Chapter 2: Application Layer

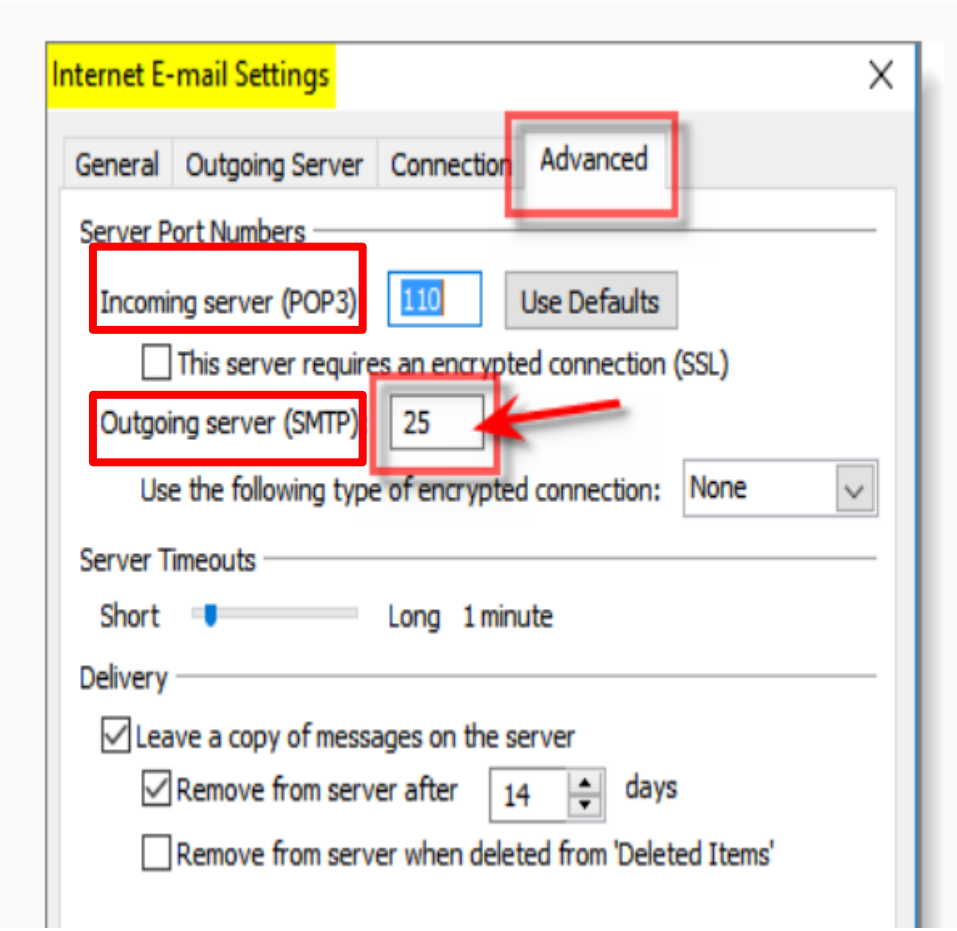
Chapter 2: Electronic Mail and DNS

Instructor: HOU, Fen

2025

Application Layer: Electronic Mail

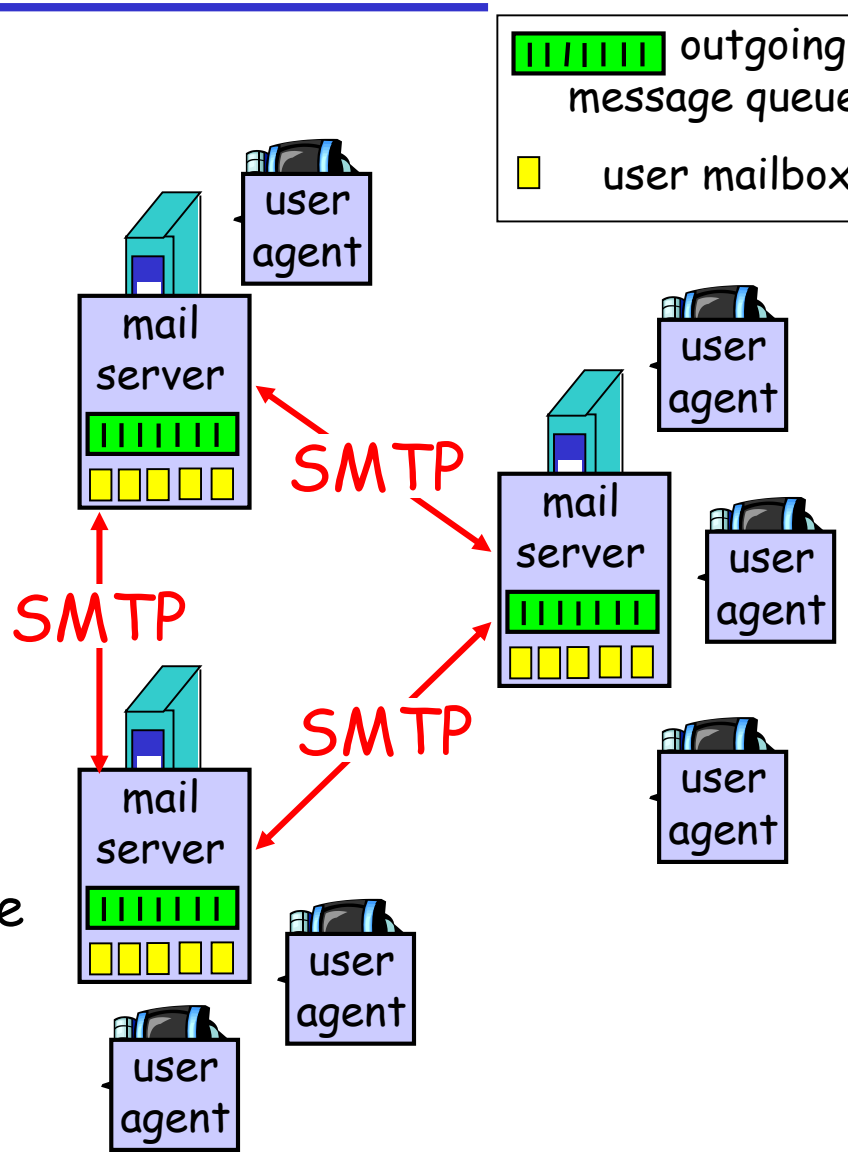
- One of the most popular and important applications in the Internet.
- E-mail is an asynchronous communication application.
- E-mail is fast, inexpensive, and easy to distribute



Application Layer: Electronic Mail

Three major components:

- user agents
  - It is called as "mail reader"
  - composing, editing, reading mail messages
  - Allows messages to be sent/retrieved to and from mail server
  - e.g., Outlook, Foxmail, etc
- mail servers: form the core of the e-mail infrastructure.
  - outgoing and incoming messages are stored on server.
  - user mailbox contains incoming messages for user
  - outgoing message queue keeps mail messages to be sent.



Application Layer: Electronic Mail

- Simple Mail Transfer Protocol: SMTP
  - A protocol used to send and receive email messages between mail servers.
    - "client" side: executes on the sender's mail server
    - "server" side: executes on the recipient's mail server
    - Both the client and server sides of SMPT run on every mail server
  - SMTP uses TCP as its underlying transport protocol to provide the reliable data transfer service.

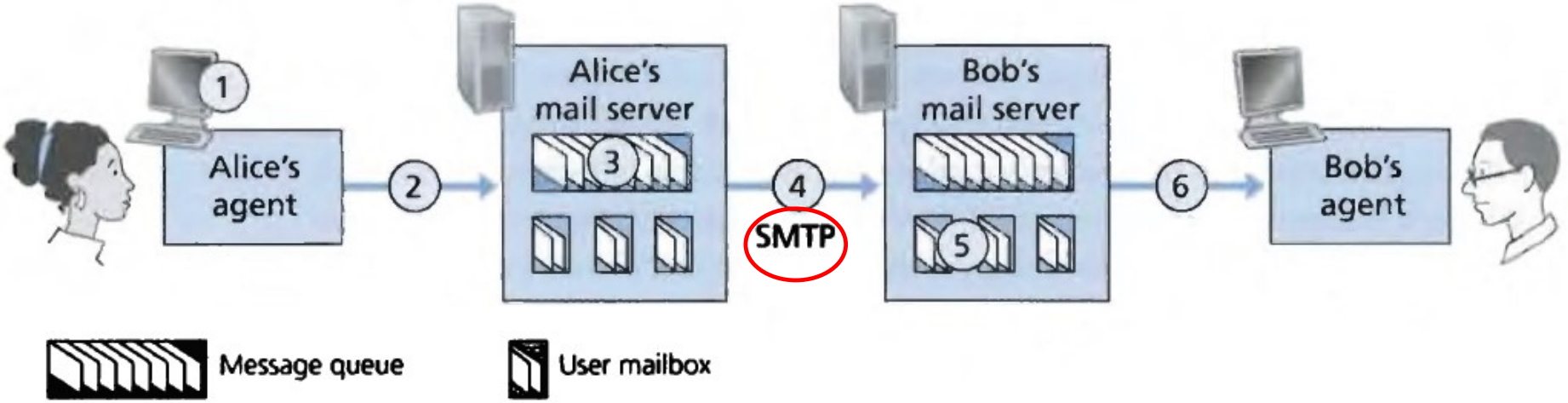


Figure 2.17 ♦ Alice sends a message to Bob

Scenario: Alice sends message to Bob

1) Alice uses her user agent (UA) to compose message.
2) Alice's UA sends message to her mail server; message placed in message queue.
3) Client side of SMTP opens TCP connection with Bob's mail server.
4) SMTP client sends Alice's message over the TCP connection.
5) Bob's mail server places the message in Bob's mailbox.
6) Bob invokes his/her user agent to read message.

