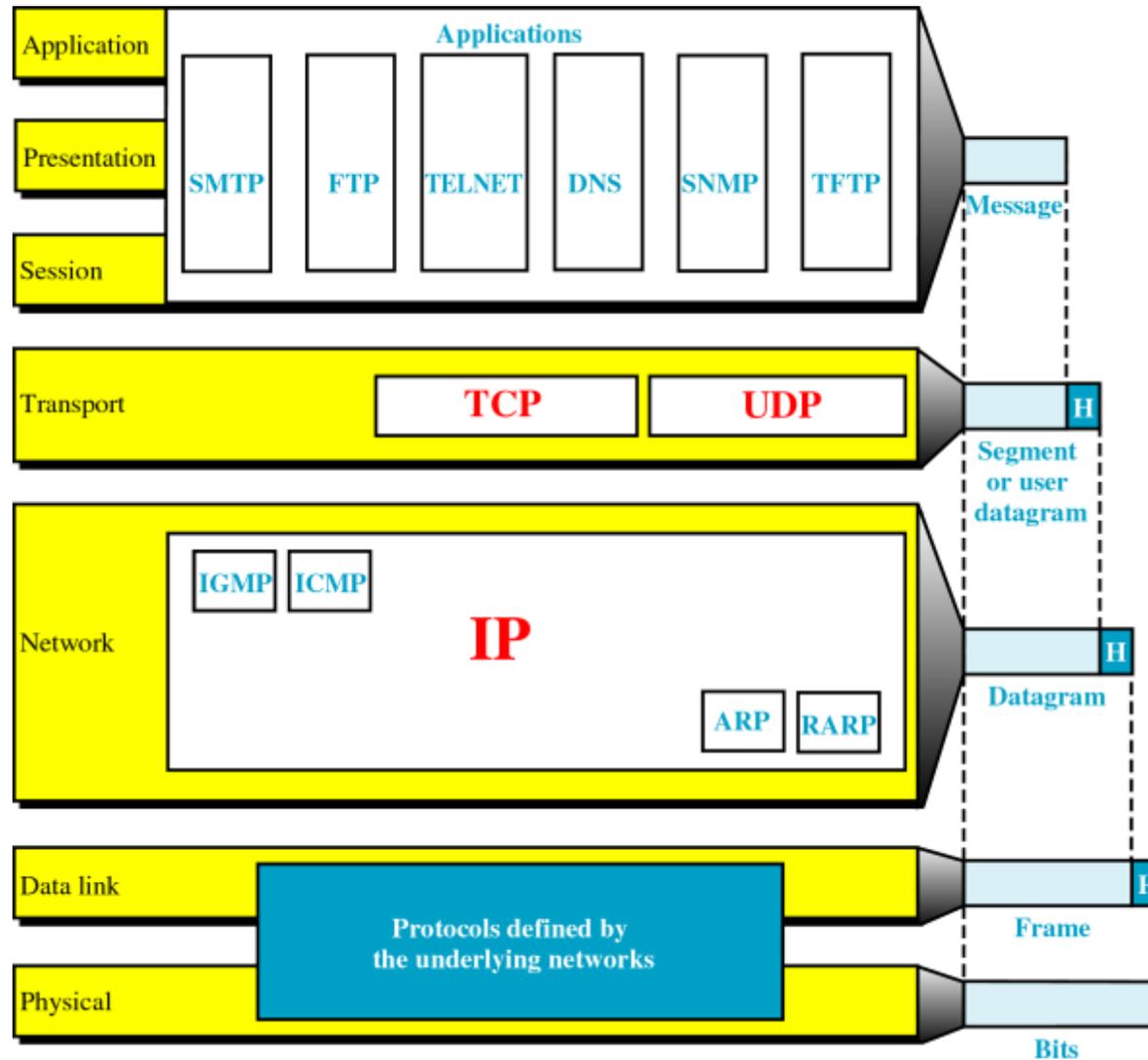


**APPLICATIONS LAYER :
WWW (HTTP, HTML),
E-MAIL(SMTP), SNMP**

**Dr. Ravindra Nath UIET CSJM
University Kanpur**

1

TCP/IP and the OSI Model



WORLD WIDE WEB

- World Wide Web (or WWW) is the most common medium for sharing information by people around the world.
- The basic idea is that one can download and look at *hypertext* multi-media documents from anywhere in the world using the Internet.
- The clients for WWW are called browsers. Several easy to use *browsers* such as Netscape Communicator and Internet Explorer are available.
- The clients obtain the information by interacting with the *HTTP servers* so called because HTTP is the protocol used by the clients and the servers to interact with each other.

The **WWW** today is a distributed client/server service, in which a client using a browser can access a service using a server. However, the service provided is distributed over many locations called sites.

Architecture of WWW

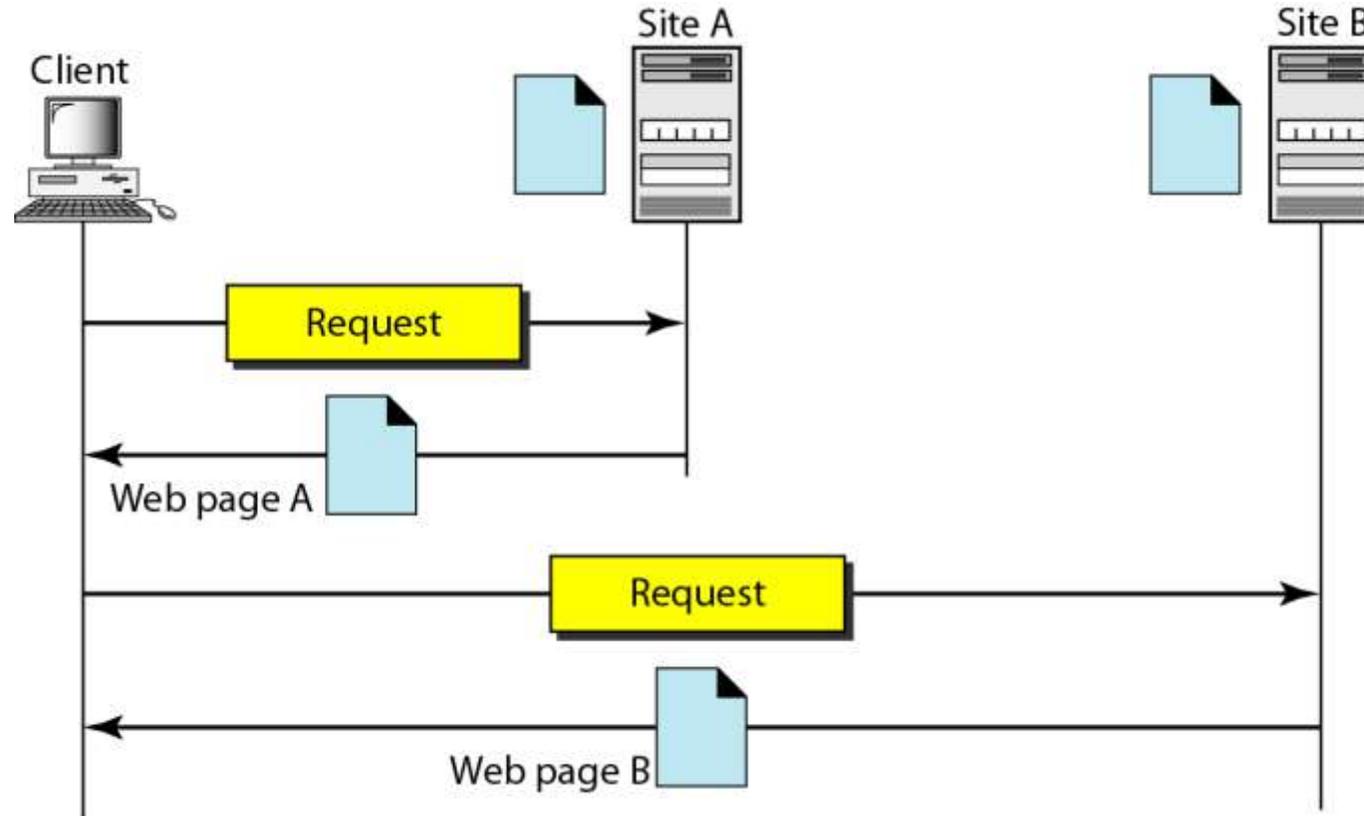


Figure 25-28

Distributed Services

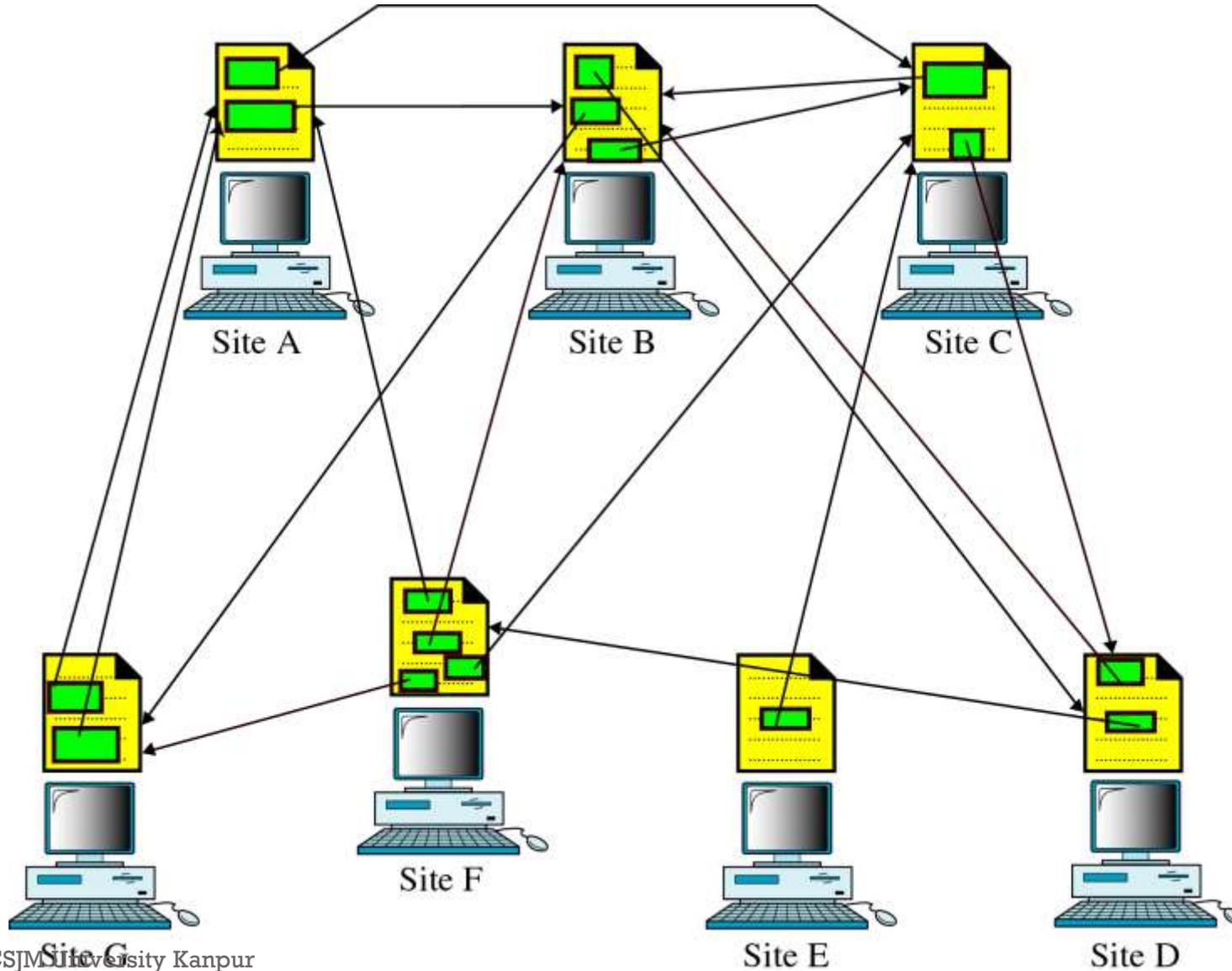
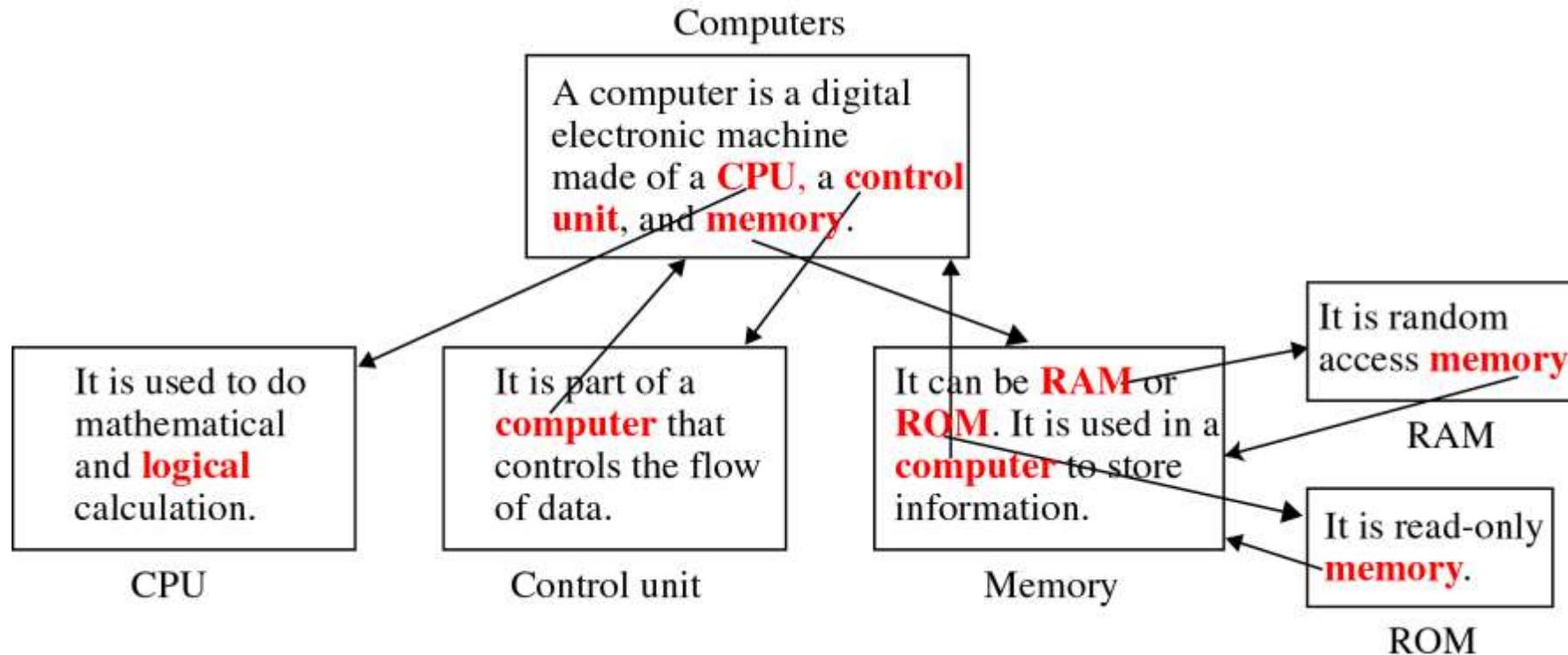
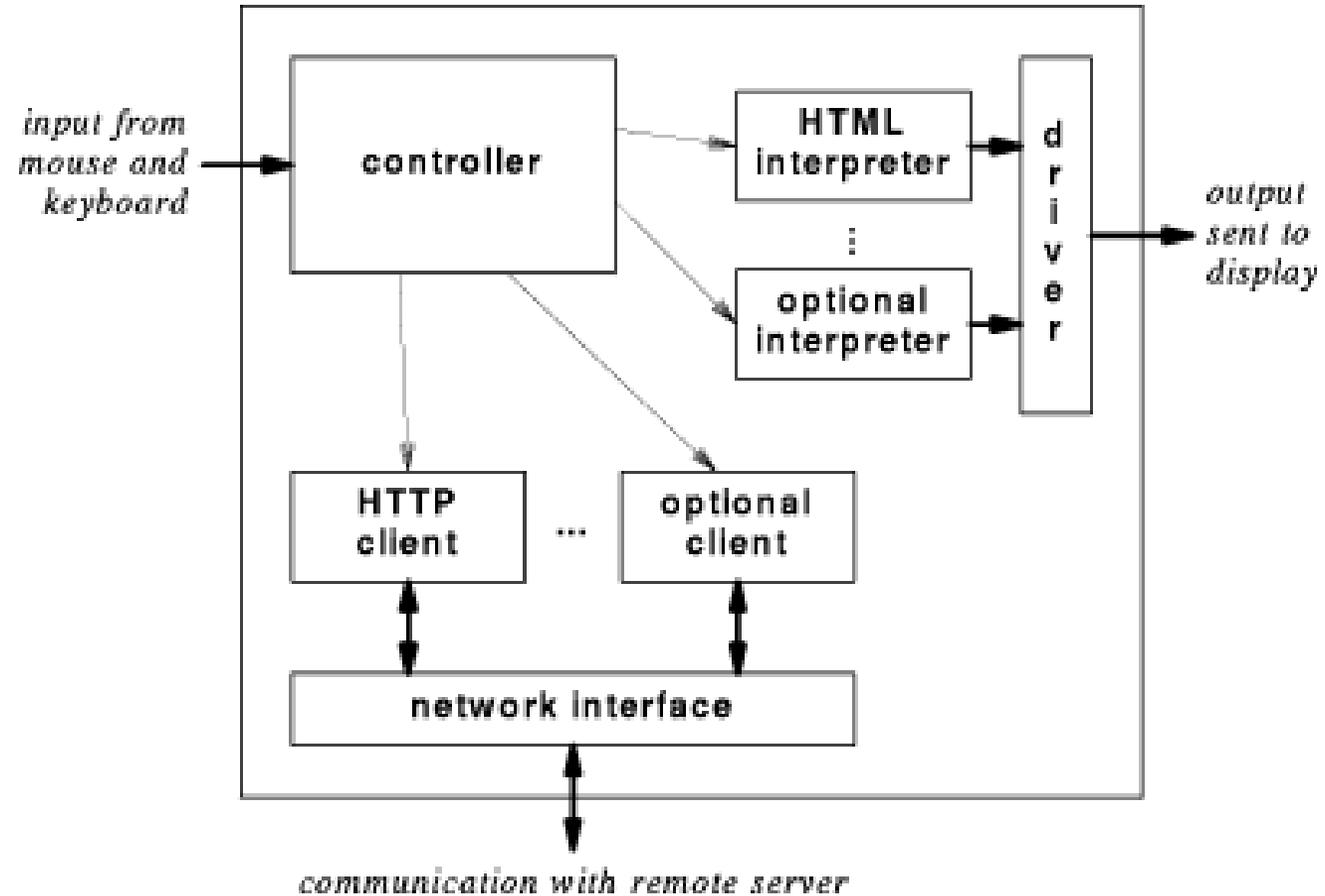


