Indian Institute of Technology Kharagpur

Electronic Mail

Prof. Indranil Sen Gupta
Dept. of Computer Science & Engg.
I.I.T. Kharagpur, INDIA

Lecture 10: Electronic mail

On completion, the student will be able to:

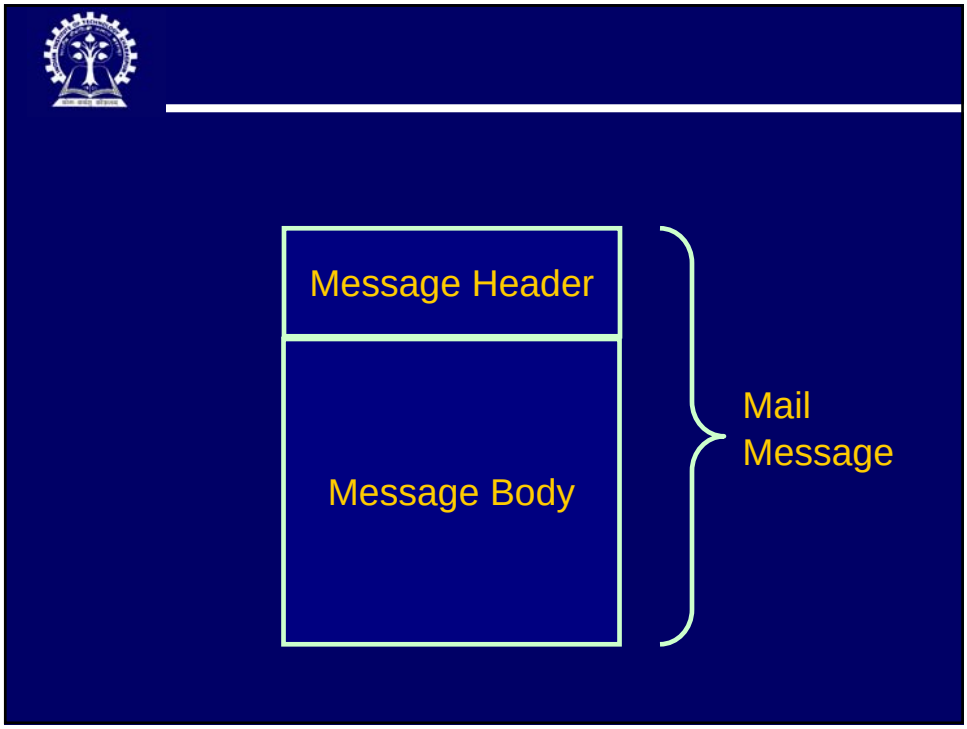
1. Describe the overall architecture of the email system.
2. Explain the functions of user agents and message transfer agents.
3. Explain the functions of SMTP and MIME protocols.
4. Demonstrate the sending of mail using raw SMTP commands.
5. Interpret the email header fields.
6. Explain the function of the POP3/IMAP protocol.

Electronic Mail

- Most widely used application on the Internet.
- For sending mails:
  - Simple Mail Transfer Protocol (SMTP)
  - Multi-purpose Internet Mail Extension (MIME)
- For receiving mails:
  - Post office protocol version 3 (POP3)
  - Internet mail access protocol (IMAP).

