Attachment C

Email Environment

February 8, 2005
# Table of Contents

- **Purpose** .......................................................................................................................... 1
- **Overview of Existing Products** .......................................................................................... 1
- **Mail Flow: @umich.edu** ................................................................................................... 2
- **Mail Flow: @somedomain.umich.edu** .................................................................................. 6
- **Anti-Virus and Anti-Spam Solutions** ................................................................................. 6
- **IT Central Services E-Mail** ................................................................................................. 7
  - **Mail Gateways** .................................................................................................................. 7
    - **Connection Setup and the U-M Block List** .................................................................. 7
    - **Sender Negotiation and the Sender’s E-mail Address** ............................................. 7
    - **Recipient Negotiation and Realtime Blackhole Lists (RBL’s)** ............................... 7
    - **Body Transfer and Anti-Virus** ..................................................................................... 7
  - **Front-end Servers** ............................................................................................................ 8
  - **Back-end Servers** ........................................................................................................... 8
  - **Sieve** ................................................................................................................................ 8
- **Departmental Servers** ........................................................................................................ 8
- **Summary for IT Central Services E-Mail** .......................................................................... 9
  - **E-mail Statistics** .............................................................................................................. 9
  - **Current Version Information** .......................................................................................... 10
- **Client Access** .................................................................................................................... 10
- **Outbound Mail** .................................................................................................................. 11
- **Departmental Mail Infrastructures** .................................................................................... 11
  - **A Sample Departmental Implementation** ....................................................................... 11
  - **UM Health System** ......................................................................................................... 13
  - **Campus Exchange Service** ............................................................................................. 13
  - **Department to Department Communications** ............................................................... 13
- **Future** .................................................................................................................................. 14
  - **Front-end Servers** .......................................................................................................... 14
  - **Authenticated SMTP** ...................................................................................................... 14
Purpose
This document describes, in general, the e-mail architecture implemented at the University of Michigan circa February 2005. The purpose of this document is to assist vendors responding to the University’s RFP for anti-virus, anti-spam, anti-phishing, and host hardening (e.g., anti-spyware and buffer overflow protection) solutions. In addition to helping vendors understand the current e-mail infrastructure, vendors can use the information to describe:

- how the vendor solution(s) would “fit” into the existing architecture
- which components of the existing architecture the vendor solution(s) would replace
- any architectural changes that would be recommended or needed in order to effectively support a proposed vendor solution

Overview of Existing Products
The University and its many departments support numerous e-mail infrastructure technologies including, but not limited to:

- **Message stores:**
  - Cyrus ([http://www.cmu.edu/computing/cyrus/](http://www.cmu.edu/computing/cyrus/))

- **Message Transfer Agents (MTAs):**
  - SendMail ([http://www.sendmail.org/](http://www.sendmail.org/))
  - SiMTA ([http://www.itd.umich.edu/umce/features/2004/simta.html](http://www.itd.umich.edu/umce/features/2004/simta.html))

- **Message stores + MTAs:**
  - Pegasus ([http://www.pmail.com/](http://www.pmail.com/))

Mail clients include, but are not limited to:

- Apple's Mail.app
- Entourage
- Eudora
- Groupwise
- IMP (for [http://mail.umich.edu](http://mail.umich.edu))
- Microsoft Outlook
- Microsoft Outlook Web Access
- Mozilla Thunderbird
- Mulberry
- Netscape Mail
- Pine
- Unix/Linux

Server and client platforms include, but are not limited to:

- Macintosh
- Windows

Server-side anti-spam products include, but are not limited to:

- BrightMail
- DSPAM
- MAPS
- Microsoft Exchange Intelligent Message Filter
- SpamAssassin

Anti-virus products include, but are not limited to:

- Antigen
- Norton Antivirus
- McAfee

Anti-Spyware products may include, but are not limited to:

- Ad-Aware
- Spybot Search & Destroy
- Microsoft Anti-Spyware Beta
Anti-Phishing products
- None

Host hardening and buffer overflow protection products may include, but are not limited to:
- ProPolice
- Qwik-Fix
- StackGuard

**Mail Flow: @umich.edu**

Mail sent from the Internet to an @umich.edu recipient enters the University e-mail infrastructure through one of ten mail gateways (MX machines) supported by IT Central Services (ITCS). These mail gateways (or mail relays) are responsible for forwarding inbound e-mail to the appropriate destination.
Forwarding information is maintained in the U-M Online Directory (UMOD). UMOD is implemented using OpenLDAP (http://www.openldap.org/). For example, if the UMOD directory specifies that mail for bjensen@umich.edu should be forwarded to bjensen@engin.umich.edu, then the gateway server that received mail addressed to bjensen@umich.edu would — after querying UMOD — forward the mail via SMTP to an SMTP server associated with the engin.umich.edu domain.

