Electronic Mail

Prof. Indranil Sen Gupta

Professor, Dept. of Computer Science & Engineering
Indian Institute of Technology
Kharagpur

Introduction

- Most heavily used application on the Internet.
- Simple Mail Transfer Protocol (SMTP)
  - Uses TCP/IP
  - Delivery of simple text messages
- Multi-purpose Internet Mail Extension (MIME)
  - Delivery of other types of data
  - Voice, images, video clips

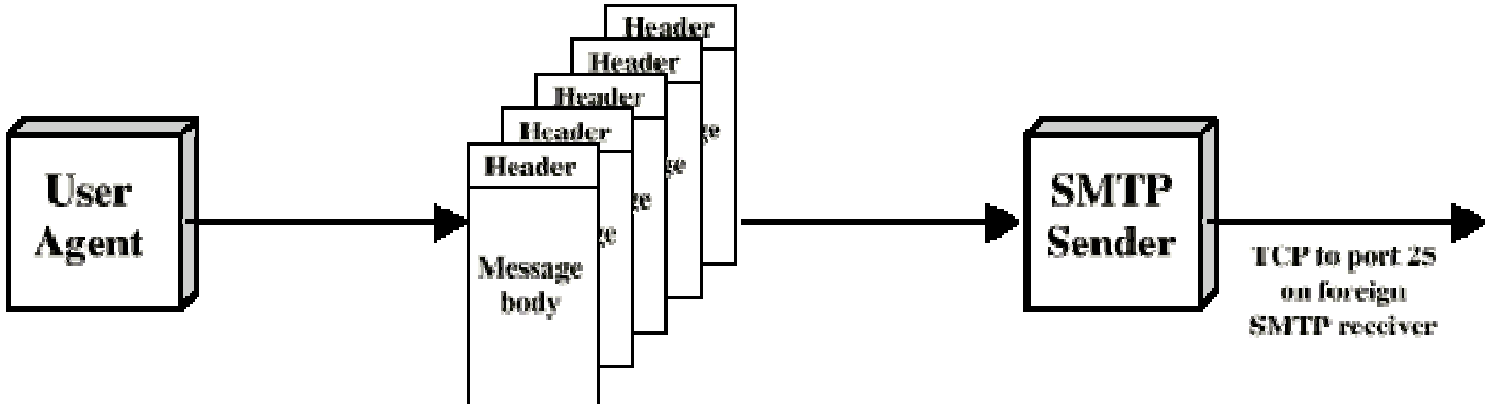
Simple Mail Transfer Protocol (SMTP)

- Based on RFC 821.
- Not concerned with format of messages or data.
  - Transmits simple text messages only.
- SMTP uses information written on envelope of mail.
  - Message header
- Does not look at contents.
  - Message body
- Except:
  - Standardize message character set to 7 bit ASCII.
  - Add log information to start-of-message.
    - Shows path taken.

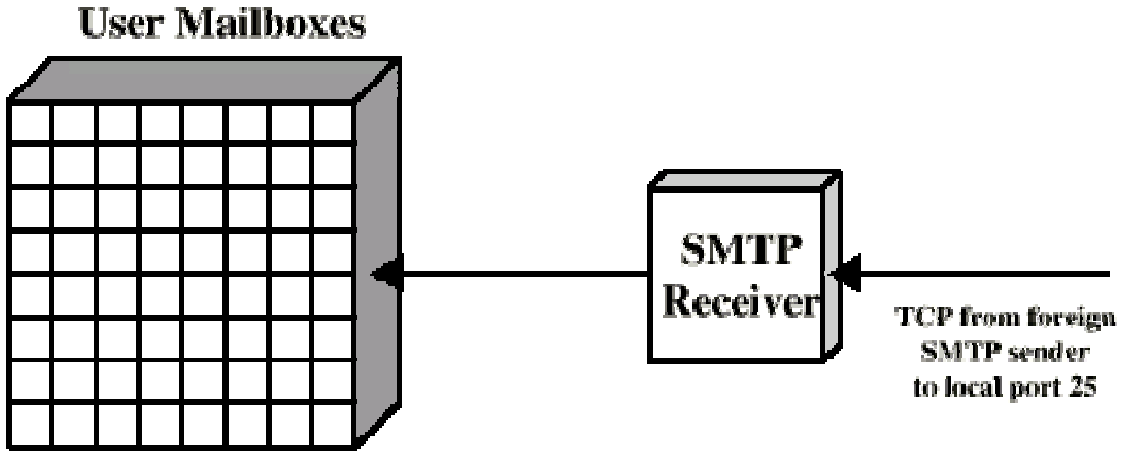
Basic Operation

- Mail created by user agent program (mail client).
  - Message consists of:
    - Header containing recipient's address and other information.
    - Body containing user data.
- Messages queued and sent as input to SMTP sender program.
  - Typically a server process (daemon on UNIX).

