Last Lecture

- Start the *Application Layer*
- DNS

This Lecture

- SMTP

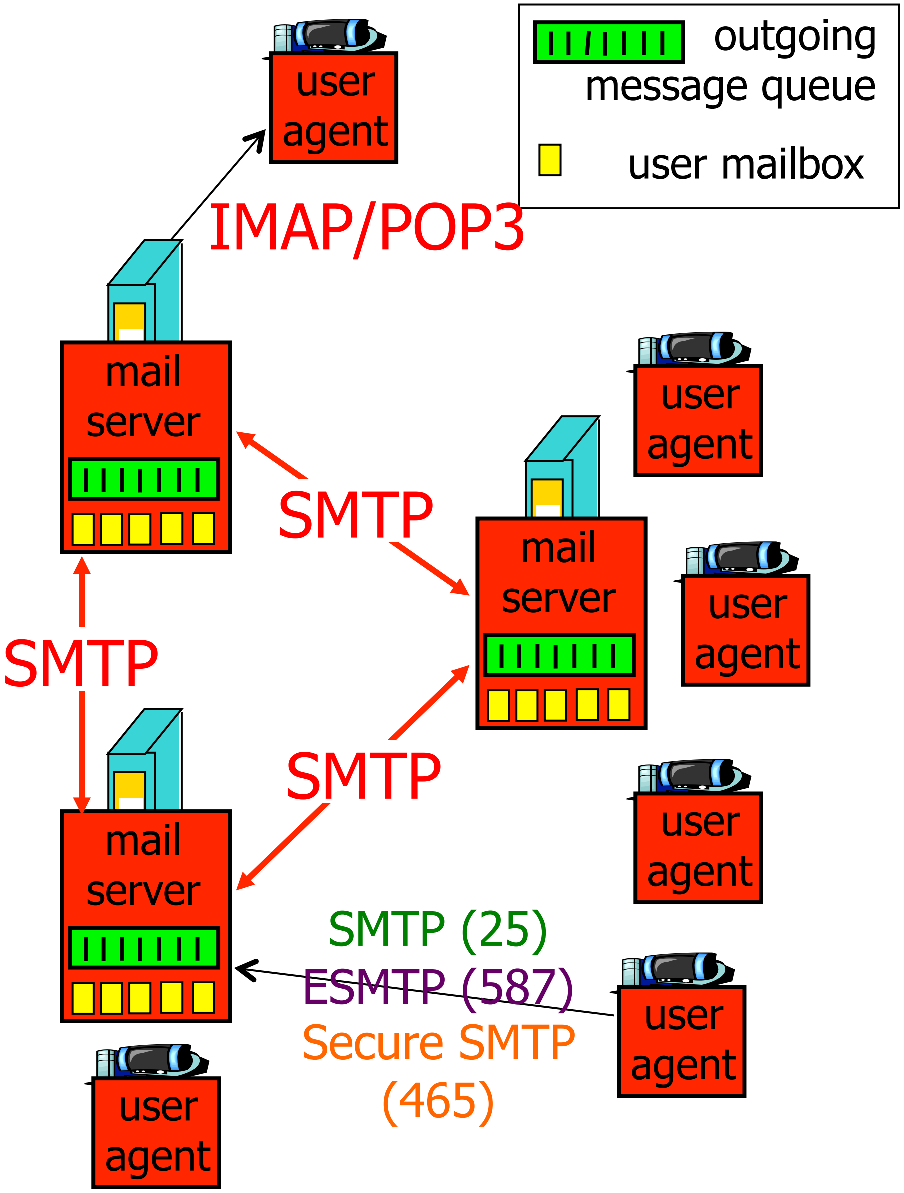
Electronic Mail Infrastructure

Four major components:

- User agents
- Mail servers
- SMTP
- Mail access protocol IMAP/POP3

User Agent ("Mail Reader")

