

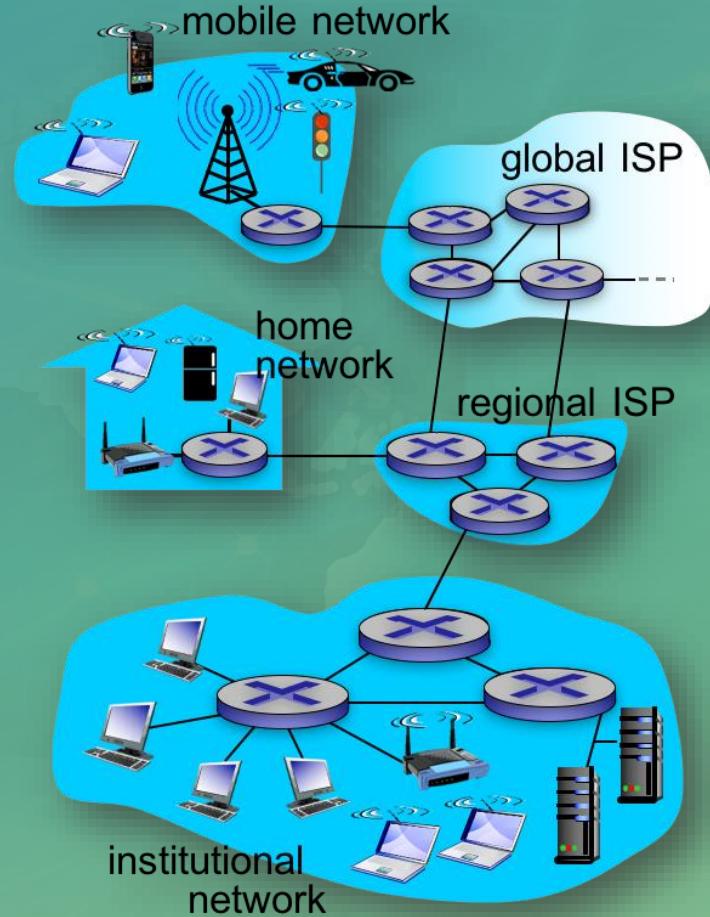


IT302 INTERNET & WEB TECHNOLOGY

INTRODUCTION

- The Internet—a global internetwork of computers.
- Internetwork (internet) means multiple networks connected together, using a standard protocol suite.
- The standard protocol suite is TCP/IP for Internet.
- Internet refers to the collection of hosts around the world that can communicate with each other using TCP/IP.
- There are thousands of networks, tens of thousands of computers, and billions of users on the Internet.

INTRODUCTION



EVOLUTION OF THE INTERNET

- The Internet was made possible by the convergence of computing and communications technologies.
- The ARPANET: 1967–1972

- ✓ ARPANET (Advanced Research Projects Agency Network) .
- ✓ Based on the packet switching technology.
- ✓ The network was owned and administered by a single entity – ARPA.

EVOLUTION OF THE INTERNET

- The Internet was made possible by the convergence of computing and communications technologies.
- The ARPANET: 1967–1972
- The TCP/IP-based ‘internetwork’: 1973–1983

- ✓ The ‘internetworking’ project began in late 1973.
- ✓ TCP/IP became the basis of the new ‘network of networks’.
- ✓ By the middle of 1983, every ARPANET host was running TCP/IP.
- ✓ 1983 marks the beginning of the Internet .

EVOLUTION OF THE INTERNET

- The Internet was made possible by the convergence of computing and communications technologies.
- The ARPANET: 1967–1972
- The TCP/IP-based ‘internetwork’: 1973–1983
- Transition from a military/research network to a ‘civilian’ one: 1983–1995

- ✓ Internet was split into ARPANET and MILNET.
- ✓ ARPA funded various operators to create TCP implementations for various operating systems.
- ✓ NSFNET (National Science Foundation Network) provided backbone.

EVOLUTION OF THE INTERNET

- The Internet was made possible by the convergence of computing and communications technologies.
- The ARPANET: 1967–1972
- The TCP/IP-based ‘internetwork’: 1973–1983
- Transition from a military/research network to a ‘civilian’ one: 1983–1995
- The commercial Internet (1995–present)

INTERNET TODAY

- The Internet is made up of many wide and local area networks joined by connecting devices and switching stations.
- It is difficult to give an accurate representation of the Internet because it is continuously changing.
- An individual or a network can be the part of Internet by getting a Internet connection.
- Internet connection is provided by Internet service providers (ISPs).

INTERNET TODAY

- There are several type of ISPs:
 - Backbone ISPs
 - ✓ Backbone ISPs are created and maintained by specialized companies.
 - ✓ These backbone networks are connected by network access points (NAPs).
 - ✓ Backbone ISPs normally operate at a high data rate.

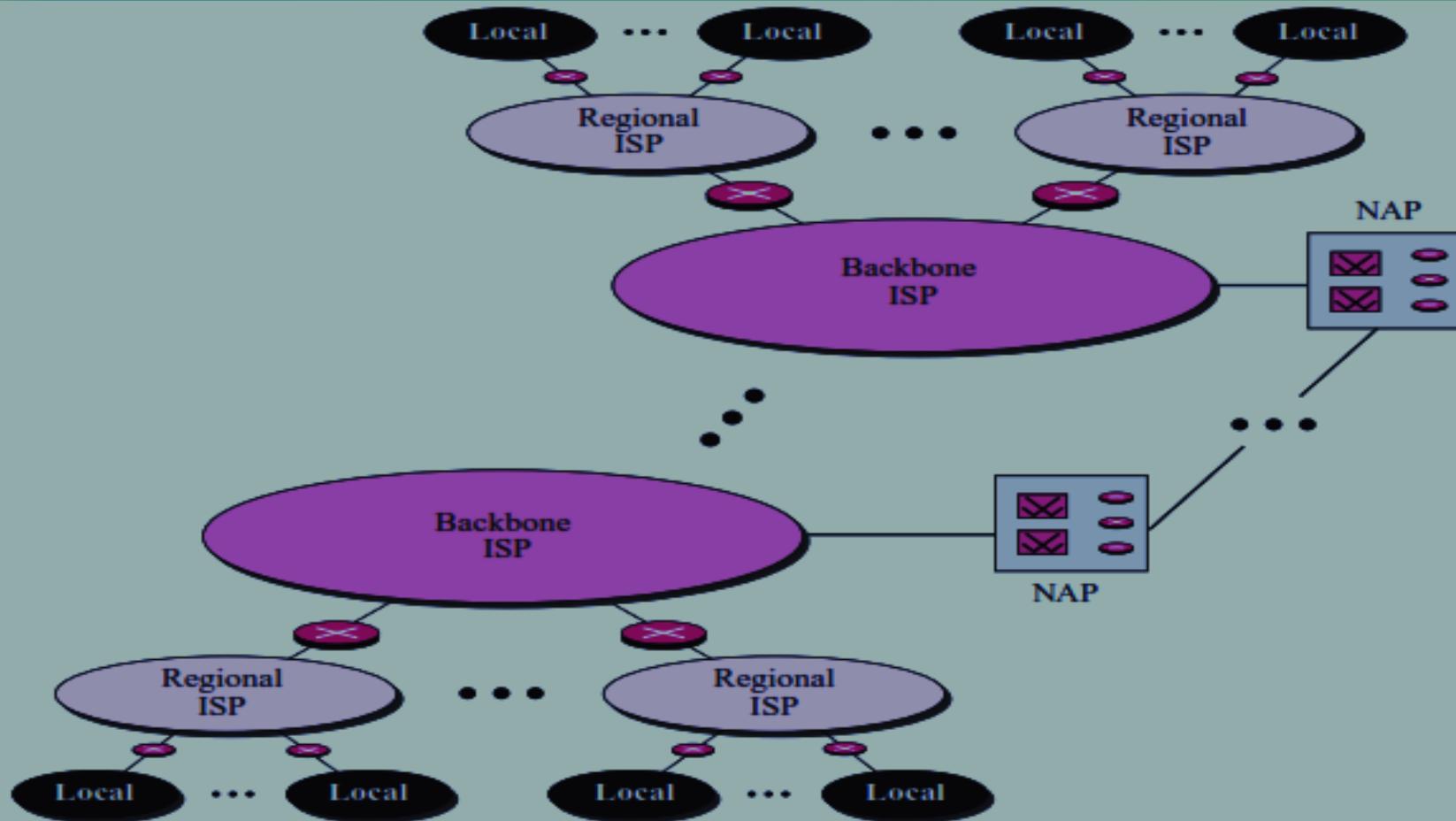
INTERNET TODAY

- There are several type of ISPs:
 - Backbone ISPs
 - Regional ISPs
 - ✓ These are connected to one or more backbone ISPs.
 - ✓ They are at the second level of hierarchy with a lesser data rate.

INTERNET TODAY

- There are several type of ISPs:
 - Backbone ISPs
 - Regional ISPs
 - Local ISPs
 - ✓ Local ISPs provide direct service to the end users.
 - ✓ The local ISPs can be connected to regional ISPs or directly to backbone ISPs.

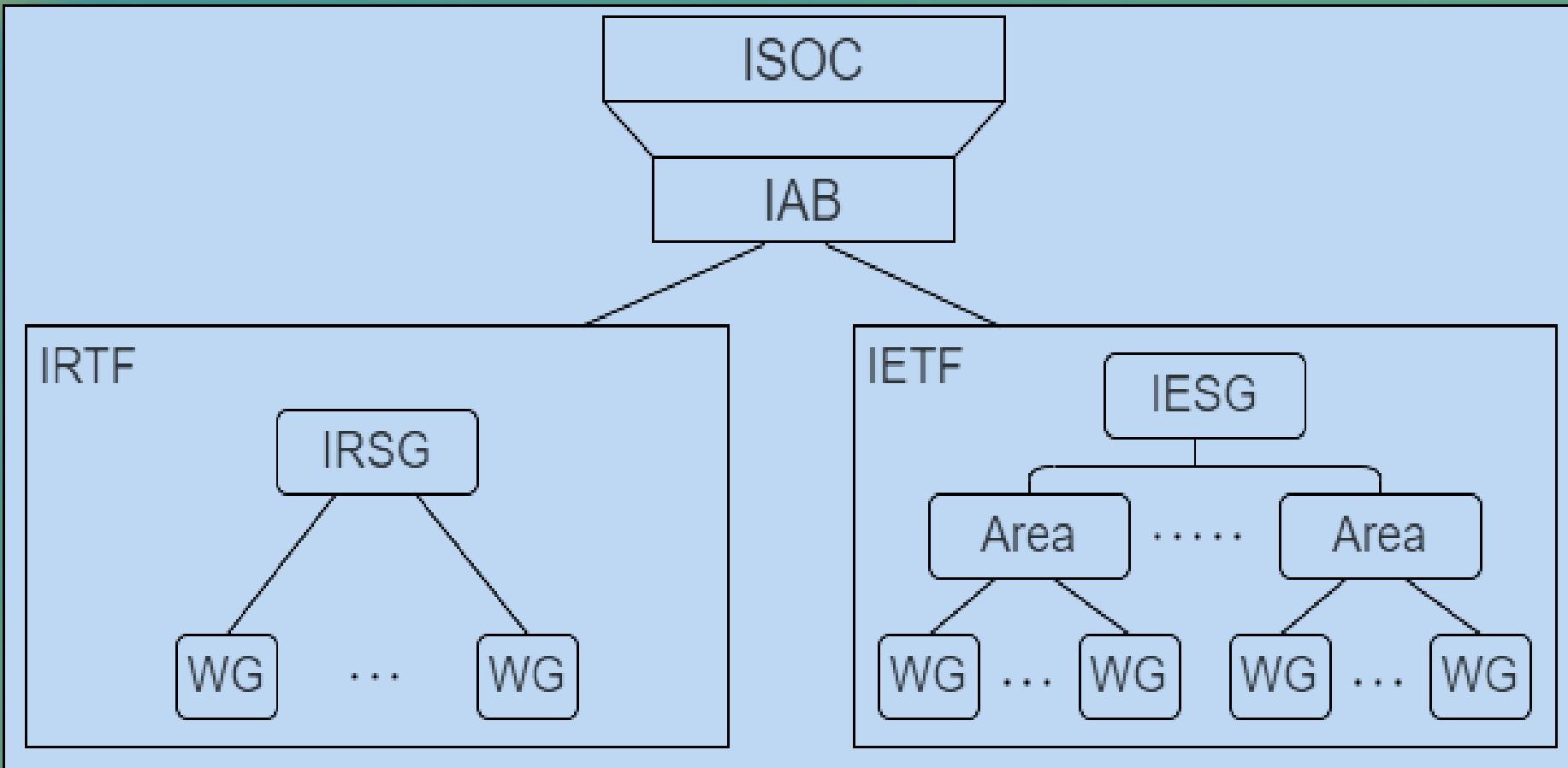
INTERNET TODAY



INTERNET ADMINISTRATION

- ❑ Nobody owns Internet.
- ❑ Internet Administration is basically a group that coordinates and guides the Internet with its growth and development.
- ❑ It makes sure that all the protocols are followed by the devices and network for the smooth functioning of the internetwork.

