Chapter-3
Application Layer

Introduction

- Application layer is the Layer 4(topmost layer) of the TCP/IP reference model, in which network-aware, user-controlled software is implemented—for example, e-mail, file transfer utilities, and terminal access.
- Some of protocols that run at the application layer include File Transfer Protocol (FTP), Hypertext Transfer Protocol (HTTP), DNS, SMTP, POP, telnet, and similar protocols that can be implemented as utilities the user can interface with.

Principle of application layer protocol

- An application layer protocol defines how an application's processes, running on different end systems, pass messages to each other.
- In particular, an application layer protocol defines:
  i. The types of message exchanged. Example: request message and response message.
  ii. The syntax of the various message types. i.e. the field in the message and how the fields are represented.
  iii. The semantics of the fields. i.e. the meaning of the information in the fields.
  iv. Rules for determining when and how a process sends message and response to a message.
- Example: The application layer protocol of email application define how message are passed between servers, how the content of certain part of mail message (header) are to be interpreted etc.

1. World Wide Web and HTTP (Hypertext transfer protocol)

- The Hypertext Transfer Protocol (HTTP), the Web's application-layer protocol which is implemented in both client program (web browser) and server program (web server).
- The client program and server programs, executing on different end systems, talk to each other by exchanging HTTP messages.
- HTTP defines how Web clients (i.e., browsers) request Web pages from servers (i.e., Web servers) and how servers transfer Web pages to clients.
- HTTP defines the structure of these messages and how the client and server exchange the messages.

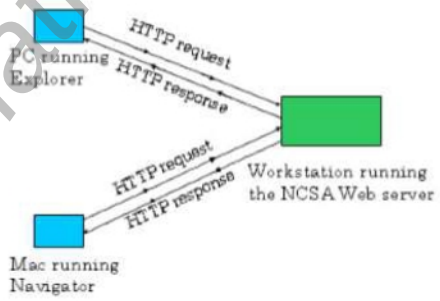


Figure 2.2-1: HTTP request-response behavior

- When a user requests a Web page (e.g., clicks on a hyperlink), the browser sends HTTP request messages for the objects (Jpeg image, audio file) in the page to the server. The server receives the requests and responds with HTTP response messages that contain the objects.
- For this HTTP client first initiate a TCP connection a TCP with the server. Once the connection is established, the browser and the server processes access TCP through their **socket** interfaces.
- The client sends HTTP request messages into its socket interface and receives HTTP response messages from its **socket** interface.
- Similarly, the HTTP server receives request messages from its socket interface and sends response messages into the socket interface.
- HTTP is also called **as stateless protocol** because HTTP server maintain no information about client.
- 80 is the default port for HTTP.

HTTP connection

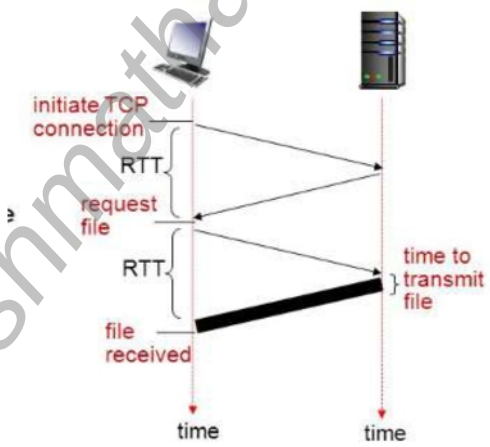
- There are two types of HTTP connection
    1. Persistent connection
    2. Non persistent connection

1. Non persistent connection

- The HTTP connection in which at most one object is sent over a TCP connection is called non persistent HTTP connection.
- Suppose a web page consists of some text and reference to 10 JPEG image. Then the steps of transferring this webpage from server to client using non persistent HTTP connection are
    i. HTTP client initiate TCP connection to HTTP server at port 80.
    ii. HTTP server waits for TCP connection at port 80 and accept it if any connection available.
    iii. HTTP client sends HTTP request message containing URL into TCP connection socket.
    iv. HTTP server receives request message then form response message containing requested object and sends message back to the socket.
    v. HTTP server **close** the connection.
    vi. HTTP client receive response message and get required web page. But while parsing, if finds 10 addition referenced JPEG image and text image.
    vii. So entire above steps are repeated for each of 10 JPEGE image object.
- Since, TCP connection is closed after the server sends each object, the connection does not persist for other object hence called as non-persistent connection.

