An e-mail quarantine with open source software

Using amavis, qpsmtpd and MariaDB for e-mail filtering

Daniel Knittel
Dirk Jahnke-Zumbusch

HEPiX fall 2016
NERSC, Lawrence Berkeley National Laboratory
United States of America
October 2016
e-mail services at DESY

> DESY is hosting 70+ e-mail domains, most prominent:
  - desy.de — of course :)
  - xfel.eu — European XFEL
  - belle2.org — since summer 2016
  - cfel.de — Center for Free-Electron Laser Science
  - cssb-hamburg.de — Center for Structural Systems Biology

> mixed environment of open source software and commercial products
  - Zimbra network edition with web access and standard clients (Outlook, IMAP, SMTP)
  - Postfix for MTAs
  - SYMPA for mailing list services
  - Sophos and Clearswift’s MIMEsweeper for SMTP

> currently ~6.500 fully-fledged mailboxes, some 1000s extra with reduced functionality (e.g. no Outlook/ActiveSync/EWS access)

> daily ~300.000 delivered e-mails
DESY e-mail infrastructure

1 DMZ filtering
- restrictive filtering, reject e-mails from very suspicious MTAs
- 1b soon: DESY’s NREN (DFN) will be integrated into e-mail flow with virus- and SPAM-scanning

2 filter for bad content ➔ suspicious e-mails into quarantine

3 2nd-level SPAM-scan based on mail text and own rules

4 distribution of e-mails to mailbox servers, mailing list servers or DESY-external destinations

5 throttling of e-mail flow to acceptable rates (individual vs. newsletter)
  - think “phishing” ➔ high rates trigger an alarm

mixed HW/VM environment
**e-mail at DESY – attachment filtering & quarantine**

> policy: e-mail traffic is filtered

- block “bad” e-mails in the first place
  - viruses are blocked
  - executable content is blocked

- also block e-mails originating from DESY if they contain malicious or suspicious content

- up to now: commercial solution

> additional measures

- mark e-mails with a high SPAM-score (2nd-level SPAM-filtering)
- monitor outgoing e-mail-flow
- throttle if over a specific rate
  - this is sender-specific and customizable (e.g. individuals vs. newsletters)
- think “Friday phishing peak”
filtering on our own – how it started

> commercial solution in place
  - >10 years
  - working mostly w/o problems ➔ detect and quarantine if applicable
  - MS-Windows based product
  - some functional deficiencies
    - RAR archives as well as Excel files often wrongly classified (false positive)
    - no e-mail header insertions (e.g. „Auto-Submitted: auto-generated“)

  - end-of-life foreseeable ➔ transition to another product necessary

  - some 10k€ annual maintenance
  - virus scanner is separate technically and in terms of licensing

> while we were musing about a possible successor…

> … at the end of 2015 increasing number of MS-Office macro viruses ➔ „Locky“

> our anti-virus scanner cannot scan documents for VBA macros

> standard policy for MS-Word documents is not to allow macros, but…
  - … it is hard for users to resist not to enable them
  - … the first malicious document occurred with an invoice stating to be from one of our business partners

> conclusion
  - contact our product vendor ➔ no luck
  - fast reaction needed
  - start to code a solution on our own

> result
  - a Postfix content filter
  - which uses the „oletools“
  - adds some header information to the e-mail
  - let the commercial solution do the quarantine work
transistion to (mostly) open source software

- idea: shift to open source software
drop-in replacement

- use well-proven software

- extensible by our needs, if necessary

- components we did need
  - e-mail transport
  - e-mail decomposition
  - classification of attachments
  - scanning for viruses or other unwanted content
  - quarantining still wanted
  - manual unblocking by postmaster and other workflow

- and perhaps some additional functionality?

- would ClamAV’s detection be good enough?
filtering building blocks

attachment scanner

postfix₁ ➔ amavis

Sophos + ClamAV

> postfix queues ➔ decoupling systems
> amavis
  - decomposes e-mails
  - classifies attachments
> Sophos (commercial AV scanner) and ClamAV scanning for viruses
  - Sophos runs as a service
  - latency between first occurrence of a virus and availability of AV-signatures

quarantine

postfix₂ ➔ dquarantine

MariaDB

> postfix queues ➔ decoupling systems
> self-made: dquarantine
  - stores e-mails incl. their attachments
  - releases e-mails by postmaster actions
  - feeds ClamAV with DESY-created signatures for unwanted attachments
  - notifies by sending e-mails
  - e-mail template mechanism for different communication partners
inner parts of the quarantine

qpsmtpd
- provides stable core SMTP features
- functionality is extended by plug-ins
- is implemented in PERL ➔ fine for string-handling
- used by apache.org, perl.org, cpan.org

MariaDB holds metadata

CLI for now

standard exception handling
- temporary failure leaves messages in queues ➔ scan later
conclusion

> mostly open source software solution
  > commercial virus scan engine still needed
> balance between rejecting / accepting mails
  > reject mail most obvious to be „bad“ → virus
  > still accept mails with attachments which may pose a problem
    → for user acceptance
    → accept & scan later
> dubious e-mails quarantined
  > released on user request
  > manually by one postmaster / double-check
  > some „easy-release“ mechanism possible
> extensible solution
  > on our own, but presumably more quickly than any vendor support
  > own ClamAV signatures
> covering legal demands
thank you for your attention
useful links

<table>
<thead>
<tr>
<th>Software</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amavis</td>
<td><a href="https://amavis.org/">https://amavis.org/</a></td>
</tr>
<tr>
<td>ClamAV</td>
<td><a href="https://www.clamav.net/">https://www.clamav.net/</a></td>
</tr>
<tr>
<td>oletools</td>
<td><a href="http://www.decalage.info/python/oletools">http://www.decalage.info/python/oletools</a></td>
</tr>
<tr>
<td>Postfix</td>
<td><a href="http://www.postfix.org/">http://www.postfix.org/</a></td>
</tr>
</tbody>
</table>