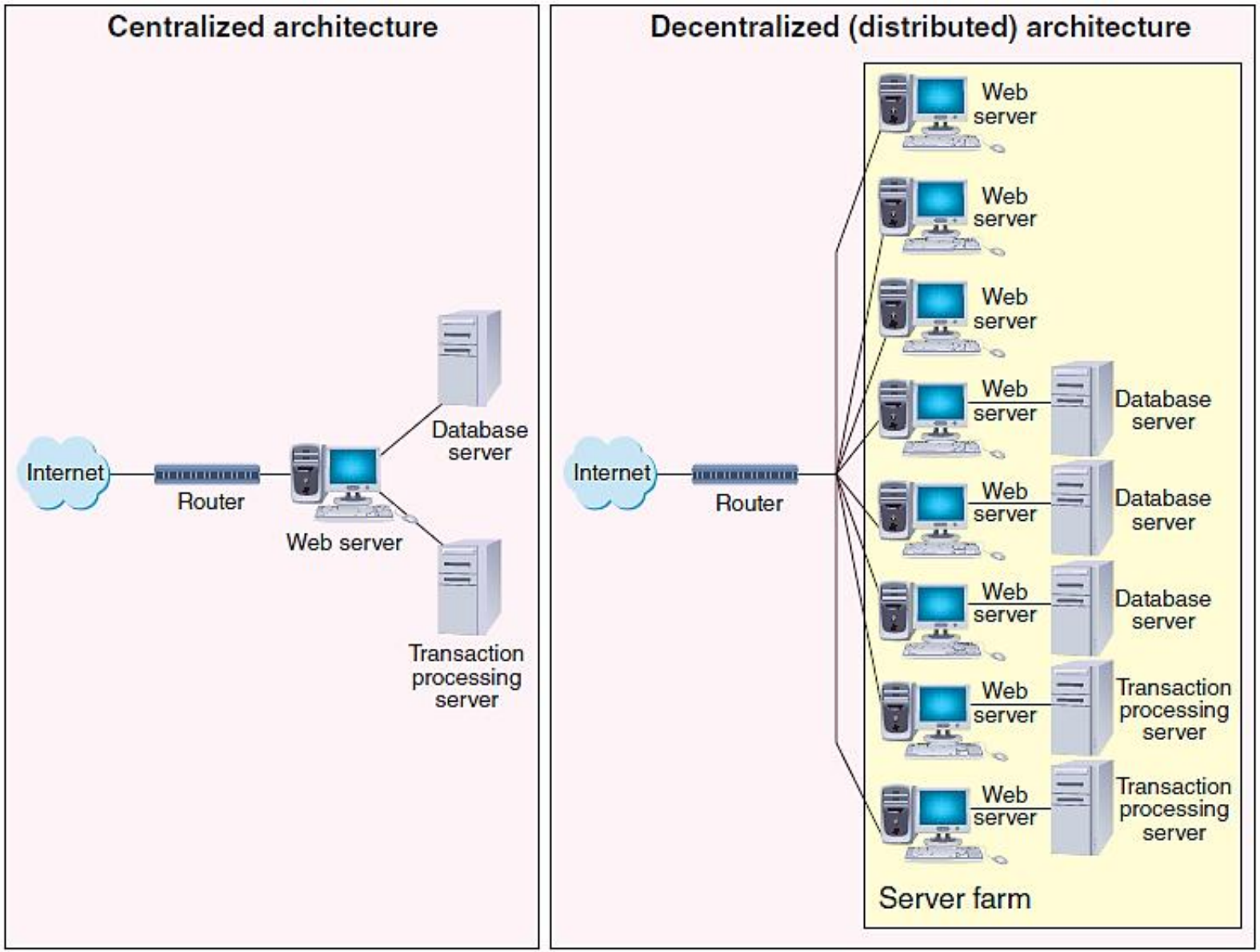



**FIGURE 3-10** Centralized and decentralized Web site architectures
E-Business, Tenth Edition

# Web Server Hardware Architectures (cont'd.)

- Load-balancing systems
  - **Load-balancing switch**
    - Network hardware monitoring server workloads
    - Assigns incoming Web traffic to the server with most available capacity
  - Simple load-balancing system
    - Traffic enters through site's router
    - Encounters load-balancing switch
    - Directs traffic to best Web server

