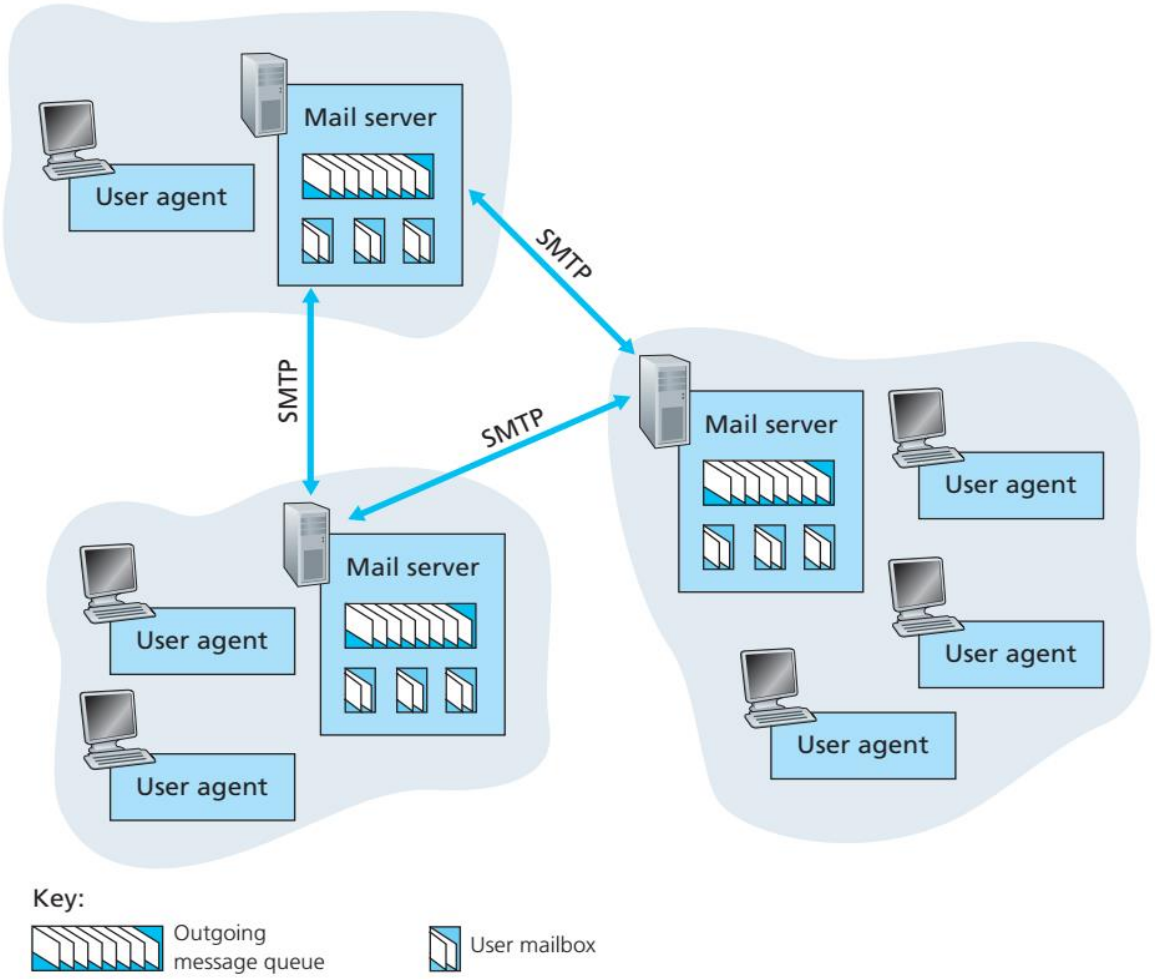
Module 2: Application Layer (Lecture – 3)

Dr. Nirnay Ghosh
Assistant Professor
Department of Computer Science & Technology
IIEST, Shibpur