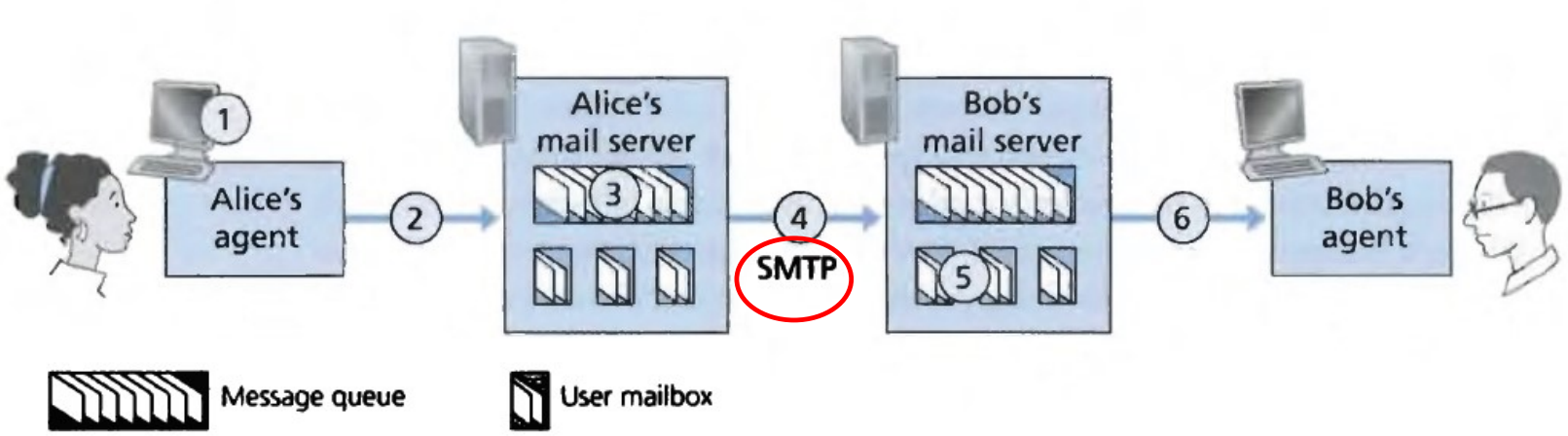
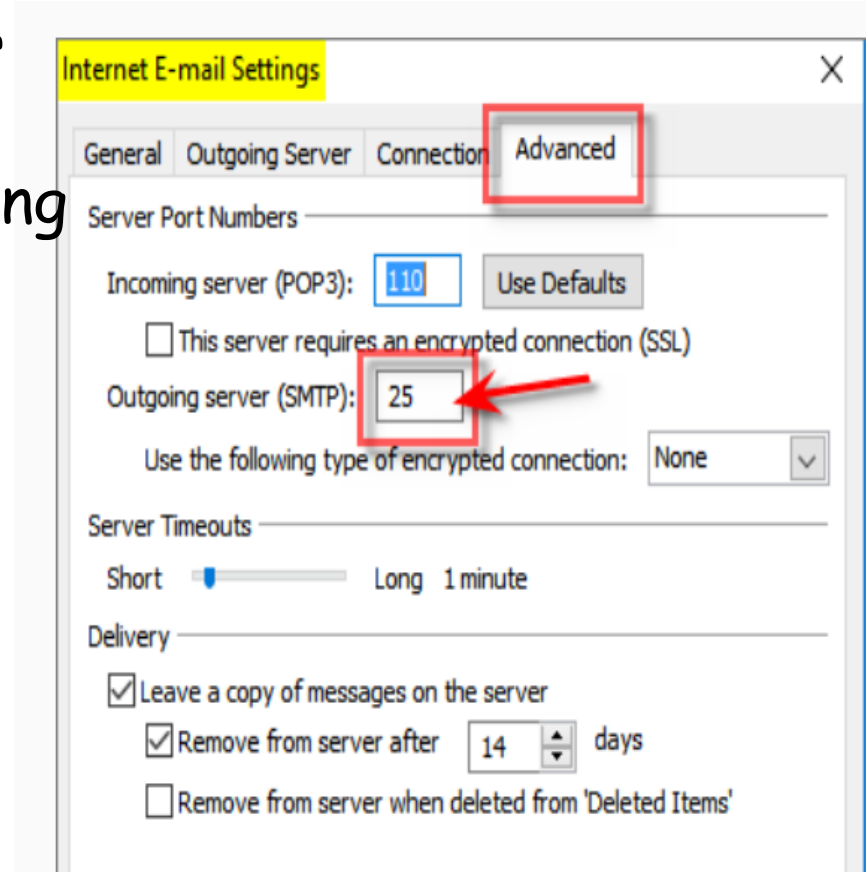


Figure 2.17 ♦ Alice sends a message to Bob

Electronic Mail: SMTP

- Client side of SMTP mail server establishes a TCP connection to the recipient's SMTP server using Port 25
- TCP three phases in message transfer
  - handshaking (greeting)
  - transfer of messages
  - closure
- command/response interaction
  - commands: Client issues such as HELO, MAIL FROM, etc.
  - response: status code and phrase

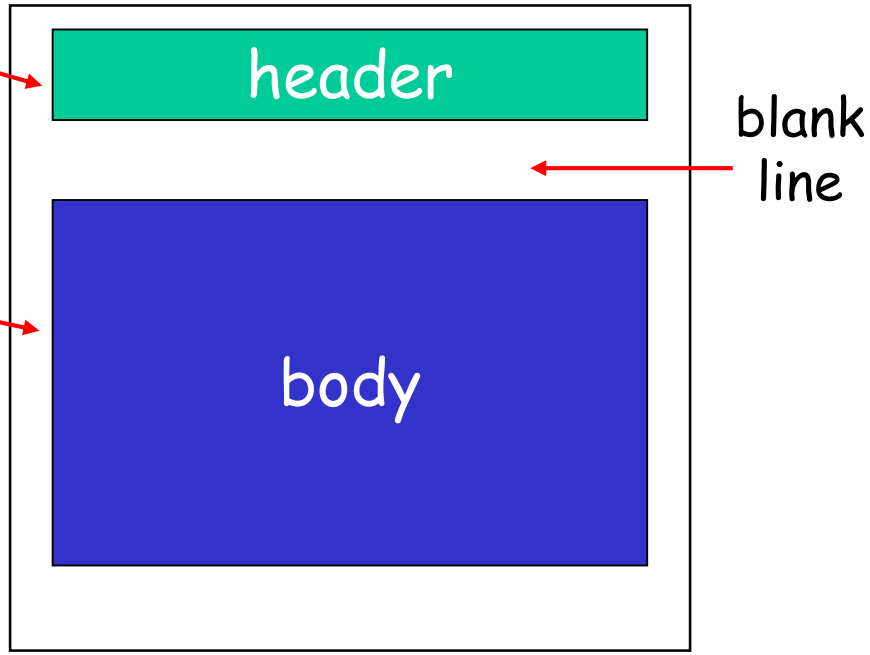


Mail message format

SMTP: protocol for exchanging email messages

Text message format:

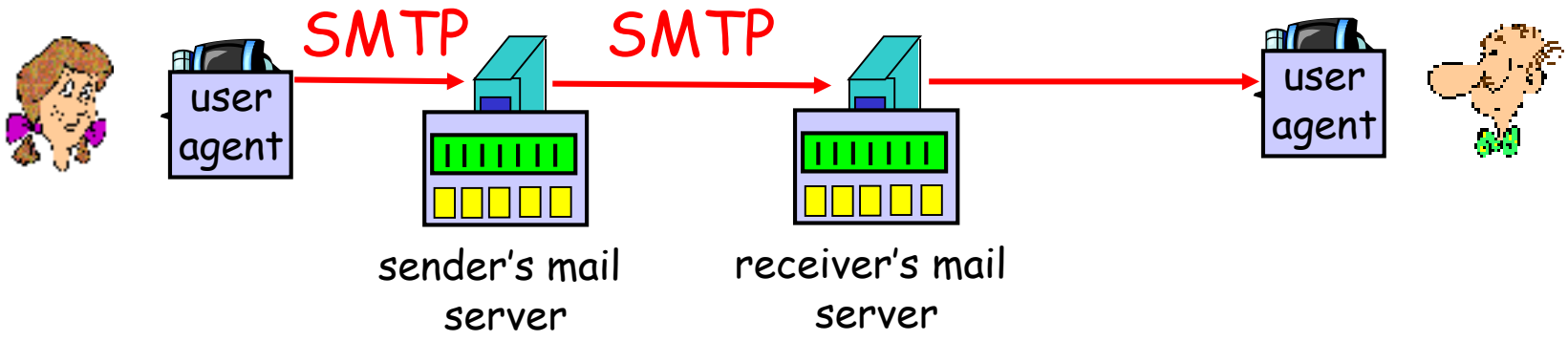
- header lines, e.g.,
  - To:
  - From:
  - Subject:
- body
  - the "message", ASCII characters only. That is, SMTP strict that all messages including the head and body of email are sent in 7-bit ASCII code



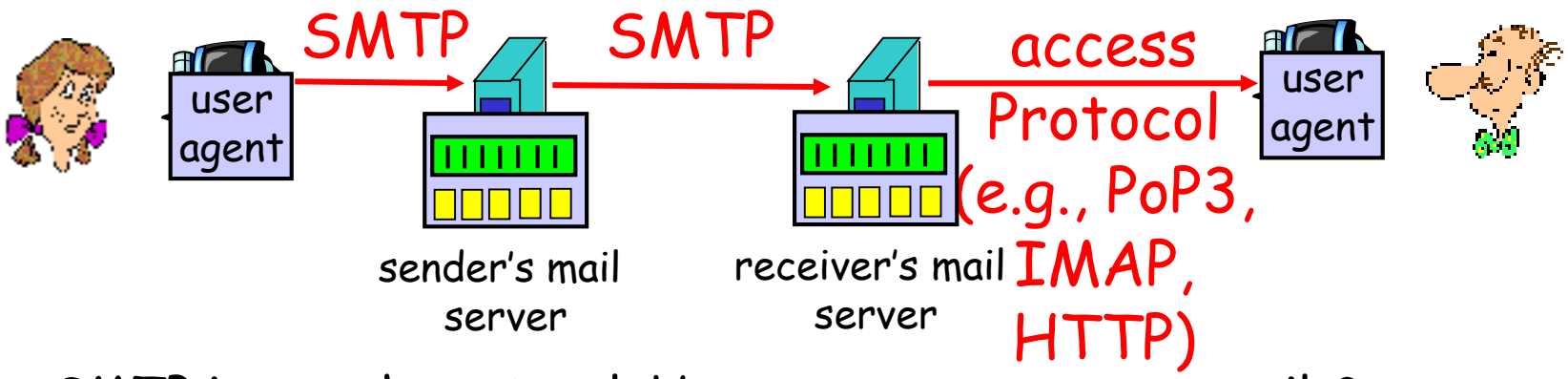
SMTP is a kind of Pull Portocol

Difference

- HTTP: pull protocol. To pull the information from the server.
- SMTP: push protocol. To push the file to the receiving mail server.



Mail access protocols



- SMTP is a push protocol. How can a user access emails?
- Mail access protocol: retrieval from mail server to local computer
  - POP: Post Office Protocol
    - Authorization (agent----server) and download
  - IMAP: Internet Mail Access Protocol
    - More features
    - Manipulation of stored messages on server
  - HTTP: Web-based Email, such as Gmail.

Add Account

POP and IMAP Account Settings

Enter the mail server settings for your account.

User Information

Your Name: Brad Johnson

Email Address: brad@myemail.com

Server Information

Account Type: POP3

Incoming mail server: mail.papamail.net

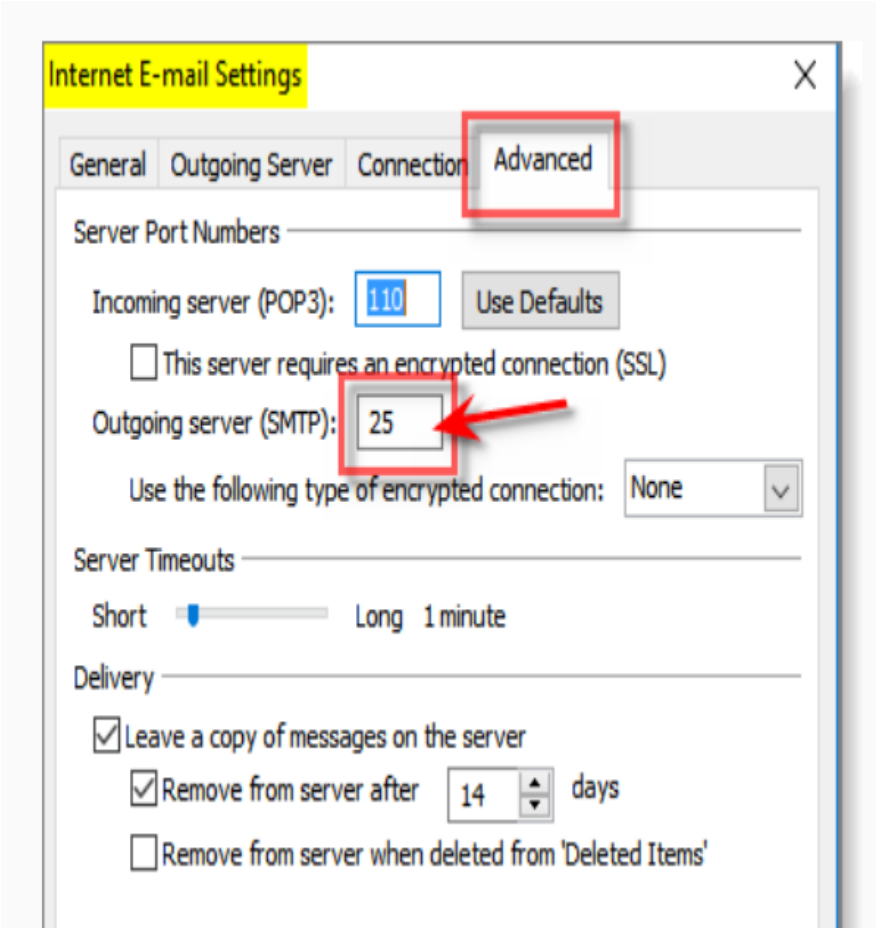
Outgoing mail server (SMTP): mail.papamail.net

Logon Information

User Name: brad@myemail.com

Password: **********

☑ Remember password



Comparison of HTTP and SMTP

Common characteristics

- Both are client-and-server Mode
- Both use the reliable data transfer service of TCP
- Include persistent connection

Difference

- HTTP: pull protocol. To pull the information from the server.
- SMTP: push protocol. To push the file to the receiving mail server.
- SMTP has 7-bit ASCII restriction. HTTP does not have this kind of restriction.
- HTTP: each object encapsulated in its own response message
- SMTP: place all of message's objects into one message

Summary: E-mail System

- Three main component of E-mail system
  - Users agents, mail server, and SMTP
- SMTP and the comparison with HTTP.
- SMTP mail message format
- Mail access protocols
  - PoP, IMAP, HTTP

Chapter 2: Application Layer

Our goals:

- conceptual, implementation aspects of network applications
  - transport-layer service models
  - client-server paradigm
  - peer-to-peer paradigm
- learn about protocols by examining popular application-level protocols
  - HTTP
  - SMTP / POP3 / IMAP
  - DNS (Domain Name System)

DNS: Domain Name System

The ways to identify each person

- Given and family name
- ID card number
- Passport number

The ways to identify a host

- hostname, e.g., www.umac.mo, www.google.com
  - Easy to be remembered, human beings prefer
  - Provide little information about the location of this host
  - Difficult to be processed by router
- IP address (IPV4: 4 bytes (i.e.,32 bit) such as 121.7.106.83)
  - Fixed length, hierarchical structure,
  - Easy to process by router, routers prefer
  - Different to be remembered