Figure 25-29

Hypertext



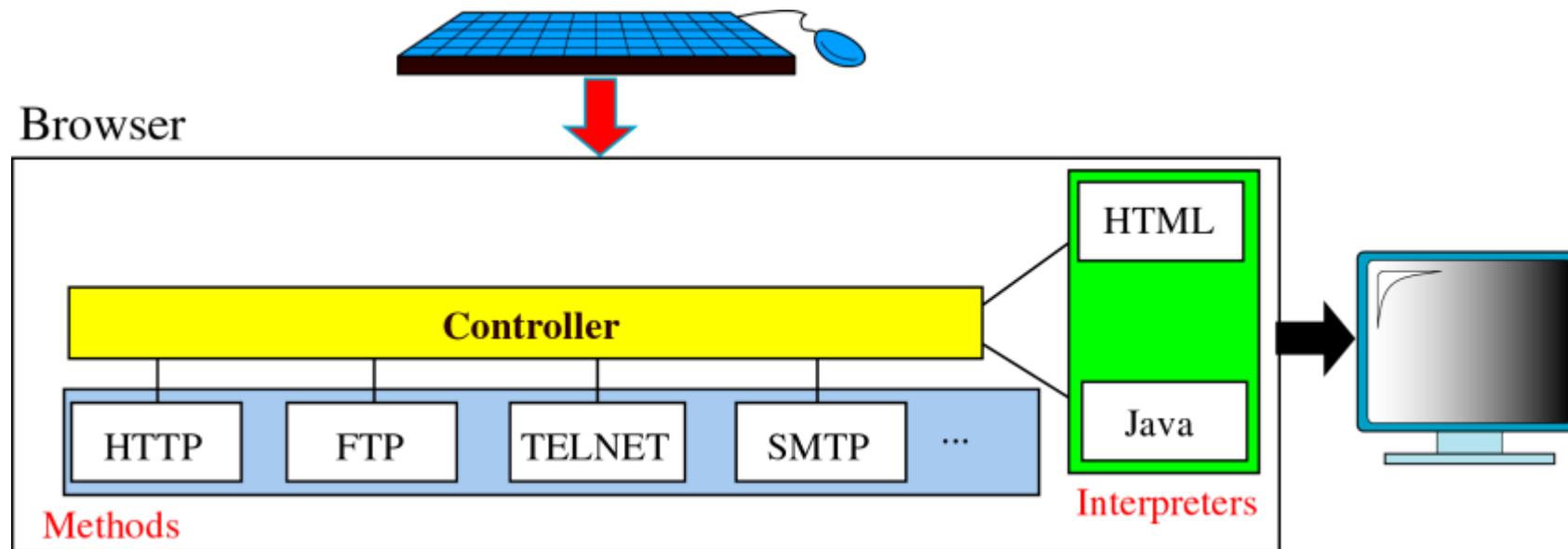
Major components of a Web browser



Caching : local cache on disk, Proxy

Figure 25-30

Browser Architecture



THE BASIC WEB TERMINOLOGY

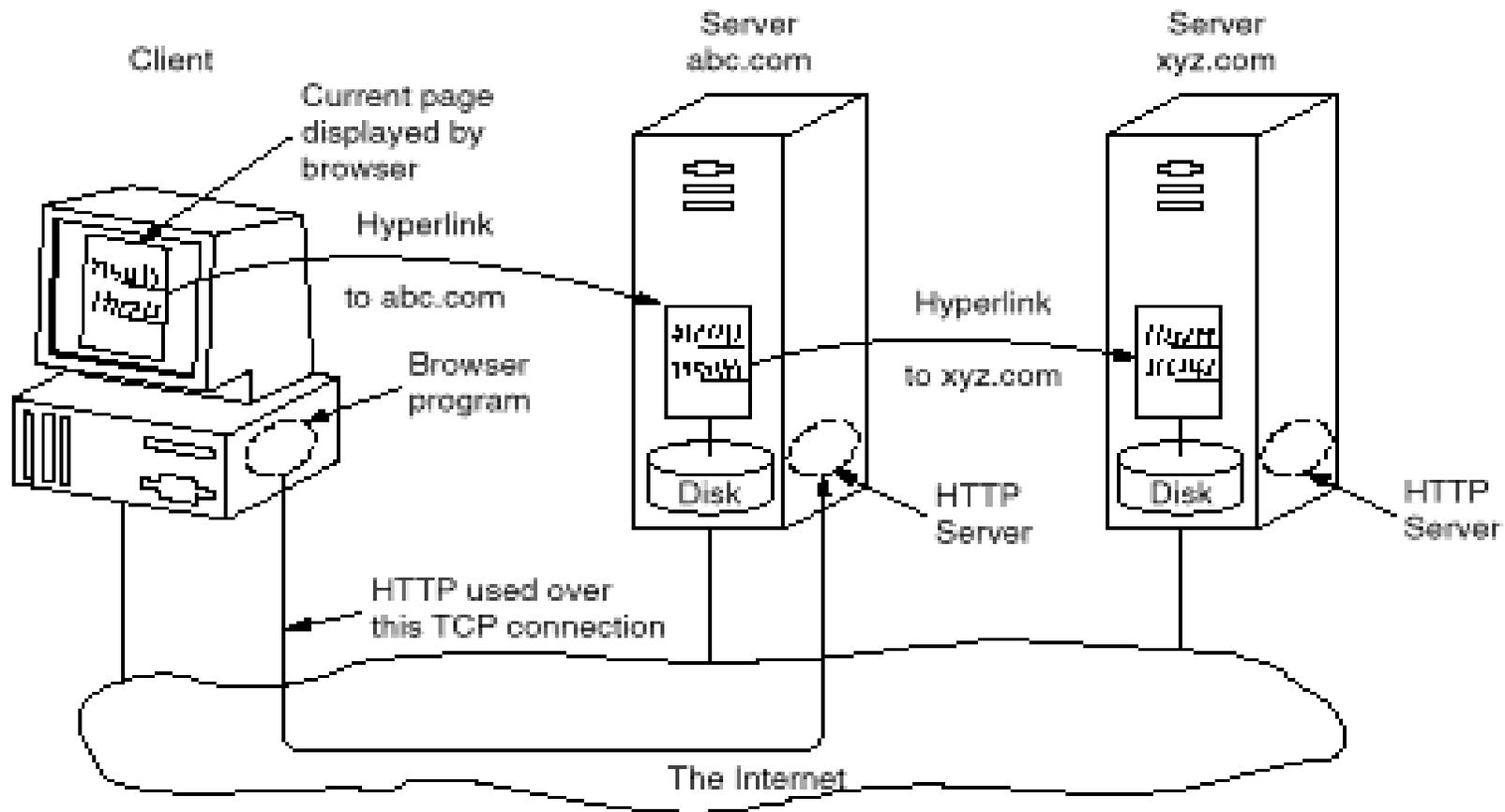
- Internet is the global backbone.
- URLs (Universal Resource Locator) are used to globally name and access all web resources.
- HTTP (HyperText Transfer Protocol) is used to retrieve URL-named resources.
- HTML (HyperText Markup Language) is used to describe the logical structure of documents.
- Web browsers are the client programs.

STEPS IN A SIMPLE WEB INTERACTION

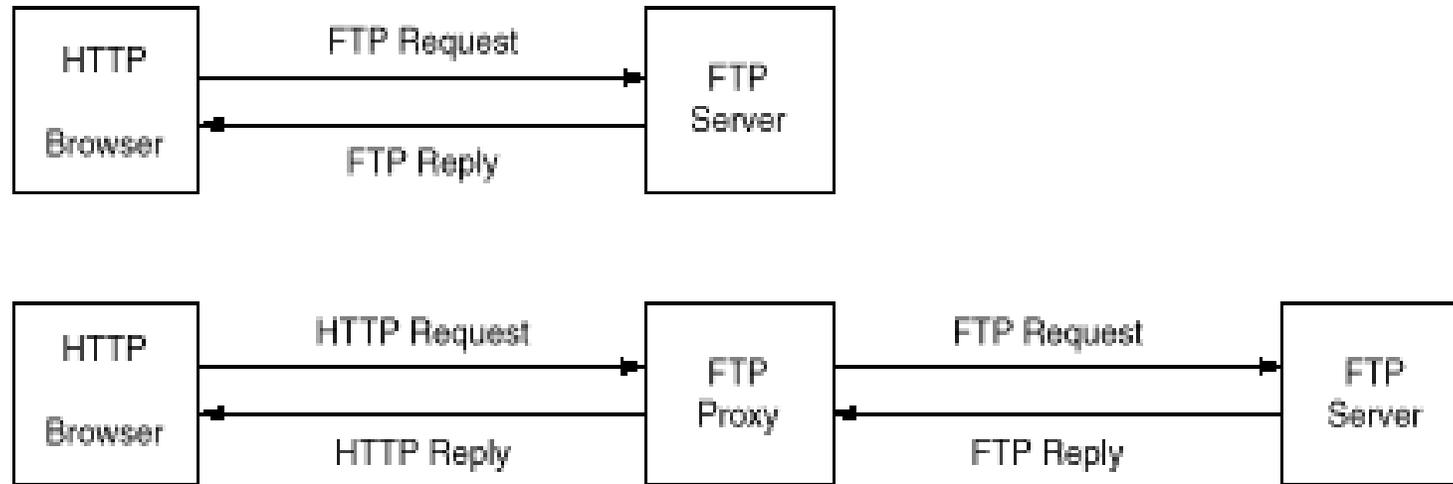
- The user specifies the URL of the document she wishes to view in the browser.
- The browser sends an HTTP request for the URL to the server.
- The server sends back the HTML document to the browser.
- The browser interprets and formats the HTML document and displays it on the screen.

Ex: <http://www.w3.org/hypertext/WWW/Project.html>

- 1. The browser determines the URL**
- 2. The browser asks the DNS for the IP of www.w3.org**
- 3. DNS replies with 18.23.0.23**
- 4. Browser makes TCP connection to port 80 on 18.23.0.23**
- 5. Sends a GET /hypertext/WWW/Project.html command**
- 6. The www.w3.org server sends the file Project.html**
- 7. The TCP connection is released.**
- 8. The browser the displays the text in Project.html**
- 9. The browser fetches and displays all images in Project.html**



The parts of the Web model.



(a) A browser that speaks FTP. (b) A browser that does not.

HYPertext TRAnSFER PRoTocol

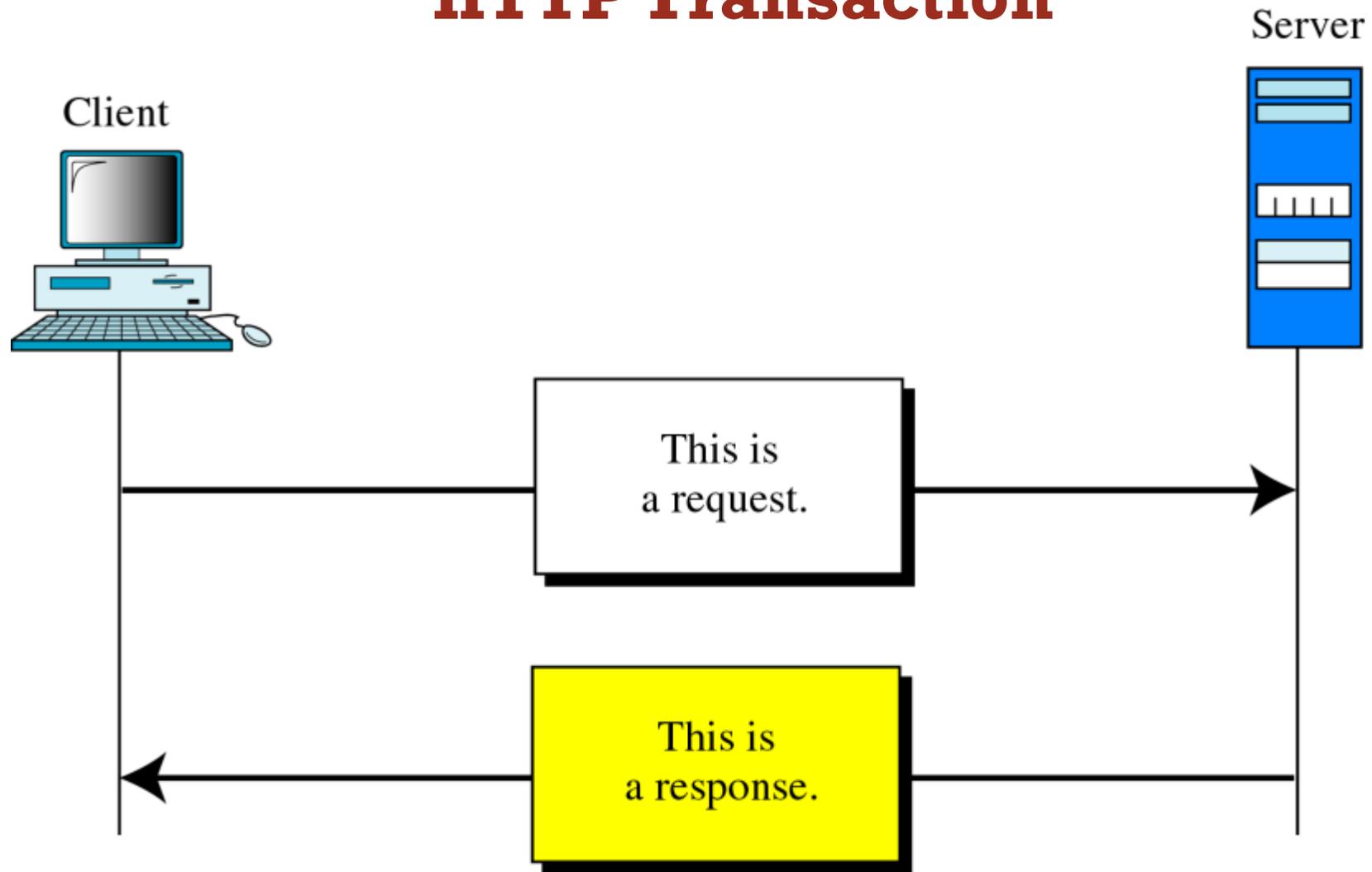
- The World Wide Web is designed around a client-server technology.
- HTTP is a set of rules that defines how Web browsers and servers communicate with each other over the network connection.
- The simplest command is **GET** which gets a document.
- HTTP also supports
 - POST to send information to some program on server side.
 - Storing a web page.
 - Deleting the web page.
 - Authentication, if the access is to be restricted.
 - Cache related directives.

Method	Description
GET	Request to read a Web page
HEAD	Request to read a Web page's header
PUT	Request to store a Web page
POST	Append to a named resource (e.g., a Web page)
DELETE	Remove the Web page
LINK	Connects two existing resources
UNLINK	Breaks an existing connection between two resources

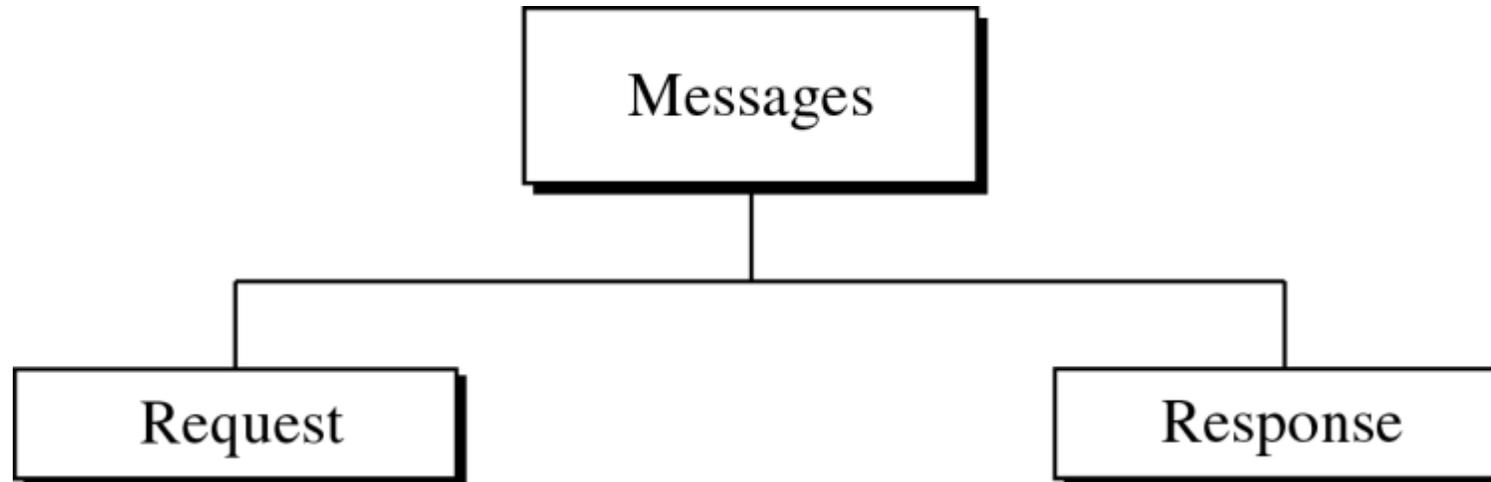
The built-in HTTP request methods.