SMTP Mail Flow



(a) Outgoing Mail



(b) Incoming Mail

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 be listed in 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 delivery complete, sender deletes destination from list for that message.
- When all destinations processed, message is deleted.

Possible Errors

- Host unreachable
- Host out of operation
- TCP connection fail during transfer
- Sender can re-queue mail
  - Give up after a period
- Faulty destination address
  - User error
  - Target user address has changed
  - Redirect if possible
  - Inform user if not

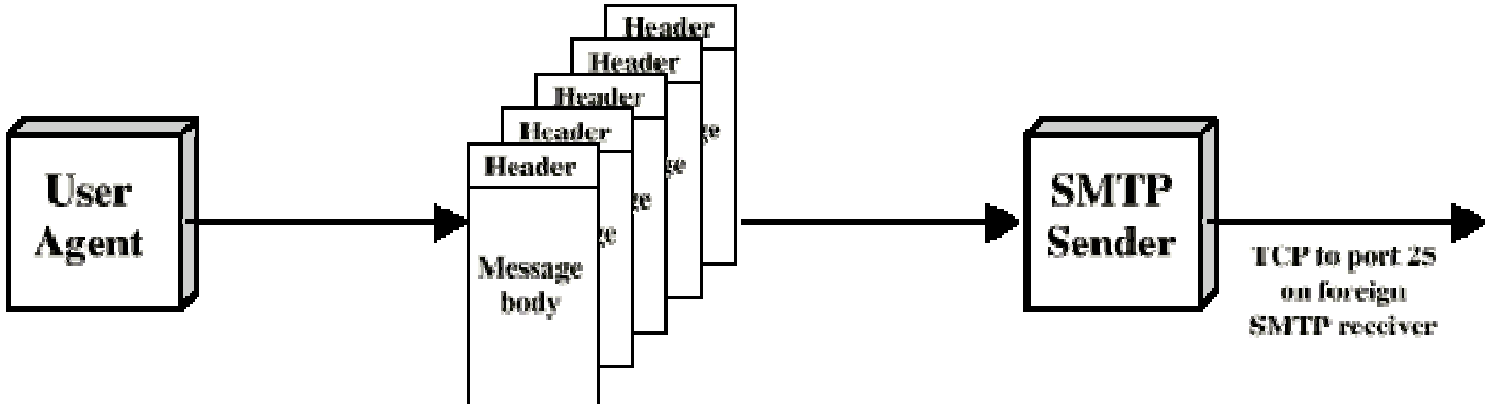
SMTP Receiver

- 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
- Sender responsible for message until receiver confirm complete transfer.
  - Indicates mail has arrived at host, not user.

