

Lecture 8: Electronic Mail

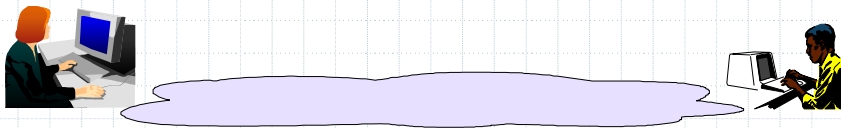
Email architecture
Protocols: SMTP, POP3, IMAP
Message formats: RFC822, MIME

Prof. Yechiam Yemini (YY)
Computer Science Department
Columbia University

Copyright, Yechiam Yemini 1993-03

Goal: Historical Perspective On Applications

- Email laid the foundations for the web
- What is a message?
 - An email text memo --> A tagged document (multimedia document, VLSI design file)
- Mail = Sender-initiated file transfer
- Web = Recipient-initiated file transfer

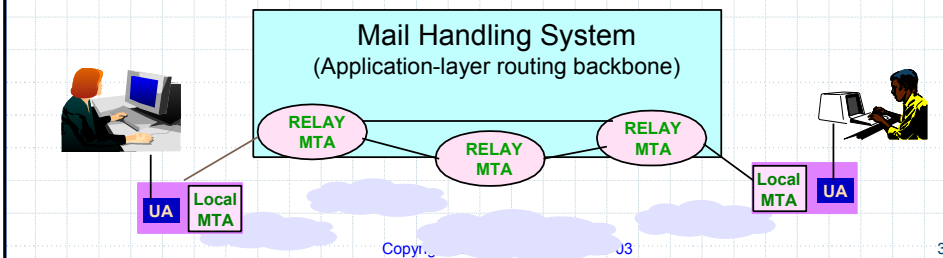


Copyright, Yechiam Yemini 1993-03

2

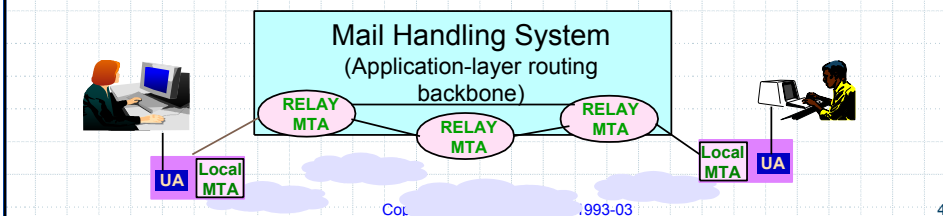
Architecture of Internet Mail

- User Agent (UA): sender & recipient functions
 - Create/present messages; transfer message to mail handling system
 - Why is a UA needed? (vs. an editor and a message file directory)
- Mail Handling System (MHS): routes msgs
 - A network of Mail Transfer Agents (MTA)
 - Why is an MHS needed? (vs. transferring message files between UAs)
- Document encoding: define message structure
 - Msg formats (RFC822, Multipurpose Internet Mail Extensions (MIME))
 - Privacy & security (PEM, PGP)
 - Why is a message structure needed? vs. a tagged document