DNS: Domain Name System

Internet hosts: two ways to identify a host

```
C:\>ipconfig/all

Windows IP Configuration

   Host Name . . . . . . . . . . . . : PC09285
   Primary Dns Suffix  . . . . . . . : pclan.umac.mo
   Node Type . . . . . . . . . . . . : Hybrid
   IP Routing Enabled. . . . . . . . : No
   WINS Proxy Enabled. . . . . . . . : No
   DNS Suffix Search List. . . . . . : pclan.umac.mo
                                       umac.mo

Ethernet adapter Local Area Connection:

   Connection-specific DNS Suffix  . : umac.mo
   Description . . . . . . . . . . . : Intel(R) Ethernet Connection I217-LM
   Physical Address. . . . . . . . . : 4C-CC-6A-10-E4-D3
   DHCP Enabled. . . . . . . . . . . : Yes
   Autoconfiguration Enabled . . . . : Yes
   Link-local IPv6 Address . . . . . : fe80::cd21:6fb9:740f:1ea2%11(Preferred)
   IPv4 Address. . . . . . . . . . . : 161.64.202.1(Preferred)
   Subnet Mask . . . . . . . . . . . : 255.255.255.0
   Lease Obtained. . . . . . . . . . : Monday, September 02, 2019 13:43:15
   Lease Expires . . . . . . . . . . : Wednesday, September 04, 2019 18:44:00
   Default Gateway . . . . . . . . . : 161.64.202.254
   DHCP Server . . . . . . . . . . . : 161.64.186.231
   DHCPv6 IAID . . . . . . . . . . . : 239914090
   DHCPv6 Client DUID. . . . . . . . : 00-01-00-01-20-2B-0C-64-4C-CC-6A-10-E4-D3

   DNS Servers . . . . . . . . . . . : 161.64.36.64
                                       161.64.36.65
                                       161.64.36.66
   NetBIOS over Tcpip. . . . . . . . : Enabled
```

DNS: Domain Name System

- DNS is used to conduct the translation between hostname and IP address
  - Applications running at a host need to translate a hostname to its IP address (identifier is IP address + port number)
  - DNS is used by most internetworking software, such as web browsers and electronic mail programs, to locate servers and to resolve, or map, a user-friendly name of a computer to its IP address.
  - Host aliasing: translation from alias name (别名) to canonical(正规的、权威的）name.
    - Two types of hostnames
      - Canonical name: formal hostname, e.g., relay1.west-cost.google.com
      - Alias name: easy to remember, e.g., google.com
  - Mail server aliasing: translation from alias name to canonical name of a mail server
    - Two types of mail servers
      - Canonical name: e.g., relay1.west-coast.hotmail.com
      - Alias name: e.g., imap.gmail.com, bob@gmail.com

DNS (Domain Name System)

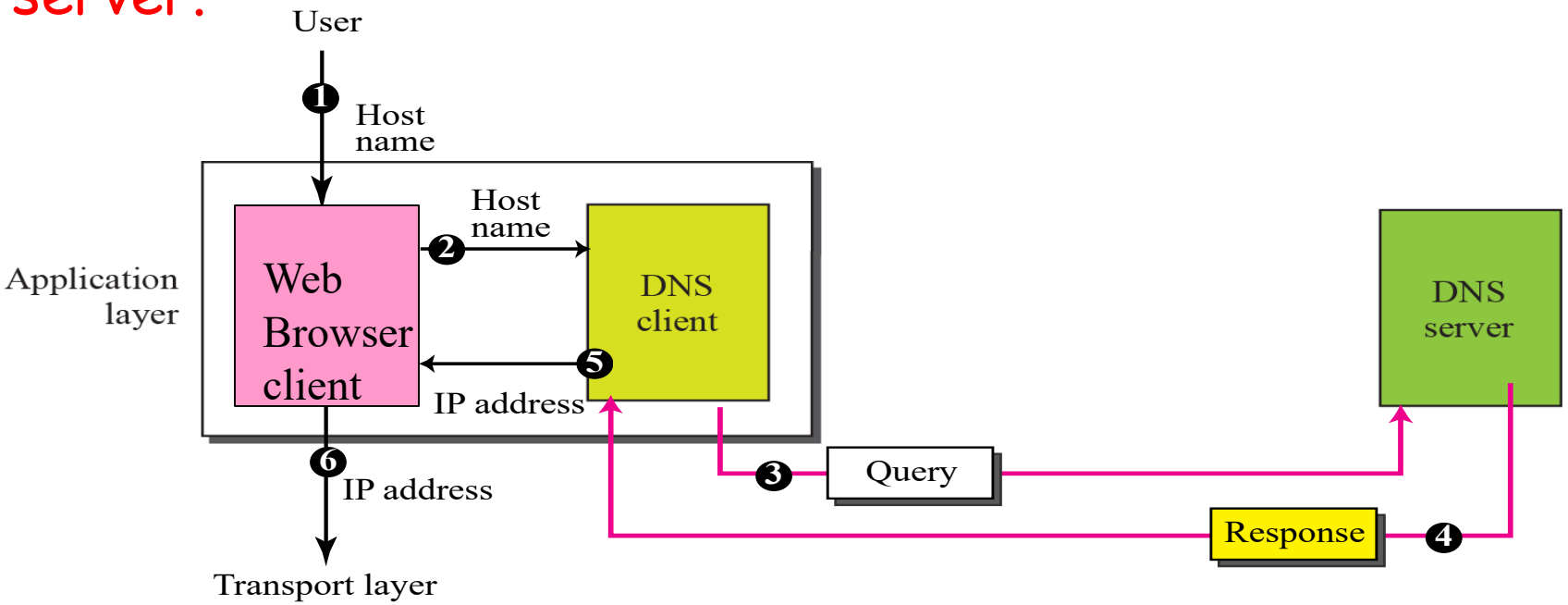
- Load distribution
  - Replicated web servers for busy sites: a set of IP addresses for one canonical name
    - Busy sites are replicated over multiple web servers.
    - Each replicated server run on a different end system and have a different IP address.
    - But all replicated servers are associated with one canonical name.
  - DNS rotation distributes the traffic among the replicated servers.
    - When clients make a DNS query for a busy site, the server responds with the entire set of IP addresses, but rotates the ordering of the addresses within each reply。

Why DNS?

- Convenience: User-friendly names are easier for people to remember than numerical IP addresses.
- Consistency: IP addresses may change, but the server names can remain constant.
- Simplicity for load distribution: With DNS, it is simple to distribute the traffic among multiple servers for a busy website.

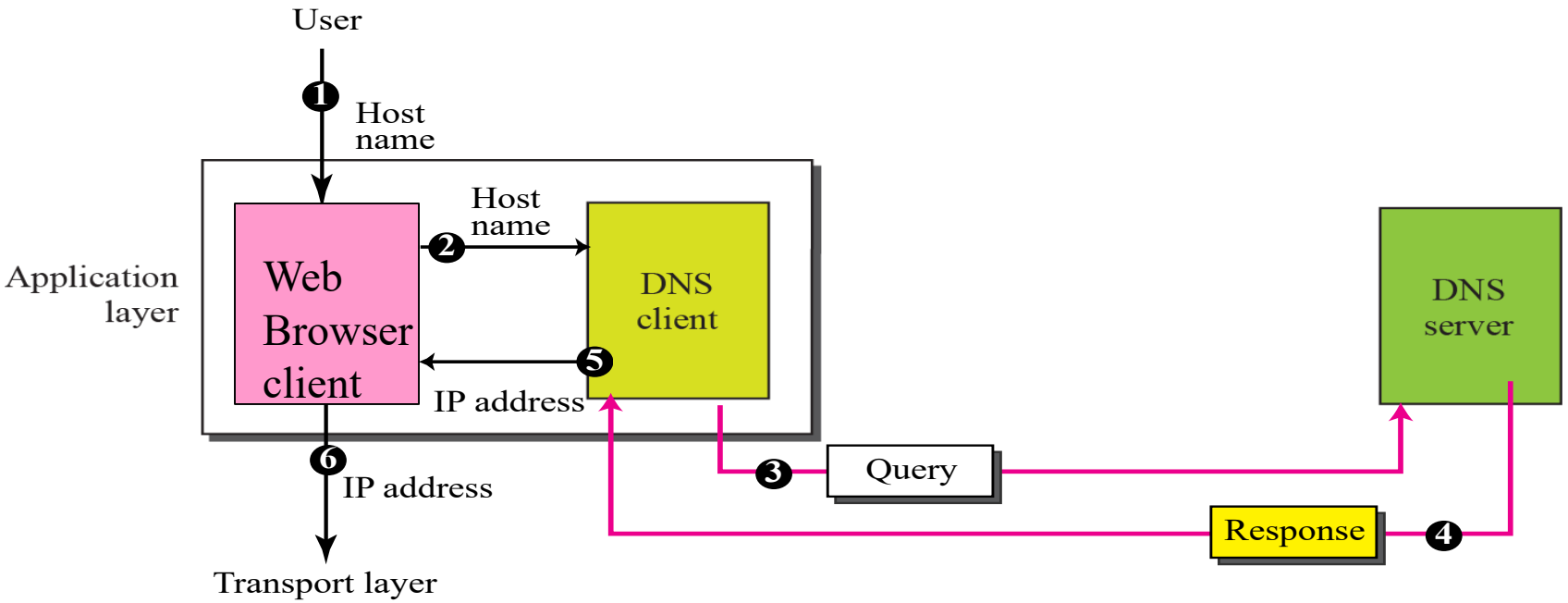
How Does DNS Work?