Electronic Mail in the Internet

- Email: an asynchronous communication medium
- Modern e-mail has many powerful features: messages with attachments, hyperlinks, HTML-formatted text, embedded photos
- Three major components of Internet e-mail system are
  - User agent: allows users to read, reply to, forward, save, and compose messages (e.g., Microsoft Outlook, Apple Mail, etc.)
  - Mail server: core of the e-mail infrastructure – houses the mailbox of each recipient – manages and maintains messages that have been sent to the recipient
  - Simple Mail Transfer Protocol (SMTP): principle application-layer protocol for Internet electronic mail
- If the recipient's mail server is down
  - The message is hold in the message queue hosted by the sender's mail server
  - Transfer attempts are made after every 30 minutes (tentative)
  - Server removes the message from the queue (notifying the sender) if transfer attempts fail for several days

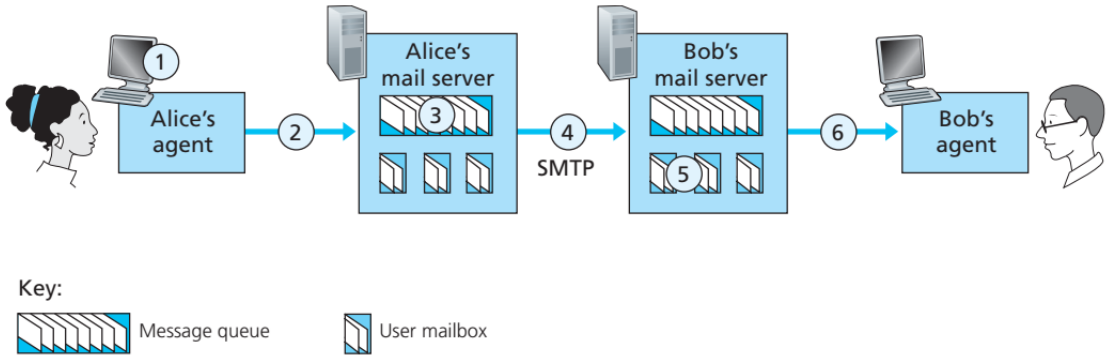


Internet E-mail System

- Typical path followed by a message:
  - *Sender's user agent → Sender's mail server → Recipient's mail server → Recipient's mailbox → Recipient's user agent*

Simple Mail Transfer Protocol (SMTP)

- Transfer messages from sender's mail server to recipient's mail server
- Does not use any intermediate mail servers even though the user agents are geographically dispersed
- Uses reliable transfer service of TCP to deliver message from sender's mail server to receiver's mail server
- Two sides – client side (executes on sender's mail server) and server side (executes on recipient's mail server)
- Basic operation of SMTP:
  - Sender's user agent send the composed message to her mail sever, where it is placed in a message queue
  - The client side of SMTP opens a persistent TCP connection on port 25 at the server side
    - Uses the same connection if the sending mail server has several messages to send to the receiving mail server



Message Exchange using SMTP

- SMTP handshaking phase takes place
- The client side of SMTP sends sender's message into the TCP connection
- The server side of SMTP receives the message and places it in recipient's mailbox
- Recipient invokes his user agent to read the message

- SMTP Handshaking Phase
  - Client side issues commands: *HELO* (an abbreviation for HELLO), *MAIL FROM, RCPT TO, DATA, QUIT (to close connection after all messages are sent)*
  - Client side also sends a single isolated period (.) to indicate the end of message
  - Server side issues replies to each command – each reply has a code and some (optional) English language explanation

SMTP Commands & Messages

```
S: 220 hamburger.edu
C: HELO crepes.fr
S: 250 Hello crepes.fr, pleased to meet you
C: MAIL FROM: <alice@crepes.fr>
S: 250 alice@crepes.fr ... Sender ok
C: RCPT TO: <bob@hamburger.edu>
S: 250 bob@hamburger.edu ... Recipient ok
C: DATA
S: 354 Enter mail, end with "." on a line by itself
C: Do you like ketchup?
C: How about pickles?
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 hamburger.edu closing connection
```