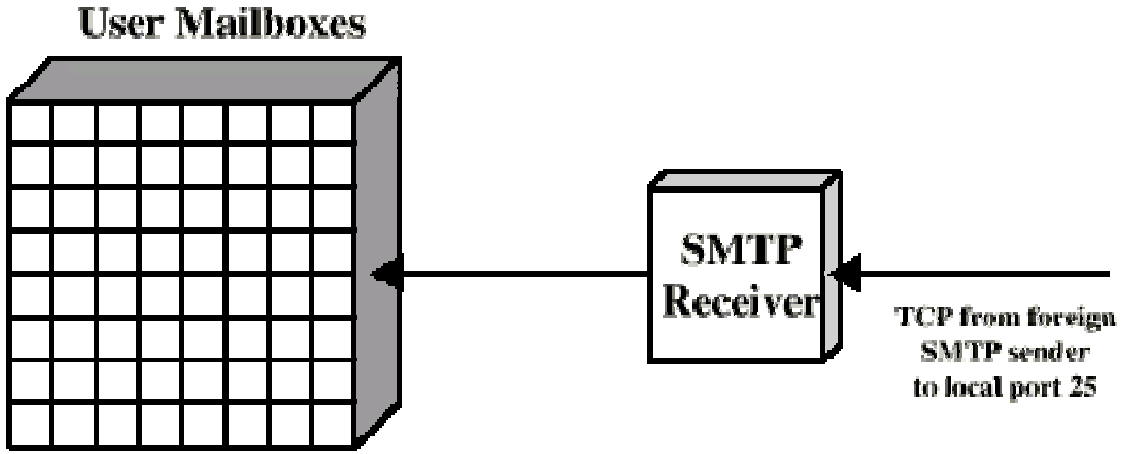
SMTP Forwarding

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

SMTP Mail Flow



(a) Outgoing Mail



(b) Incoming Mail

SMTP System Overview

- Commands and responses 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

- **Leading digit indicates category.**
  - **Positive completion reply (2xx)**
  - **Positive intermediate reply (3xx)**
  - **Transient negative completion reply (4xx)**
  - **Permanent negative completion reply (5xx)**

Operation Phases

- Connection setup
- Exchange of command-response pairs
- Connection termination

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

Mail Transfer

- Sender may send one or more messages to receiver.
- MAIL command identifies originator.
  - Gives reverse path to be used for error reporting.
  - Receiver returns 250 OK or appropriate fail/error message.
- One or more RCPT commands identifies recipients for the message.
  - Separate reply for each recipient.
- DATA command transfers message text.
  - End of message indicated by line containing just period (.)

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

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@hotmail.com>

S: 250 OK

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).**

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/110.
  - Sends user name and password.
  - Can access the mails, one by one.
  - Two modes:
    - Delete mode – mails deleted as they are read
    - Keep mode – mails remain in the mailbox

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 can not transmit nontext / executables.
  - Uuencode and other schemes are available.
    - Not standardized.
- Cannot transmit text including 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**
- **Number of *content formats* defined.**
- **Transfer encoding defined.**

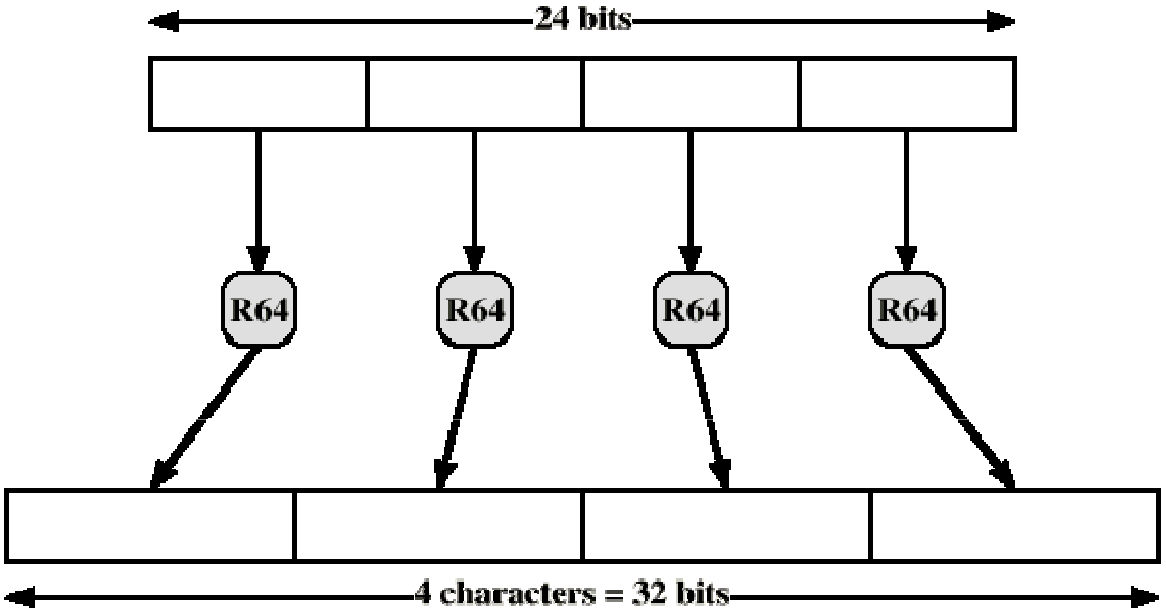
Content Types

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

MIME Transfer Encodings

- Reliable delivery across wide largest range of environments.
- Content transfer encoding field:
  - Six values
  - Three (7bit, 8bit, binary) no encoding done
    - Provide info about nature of data
- Quoted-printable
  - Data largely printable ASCII characters.
  - Non-printing characters represented by hex code.
- Base64
  - Maps arbitrary binary input onto printable output.
- X-token
  - Named nonstandard encoding.

Base 64 Encoding



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