- Before a network application (e.g., web browser) sends a message to its receiving host, it should know the IP address of the receiving host.
- The application invokes the client side of DNS, specifying the hostname that needs to be translated.
- Then, DNS client sends a query message to a DNS server.



How Does DNS Work?

- After a time delay, DNS client receives a DNS reply message including the IP address for this hostname.
- The IP address is then passed to the invoking application.
- Once the browser receives the IP address from DNS, it can initiate a TCP connection to the HTTP server process (IP address + port 80)



How to Implement DNS (Domain Name System)

Centralized DNS? One DNS server contains all the mappings between the hostnames and IP addresses

- single point of failure
- Limit the traffic volume
- distant centralized database---a single DNS server cannot be "close to" all end systems, which will leads to long delay for some end systems.
- Maintenance: this centralized database will be huge and has to be updated frequently to account for every new host.

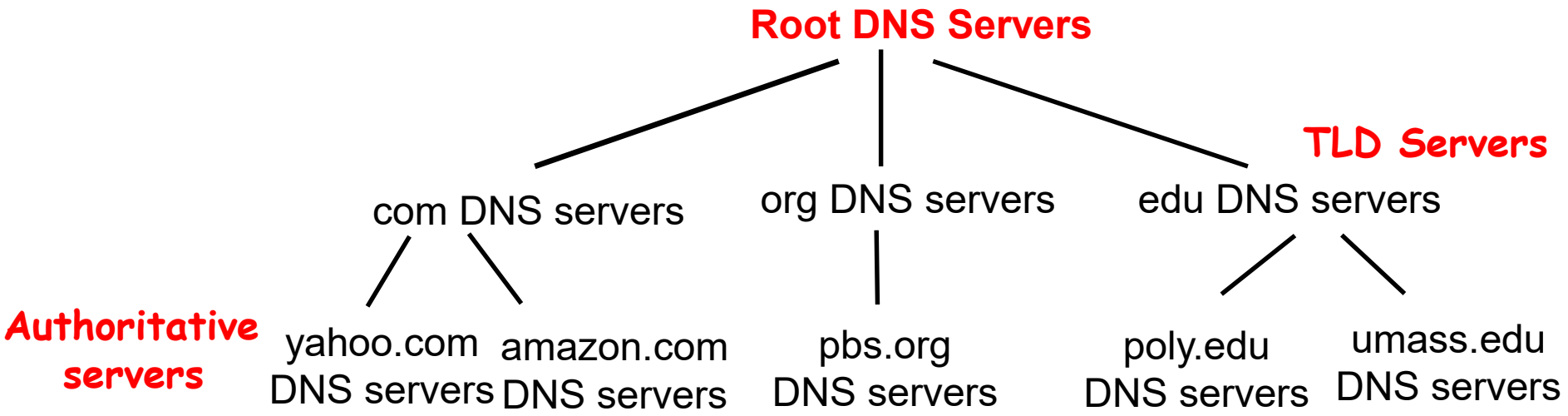
Distributed DNS: DNS consists of a large number of DNS servers distributed around the globe and specify how the DNS server and querying hosts communicate.

DNS (Domain Name System)

- The DNS is
  - A distributed database system implemented in the hierarchy of DNS servers. (scalability and reliability)
  - An application-layer protocol that allows hosts to query the distributed database to resolve names(address/name translation)
  - DNS uses the UDP as its underlying transport layer protocol.

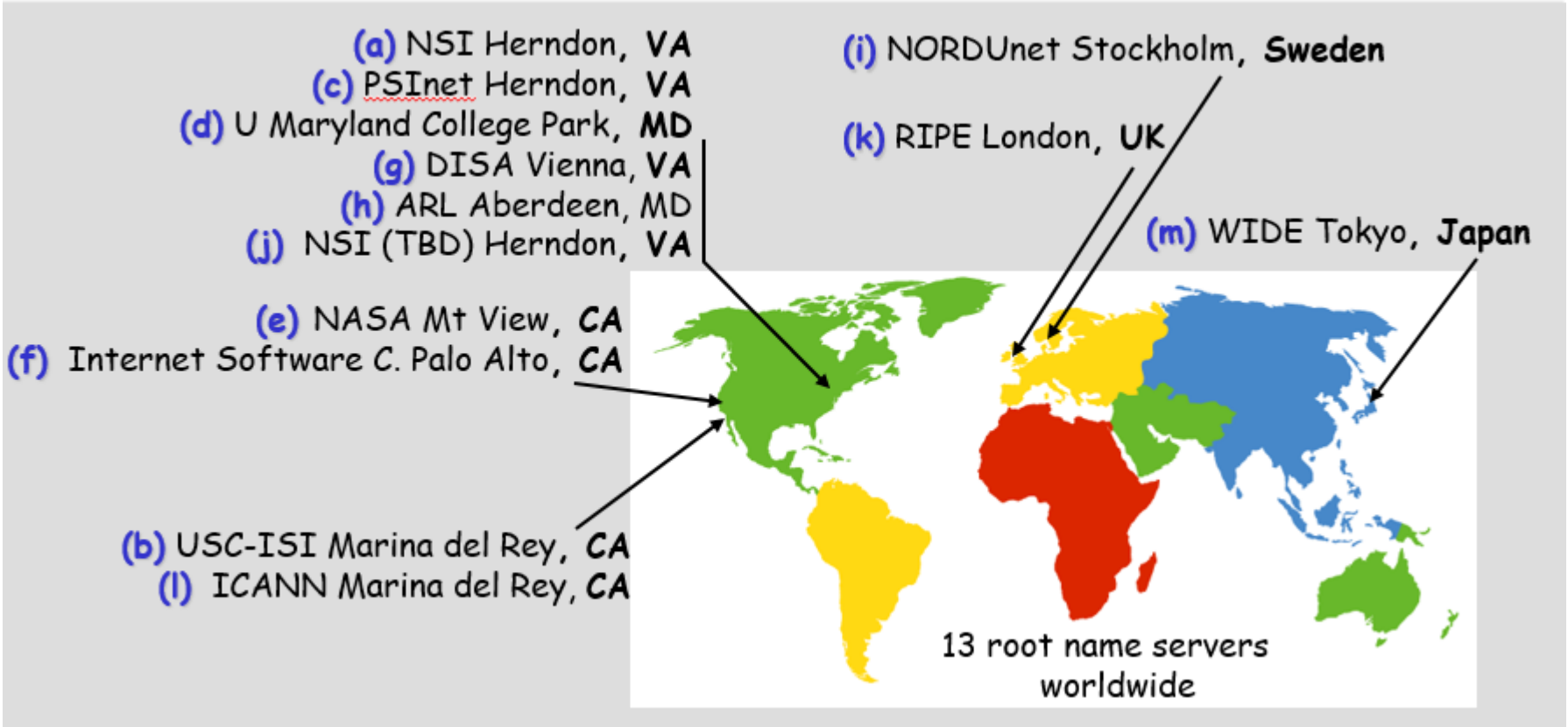
DNS: Distributed, Hierarchical Database

Three classes of DNS servers (Root DNS servers, TLD (top-level Domain) servers, and authoritative servers) organize in hierarchy.



Root DNS server

- In the Internet, there are 13 root DNS servers labeled A throughput M in the following figures, most of which are located in North America.
- Root DNS servers typically do not contain hostname to IP mappings; they contain mappings for locating top-level domain (TLD) servers.



TLD DNS Server

- Top-level domain (TLD) servers are responsible for com, org, net, edu, int. etc, and all top-level country domains uk, fr, ca, cn,jp.
  - Com: Commercial.
  - Edu: educational institute
  - Int: international organization
  - Ac:academic research institute
  - Gov: US federal government
  - Net:network provider
  - Jp: Japan, uk:England; ca: Canada; cn: China; etc.
- TLD servers typically do not contain hostname to IP mappings; they contain mappings for locating authoritative servers.