Drawback

- Non persistent connection has mainly two drawback
    i. Overhead to server for establishing and maintaining connection for each requested object.
    ii. Require two RTT (Round Trip Time: time to move packet from client to server then back to client) per object. First RTT for establishing connection and second RTT for requesting and receiving object.



Response Time= 2 RTT + transmit time

2. Persistent connection

- In case of persistent connection, the server leave connection open after sending response so that subsequent HTTP message between same client and server are sent over same connection.
- So here, a client can send request as soon as it encounters additional referenced object.
- Hence it requires only one RTT for initiating TCP connection for all the referenced object.
- It is further categorized as
    i. Persistent connection with pipelining
        - Here the client can issue a request as soon as it encounters reference.

ii. Persistent connection without pipelining

- Here the client issues a new request only when the previous response has been received.

FTP (File transfer protocol)

- FTP (File Transfer Protocol) is a protocol for transferring a file from one host to another host. Figure 2.3-1: FTP moves files between local and remote file systems.

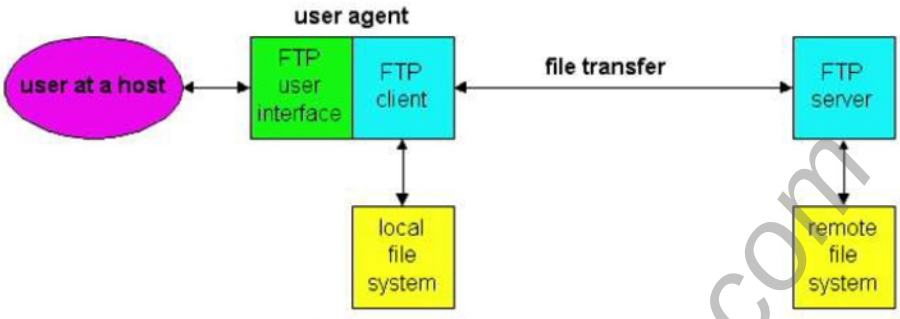


Figure 2.3-1: FTP moves files between local and remote file systems.

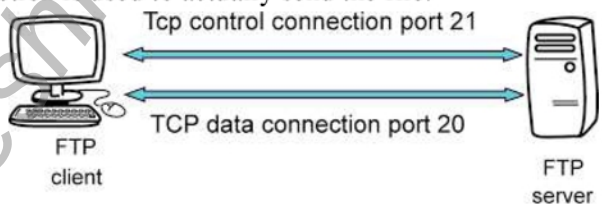
- An FTP client is an application that can issue FTP commands to an FTP server, while an FTP server is a service or daemon running on a server that responds to FTP commands from a client.
- For any user to access the remote computer, the user first provides the hostname of the remote host, which causes the FTP client process in the local host to establish a TCP connection with the FTP server process in the remote host. The user then provides the user identification and password, which get sent over the TCP connection as part of FTP commands. Once the server has authorized the user, the user copies one or more files stored in the local file system into the remote file system (or vice versa).
- Get "filename" FTP command is used to get file from server to client. Put "filename" FTP command is used to write new file into server from client.

Similarity Between FTP and HTTP

- Both are file transfer protocol
- Both run on top of TCP i.e. transport layer protocol which is connection oriented and reliable data transfer protocol.

Difference between FTP and HTTP

3. HTTP uses a single TCP connection but FTP uses two parallel TCP connection (control connection and data connection) to transfer a file. Control connection is used for sending control information like username, password, command to get and put file etc. The data connection is used to actually send the file.