- Composing, editing, reading mail messages
- E.g., Eudora, Outlook, Pine, Thunderbird, Apple Mail



Typical Scenario: Alice Emails Bob

1) Alice uses UA to compose message to bob@someschool.edu

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

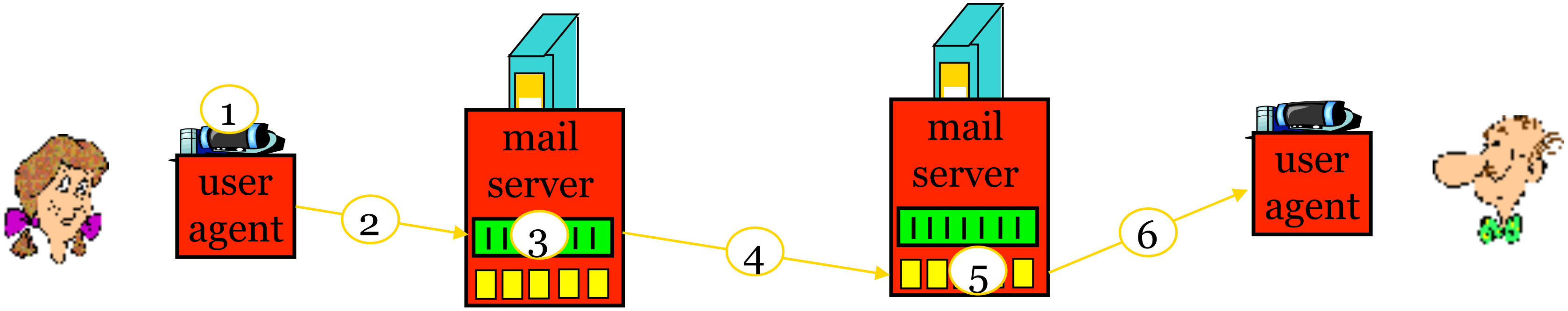
How does it know IP of Bob's mail server?

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 retrieve the message

There are often more than 1 mail server on the path (follow MX preference)

