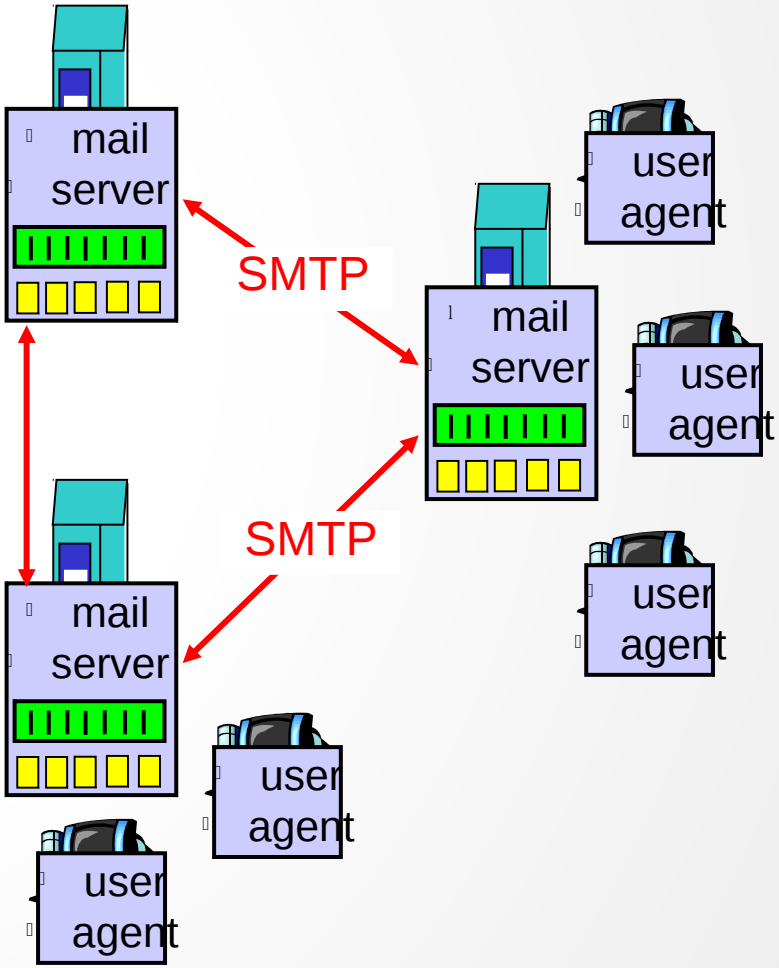
Electronic Mail (Email)

Manas Ranjan Lenka
School of Computer Engineering,
KIIT University

Three Major Components

- User Agents (Mail Readers)
- Mail Servers
- SMTP protocol



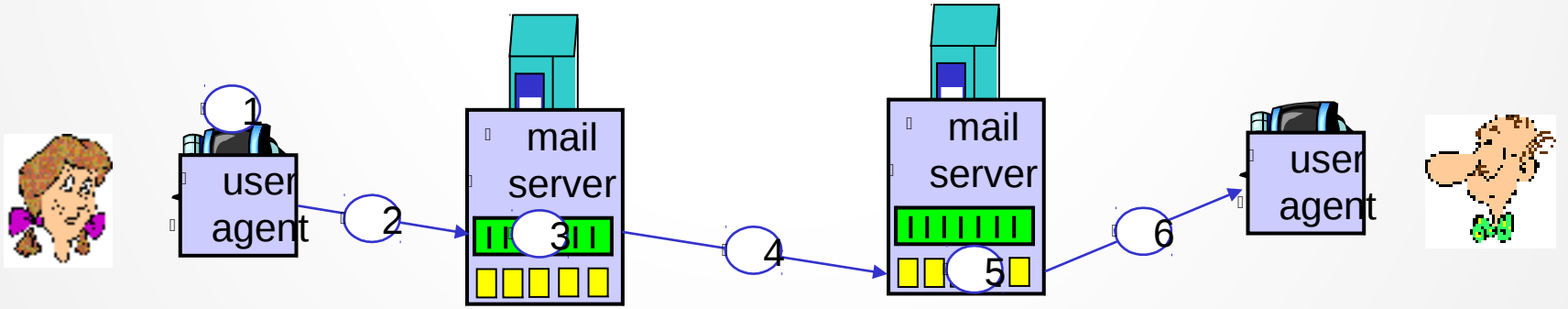
Three Major Components

- User Agents (Mail Readers)
  - Local programs that help people read, send and manage emails
  - Can be text based or GUI based
  - E.g. pine, elm, Microsoft Outlook, Mozilla Thunderbird
- Mail Servers
  - **mailbox** contains incoming messages for user
  - **message queue** of outgoing (to be sent) mail messages
  - **SMTP protocol** between mail servers to send email messages
    - client: sending mail server
    - “server”: receiving mail server