Figure 25-23

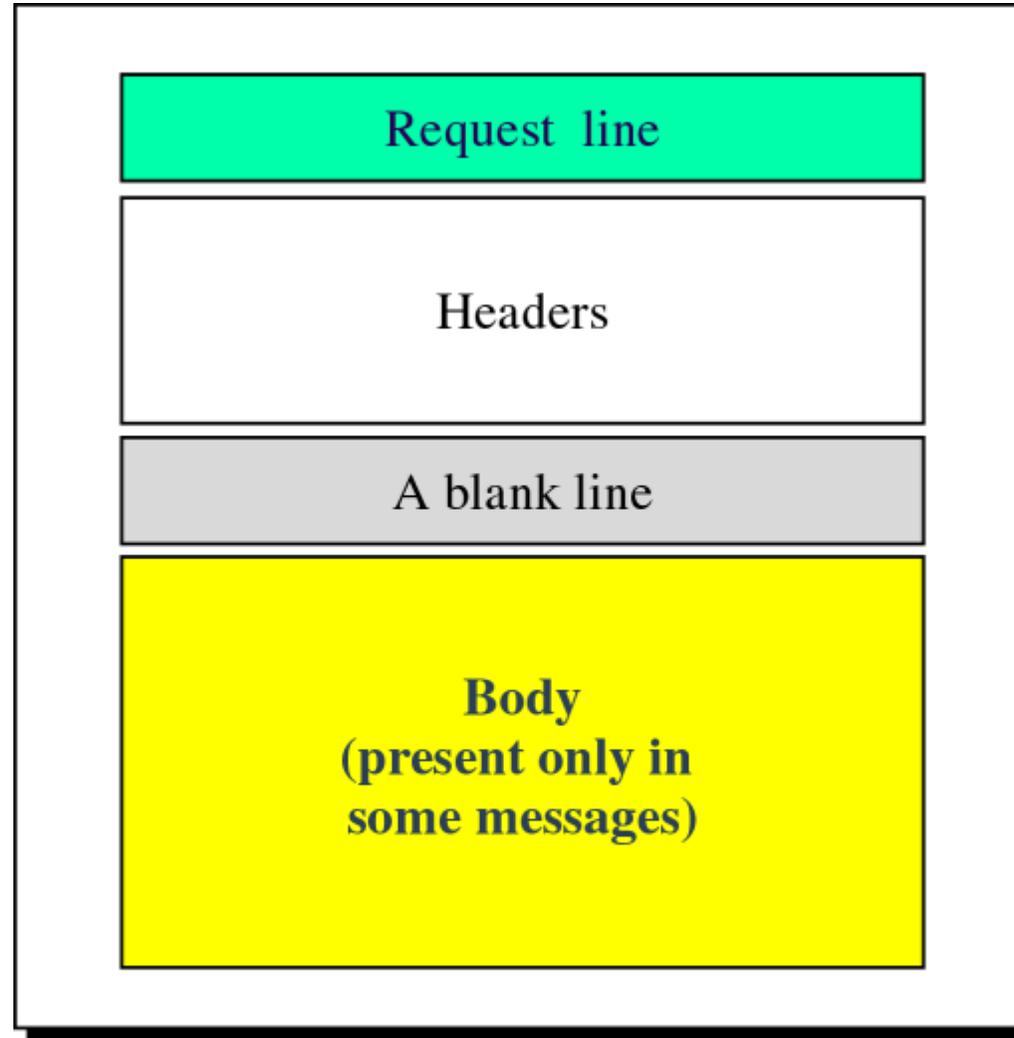
HTTP Transaction



Message Categories

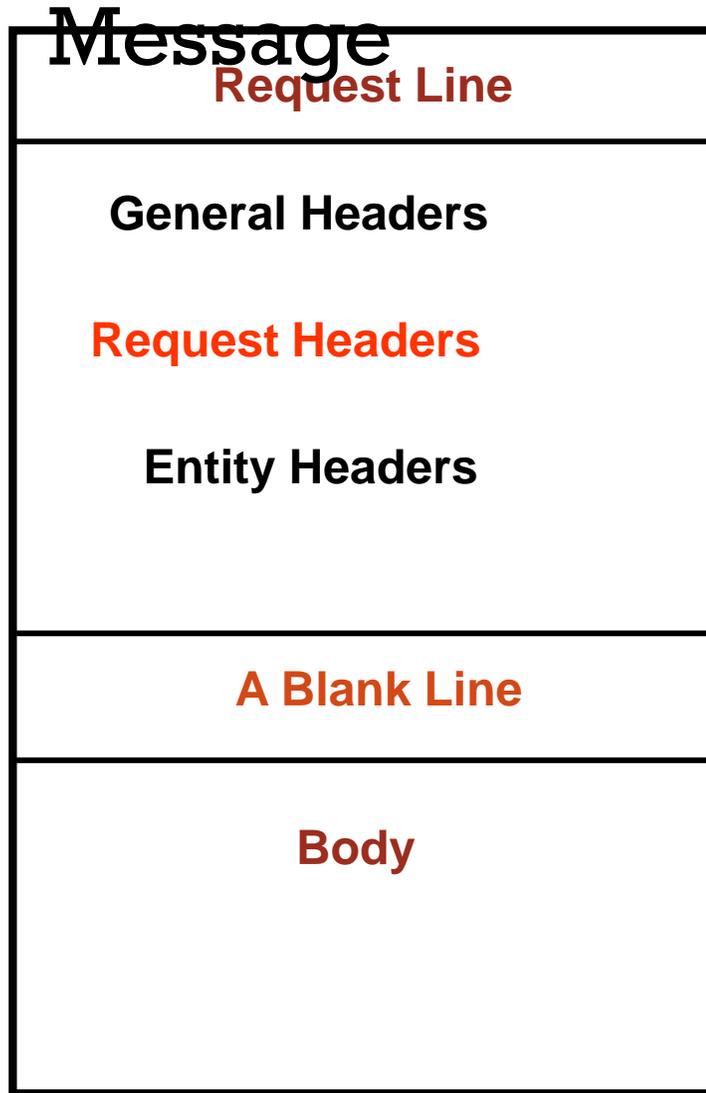


Request Message

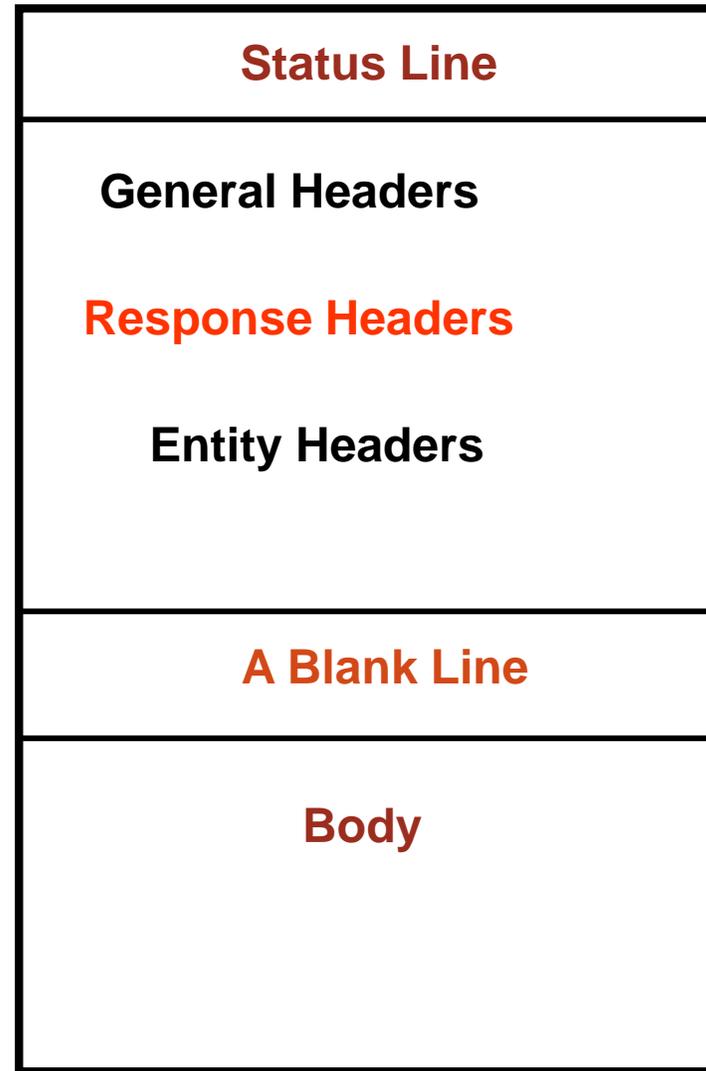


Request message

Request Message



Response



UNIFORM RESOURCE LOCATOR (URL)

- Much of the information on the Internet is organized into pages.
- Every single page is kept distinct in a world of electronic pages by using URLs.
- Each page has a unique URL.
- URL is a string that provides location and access information of the page.
- URL is also used to launch certain applications (protocols), for example, email, news, etc.

URL COMPONENTS

■ Protocol

- Identifies a manner for interpreting computer information.
- Many Internet pages use HTTP or FTP.

■ Server

- Identifies the computer system that stores the information.
- Each server has a unique address.

■ Port Number

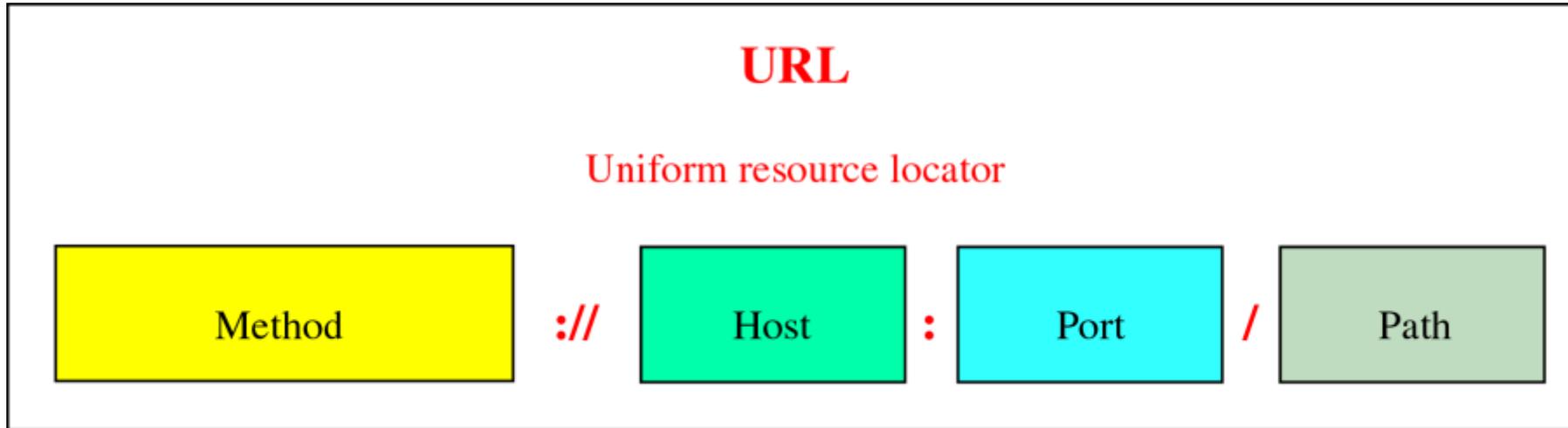
- What port number the server software is running on.
- Mostly missing from URL, since defaults are used.

■ Pathname

- Identifies the location of an item on the server.

Figure 25-27

URL



URL - AN EXAMPLE

- Here's a sample URL -

http://home.netscape.com:80/index.html

- URL text has components that specify the protocol, the server, port number and pathname of the item.

- The protocol is followed by a colon - **http:**

- The server is preceded by two slashes -

//home.netscape.com

- The port number is preceded by a colon and followed by a slash - **:80/**

- Each segment of the pathname is preceded by a single slash - **/index.html**

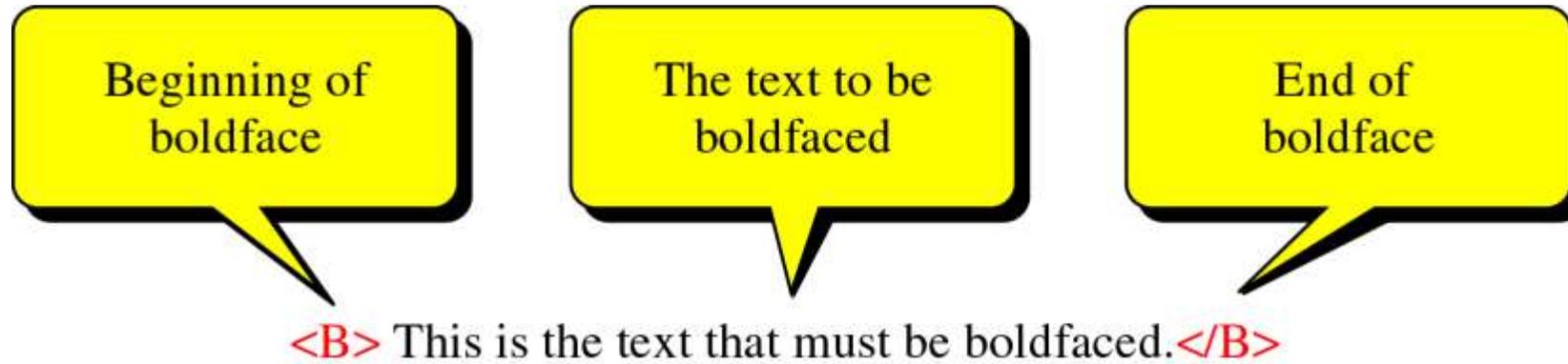
HYPertext MARKUP LANGUAGE

- HTML is a collection of markup tags, defining various components of a WWW document.
- Is simple and platform-independent.
- Supports the hypermedia features.
- Documents are plain text files.

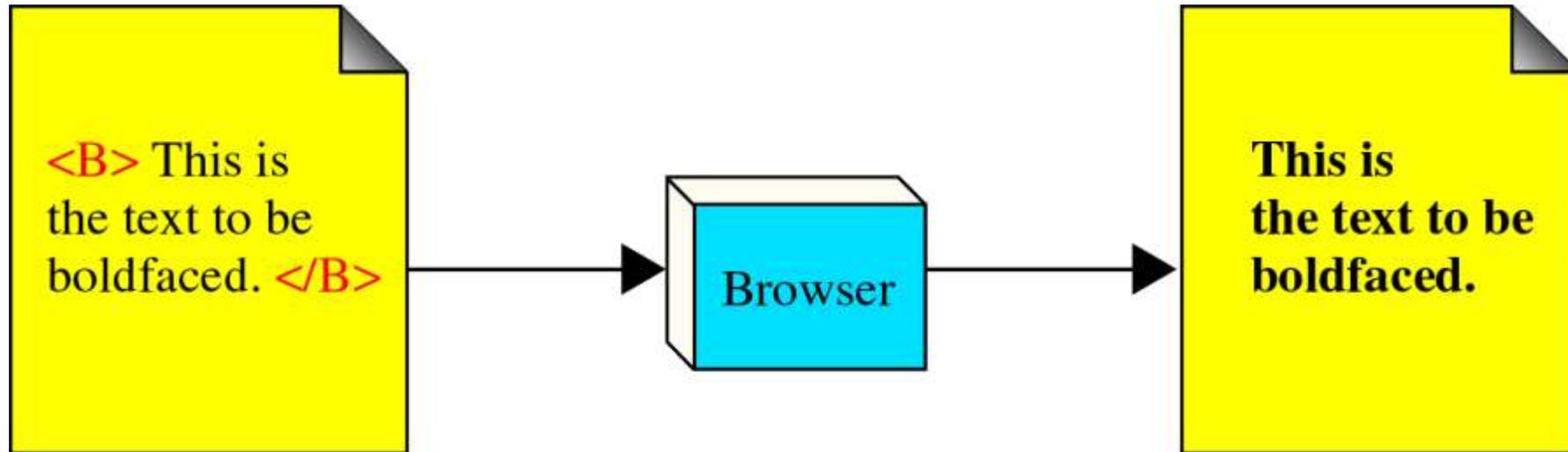
Tag	Description
<HTML> ... </HTML>	Declares the Web page to be written in HTML
<HEAD> ... </HEAD>	Delimits the page's head
<TITLE> ... </TITLE>	Defines the title (not displayed on the page)
<BODY> ... </BODY>	Delimits the page's body
<Hn> ... </Hn>	Delimits a level <i>n</i> heading
 ... 	Set ... in boldface
<I> ... </I>	Set ... in italics
 ... 	Brackets an unordered (bulleted) list
 ... 	Brackets a numbered list
<MENU> ... </MENU>	Brackets a menu of items
	Start of a list item (there is no)
 	Force a break here
<P>	Start of paragraph
<HR>	Horizontal rule
<PRE> ... </PRE>	Preformatted text; do not reformat
	Load an image here
 ... 	Defines a hyperlink

A selection of common HTML tags. Some have additional parameters.

Boldface Tags



Effect of Boldface Tags



Beginning and Ending Tags

```
< TagName Attribute = Value Attribute = Value ... >
```

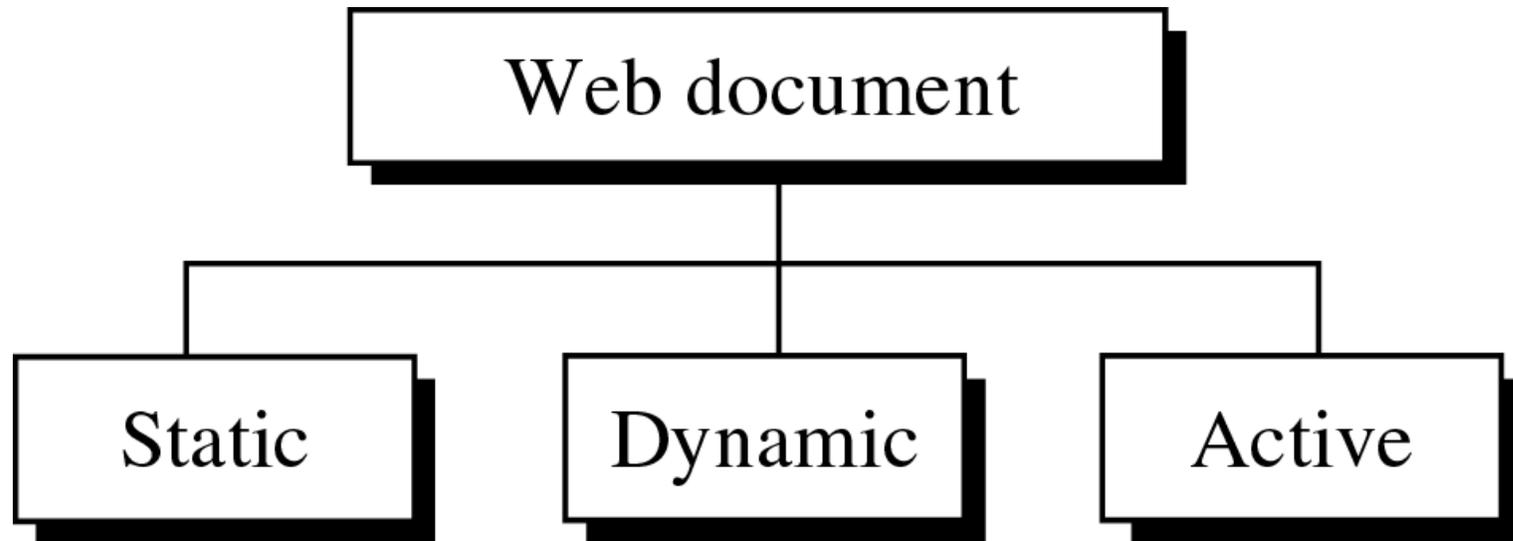
a. Beginning tag

```
< /TagName >
```

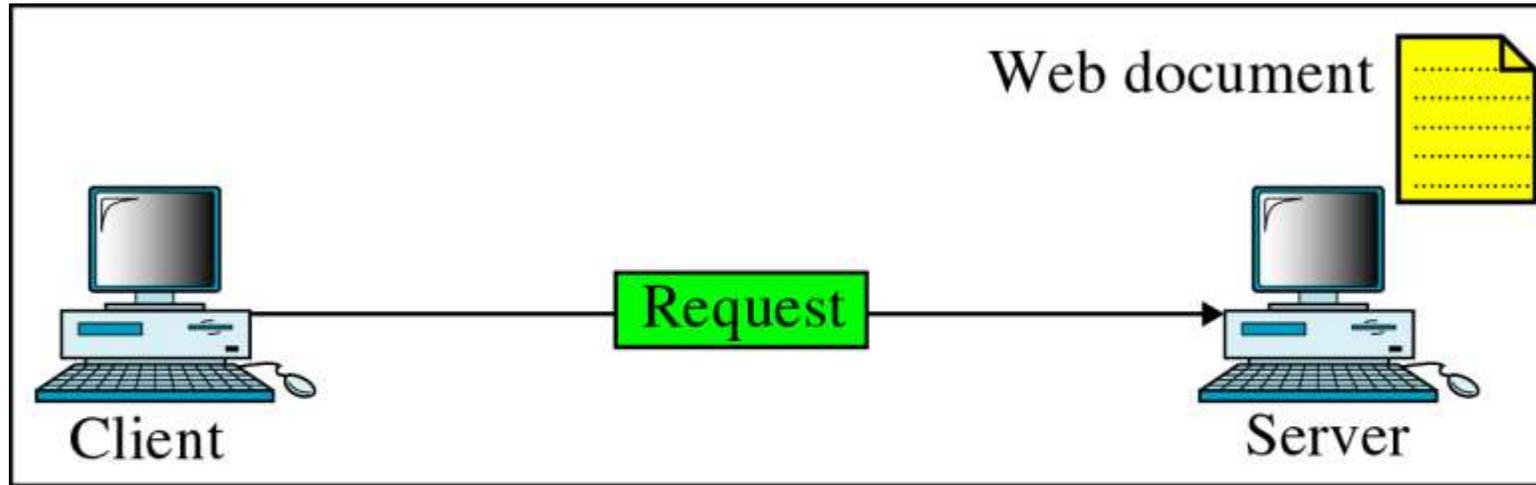
b. Ending tag

```
C: telnet www.w3.org 80
T: Trying 18.23.0.23 ...
T: Connected to www.w3.org.
T: Escape character is '^'.
C: GET /hypertext/WWW/TheProject.html HTTP/1.0
C:
  S: HTTP/1.0 200 Document follows
  S: MIME-Version: 1.0
  S: Server: CERN/3.0
  S: Content-Type: text/html
  S: Content-Length: 8247
  S:
  S: <HEAD> <TITLE> The World Wide Web Consortium (W3C) </TITLE> </HEAD>
  S: <BODY>
  S: <H1> <IMG ALIGN=MIDDLE ALT="W3C" SRC="icons/WWW/w3c_96x67.gif">
  S: The World Wide Web Consortium </H1> <P>
  S:
  S: The World Wide Web is the universe of network-accessible information.
  S: The <A HREF="Consortium/"> World Wide Web Consortium </A>
  S: exists to realize the full potential of the Web. <P>
  S:
  S: W3C works with the global community to produce
  S: <A HREF="#Specifications"> specifications </A> and
  S: <A HREF="#Reference"> reference software </A> .
  S: W3C is funded by industrial
  S: <A HREF="Consortium/Member/List.html"> members </A>
  S: but its products are freely available to all. <P>
  S:
  S: -
```