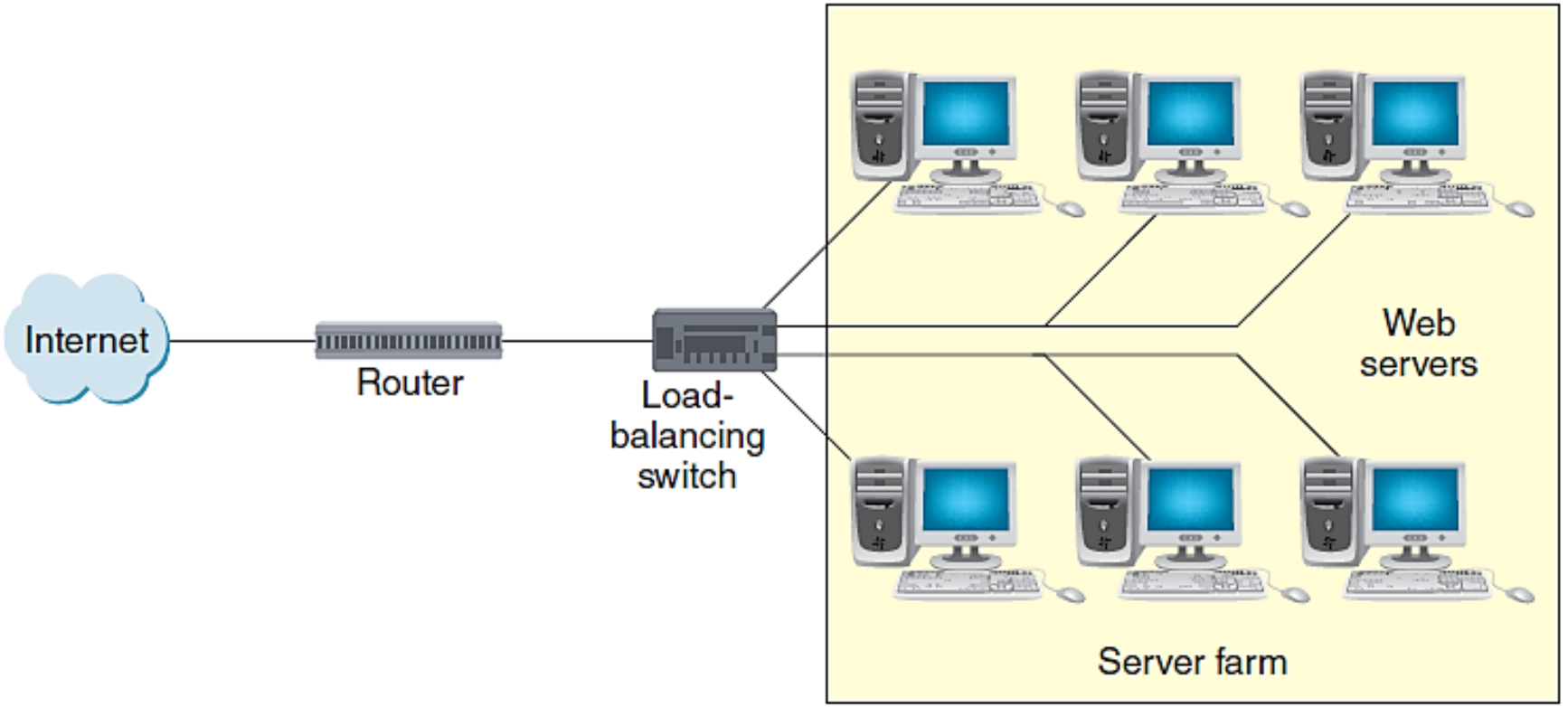





**FIGURE 3-11** Basic load-balancing system

# Web Server Hardware Architectures (cont'd.)

- Load-balancing systems (cont'd.)
  - More complex load-balancing systems
    - Incoming Web traffic enters from two or more routers
    - Directed to groups of dedicated Web servers