INTERNET ADMINISTRATION





ARCHITECTURE OF INTERNET

ARCHITECTURE OF INTERNET

- The Internet architecture is also called the TCP/IP architecture.
- A 4-layer model TCP/IP network model provides the logical architecture to Internet.
- TCP/IP is a hierarchical protocol made up of interactive modules.
- The term hierarchical means that each upper level protocol is supported by one or more lower level protocols.

THE TCP/IP PROTOCOL SUITE

Application Layer

HTTP FTP DNS SMTP Telnet

Transport Layer

Transmission Control Protocol (TCP)
User Datagram Protocol (UDP)

Network Layer

Internet Protocol (IP)

Network Interface Layer

Ethernet \ Token Ring\ Other Link Layer

Layers

Protocols

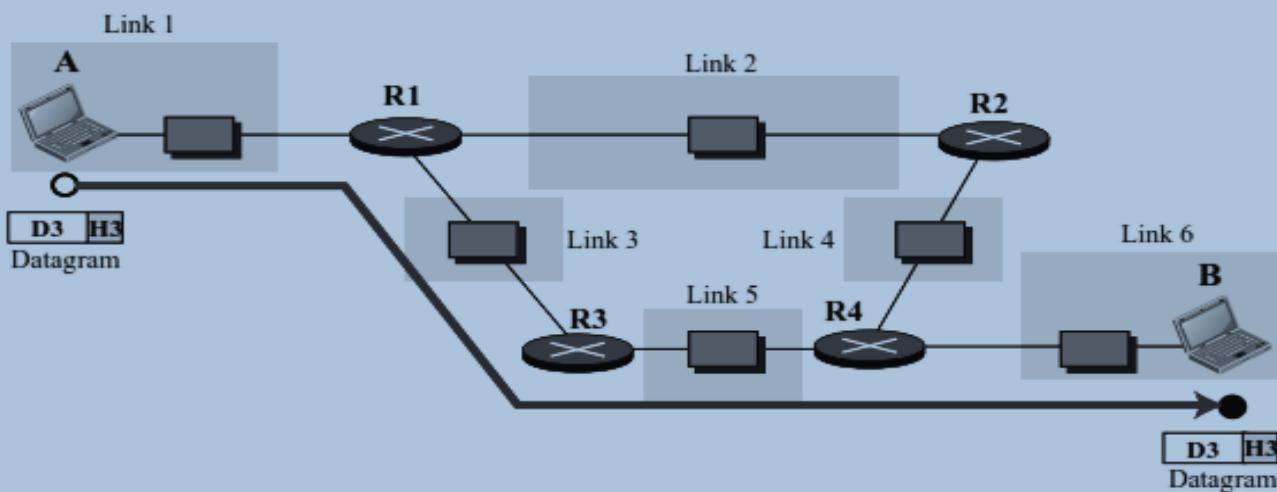
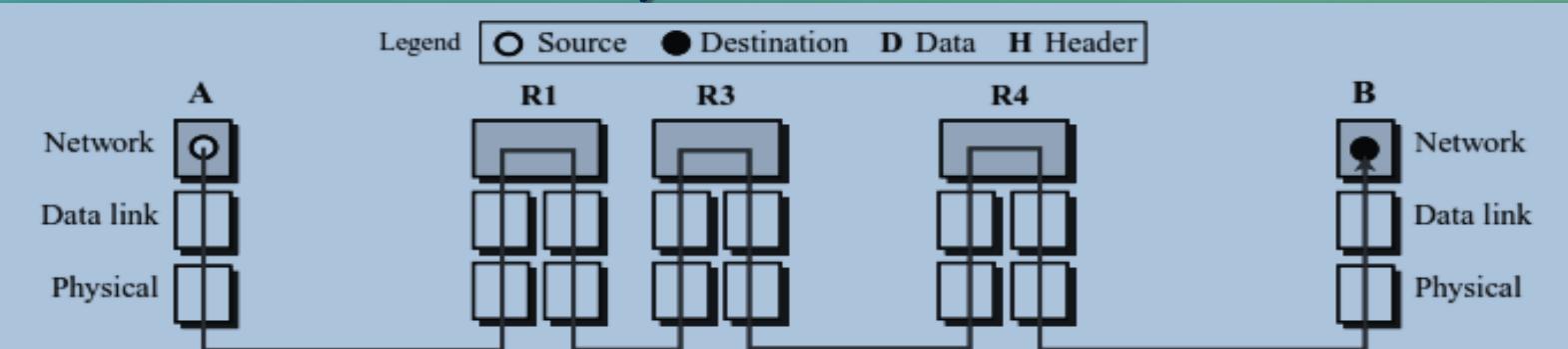
THE TCP/IP PROTOCOL SUITE

■ The Internet/Network Layer

- The Internet Protocol (IP) is used at Internet Layer.
- IP transports data in packets called datagrams.
- Datagrams can travel along different routes and can arrive out of sequence or be duplicated.
- IP does not keep track of the routes and has no facility for reordering datagrams.
- IP uses a logical address (IP Address) for unique identification of each host in Internet.

THE TCP/IP PROTOCOL SUITE

The Internet/Network Layer



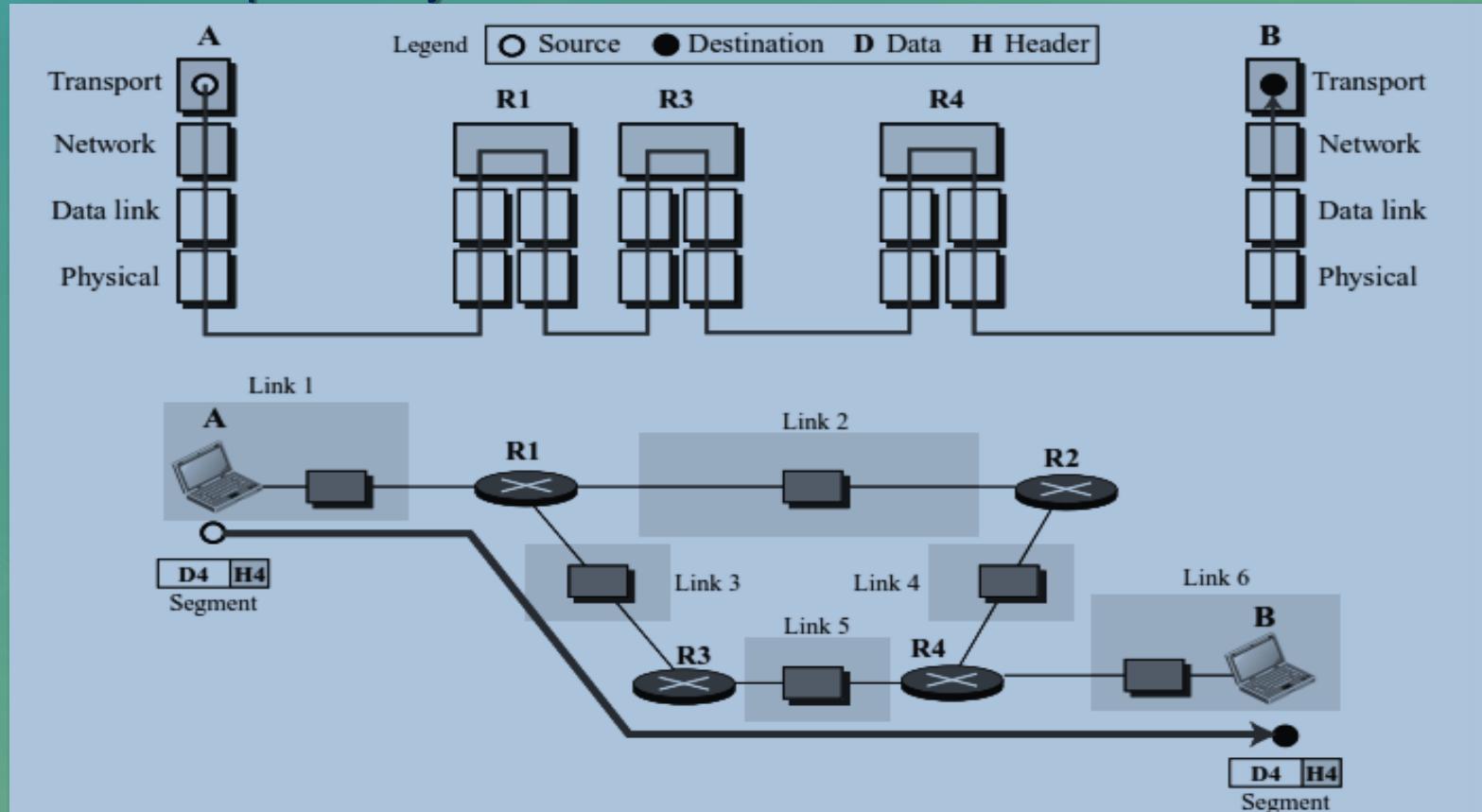
THE TCP/IP PROTOCOL SUITE

■ The Transport Layer

- The User Datagram Protocol (UDP) and Transmission Control Protocol (TCP) are used at Transport Layer.
- The transport layer is responsible for delivering the whole message called segments.
- The segments need to be broken into datagrams.
- Each segment have a sequence number for reordering.
- Datagrams are assembled into a segment at destination by Transport Layer.