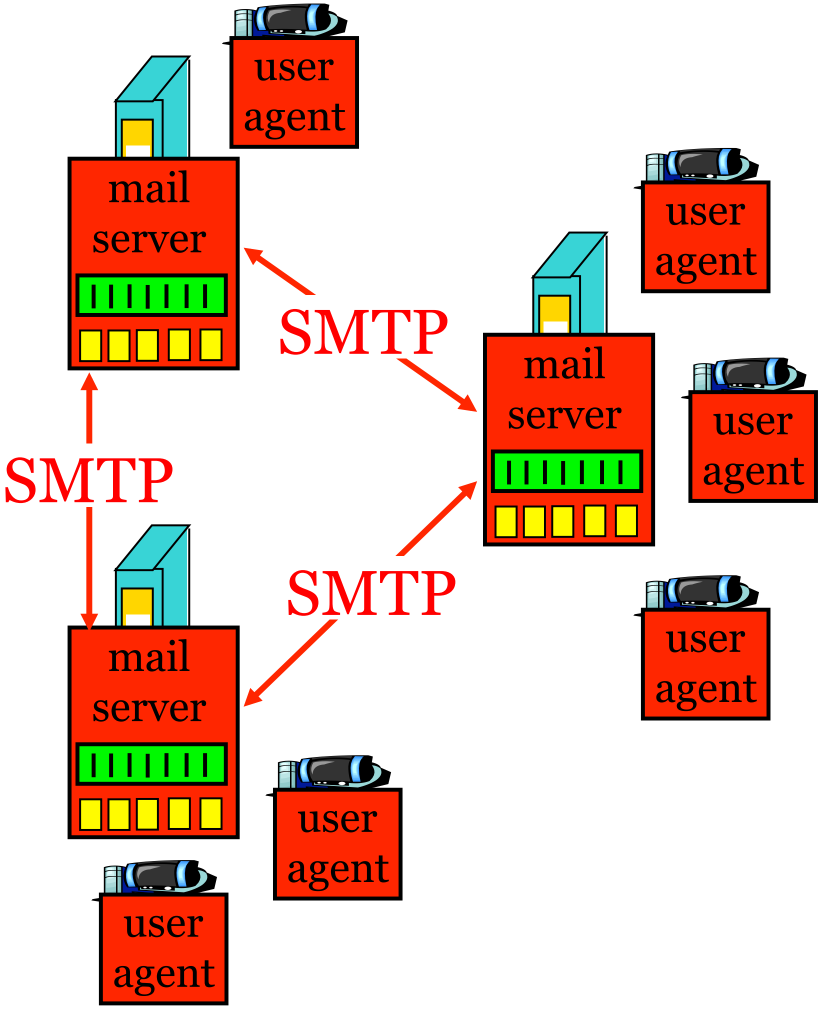


Mail Servers

Typical functionalities

- *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



SMTP

- Uses TCP to reliably transfer email message from client to server, port 25
- Three phases of transfer
  - handshaking (greeting)
  - transfer of messages
  - closure
- Command/response interaction
  - Commands (or Verb): ASCII text
  - Response : status code and phrase
    - 200-399: acceptance
    - 400-499: temporary rejection
    - 500-599: permanent rejection
- Messages must be in 7-bit ASCII

Sample SMTP Interaction

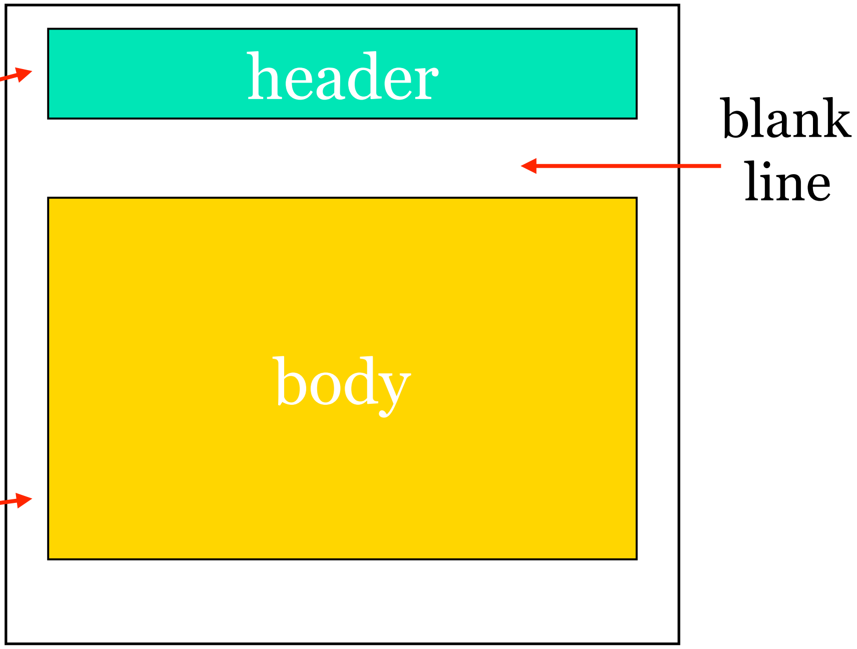
```
[hungngo@saigon] ~ $ telnet ubmx.buffalo.edu 25
Trying 128.205.5.197...
Connected to ubmx.buffalo.edu.
Escape character is '^]'.
220 mxB.acsu.buffalo.edu ESMTP Prefixe
HELO buffalo.edu
250 mxB.acsu.buffalo.edu
MAIL FROM: <hungngo@buffalo.edu>
250 2.1.0 Ok
RCPT TO: <my_email@gmail.com>
554 5.7.1 <my_email@gmail.com>: Relay access denied
RCPT TO: <hungngo@buffalo.edu>
250 2.1.5 Ok
DATA
354 End data with <CR><LF>.<CR><LF>
This is just a test
.
250 2.0.0 Ok: queued as 7FE8B2889
QUIT
221 2.0.0 Bye
Connection closed by foreign host.
```