4. HTTP use 80 as port number but FTP uses 20 and 21 as port number.
5. FTP server maintain state about user but HTTP is stateless i.e. it doesn't have to keep track of users.

DNS

- IP address are tough for human to remember and impossible to guess. Domain Name System are usually used to translate a hostname or Domain name into an IP address.
- The DNS is a way that the internet domain names are translated into internet protocol (IP) address.
- DNS automatically converts the host name(pu.edu.np) we type in our web browser address bar to the IP address 202.2.106.77 of webserver hosting the site.
- DNS is an application layer protocol.

Working principle

- DNS implement a distributed database to store the hostname and address information for all public host on the internet as hierarchy of many name servers.
- When a client like web browser issues a request involving the internet hostname, a piece of software called DNS resolver first contact a DNS server to determine IP address.
- If the server doesn't contain the needed mapping, it will forward the request to next higher level in the hierarchy.
- After several forwarding, the IP address for the given host eventually arrive at the resolver, that in turn complete the request over internet protocol(IP) address.

Note:

Dos Command: nslookup

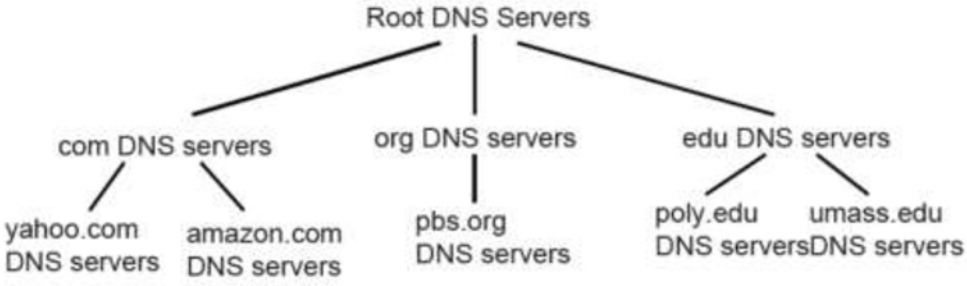
```
C:\Users\Bheeshma>nslookup google.com
Server:  UnKnown
Address:  192.168.1.1

Non-authoritative answer:
Name:    google.com
Addresses:  2404:6800:4002:803::200e
          216.58.221.46
```

Here the firs two line indicates the server name and ip address of DNS server which the request was directed.

Services provided by DNS

1. Translation of hostname to IP address
    - People prefer mnemonics that IP address.
    - DNS translates the hostname to IP address.
2. Host aliasing
    - A host with a complicated host name can have one or more alias name.
    - For example, a hostname such as relay1.west-coast.enterprise.com could have, say, two aliases such as enterprise.com and www.enterprise.com.
3. Mail server aliasing
    - It can be used to provides simple mnemonics email address instead of using more complicated email address.
    - Example: canonical host name = relay1.west-coast.hotmail.com
      Alias= bob@hotmail.com.
4. Load distribution
    - DNS is also being used to perform load distribution among replicated servers, such as replicated Web servers.
    - Busy sites, such as cnn.com, are replicated over multiple servers, with each server running on a different end system, and having a different IP address.
    - For replicated Web servers, a set of IP addresses is thus associated with one canonical hostname.
    - The DNS database contains this set of IP addresses.
    - When clients make a DNS query for a name mapped to a set of addresses, the server responds with the entire set of IP addresses, but rotates the ordering of the addresses within each reply.