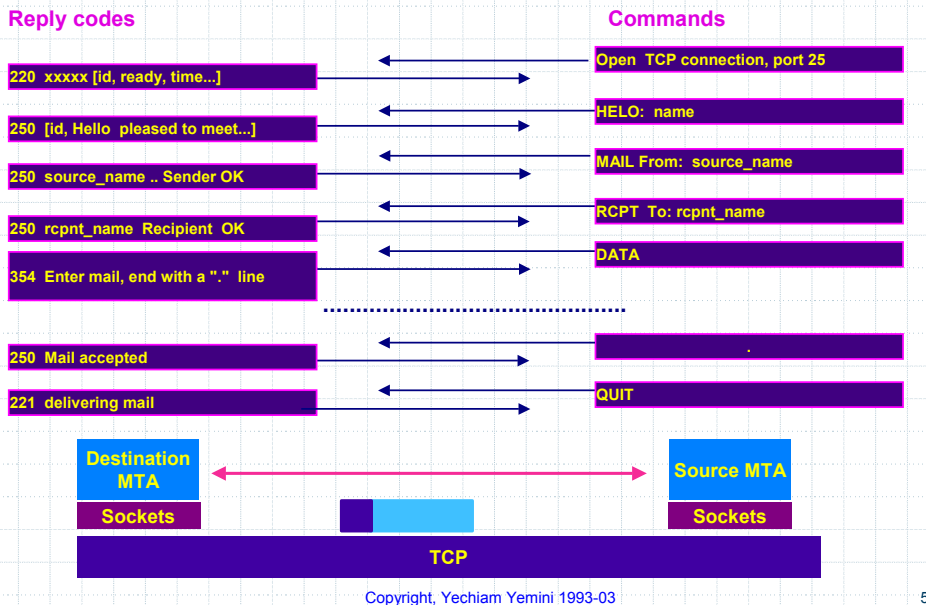


Mail Transfer With SMTP

- SMTP: relay msgs between peer MTAs
 - Reliable & efficient delivery of message files
- Uses DNS services to identify target MTA
 - MX (Mail eXchange) records specify names of MTA serving a particular name
 - Reliability through replication; multiple servers can deliver mail to destination
- Uses TCP to transport message files



SMTP By Example



SMTP By Example

Operations

HELO <domain_name> --Begin relay transaction
 MAIL FROM: <address> -- Identify message source
 RCPT TO: --Identify recipient
 DATA --Begin message file transfer
 . --End of message file transfer
 RSET --Reset a transaction NOOP--Ping the peer MTA, QUIT

Optional

SEND FROM:, SOML FROM:, SAML FROM:--Various delivery forms
 VRFY <string> --Verify mailbox name
 EXP <string>-- Expand a string to list of recipients
 TURN --Pass turn to server to deliver mail to client MTA
 HELP "<string>"

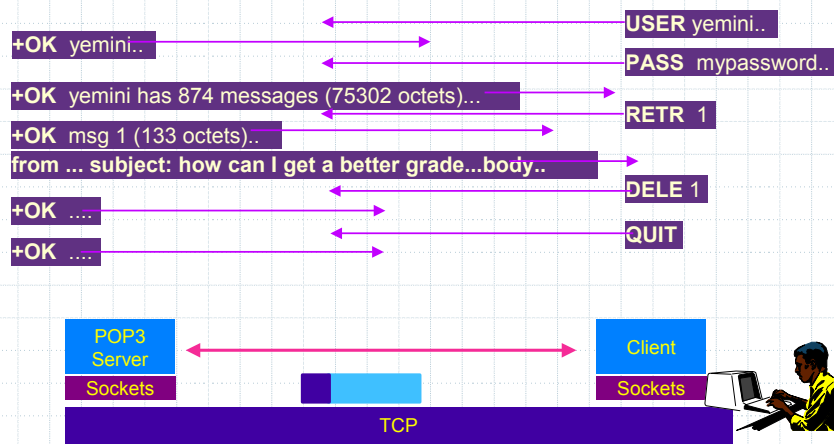
Reply Codes

Completion (first digit): 1 = positive preliminary, 2=positive complete, 3=positive intermediate, 4=transient negative, 5= permanent negative
 Category (second digit): 0= syntax issues, 1=informational, 2=connections, 5=application specific
 Instance (third digit)
 220 MTA available, 250 Identity accepted, 251 Relaying to recipient, 354 Ready for contents
 450 Mailbox busy, 421 MTA unavable, 500 Syntax err, 550 No such rcpt, 552 Exceeded storage

Mail Retrieval by Clients: POP3

- Why are MTAs insufficient?
 - User machines are not always accessible
 - Maintaining mail servers is not simple
- Functions
 - Retrieve/send mail messages from/to Mail server
 - Update mailbox at server
 - Authenticate

An Example POP3 Session



IMAP4: Server-centric email

- Rich server functions
 - Authentication
 - Create/delete and manage multiple mailboxes (rename, select, expunge, checkpoint...)
 - Acknowledgment of operations completion (or not)
- Rich mailbox manipulation primitives
 - Select/find msgs by numerous headers field contents
 - Apply operations to selected messages or their headers
 - Tagging & operations on parts of composite messages
- Contrast with POP3
 - Maximally-featured mailbox/message manipulations vs. Mail file downloading
- Why have server-centric email?
 - Contrast with SMTP (peer-peer) and POP3 (client-centric)