SMTP (Basic) Mail Message Format

SMTP: protocol for exchanging email msgs

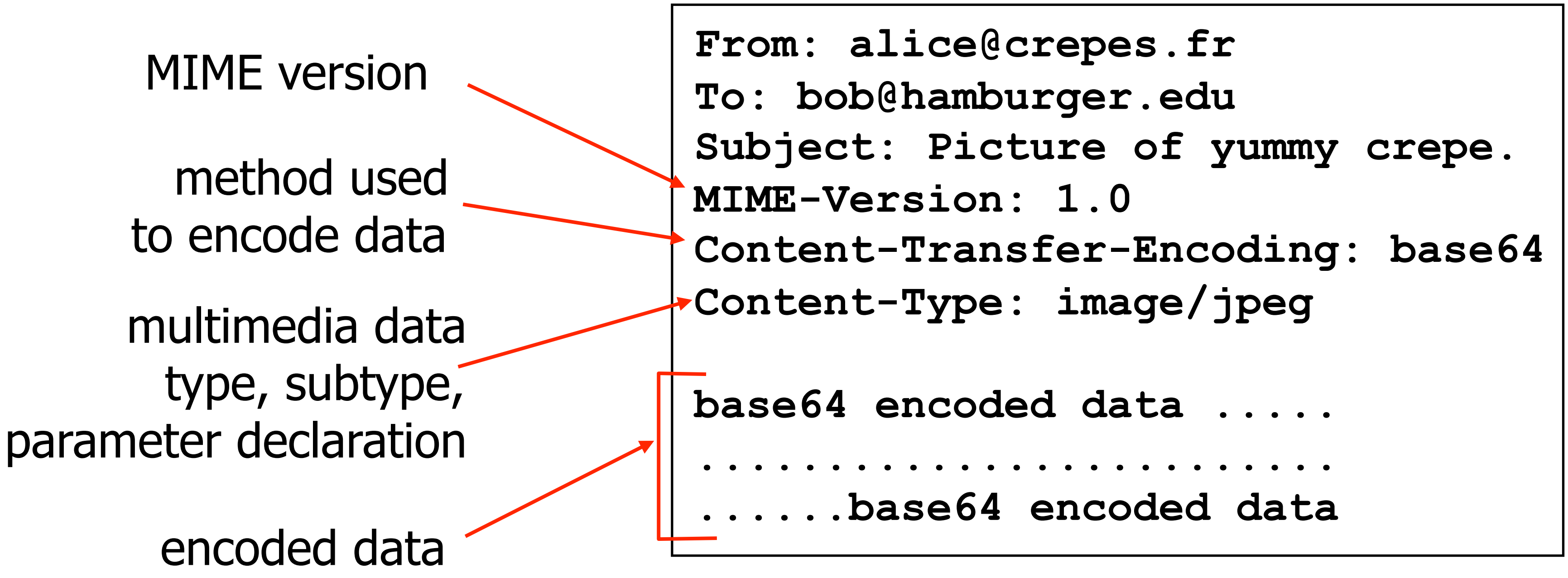
RFC 822: standard for text message format:

- header lines, e.g.,
  - To:
  - From:
  - Subject:
  - *Header lines are different from SMTP commands*!
- body
  - the "message", **7-bit** ASCII characters only



Message Format: Multimedia Extensions

- MIME: multimedia mail extension, RFC 2045, 2056
- Additional lines in msg header declare MIME content type



