amavisd-new
advanced configuration and management

Mark Martinec

Institut “Jožef Stefan”

http://www.ijs.si/software/amavisd/

Agenda

- what it is
- performance / benchmark / tuning
- policy banks
- lookups, SQL, banning rules
- regular maintenance and monitoring
- tips & tricks

amavisd-new - what it is?

- interfaces between MTA and virus checkers and/or SpamAssassin
- decodes/unpacks mail and checks parts
- quarantines malware
- logging/reporting: to SQL (new with 2.3)

why is it popular?

- reliable:
    - checks status of every operation, internal asserts
    - in case of failure mail stays with MTA
- adheres to standards (SMTP, MIME, DSN, ...)
- reasonably fast, reasonably feature-rich
- can run chroot-ed
- GPL license
- 800+ downloads of 2.3.0 in the first two days after a release

AMaViS history

shell program:

- 1997 Mogens Kjaer, Juergen Quade
- 1998-01-17 AMaViS 0.1 (Christian Bricart) - 300 lines
  "AMaViS - A Mail Virus Scanner"
- 1998-12 AMaViS 0.2.0-pre1
- 1999-07 AMaViS 0.2.0-pre6 (Rainer Link, Chris Mason)
- 2000-10 AMaViS 0.2.1 (Christian Bricart)

Perl program:

- 2000-01 Amavis-perl (Chris Mason)
- 2000-08 Amavis-perl-8
- 2000-12 Amavis-perl-10
- 2001-04 Amavis-perl-11 (split> amavisd)
- 2003-03 Amavis-0.3.12 (Lars Hecking)

AMaViS history

Perl daemon:

- 2001-01 daemonisation (Geoff Winkless)
- 2001-04 amavisd-snapshot-20010407 (Lars Hecking)
- 2001-07 amavisd-snapshot-20010714
- >2002-04 amavisd-snapshot-20020300 (split> amavisd-new)
- 2003-03 amavisd-0.1 2100 lines

Perl, modular re-design

- 2002-03 amavis-ng-0.1 (Hilko Bengen)
- 2003-03 amavis-ng-0.1.6.2 (Hilko Bengen)

amavisd-new 3+ years of development (7 years of tradition)

Perl daemon, pre-forked, Net::Server

- 2002-03 amavisd-new-20020329 (Mark Martinec)
- 2002-04 amavisd-new-20020418
- 2002-04 amavisd-new-20020424
- 2002-05 amavisd-new-20020517
- 2002-06 amavisd-new-20020630
- 2002-11 amavisd-new-20021116
- 2002-12 amavisd-new-20021227
- 2003-03 amavisd-new-20030314
- 2003-06 amavisd-new-20030616
- 2003-11 amavisd-new-20030616-p6 10.000 lines
- 2004-06 amavisd-new-20030616-p10
- 2004-07 2.0
- 2004-08 2.1.0
- 2004-08 2.1.1
- 2004-09 2.1.2
- 2004-11 2.2.0
- 2004-12 2.2.1
- 2005-04 2.3.0
- 2005-05 2.3.1 15.000 lines

performance:
benchmarking platform

- dual AMD Opteron 246, 2 GHz
- 2 GB memory
- ATA-100 and SCSI-3 disk
- FreeBSD 5.4, 64-bit

- Perl 5.8.6
- SA 3.0.3, SA 3.1(cvs)
- amavisd-new-2.3.1

performance: benchmarking setup

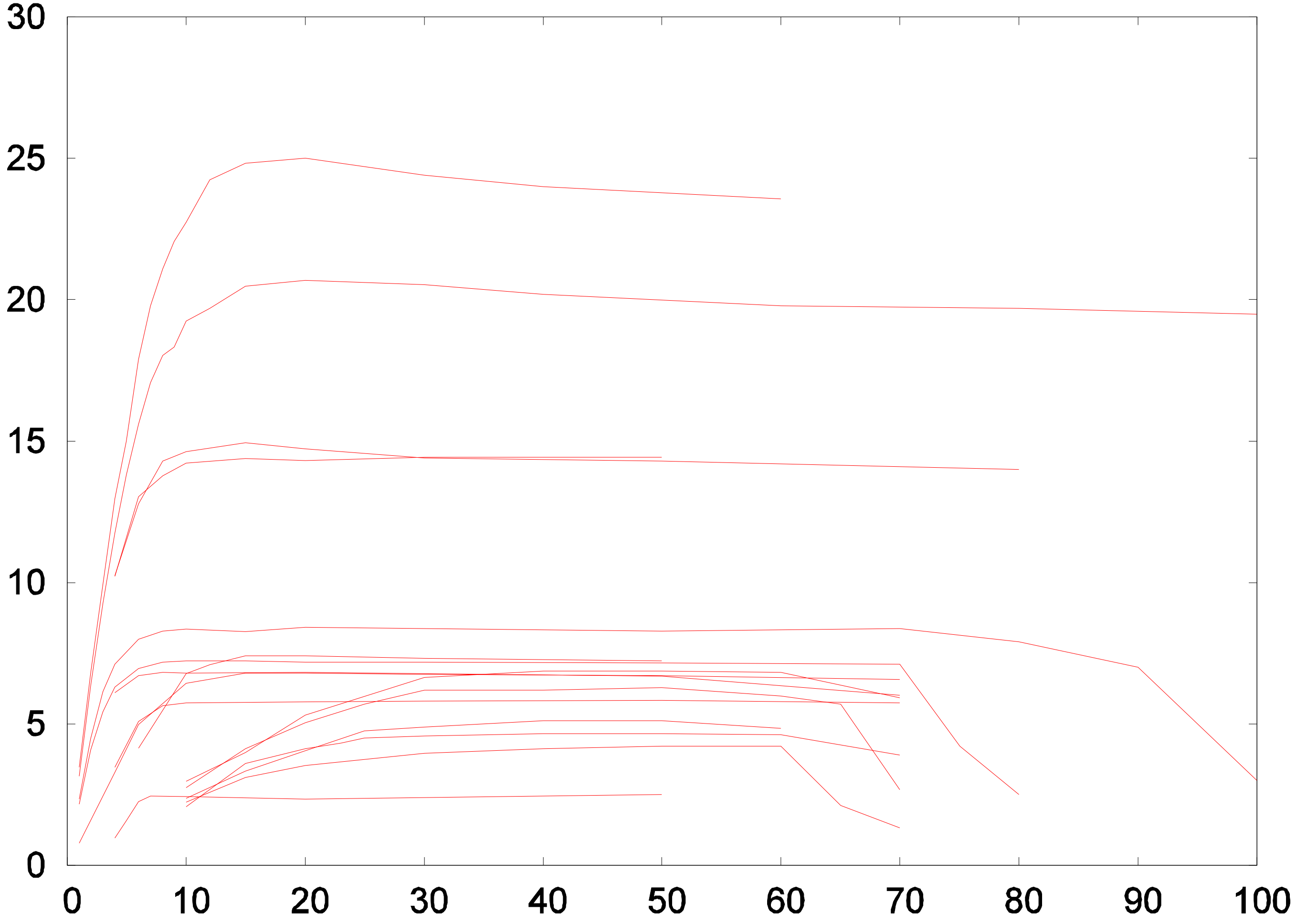
- SMTP source and sink on a separate host
- *smtp-sink* instrumented with clock, shows transactions/s
- smtp source:

  dedicated Postfix with spool on md for real mail,

  smtp-source for raw Postfix baseline measurements
- 1500 mail messages: real mail, random 24h sample, cca 20% quarantined, all delivered (*_lovers, tagged)
- MySQL server on the same host