IMAP4 Example

```
S: * OK IMAP4 Service Ready
C: a001 login mrc secret
S: a001 OK LOGIN completed
C: a002 select inbox
S: * 18 EXISTS
S: * FLAGS (\Answered \Flagged \Deleted \Seen \Draft)
S: * 2 RECENT
S: * OK [UNSEEN 17] Message 17 is the first unseen message
S: * OK [UIDVALIDITY 3857529045] UIDs valid
S: a002 OK [READ-WRITE] SELECT completed
C: a003 fetch 12 full
S: * 12 FETCH (FLAGS (\Seen) INTERNALDATE "14-Jul-1993 02:44:25 -0700"
RFC822.SIZE 4282 ENVELOPE ("Wed, 14 Jul 1993 02:23:25 -0700 (PDT)"
"IMAP4 WG mtg summary and minutes"
(("Terry Gray" NIL "gray" "cac.washington.edu"))
((NIL NIL "minutes" "CNRI.Reston.VA.US")
("John Klensin" NIL "KLENSIN" "INFOODS.MIT.EDU")) NIL NIL
"<B27397-0100000@cac.washington.edu>")
BODY ("TEXT" "PLAIN" ("CHARSET" "US-ASCII") NIL NIL "7BIT" 3028 92))
S: a003 OK FETCH completed
.....
C: a006 logout
S: * BYE IMAP4 server terminating connection
S: a006 OK LOGOUT completed
```

Message Formats (RFC822)

- Messages are text memos
 - Headers provide delivery information
 - Body provides text
- Syntax is formally defined by BNF
 - A message is a text string with a simple structure
 - BNF defines a regular grammar to generate/parse msgs

```
Return-Path: <stranger@mail.mycompany.com>
From: john_stranger@mail.mycompany.com
Date: Thu, 17 Oct 96 16:35:48 PST
To: yemini@cs.columbia.edu
Cc: netbook@cs.columbia.edu
Subject: Please help
```

Sample Top-Level Syntax Definitions

```
message= fields *( CRLF *text)
                ; Everything after first null line is msg body
fields=dates      ; Creation time,
    source        ; author id & one
    1*destination; address required
    *optional-field; others optional
source=[ trace ] ; net traversals
    originator   ; original mail
    [ resent ]  ; forwarded
trace= return     ; path to sender
    1*received   ; receipt tags
return= "Return-path" ":" route-addr ; return address
received="Received" ":" ; one per relay
    ["from" domain] ; sending host
    ["by" domain]   ; receiving host
    ["via" atom]    ; physical path
    *("with" atom)  ; link/mail protocol
    ["id" msg-id]   ; receiver msg id
    ["for" addr-spec]; initial form
    ":" date-time   ; time received
```

Sample Return Header

```
Return-Path: <Mailer-Daemon@pop-pe.rnp.br>
Date: Fri, 4 Oct 96 20:18:06 EST
From: Mail Delivery Subsystem <Mailer-Daemon@pop-pe.rnp.br>
Subject: Returned mail: Service unavailable
To: <yemini@cs.columbia.edu>
----- Transcript of session follows -----
Connected to rosa.pop-pe.rnp.br:
>>> HELLO pop-pe.rnp.br
<<< 553 pop-pe.rnp.br host name configuration error
554 <dafi@di.ufpe.br>... Service unavailable
----- Unsent message follows -----
Return-Path: <yemini@cs.columbia.edu>
Received: from setton.cs.columbia.edu by pop-pe.rnp.br (4.1/SMI-4.1)
id AA14076; Fri, 4 Oct 96 20:18:06 EST
Received: from hudson.cs.columbia.edu (hudson.cs.columbia.edu [198.29.104.123]) by
setton.cs.columbia.edu (8.6.12/8.6.12) with SMTP id TAA16972 for <dafi@di.ufpe.br>; Fri, 4
Oct 1996 19:07:42 -0400
Date: Fri, 4 Oct 1996 19:07:42 -0400
Message-Id: <199610042307.TAA16972@hudson.cs.columbia.edu>
X-Sender: yemini@cs.columbia.edu
X-Mailer: Windows Eudora Pro Version 2.1.2
Mime-Version: 1.0
Content-Type: text/plain; charset="us-ascii"
To: Danilo Florissi <dafi@di.ufpe.br>
From: Yechiam Yemini <yemini@cs.columbia.edu>
Subject: Thank You For the Fun in Brazil
```