MIME version

method used to encode data

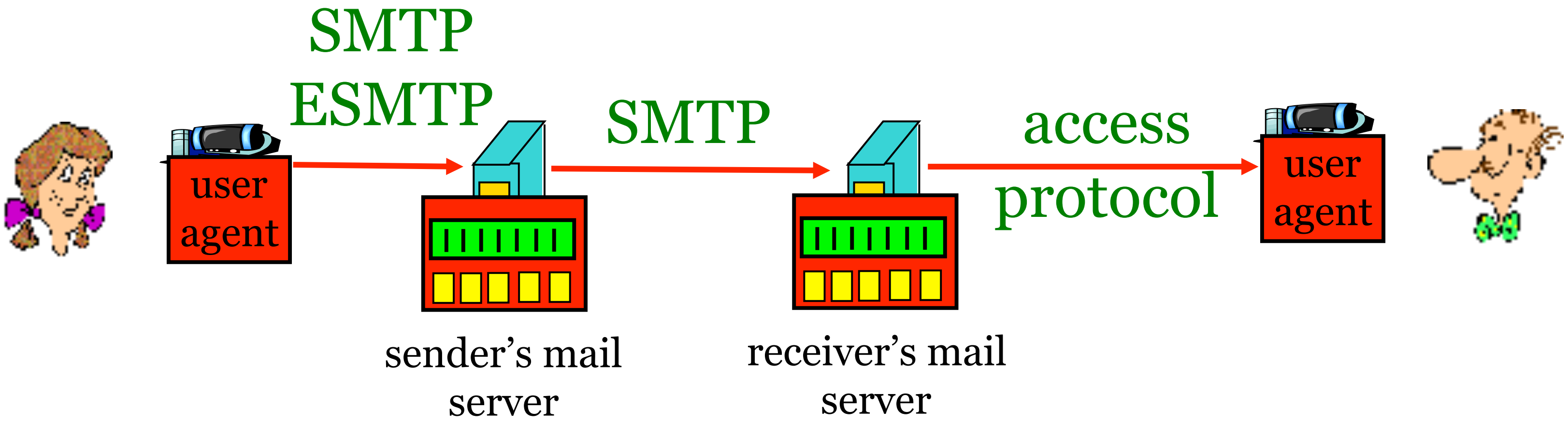
multimedia data type, subtype, parameter declaration

encoded data

```
From: alice@crepes.fr
To: bob@hamburger.edu
Subject: Picture of yummy crepe.
MIME-Version: 1.0
Content-Transfer-Encoding: base64
Content-Type: image/jpeg

base64 encoded data .....
.........................
......base64 encoded data
```

Mail Access Protocols



- Mail access protocol: retrieval from server
  - **POP**: Post Office Protocol [RFC 1939]
    - authorization (agent <-->server) and download
  - **IMAP**: Internet Mail Access Protocol [RFC 1730]
    - more features (more complex)
    - manipulation of stored messages on server
  - **HTTP**: gmail, Hotmail, Yahoo! Mail, etc.

POP3 Protocol

Authorization phase

- client commands:
  - `user`: declare username
  - `pass`: password
- server responses
  - `+OK`
  - `-ERR`

```
S: +OK POP3 server ready
C: user bob
S: +OK
C: pass hungry
S: +OK user successfully logged on
```

Transaction phase, client:

- `list`: list message numbers
- `retr`: retrieve message by number
- `dele`: delete
- `quit`

```
C: list
S: 1 498
S: 2 912
S: .
C: retr 1
S: <message 1 contents>
S: .
C: dele 1
C: retr 2
S: <message 1 contents>
S: .
C: dele 2
C: quit
S: +OK POP3 server signing off
```

POP3 and IMAP

More about POP3

- Previous example uses "download and delete" mode.
- Bob cannot re-read e-mail if he changes client
- "Download-and-keep": copies of messages on different clients
- *POP3 is stateless across sessions*

IMAP

- Keep all messages in one place: the server
- Allows user to organize messages in folders
- IMAP keeps user state across sessions:
  - names of folders and mappings between message IDs and folder name

SMTP: a Cinderella Story

- A simple idea (1971) changes our lives forever!