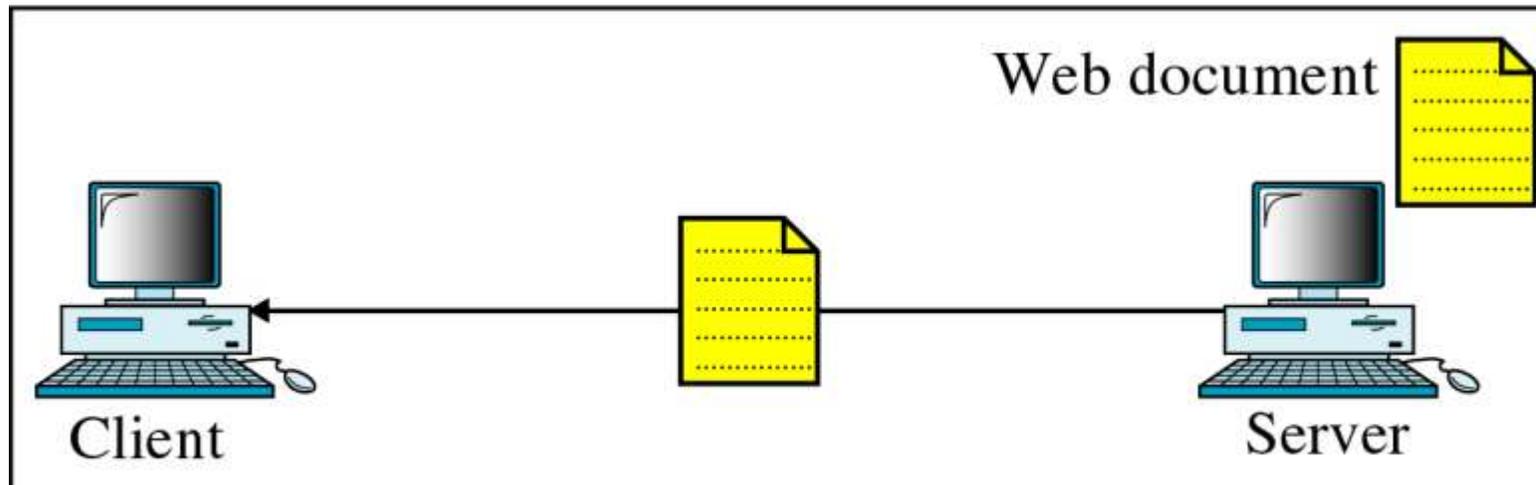
Categories of Web Documents



Static Document



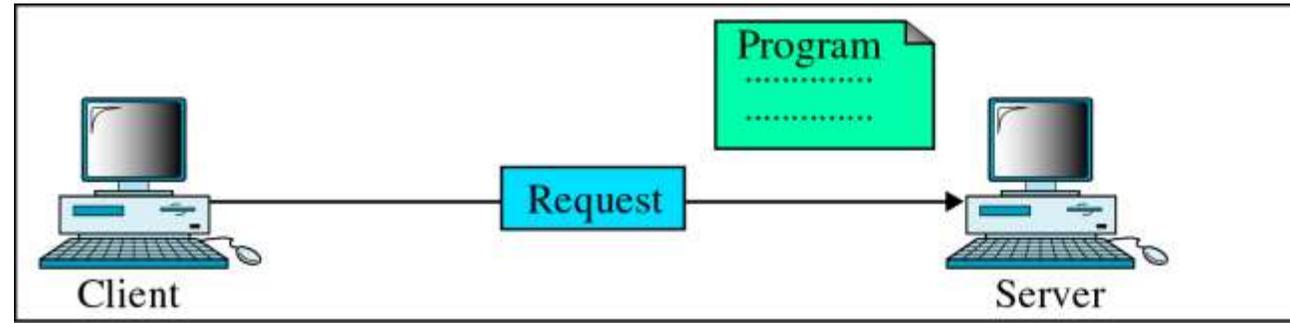
a. Request



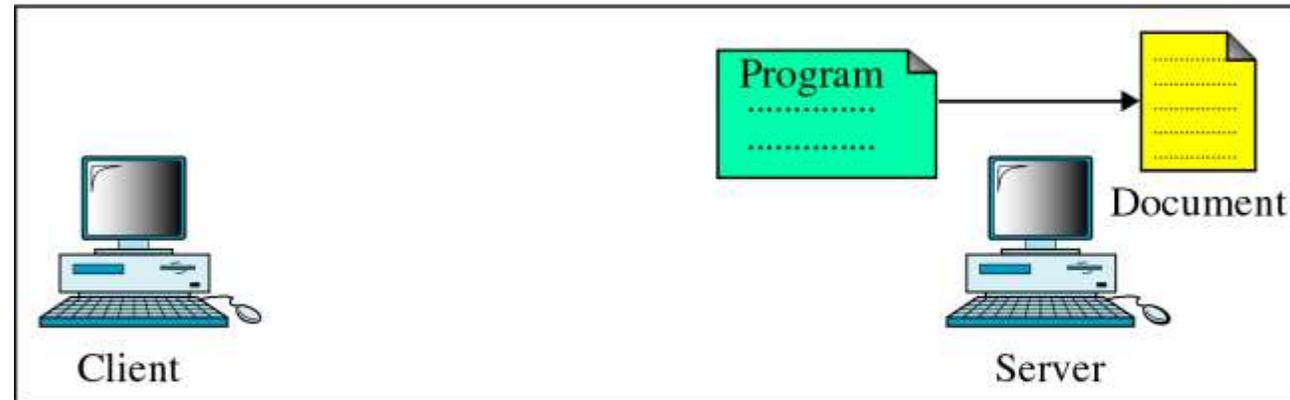
b. Response

Figure 25-36

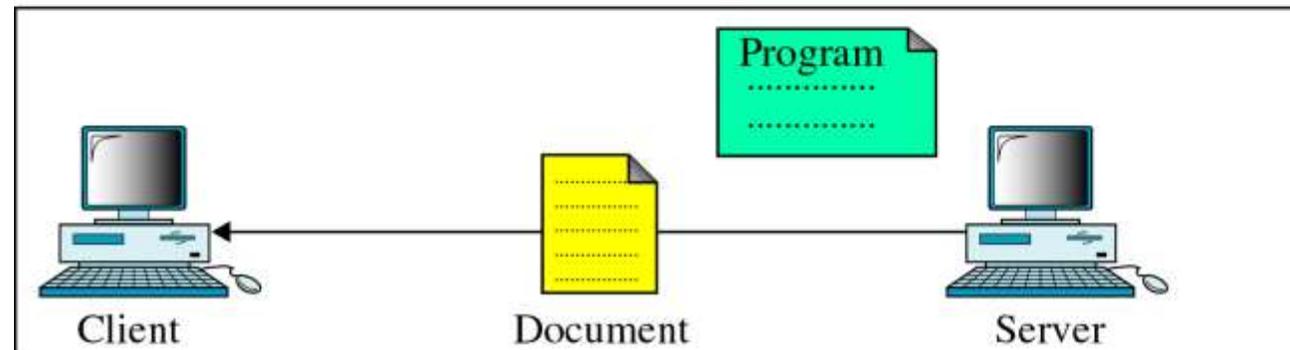
Dynamic Document



a. Request for running a program



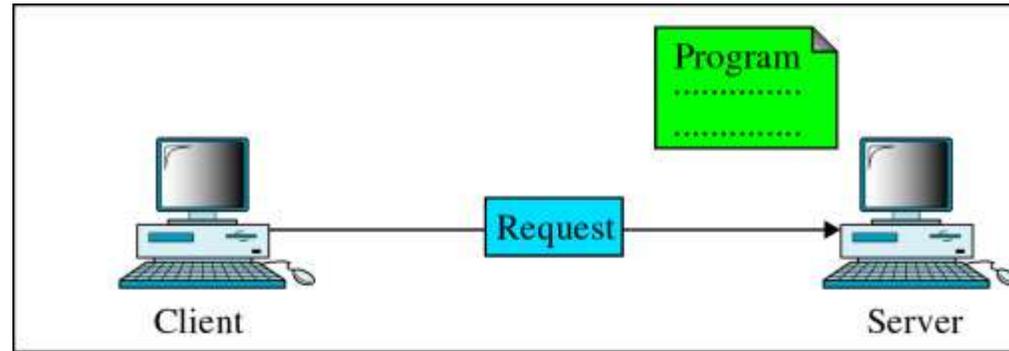
b. Running the program and creating the document



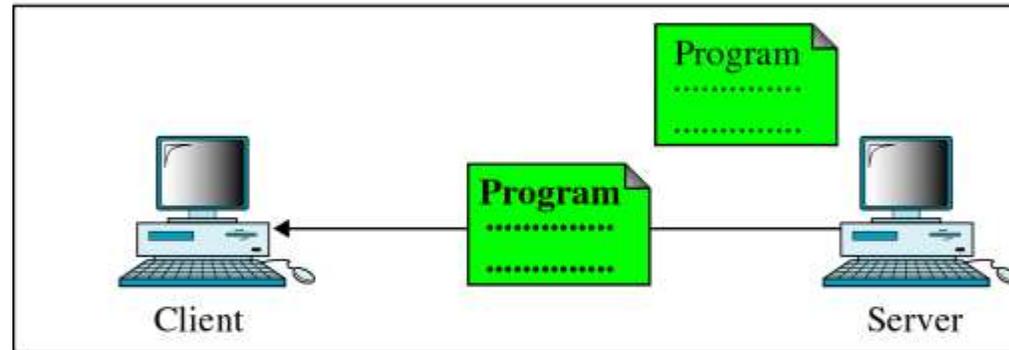
c. Response

Figure 25-37

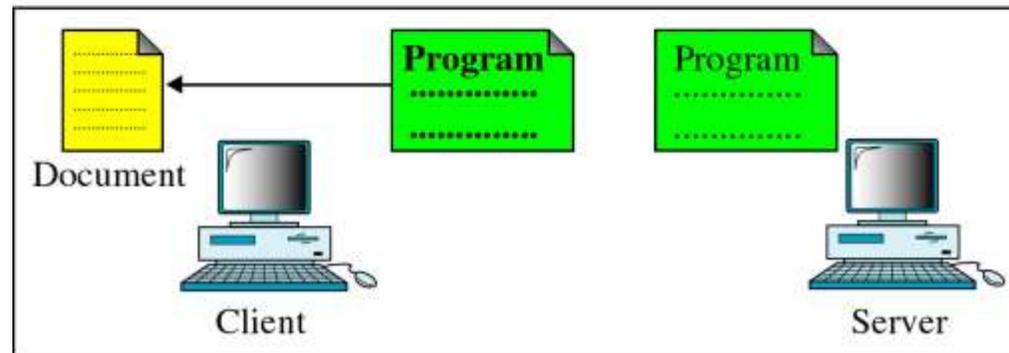
Active Document



a. Request for a copy of a program



b. Sending a copy of the program



c. Running the program and creating the document

Table 27.2 Status codes

<i>Code</i>	<i>Phrase</i>	<i>Description</i>
Informational		
100	Continue	The initial part of the request has been received, and the client may continue with its request.
101	Switching	The server is complying with a client request to switch protocols defined in the upgrade header.
Success		
200	OK	The request is successful.
201	Created	A new URL is created.
202	Accepted	The request is accepted, but it is not immediately acted upon.
204	No content	There is no content in the body.

Table 27.2 Status codes (continued)

<i>Code</i>	<i>Phrase</i>	<i>Description</i>
Redirection		
301	Moved permanently	The requested URL is no longer used by the server.
302	Moved temporarily	The requested URL has moved temporarily.
304	Not modified	The document has not been modified.
Client Error		
400	Bad request	There is a syntax error in the request.
401	Unauthorized	The request lacks proper authorization.
403	Forbidden	Service is denied.
404	Not found	The document is not found.
405	Method not allowed	The method is not supported in this URL.
406	Not acceptable	The format requested is not acceptable.
Server Error		
500	Internal server error	There is an error, such as a crash, at the server site.
501	Not implemented	The action requested cannot be performed.
503	Service unavailable	The service is temporarily unavailable, but may be requested in the future.

SEARCH ENGINES

- A tool to locate required information from the vast resources of the WEB.
- It has the following components:
- Communication between user and search engines (submitting query, CGI)
- Searching
 - * Building/updating of databases and indexing. (A regular and periodic activity.)
 - * Searching indexed databases (query processing).

WEB ROBOTS

- A robot is a program that automatically traverses the Web's hypertext structure by retrieving a document, and recursively retrieving all documents that are referenced.
- Normal Web browsers are not robots, because they are operated by a human, and they do not automatically retrieve referenced documents (other than inline images).
- Other similar programs are:
Spiders, Worms, Web Crawlers, Web Ants.

ROBOT APPLICATIONS

- **Resource Discovery.** Building searchable databases. Any manual effort would be futile due to vast/diverse amount of information available.
- **Link Validation.** Dead links are difficult to detect. Now, robots that perform regular updates on a database can help maintain the validity of its contents as well as its structure.
- **Site Mirroring.** Entire directories from a site are recursively copied by FTP and maintained periodically by copying files that have been changed or added. This is tedious and time consuming.
- **Monitoring ``What's New.''**
- **Statistical Analysis.**

WHAT IS E-MAIL

- Communication using computer networks.
- Information can be sent by one person on a computer to a person on another computer anywhere on the network.
- Information is usually text, but other kind of information (graphics, binary files, even speech and video) is possible.
- The message can also be encrypted for security.
- E-mail can be addressed to multiple persons at the same time.

Figure 25-11

SMTP Concept

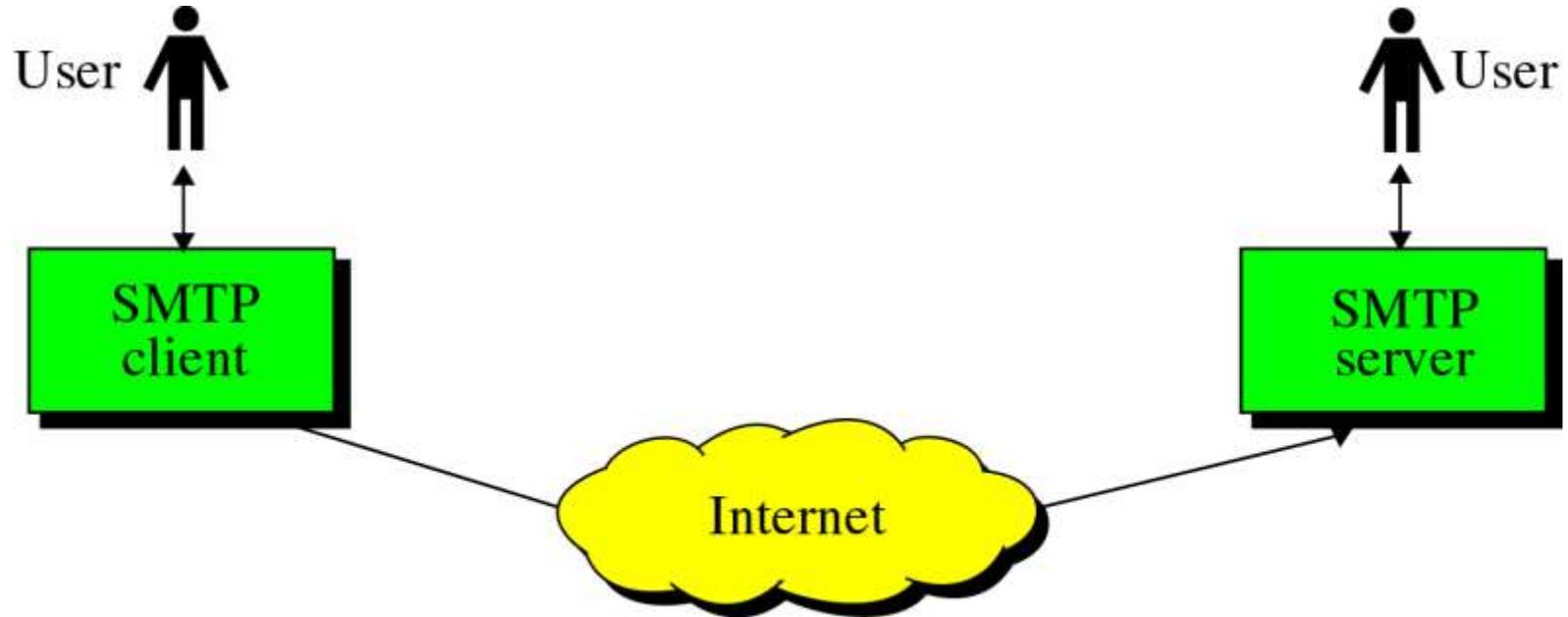
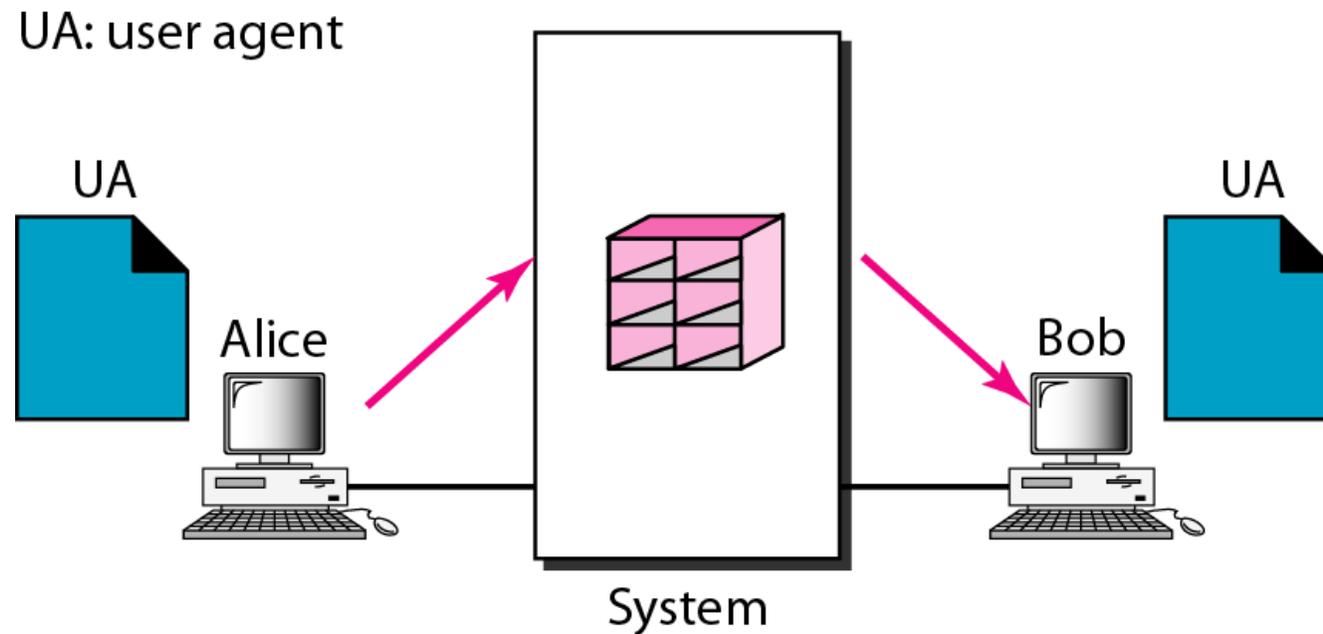