Multipurpose Internet Mail Extensions (MIME)

- The need
 - Structured message objects (e.g., multimedia mail, user-defined document structures)
 - Full compatibility with SMTP/RFC822
- Idea: extend RFC 822 headers
- MIME header fields
 - Mime-Version: ; version number
 - Content-Type: ; defines the message structure
 - Content-Transfer-Encoding: NVT ASCII, quoted printable, ... binary, ...
 - Content-Id:
 - Content-Description:

Functions

- Pass richer message objects
 - Text, formatted text, image, video, audio....
- Support multimedia objects
 - Synchronize presentation of text, image, video, audio
- Support active components
 - E.G., A message can retrieve a file from an FTP server
- Enable open-ended addition & registration of new object types

Content Type

- Text: various forms of texts (simple->rich->richer)
- Multipart: multiple objects; mixed(sequential), parallel...
 - Can be used to nest composite objects
- Message: RFC822, partial (fragment of a message), external-body (pointer to a message)
- Application: octet-stream (binary), oda (ISO open document interchange format), postscript
- Media objects
 - Image: jpeg, gif, G3fax
 - Audio: basic (ISDN mu-law codec)
 - Video: mpeg
 - Voice-mail messages
- Access type: Invoke methods (e.g., FTP) to retrieve object
- New Types: Register with IANA

Sample Applications

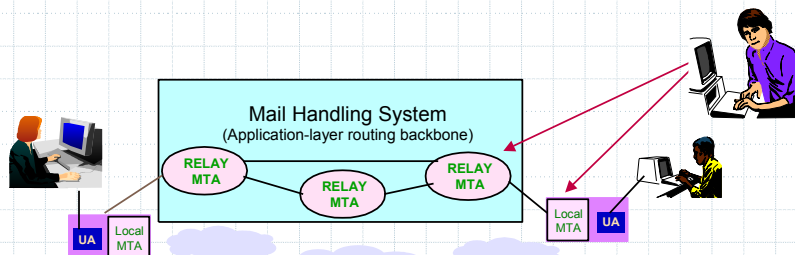
- Multimedia mail
 - Create a multipart document, possibly defining parallel play of audio, video and text
- Retrieve documents (predecessor of URL)
 - Create a message object with External-Body providing reference to document
- Pass programs (predecessor of CGI)
 - Transmit postscript documents
 - Pass a Tcl/Tk form
 - Transmit HTML documents
 - Will need html-enabled mail user agent
 -

```
To: class
Subject: Assignment 2
Mime-Version: 1.0
Content-Type: Message/External-Body
  access-type="mail-server"
  server="mail@cs.columbia.edu"
```

```
To: class
Subject: Assignment 2
Mime-Version: 1.0
Content-Type: Message/External-Body
  name="assignment2.txt"
  site="cs.columbia.edu"
  access-type="anon-ftp"
  directory="cs4119"
```

Email Insecurities

- Virus distribution through MIME attachments
- SPAM
- Other
 - Intercept POP3 messages to compromise users
 - Gain info about domain through SMTP queries (VRFY, EXPND)
 - ...



Focusing on the Fundamentals

- Mail & Web are dual service forms
 - Mail provides sender-controlled file dispatch
 - Web provides receiver-controlled file retrieval
- What is a "message"?
 - Tagged document
- Why not unify both?
 - A protocol to transfer files from storage to storage
 - A protocol for processing tagged documents

Rethinking Email

- Can email be redesigned to:
 - Eliminate SPAM and viruses?
 - Enable us to filter and prioritize messages (manage large message spaces)?
 - Simplify email administration
 - Enable organizations to manage their responsibility for email communications