THE TCP/IP PROTOCOL SUITE

The Transport Layer



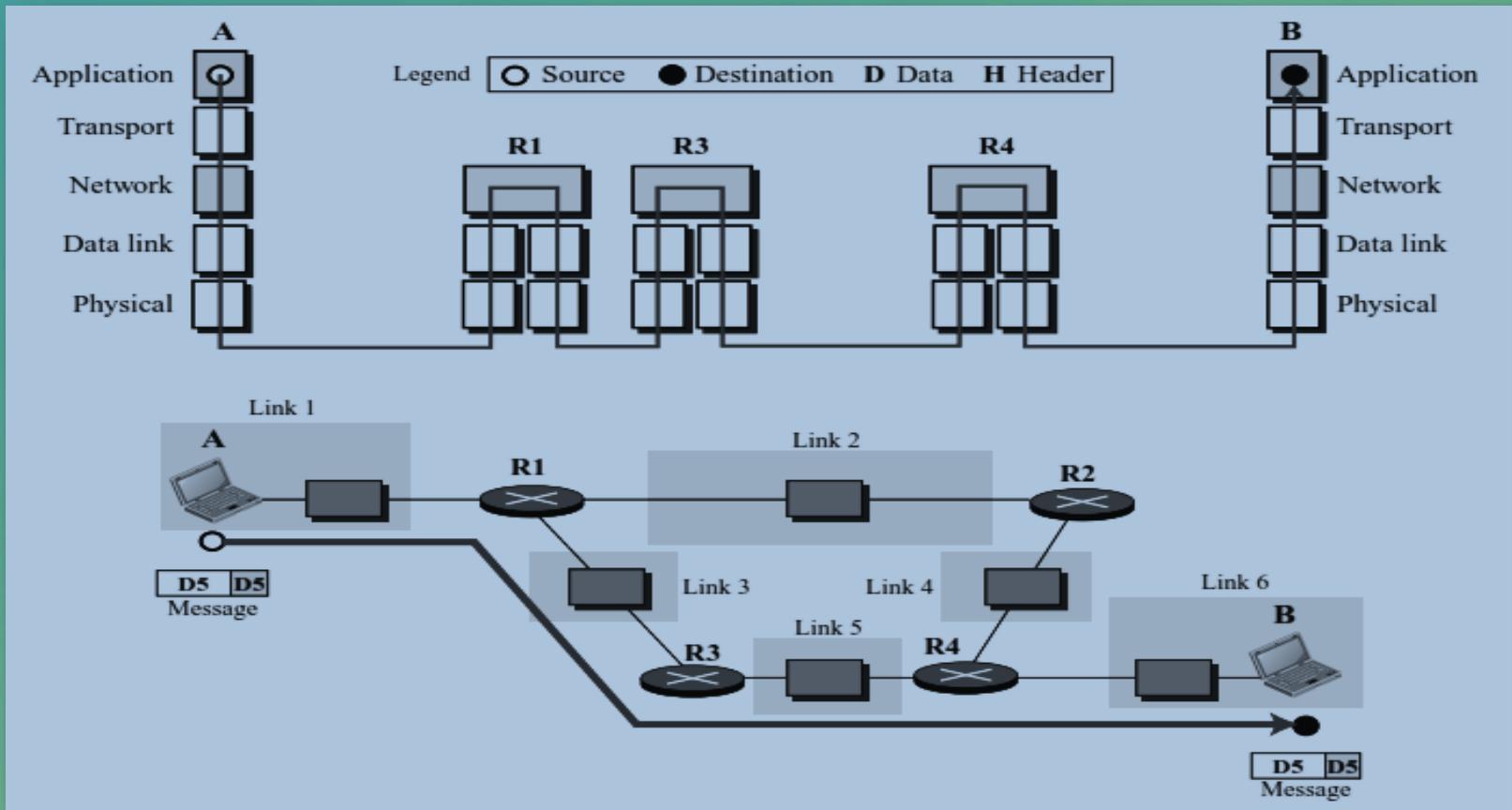
THE TCP/IP PROTOCOL SUITE

■ The Application Layer

- The application layer allows a user to access the services of Internet.
- Many protocols are defined at this layer to provide services such as
 - ✓ electronic mail,
 - ✓ file transfer,
 - ✓ accessing the World Wide Web etc.

THE TCP/IP PROTOCOL SUITE

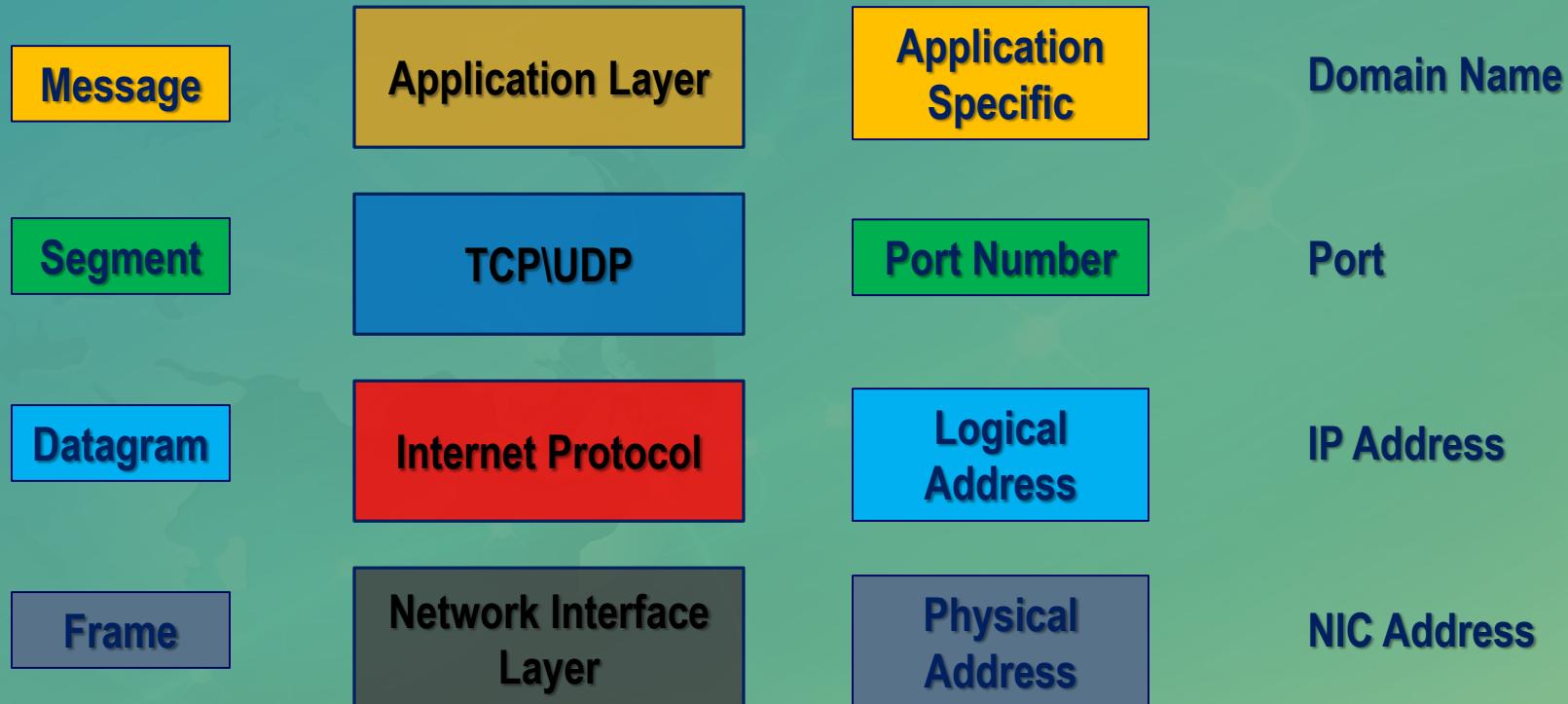
The Application Layer



ADDRESSING IN INTERNET

- Four levels of addresses are used in Internet:
 - Physical address
 - Logical address
 - Port address
 - Application-specific address
- Each address is related to a one layer in the TCP/IP architecture.

ADDRESSING IN INTERNET



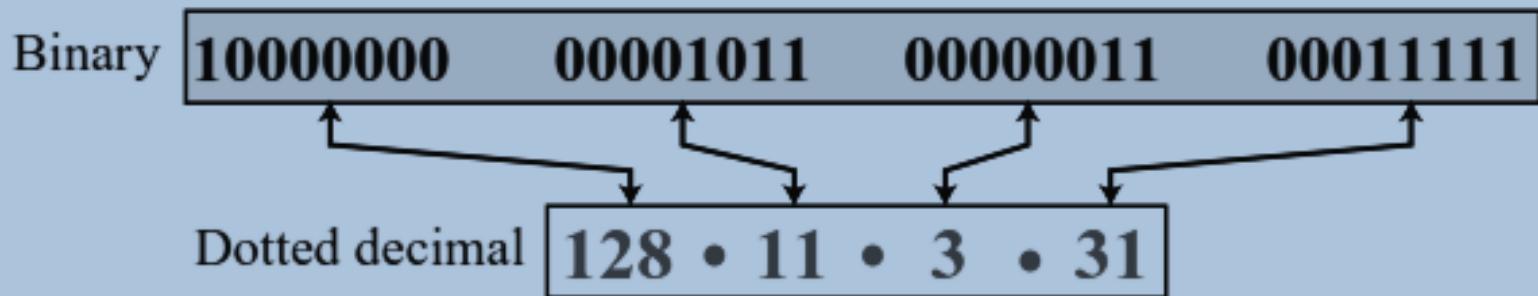
ADDRESSING IN INTERNET

Logical Address:

- Different networks can have different physical address formats.
- A universal addressing system is needed in which each host can be identified uniquely in Internet.
- Internet uses a 32-bit long IP address that uniquely identify a host in Internet.
- No two publicly addressed and visible hosts on the Internet can have the same IP address.

ADDRESSING IN INTERNET

Logical Address:



- Internet uses a 32-bit long IP address that uniquely identify a host in Internet.
- No two publicly addressed and visible hosts on the Internet can have the same IP address.

ADDRESSING IN INTERNET

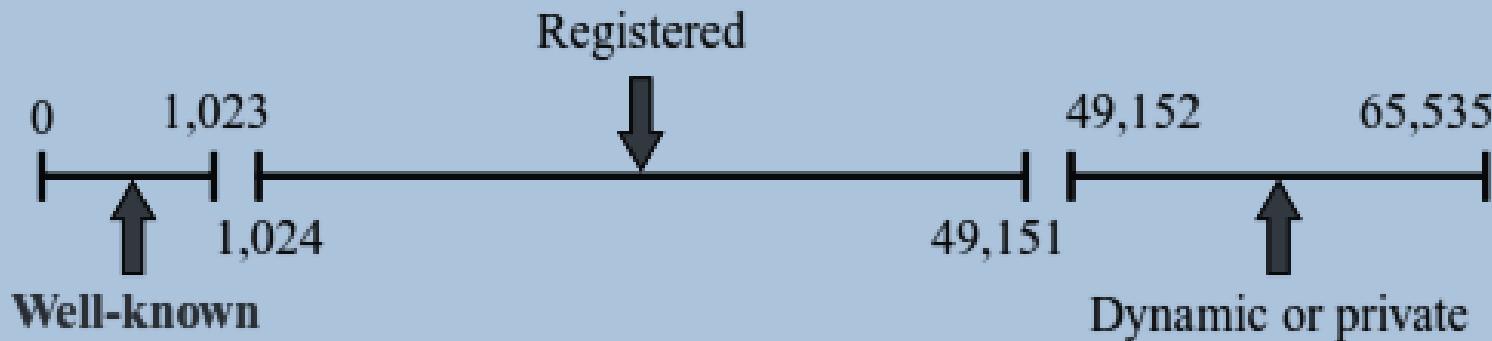
■ Port Address:

- The IP address is necessary for a quantity of data to travel from a source to the destination host.
- A multitasking system can execute multiple process simultaneously and each process can communicate over Internet with another process.
- In the TCP/IP architecture, the label assigned to a process is called a port address.
- A port address in TCP/IP is 16 bits in length, represented by one decimal number.

ADDRESSING IN INTERNET

Port Address:

- The IP address is necessary for a quantity of data to travel from a source to the destination host.



- A port address in TCP/IP is 16 bits in length, represented by one decimal number.



INTRODUCTION TO WORLD WIDE WEB

SERVICES OF INTERNET

Internet services can be divided into four groups:

➤ Communication Services

- ✓ Electronic mail (E-mail) .
- ✓ Telnet : Used to log on to a remote computer.
- ✓ Internet Relay Chat (IRC)
- ✓ Internet Telephony (VoIP)
- ✓ Instant Messaging

SERVICES OF INTERNET

- Internet services can be divided into four groups:
 - Communication Services
 - Information Service
- ✓ File Transfer Protocol (FTP)
 - ✓ Gopher : Used to search, retrieve, and display documents on remote sites.

SERVICES OF INTERNET

Internet services can be divided into four groups:

- Communication Services
- Information Service
- Web Service

- ✓ A web service is any piece of software that makes itself available over the Internet.
- ✓ A web service is a collection of open protocols and standards for exchange of application logic.

SERVICES OF INTERNET

Internet services can be divided into four groups:

- Communication Services
- Information Service
- Web Service
- World Wide Web

- ✓ **World Wide Web (WWW), byname the Web, the leading information sharing service of the Internet.**
- ✓ **The Web gives users access to a vast array of documents that are connected to each other by means of hypertext or hypermedia links**

WORLD WIDE WEB

- WWW is an Internet service that allows computer users to access web-based applications and to locate and view multimedia-based documents on almost any subject over the Internet.
- In 1989, Tim Berners-Lee of CERN (the European Organization for Nuclear Research) began to develop a technology for sharing scientific information via hyperlinked text documents.
- The WWW today is a **distributed client-server service**, in which a client using a browser can access a service using a server.

