Manipulating Email with MIMEDefang

perl -pe ’s/xxx/xxx/magic’ <your email>

Sysadmin Miniconf, LCA 2006

Monday, 23 January 2006

Mark Suter, Miju Systems

<mark.suter@miju.com.au>

Copyright 2006 Mark Suter

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.1 or any later version published by the Free Software Foundation; with no Invariant Section, with no Front-Cover Texts, and with no Back-Cover Texts.

The "Spartan" Swiss Army Knife image is used with the kind permission of Victorinox.

Slides @ http://zwitterion.org/talks/

Ethics of Filtering Email

Hopefully, fairly standard stuff:

- Breaks the gentleman's agreement over email
- Have a written policy (often, just one page)
- Let it be known what you're doing and why
- www.sage-au.org.au/ethics.html

IANAL, YSSARL

Filtering email

Filtering is an instrument of policy. Our policy might require

- dealing with viri
- dealing with UBE/UCE
- adding disclaimers
- removing "inappropriate content"
- whatever else the organisation needs

Email Overview - where to filter?

IMAP Internet Mail Access Protocol
LDA Local Delivery Agent
MTA Mail Transport Agent
MUA Mail User Agent
POP Post Office Protocol
SMTP Simple Mail Transfer Protocol

Filtering in the MTA

- Often done by adding an extra MTA
- Once per message, not once per recipient (compare to LDA like procmail)
- Can filter even when just relaying (no need to deliver)
- Very powerful SMTP transaction manipulation

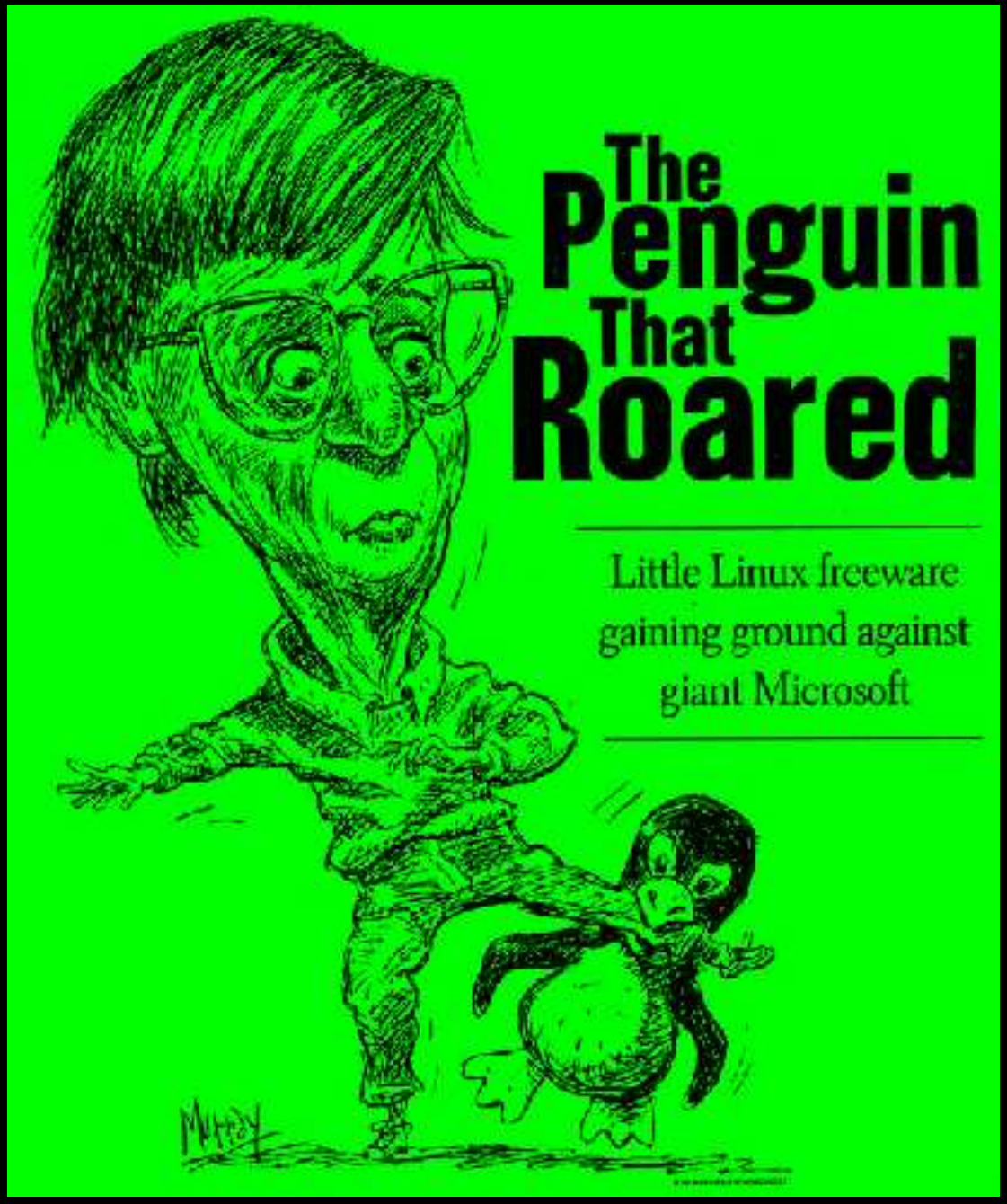
Milter (Mail fiLTER)

- Protocol+library for filtering in the sendmail
- Part of Sendmail since 8.10, works well in 8.13.5
- Allows filters to affect SMTP:
  - accept
  - reject
  - discard
  - copy
  - alter
  - ...
- More information at www.milter.org

Roaring Penguin's MIMEDefang

- Based on Perl - the Swiss Army Knife ;)
- Uses Milter to interface with Sendmail
- Widely used
- Can handle very large loads (millions of emails/day)
- GNU/Linux, FreeBSD, Solaris, OS X, Tru64, HP-UX, AIX, etc

"Roaring Penguin" - What's in a name?