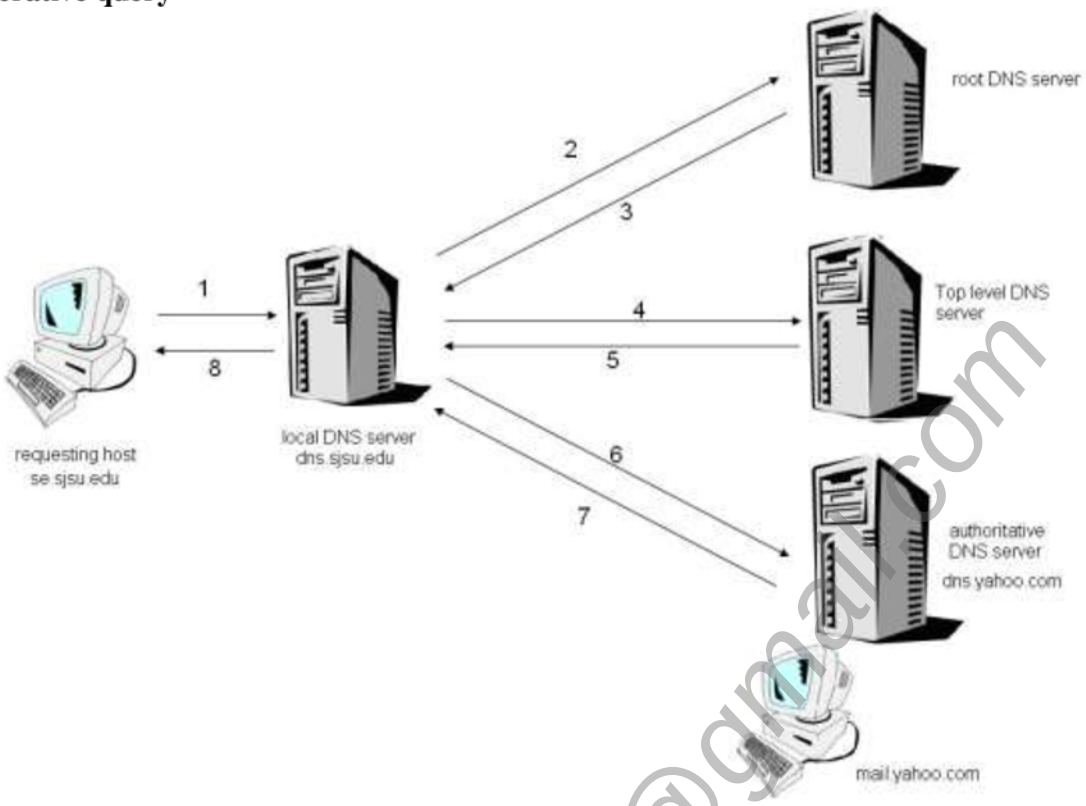
A distributed Hierarchical database



- DNS implement distributed database to store the host name and the address information for all public host on the internet as a hierarchy of many name servers.
- A single server doesn't provide mapping for all host in the internet.
- There are three classes of DNS server named as
  1. Root DNS server
  2. Top level DNS server
  3. Authoritative DNS server