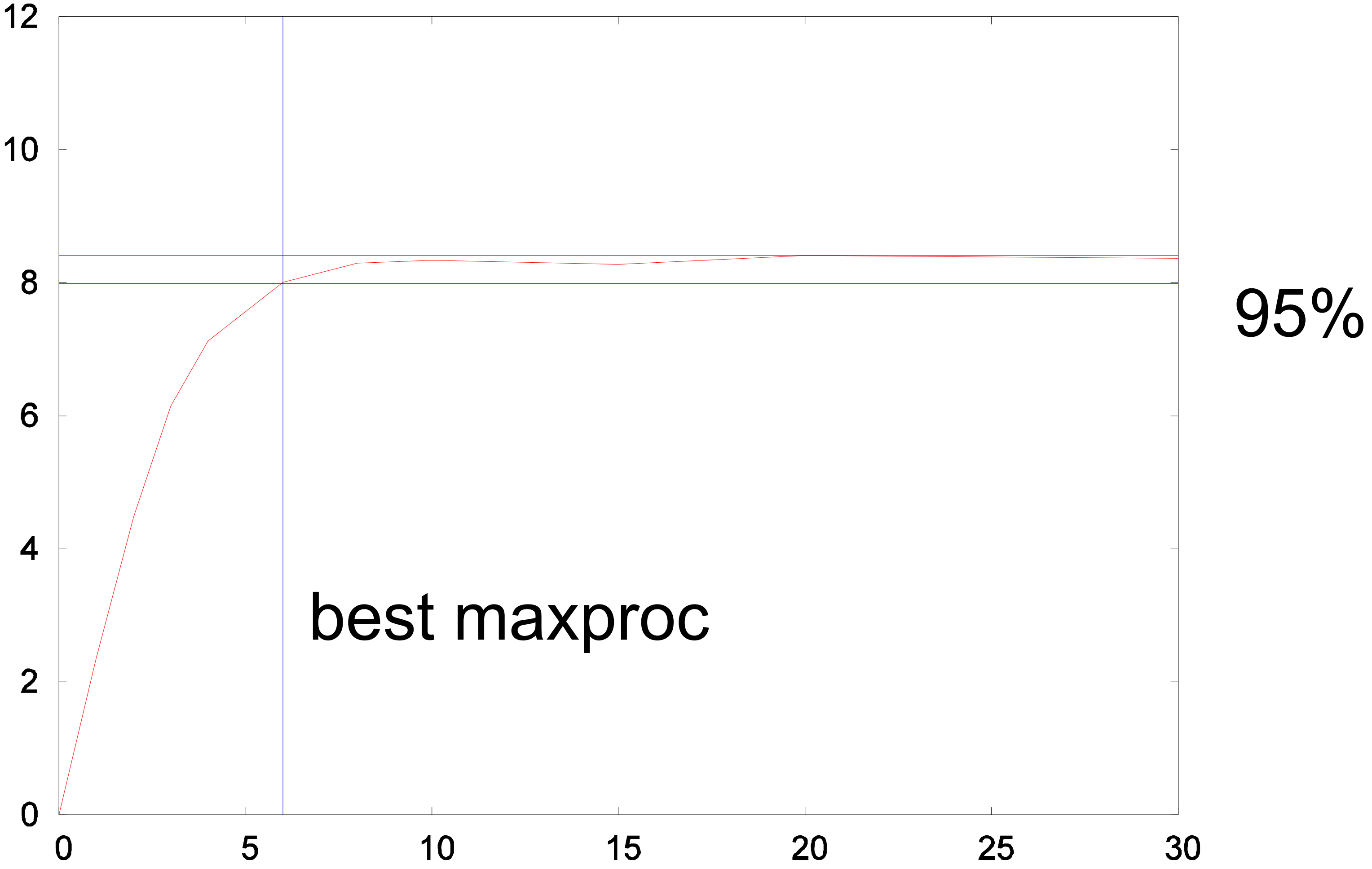
performance – general idea

msgs/s *vs.* maxproc



performance – general idea

msgs/s *vs.* maxproc

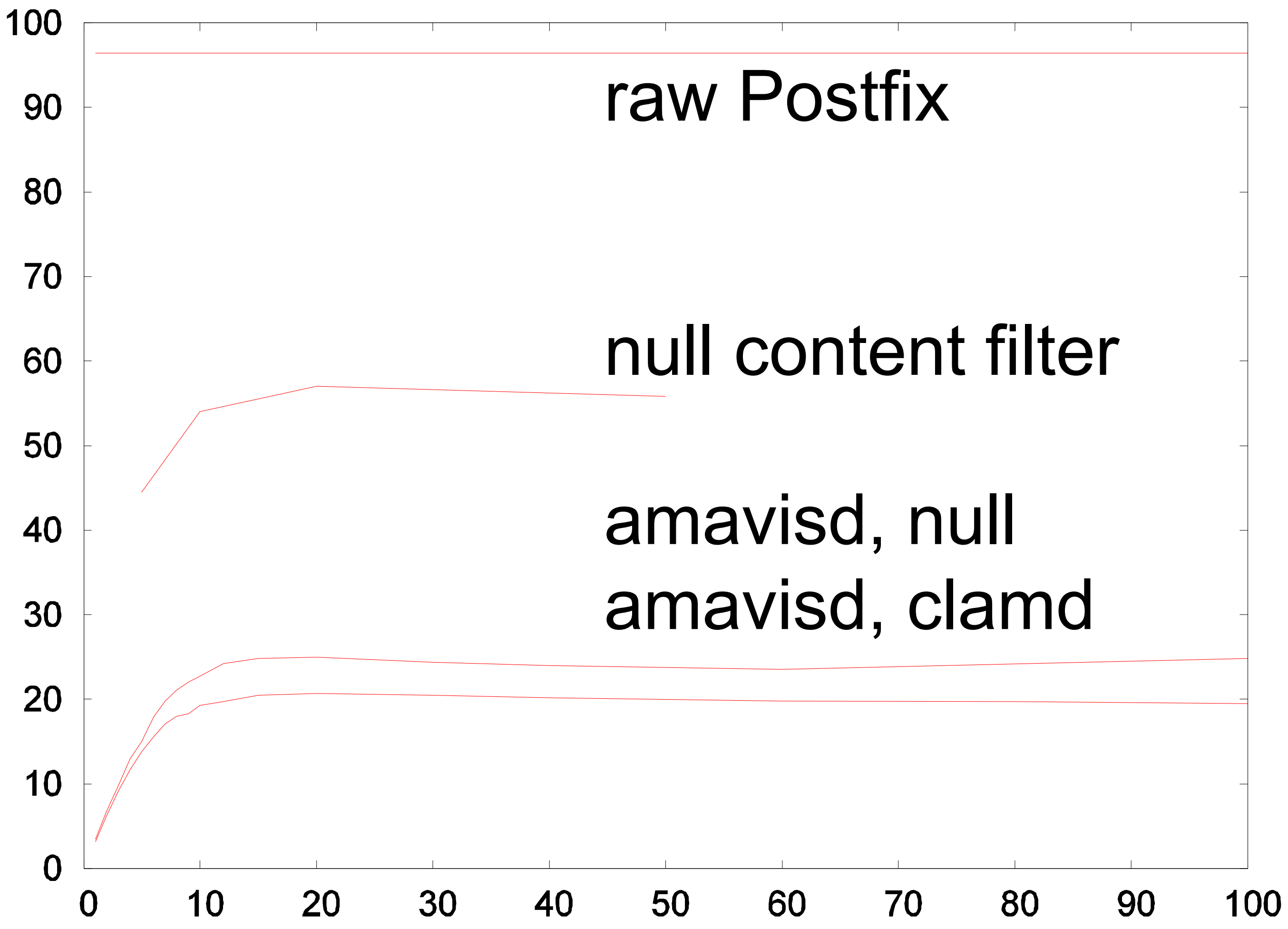


performance: Postfix baseline

- 96.5 SMTP transactions per second (subject to disk speed)
- just enabling Postfix content filtering (null filter) drops mail throughput to 60%
- every mail hits the disk twice