Figure 26.6 *First scenario in electronic mail*



**When the sender and the receiver of an e-mail are on the same system,
we need only two user agents.**

Figure 25-12

UAs and MTs

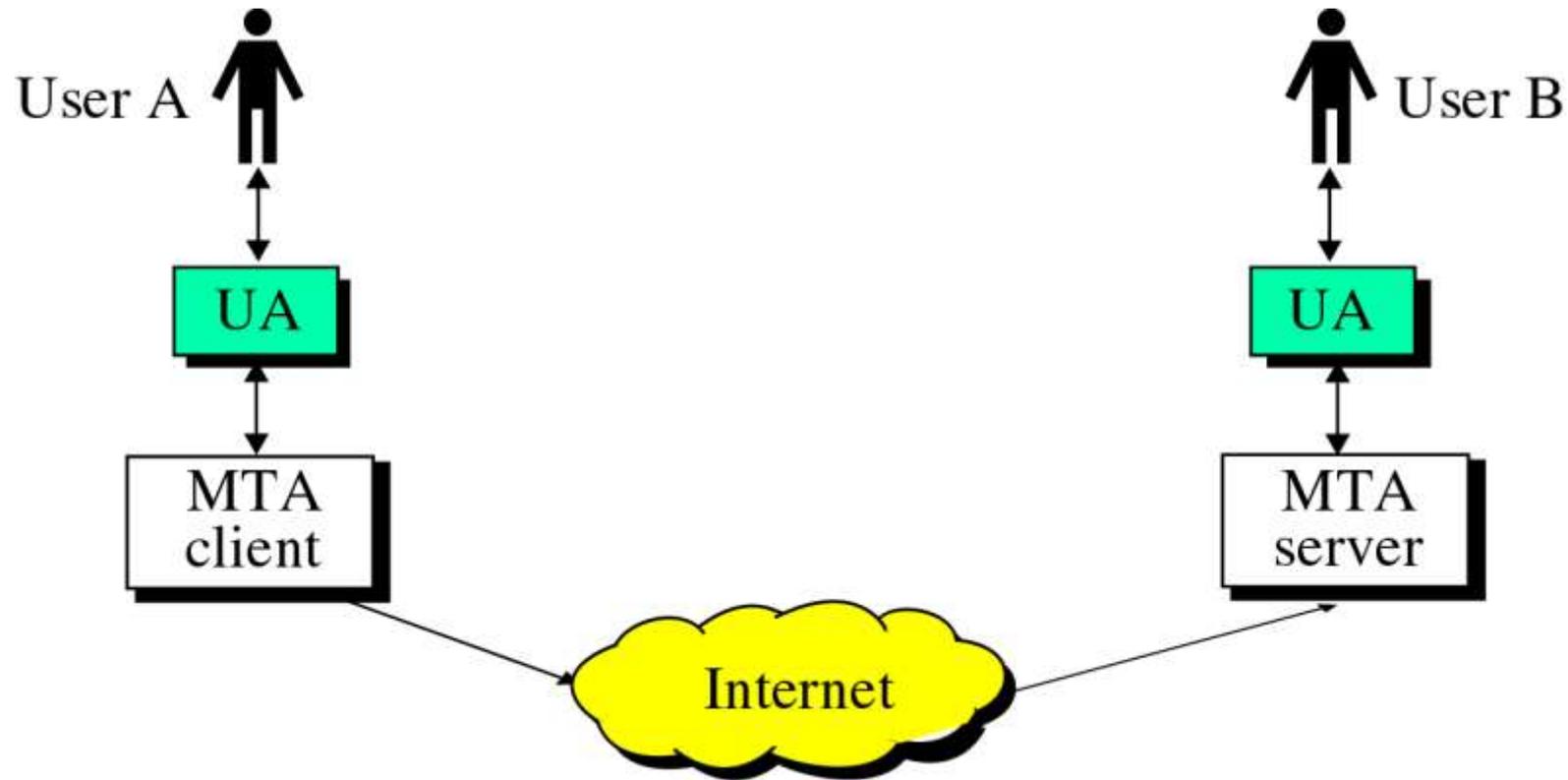
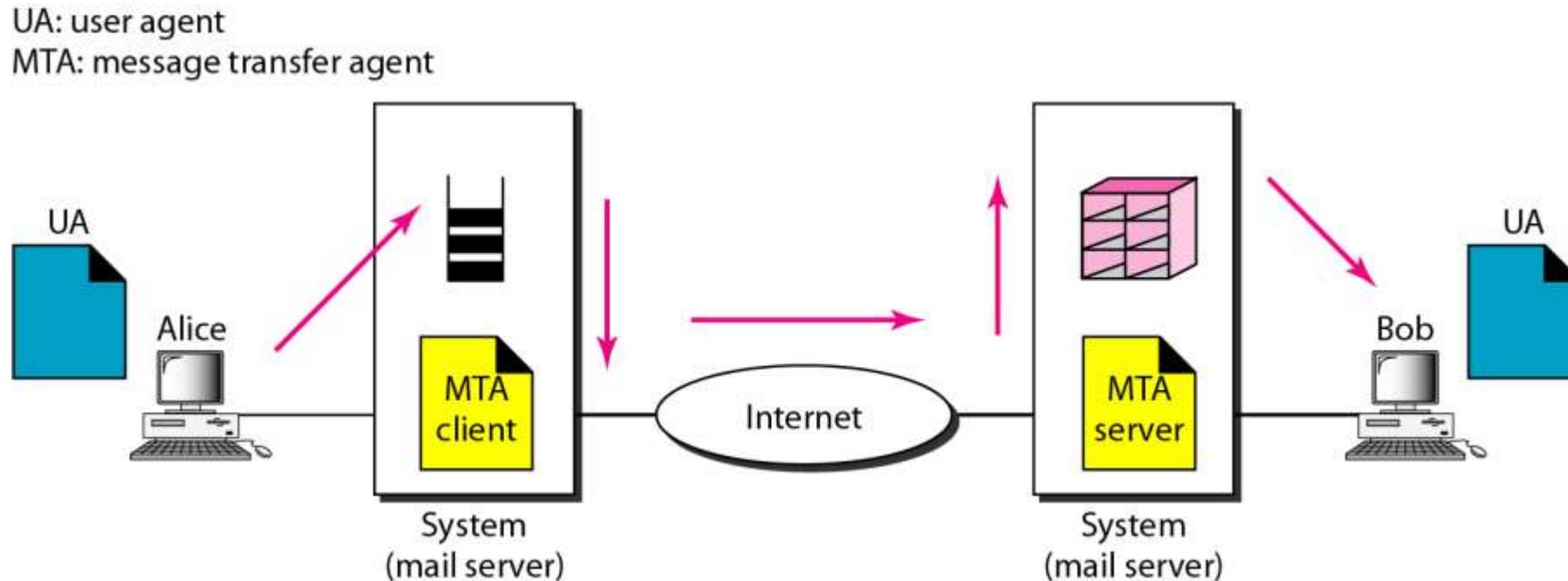


Figure 26.7 *Second scenario in electronic mail*



When the sender and the receiver of an e-mail are on different systems, we need two UAs and a pair of MTAs (client and server).

Figure 25-13

Relay MTAs

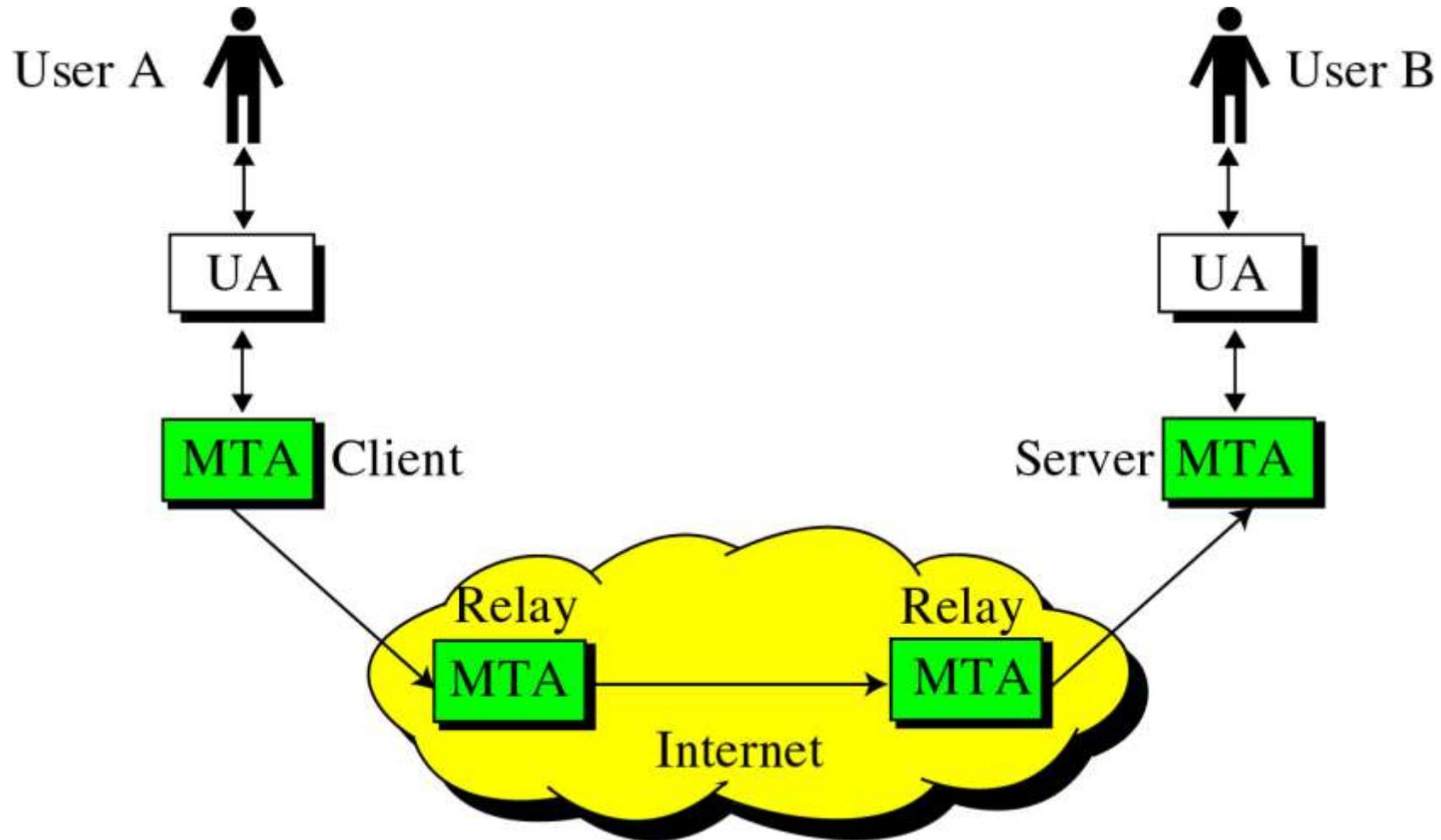


Figure 26.8 *Third scenario in electronic mail*

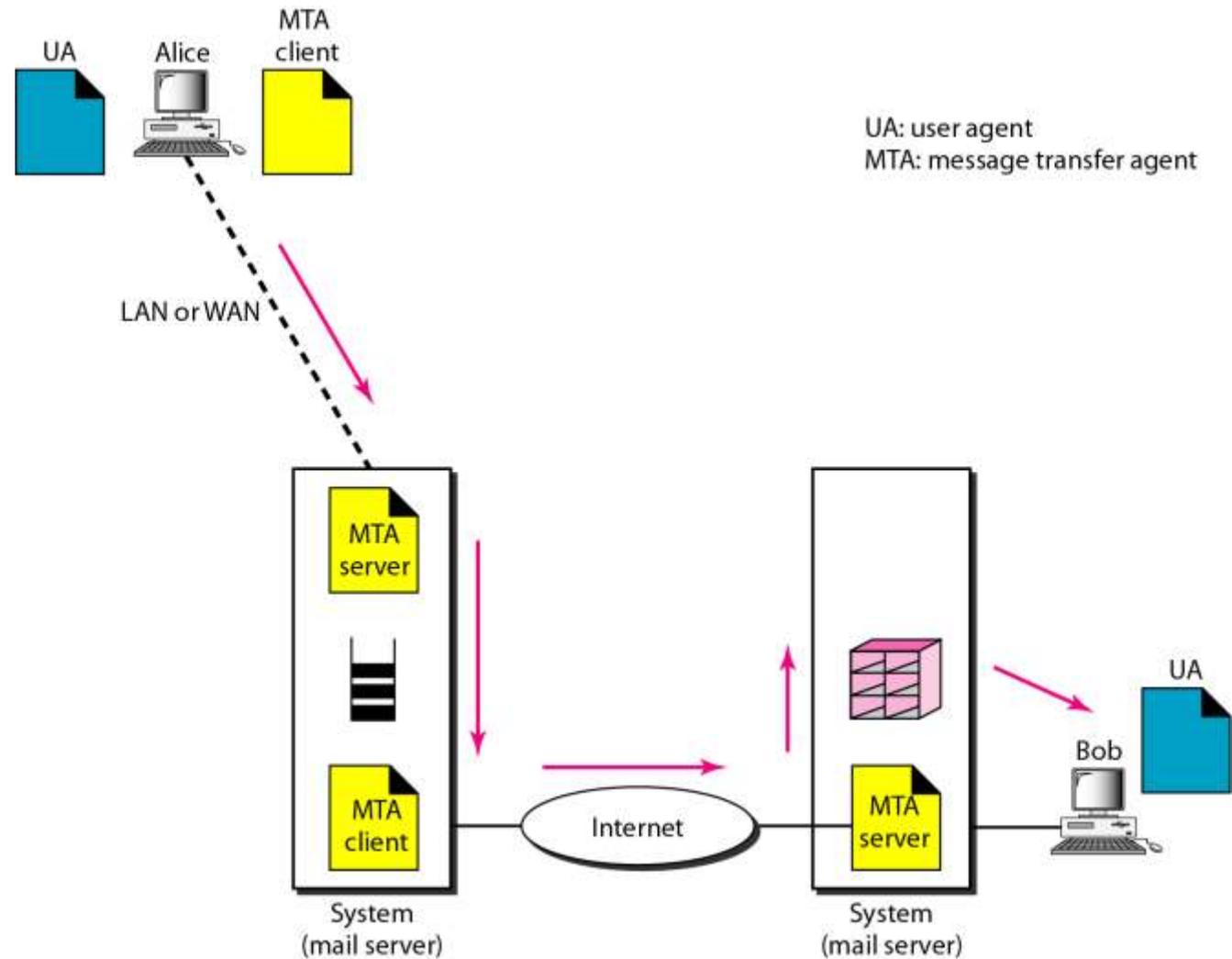
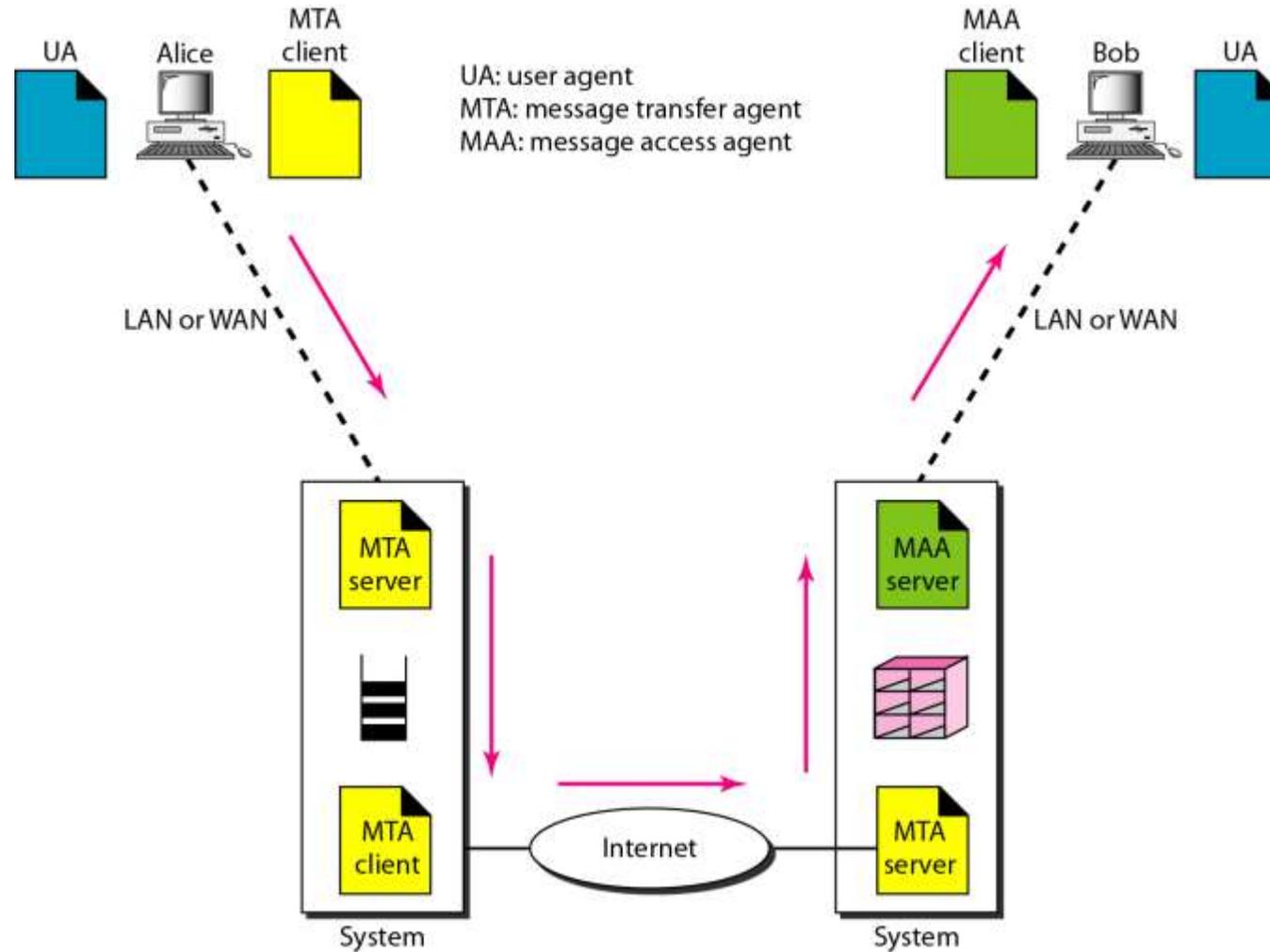