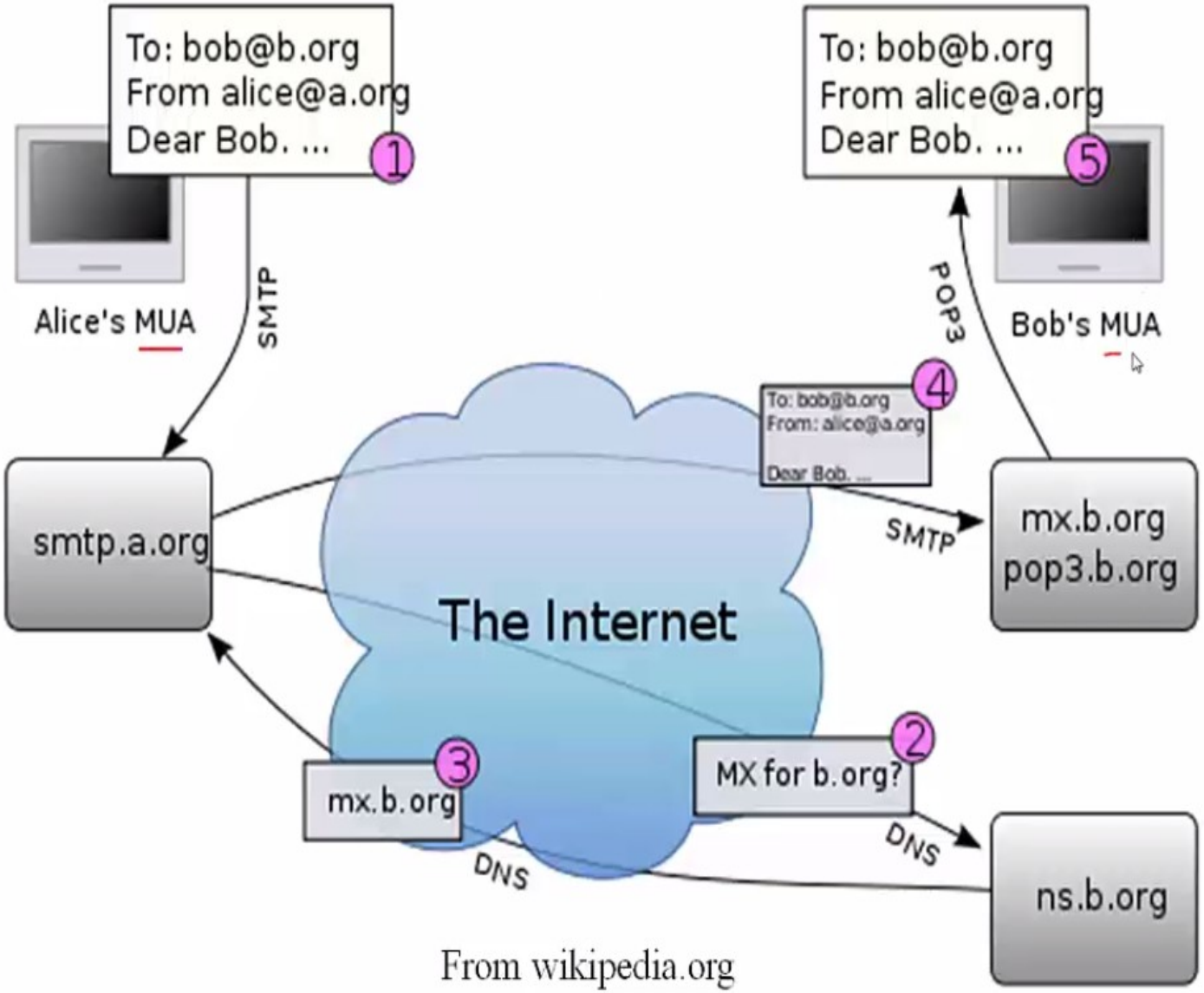
Architecture

Scenario: Alice sends message to Bob

1) Alice uses UA to compose message and set the “to” field value to `bob@someschool.edu`
2) Alice’s UA sends message to her mail server; Sending mails is a PUSH operation(SMTP protocol); 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 user agent to read message; Receiving mails is a PULL operation (POP3, IMAP, HTTP)