WORLD WIDE WEB



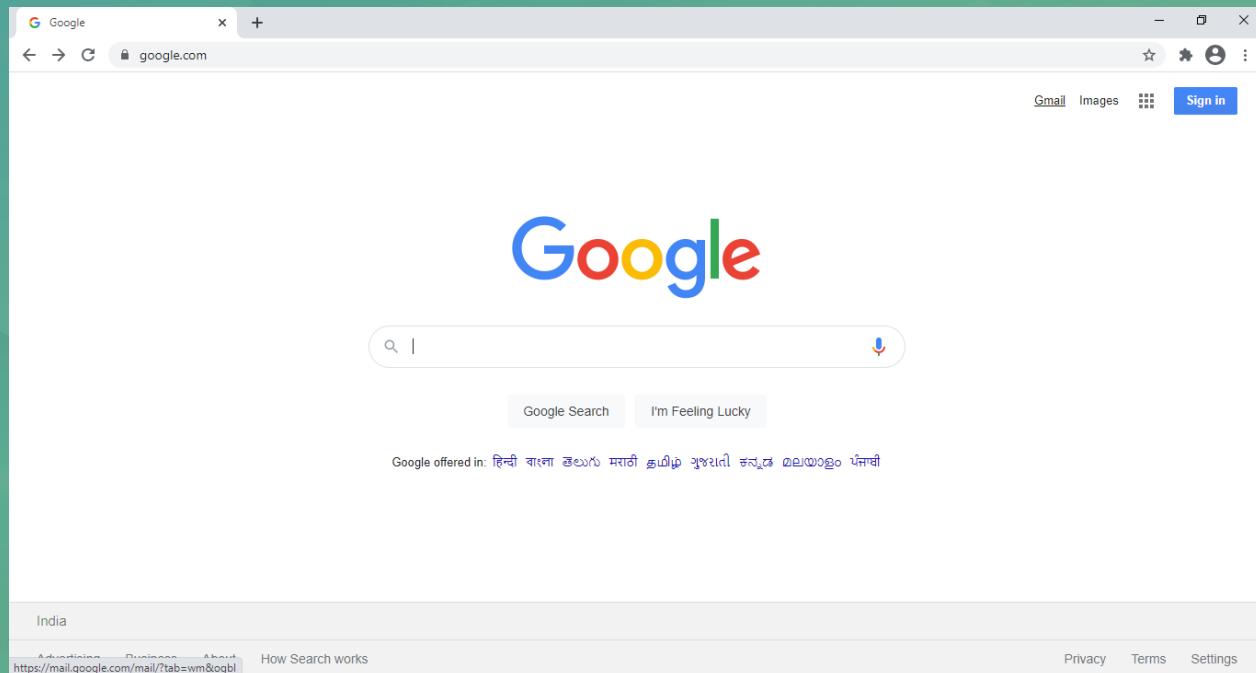
WEB CLIENT (BROWSER)

- Web browser is an application software that interpret and display a Web document.
- Primary Functions:
 - It takes the user inputs in the form of Uniform Resource Locator (URL).
 - Displays web pages to the client.
 - It allows the user to interact with the web pages.

WEB CLIENT (BROWSER)

Anatomy of Web Browsers

- Title Bar
- Toolbar
- Address Field
- RSS Button
- Search Box
- Status Bar
- Tabs
- Bookmarks



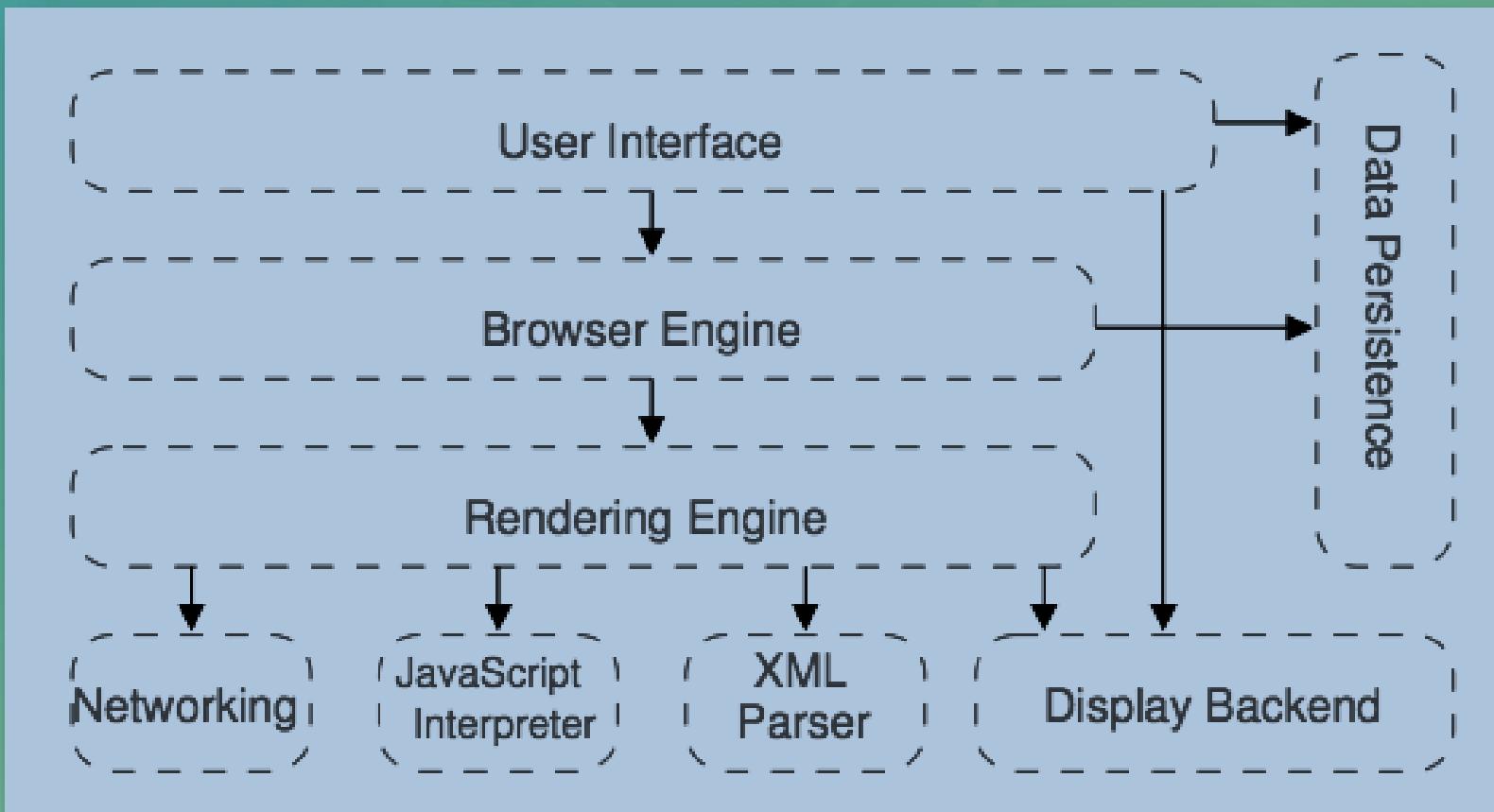
WEB CLIENT (BROWSER)

■ Architecture of a Browser

- A browser is a group of structured codes.
- These codes are separated in to different components.
 - User Interface
 - Browser Engine
 - Rendering Engine
 - Networking
 - Java Script Interpreter & UI Backend
 - Data Storage

WEB CLIENT (BROWSER)

Architecture of a Browser



WORLD WIDE WEB



WEB SERVER

- Web Server is a computer where the web content is stored.
- The Web page is stored at the server.
- Each time a client request arrives, the corresponding document is sent to the client.
- Servers normally store requested files in a cache in memory to improve efficiency.
- A server can answer more than one request at a time.
- E.g. Apache Tomcat, Microsoft Internet Information Server.

WORLD WIDE WEB



UNIFORM RESOURCE LOCATOR (URL)

- URIs (Uniform Resource Identifiers) identify resources on the Internet.
- URL is a subset of URIs that identify the network location of a resource.

Protocol

::/

Host

:

Port

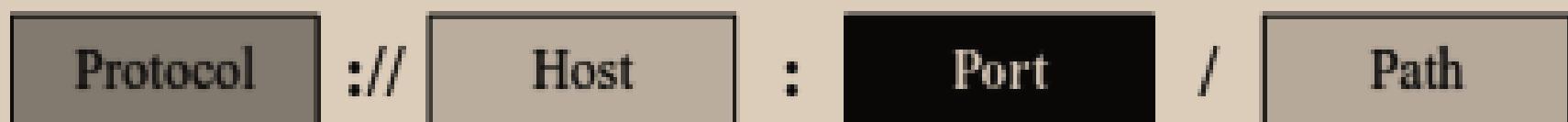
/

Path

Client-server application program used to retrieve the document. E.g. HTTP, FTP

UNIFORM RESOURCE LOCATOR (URL)

- URIs (Uniform Resource Identifiers) identify resources on the Internet.
- URL is a subset of URIs that identify the network location of a resource.



Domain name of the computer on which the information is located.
E.g. www.google.com

UNIFORM RESOURCE LOCATOR (URL)

- URIs (Uniform Resource Identifiers) identify resources on the Internet.
- URL is a subset of URIs that identify the network location of a resource.

Protocol

::/

Host

:

Port

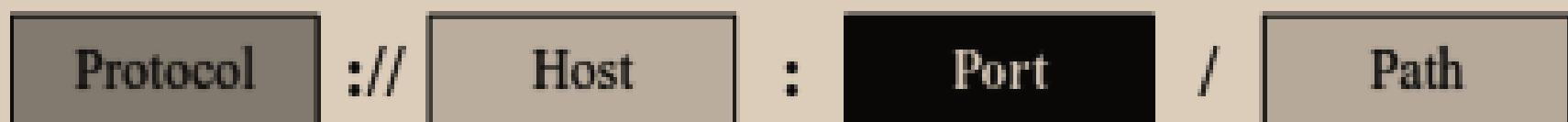
/

Path

The URL can
optionally contain the
port number of the
server.

UNIFORM RESOURCE LOCATOR (URL)

- URIs (Uniform Resource Identifiers) identify resources on the Internet.
- URL is a subset of URIs that identify the network location of a resource.



- <https://www.google.com/index.html>

The path defines the complete file name where the document is stored in the directory system.