SMTP Vs. HTTP

- Similarities:
  - SMTP: transfers files (i.e., mail messages) from sender's mail server to recipient's mail server
  - HTTP: transfers files (also called objects) from a Web server to a Web client
  - Both use persistent TCP connections
- Differences:

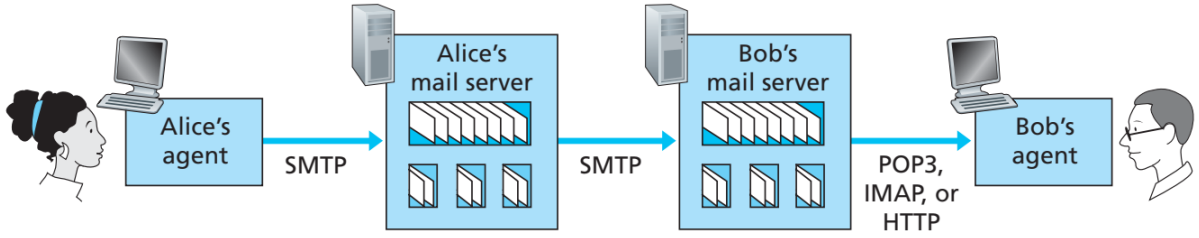
SMTP	HTTP
Mainly a push protocol (sending mail server pushes the message to the receiving mail server)	Mainly a pull protocol (someone loads information in the Web server and HTTP is used to pull information from the server)
Requires each message to be in 7-bit ASCII format	Does not impose this restriction
Places all of the message's objects into one object	Encapsulates each object in its own HTTP response message

Mail Access Protocols

- Uses a client-server architecture
- User agent (mail client): runs on the local PC (e.g., office PC, laptop, smartphone)
  - Enables viewing multimedia messages and attachments
- Mailbox: stored on another always-on shared mail server
- Mail server: typically maintained by the user's ISP (e.g., university or company)
- Limitation of SMTP in mail access:
  - Follows push protocol to send e-mail messages from sender's user agent to sender's mail server
  - Relays e-mail message to recipient's mail server
  - Cannot pull the message from recipient's mail server to recipient's user agent
- Mail access protocols used to transfer mail from recipient's mail server to recipient's user agent
  - Example: Post Office Protocol – Version 3 (POP3), Internet Mail Access Protocol (IMAP), Web mail

Post Office Protocol – Version 3 (POP3)

- Simple mail access protocol
- User agent (client) opens a TCP connection to the POP3 mail server on port 110
- POP3 servers: maintains some state information during a POP3 session
  - Keeps track of which messages have been marked deleted
  - Does not carry state information across POP3 sessions
- With the TCP connection established, POP3 progresses through three phases:
  - Authorization: user agent sends a username and a password (in clear text) to authenticate himself/herself – has two principle commands: *user <username>* and *pass <password>*
  - Transaction: user agent retrieves messages – can mark messages for deletion, remove deletion marks, and obtain mail statistics – the sequence of commands issued are: *list, retr, dele*
  - Update: occurs after the client has issued the *quit* command ending the POP3 session – mail server deletes the messages that were marked for deletion

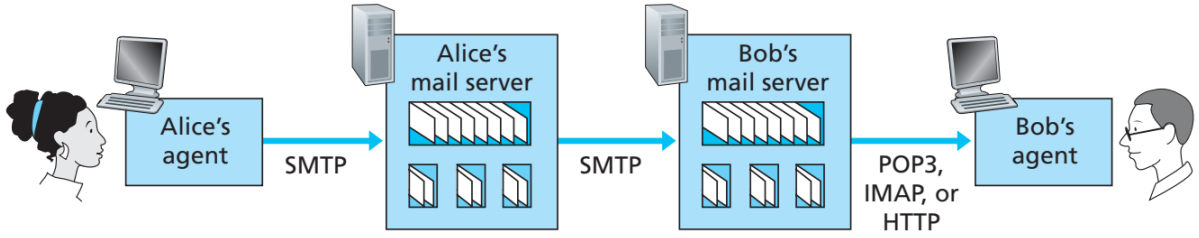


E-mail Protocols and their Communicating Entities

- POP3 Transactions
  - User agent issues commands and the server responds to each command
  - Two possible responses:
    - +OK (followed by server-to-client data): indicates the previous command was fine
    - -ERR: indicates the something was wrong with the previous command
  - User agent can configured to two modes: "download and delete" & "download and keep"
  - Commands issued during transaction phase:
    - *list*: mail server lists the size of each stored message
    - *retr*: mail server retrieves the message for user agent
    - *dele*: deletes each message from the server

