Enhancing E-mail Reliability

Patrick Chu
Supervisor: Dr. Tim Moors
Assessor: Dr. Mohammad Rezvan
Outline

- Background
  - Is E-mail reliable?
  - What have been done?
- Strategies
  - End-users enhancement
  - Alternatives
- Project Proposal
- Project Schedule
Electronic Mail

- Invented in late 1971 by Ray Tomlinson
- Developed as network mail for the ARPANET originally
- Evolved in 1993 on the Internet
  - America Online and Delphi connected their proprietary e-mail systems to the Internet
  - Large scale adoption of Internet e-mail as a global standard began
E-mail (cont.)

- Advantages:
  - Simple
  - Fast
  - Inexpensive
  - Asynchronous

- Question:
  - “Is it always reliable?”
Scope of Thesis

- Enhance reliability of e-mails towards end-users
  - Develop a system to report e-mail delays and losses in real time
  - Detection of lost mails
  - Debugging tool to detect possible location of errors
  - Plug-ins for popular mail clients
Review of E-mail System

Key Elements of the Internet Mail System:

- **Mail User Agent (MUA)** — The client program in which the user sends and receives mail.

- **Mail Transfer Agent (MTA)** — A mail server program which enables email transfers from one machine to another. N.B. MTAs are not “Mail Transport Agents”, because they do not actually deliver mail themselves.

- **Mail Delivery Agent (MDA)** — A program which the MTA uses to put messages into a user's mailbox or to transport mail to another MTA.

- **Mail Retrieval Agent (MRA)** — A program or service which fetches messages from a mailbox on a remote server and passes them to a MUA.
Review of E-mail System

http://ebusiness.gbdirect.co.uk/howtos/mail-system.html
Mail Protocols

- SMTP
  - RFC 2821
  - Use between MTAs to forward mails
- POP3/IMAP
  - RFC 1939, RFC 3501
  - Use to retrieve mails from server
- HTTP
  - Web protocol to access mailbox through web interface
- All base on **reliable** (TCP/IP) connections
Is E-mail Reliable?

- An E-mail dependability thesis has been done in 2004 by Anthony Lang (Telecommunication, UNSW)
- **Reliable** mail services (Hotmail, Yahoo!, UNSW) lost a few percentages of mails
- **Unreliable** services lost at most 10% of mails and unrecoverable
Why mails get lost?

- MTAs fail to deliver/receive mails:
  - System crashes
  - Network problems
  - Spam/anti-virus filters drop valid mails

- MUAs fail to send/retrieve mails:
  - Wrong receipt address
  - Target mailbox full
  - Error during retrieving mails from server
Existing Reliability Enhancement

- **Physical/Hardware**
  - Backup mail server
  - Proxy server
  - Redistribute services

- **Protocol/Software**
  - Alias in Mail eXchanger of Domain Name System (MX field in DNS)
  - Effective Spam filtering
  - Delivery Status Notification (DSN)
  - Message Disposition Notification (MDN)
  - Pre-send test mail
Backup Mail Server

- Secondary mail server (or more)
- Take over the send/receive process when primary server goes down
- Synchronize mails and logs when primary one comes online again
- Advantages:
  - Fully automatic
  - No mail is lost due to server crashes
- Drawback:
  - Extra Maintenance
Backup Mail Server

http://www.webservio.com
Proxy Server

- Mails are sent to a proxy mail server instead of real MTAs
- Proxy server can communicate with MTAs
- Mails are forwarded to lowest loading server

Advantages:
- Load balancing
- Transparent to users

Drawback:
- Extra Maintenance
- Single point of failure
Proxy Server
Redistribute Services

- Redistribute spam and anti-virus scanning process to other physical servers

- Advantages:
  - Control workload of different servers
  - May lower overall system cost
  - Can be combined with proxy load balancing

- Drawback:
  - Extra maintenance
Redistribute Services

- **Example:** Test mail From Hotmail to UNSW (mail header shown)