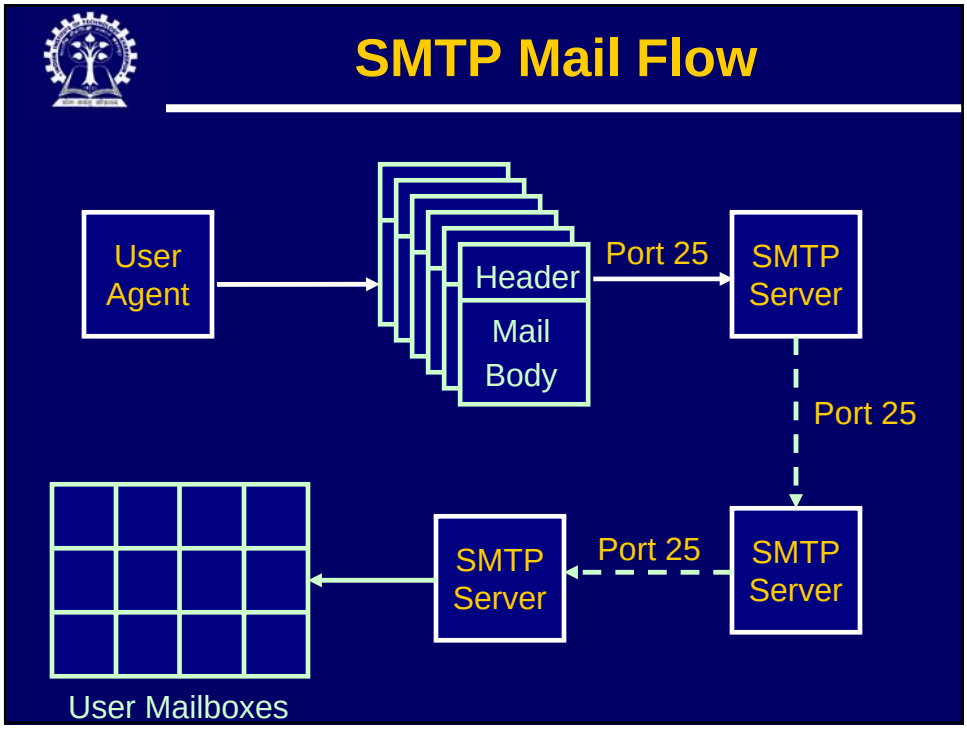
Simple Mail Transfer Protocol

- Based on RFC 821.
- Transmits simple text messages only.
  - 7-bit ASCII format.
- Uses information written on envelope of mail.
  - Message header.
  - Contains recipient address and other information.
- Does not look at contents.
  - Message body.



Basic Operation

- **Mail is created by user agent program (mail client).**
- **Messages queued and sent as input to SMTP sender program.**
  - **Typically a server process.**
  - **Daemon on UNIX.**
    - **sendmail or qmail**



Mail Message Contents

- **Each queued message has:**
  - **Message text**
    - **RFC 822 header with message envelope and list of recipients.**
    - **Message body, composed by user.**
  - **A list of mail destinations**
    - **Derived by user agent / SMTP server from header.**
    - **May require expansion of mailing lists.**

SMTP Sender

- Takes message from queue.
- Transmits to proper destination host.
  - Via SMTP transaction.
  - Over one or more TCP connections to port 25.
- When all destinations processed, message is deleted.

Optimization

- If message is sent to multiple users on a given host, it is sent only once.
  - Delivery to users handled at destination host.
- If multiple messages are ready for given host, a single TCP connection can be used.
  - Saves overhead of setting up and dropping connection.

Possible Errors

- Host unreachable
- Host out of operation
- TCP connection fail during transfer
- Faulty destination address
  - User error
  - Target user address has changed
  - Redirect if possible
  - Inform user if not

Sender can re-queue mail

- Give up after a period

SMTP Protocol - Reliability

- Used to transfer messages from sender to receiver over TCP connection.
  - Uses port number 25.
- Attempts to provide reliable service.
- No guarantee to recover lost messages.
- No end-to-end ACK to sender.
- Error indication report not guaranteed.