performance: baseline

msgs/s *vs.* maxproc



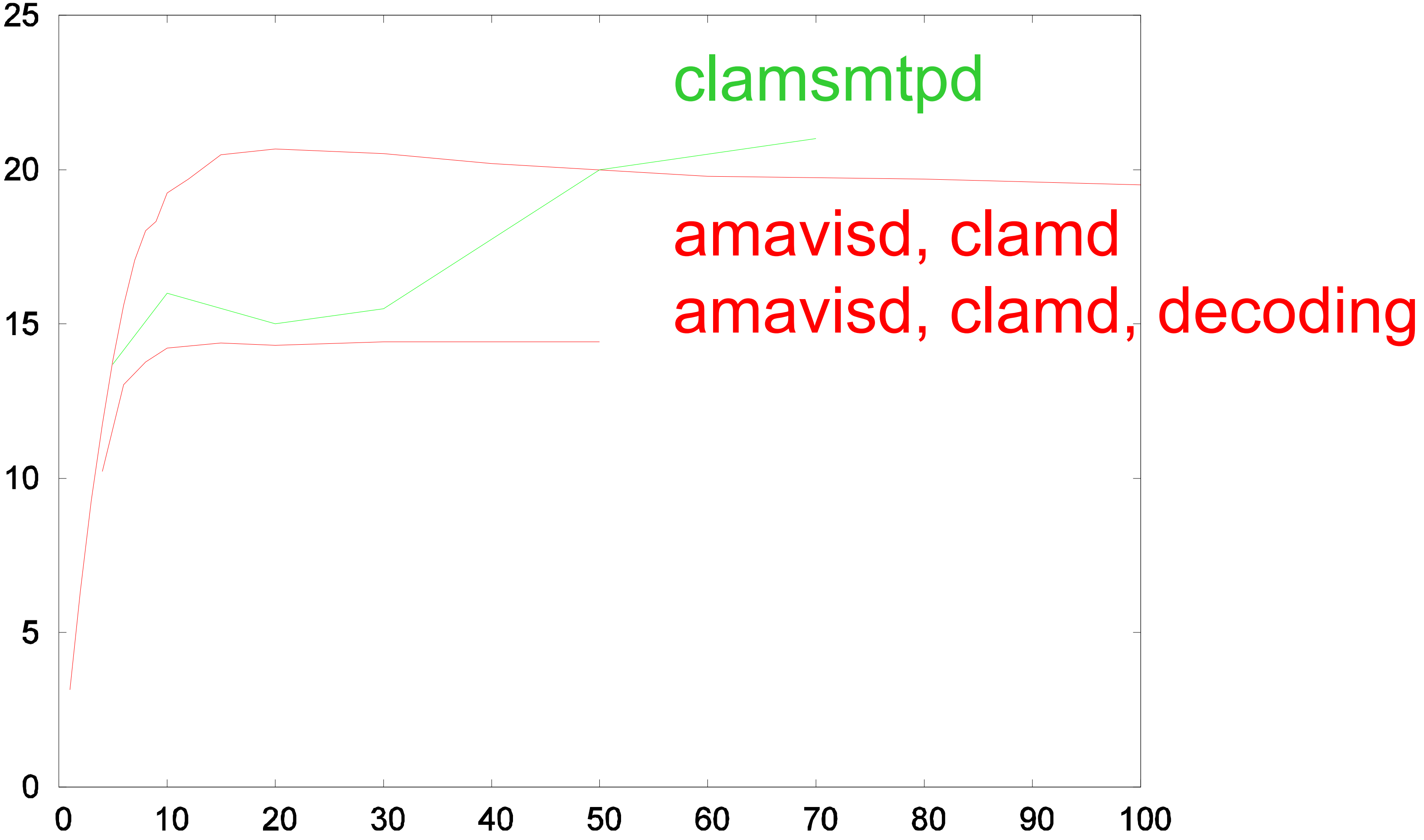
performance:
amavisd baseline

- inserting amavisd with all checks and external decoding *disabled* drops throughput to 1/4 of raw Postfix throughput (additional 1/2)
- one additional data transfer, MIME decoding
- optimum maxproc is 12 processes

(at 95 % max throughput)

performance: plain virus checking

msgs/s *vs*. maxproc

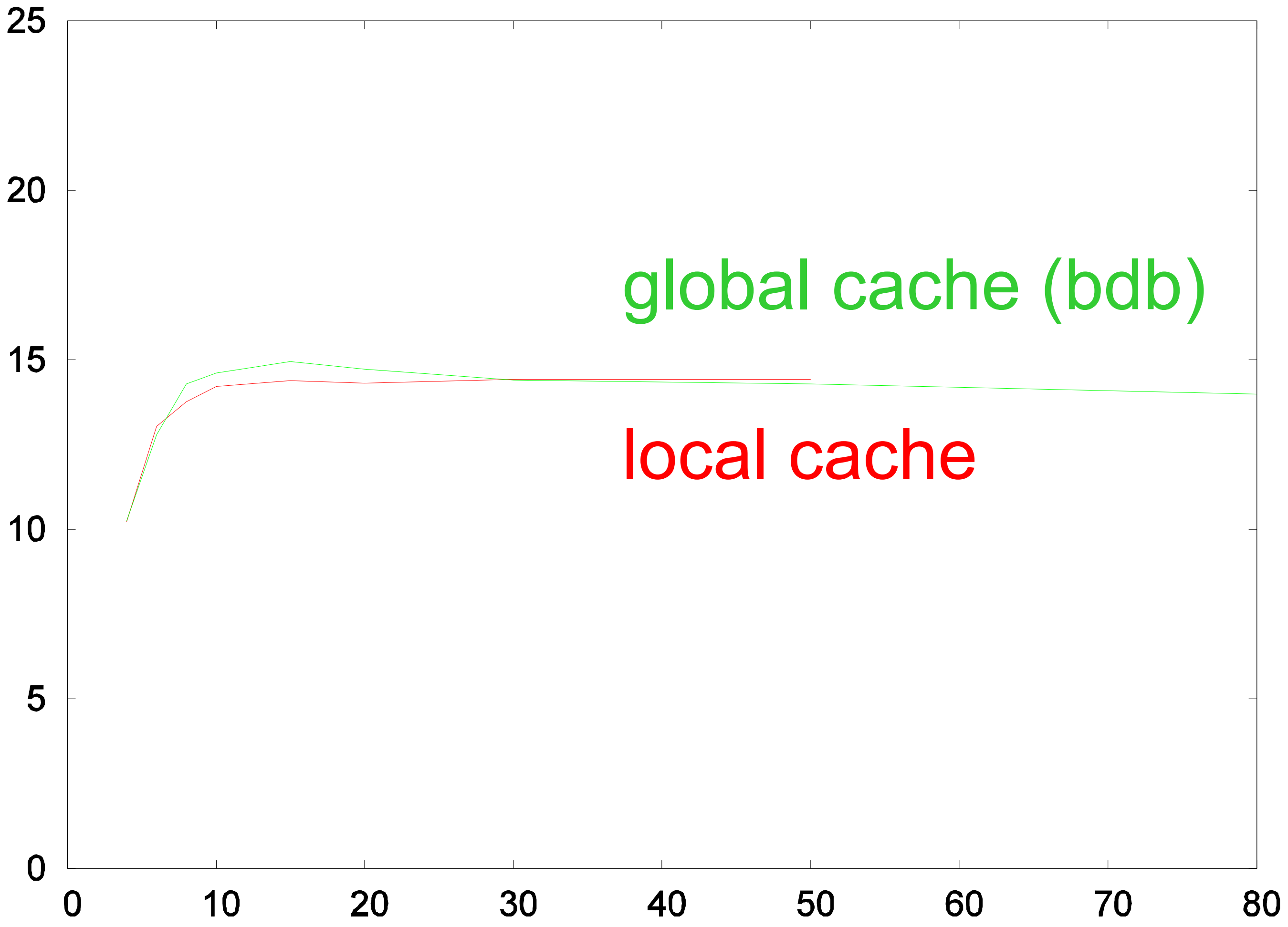


performance:
enabling bdb with a global cache

- global cache (more than) pays for itself and for bdb usage
- benefit: statistics counters, nanny (sanity database)
- no apparent database lock contention

performance: global cache

msgs/s *vs.* maxproc



performance:
does RAM-disk help?

