

INTERNET Email - an Introduction and Overview

Jim Jackson <jj@franjem.org.uk>

November 13, 2006

Contents

1 In the Beginning	3
2 What is an Email?	4
3 Email Header Fields	5
4 Hello World!	6
5 Getting Email from here to there	7
6 The SMTP conversation	8
7 Envelope Sender and Recipient(s)	9
8 Email Delivery	10
9 Reading Email	11
10 Post Office Protocol - POP	12
11 Internet Message Access Protocol - IMAP	13
12 Web Based Mail reading	14
13 Typical Mail Services	15

This document can be found at <http://www.comp.leeds.ac.uk/jj/linux/InternetEmail.html>

1 **In the Beginning**

- Email was one of the first Internet Applications
 - Originally specified in RFC's in August 1982
 - Current Internet Email builds on and expands the early standards, but is essentially much the same
-

2 What is an Email?

- An Email is a file or sequence of bytes that matches the format described in RFC2822, the update to the original RFC822.
 - Such a file has 2 parts:
 - a Header consisting of a sequence of fields have values that give information about the Email in a machine readable format, e.g. To:, From:, Date:
 - an optional body consisting of lines of text
 - A blank line separates the header from the body
-

3 Email Header Fields

- An Email header field consists of the fieldname, a colon (':') and the value of the field.
 - Some field values are fairly free form, e.g. Subject:
 - others have a strict structure, e.g. Email addresses in To:, From:, Cc: fields
-

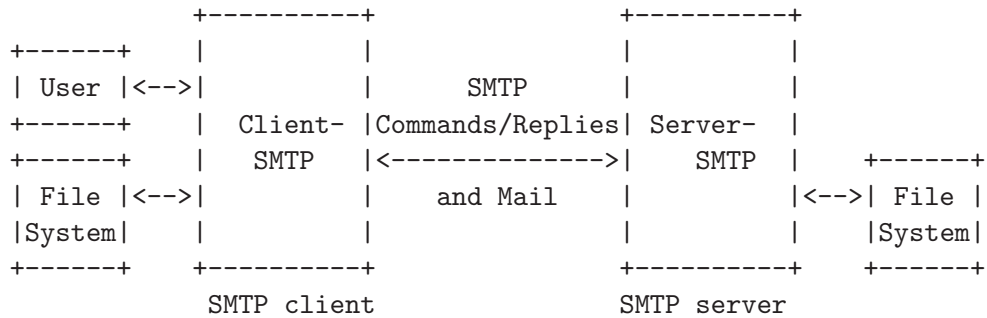
4 Hello World!

From: Jim Jackson <jj@franjem.org.uk>
To: all@world.com
Subject: Message to Everyone

Hello World!

5 Getting Email from here to there

- Email is sent using **Simple Mail Transport Protocol** aka **SMTP** . First defined in RFC821 in Aug 1982, currently defined by RFC2821
- The model.....



6 The SMTP conversation

```

                <----- 220 Greeting message
HELO system.name ----->
                <----- 250 servername pleased to meet you
Mail From: user@add.com
                ----->
                <----- 250 user@add.com... sender ok
Rcpt To: name@host.edu
                ----->
                <----- 250 name@host.edu... Recipient ok
Data
                ----->
                <----- 354 Enter mail, end with "."
                        on a line by itself
RFC2822 stream of bytes...
                ----->
.
                ----->
                <----- 250 RefID Message Accepted
quit
                ----->
                <----- 221 system.name closing connection
```

7 Envelope Sender and Recipient(s)

- The email address used in the SMTP *Mail From:* line, the *envelope sender*, is not necessarily the same as the address in the RFC2822 *From:* field of the Email.
- However, the *envelope sender* **is** the address to which any Delivery Status Notifications get delivered
- The *envelope recipient(s)* in the SMTP *Rcpt To:* fields are also not necessarily the same as the recipients in the Message's *To:* and *Cc:* fields.

Think of the example of an email to an email list. It is composed *To: elist@address*, gets to the elist site. The elist software then sends out the email message to all the people on the list, changing the *envelope recipients*, and changing the *envelope sender* to be the address of the elist admin - so they get any DSN's.

spammers just love the **simple** in _SMTP :-)

8 Email Delivery

- Email eventually reaches an **SMTP** server that actually delivers the email. How this is done is not actually defined. There are several commonly used formats for storing delivered mail:
 - Mbox format - server prefixes the RFC822 message with a one line special header and appends this to a file designated as the users **INBOX** , see <http://en.wikipedia.org/wiki/Mbox>
 - Maildir format - each message is kept in a separate file - does away with need for file-locking, see <http://en.wikipedia.org/wiki/Maildir>
 - Proprietary Database usage e.g. Microsoft Exchange
 - Mbox and Maildir are well known text based formats and make it easy to interface with Email Reading software. Standard Text tools can be pressed-ganged into use diagnosing and fixing mail problems.
-

9 Reading Email

- User's mail software can use one or more of the following methods:
 - direct access to the **Mbox** file or **Maildir** directory the email is delivered into.
 - use **POP** - RFC1939
 - use **IMAP** - RFC3501
 - use a Web Interface to access the mail
 - Direct access is often not possible. **POP/IMAP** are client/server remote access methods - **POP** being the most commonly available.
-

10 Post Office Protocol - POP

- A simple protocol for picking up mail from a remote mail server. Very commonly implemented.
 - provides facilities for
 - authentication before allowing access
 - querying status of mailbox
 - getting summary list of contents of mailbox
 - retrieving messages from mailbox
 - deleting messages in mailbox
 - most POP clients do not make full use of facilities POP provides - they usually just provide a service to download new messages, and to then optionally delete them.
 - the downside to POP is that once the email has been downloaded it tends to be only available on the machine to which it has been downloaded.
 - see http://en.wikipedia.org/wiki/Post_Office_Protocol
-

11 Internet Message Access Protocol - IMAP

- A more complex protocol for accessing and manipulating incoming mail and mail folders on the server.
 - besides the basic facilities provide by POP, IMAP allows users to have mail folders reside on the server. An IMAP client can move mail between remote folders, or between local and remote folders, or even between folders on different IMAP servers.
 - IMAP is expensive for ISPs to provide and hence is usually only available as a paid for extra service. It big advantage is that the user has access to email in the server side folders from any machine they use, where ever they are. But messages can be downloaded whenever they are needed locally.
 - see http://en.wikipedia.org/wiki/Internet_Message_Access_Protocol
-

12 Web Based Mail reading

- The simplest client side solution - usually only a standard web browser is required.
 - The mail resides remotely on the service providers servers, and is accessible from anywhere, including Internet Cafes etc.
 - Downside is you are stuck with a fixed interface and facilities, the provider limits the amount of mail you can keep on their servers, dealing with attachments etc can be less than easy.
 - many providers support Web access as well as POP access.
-