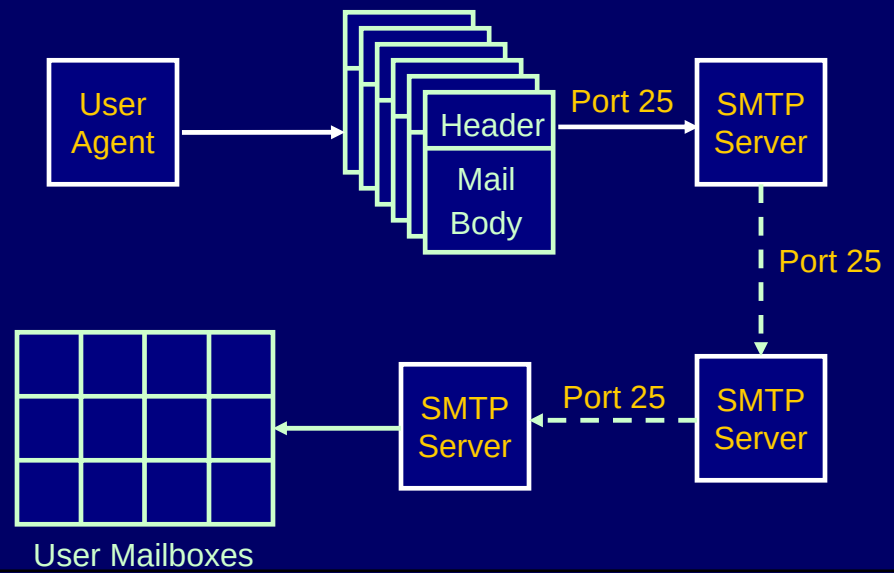
SMTP Receiver

- Accepts arriving message.
- Places in user mailbox or copies to outgoing queue for forwarding.
- Receiver must:
  - ➢ Verify local mail destinations.
  - ➢ Deal with errors
    - ▪ Transmission
    - ▪ Lack of disk space

SMTP Forwarding

- Mostly direct transfer from sender host to receiver host.
- May go through intermediate mail servers via forwarding capability.
  - ➢ Sender can specify route.

SMTP Mail Flow



SMTP System Overview

- Commands and responses exchanged between sender and receiver.
- Initiative with sender.
  - Establishes TCP connection.
- Sender sends commands to receiver.
  - e.g. HELO <domain><CRLF>
- Each command generates exactly one reply.
  - e.g. 250 requested mail action ok; completed.

SMTP Replies

- Starts with 3-digit code.
- Leading digit indicates category.
  - 2xx -- Positive completion reply
  - 3xx -- Positive intermediate reply
  - 4xx -- Transient negative completion reply
  - 5xx -- Permanent negative completion reply

Operation Phases

a) Connection setup
b) Exchange of command-response pairs
c) Connection termination

a) Connection Setup

- Sender opens TCP connection with receiver.
- Once connected, receiver identifies itself.

  220 <domain> service ready
- Sender identifies itself.

  HELO
- Receiver accepts sender's identification.

  250 OK
- If mail service not available, the second step above becomes:

  421 service not available

b) Mail Transfer Commands

- The MAIL FROM command identifies originator.
  - Gives reverse path to be used for error reporting.
  - Receiver returns 250 OK or appropriate failure / error message.

- One or more RCPT TO commands identify recipients for the message.
  - Separate reply for each recipient.
- The DATA command transfers message text.
  - End of message indicated by a line containing just period (.)

c) Closing Connection

- Two steps:
  - Sender sends QUIT and waits for reply.
  - Then initiate TCP close operation.
- Receiver initiates TCP close after sending reply to QUIT.

An Example SMTP Session

- How to connect to an SMTP server?

  telnet servername 25

  - A TCP connection gets established over port number 25.
  - The telnet client and the mail server can now start a dialogue.

An Example SMTP Session

S: 220 hotmail.com Simple Mail Transfer Service Ready
C: HELO yahoo.com
S: 250 hotmail.com