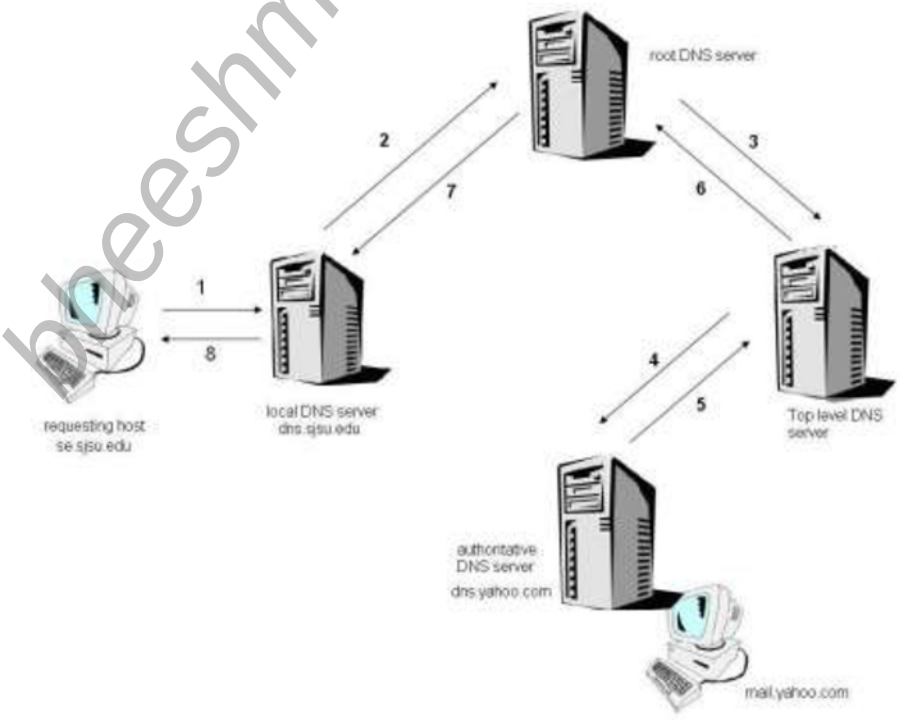
DNS Queries for Name Resolution

i. Iterative query



- Suppose a DNS client want to determine IP address for www.amazon.com
- First client queries one of the root server.
- The root server returns the IP address for TLD server for top level domain .com
- The client then contacts one of these TLD server, which returns the IP address for authoritative server for amazon.com
- Finally, the client contacts one of the authoritative server for amazon.com which returns the ip address for the host name www.amazon.com.

ii. Recursive Query



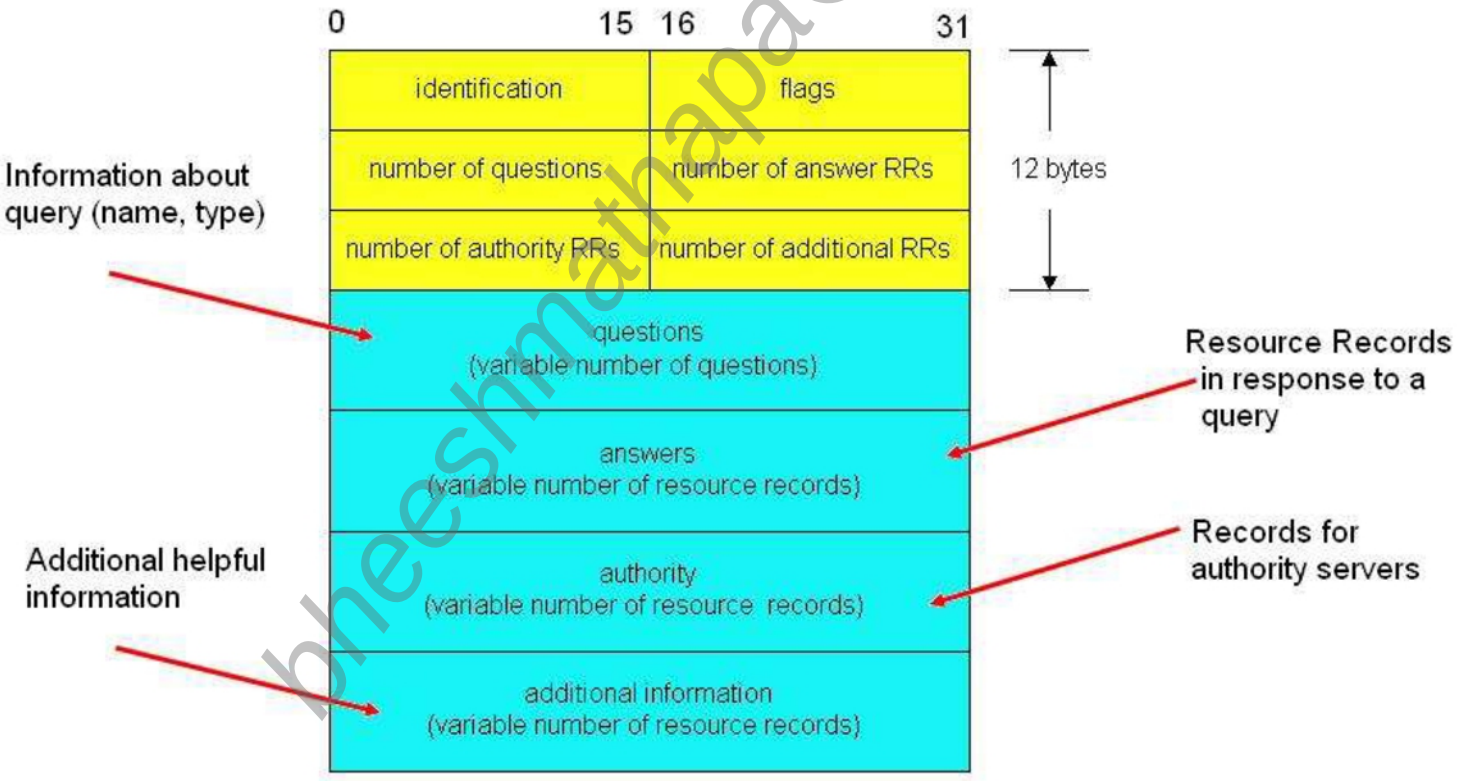
Explain yourself……..