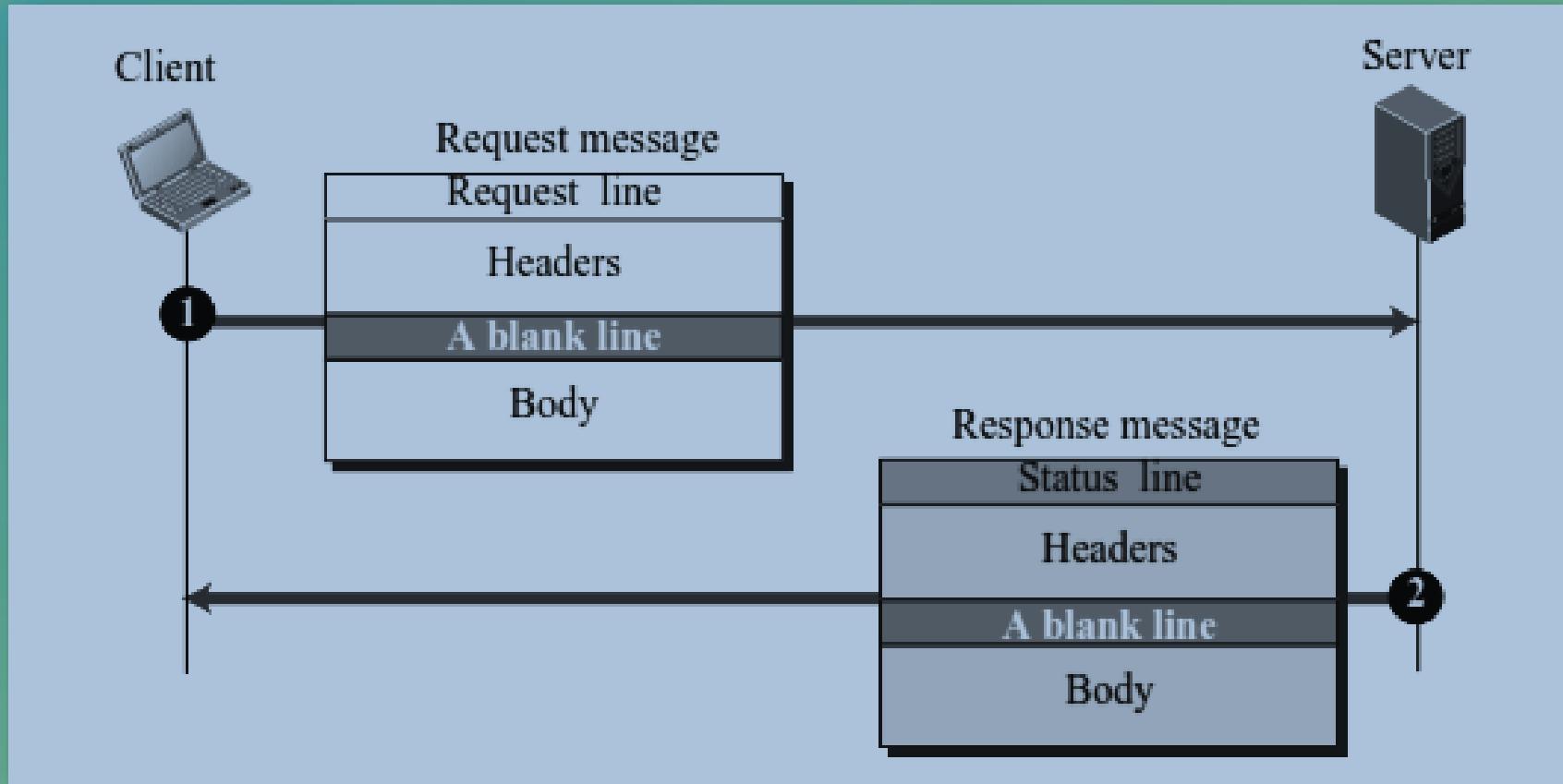
WORLD WIDE WEB



HYPertext Transfer Protocol (HTTP)

- HTTP is a protocol used mainly to access data on the World Wide Web.
- HTTP uses the services of TCP on well-known port 80.
- HTTP itself is a stateless protocol.
- The client initializes a HTTP transaction by sending a request. The server replies by sending a response.

HYPertext Transfer Protocol (HTTP)



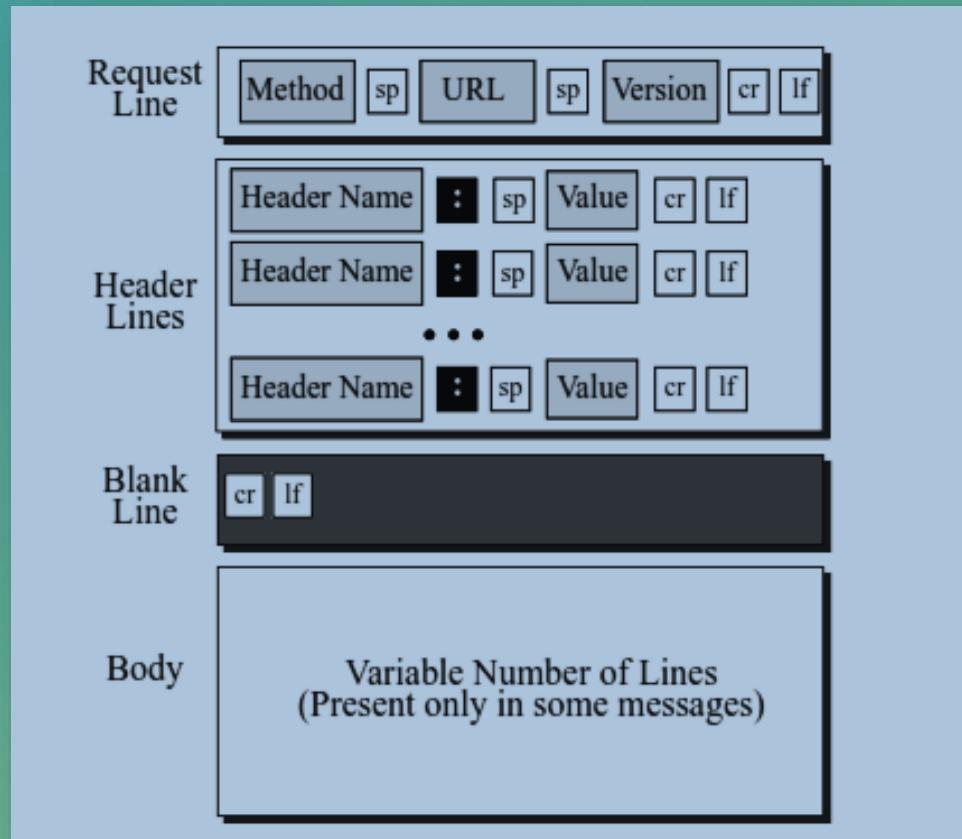
HYPertext Transfer Protocol (HTTP)

- Type of messages:

- Request Message
- Response Message

HYPertext Transfer Protocol (HTTP)

Request Message Format



HYPertext Transfer Protocol (HTTP)

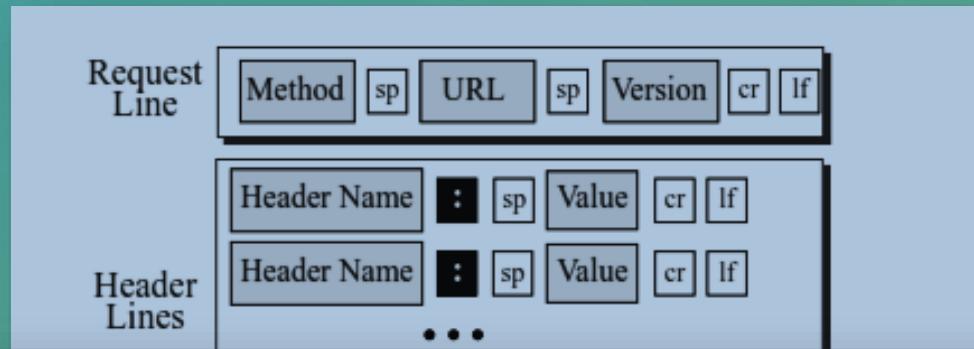
Request Message Format



<i>Method</i>	<i>Action</i>
GET	Requests a document from the server
HEAD	Requests information about a document but not the document itself
POST	Sends some information from the client to the server
PUT	Sends a document from the server to the client
TRACE	Echoes the incoming request
CONNECT	Reserved
DELETE	Remove the Web page
OPTIONS	Enquires about available options

HYPertext Transfer Protocol (HTTP)

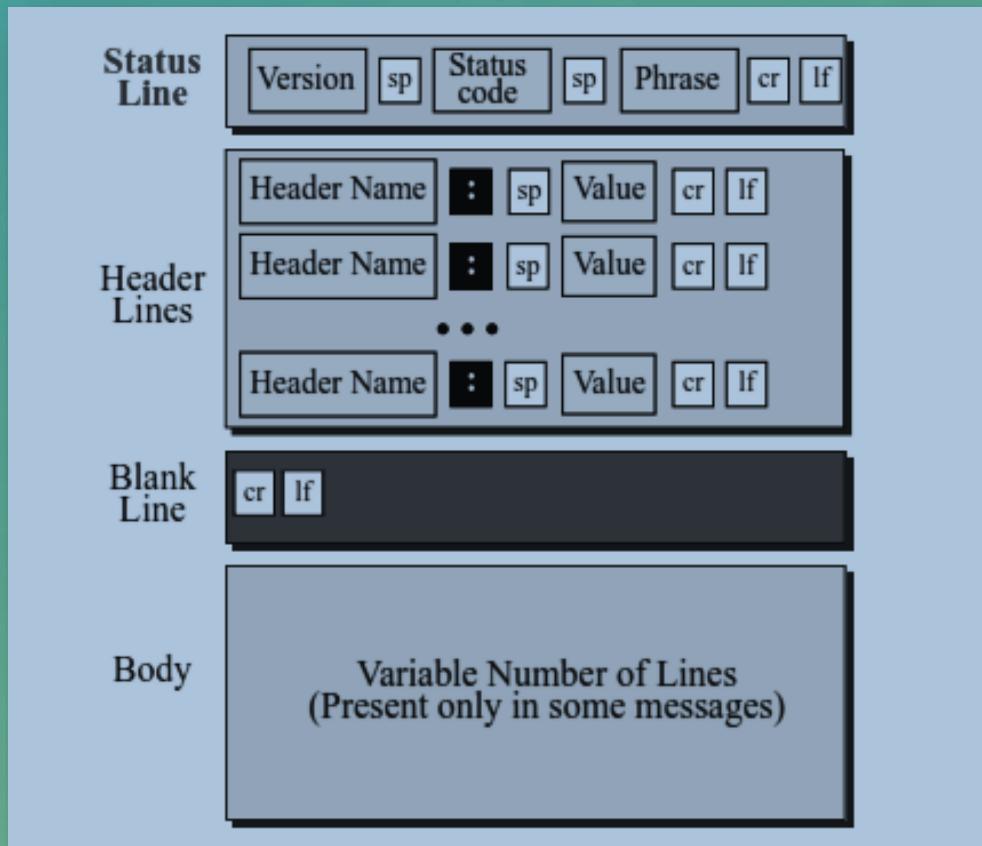
Request Message Format



<i>Header</i>	<i>Description</i>
User-agent	Identifies the client program
Accept	Shows the media format the client can accept
Accept-charset	Shows the character set the client can handle
Accept-encoding	Shows the encoding scheme the client can handle
Accept-language	Shows the language the client can accept

HYPertext Transfer Protocol (HTTP)

Response Message Format



HYPertext Transfer Protocol (HTTP)

Response Message Format

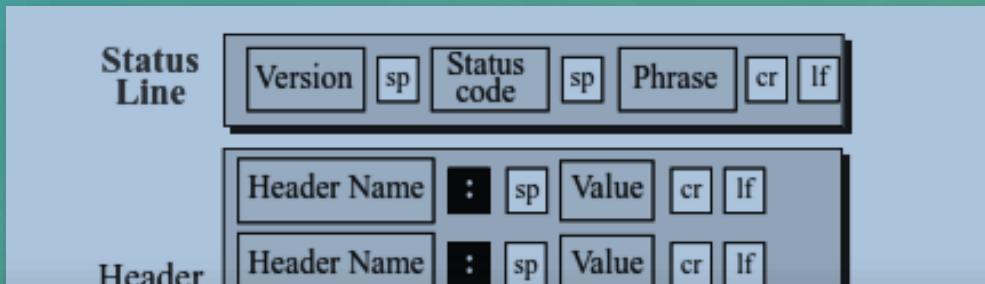


<i>Status Code</i>	<i>Status Phrase</i>	<i>Description</i>
Informational		
100	Continue	The initial part of the request received, continue.
101	Switching	The server is complying to switch protocols.
Success		
200	OK	The request is successful.
201	Created	A new URL is created.
202	Accepted	The request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

		upon.
204	No content	There is no content in the body.
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 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.

HYPertext Transfer Protocol (HTTP)

Response Message Format



<i>Header</i>	<i>Description</i>
Date	Shows the current date
Upgrade	Specifies the preferred communication protocol
Server	Gives information about the server
Set-Cookie	The server asks the client to save a cookie
Content-Encoding	Specifies the encoding scheme
Content-Language	Specifies the language

HYPertext Transfer Protocol (HTTP)

Example



HYPertext Transfer Protocol (HTTP)

■ Type of HTTP connections

➤ Nonpersistent Connection

- One TCP connection is made for each request/response.