C: MAIL FROM: <isg@yahoo.com>
S: 250 OK

C: RCPT TO: <myfriend@hotmail.com>
S: 250 OK

C: RCPT TO: <somebody@rediffmail.com>
S: 250 OK

An Example SMTP Session

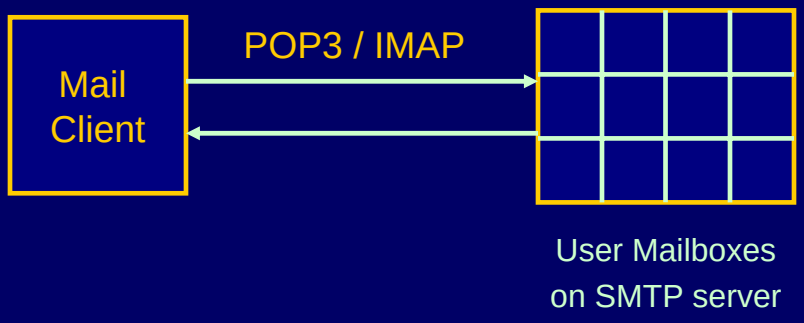
C: DATA
S: 354 Start mail input; end with (.)
C: … actual contents of the message …
C: ……
C: ……
C: .
S: 250 OK

C: QUIT
S: 221 hotmail.com Service closing transmission channel

Mail Access Protocols

- Two mail access protocols are widely used:
  1. Post Office Protocol, version 3 (POP3)
  2. Internet Mail Access Protocol version 4 (IMAP4).

What do they do?



POP3

- **The client POP3 software is installed on the recipient machine, and the server POP3 software installed on mail server.**
  - **The client (user agent) opens a connection with the server on TCP port number 110.**
  - **Sends user name and password.**
  - **Can access the mails, one by one.**

POP3 (contd.)

- Two modes:
  - Delete mode – mails deleted as they are read
  - Keep mode – mails remain in the mailbox
- POP3 has commands for:
  - Log in
  - Log out
  - Fetch messages
  - Delete messages

IMAP4

- Provides the following extra features:
  - A user can check the email header before downloading.
  - A user can search the contents of the email for a specific string prior to downloading.
  - A user can create, delete, or rename mailboxes on the mail server.
  - A user can create a hierarchy of mailboxes in a folder for email storage.

Multipurpose Internet Mail Extension (MIME)

- SMTP cannot transmit non-text messages.
  - Solutions (like uuencode) exists on some systems, but are not standardized.
- Cannot transmit text that includes international characters (e.g. â, å, ä, è, é, ê, ë).
  - Need 8 bit ASCII.

- Servers may reject mail over certain size.
- Some SMTP implementations do not adhere to standard.
  - CRLF, truncate or wrap long lines, removal of white space, etc.

Overview of MIME

- Five new message header fields:
  - MIME-version
  - Content-type
  - Content-transfer-encoding
  - Content-Id
  - Content-description
- A number of *content types* and *transfer encoding* formats have been defined.

Content Types

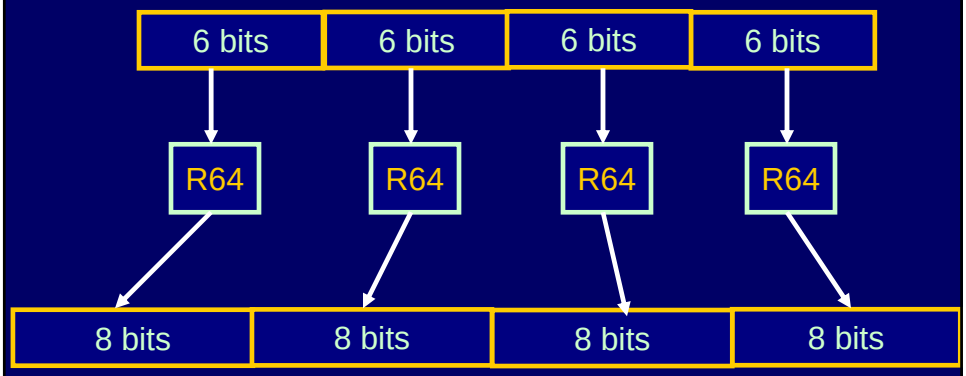
- Text body
- Multipart
  - Mixed, Parallel, Alternative
- Message
  - RFC 822, Partial, External-body
- Image
  - jpeg, gif
- Video
  - mpeg
- Audio
  - Basic
- Application
  - Postscript
  - octet stream

MIME Transfer Encodings

- Specifies how the mail body is wrapped for transmission.
- Content transfer encoding field can have six possible values.
  - 7bit, 8bit, binary: no encoding done for these three.
    - Provide information about nature of data.