Forwarding addresses of the form @mail.umich.edu indicate that the mail should be forwarded to the back-end mailbox servers. Mail destined for the back-end mailbox servers is passed from a umich.edu gateway server to one of eight mail.umich.edu servers that front-end the back-end mail store. In order to determine the specific back-end server that is hosting a given user’s mailbox, the front-end servers consult a mapping database (or cache thereof) using the mailbox update (MUPDATE) protocol (RFC 3656). After determining which server hosts the recipient’s mailbox, the e-mail is delivered to the back-end mailbox server via the LMTP protocol (RFC 2033). The mail.umich.edu front-end servers, back-end stores, MUPDATE database, and LMTP delivery are implemented using Carnegie Mellon’s Cyrus IMAP e-mail system (http://asg.web.cmu.edu/cyrus/).
It is possible for a single user to have multiple forwarding addresses, and users are allowed to modify their forwarding addresses. If the forwarding address is outside of the umich.edu domain (e.g., a hotmail account), then the gateway servers abide, acting as a simple mail relay.

Of course, mail can also be sent to somegroup@umich.edu in which case the software running on the gateway servers relies on UMOD to determine all members of the recipient group and performs the forwarding logic discussed above for each member. Other processing also occurs during group expansion. For example, mail sent from a non-member to a “Members Only” group can be rejected. Similarly, duplicate recipients resulting from multiple group memberships can be appropriately condensed.
Mail Flow: @somedomain.umich.edu

Note that it is possible for inbound e-mail to bypass the mail gateway servers altogether. This can occur when the recipient address is qualified beyond the umich.edu domain (and the MX records for the destination domain do not redirect mail to the gateway servers). For example, mail sent from the Internet to bjensen@engin.umich.edu is handled by an MX server associated with the engin.umich.edu domain and is not processed by the mail gateway servers. Similarly, mail sent to Bo@mail.umich.edu is handled by the front-end servers without passing through the gateway servers:

Thus, there are multiple ingress points for e-mail to enter the University and — unless otherwise noted — there are no firewalls that restrict, redirect, proxy, or otherwise filter this inbound message flow.

Anti-Virus and Anti-Spam Solutions

Anti-virus and anti-spam solutions are implemented at various points in the architecture described above. The following sub-sections provide additional detail about the architectural components and anti-virus or anti-spam solutions implemented centrally as well as at the departmental level.
IT Central Services E-Mail

The following sub-sections provide additional detail about the anti-virus and anti-spam solutions implemented on the mail gateways, front-end servers, and back-end servers maintained by IT Central Services.

Mail Gateways

The mail gateway software is not sendmail. Instead, the message transfer agent running on the mail gateways was developed at U-M and is called SiMTA for Simple Internet Mail Transfer Agent (http://www.itcs.umich.edu/umce/features/2004/simta.html). In addition to enabling integration with UMOD (for mail forwarding), SiMTA also allows anti-spam and anti-virus solutions to be invoked during various phases of the SMTP protocol as discussed in the following sub-sections:

Connection Setup and the U-M Block List

During the SMTP connection setup phase, SiMTA can reject an e-mail based on the source IP address of the sender. To support this operation, the University currently maintains an internal list of problematic hosts. Gateway servers query the list and reject any inbound mail sourced from IP addresses on the U-M block List.

Sender Negotiation and the Sender's E-Mail address

If a message is not immediately rejected based on the source IP address, SiMTA examines the sender's e-mail address during the SMTP sender negotiation phase. If the domain of the sender cannot be resolved via DNS, the mail is rejected.

Recipient Negotiation and Realtime Blackhole Lists (RBL's)

An anti-spam solution from the Mail Abuse Prevention System (MAPS) called RBLPlus (http://www.mail-abuse.com) is implemented on the mail gateways. End users (or mail groups and distribution lists) can “opt in” for the service. This is based on several IP-address-based lists of suspected spammers. The mail of recipients who have “opted in” — as indicated by an entry in the UMOD directory — is subject to rejection based on the MAPS block list. Note that user-level “opt ins” do not affect group-level “opt ins.” Thus, if USERA belongs to GROUP1 and USERA has “opted-in” for the block list service but GROUP1 has not, then mail sent to USERA via GROUP1 will not be subject to the MAPS block list. The reverse is also true. If GROUP1 has “opted-in” but USERA (a member of GROUP1) has not, mail sent to USERA via GROUP1 will be subject to the MAPS block list.

For users who opt-in, during the SMTP recipient negotiation phase, SiMTA invokes a process that queries a locally cached copy of the MAPS block lists and rejects mail from such sources. Note that — unlike the U-M block list — the MAPS anti-spam process is invoked during the SMTP recipient negotiation phase because of the “opt in” nature of the service.