- writing fs metadata can be expensive
- amavisd reuses *work* and *parts* directory, and a temporary file *email.txt*
- MIME::Parser reuses its temporary file

- depends on fs, fs options and disk caching

performance:
does RAM-disk help?

Recent Slashdot article: http://www.livejournal.com/~brad/2116715.html

> *Most RAID cards* ***lie*** *(especially LSI ones), some OSes lie (rare), and most disks lie (doesn't matter how expensive or cheap they are). They lie because their competitors do and they figure it's more important to look competitive because the magazines only print speed numbers, not reliability stats.*

FreeBSD Handbook: 11.12 Tuning Disks

- Soft Updates guarantees filesystem consistency in the case of a crash but could very easily be several seconds behind updating the physical disk (*very appropriate for amavisd work area*, *instant recovery*)

man page ATA(4):

- *hw.ata.wc* sysctl variable set to 1 to enable Write Caching, 0 to disable (default is enabled). Can cause data loss on power failures.

performance:
does RAM-disk help?

- measuring setup: md 0.5 GB (out of 2GB) on FreeBSD: /tmp, /var/tmp, /var/amavis/tmp, clam-tmp

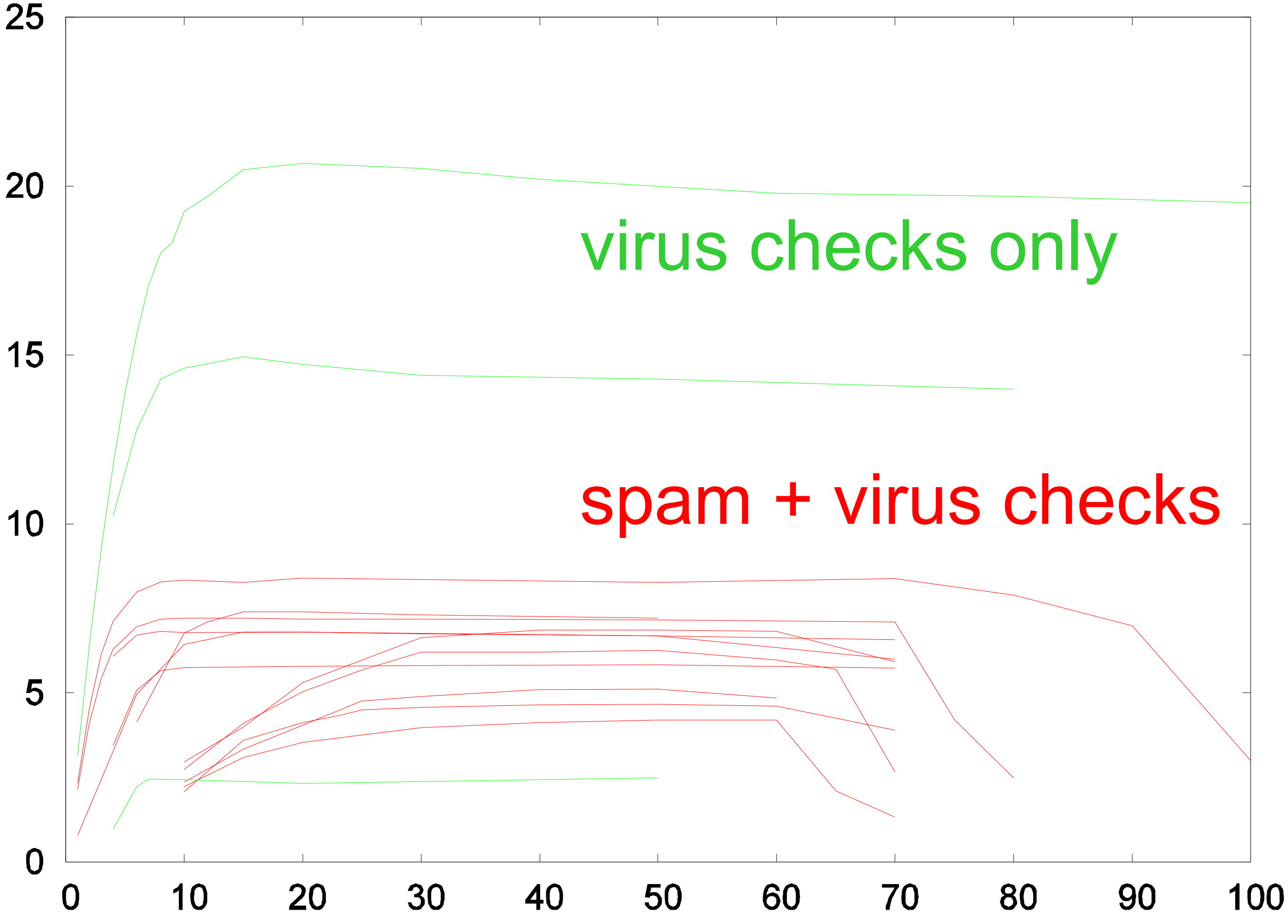
- no change in throughput (below 1%)
compared to UFS2 file system on a ATA disk
with write-cache enabled

- below 5% change on SCSI disk with soft updates
and no SA checks, practically no change with SA enabled