Figure 26.9 *Fourth scenario in electronic mail*

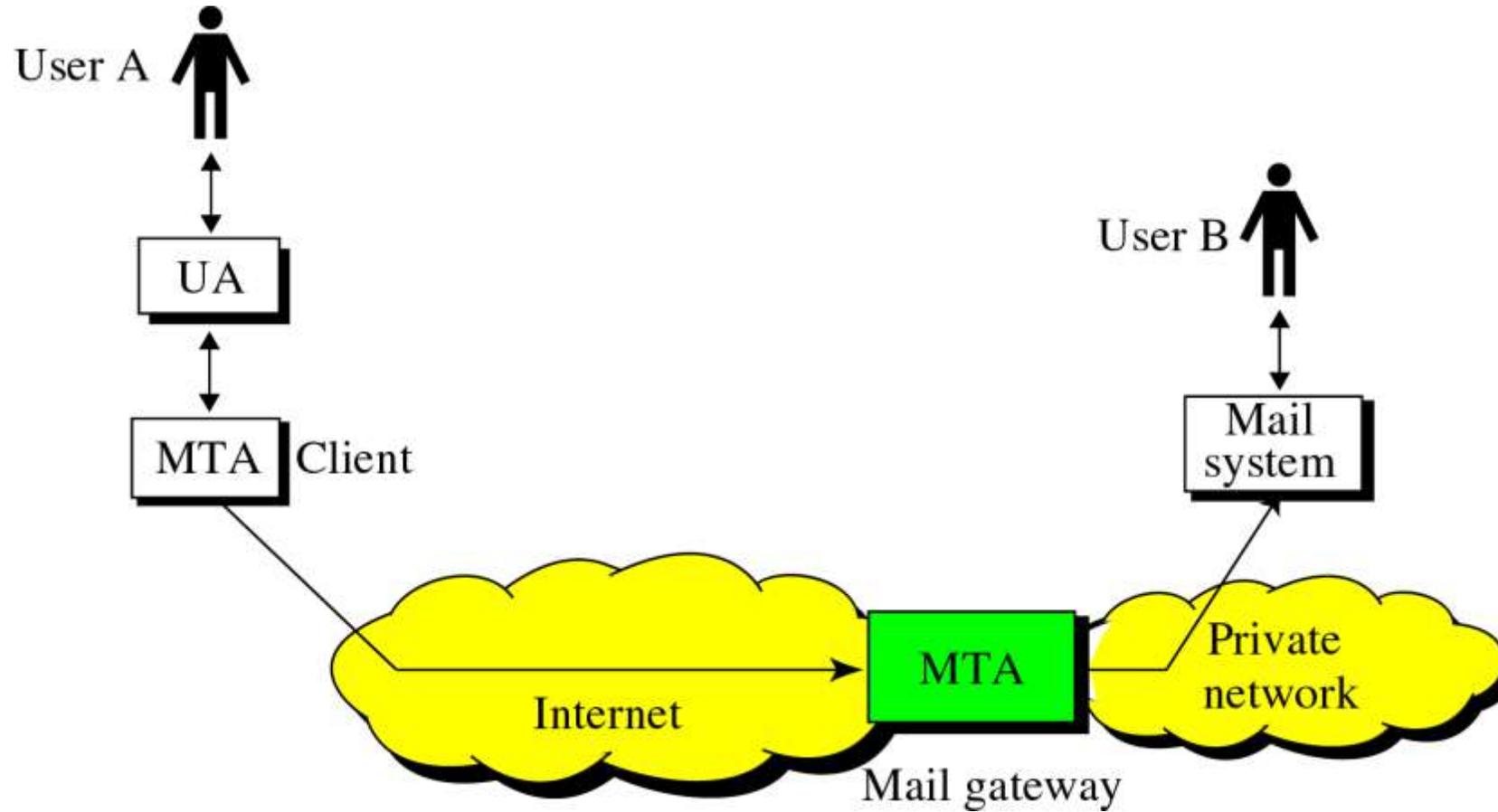


When both sender and receiver are connected to the mail server via a LAN or a WAN, we need two UAs, two pairs of MTAs and a pair of MAAs.

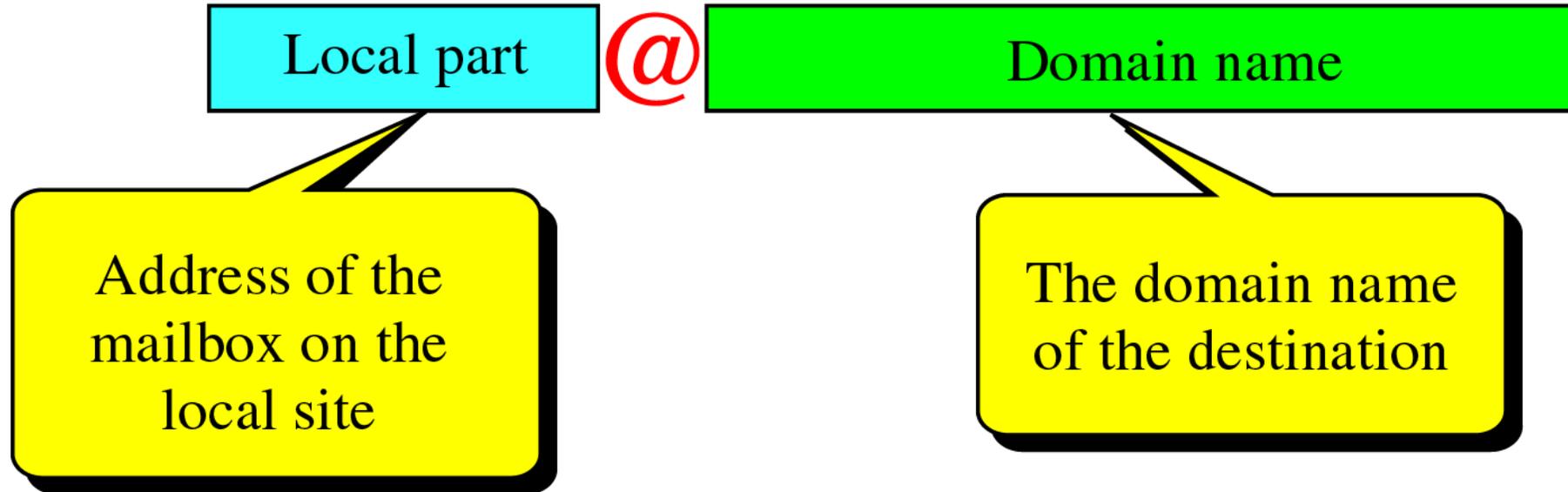
This is the most common situation today.

Figure 25-14

Mail Gateways



E-mail Address



HOW DOES IT WORK ?

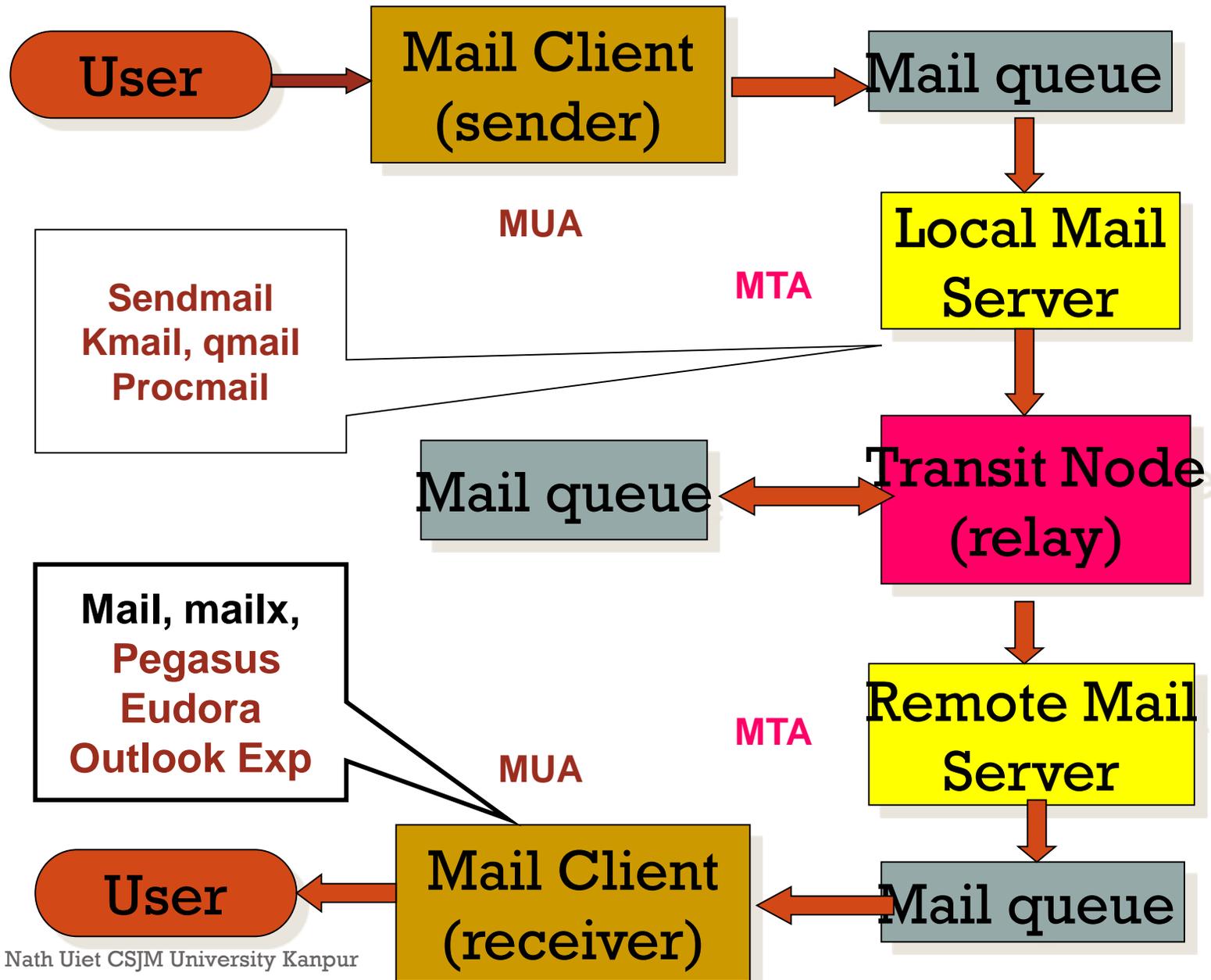
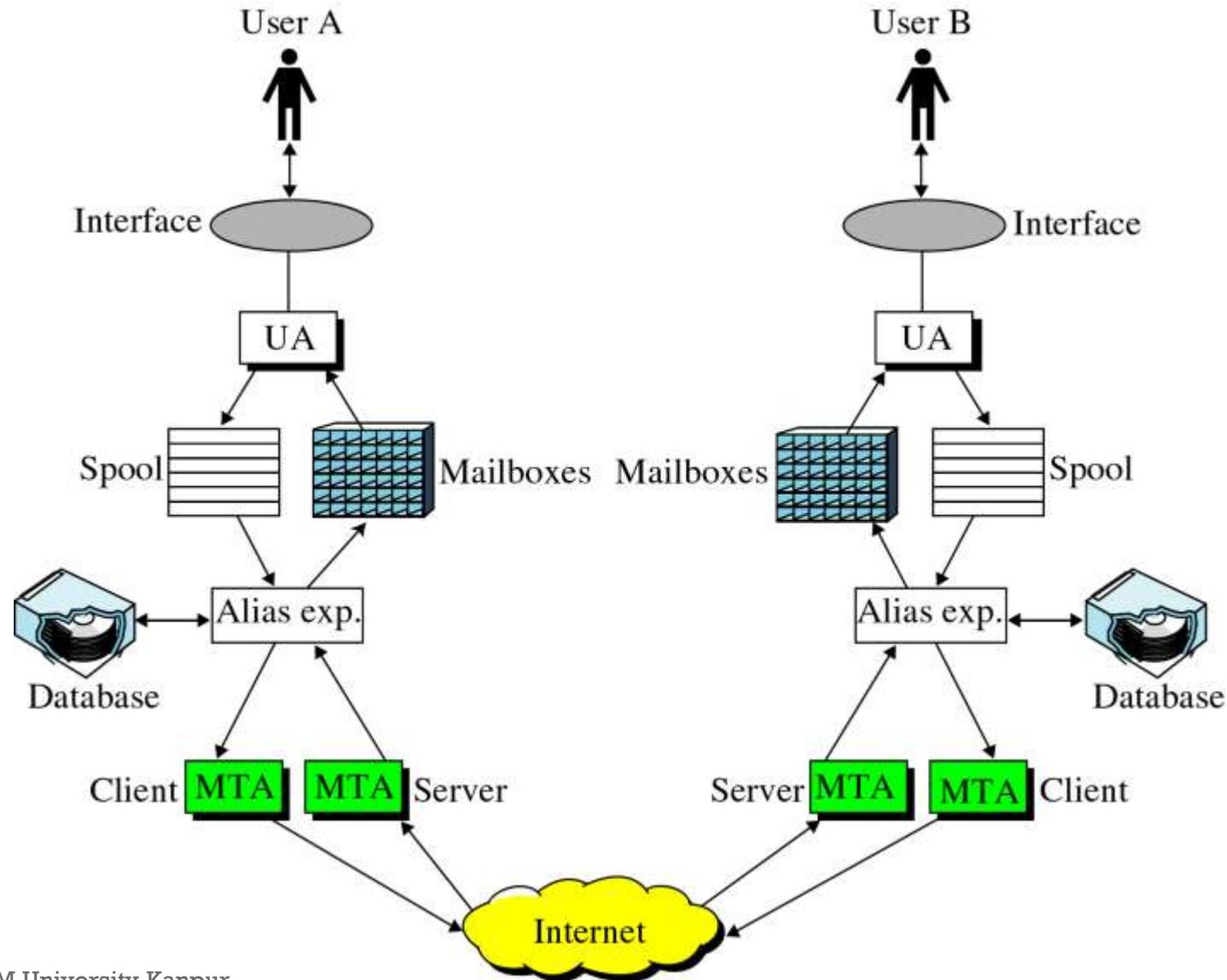


Figure 25-16

The Entire E-mail System



Header	Meaning
To:	Email address(es) of primary recipient(s)
Cc:	Email address(es) of secondary recipient(s)
Bcc:	Email address(es) for blind carbon copies
From:	Person or people who created the message
Sender:	Email address of the actual sender
Received:	Line added by each transfer agent along the route
Return-Path:	Can be used to identify a path back to the sender

X-Charset
X-Mailer
X-Sender
X-Face

Date:	The date and time the message was sent
Reply-To:	Email address to which replies should be sent
Message-Id:	Unique number for referencing this message later
In-Reply-To:	Message-Id of the message to which this is a reply
References:	Other relevant Message-Ids
Keywords:	User chosen keywords
Subject:	Short summary of the message for the one-line display