Authoritative DNS Server

- Authoritative DNS server: It is also called as 2nd-level domain DNS server. It provides mapping from hostname to IP address for organization's servers (e.g., Web server, mail server).
  - Store the mapping from hostname to IP address.
  - Each organization with publicly accessible hosts (e.g., Web servers or mail servers) on the Internet must have records that map the name of these hosts to IP addresses.
  - Each organization can implement its own authoritative DNS server to hold these records, it also can have these records stored in an authoritative DNS server of some service provider.

Local DNS Server

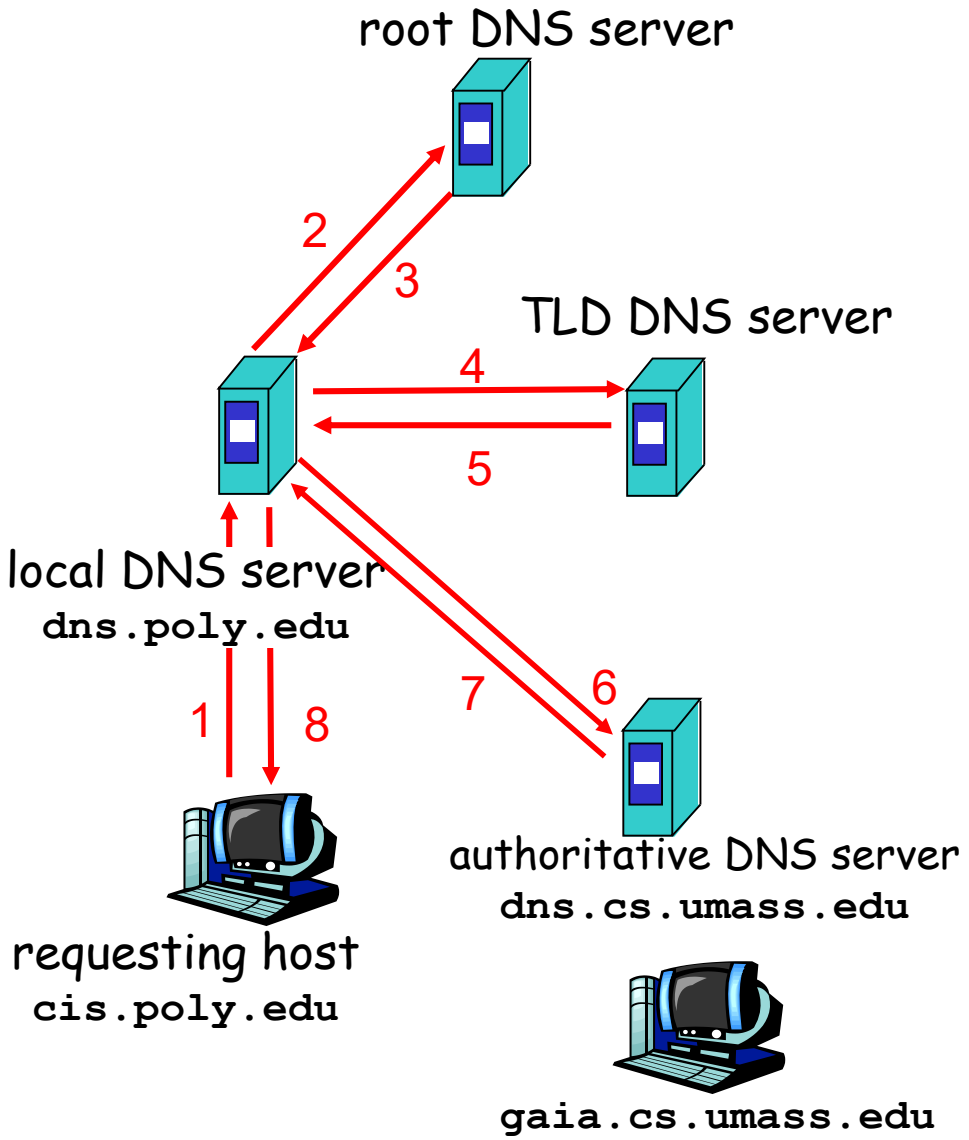
- A local DNS server is also called a default name server. When a host makes a DNS query, query is sent to its local DNS server
  - A local DNS server does not strictly belong to the hierarchy of servers.
  - A local DNS server acts as a proxy, forwards query into hierarchy of servers.
  - Reduces lookup latency for commonly searched hostnames
- The IP address of your local DNS server can be easily determined by using the command ipconfig /all in Windows OS

Local DNS Server

- Each ISP (residential ISP, company, university, etc) has a local DNS server.
- When a host connects to Internet through a ISP, the ISP provides the host with the IP addresses of its local DNS server.
- A host's local DNS server is typically "close to" the host.
  - For an institutional ISP, at the same LAN
  - For a residential ISP, no more than a few routers.

Interaction of the various DNS servers

- Host cis.poly.edu wants to know the IP address for gaia.cs.umass.edu
- DNS lookup process:
  - The host sends a DNS query msg to its local DNS server
  - Local DNS server forwards the query msg to a root DNS server.
  - The root DNS server returns a list of IP addresses for TLD servers responsible for edu.

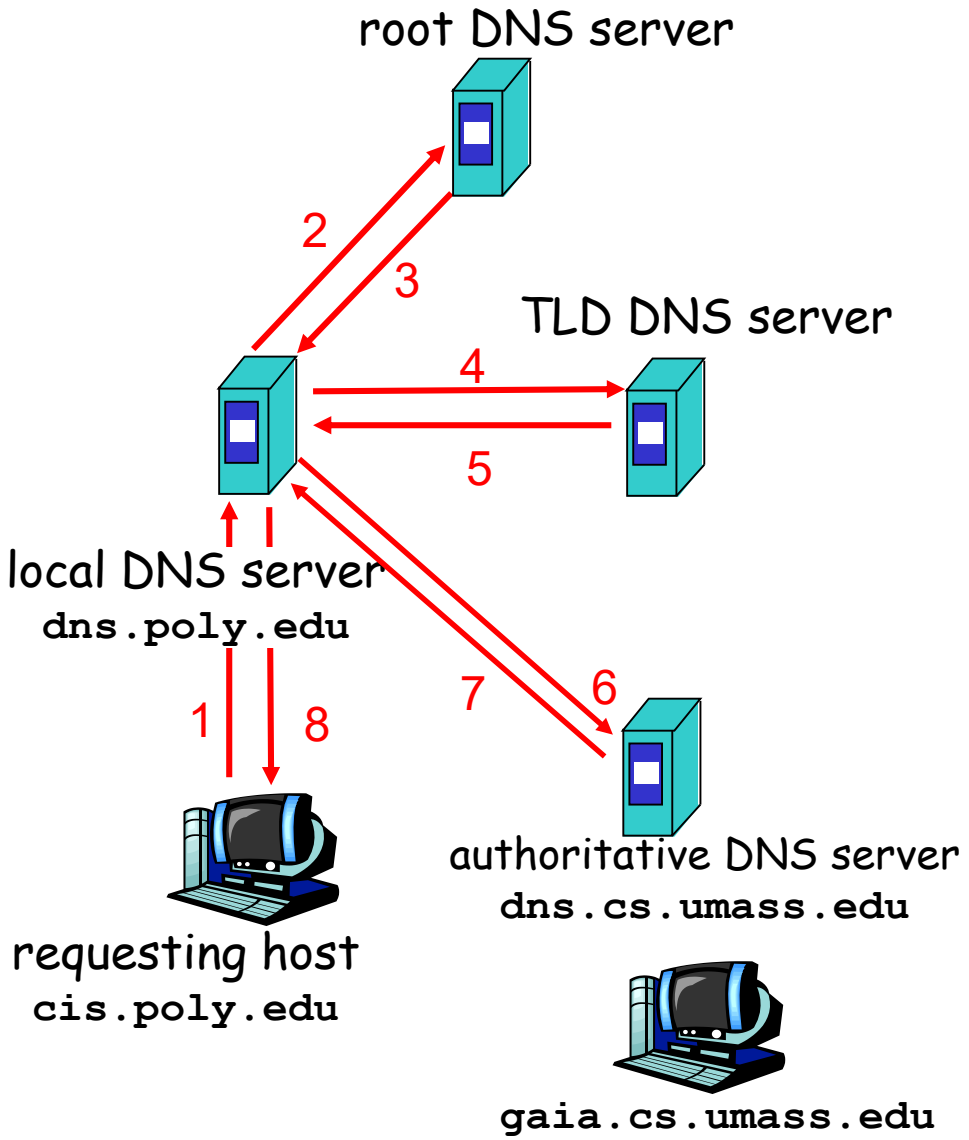


Root DNS servers

- Root DNS server is contacted by local DNS server that can not resolve name
- Root DNS server:
  - gets name mapping: Root DNS server knows the intermediate name server. That is, it knows who to contact such that the authoritative name server can be found.
  - returns mapping to local DNS server

Interaction of the various DNS servers

- The local DNS server resends the query message to one of these TLD servers
- The TLD server takes the note of the umass.edu suffix and responds with the IP address of the Authoritative DNS server for the University of Massachusetts.
- The local DNS server resends the query message to this authoritative DNS server and gets response with the IP address of gaia.cs.umass.edu.