DNS record

- The DNS server includes resource record that provide hostname to IP address mapping.
- A resource record has four fields: (Name, Value, Type, TTL)
- TTL is the time to live the resource record.
- Name and Value depends on type
- If type=A, then name is host name and value is IP address for hostname. Example:
  (relay1.bar.foo.com, 192.168.1.5, A) is a type of A Record
- If Type=NS, then Name is a domain (such as foo.com) and Value is the hostname of a server that knows how to obtain the IP addresses for hosts in the domain. This record is used to route DNS queries further along in the query chain. Example
  (foo.com, dns.foo.com, NS) is a Type NS record.
- If Type=CNAME, then Value is a canonical hostname for the alias hostname Name. Example
  (foo.com, relay1.bar.foo.com, CNAME) is a CNAME record.
- If Type=MX, then Value is a hostname of a mail server that has an alias hostname Name. Example,
  (foo. com. mail.bar.foo.com, MX) is an MX record.
  MX records allow the hostnames of mail servers to have simple aliases

DNS Message

These are the only two kinds of DNS messages. Both request and reply messages have the same format as shown in figure below.



- **The first 12 bytes is the header section, which has a number of fields.**
- **The first field is a 16-bit number that identifies the query.** This identifier is copied into the reply message to a query, allowing the client to match received replies with sent queries.

- There are a number of flags in the flag field (16 bit) each of which stores 1 and 0. For example, **A one-bit query reques/reply flag indicates whether the message is a request (0) or a reply (1). A one-bit authoritative flag is set in a reply message when a name server is an authoritative server for a queried name.** A one-bit recursion-desired flag is set when a client (host or name server) desires that the name server to perform recursion when it doesn't have the record. A one-bit recursion available field is set in a reply if the name server supports recursion.
- **In the header, there are also four "number of" fields.** These fields indicate the number of occurrences of the four types of "data" sections that follow the header.
- **The question section contains information about the query that is being made. This section includes**
  - **i. A name field that contains the name that is being queried,**
  - **ii. A type field that indicates the type of question being asked about the name (e.g., a host address associated with a name - type "A", or the mail server for a name - type "MX").**

- **In a reply from a name server, the answer section contains the resource records i.e. DNS record for the name that was originally queried**. Recall that in each resource record there is the Type (e.g., A, NS, CSNAME and MX), the Value and the TTL. A reply can return multiple RRs in the answer, since a hostname can have multiple IP addresses (e.g., for replicated Web servers, as discussed earlier in this section).
- **The authority section contains records of other authoritative servers.**
- **The additional section contains other "helpful" records. For example, the answer field in a reply to an MX query will contain the hostname of a mail server associated with the alias name Name eg. Foo.com. The additional section will contain a Type A record providing the IP address for the canonical hostname of the mail server**.