- According to pingdom.com:
  - 1.4 billion – The number of email users worldwide.
  - 247 billion – The number of emails sent per day in 2009.
  - 90 trillion – The total number of emails sent in 2009
  - 81% – *The percentage of emails that are spam.*

- Remember Sabeer Bhatia? You could do the same.

Our Cinderella is Not Very Smart

RFC 2821	S	Simple Mail Transfer Protocol (SMTP)
RFC 1123	S	Requirements for Internet hosts - application and support
RFC 974	S	Mail routing and the domain system (MX records)
RFC 1869	S	SMTP Service Extensions
RFC 1870	S	SMTP Service Extension for Message Size Declaration
RFC 1652	D	SMTP Service Extension for 8bit-MIMEtransport
RFC 3030	P	SMTP Service Extensions for Transmission of Large and Binary MIME Messages
RFC 1845	E	SMTP Service Extension for Checkpoint/Restart
RFC 1846	E	SMTP 521 Reply Code
RFC 2920	S	SMTP Service Extension for Command Pipelining
RFC 1985	P	SMTP Service Extension for Remote Message Queue Starting (ETRN)
RFC 2645	P	On-Demand Mail Relay (ODMR) SMTP with Dynamic IP Addresses
RFC 2852	P	Deliver By SMTP Service Extension
RFC 2034	P	SMTP Service Extension for Returning Enhanced Error Codes
RFC 3464	P	An Extensible Message Format for Delivery Status Notifications (DSNs)
RFC 3463	D	Enhanced Mail System Status Codes
RFC 3461	P	SMTP Service Extension for Delivery Status Notifications
RFC 3462	P	Multipart/Report Content Type for the Reporting of Mail System Administrative Messages
RFC 2476	P	Message Submission
RFC 2554	P	SMTP Service Extension for Authentication
RFC 2505	B	Anti-Spam Recommendations for SMTP MTAs
RFC 2442	I	Batch SMTP Media Type
RFC 1047	I	Duplicate messages and SMTP
RFC 1090	I	SMTP on X.25

The Headache

- To filter or not to filter, that’s the problem!
- The number of spam emails sent in 2009 (assuming 81% are spam) is … *73 trillion*
- Note:
  - “Spam” ® is a registered trademark of a meat product made by Hormel
  - “Spam” comes from a Monty Python sketch
- What’s the root cause of spamming?

Main Problem

Botnet!

NY Times, Jan 23, 2009:

Worm Infects Millions of Computers Worldwide

“A new digital plague has hit the Internet, infecting millions of personal and business computers in what seems to be the first step of a multistage attack. The world’s leading computer security experts do not yet know who programmed the infection, or what the next stage will be.

In recent weeks a worm, a malicious software program, has swept through corporate, educational and public computer networks around the world. Known as Conficker or Downadup, it is spread by a recently discovered Microsoft Windows vulnerability, by guessing network passwords and by hand-carried consumer gadgets like USB keys.

…”

Much more about botnets later in the course

Some Partial Solutions

- Authentication (SMTP over SSL/TLS)
  - Users
  - Mail servers
  - How to trust people/servers? A “trusted” third party causes other problems.
- Rewrite SMTP
  - Key CS phrase: “Backward compatibility”
  - Currently there are millions of SMTP servers on the net
  - Took IETF > 6 years to decide that spam is harmful and formed a “research group” to … study solutions
- Ad Hoc extensions to SMTP (e.g. TEOS)
- Microsoft: why don’t we amend DNS?

More Partial Solutions

- "Challenge-response" technology
  - SpamArrest.com, Mail-block.com, iPermitMail.com
  - Poses problems on its own
- RFC 2505:
  - Do not relay
  - Use DNS information (hopefully with secure DNS – RFC2065 – which makes IP spoofing much harder)
    - This is currently in use on many MTAs
    - Another problem: DoS on DNS servers
    - Delays: waiting for DNS response
    - …