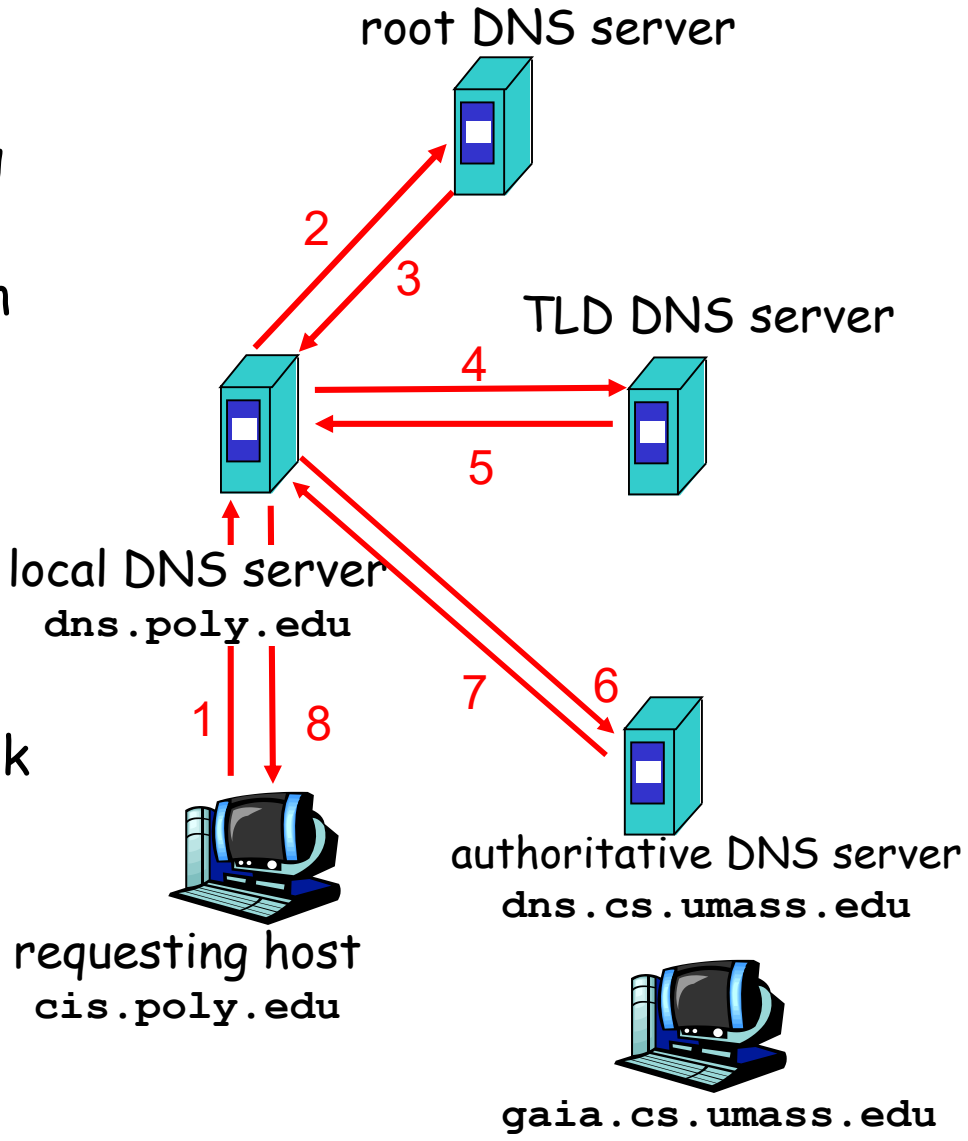
Recursive and iterative queries

Recursive（递归） query:

- The query from the requesting host to the local DNS server.
- puts burden of name resolution on contacted name server (i.e. local DNS server)

Iterative（迭代） query:

- contacted server replies with name of server to contact
- "I don't know this name, but ask this server"
- These replies are directly returned to local DNS server

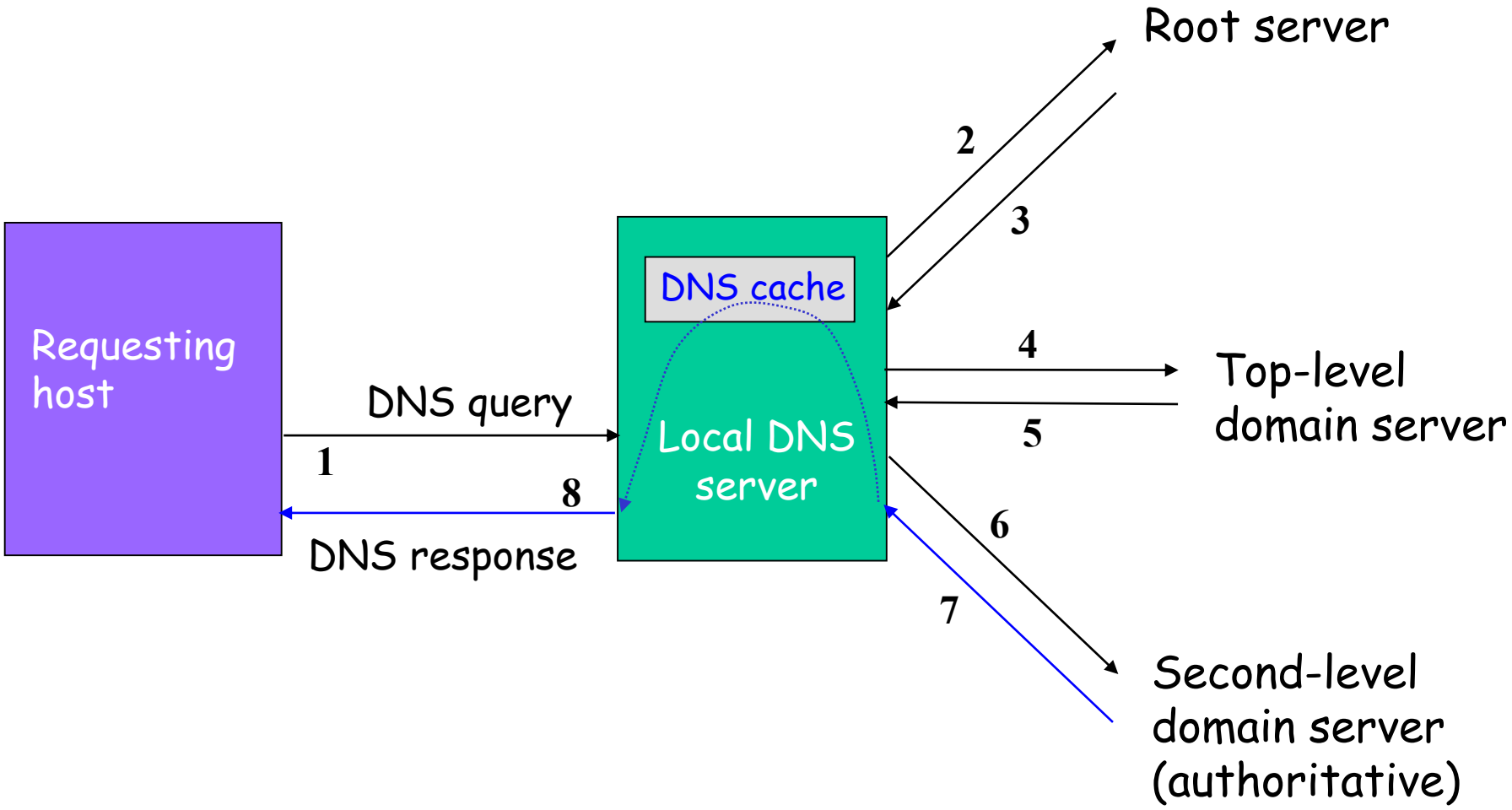


DNS: caching and updating records

- DNS caching is to improve delay performance by reducing the number of DNS messages.
- Once (any) DNS server learns a hostname-to-IP address mapping, it *caches* the mapping
  - Local DNS servers also cache IP addresses of TLD servers
    - Thus root DNS servers are not often visited
  - Cache entries timeout after some time (a configurable parameter, i.e., set to two days)