- likely benefit with some other file systems

performance: SpamAssassin

msgs/s *vs*. maxproc

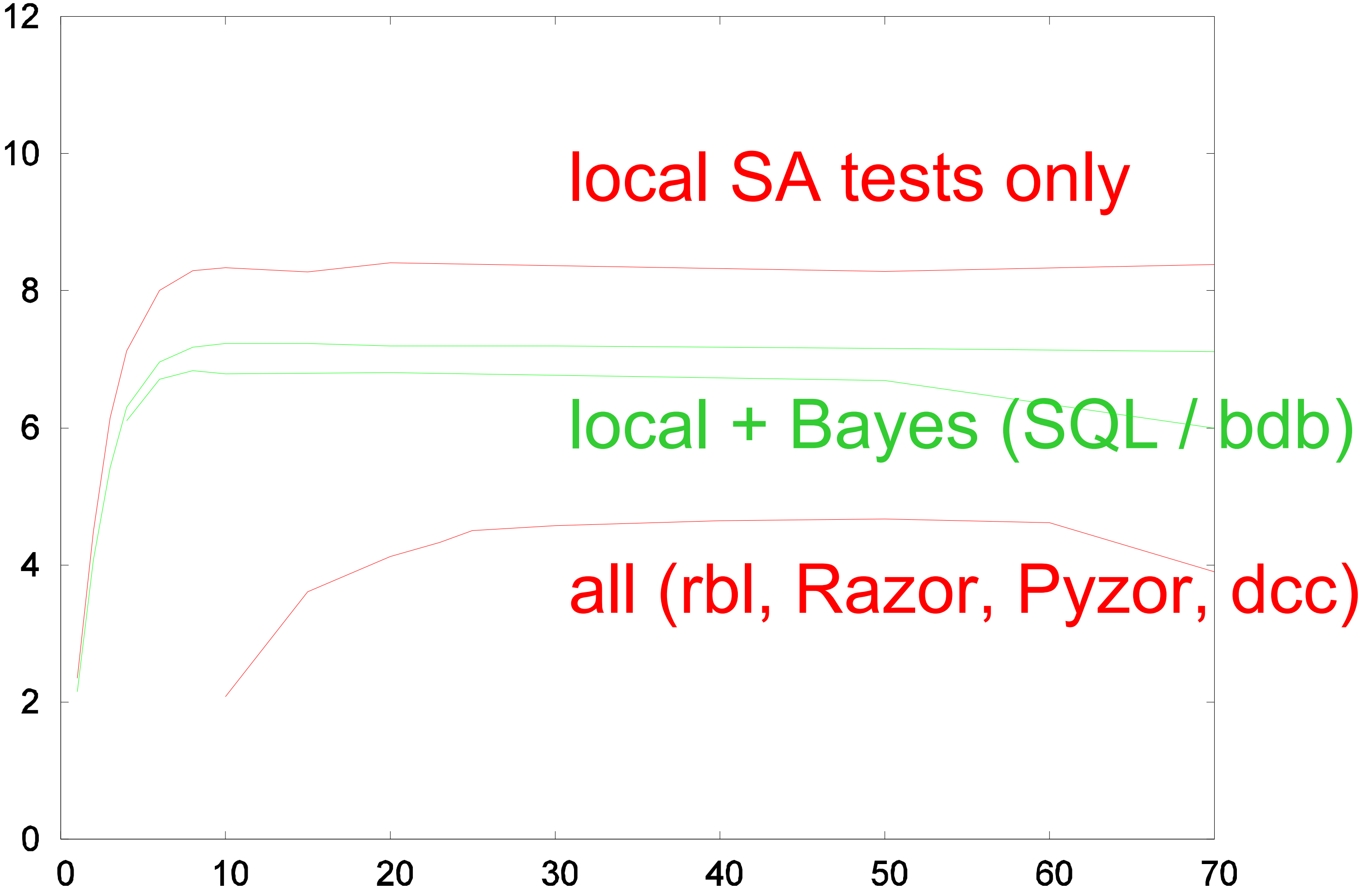


performance: Bayes & AWL db

- costs 14%
- database choice
  - BerkeleyDB
  - MySQL InnoDB 6% higher throughput than bdb
  - MySQL MyISAM faster(?), may need REPAIR TABLE

performance: SpamAssassin

msgs/s *vs.* maxproc



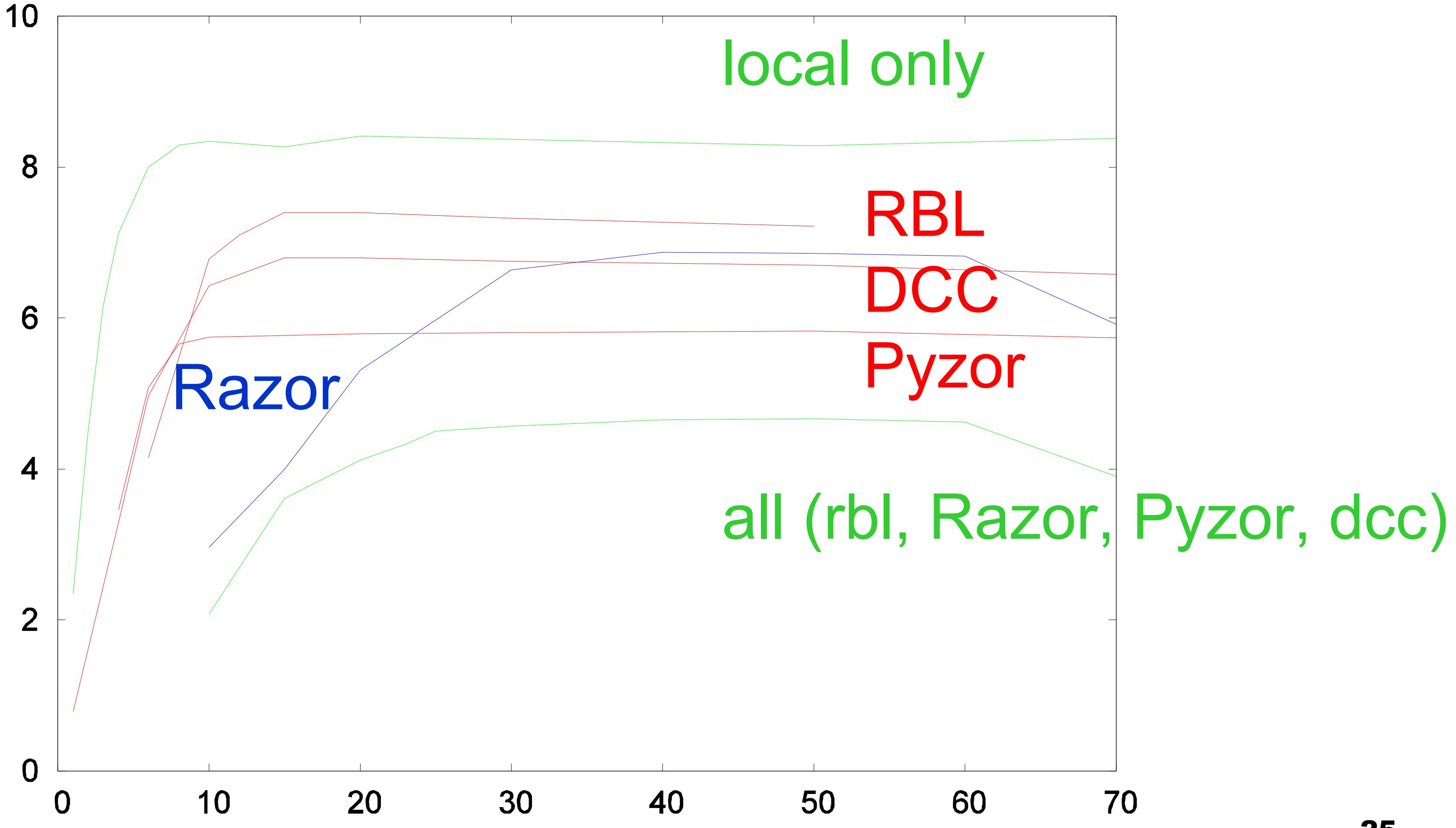
performance: Razor2, Pyzor, DCC

- - Razor2 has high latency and requires more parallel processes
- - Pyzor is more resource-hungry
- - DCC is low latency like Pyzor and uses low resources like Razor2