Internet Mail Access Protocol (IMAP)

- Drawback of POP3:
  - Does not provide any means for a user to create folder on remote server and assign messages to folder
- IMAP: mail access protocol – significantly more complex than POP3
- IMAP server: associate each message with a folder
  - INBOX: when a message first arrives, it is associated with the recipient's INBOX folder
  - Recipient can move messages into a new, user-created folder, read the message, delete the message, as so on
  - Provides commands to search remote folders for message matching specific criteria
- IMAP server maintains user state information across IMAP sessions
  - State information: names of the folders and which messages are associated with which folder, and so on

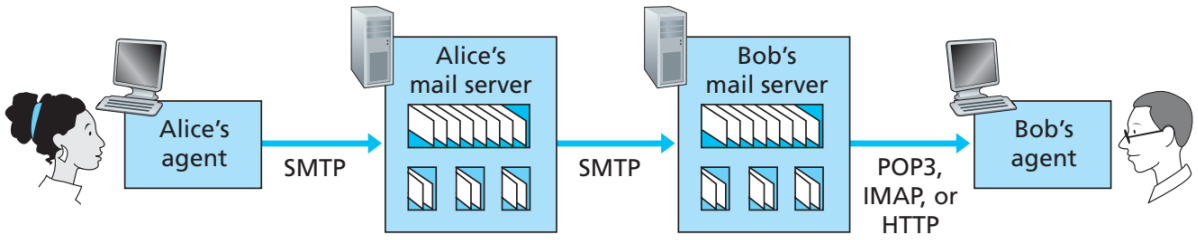


E-mail Protocols and their Communicating Entities

- It has commands that permit a user to obtain components of messages
  - User can obtain just the message header of a message or just one part of a multipart *Multipurpose Internet Mail Extensions (MIME)* message
  - Useful when there is low bandwidth connection between the user agent and its mail server

Web-based E-Mail (Webmail)

- Users nowadays send and access e-mail messages through Web browsers
  - Hotmail introduced Web-based access in mid-1990s
  - Also provided by Google, Yahoo!, as well as every major university and corporation
- User agent: ordinary Web browser
- User communicates with his/her remote mailbox via HTTP
- E-mail message is sent from recipient's mail server to his/her browser using the HTTP protocol rather than POP3 or IMAP
- Sender's message is sent from his/her browser to his/her mail server over HTTP rather than over SMTP
- Web-based e-mail servers can still send messages to, and receive messages from other mail servers using SMTP



E-mail Protocols and their Communicating Entities

Domain Name System (DNS)

- DNS – the Internet's directory service
- Translates user supplied hostname (e.g., cnn.com, www.google.com, iiests.ac.in, etc.) to IP addresses
- Commonly used by other application-layer protocols (HTTP, SMTP, and FTP)
- DNS protocol: runs over UDP and uses port 53
  - Distributed database
  - Implemented in a hierarchy of DNS servers
  - Allows hosts to query the distributed database
- DNS server: Unix machines running the Berkeley Internet Name Domain (BIND) software

- Steps involved in obtaining IP address of a requested Web server using DNS:
  - The requesting machine runs the client side of the DNS application
  - The browser extracts the hostname from the URL and passes it to the client side of the DNS application
  - The DNS client sends query containing the hostname to a DNS server
  - The DNS client eventually receives a reply which includes the IP address of the hostname
  - Once the browser receives the IP address from DNS, it can initiate a TCP connection to the Web server process located at port 80 at that IP address
  - Most often the desired IP address is cached in a "nearby" DNS server – reduces DNS network traffic as well as the average delay