  - Quoted-printable
    - Data mostly printable ASCII characters.
    - Non-printing characters represented by hex code.
  - Base64
    - Maps arbitrary binary input onto printable output.
  - X-token
    - Named nonstandard encoding.

Base64 Encoding



- **Expands the message by 33%.**
- **Uses the symbols A..Z, a..z, 0..9, +, /**

MIME Header Example

From: Indranil Sengupta <isg@iitkgp.ac.in>
To: Jaswinder Ahuja <jassi@cadence.com>
Subject: Simple Message
MIME-Version: 1.0
Content-type: multipart/mixed; boundary="simple boundary"

This is the preamble. It is to be ignored, though it is a handy place for mail composers to include an explanatory note.--simple boundary

This is implicitly typed plain text. It does NOT end with a linebreak.
-- simple boundary
Content-type: text/plain; charset=us-ascii

This is explicitly typed plain ASCII text. It DOES end with a linebreak.

--simple boundary--
This is the epilogue. It is also to be ignored.

Another MIME Example

From: Indranil Sengupta <isg@iitkgp.ac.in>
To: Jaswinder Ahuja <jassi@cadence.com>
Subject: Formatted text mail
MIME-Version: 1.0
Content-type: multipart/alternative; boundary=boun42

--boun42

Content-type: text/plain; charset=us-ascii

... plain text version of message goes here ...

--boun42
Content-type: text/enriched

... RFC1896 text/enriched version of the same message goes here ...
--boun42--

End of Lecture 10

SOLUTIONS TO QUIZ QUESTIONS ON LECTURE 9

Quiz Solutions on Lecture 9

1. What is an iterative server??

   It is one for which the client requests are processes sequentially, one at a time.

2. What is a concurrent server?

   It is one where multiple client requests can be handled at the same time. Multiple copies of server are created for the purpose (using process or thread).

Quiz Solutions on Lecture 9

3. Which of client or server must start first?

   The server program must start first.

4. What are the components of a socket?

   One of the following:

   - Protocol, local IP, local port
   - Protocol, remote IP, remote port

   A socket is also called half-association.

Quiz Solutions on Lecture 9

5. What is a domain?

   A set of computers related by either their geographic location, or the way they function.

6. What are named and nslookup?

   named: domain name server running under the UNIX system.

   nslookup: command to invoke DNS name resolver.

Quiz Solutions on Lecture 9

7. How does recursive name resolution works?

   Host sends a request to a DNS server, which recursively forwards it to other DNS servers. The response is sent back to the host along the same path.

8. How can you connect to a server xyz.com over port number 1234?

   telnet xyz.com 1234

Quiz Solutions on Lecture 9

9. Why does ftp use more than one port numbers for communication?

   One port number (21) is used for control connection, and another one for data transfer.

10. What are the functions of the hash and bin commands in ftp?

    hash: indicates progress of file transfer

    bin: selects binary mode of data transfer

QUIZ QUESTIONS ON LECTURE 10

Quiz Questions on Lecture 10

1. What are the basic drawbacks of SMTP?
2. Which port number do SMTP servers use for accepting client requests?
3. Why does MIME does not have any port number associated with it?
4. Under what condition can a SMTP server also act as a mail client?
5. What are the purposes of the “MAIL FROM” and “RCPT TO” commands in SMTP?
6. What is the difference between Cc and Bcc in the SMTP header?

Quiz Questions on Lecture 10

7. Why is IMAP preferred over POP3?
8. A message of size 3000 bytes is encoded using Base64 scheme. What will be the size of the encoded message?
9. Is it mandatory for the DNS server to run on the same machine that runs the SMTP server?
10. How are mail attachments handled in MIME?