Body Transfer and Anti-Virus

Before e-mail is forwarded on to its destination, SiMTA runs a shell script that invokes the McAfee anti-virus software to scan the content of the e-mail for viruses. If McAfee detects a virus, the name of the virus is compared to an internal list maintained by the university to decide whether to reject or accept and discard the message. Viruses typically transmitted through e-mail are discarded, while all others are rejected.
Front-end Servers

DSPAM

Content-based anti-spam processing currently takes place on the front-end servers using a product called DSPAM (http://www.nuclearelephant.com/projects/dspam/). DSPAM uses statistical (Bayesian) filtering to adaptively learn what is and is not spam based on end user training. Authorized users “train” DSPAM by sending appropriate mail to spam.username@mail.umich.edu and notspam.username@mail.umich.edu. The samples are “tokenized” and stored in a MySQL database separate from the front-end server machines. The algorithms currently used by DSPAM to distinguish spam from non-spam are Graham-Bayesian, Burton-Bayesian and Graham’s p-value algorithms. Mail classified by DSPAM as spam is tagged so that it can be processed by Sieve rules on the back-end servers.

Back-end Servers

Sieve

Sieve rules run on the back-end servers in order to place e-mail tagged by DSPAM into the user’s “Junk Mail” folder.

Departmental Servers

Note that mail forwarded from the umich.edu gateway servers to departmental e-mail servers is not content-filtered for spam (since DSPAM is currently implemented on the front-end servers). Therefore, most departmental e-mail infrastructures — some of which are very large — implement their own anti-spam solutions including, but not limited to, Brightmail and the Microsoft Exchange Intelligent Message Filter. A sample departmental mail infrastructure that provides its own anti-virus and anti-spam solutions is described in detail later in this document.
Summary for IT Central Services E-Mail

The anti-virus and anti-spam solutions described above are:

![Image of anti-virus and anti-spam components](image)

**Figure 3: Anti-virus and anti-spam components**

### IT Central Services E-mail Statistics

Currently, the umich.edu gateway servers operated by IT Central Services accept approximately 1.5 million messages (and connections) per day with a maximum peak rate on any individual server of 350 messages per second. Of these 1.5 million daily messages, approximately 633,000 (42%) are abandoned due to anti-virus, anti-spam, and other filtering mechanisms. Of the 633,000 e-mails that are abandoned daily, approximately 33,000 contain viruses.

The 24 back-end mailbox servers store mail for 76,000 accounts in 650,000 mailboxes. This is because every folder (e.g., “inbox”, “junk”, “personal”, “shared”) that a user has constitutes a mailbox.

Numbers for all e-mail entering the University at various ingress points do not exist.
Current Version Information

- The Linux servers hosting the gateways, front-end servers, and back-end stores are running a 2.4 kernel along with version 2.3.3 of the GLIBC C library.
- DSPAM is version 3.2
- Open LDAP is version 2.1.26
- Cyrus is version 2.28

Client Access

Users who do not use a departmental mail infrastructure can access the IT Central Services mail infrastructure (mail.umich.edu) by configuring their mail clients to use the front-end mail servers which support POP3 and IMAP4 client access over SSL. These are the same front-end servers that move mail between the gateway servers and the back-end stores. Web-based access to mail.umich.edu e-mail is also provided via the Internet Message Program [IMP] (http://www.horde.org/imp/).

Figure 4: Client access methods for mail.umich.edu
Outbound Mail

Users who do not use a departmental mail infrastructure can send mail via the IT Central Services infrastructure by configuring their mail clients to use an smtp.itd.umich.edu Sendmail server. McAfee anti-virus software is used to prevent infected outbound mail from being delivered. All outbound virus-infected mail is rejected.

![Diagram of outbound mail via ITCS infrastructure](image-url)

Figure 5: Outbound mail via ITCS infrastructure

Outbound mail from the IMP servers is typically delivered directly to the target address. If the SMTP receiver for the target address is down, the outbound mail is off-loaded onto another “private relay” server where it is queued for future delivery.

Departmental Mail Infrastructures

Previous sections of this document have focused on the architecture, anti-virus, and anti-spam solutions implemented by the umich.edu mail gateways acting in concert with the mail.umich.edu technologies (Cyrus, SI3MTA, DSPAM, etc.). However, numerous departments within the University implement their own e-mail infrastructures. Indeed, some of these “departments” — such as UM-Flint, UM-Dearborn, and the UM Health System — are enterprises in their own right. This section provides a detailed example of one particular departmental e-mail system along with some high-level information on other departmental efforts.