Return-Path: testmail@hotmail.com
Received: from smtp3.unsw.edu.au ([unix socket])
    by thunderbird (Cyrus v2.1.12) with LMTP; Mon, 05 Sep 2005 16:03:52 +1000
X-Sieve: CMU Sieve 2.2
Received: from smtp-dist.unsw.edu.au (smtp-dist-01.services.comms.unsw.EDU.AU [149.171.97.16])
    by smtp3.unsw.edu.au (8.11.2/8.11.2) with ESMTP id j8563pw18663
    for <test_student@pop3.student.unsw.edu.au>; Mon, 5 Sep 2005 16:03:51 +1000 (EST)
Received: from localhost (avspam-01.services.comms.unsw.edu.au [149.171.100.16])
    by smtp-dist.unsw.edu.au (8.13.1/8.13.1) with ESMTP id j8563oTp023828
    for <test_student@student.unsw.edu.au>; Mon, 5 Sep 2005 16:03:50 +1000 (EST)
Received: from smtp.unsw.edu.au ([127.0.0.1])
    by localhost (slag [127.0.0.1]) (amavisd-new, port 10025) with ESMTP
    id 13116-09 for <test_student@student.unsw.edu.au>;
    Mon, 5 Sep 2005 16:03:48 +1000 (EST)
Received: from hotmail.com (bay21-f13.bay21.hotmail.com [65.54.233.102])
    by smtp.unsw.edu.au (8.13.1/8.13.1) with ESMTP id j8563LVB026070
    for <test_student@student.unsw.edu.au>; Mon, 5 Sep 2005 16:03:27 +1000 (EST)
Received: from mail pickup service by hotmail.com with Microsoft SMTPSVC;
    Sun, 4 Sep 2005 23:03:20 -0700
Message-ID: BAY21-F13321938C106532E03894BAAA40@phx.gbl
Received: from 206.112.112.60 by by21fd.bay21.hotmail.msn.com with HTTP;
    Mon, 05 Sep 2005 06:03:19 GMT

......
(message body)
......
Alias in DNS record

- **Domain Name System (DNS)** record stores information about servers in a domain.
- SMTP relies on DNS to figure out mail server on destination domain.
- Each “**MX**” **field in DNS** corresponds to a Mail eXchange with its **priority**.
- MTAs or MTUs can select the best available servers base on priority and response time.

**Advantages:**
- No proxy server is needed
- Easy configuration

**Drawback:**
- Expose mail servers
Alias in DNS record

; zone file fragment
IN MX 10 mail.example.com.
....
mail IN A 192.168.0.4
IN A 192.168.0.5
IN A 192.168.0.6

; zone file fragment
IN MX 10 mail.example.com.
IN MX 10 mail1.example.com.
IN MX 10 mail2.example.com.
....
mail IN A 192.168.0.4
mail1 IN A 192.168.0.5
mail2 IN A 192.168.0.6
Effective Spam Filtering

- Use open-source developed solution to enhance spam filtering
  - Sender Policy Framework (SPF)
  - Sender ID
  - DomainKeys
- Reject unauthenticated mails
- Advantages:
  - Fast and effective strategies
  - Free up resources in mail servers to process valid messages
- Drawback:
  - Not widely supported by mail servers
Effective Spam Filtering

The sender of this message, philipchukamfai@hotmail.com, could not be verified by Sender ID. Learn more about Sender ID.

From: Chufai <philipchukamfai@hotmail.com>
Sent: Friday, October 7, 2005 6:46 PM
To: @hotmail.com
Subject: New Birthday Calendar

Content-Type: text/plain; charset=us-ascii
Content-Transfer-Encoding: 7bit
Delivery Status Notification (DSN)

- Defined in RFC 3461
- Extension to SMTP
- Request for an **acknowledgement** / **notification** of delivery status (success / failure / delay) from the last MTA that handled the mail

**Advantage:**
- Clear and effective enhancement

**Drawback:**
- **Limited support** from majority of MTAs
Delivery Status Notification (DSN)

----- The following recipients were processed by MTA twango.tiscalinet.it.

prova.prova@tiscalinet.it, Action Failed, Status: 5.2.2 (mailbox full)
Remote MTA mples2b.mx.tiscalinet.it: SMTP Diagnostic 552 RCPT <prova.prova@tiscalinet.it> ERROR. Exceeded storage allocation

---- Message SMTP Diagnostic: 554 DATA Transaction failed, no recipients given.
Message Disposition Notification (MDN)

- Defined in RFC 3798
- Known as "Acknowledgement" or "Return Receipts"
- Recipient MUA is requested to generate a receipt after user has read the message

**Advantages:**
- Easy to implement
- MTAs not involved

**Drawback:**
- Not transparent to users
- Privacy issue
Message Disposition Notification (MDN)
Pre-send Test Mail

- Send **small draft mail** to pre-setup mailbox in destination domain
- Test if draft mail can get through **spam and anti-virus filter**, also test for server availability
- Send the actual mails if successful

**Advantage:**
- No need for DSN and MDN

**Drawback:**
- Difficulties in mailbox setup
- **Limited Usage**
### Pre-send Test Mail

#### Delivery Assurance Dashboard

<table>
<thead>
<tr>
<th>Announcement (max 5)</th>
<th>More</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/16/2005 18:50</td>
<td>September 19th we will begin monitoring bulk for Cox, Excite and Wannado.co.UK, BLA history will be defaulted to show only active listings and Campaign Preview will have lots of new features!</td>
<td></td>
</tr>
</tbody>
</table>