PROTOCOLS

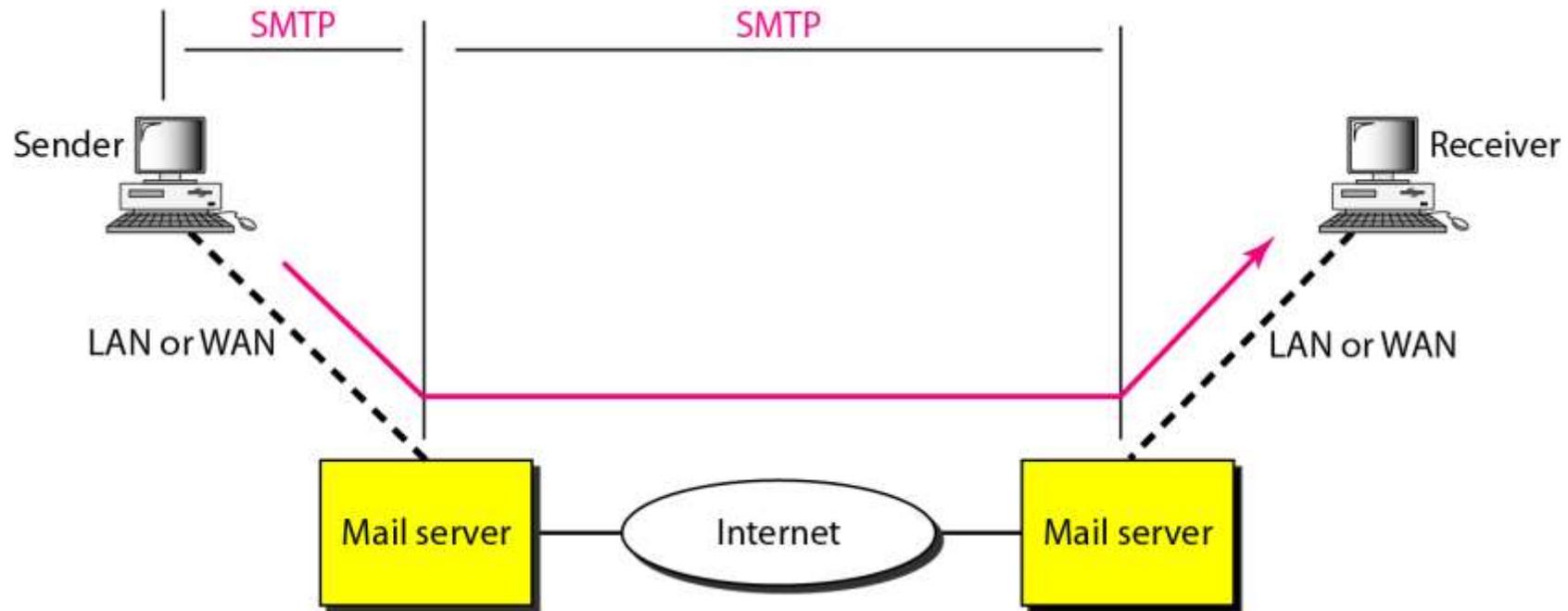
SMTP: Simple Mail Transfer Protocol

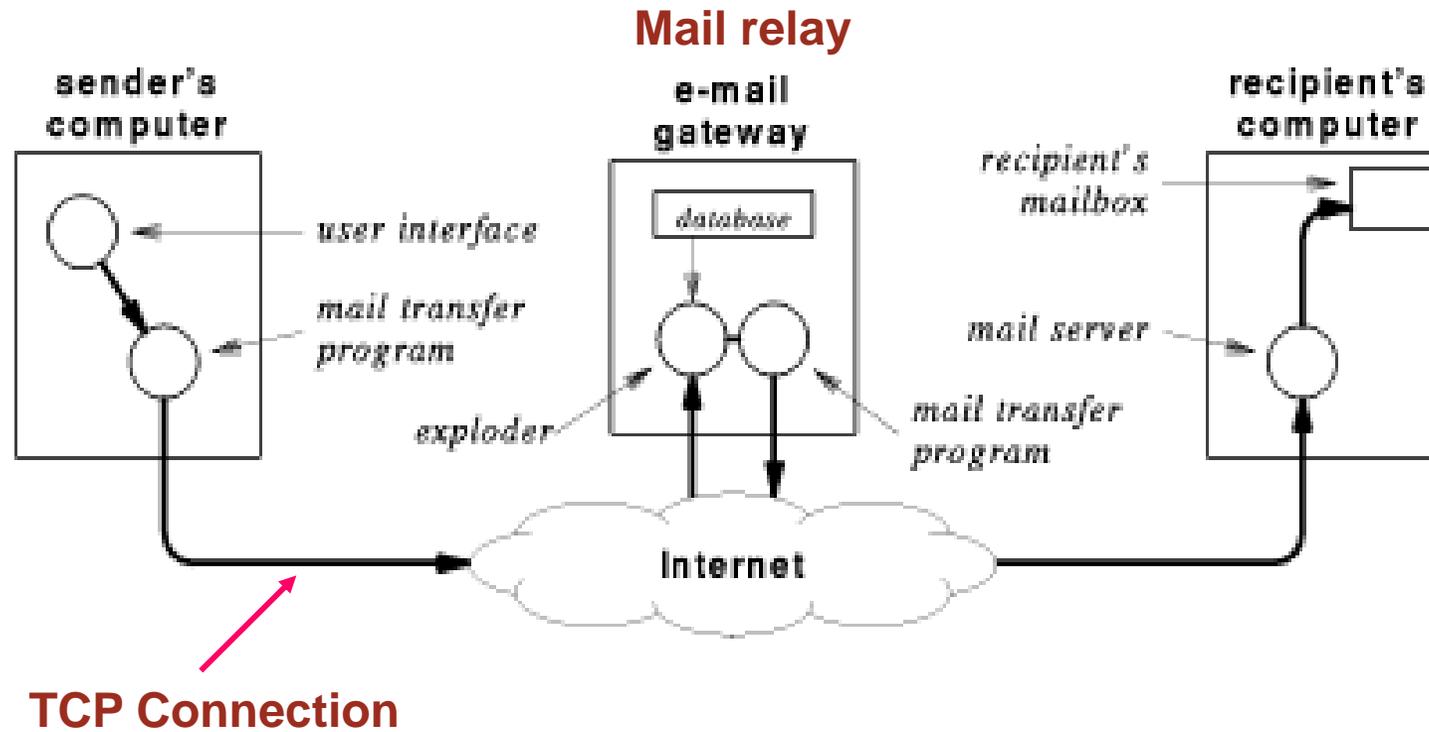
- Used by clients to send email to a server.
- Used by a server to forward mail to another server.
- Commands to specify the sender, recipients, and message.
- Text based interface, causing security problems.
- Allows only ASCII messages.

ESMTP: Extended Simple Mail Transfer Protocol

- Allows 8-bit data as part of message (but not widely implemented).
- Supports notification on receipt.

Figure 26.16 *SMTP range*





Exploder : forwards mail to mailing list

PROTOCOLS (CONTD.)

- **POP: Post Office Protocol**
 - To copy mail from server to the local disk.
 - Authentication using userid and password.
- **IMAP Interactive Mail Access Protocol**
 - Server maintains mailbox and folders.
 - Can be accessed from anywhere.
 - Email repository can be looked upon as a relational database with commands to search and retrieve, etc.

Figure 26.19 **POP3 and IMAP4**

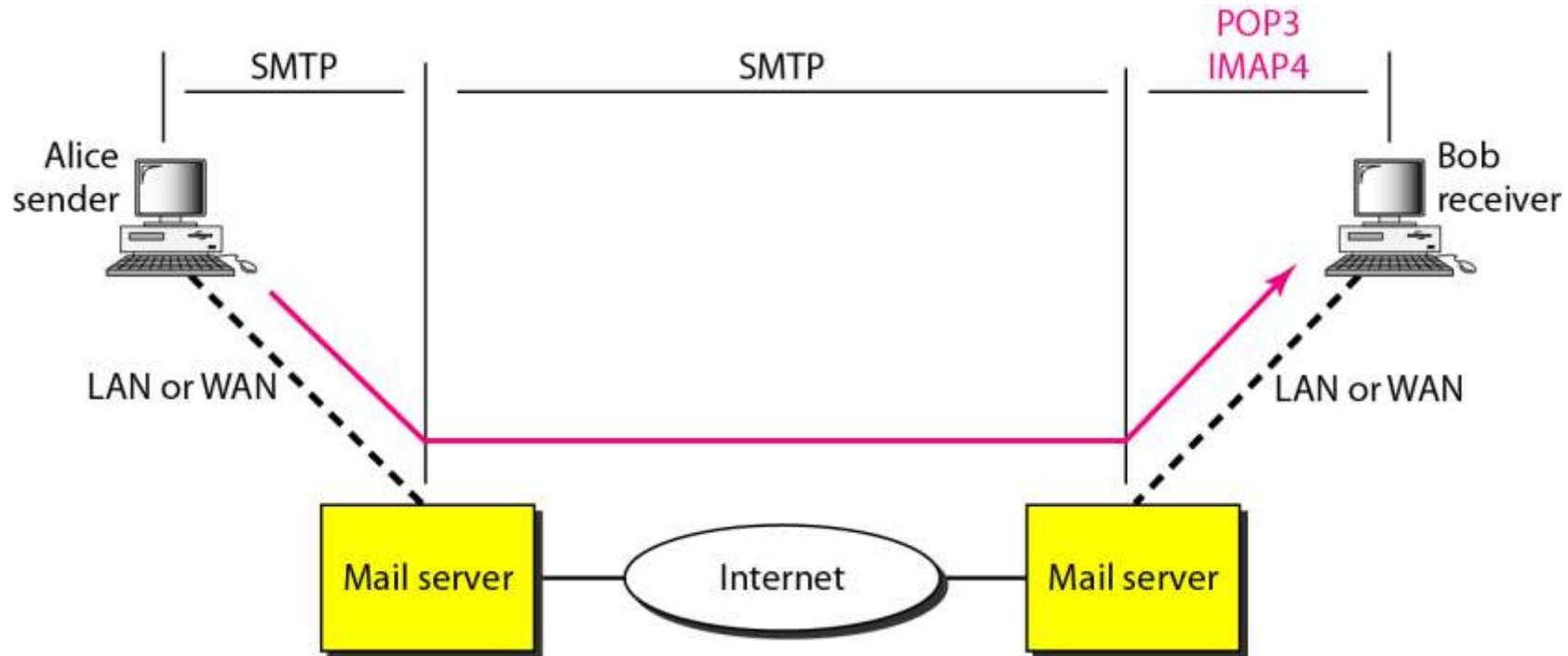
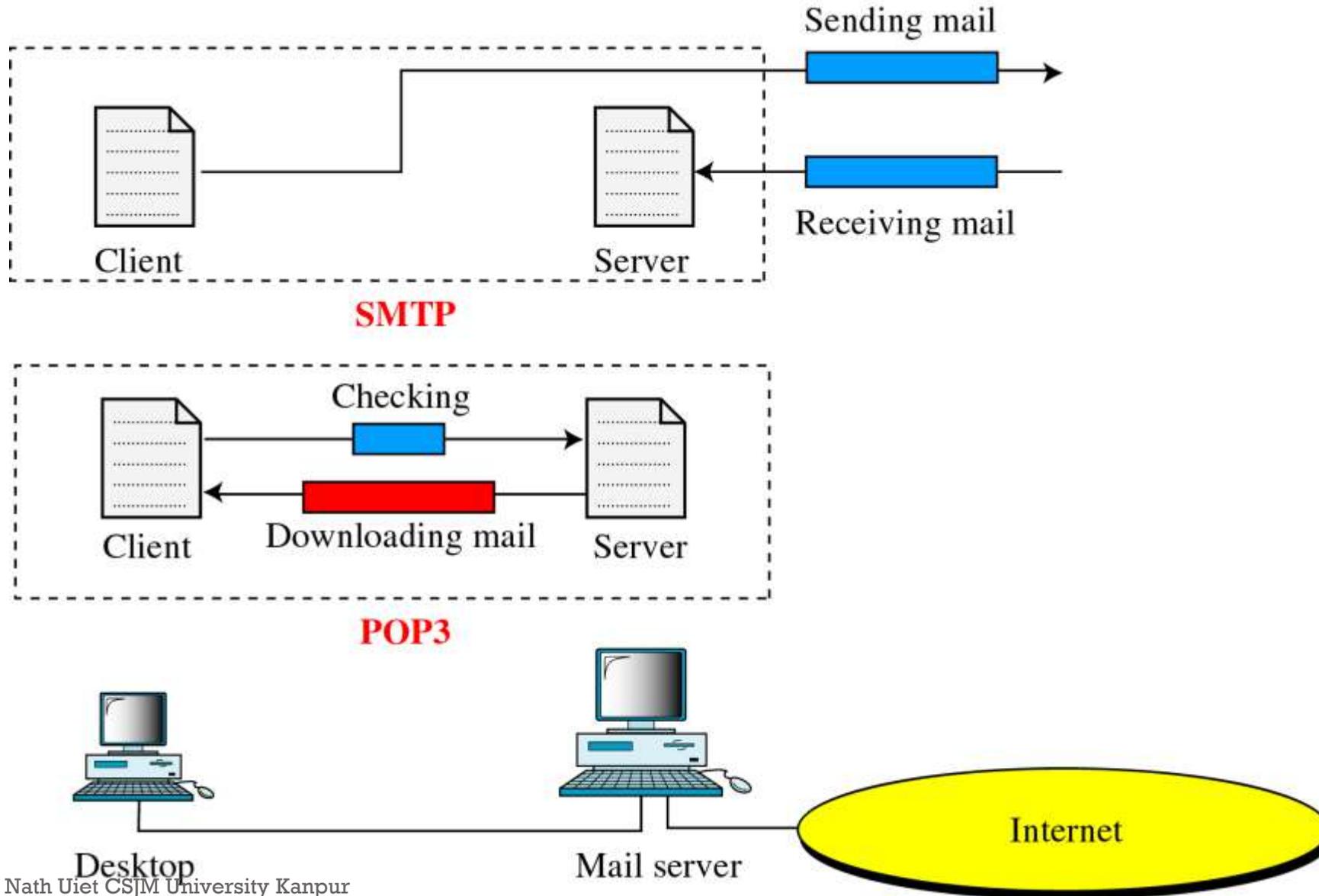
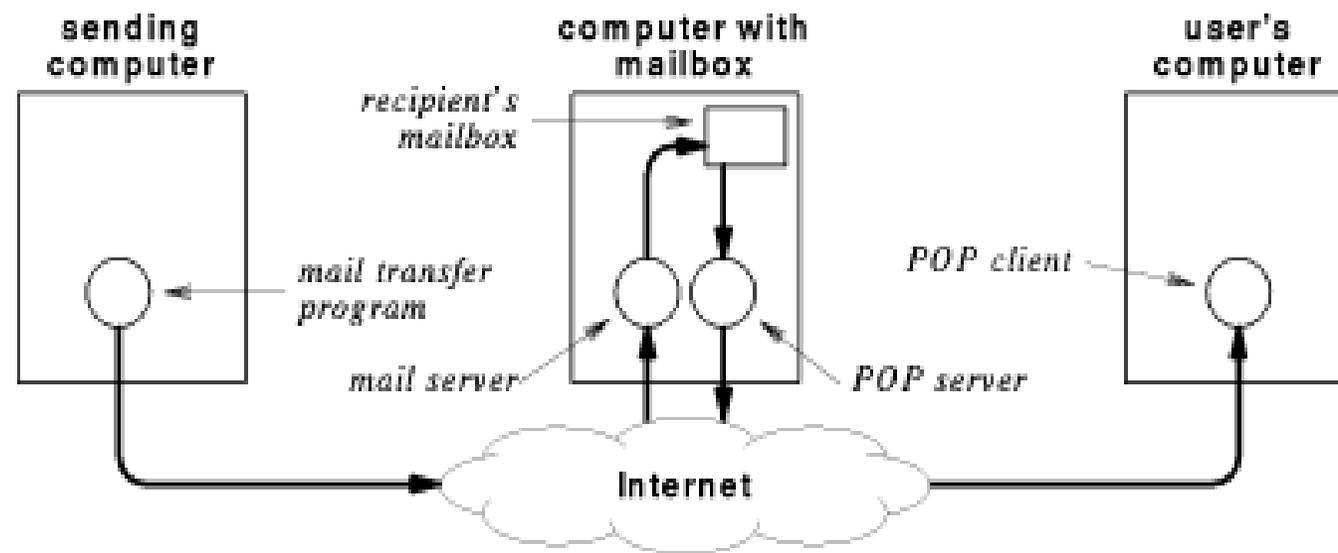


Figure 25-18

POP3 and SMTP





SENDING BINARY FILES

- Examples: data files, audio, video, languages other than English.
- Mail protocol (SMTP) allows only ASCII files to be sent in e-mail
- Encode the binary file into ASCII form:
uuencode <parameters>
- Receiver does the decoding:
uudecode
- If the mail client is MIME-compatible, then one can simply ``attach" a binary file to the mail.

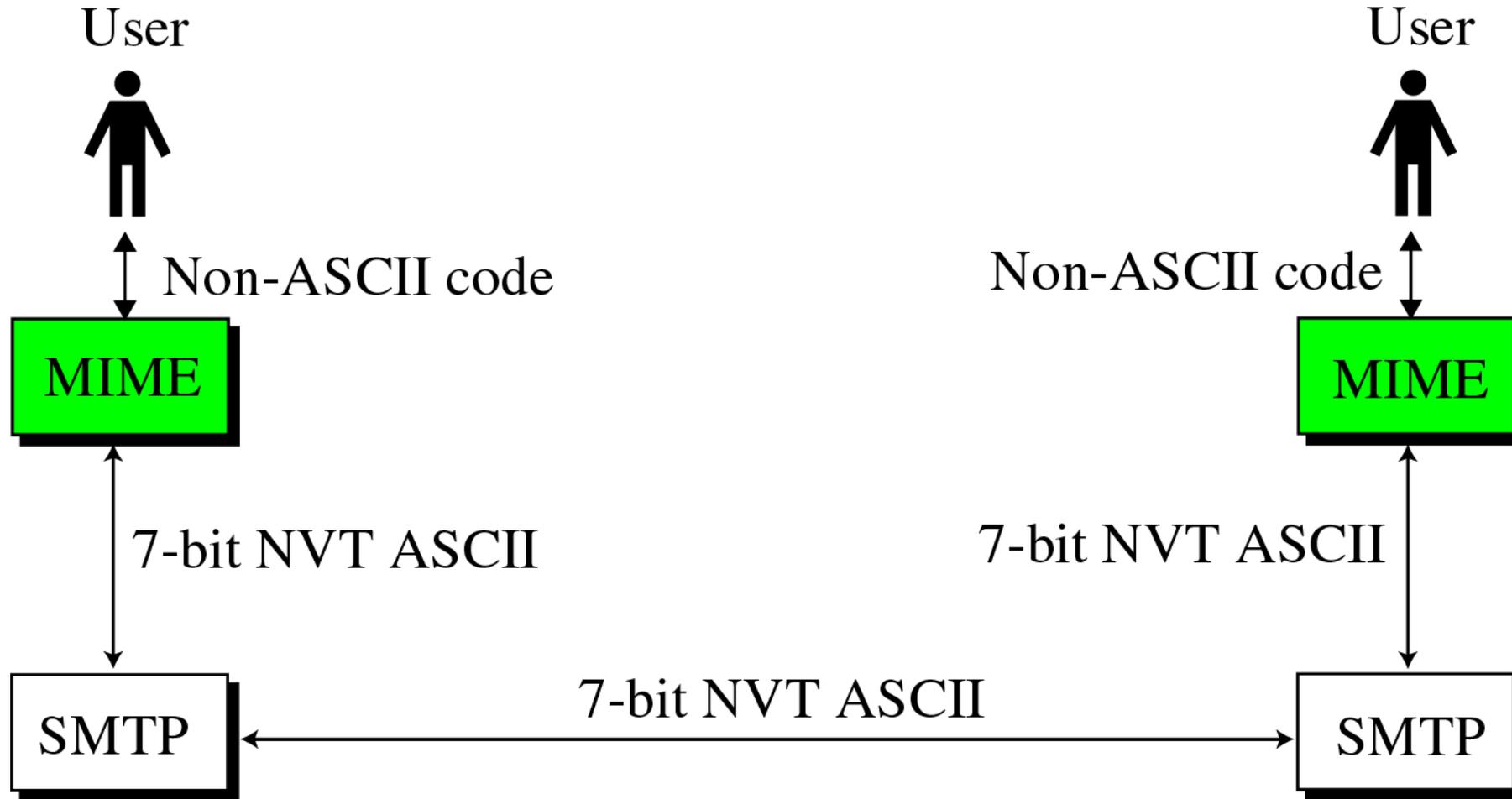
MIME

Multipurpose Internet Mail Extension.

- To allow attaching of all kinds of documents to regular e-mail, including multimedia documents.
- Content-Type in the header specifies the type of file (e.g., audio or gif).
- Content-Transfer-Encoding specifies how the binary file has been converted to an ASCII file.
- There can be multiple parts, each with its own header.
- MIME does not require any extensions to SMTP such as, ESMTP or use of non-ASCII characters.
- For reading MIME mail, the client software should be MIME-compliant, no change is required in any of the mail servers (like **sendmail**).

Figure 25-17

MIME



Header	Meaning
MIME-Version:	Identifies the MIME version
Content-Description:	Human-readable string telling what is in the message
Content-Id:	Unique identifier
Content-Transfer-Encoding:	How the body is wrapped for transmission
Content-Type:	Nature of the message

. RFC 822 headers added by MIME.

Type	Subtype	Description
Text	Plain	Unformatted text
	Richtext	Text including simple formatting commands
Image	Gif	Still picture in GIF format
	Jpeg	Still picture in JPEG format
Audio	Basic	Audible sound
Video	Mpeg	Movie in MPEG format
Application	Octet-stream	An uninterpreted byte sequence
	Postscript	A printable document in PostScript
Message	Rfc822	A MIME RFC 822 message
	Partial	Message has been split for transmission
	External-body	Message itself must be fetched over the net
Multipart	Mixed	Independent parts in the specified order
	Alternative	Same message in different formats
	Parallel	Parts must be viewed simultaneously
	Digest	Each part is a complete RFC 822 message

The MIME types and subtypes defined in RFC 1521.

From: elinor@abc.com
To: carolyn@xyz.com
MIME-Version: 1.0
Message-Id: <0704760941.AA00747@abc.com>
Content-Type: multipart/alternative; boundary=qwertyuiopasdfghjklzxc
Subject: Earth orbits sun integral number of times

This is the preamble. The user agent ignores it. Have a nice day.

