Email Administration

Don Porter

What is an email?
- A simple message format:
  - Envelope (server-internal routing info, not user-visible)
  - Headers (basically the history of the message handling)
  - These are viewable in your email app
  - Body (the text you see in your email app)

Mail Transport Agent (MTA)
- Accepts emails from the Internet
  - Delivers to local users
  - Or sends outgoing messages
- Every site that accepts email runs one
  - Identified with an MX record in DNS
- Listens on Port 25

Simple MTA Example

Example email header:

Theme
- Email systems started with a pretty simple design
  - Everyone trusted each other, it was nice
- Then spam came along
  - Lots of complexity and distrust to try to reduce spam
Simple MTA Example, redux

From: don@google.com
To: porter@stonybrook.edu

MTA

25

Da' Internet

storybrook.edu

25

MTA

googie.com

25

U A

25

Da’ Internet

storybrook.edu

25

MTA

googie.com

25

MTA

Where is my inbox?

• On Unix/Linux, it is either a file or directory
• 2 Popular formats:
  – mbox – single file
  – maildir – directory with one file per-message
• Literally, a file like: /home/porter/mail/mbox
  – Each message has a special delimiter between it
• Maybe shared with other machines over NFS
  – Users can get their mail without logging into MDA machine

Simple MTA Example

From: don@google.com
To: porter@stonybrook.edu

MTA

25

Da’ Internet

storybrook.edu

25

MTA

googie.com

25

MTA

So what is Apple Mail, or Outlook?

• A User Agent (UA)
• Usually provides a program to type emails
• Ultimately packaged and sent to an MTA using SMTP

So how does email get to my inbox?

• Once an email reaches its destination MTA,
  – Handed off to a mail delivery agent (MDA)
  – MDA can be the same program or a different daemon as MTA
• MDA finds inbox for the recipient and stores the email

SMTP

• Simple Mail Transport Protocol
• It really is simple.
• Main operations:
  – Send a message
  – Check if an address is valid
  – Expand an address (for lists and forwarding)
• Email basically works by lots of MTA servers passing messages to each other
Old-School Mail Reading
• Programs like mutt, pine, elm would read these mail files directly on a server
  – You could even open them in vi or emacs!

Text-based email reading (mutt)

From: http://greek0.net/mutt.html

Pointy-Clicky Mail Reading
• Not everyone runs a command line on a Linux server
• Mail for the rest of us?
  – 2 popular protocols: POP and IMAP
  – Make a nice GUI app, like Thunderbird or Apple Mail
  – Download inbox using POP or IMAP
  – Still send using SMPT
• Access Agent (AA) – serves POP or IMAP

POP vs. IMAP
• Post Office Protocol (POP)
  – Download everything and (usually) delete it from inbox on server
  – Designed when people had one PC
• Internet Mail Access Protocol (IMAP)
  – Sync a local copy of mailbox with server
  – Work offline and sync later
  – Multiple clients
    • Laptop, desktop, iphone, ipad
  – Better choice for most modern users

Recap so far…
• User Agent: Program for human user
• Sending mail?
  – Protocol: SMTP
  – Emails exchanged by MTAs
  – Delivered to inboxes by MDAs
  – Inboxes are just simple files (mostly)
• Receiving mail?
  – Read inbox through an AA
  – Protocols: POP or IMAP
SMTP Review

- Main operations:
  - Send a message
  - Check if an address is valid
  - Expand an address (for lists and forwarding)

- Anything missing?

Are my emails private?

- No
- Sent in plain text over internet, inbox in plain text
- If you want privacy, use encryption in your UA
  - PGP/GPG are good choices

How do I know email came from sender?

- You don’t
- Anyone can put any name/email address in the ‘from’ field
  - Of course, replies may not go to you…
- This is why people who care use PGP/GPG to sign messages

Spam

- Junk email
  - Unwanted ads, harassing emails, etc.
  - Often selling illegal products
- Why is it called spam?
  - A massive nuisance
    - But also a massive business
- A lot of modern system organization designed to limit/prevent spam
  - We will refine previous model of email to cope with spam

Example Recap

- User downloads software
  - Software includes “trojan horse” malware
  - Malware connects to email servers and sends spam from user’s computer over SMTP
  - Other SMTP servers then relay the email throughout the internet
- Ideas how to fix this?
Refinement 1

- SMTP servers (as presented) are dumb
  - Accept and relay mail from anyone!
- Idea: Only send mail from authenticated users
- Mail Submission Agent (MSA)
  - Basically, provides a log-in system, and then forwards mail to MTA
  - MTA configured to only relay mail from MSA
  - MSA often listens on port 587

Putting it together

Refinement, recap

- In the “good old days”, any email server would send your mail for you
  - Made email (and spam) easy
- Now, only servers that know you will relay email for you
  - Spam program at least has to steal your account info now

Problem 2

- The first example was about relaying
  - Using Stonybrook.edu to relay email to lots of other addresses (gmail, hotmail, etc)
  - Fixed in part by:
    - Only accepting new email for legitimate users in Stonybrook.edu
    - Only relaying email from authenticated Stonybrook.edu users
- What about spam for users of Stonybrook.edu?

Refinement 2

- Be choosier about who you accept email from
- Whitelists: known-good email servers (accept)
- Blacklists: known-bad email servers (reject)
  - Services that provide both for administrators
- Unknown servers? (Gray lists)
  - Reject: May lose email
  - Delay: Spammers often impatient

Refinement 3: Spam Detection

- Before the MDA delivers spam, run it through a filter
- If it passes, deliver to inbox
- If it fails, put in Junk folder
- How do spam filters work?
2 Strategies for Spam Filters

- Fuzzy matching against known spam
  - Verbatim matches foiled by including time of day
  - Services track these things for admins
- Bayesian learning filters
  - Users mark things as ‘spam’ or ‘ham’
  - Basically training a fancy-dancy machine learning classifier
  - Classifier applied to new mail
  - Learns user preferences over time

2 Phase Spam Prevention

- My advice (and the book’s):
  - Do quick checks at the MTA (white list, black list)
  - Then do detailed checks (spam filtering) in MDA
- Why?
  - Quickly drop obvious garbage
  - Shift load for heavier-weight scanning to MDA

Summary

- MTAs exchange mail all over internet
  - Only relay outgoing mail from MSA (to prevent spam)
  - Only accept incoming mail from white/gray list (to prevent spam)
- MDA delivers mail
  - After it passes a filter (to prevent spam)
  - Stores it in a simple inbox file
- UA gets mail using IMAP/POP
  - Sends it via MSA using SMTP