#### Mailbox Monitor

<table>
<thead>
<tr>
<th>Campaign Performance (in order of date first seen)</th>
<th># of Campaigns</th>
<th>Inbox %</th>
<th>Bulk %</th>
<th>Total Received %</th>
<th>Missing %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Today</td>
<td>9</td>
<td>83.0</td>
<td>15.0</td>
<td>98</td>
<td>2.0</td>
</tr>
<tr>
<td>My Campaigns</td>
<td>9</td>
<td>85.7</td>
<td>14.0</td>
<td>99.7</td>
<td>0.4</td>
</tr>
<tr>
<td>All Clients</td>
<td>--</td>
<td>87.1</td>
<td>3.3</td>
<td>90.4</td>
<td>9.7</td>
</tr>
<tr>
<td>Yesterday</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My Campaigns</td>
<td>72</td>
<td>85.3</td>
<td>14.1</td>
<td>99.4</td>
<td>0.6</td>
</tr>
<tr>
<td>All Clients</td>
<td>--</td>
<td>88.0</td>
<td>3.1</td>
<td>91.1</td>
<td>9.0</td>
</tr>
<tr>
<td>Last 7 Days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My Campaigns</td>
<td>282</td>
<td>85.3</td>
<td>14.2</td>
<td>99.5</td>
<td>0.6</td>
</tr>
<tr>
<td>All Clients</td>
<td>--</td>
<td>89.0</td>
<td>2.7</td>
<td>91.7</td>
<td>8.3</td>
</tr>
</tbody>
</table>

#### Delivery Alert Activity (within last 7 days)

<table>
<thead>
<tr>
<th>Alert Time</th>
<th>0%</th>
<th>25</th>
<th>50</th>
<th>75</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

http://www.returnpath.biz/delivery/monitor/dashboard
Summary of Enhancement

- **Physical/Hardware**
  - Backup mail server
  - Proxy server
  - Redistribute services

- **Protocol/Software**
  - Alias in Mail eXchanger of Domain Name System (MX field in DNS)
  - Effective Spam filtering
  - Delivery Status Notification (DSN)
  - Message Disposition Notification (MDN)
  - Pre-send test mail
Summary of Enhancement

- None of them are perfect
- Responsibility of delivery is passed to MTAs
- Most of enhancement schemes need to be carried out in MTAs
- End-users are passive
- Demand of developing software tool for end-users
Can we do more?

- Develop next generation mail protocols
  - Hard to deploy
  - **Backward capability** (e.g. IPv4 to IPv6)

- Develop **end-users** mail tools
  - Operate on MUAs, independent of MTAs
  - Assist user to **detect, trace and retransmit lost mails**
  - Easy and fast to deploy
  - **Aim of this thesis!**
Scope of Thesis

- **Enhance reliability** of e-mails towards end-users:
  - Develop a system to **report e-mail delays and losses** in real time
  - **Detection** of lost mails by extensions (e.g. sequence numbers)
  - Debugging tool to **detect possible location of errors** (e.g. traceroute-like tool)
  - **Plug-ins** for popular mail clients
Further E-mail Enhancement

- Strategies on end-users enhancement:
  - **Sender Side** Mechanisms
  - **Recipient Side** Mechanisms
  - **Joint Client-Server** Mechanisms
Further Enhancement (cont.)

- **Sender Side Mechanisms**
  - Persistence e-mail sending
  - Automatic (silent) forwarding
  - Authenticating e-mail warnings
  - Replies tracking
  - SMTP traceroute tool

- **Recipient Side Mechanisms**
  - Abnormal lack of mails
  - Mails linking

- **Joint Client-Server Mechanisms**
  - Meta-data Restoring
Sender Side Mechanisms

- **Persistence e-mail sending**
  - Common MUAs pass responsibility of transmitting e-mails to configured mail servers
  - Give up when servers report an error
  - Assisting program can be implemented to **automatically retry** sending mails after the errors are identified
Sender Side Mechanisms

- **Automatic (silent) forwarding**
  - **Automatic forward** the e-mails to a pre-configured *secondary address* for particular recipient if primary one failed
  - Process is not able to be carried out if errors in primary one is not detected
Sender Side Mechanisms

- **Replies tracking**
  - Mails can be identified using *unique message-ID* field in their header
  - Further actions such as raising warning or resending should be performed by user if a message fail to receive reply after certain period of time
Sender Side Mechanisms

- **Authenticating e-mail warnings**
  - Filter out e-mail warnings that are not related to e-mails that the user has sent
  - Error checking can be done on particular e-mails
  - Common MUAs display all errors
Sender Side Mechanisms

