Internet Mail — Past, Present, and (a bit of) the Future

Eric Allman
Chief Science Officer
Sendmail, Inc.
March 2009
Introduction and Outline

- A history of email and related technologies from the very early days through present
- Some observations about spam, email security, authentication, and reputation
- Very fuzzy speculations about the future of email and messaging in general
Internet Pre-History
1968–1981

Arpanet: 0–213 hosts
12/68: first hypertext demo

9/69: first Arpanet node
Genesis of the Arpanet

- Sprung into being September 2, 1969
- One host at UCLA
- No one to talk to and no place to go....
THE ARPA NETWORK

SEPT. 1969

FIGURE 6.1 Drawing of September 1969
(Courtesy of Alex McKenzie)
Genesis of the Arpanet

- Sprung into being September 2, 1969
- One host at UCLA
- No one to talk to and no place to go....
- Soon, hosts added at Stanford Research Institute, University of California Santa Barbara, and University of Utah (one per month)
Arpanet Late 1969 (Logical View)
12/68: first hypertext demo
9/69: first Arpanet node
Late '71: first email sent
Ray Tomlinson — the Real Father of Email
Ray Tomlinson

- Really is the Father of Email
- Bolt Beranek & Newman
- Email just append-only file transfer to a special file
- Late 1971: linked BBNA and BBNB
The First Two Email Hosts
12/68: first hypertext demo
9/69: first Arpanet node
Late '71: first email sent
11/71: UNIX 1.0
Ken Thompson and Dennis Ritchie
1973

Beginning of a truly exciting time at Berkeley
Bill Joy
Beastie
12/68: first hypertext demo
9/69: first Arpanet node
Late '71: first email sent
4/77: RSA
11/71: UNIX 1.0
Fall '74: “Merkle Puzzles”
The ARPANET at Berkeley

- ARPANET connection to the INGRES PDP-11/70 for Distributed Database Research (9600 baud!)
- Everyone in the CS Division wanted an account
- PDP-11 couldn’t handle that many simultaneous logins (not enough memory or RS-232 ports)
A PDP-11/70 (but not ours)
When Computers had Switches and Lights…
What people really wanted was email, not full access (they didn’t need telnet or FTP)
BerkNET linked internal machines using RS-232 lines (Eric Schmidt)
Also had a UUCP connection from Ernie CoVAX (main department machine)
Delivermail: forward email between nets (quick hack)
Released in 4.0 BSD (October 1980); very popular
ARPANET GEOGRAPHIC MAP, OCTOBER 1980

SATELLITE CIRCUIT
○ IMP
☐ TIP
△ PLURIBUS IMP
◊ PLURIBUS TIP
leştirme C30

(NOTE: THIS MAP DOES NOT SHOW ARPA’S EXPERIMENTAL SATELLITE CONNECTIONS)
NAMES SHOWN ARE IMP NAMES, NOT (NECESSARILY) HOST NAMES
Internet History
1981–1988

213–33,000 hosts
10/81: I start implementing SMTP in sendmail

4/82: 4.1a BSD ships with sendmail

8/82: RFC821, 822 published

1/1/83: Internet “flag day”

9/83: 4.2 BSD ships
Beastie in color
Internet History
1988–1993

33,000–1,300,000 hosts
11/88: Internet worm

3/89: Tim Berners-Lee proposes WWW
Tim Berners-Lee
11/88: Internet worm

3/89: Tim Berners-Lee proposes WWW

7/91: CIX founded

6/92: MIME
Internet History
1993–1996

1,300,000–12,900,000 hosts
2/93: NCSA Mosaic released
6/93: sendmail 8 released
3/94: Netscape founded
4/94: “green card” spam
Spam

- Network previously largely cooperative
- Flame wars all too common but isolated
- Spam not new: September 13, 1904 — unsolicited commercial email via telegraph
- DEC spam: May 1, 1978
- “Make Money Fast” chain letter: 1988
- Attitudes change: “Greed is Good” comes to the Internet
- Canter and Siegel (Green Card Spam) were unapologetic
2/93: NCSA Mosaic released
6/93: sendmail 8 released
9/95: Sendmail Consortium
3/94: Netscape founded
4/94: “green card” spam
4/94: WebCrawler
11/94: “Good Times” pseudo-virus
10/95: S/MIME
Internet History
1996–2000

12,900,000–109,000,000 hosts
7/96: Hotmail founded
5/97: Exchange gets SMTP
10/97: Yahoo! founded
12/97: Microsoft acquires Hotmail
3/98: Sendmail, Inc. founded
3/99: Melissa virus
1/00: encryption policy relaxed
3/00: milter released
4/00: I Love You virus
Internet History
2001–2009

109,600,000–439,000,000 hosts
Internet June 1999

Burch/Cheswick map of the Internet showing the major ISPs. Data collected 28 June 1999

Copyright (C) 1999, Lucent Technologies
Messaging, Spam, Security, and Authentication
Email Spamming

- **Economic issue:** $C$(sending) $\ll$ $C$(receiving)
  - Note: cost function $C$ is not denominated in $\$ or ¥
  - Can be CPU time, memory use, human time, etc.