➤ Persistent Connection

- In a persistent connection, the server leaves the connection open for more requests after sending a response.
- HTTP version 1.1 specifies a persistent connection by default.

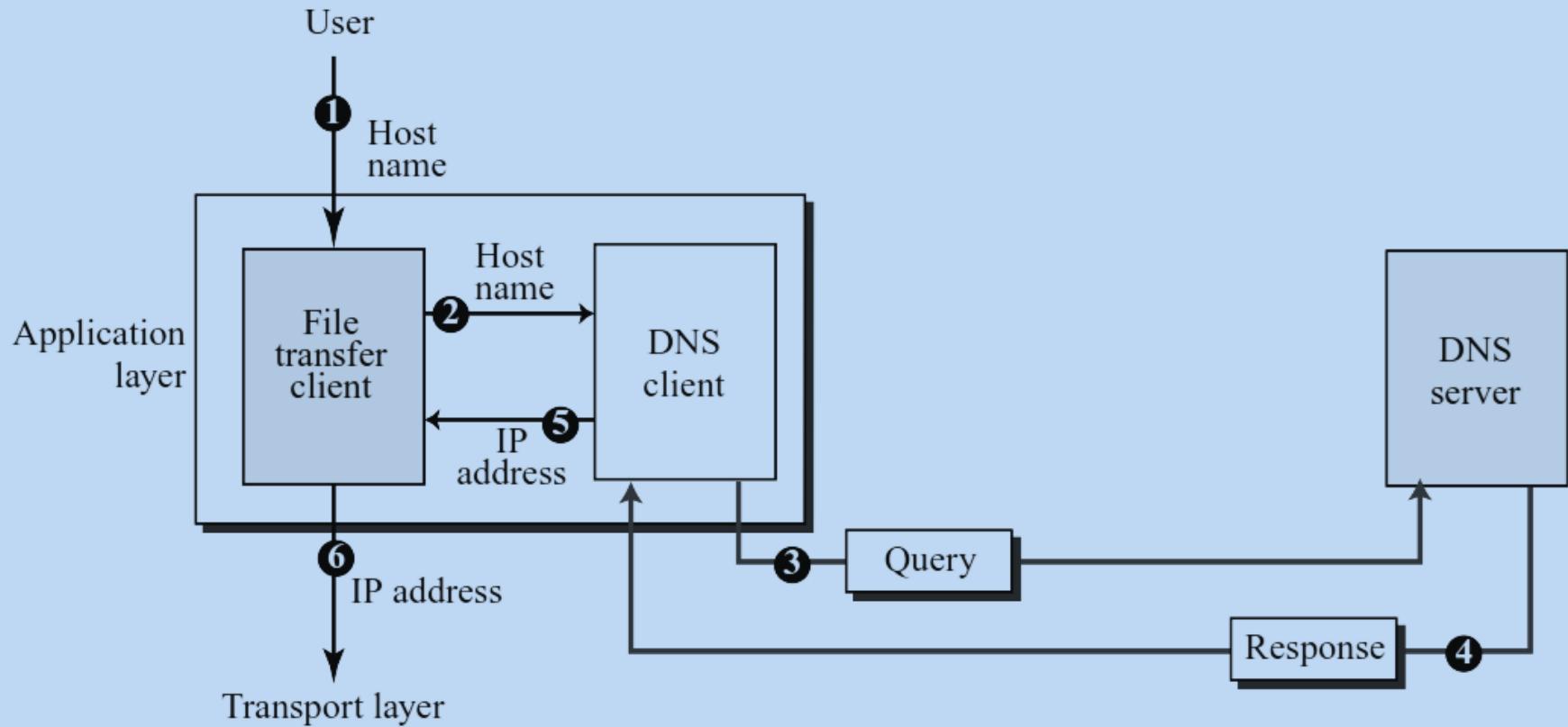
DOMAIN NAME SYSTEM (DNS)

- DNS is a client/server application program used to help other application layer programs.
- DNS is used to map a host name in the application layer to an IP address in the network layer.
- WWW uses host name as address of web server at application layer.
- Users can remember name easily.
- TCP/IP protocols uses the IP address of web server to uniquely identify a host.

DOMAIN NAME SYSTEM (DNS)

- Earlier in Internet, host file is used to map host name with corresponding IP address of web servers.
- Size of Internet is very large, so maintaining a host file at each is not feasible for mapping.
- Storage of a large file at a single centralized system creates huge amount of traffic on the Internet.
- Domain Name System (DNS) is used now a days for mapping host name with IP address.
- It uses DNS protocol.

DOMAIN NAME SYSTEM (DNS)



DNS : NAME SPACE

- DNS uses name space to provide unique name to each host in Internet.
- A name space can be organized in two ways:
 - Flat Name Space
 - A name in this space is a sequence of characters without structure.
 - It can not be used in Internet.
 - Requires a centralized authority to assign names.
 - Hierarchical Name Space

DNS : DOMAIN NAME SPACE

- A domain name space was designed for Internet that uses a hierarchical name space.
- The names are defined in an inverted-tree structure with the root at the top.
- The tree can have only 128 levels: level 0 (root) to level 127.
- Each node in the tree has a label : a string with a maximum of 63 characters.
- The root label is a null string.
- Children of a node should have different labels.

DNS : DOMAIN NAME

- Each node in the DNS tree has a domain name.
- A full domain name is a sequence of labels from the node upto root separated by dots (.).
- The last character of a domain name is a dot.
- Fully Qualified Domain Name (FQDN)
 - If a label is terminated by a null string
 - A DNS server can only match an FQDN to an address.
- Partially Qualified Domain Name (PQDN)
 - If a label is not terminated by a null string

DNS : DOMAIN

- A domain is a subtree of the domain name space.
- The name of the domain is the name of the node at the top of the subtree.
- A domain may itself be divided into domains (sub-domain)

DNS : DISTRIBUTION OF NAME SPACE

- ❑ It is not possible to store all domain name on a single server.
- ❑ It will increase load on a single server.
- ❑ It is not reliable because any failure makes the data inaccessible.
- ❑ So, domain names are distributed among many computers called DNS Servers.
- ❑ Each server can be responsible (authoritative) for either a large or small domain.

DNS : DISTRIBUTION OF NAME SPACE

- A server is responsible for or has authority over is called a zone.
- Name space is divided among various hierarchical zones.
- The server have a database called a zone file and keeps all the information for every node under that domain.
- A root server is a server whose zone consists of the whole tree.
- There are several root servers, each covering the whole domain name space.
- The root servers are distributed all around the world.

DNS : DISTRIBUTION OF NAME SPACE

■ DNS defines two types of servers:

➤ Primary Server

- ✓ It is a server that stores a file about the zone for which it is an authority.
- ✓ It is responsible for creating, maintaining, and updating the zone file.

➤ Secondary Server

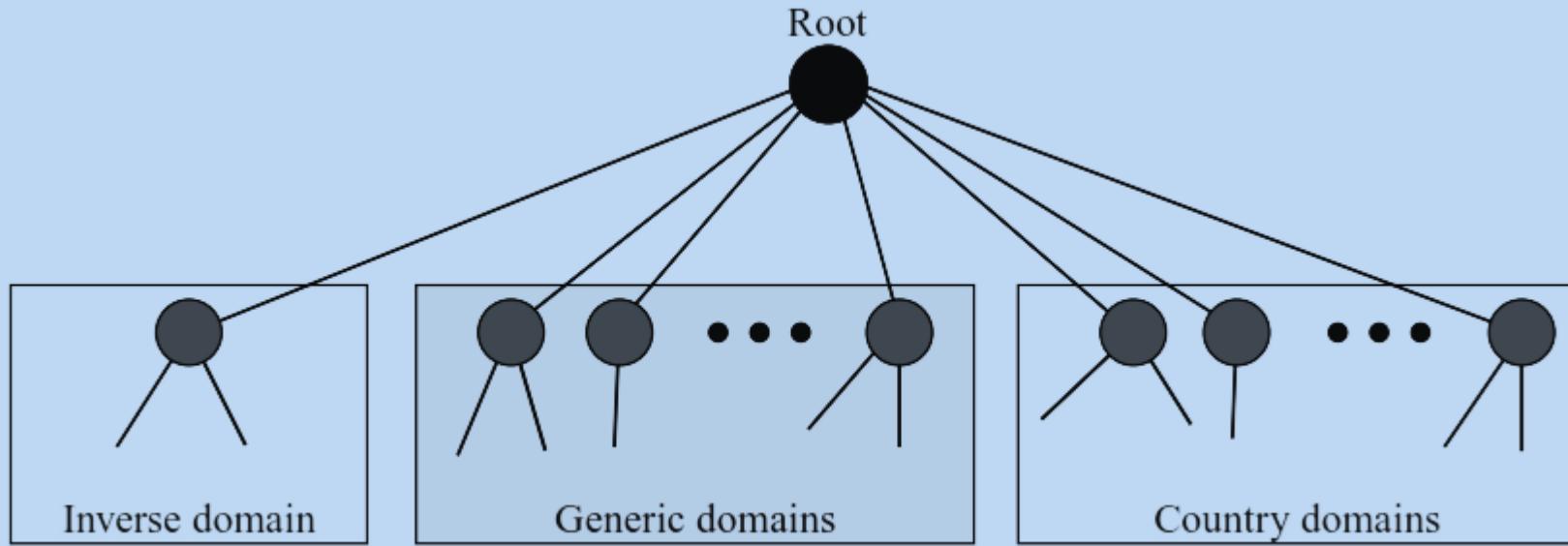
- ✓ It is a server that loads the complete information about a zone from another server.
- ✓ The secondary server neither creates nor updates the zone files.

DNS IN INTERNET

- DNS is a protocol that can be used in different platforms.
- The domain name space (tree) is divided into three different sections in Internet:
 - Generic Domains
 - Country Domains
 - Inverse Domain

DNS IN INTERNET

- DNS is a protocol that can be used in different platforms.
- The domain name space (tree) is divided into three different sections in Internet:



DNS IN INTERNET

■ Generic Domains

- The generic domains define registered hosts according to their generic behavior.