Licensing

It’s all Free Software:

- Sendmail
  - BSD License
    - ftp.sendmail.org/pub/sendmail/LICENSE
  - But, dual licensed by Sendmail, Inc (a bit like MySQL AB)
    - www.sendmail.org/license-info.html
- Perl
  - Perl Artistic License
    - www.perl.com/language/misc/Artistic.html
- MIMEDefang
  - GNU General Public License
    - www.roaringpenguin.com/penguin/open_source_mimedefang.php

MIMEDefang Architecture

- sendmails talk to threaded mimedefang (socket)
- mimedefang talks to multiplexor (socket)
- multiplexor manges pool of slaves (pipes)
- mimedefang.pl "uses" mimedefang-filter (we edit)

MIMEDefang Advantages (1 of 2)

- Easy: We only need to write Perl. CPAN!
- Robust: slaves are independent
- Efficient: Slaves are pre-forked
- Scalable: Known cost per slave (can be ulimited)

MIMEDefang Advantages (2 of 2)

- Almost all the work is already done:
  - Anti-Virus hooks (e.g. ClamAV, Sophos)
  - SpamAssassin hooks
  - MIME processing is handled
- A fairly complete interface to Milter

Simple Mail Transport Play

S: 220 server.example.com ESMTP
C: EHLO client.example.com
S: 250 OK ..... filter_relay
C: MAIL FROM:<sales@example.com>
S: 250 OK ..... filter_sender
C: RCPT TO:<support@example.com>
S: 250 OK ..... filter_recipient
C: DATA
S: 354 Start Input
C: Help! I can’t send emails!
C: .
S: 250 OK ..... filter_begin, filter, filter_end
C: QUIT
S: 221 Closing connection

filter_relay

- Optional (we've got sendmail's acls)

```
sub filter_relay ($$$) {
    my ($ip, $name, $helo) = @_;
    if (rand < 0.75) {
        return (0, "Sorry, the magic eight ball doesn't like you!");
    }
    return (1, "ok");
}
```

filter_sender

- Again, optional

```
sub filter_sender ($$$$) {
    my ($sender, $ip, $hostname, $helo) = @_;
    if ($sender =~ /^<?spammer@badguy.com>?$/i) {
        return (0, 'Sorry; spammer@badguy.com is blacklisted.');
    }
    return (1, "ok");
}
```

filter_recipient

- Still optional

```
sub filter_recipient ($$$$$$) {
    my ($recipient, $sender, $ip, $hostname, $first, $helo) = @_;
    if ($sender =~ /^<?spammer@badguy.com>?$/i) {
        if ($recipient =~ /^<?postmaster@mydomain.com>?$/i) {
            return (1, "ok");
        }
        return (0, 'Sorry; spammer@badguy.com is blacklisted.');
    }
    return (1, "ok");
}
```

filter_begin

- Called once the email is ready to be processed

```
sub filter_begin () {
    # ALWAYS drop messages with suspicious chars in headers
    if ($SuspiciousCharsInHeaders) {
        action_quarantine_entire_message();
        action_notify_administrator("Suspicious characters");
        # Do NOT allow message to reach recipient(s)
        return action_discard();
    }
    # Scan for viruses if any virus-scanners are installed
    my($code, $category, $action) = message_contains_virus();
    $FoundVirus = ($category eq "virus");
}
```

filter

- Called for each MIME part

```
sub filter ($$$$) {
    my ($entity, $fname, $ext, $type) = @_;
    return if message_rejected();
    if ($FoundVirus) {
        my $ScannerMessages = "";
        my ($code, $category, $action) = entity_contains_virus($entity);
        if ($category eq "virus") {
            return action_quarantine($entity, "$ScannerMessages\n");
        }
    }
    return action_accept();
}
```

filter_end

- Called after all parts have been processed

```
sub filter_end ($) {
    my($entity) = @_;
    return if message_rejected();
    if (-s "INPUTMSG" < 100*1024) {
        my ($hits, $req, $names, $report) = spam_assassin_check();
        if ($hits >= $req) {
            # X-Spam-Score: 6.8 (******)
            my $score = "*" x int $hits;
            action_change_header("X-Spam-Score", "$hits ($score)");
        }
    }
}
```

Possibilities are endless

- Sendmail macros are available (remember the bat book?)
- Replace attachments with URLs to save bandwidth
- Archive emails *your* way
- Anything that can be coded:
  - Always true; however Perl+CPAN make this easier.

Sendmail

- One modification to sendmail.mc:

```
INPUT_MAIL_FILTER(
    ‘mimedefang’,
    ‘S=unix:/var/spool/MIMEDefang/mimedefang.sock,
        F=T,
        T=C:5m;S:2m;R:2m;E:5m’
)
```

- Needs libmilter:
  - libmilter-dev package on Debian GNU/Linux
  - Simple compile time option for sendmail

More Information (1 of 2)

In our Internet Age, the skill is finding the *useful* information. Of course, there are the homepages for the key components of the system I've discussed.

- www.sendmail.org
- www.perl.org
- www.mimedefang.org

More Information (2 of 2)

MIMEDefang has a very good mailing list. The main developer, David F. Skoll, is active on this list.

- lists.roaringpenguin.com/mailman/listinfo/mimedefang
- lists.roaringpenguin.com/pipermail/mimedefang/

Any more questions?

"Ask the next question"
Theodore Sturgeon (1918-1985)