- Pyzor maxproc95 = 8
- DCC maxproc95 = 11
- Razor2 maxproc95 = 30

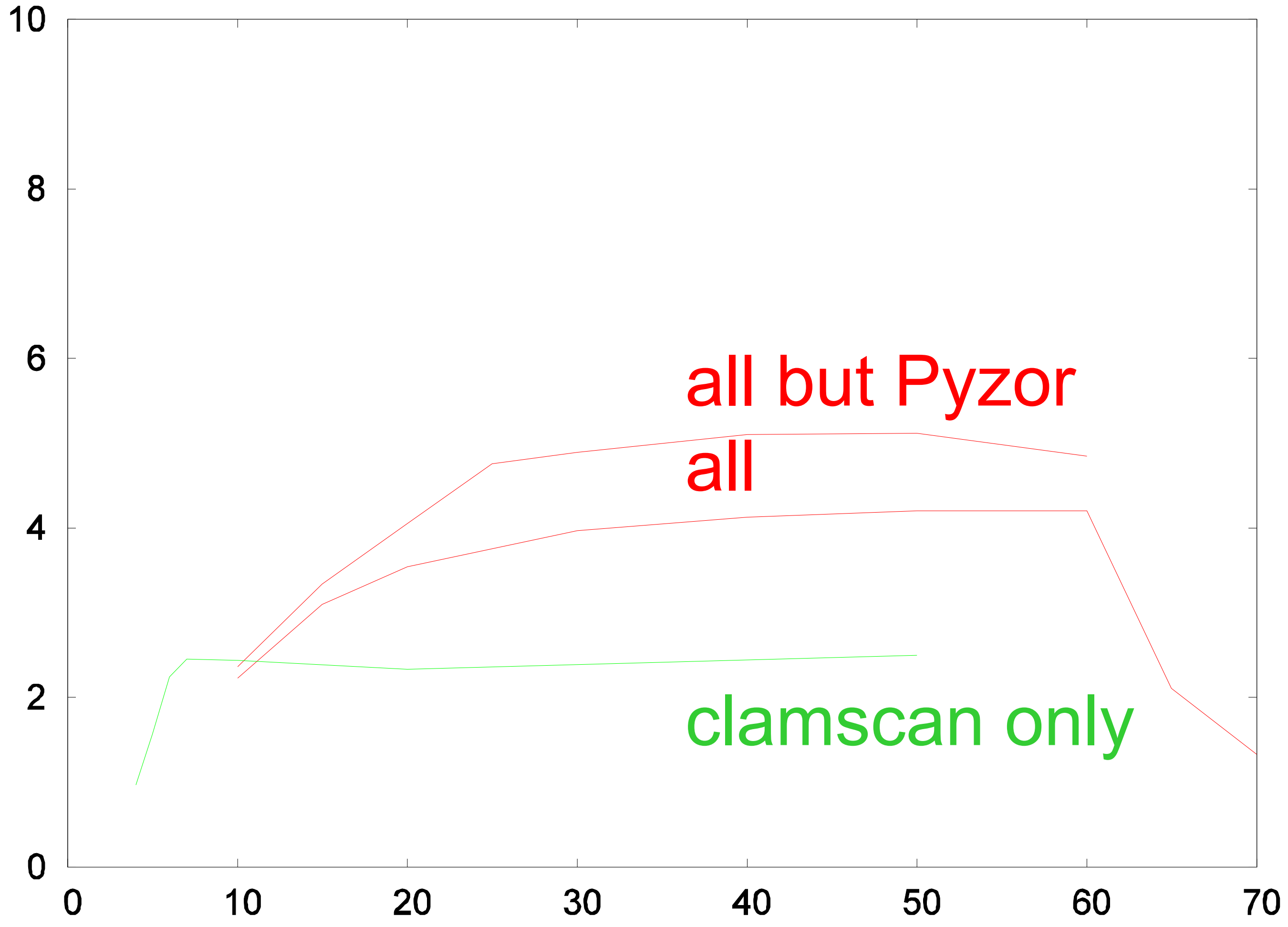
performance: SpamAssassin

msgs/s *vs*. maxproc



performance: can it get worse?

msgs/s *vs.* maxproc



performance: experience

Julian Rendon reports:

> *I'm using amavisd-new in 2 Sparc servers processing*
> *each one more than 1.3 Millions mails a day,*
> *serving as a mailhub gateway without local users.*
>
> *I found a 39 max_server concurrency to be good for my hardware.*
> *We currently process more than 32 GB of mail a day.*
> *Servers are 2 SF280R, each one with 2 1.2 GHz processors,*
> *2 MB RAM and fiber-channel disks.*

tuning

optimal number of processes at > 95% of max throughput

- 12 - no decoding, no checking
- 12 - no decoding, clamd

- 8 - decoding, clamd
- 6 - decoding, clamd, SA (local)
- 6 - decoding, clamd, SA (local, bayes)
- 12 - decoding, clamd, SA (rbl; no bayes)
- 23 - decoding, clamd, SA (rbl,razor,pyzor,dccproc; no bayes)
- 30 - decoding, clamd, SA (razor only)

- 32 - everything, 4 msgs per second, 1 GB of memory would suffice
- (20 - everything, 3.55 messages per second)

Need 5 msgs/s instead of 4? Drop Pyzor.

tuning

- memory
  - memory: RSS/VSZ = 60% is memory-resident
  - cca 60% of a process' RSS is shared
  - cca 30 MB real memory for a 100 MB virtual memory process
- 1 GB: 25 processes - just manages to reach optimum with all checks enabled
- 2 GB: 60 processes - plenty of headroom

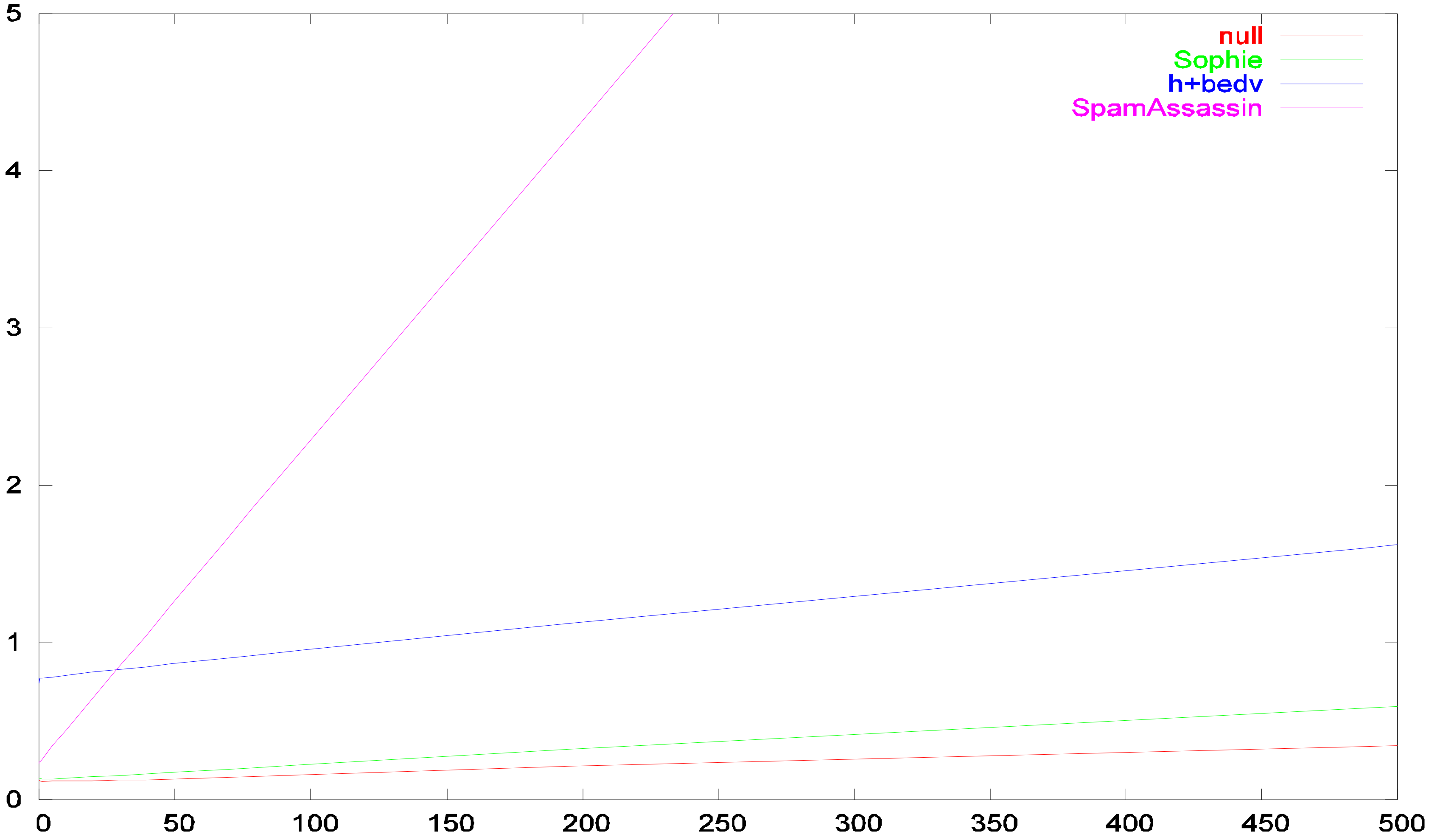
tuning: general

- separate disks for MTA spool and amavisd-new work area
- avoid slow command-line virus scanners
- Linux syslogd: disable sync for MTA and amavisd logs
- some RulesEmporium (SARE) rulesets expensive
- turn on *$quarantine_subdir_levels* (2.3.0)
- separate MTA and amavisd hosts
- split load through multiple MX records