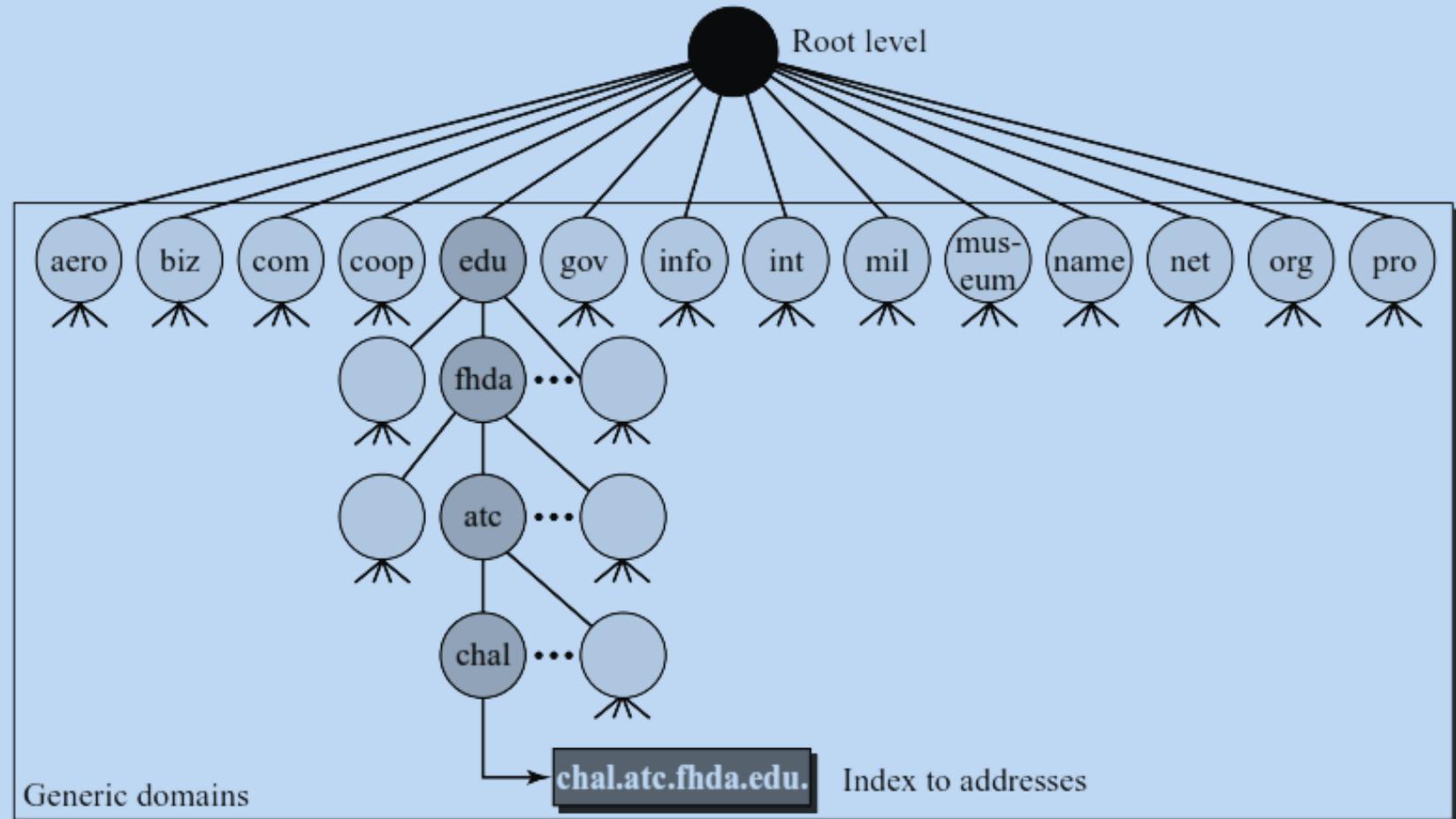
■ Country Domains

- The country domains section uses two-character country abbreviations.

■ Inverse Domain

- The inverse domain is used to map an address to a name.

DNS IN INTERNET



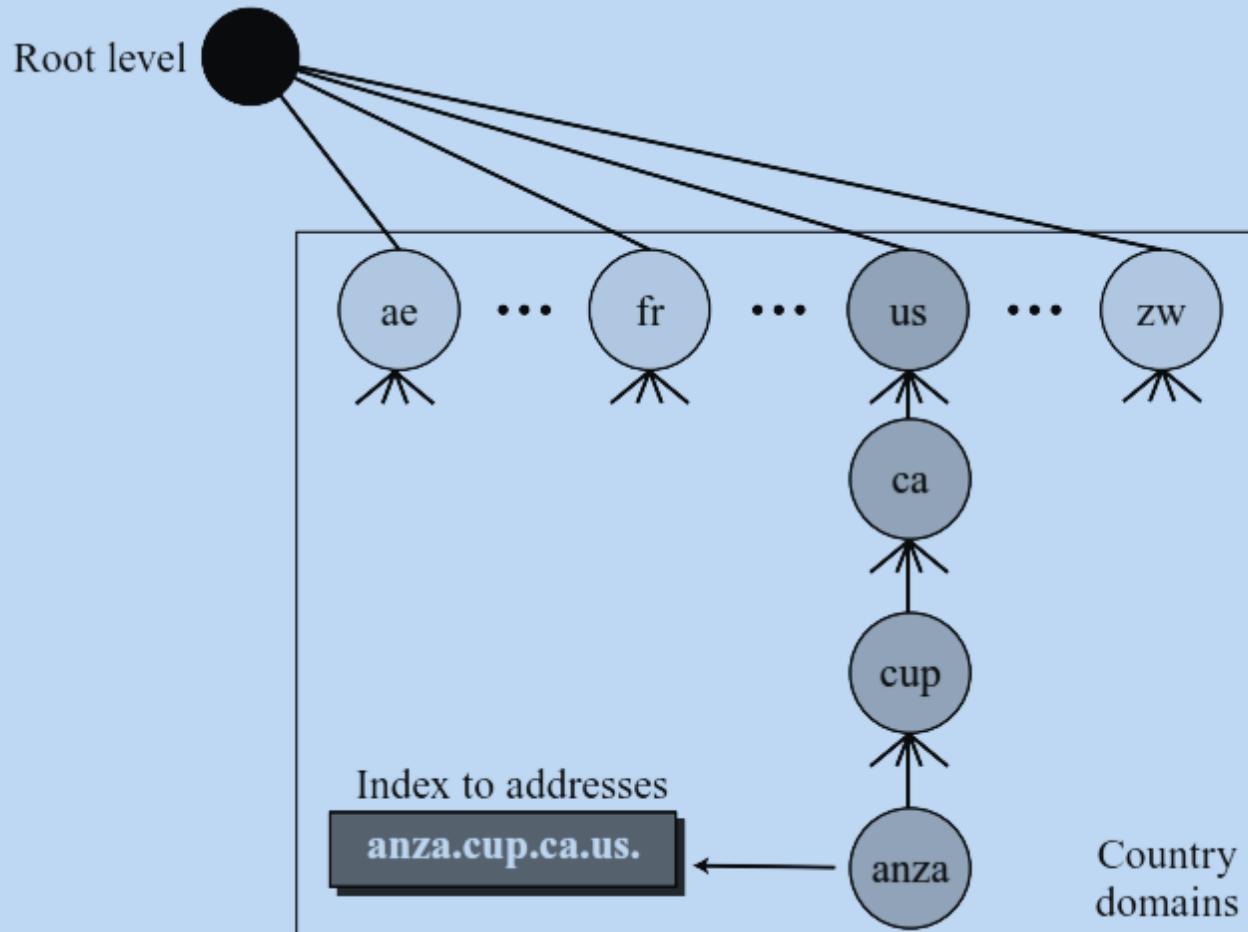
DNS IN INTERNET



Generic

<i>Label</i>	<i>Description</i>
aero	Airlines and aerospace companies
biz	Businesses or firms (similar to “com”)
com	Commercial organizations
coop	Cooperative business organizations
edu	Educational institutions
gov	Government institutions
info	Information service providers
int	International organizations
mil	Military groups
museum	Museums and other non-profit organizations
name	Personal names (individuals)
net	Network support centers
org	Nonprofit organizations
pro	Professional individual organizations

DNS IN INTERNET



DNS IN INTERNET

General

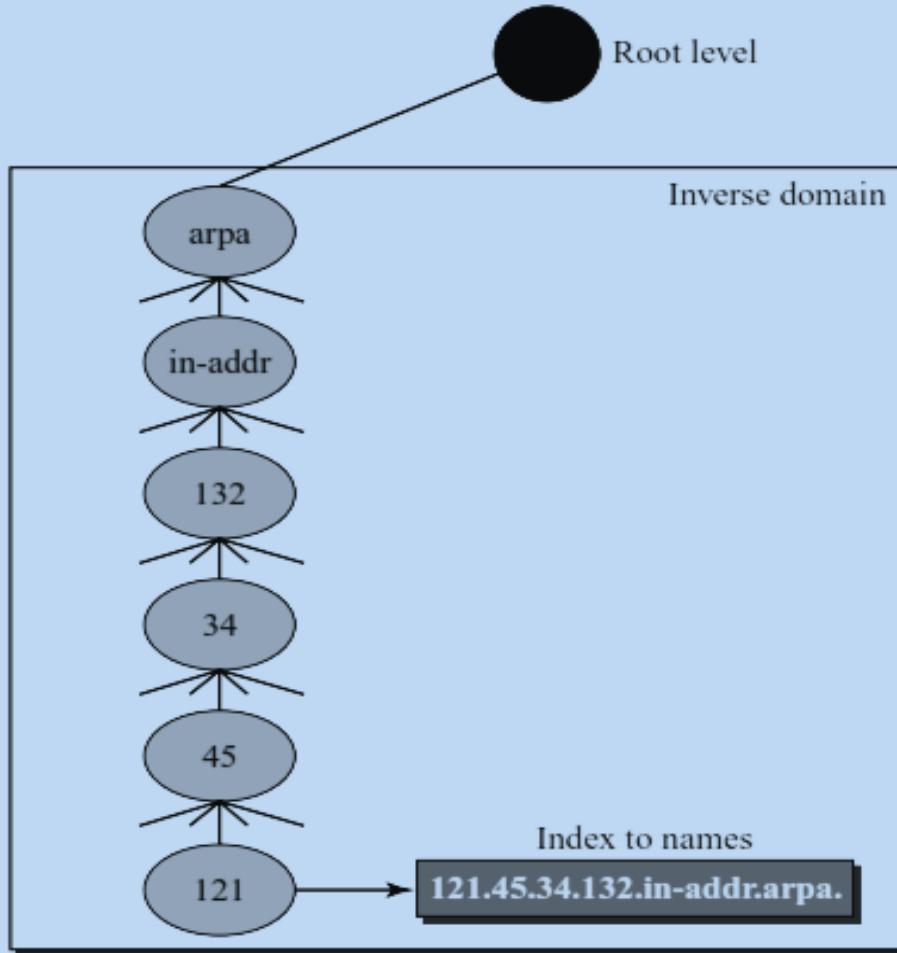
- The domain tree starts from the root level.

Countries

- The domains are divided into country codes.

Inverse

- The names are converted into IP address.



hosts according

character country

address to a

DNS : DNS RESOLUTION

- Mapping a name to an address or an address to a name is called name-address resolution.
- A host that needs to map an address to a name or a name to an address calls a DNS client called a resolver.
- The resolver accesses the closest DNS server with a mapping request.
- After the resolver receives the mapping, it verifies and delivers the result to the process that requested it.

DNS : DNS RESOLUTION

■ Resolver can do two type of mapping:

➤ Names to Addresses

- ✓ The resolver gives a domain name to the server and asks for the corresponding address.

➤ Addresses to Names

- ✓ The resolver sends a inverse DNS query to the server and asks for the corresponding domain name.
- ✓ This is also called PTR query.
- ✓ The IP address is reversed and two labels, in-addr and arpa, are appended to create a domain acceptable by the inverse domain section.

DNS : DNS RESOLUTION

- A resolution can be of two types:

- Recursive Resolution
- Iterative Resolution

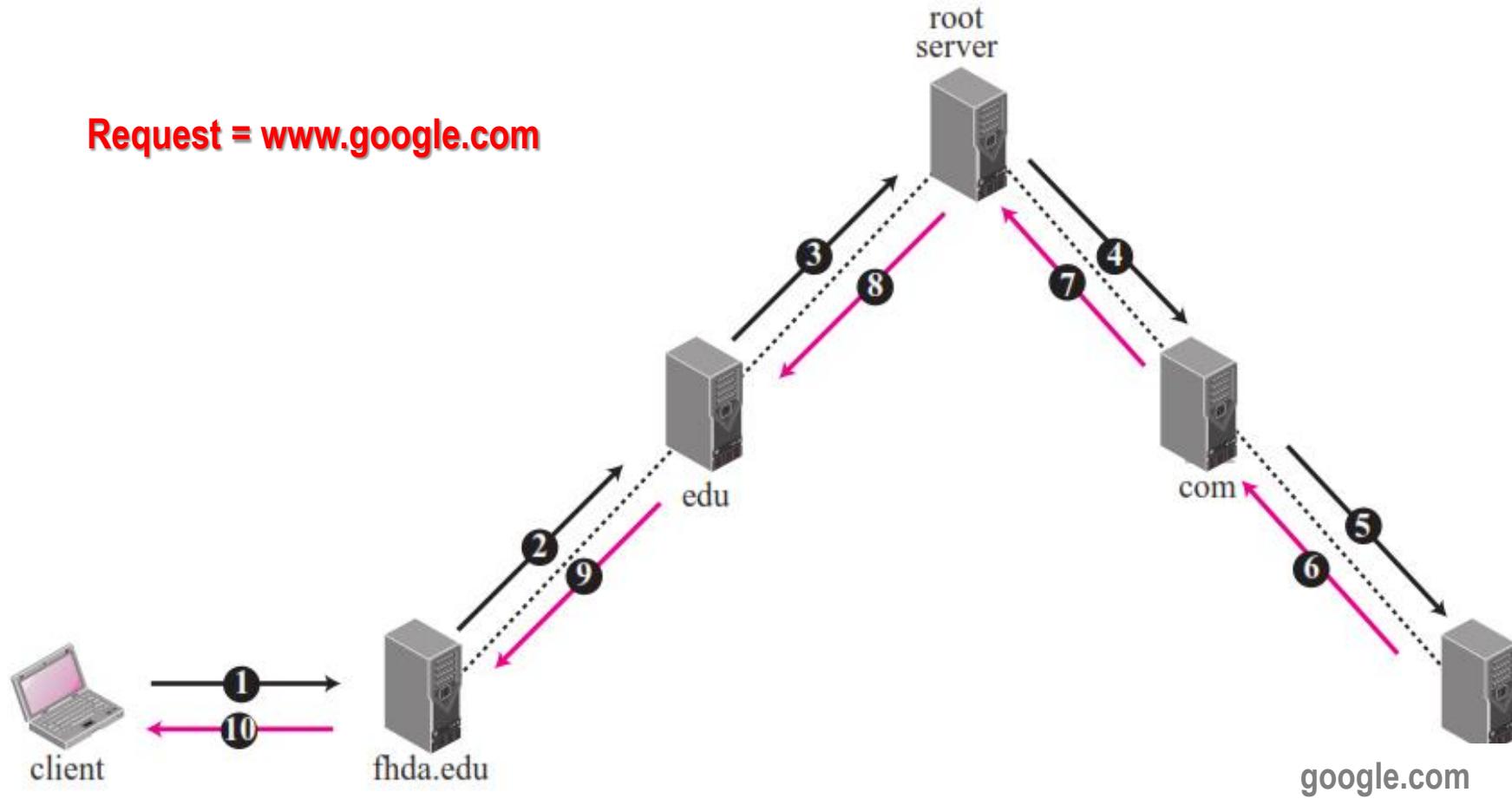
DNS : DNS RESOLUTION

█ Recursive Resolution

- The client (resolver) asks for a recursive answer from a nearest local DNS server.
- Client requires the Local Server to give either the requested mapping or an error message.
- If the server is the authority for the domain name, it checks its database and responds.
- If the server is not the authority, it sends the request to another server (the parent usually) and waits for the response.
- When the query is resolved, the response travels back until it reaches the requesting client.

DNS : DNS RESOLUTION

Request = **www.google.com**

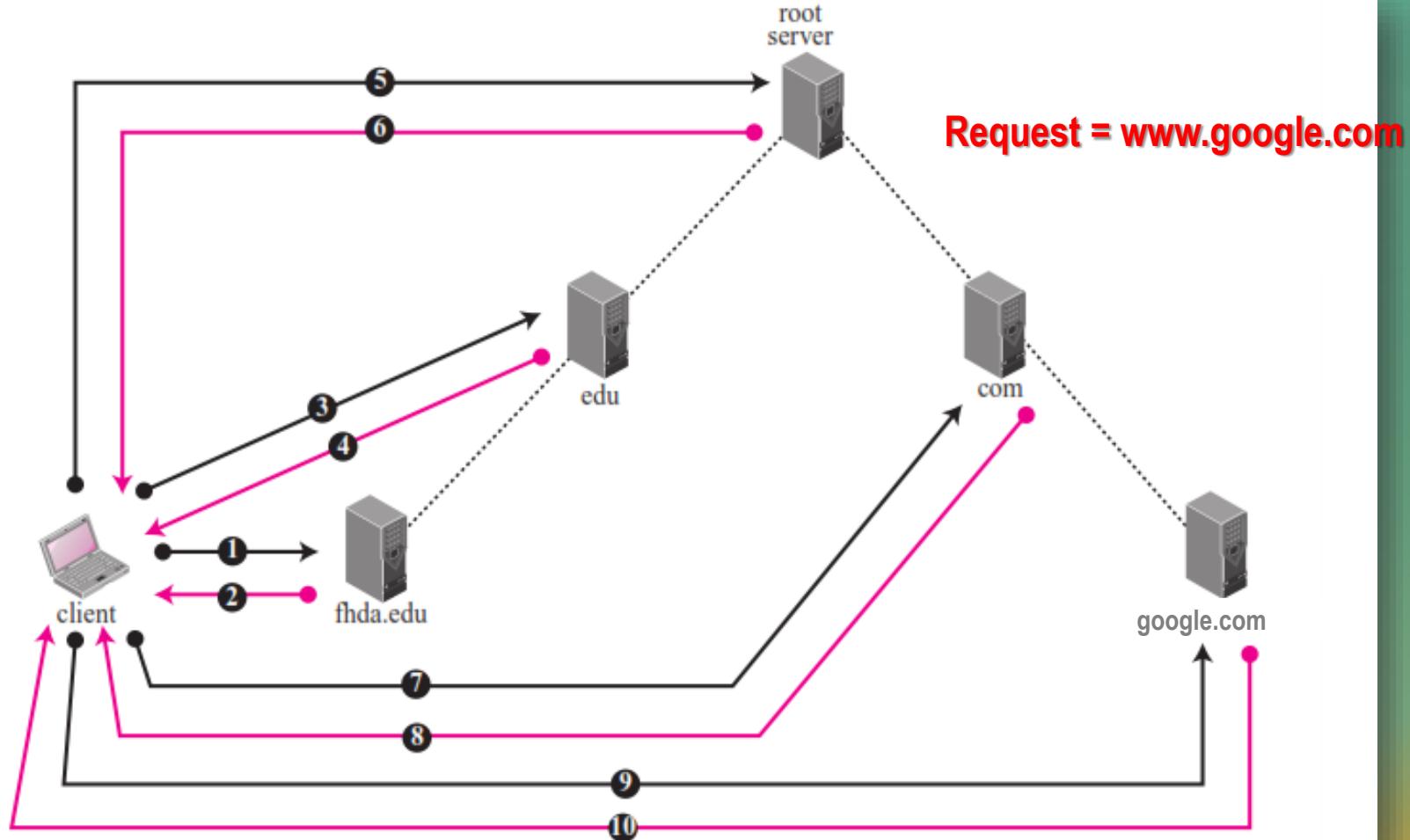


DNS : DNS RESOLUTION

■ Iterative Resolution

- The client does not ask for a recursive answer, then the mapping can be done iteratively.
- If the server is an authority for the name, it sends the answer.
- If it is not, it returns (to the client) the IP address of the server that it thinks can resolve the query.
- The client is responsible for repeating the query to this second server.
- This process is called iterative because the client repeats the same query to multiple servers.

DNS : DNS RESOLUTION



DNS : DNS RESOLUTION

■ Caching Mechanism

- A server stores the response from another server in cache memory before sending to client.
- If the same or another client asks for the same mapping, it can check its cache memory and resolve the problem.
- The response is marked as unauthoritative to inform the client that the response is from Cache.
- The caching can be problematic, when server caches the mapping for a long time and the mapping gets outdated.
- To counter this, TTL (Time-to-Live) techniques are used.

WORLD WIDE WEB



WEB DOCUMENT

- A document available on WWW to share information that includes text, graphics, audio, video and hyperlinks.
- Web documents are stored at web server but viewed by web browser.
- Web document is also called web page.
- Collection of linked web pages on a web server is known as website.
- There is unique Uniform Resource Locator (URL) is associated with each web page.

WEB DOCUMENTS

- The documents in the WWW can be grouped into three broad categories:
 - Static Web Page
 - Dynamic Web Page
 - Active Web Page

WEB DOCUMENTS

■ Static Web Page

- Fixed-content documents stored at web server.
- The contents of the file are determined when the file is created, not when it is used.
- When a client accesses the document, a copy of the document is sent.
- Static documents are prepared using markup language.
- E.g. HTML, XML, XSL, XHTML

WEB DOCUMENTS

Static Web Page

Client



1

Request



Static document

Server



2



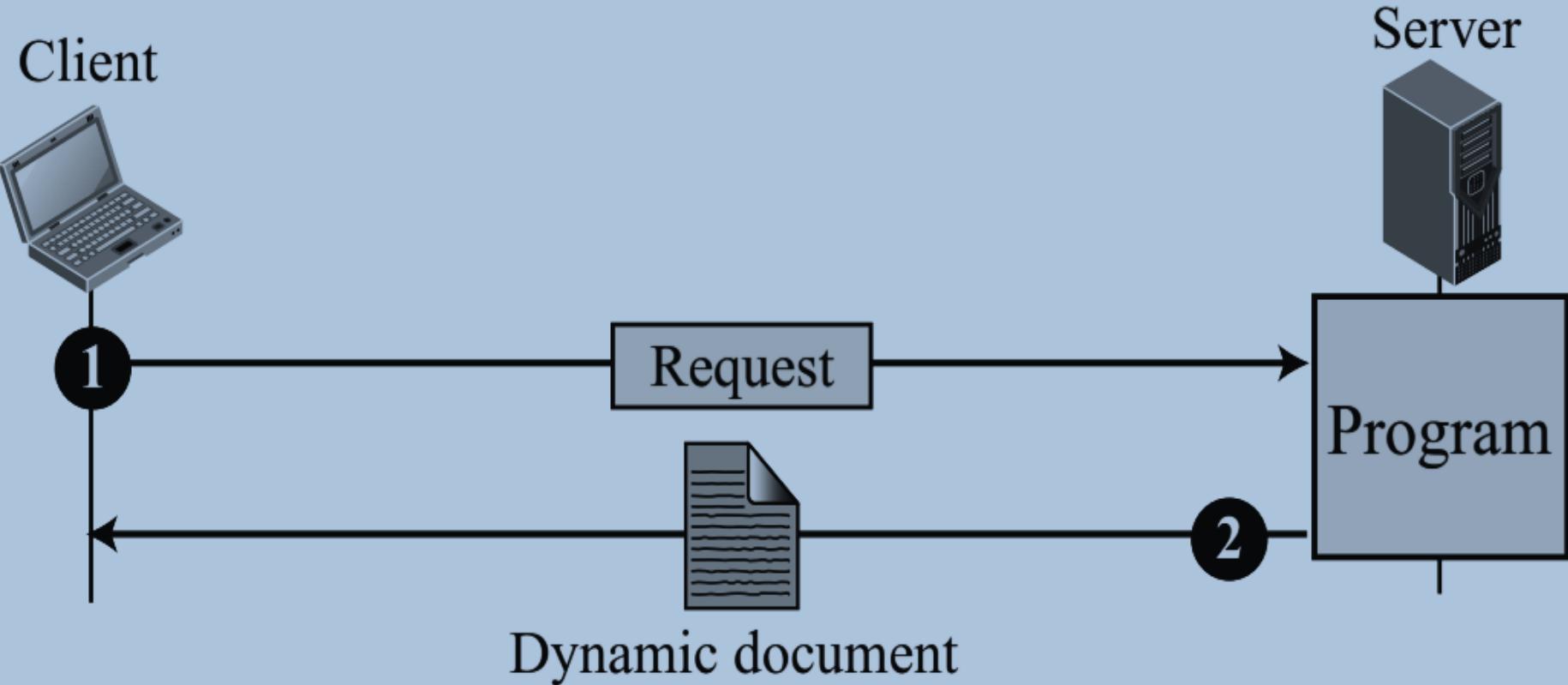
WEB DOCUMENTS

■ Dynamic Web Page

- Content of a web page change with time.
- Created by a Web server whenever a browser requests the document.
- Web server runs an application program or a script that creates the dynamic web page.
- A fresh document is created for each request, so content of vary from one request to another.
- Web server uses a server-side technology to create dynamic web page.

WEB DOCUMENTS

Dynamic Web Page



WEB DOCUMENTS

Active Web Page

- Content are generated at client side using a programming language.
- When a browser requests an active document, the server sends a copy of the document or a script.
- Programs or scripts are embedded in static web pages.
- Active documents are sometimes referred to as client-site dynamic documents.
- Uses client side technology like Applet, Scripting.

WEB DOCUMENTS

Active Web Page