- **Possible approaches**
  - Content filtering (reaching limits; doesn’t fix cost function)
  - ePostage (infrastructure & acceptance problems)
  - Challenge-Response (poor scaling; user confusion)
  - HashCash (useless with zombie farms)
  - Graylisting (easy to defeat; broken by server farms)
  - Authentication (insufficient by itself)
Phishing

- Try to trick someone into giving up private information, sending money, etc.
- Not limited to email
- Nothing new, just easier to do
  - “Nigerian” or “419” attacks predated email
  - Shysters prey on elderly
- “Shotgun” phishing overlaps with spam
- “Spear phishing” changes the rules
  - Can spend significant money targeting someone
  - They usually have significant information about you
- Attackers often try to pretend to be someone you trust
  - Authentication can really help with this
Issues with Authentication

- Choosing the standard(s)
  - Path-based (SenderID/SPF)
  - Signature-based (DKIM, DomainKeys)
  - Neither authenticates users, just SMTP nodes
  - Both can break on non-malicious messages

- How to achieve adoption?
  - Sender- or Recipient-driven?
  - Senders seem to be the primary driver

- Status of unsigned email
  - Unsigned mail must remain legal during transition
  - Author Domain Signing Practices (ADSP) tells how to treat unsigned mail (DKIM only)

- Authentication by itself is not enough
DKIM details

- DKIM contains two specifications
  - “Base” signing specification — how to sign an individual message
  - “Signing Practices” — how to interpret unsigned messages
- Base spec (RFC4871) published May 2007
  - Signs body and selected headers using keypairs
  - Public key management done in DNS
- Signing Practices is very controversial
  - Even the name changed a few times — ultimate name: Author Domain Signing Practices (ADSP)
  - Finally got watered down to the point where there was nothing controversial left
Issues with Reputation

- Negative reputation is well known
  - E.g., DNS blackhole lists
  - Can change very quickly

- Positive reputation is hard without “breaking” the world we know today
  - Does “presumption of innocence” become “presumption of guilt”?  
  - How does the smaller player join the club?

- Will we go to a “closed society” email model?
  - New domains will have no reputation, so recipients may be unwilling to accept their mail
  - This is (sort of) what Challenge-Response does

- Accreditation for a fee?
Other Messaging Not Immune from Abuse

- **SPIM (Instant Messaging Spam) becoming more common**
  - Authentication helps, but too easy to get accounts
- **Social Networking sites being targeted**
  - MySpace and Facebook have both been targeted
  - Often successful because messages seem to come from “friends”
  - Particularly good for phishing
- **SPIT (Internet Telephony Spam) is a growth industry**
Some Speculations on the Future
Email of the Future — Short Term Predictions

- Authenticated email with (some) sender reputation, growing with time
- Slow movement toward Email Address Internationalization (EAI IETF Working Group)
  - Downgrading is very hard to do and may not happen at all
- Young people moving off SMTP-based email, using IM, Social Networking, micro-blogging instead
Email of the Future — Medium Term Predictions

- **SMTP is dead, long live SMTP (?)**
  - Often proposed and predicted, but so far SMTP survives
  - Could fix some problems, but…
  - SMTP (with extensions) is “good enough”

- **UTF-8 everywhere**
  - Probably will happen; not hard to do (except for downgrading)

- **Most email will be at hosted providers**
  - Getting ever harder to build and operate a mail system due to increasing challenges and demands
Random Longer-Term Predictions

- Casual messaging will continue to move away from traditional email
  - People crave instantaneous gratification
  - Fits better with mobile usage
- Email will not die
  - Too well suited to business
  - Need for longer, more considered messages
  - Security and regulatory constraints
  - Cheaper than texting (for now; this cost is artificial)
  - IM doesn’t work well across time zones
- Distinction between email, IM, voice will blur
- Interesting work: Computer/Human interface
The Unified Messaging Mess (non-email)

- Too many messaging services and applications with too much overlap, not enough compat

<table>
<thead>
<tr>
<th>Service/App</th>
<th>IM</th>
<th>SMS</th>
<th>Voice</th>
<th>Video</th>
<th>Inc Calls</th>
<th>Out Calls</th>
<th>Call Fwd</th>
<th>VM</th>
<th>Addr Book</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adium</td>
<td>✔</td>
<td>✔</td>
<td>✘</td>
<td>✘</td>
<td>✘</td>
<td>✘</td>
<td>✘</td>
<td>✗</td>
<td>✔</td>
<td>IM</td>
</tr>
<tr>
<td>Skype</td>
<td>✔</td>
<td>+</td>
<td>✔</td>
<td>✔</td>
<td>+</td>
<td>+</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>Voice</td>
</tr>
<tr>
<td>Google Voice</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✘</td>
<td>✔</td>
<td>US Int</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>Call Forward</td>
</tr>
<tr>
<td>Broad voice</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✘</td>
<td>✔</td>
<td>☛</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>VoIP, $</td>
</tr>
<tr>
<td>Gizmo</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>+</td>
<td>+</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>VoIP soft</td>
</tr>
<tr>
<td>iChat</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✘</td>
<td>✘</td>
<td>✘</td>
<td>✔</td>
<td>✗</td>
<td>IM</td>
</tr>
</tbody>
</table>

ernie:ucbvax!decwrl!x2300::mike@berkeley.edu
Thank You

Gratuitous Advertising:
Conference on Email and Anti-Spam (CEAS)
July 16–17, 2009
Mountain View, California
http://ceas.cc