DNS: caching and updating records

- DNS caching based on a time-to-live (TTL)

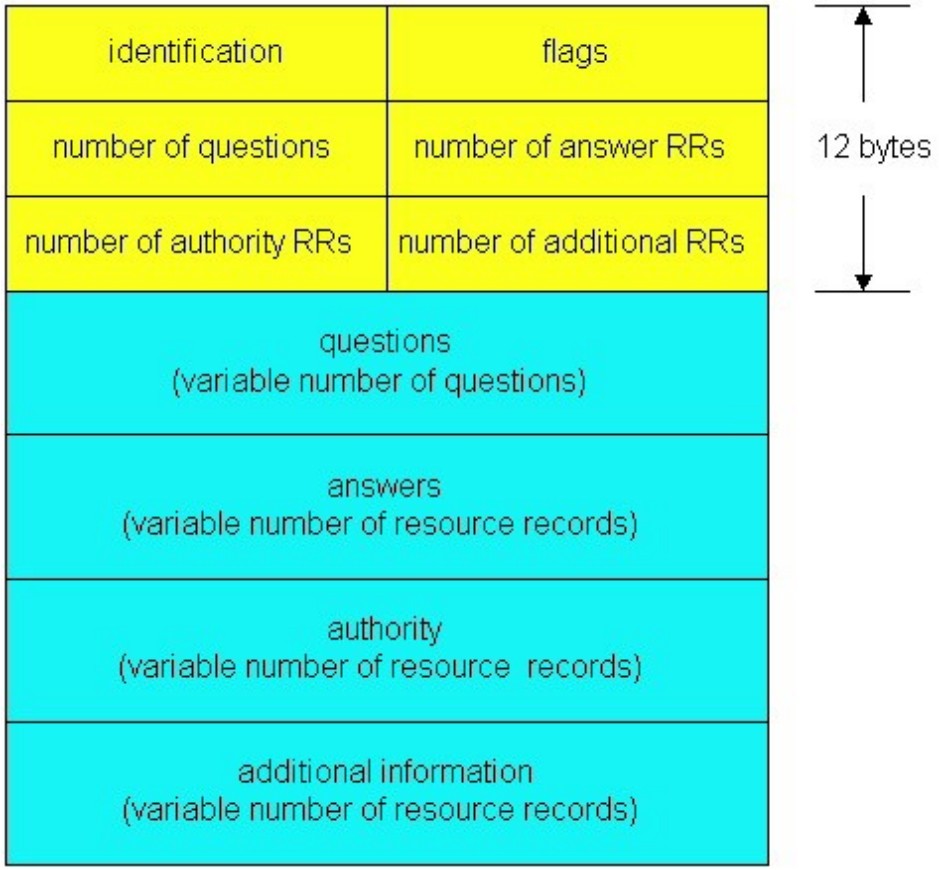


DNS Messages

DNS message : *query* and *reply* messages, both with same *message format*

msg header (12 bytes)

- identification: It is used to match the response with the query. 16 bit # for query, reply to query uses same #..
- flags:
  - Query (0) or reply (1)
  - recursion desired
  - recursion available
  - Authoritative flag

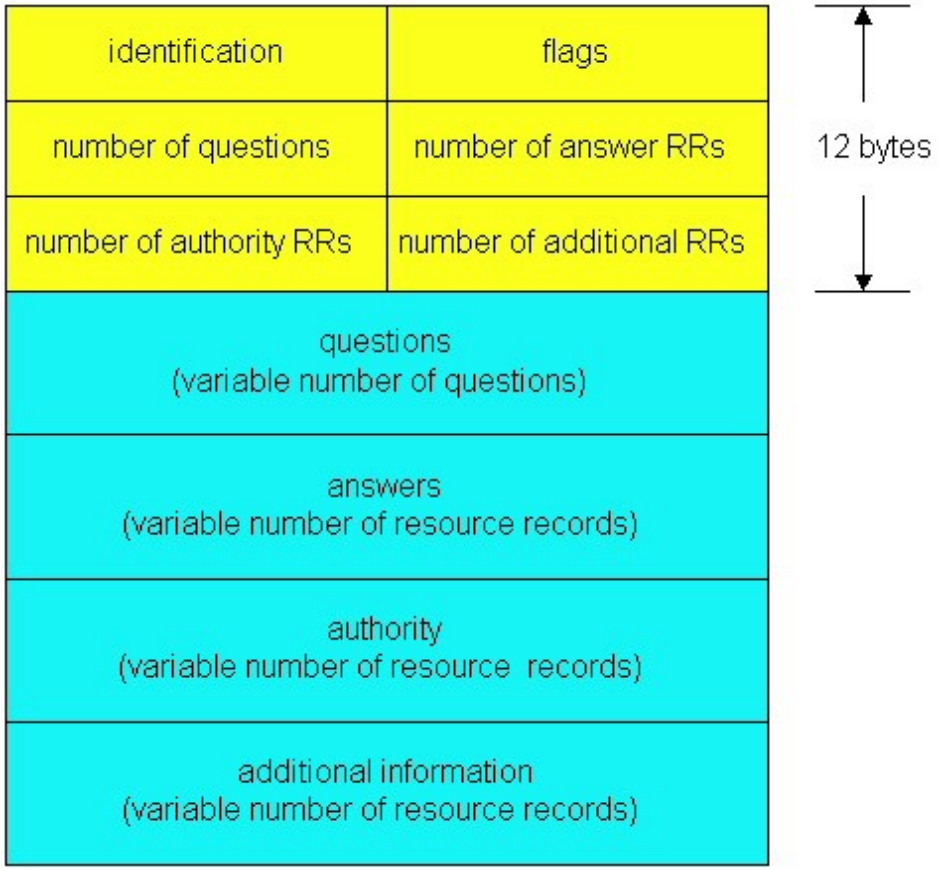


DNS Messages

DNS message : *query* and *reply* messages, both with same *message format*

msg header (12 bytes)

- Number of questions: contains the number of queries in the question section of the message.
- Number of answer resource records(RRs): contains the number of answer records in the answer section of a response mesage



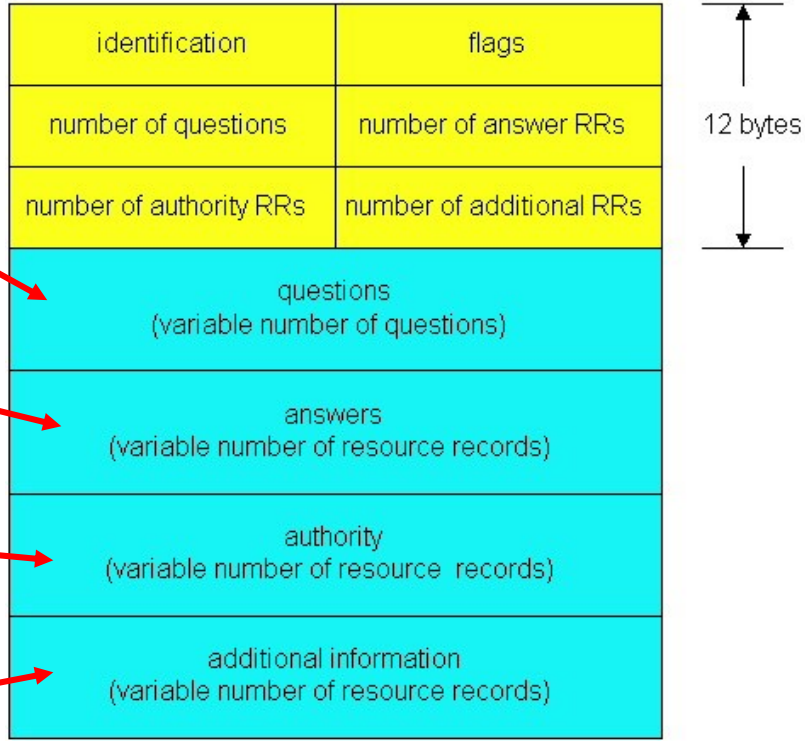
DNS Messages

Name, type fields for a query

RRs in reponse to query

records for authoritative servers

additional "helpful" info that may be used



DNS records

DNS servers store resource records (RR)

RR format: **(name, value, type, ttl)**

- Type=A
  - **name** is hostname
  - **value** is IP address
  - A type A record provides a standard hostname-to-IP address mapping

- Type=CNAME
  - **name** is an alias name
  - **value** is cannonical name for the alias name.
  - provide querying hosts the canonical name for a hostname
  - example of a CNAME record:

(www.google.com,backup2.google.com,CNAME)

- Ttl: time to live.
  - **Determines when a RR should be removed from a cache**