- **SMTP Traceroute Tool**
  - To *trace* the possible travel path of a mail
  - Use when failure of delivery with unknown error
  - Using the concept from *traceroute tool in IP*
    - Utilize the **TTL (Time To Live)** fields in header
    - Generate **ICMP** error to source when TTL reaches zero
Sender Side Mechanisms

- Similar in mail header:
  - “Received-From” header fields
  - Appended by each MTA on the path
  - Error mail with original message is bounced back to sender once threshold is reached in mid-way

- Difficulties:
  - Different MTAs on the path have different threshold limit for the header fields
Sender Side Mechanisms

- Possible improvement
  - Use tools such as **SMTP Map** to assist detection of mail software on MTAs
  - Try **default** threshold values first
  - Sending **multiple trace** mails along the path to reduce trace delay
Sender Side Mechanisms

- **SMTP Traceroute sample**
  - % smtp_trace
  - % Usage:
    - smtp_trace [recipient mail] [local mail server]
    - % smtp_trace test@unsw.edu.au hotmail.com
    - Tracing test@hotmail.com…
    - 25ms mail4.hotmail.com
    - 600ms smtp.unsw.edu.au
    - 700ms smtp-dist.unsw.edu.au
    - 1080ms smtp3.unsw.edu.au
Recipient Side Mechanisms

- **Abnormal lack of messages**
  - Gather *statistic* information about flow of messages to client
  - **Abnormally** few mails received by client *may imply* local server fault
  - Alarm can be raised to notify user and mail administrator
Recipient Side Mechanisms

- **Mails Linking**
  - Inserting *unique sequence* number on each mail from sender
  - Receiver side can detect loss of mails from *missing sequence linkage*
  - Further actions can be performed
    - Retrieve backup copies from server
    - Request sender for retransmission
Joint Client-Server Mechanisms

- Cooperate to enhance reliability
- Modify mail server to maintain meta-information of mails
  - Header
  - Timestamp
  - Message-ID
- Client can access meta-data in server to detect if there are lost messages on local machine
- Further actions can be performed
  - Retrieve backup copies from server
  - Request sender for retransmission
Alternative Solutions

- **Direct mail delivery**
  - Send *directly* to recipient mailbox
  - No MTAs are involved
  - Errors are known instantly
  - Security issues
  - Spammer and virus

- **Instant Messenger**
  - ICQ, MSN Messenger, AOL
  - Fast and Easy to use
  - Limited message length
  - May *not support* asynchronous
Thesis A Work

- SMTP traceroute tool prototype
- Mail server setup
  - “Skunksmail” (skunksmail.ee.unsw.edu.au)
  - Room 343D Electrical Engineering Building
  - Fedora Core 4 (kernel 2.6.11)
  - Sendmail as SMTP software
  - UNSW SMTP server as gateway (smtp.unsw.edu.au)
- Not open relayed
- Fully tested
Skunksmail

UNSW Campus Network

skunksmail.ee.unsw.edu.au

smtp.unsw.edu.au
Thesis B Work

- **Continue** development of end-user tools
- **Testing** on different platforms
- **Plug-in** integration
Infrastructure

Common Mail Clients

Application Layer

System Service Layer

Network API

Daemon/Background Service
Project Schedule

- Implementation
  - SMTP traceroute tool (1 month)
  - Persistence mails sending (1 week)
  - Automatic (silent) forwarding (1 week)
  - Authenticating e-mail warnings (1 week)
  - Replies Tracking (1 week)
  - Detection of abnormal messages flow (1 week)
  - Mail Linking (1 week)
  - Meta-information server (2 weeks)

- Total implementation time ~ 3 months
Project Schedule

- Testing phase
  - Test on setup mail server (2 weeks)
  - Test on common clients (2 weeks)
  - Debugging (1 week)
  - Improvement (1 week)

- Plug-in integration
  - On common mail clients (3 weeks)
  - Testing (2 weeks)

- Total duration ~ 3 months
References

- RFC 2821 Simple Mail Transfer Protocol
- RFC 3461 SMTP Service Extension for Delivery Status Notifications
- RFC 974 Mail Routing and the Domain System
- RFC 1939 Post Office Protocol - version 3
- RFC 3798 An Extensible Message Format for Message Disposition Notifications
- RFC 3501 Internet Message Access Protocol - version 4rev1
- “Backup Mail Server Service”, Webservio Inc. (http://www.webservio.com)
- “An Overview of Internet E-mail”, GBdirect (http://ebusiness.gbdirect.co.uk/howtos/mail-system.html)
- Moxy Mail Proxy (http://moxy.psychogenic.com)
- LoadMail Forwarding (http://www.load.com/loadmail/)
- “Delivery Assurance Solutions”, Return Path Inc. (http://www.returnpath.biz/delivery/)
Questions?