A Sample Departmental Implementation

The following architecture, based on Microsoft Exchange, is real even though the name of the department (unit.umich.edu) is hypothetical:
Inbound e-mail may come directly from the Internet to a department or may be forwarded from the umich.edu mail gateways. Mail sent to UserX@unit.umich.edu bypasses the umich.edu mail gateways while mail sent to UserX@umich.edu is forwarded from the umich.edu gateways (assuming UserX has specified UserX@unit.umich.edu as a forwarding address in the UMOD directory).

This department has two Exchange SMTP gateway servers that process all incoming and outgoing e-mail. These gateway servers run Vamsoft Open Relay Filter (http://www.vamsoft.com/orf/) for inbound anti-spam and Sybari Antigen (http://www.sybari.ws/) for anti-virus.

For messages that are sent directly to @unit.umich.edu (bypassing the umich.edu gateways), Open Relay Filter (ORF) running on the gateway servers first performs the following tests based on the EHLO messages from the external relaying SMTP server:

- Reverse DNS lookup
- MX record check
- Sender address check

If the domain of the sender does not resolve, or does not have an MX record (and is not on an internally maintained “allow” list), the e-mail is rejected. Similarly, if the sender’s address is blank, the e-mail is rejected. All of this happens before the e-mail is actually delivered to the gateway.

If the e-mail is not rejected (or if mail is forwarded from the umich.edu gateways), ORF receives the message and performs further filtering using several DNS deny lists including but not limited to:

- Open Relay DataBase (ORDB)
- SpamCop
- CBL
• Distributed Blacklist (DSBL)
• SpamHaus
• Custom in-house deny lists

Suspected spam is “tagged.” Outlook clients can create filtering rules for tagged mail. The DNS lookup results are cached for 12 hours to reduce bandwidth consumption and improve the performance of spam filtering.

After anti-spam processing has taken place, accepted mail is analyzed by the Sybari Antigen anti-virus product (http://www.sybari.ws/). Antigen is configured to use six simultaneous scan engines:

• Norman Data Defense
• McAfee
• Sophos
• CA InnoculateIT
• CA VET
• Kaspersky

and includes a Worm list to contain outbreaks. Infected e-mail may be rejected and purged or quarantined. E-mail with potentially dangerous extensions such as (.exe, .bat, .vbs, etc.) may be cleaned by stripping the attachment and notifying the recipient.

Antigen also runs on the mailbox servers in order to filter outbound e-mail as close to the source of potential infection as possible. This would be effective in cases when, for example, an Outlook client has contracted a mass-mailing virus from a web site. Anti-spam software is not applied to outbound mail.

This particular unit also hosts a Blackberry Enterprise Server that is not shown. The Blackberry server encrypts and forwards Blackberry user’s e-mail from the store to the Blackberry network which, in turn, routes the e-mail through the user’s wireless provider to their Blackberry device.

**UM Health System**

One of the largest e-mail infrastructures on campus is maintained by Medical Center Information Technology (MCIT) for the University of Michigan Health System (UMHS). It is based on Novell’s Groupwise and currently supports approximately 22,000 users. Two McAfee e1000 appliances are used for anti-virus and anti-spam. Spam protection includes RBL subscriptions, and content filtering combined with Spam Killer, a Bayesian filtering technology, for scoring. Unlike many other departmental infrastructures, the MCIT Groupwise infrastructure is behind a firewall.

**Campus Exchange Service**

IT Central Services is currently in the process of providing a for-fee mail service based on a Microsoft Active Directory and Exchange infrastructure. Interested colleges include the College of Literature, Science, and the Arts (LS&A), the College of Engineering, and the Ross School of Business. This infrastructure currently leverages Brightmail for anti-spam and Antigen for anti-virus.

**Department to Department Communications**

As noted earlier, mail sent from the Internet may bypass the umich.edu mail gateways by further qualifying the domain name of a recipient. For example, mail sent to bjensen@umich.edu is
processed by the umich.edu gateway servers; mail sent to bjensen@engin.umich.edu is sent directly to the College of Engineering mail servers.

Mail sent from one campus departmental mail system to another departmental mail system on campus may or may not go through the umich.edu gateway servers.

**Futures**

**Front-end Servers**
With respect to the front-end servers that sit between the umich.edu mail gateways and the mail.umich.edu back-end store (Figure 3), we expect that the SMTP anti-spam portion of these front-end servers would move onto the gateway servers themselves. This would allow the anti-spam solution to apply to mail that is forwarded to departmental units. The front-end servers would still remain as client access points.

**Authenticated SMTP**
With respect to outbound mail (Figure 5), we anticipate that authenticated SMTP will be required for clients to access the SMTP servers (smtp.itd.umich.edu).