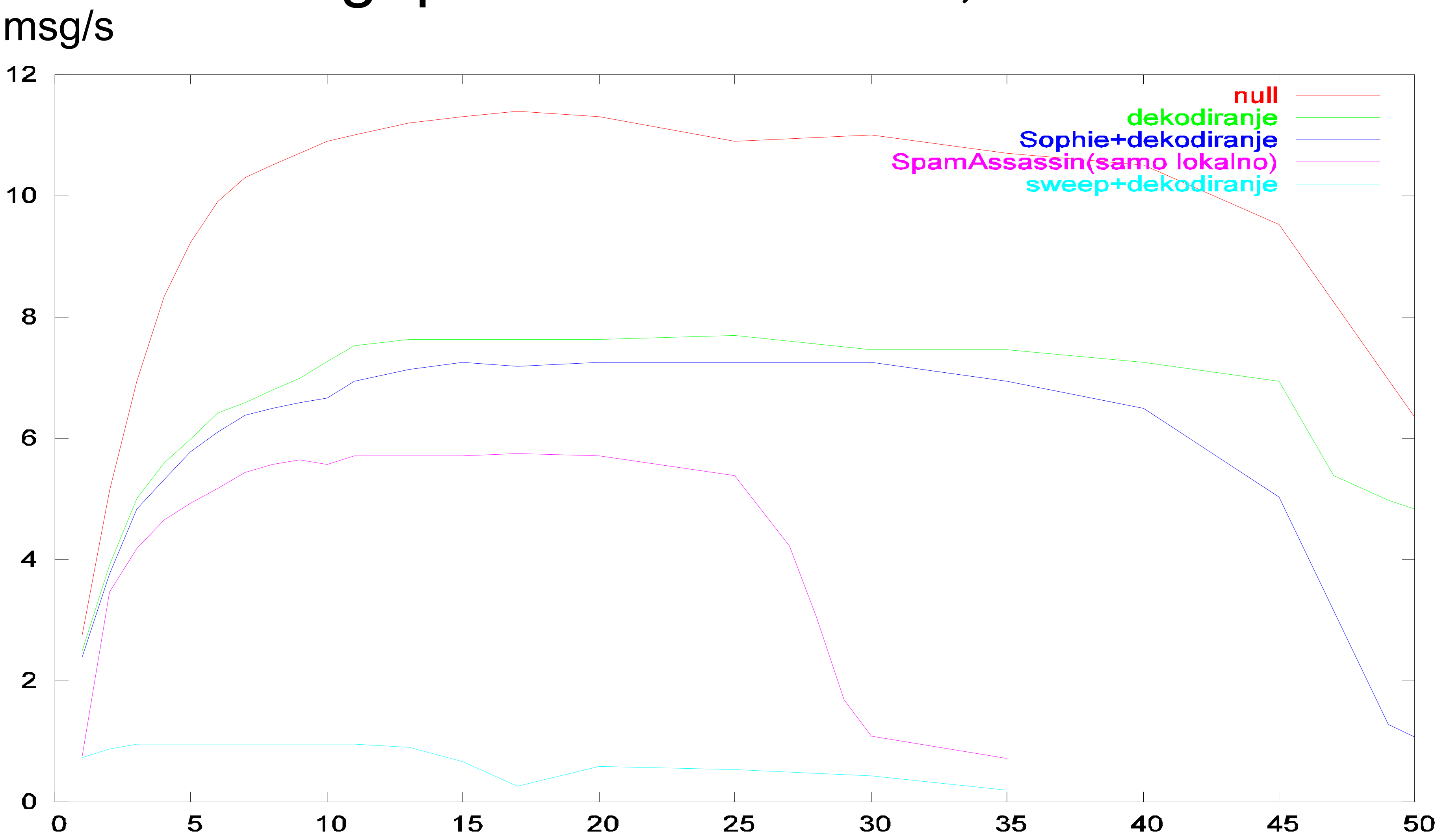
tuning: timing report

```
TIMING [total 1725 ms] -
  lookup_sql: 6 (0%)0,
  SMTP pre-DATA-flush: 1 (0%)0, SMTP DATA: 88 (5%)6,
  body_hash: 1 (0%)6, sql-enter: 4 (0%)6,
  mime_decode: 6 (0%)6, get-file-type1: 23 (1%)7,
  parts_decode: 0 (0%)8,
  AV-scan-1: 7 (0%)8, AV-scan-2: 4 (0%)8, AV-scan-3: 5 (0%)8,
  AV-scan-4: 1 (0%)9, AV-scan-5: 1 (0%)9, AV-scan-6: 0 (0%)9,
  lookup_sql: 4 (0%)9, spam-wb-list: 3 (0%)9,
  SA msg read: 0 (0%)9, SA parse: 2 (0%)9,
  SA check: 1536 (89%)98,
  update_cache: 2 (0%)98, post-do_spam: 6 (0%)99,
  deal_with_mail_size: 0 (0%)99,
  main_log_entry: 18 (1%)100,
  sql-update: 4 (0%)100, update_snmp: 1 (0%)100,
  unlink-1-files: 1 (0%)100, rundown: 0 (0%)100
```

processing time [s] / size [kB]



throughput: 2GHz celeron, 256 MB



number of processes

making MTA and amavisd-new talk to each other - input

- SMTP or LMTP or AM.PDP on input

- $inet_socket_port = 10024;
- $inet_socket_port = [10024,10026,10027];

- @inet_acl = qw(127.0.0.0/8 [::1] 192.168.1.1); *# access control*
- $inet_socket_bind = '127.0.0.1'; *# restrict to one interface*

- $unix_socketname = '/var/amavis/amavisd.sock';
  # e.g. quarantine release

making MTA and amavisd-new talk to each other - output

- SMTP or pipe on output

- $forward_method = 'smtp:[127.0.0.1]:10025';
- $notify_method   = 'smtp:[127.0.0.1]:10025';

- $forward_method = 'smtp:*:*';
- $notify_method   = 'smtp:*:10587';

  1st asterisk SMTP client peer address

  2nd asterisk incoming SMTP/LMTP session port number plus one

- $virus_quarantine_method, $spam_quarantine_method, ...

making MTA and amavisd-new talk to each other

- one MTA, one amavisd (same or separte hosts)
- multiple MTAs sharing one amavisd
- external MTA > amavisd > internal MTA
- receiving MTA > amavisd > transmitting MTA

- TCP port-based policy bank override of *$forward_method* and *$notify_method*
- allows each input channel its own forwarding route

policy banks

- sets of configuration parameters that apply
  to processing a mail message as a whole (not per-recipient)
- fast switchover from one set to another
- selected by:
  TCP port number, SMTP client IP address, sender domain
- similar to Postfix FILTER option (applies to the whole message)
- TCP port number or socket - to policy name mapping:

```
$interface_policy{'10026'} = 'INTERNAL';
$interface_policy{'10028'} = 'NET4';
$interface_policy{'10030'} = 'OFFICE';
$interface_policy{'3330'}   = 'PfTCP';
$interface_policy{'9998'}   = 'AM.PDP';
$interface_policy{'SOCK'} = 'AM.PDP';
```

