System Administration:
Week 7 Notes: Postfix, Procmail

March 20, 2006

1 Mail Protocols

1.1 POP3

POP3 (Post Office Protocol version 3) is a protocol used to retrieve mail from a remote server. Uses port 110, and port 995 when encrypted with SSL (Secure Sockets Layer), which authenticates the client via handshake when certificates are exchanged between client and server.

1.2 IMAP

IMAP (Internet Message Access Protocol) is used to retrive mail from a server, like POP3, however it does have several advantages of POP3. Unlike POP3 which connects to the server for a short period of time, IMAP allows you to stay connected to a server as long as you want and instantly retrieve messages. Unlike POP3 it also allows for more than one email client to be accessing the email box on a server simultaneously.

1.3 SMTP

SMTP (Simple Mail Transfer Protocol) is the protocol used to transmit mail over the Internet which uses port 25. The biggest problem with SMTP is the lack of an ability to authenticate the sender. Unlike POP3 and IMAP SMTP requires an established connection to receive mail. Note that even though you could use IMAP or POP3 to retrieve mail your mail client still uses SMTP to send mail.

2 MX Record

An MX Record is a record that tells which servers to send the e-mail to and the priority in which this should be done. You can find out the MX record vai **dig** commands, e.g. **dig ocf.berkeley.edu mx** or **dig berkeley.edu mx**. OCF's

main mailservers are war and sandstorm, with pestilence being the backup. Calmail on the other hand only has one mail server.

3 Spam

One of the biggest problems with email today is the abundance of spam. Below are several high-level approaches of identifying it.

3.1 SpamAssassin

SA is a Perl-based software used to successfully categorize spam. It scores each message based on the information found in the header and body with a score of 5 being default to classify a message as spam. It contains DNS-blacklists as well as training info online which it accesses.

3.2 Bayesian Spam Filtering

BSP is a process of classifying email into different categories based on the probability of a message being spam. Applications like **SpamAssassin** use it to clasify email. The basic formula is $P(spam|text) = \frac{P(text|spam)P(spam)}{P(words)}$. So in a BSP you train the filter first on $P(text|spam)$ by showing it previous email messages and then you are able to detect spam. The more you train your filter, the less spam gets through.

4 MTA

A Mail Transfer Agent is responsible for transfering email between servers. It uses the protocols discussed above.

4.1 Postfix

Postfix is an MTA which is gaining popularity over Sendmail.

4.2 Sendmail

Sendmail is the world's most popular MTA, written in Berkeley in early 1980s.

5 MDA

5.1 Procmail

Procmail is software used to deliver and classify messages (Mail Delivery Agent). It would not be a good idea to simply rely on postfix as the spam filtering. It is a better idea to use it to call SpamAssassin.

```
Basic syntax is

:0
* regex
folder_name

Trust Berkeley people not to spam you? Ehh...

:0
* ^From.*berkeley\.edu
$DEFAULT

Get special emails?

:0
* ^From.*bob@*berkeley\.edu
Chancellor

Don't you love to hate facebook?

:0
* ^From.*facebook
/dev/null


An example of a complete .procmailrc

:0
* !(To|Cc).*my_address
* !(To|Cc).*my_address_2
* !From.*my\.domain\.edu
* !From.*list.*@
spam
```

6 Email Clients

6.1 Pine

Pine (Program for Internet News and Email) was originally developed in Univ of Washington and is still used today.

6.2 Mutt

Mutt is probably the most powerful text-based client and is highly customizable. Its slogan is "mutt sucks less" due to author's belief that all mail clients are

flawed.

6.3 Mozilla Thunderbird

A nice open-source GUI-based email client with lots of cool options.