--qwertyuiopasdfghjklzxcvbnm

Content-Type: text/richtext

Happy birthday to you
Happy birthday to you
Happy birthday dear <bold> Carolyn </bold>
Happy birthday to you

--qwertyuiopasdfghjklzxcvbnm

Content-Type: message/external-body;
access-type="anon-ftp";
site="bicycle.abc.com";
directory="pub";
name="birthday.snd"

content-type: audio/basic
content-transfer-encoding: base64
--qwertyuiopasdfghjklzxcvbnm--

```

S: 220 xyz.com SMTP service ready
C: HELO abc.com
S: 250 xyz.com says hello to abc.com
C: MAIL FROM: <elinor@abc.com>
S: 250 sender ok
C: RCPT TO: <carolyn@xyz.com>
S: 250 recipient ok
C: DATA
S: 354 Send mail; end with "." on a line by itself
C: From: elinor@abc.com
C: To: carolyn@xyz.com
C: MIME-Version: 1.0
C: Message-Id: <0704760941.AA00747@abc.com>
C: Content-Type: multipart/alternative; boundary=qwertyuiopasdfghjklzxcvbnm
C: Subject: Earth orbits sun integral number of times
C:
C: This is the preamble. The user agent ignores it. Have a nice day.
C:
C: --qwertyuiopasdfghjklzxcvbnm
C: Content-Type: text/richtext
C:
C: Happy birthday to you
C: Happy birthday to you
C: Happy birthday dear <bold> Carolyn </bold>
C: Happy birthday to you
C:
C: --qwertyuiopasdfghjklzxcvbnm
C: Content-Type: message/external-body;
C:     access-type="anon-ftp";
C:     site="bicycle.abc.com";
C:     directory="pub";
C:     name="birthday.snd"
C:
C: content-type: audio/basic
C: content-transfer-encoding: base64
C: --qwertyuiopasdfghjklzxcvbnm
C: .
S: 250 message accepted
C: QUIT
S: 221 xyz.com closing connection

```

Fig. 7-47. Transferring a message from *elinor@abc.com* to *carolyn@xyz.com*.

MAILING LISTS

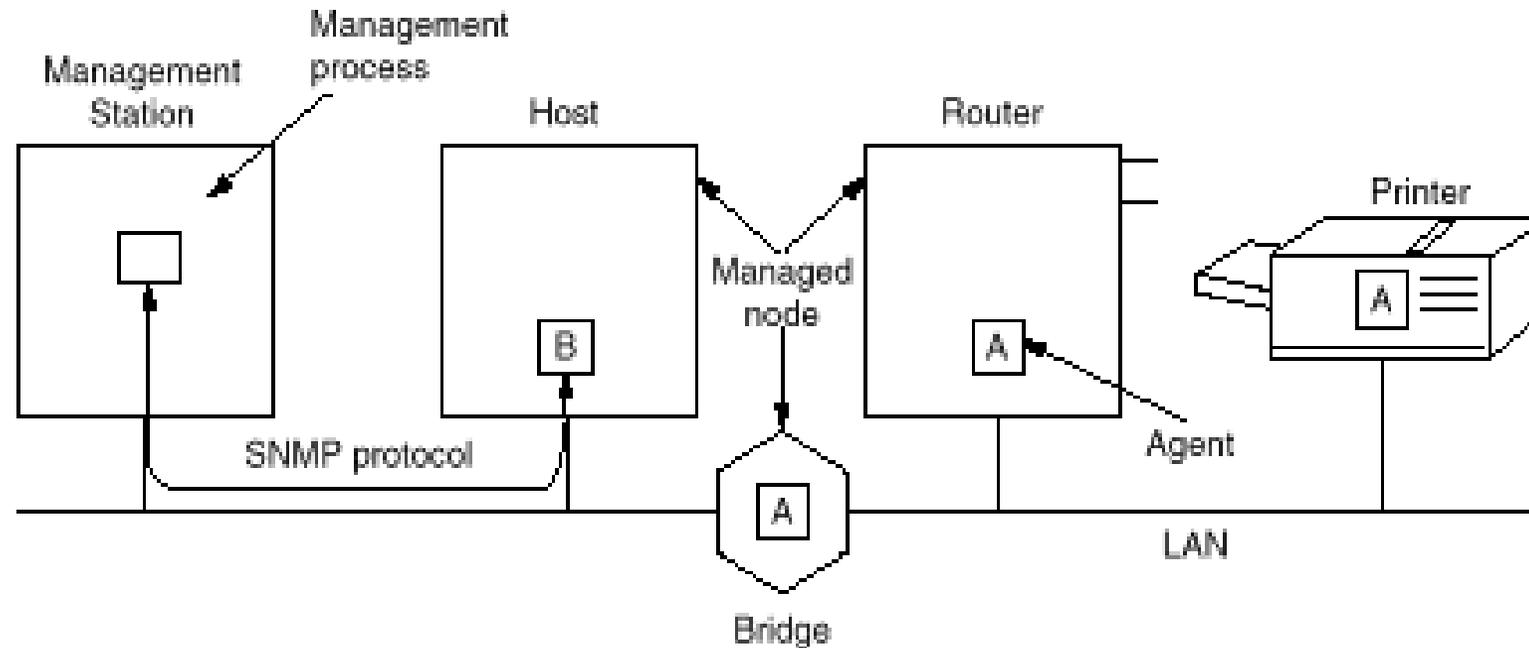
- Means to send the same message to a number of people.
- Can be managed with software called list servers.
- List servers permit:
 - automatic maintenance of list - addition, deletion, changes.
 - creating a digest of messages.
 - maintaining an archive of messages.
 - different types of access control: public, private, moderated.

EMAIL PRIVACY

- **PGP:** Pretty Good Privacy
- **PEM** Privacy Enhanced Mail
- Both provide confidentiality through symmetric key encryption.
- Symmetric key is exchanged using public keys.
- Both support authentication and integrity through digital signatures.

The Network Management Model (SNMP) Components

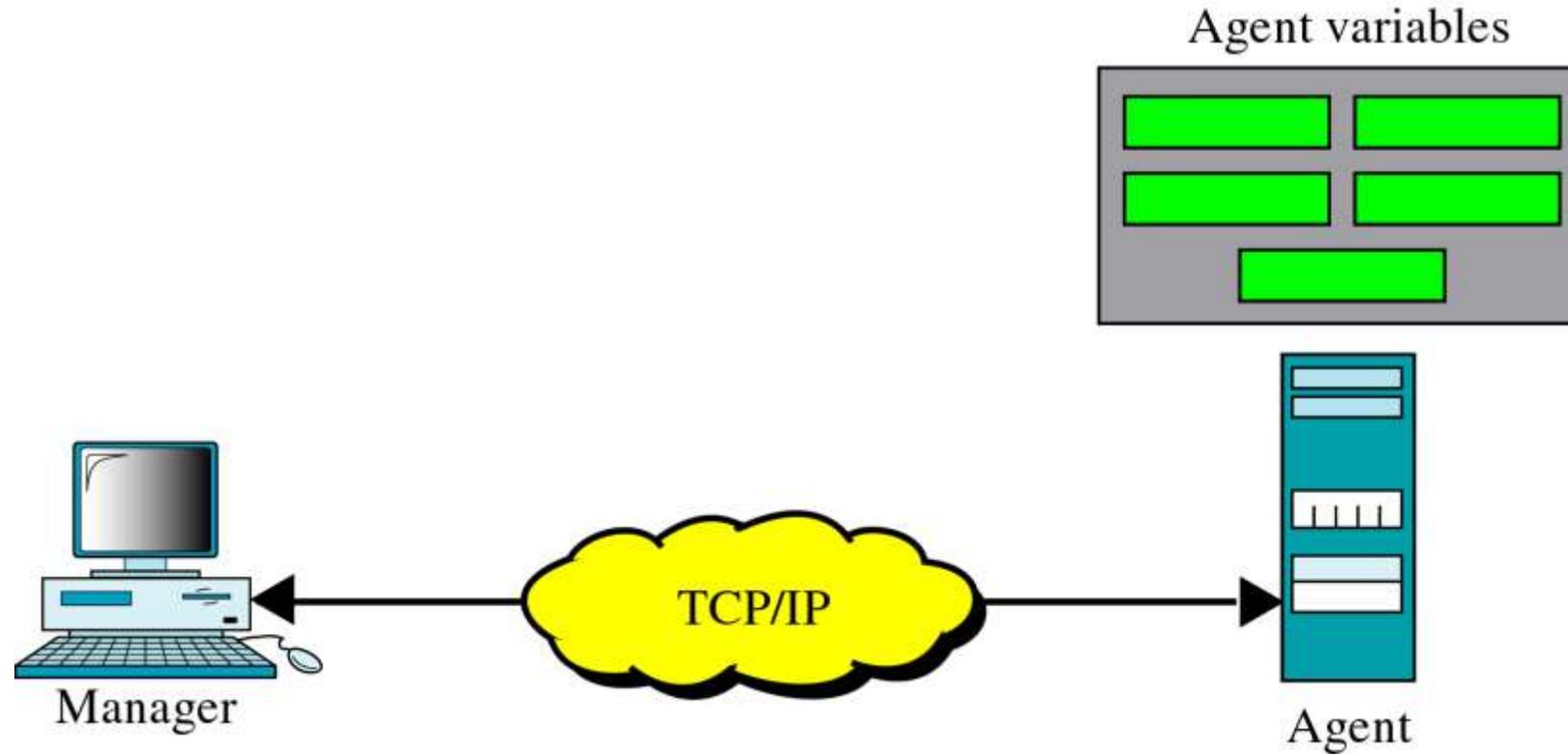
1. Managed Nodes
2. Management Stations
3. Management Information
4. Management Protocol



Components of the SNMP management model.

Figure 25-19

SNMP Concept



Internet Management Components

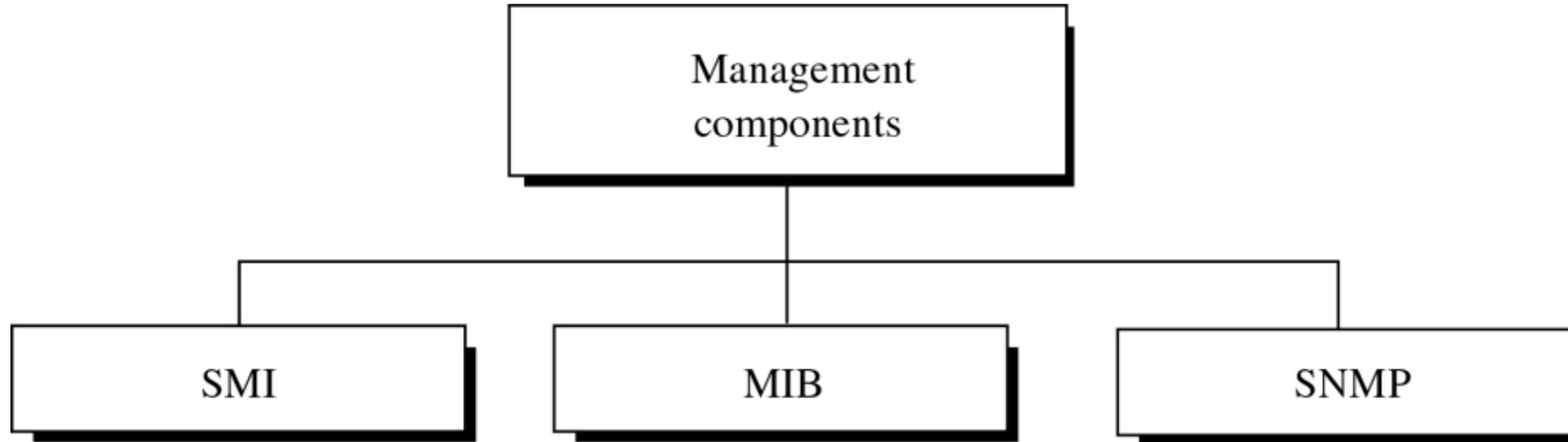


Figure 25-21

MIB

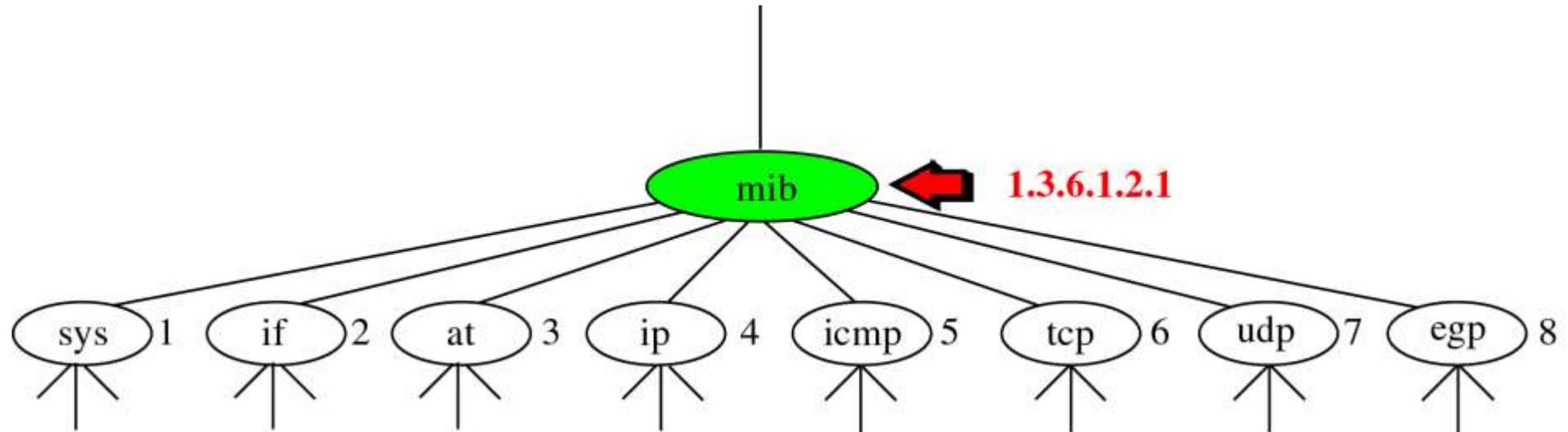
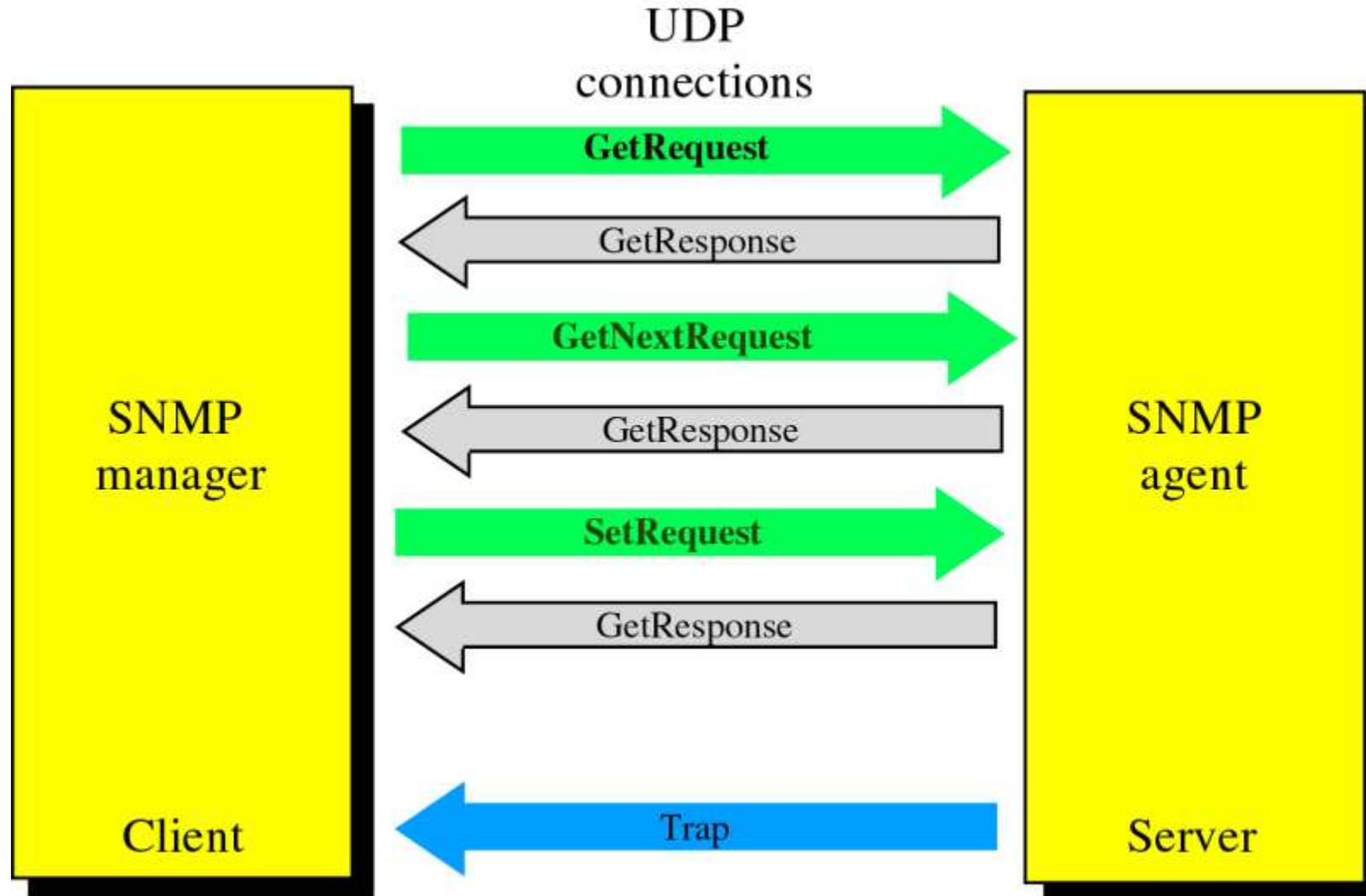


Figure 25-22

SNMP Messages



SIMPLE NETWORK MANAGEMENT PROTOCOL

- To get information about nodes in the network.
- Manage nodes by setting control variables in them.
- Advantages
 - Independent of hardware/link layer.
 - Nodes need not be physically on the same network.
- Architecture
 - Client-Server communication.
 - Database of control variables (MIB).

SNMP MANAGEMENT DATA

- **MIB: Management Information Base**
Data items a node should keep and the operations allowed on each.
- Variables are defined in a hierarchical fashion.
- A node need not keep all the defined variables.
- **SMI: Structure of Management Information**
- Set of rules used to define and identify MIB variables.
- SMI uses Abstract Syntax Notation.1 (ASN.1) for unambiguity and interoperability.

SNMP MESSAGES

- **get-request** Fetch value of the specific variable.
- **get-next-request** Fetch value of the next variable.
- **get-bulk-request** Fetch a large table.
- **get-response** Reply to a fetch operation.
- **set-request** Store the value of specific variable.
- **inform-request** Manager-to-manager message describing local MIB.
- **trap** Reply triggered by an event.

NETWORK DEBUGGING

■ Ping

- A quick way to know if a host is up.
- Can check round trip times for packets.
- Can check if packet loss is excessive.

■ Netstat

- Gives information about local host.
- Status of current connections.
- Routing table.
- Statistics.

■ Traceroute

- Can find out the path taken to a particular destination.