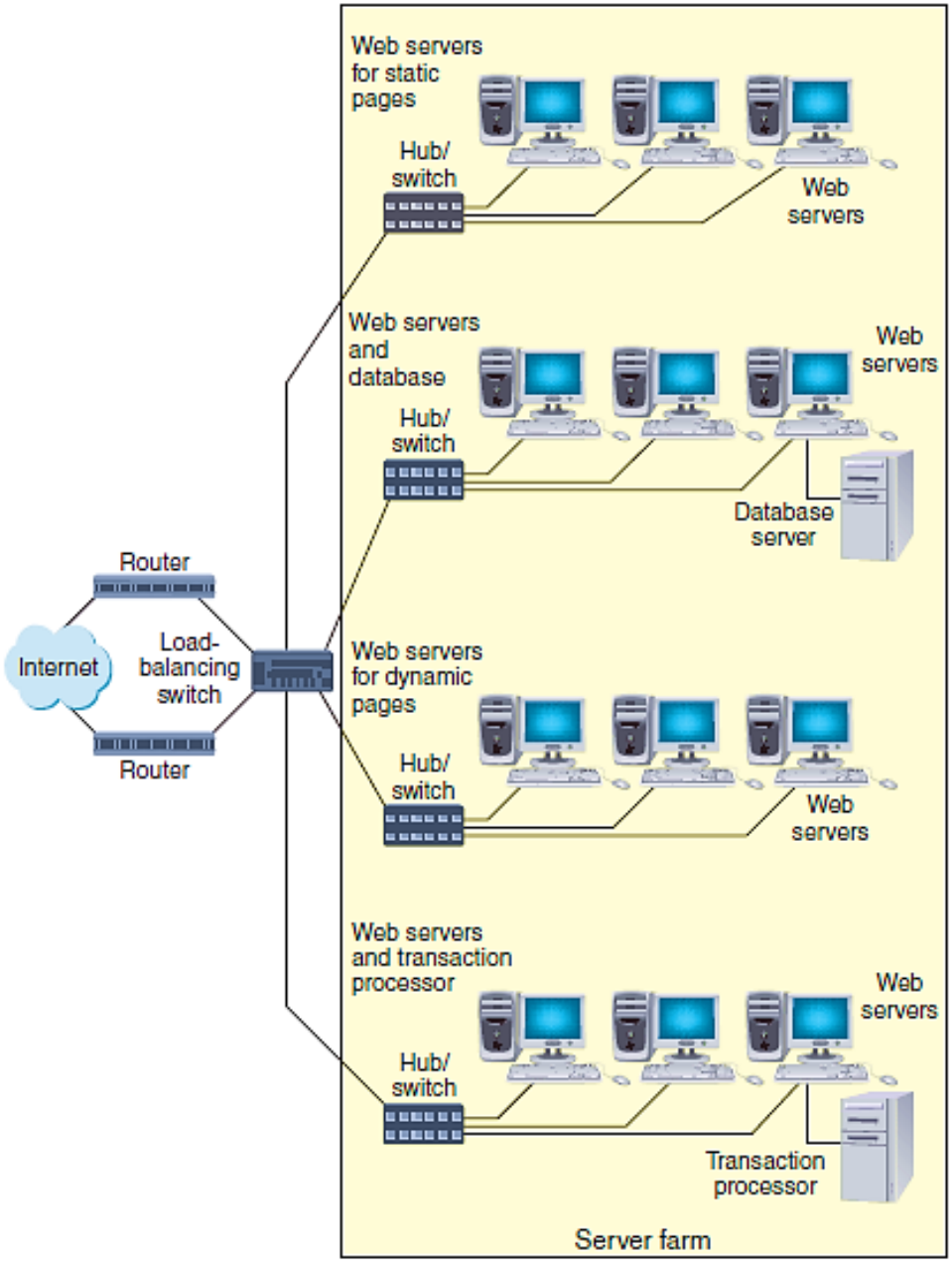



**FIGURE 3-12** Complex load balancing

# Summary

- Client/server Web architecture
  - HTTP-based tiered architectures
- Several operating systems used on Web servers
- Web server utility programs can be helpful
- E-mail has benefits and drawbacks
  - Spam problem has grown dramatically
- Understand Web server performance
  - Factors, evaluation tools, solutions
- Web server hardware
  - Important consideration in online business site design