Architecture



From wikipedia.org

- POP3, IMAP, HTTP

Mail Servers

- Each user has a mailbox in a mail server, where user messages are stored

- Mail server also maintains a message queue of outgoing messages
  - In case of failure, attempts retransmissions and informs sender if it drops the message

- Both client and server side of SMTP run on a mail server

Simple Mail Transfer Protocol (SMTP)

- Uses TCP to reliably transfer email messages
- Operates on port 25
- Three phases of transfer based on command (ASCII text) / response (status code and phrase)
  - Handshaking (greeting)
  - Transfer of messages (many messages can be sent on a given connection)
  - Close connection

Mail Access Protocols

- Message waiting in receiver’s mailbox at the receiver’s mail server

- When the receiver open her/his user agent one of the retrieval protocols given below will be used:
  - POP3
  - IMAP
  - HTTP

Post Office Protocol – V3 (POP3)

- Very simple protocol (RFC 1939), limited functionality
- Typical usage: emails retrieved, stored (at User's PC) and read offline
- Does not maintain state across sessions
- Uses TCP and operates on port 110

- Three Phases:
  - Authorization: Authentication of user via username /passwd
  - Transaction: retrieve messages, mark deletions, obtain mail statistics
  - Update: After quit, server deletes marked emails

- Disadvantage: Poor support for remote multi-machine access

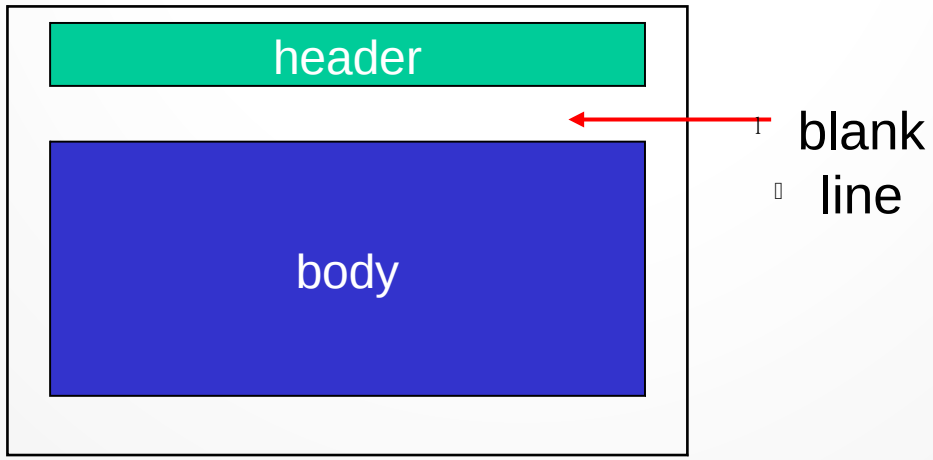
IMAP (RFC 3501)

- Emails organized in multiple folders (mailboxes) on server
- IMAP keeps user state across sessions
  - E.g. Names of folders and mappings between message IDs and folder name
- IMAP servers listen on port 143; IMAP over SSL (IMAPS) assigned port 993
- Permits multiple email clients (same user but different readers) to simultaneously connect to the mailbox
- Permits server-side searches (e.g. search for messages with subject line “networks”)

Mail Message Format

RFC 822: standard for text message format:

- Message within envelope contains header and body
  - Both are represented in ASCII text
  - Header contains control information for user agents
  - Body is for message content
  - Header separated from message body by blank line
- Header is a series of <CRLF> terminated lines
  - Contains type and value separated by a colon
    - From: , To:, Date: , Message-ID: , CC: , BCC: , Subject:, Received: , Return-Path:, Reply-To: etc.



Message Body and MIME

- Early days, email was made of English text messages, expressed in ASCII
  - No support for foreign languages
  - No support for non-text attachments (pdf, doc, jpg, audio files etc)
- Solution: Multipurpose Internet Mail Extensions (MIME)
  - Additional headers
  - Define content types and subtypes
  - Add structure to message body
  - Encoding rules for non ASCII messages (convert them to ASCII)

Headers added by MIME

Header	Meaning
MIME-Version	Identifies the MIME Version
Content-Description	ASCII string that tells what is in the message
Content-ID	Unique identifier
Content-Type	Type of data contained in the message
Content-Transfer-Encoding	How the data in message is encoded (e.g. 7bit, Base64)

Content-Type:
Text/Plain – Unformatted text
Text/Enriched – Simple formatted text
Image/Jpeg – Picture in jpeg format
Video/Mpeg – Video in mpeg format
Application/Msword – word document
Message/Mixed – Message made of independent pieces (each piece has own header line)