- two policy names are hard-wired:
  MYNETS: client IP address matches *@mynetworks* (XFORWARD)
  MYUSERS: sender matches *@local_domains_maps*

policy banks

```
$policy_bank{'NET4'} = {
  log_level => 3,
  smtpd_greeting_banner =>
      '${helo-name} ${protocol} ${product} NET4 service ready',
  notify_spam_sender_templ => read_text
     ("$MYHOME/notify_spam_sender.txt"),
};
$policy_bank{'MYNETS'} = {  # mail originating from @mynetworks
  virus_admin_maps    => ["virusalert\@$mydomain"],
  spam_admin_maps   => ["virusalert\@$mydomain"],
  spam_dsn_cutoff_level_maps => [ 15 ],
  final_spam_destiny => D_DISCARD,
  banned_filename_maps => [new_RE(
    [qr'\.[^./]*\.(exe|vbs|pif|scr|bat|cmd|com|cpl|dll)\.?$'i => 1],
    [qr'^\.(Z|gz|bz2|rpm|cpio|tar|zip|rar|arc|arj|zoo)$'      => 0],
    [qr'^\.(exe-ms)$'  => 1]) ],
};
$policy_bank{'AM.PDP'} = {
  protocol => 'AM.PDP',
  inet_acl => ['127.0.0.0/8',  '192.0.2.11'],
};
```

policy banks

poor-man's SPF (sender policy framework)

```
@score_sender_maps = ('.' => [                 # site-wide table
  { 'nobody@cert.org'       => -3.0,           # (soft)whitelist
    'owner-alert@iss.net'   => -3.0,
    'sed@sed-si.com'        =>  5.0,           # blacklist
    '.sed-si.com'           =>  2.0,
    $mydomain               =>  1.3,           # poor-man's SPF
  }
] });

$policy_bank{'MYNETS'} = {
  score_sender_maps =>
    [ @score_sender_maps,
      { '.' => [
        { $mydomain => -1.3 }                  # compensate
    ] } ],
};
```

policy banks

```
$policy_bank{'ALT'} = {
  forward_method => 'smtp:*:*',
  local_client_bind_address => '193.2.4.6',
  localhost_name => 'extra.example.com',

  defang_spam => 1,
  final_spam_destiny       => D_PASS,
  spam_tag2_level_maps => 6.32,
  spam_kill_level_maps    => 6.72,

  av_scanners => [
    ['Sophos SAVI', \&sophos_savi, "*", [0], [1], qr/^(.*) FOUND$/],
    ['Mail::ClamAV', \&ask_clamav,  "*", [0], [1], qr/^INFECTED: (.+)/],
  ],
};
```

policy banks - Postfix side

```
# incoming mail MX
192.0.2.1:smtp inet  n  -  n  -  -  smtpd
   -o content_filter=smtp-amavis:[127.0.0.1]:10040

# tcp port 587 to be used by internal hosts for mail submission
submission inet  n  -  n  -  -  smtpd
   -o content_filter=smtp-amavis:[127.0.0.1]:10042
   -o smtpd_client_restrictions=permit_mynetworks,reject

# incoming mail from fetchmail
127.0.0.1:2345 inet  n  -  n  -  -  smtpd
   -o content_filter=smtp-amavis:[127.0.0.1]:10041
   -o smtpd_client_restrictions=permit_mynetworks,reject
   -o mynetworks=127.0.0.0/8

# locally originating mail submitted on this host through a sendmail msp
pickup     fifo  n  -  n  60  1  pickup
   -o content_filter=smtp-amavis:[127.0.0.1]:10043
```

policy banks - Postfix side

```
content_filter = smtp-amavis:[127.0.0.1]:10044

smtpd_recipient_restrictions =
  reject_...
  check_client_access cidr:/etc/postfix/filter.cidr
  permit_sasl_authenticated
  reject_unauth_destination
  check_sender_access static:FILTER:smtp-amavis:[127.0.0.1]:10040
```

instead of permit_mynetworks, overrides global content_filter setting:
/etc/postfix/filter.cidr :

```
  127.0.0.0/8      FILTER smtp-amavis:[127.0.0.1]:10042
  10.0.0.0/8       FILTER smtp-amavis:[127.0.0.1]:10042
  172.16.0.0/12    FILTER smtp-amavis:[127.0.0.1]:10042
  192.168.0.0/16   FILTER smtp-amavis:[127.0.0.1]:10042
```

Represents additional information as TCP port numbers.

Until some better mechanism becomes available for passing additional information to a content filter, perhaps over XFORWARD

policy banks - Postfix side

How to bypasss content filtering for mail from particular subnets.

Postfix-only solution: restriction classes / access maps, e.g.:

```
smtpd_recipient_restrictions =
  check_client_access cidr:/etc/postfix/filter.cidr, ...
  reject..., permit_mynetworks, reject_unauth_destination, ...
```

/etc/postfix/filter.cidr :

```
  10.11.0.0/16    DUNNO
  172.16.0.0/12  DUNNO
  0.0.0.0/0         FILTER smtp-amavis:[127.0.0.1]:10024
  ::/0                 FILTER smtp-amavis:[127.0.0.1]:10024
```

static lookup tables

- associative array lookups (Perl hash):
  ('me.ac.uk' => 1, '.ac.uk' => 0, '.uk' => 1)
  or: *read_hash('/etc/mydomains-hash')*

- list lookups (acl):
  ('me.ac.uk', '!.ac.uk', '.uk')
  or: *qw(me.ac.uk !.ac.uk .uk)*
  or: *read_array('/etc/mydomains-list')*

- regular expressions list:
  new_RE(qr'[@.]example\.com$'i, qr'[@.]example\.net$'i))

- constant
  1, or: 'string'

static lookup tables

@*_maps are lists of references to lookup tables:

- @local_domains_maps = (); # empty list
- @local_domains_maps = (1); # one element list of const
- @local_domains_maps = ([".$mydomain"]); # one element list, acl
- @local_domains_maps =
  ([qw(.example.com !host.sub.example.net .sub.example.net)]) ;
- @local_domains_maps = (new_RE(qr'[@.]example\.com$'i)) ;
- @local_domains_maps = (read_hash ("$MYHOME/local_domains"));

SQL lookups:

```
CREATE TABLE users (
   id                          SERIAL PRIMARY KEY,
   priority                    integer,              -- 0 is low priority
   policy_id                   integer unsigned,
   email                       varchar(255),
   local                       char(1)
 );
CREATE TABLE policy (
   id                          SERIAL PRIMARY KEY,
   virus_lover                 char(1),
   spam_lover                  char(1),
   virus_quarantine_to         varchar(64),
   spam_quarantine_to          varchar(64),
   spam_kill_level float,
   addr_extension_spam         varchar(64),
   banned_rulenames            varchar(64)           -- comma-separated list
);
SELECT  *,  users.id
  FROM users LEFT JOIN policy ON users.policy_id=policy.id
  WHERE users.email IN (?,?,?,...)
  ORDER BY users.priority DESC
```

SQL lookups

Postfix SQL lookup table search order is:

- user+foo@example.com
- user@example.com
- user+foo
- user
- @example.com

subdomain lookups controlled by
parent_domain_matches_subdomains

amavisd-new SQL search order is sorted by field *priority:*

- user+foo@example.com
- user@example.com
- user+foo *(only if domain part is local)*
- user *(only if domain part is local)*
- @example.com
- @.example.com
- @.com
- @.

SQL lookups - per-recipient w/blacklist

static equivalents:

@whitelist_sender_maps, @blacklist_sender_maps, @score_sender_maps

puts sender and recipient in relation *wb* (white- or blacklisted sender)

```
CREATE TABLE wblist (
  rid       integer unsigned,          -- recipient: users.id
  sid       integer unsigned,          -- sender: mailaddr.id
  wb        varchar(10)                -- W or Y / B or N / space=neutral / score
);
CREATE TABLE mailaddr (
  id        SERIAL PRIMARY KEY,
  priority  integer,
  email     varchar(255) NOT NULL
);
SELECT wb
  FROM wblist LEFT JOIN mailaddr ON wblist.sid=mailaddr.id
  WHERE (wblist.rid=?) AND (mailaddr.email IN (?,?,?...))
  ORDER BY mailaddr.priority DESC
```

SQL logging and reporting:

```
CREATE TABLE msgs (
   mail_id        varchar(12),       -- long-term unique mail id
   secret_id      varchar(12),       -- secret counterpart of mail_id for releasing
   am_id          varchar(20),       -- amavisd log id
   time_num       integer unsigned,  -- rx_time: second since Unix epoch
   time_iso       char(16),          -- rx_time: ISO8601 UTC ascii time
   sid            integer unsigned,  -- sender: maddr.id
   policy         varchar(255),      -- policy bank path (like macro %p)
   client_addr    varchar(255),      -- SMTP client IP address (IPv4 or v6)
   size           integer unsigned,  -- message size in bytes
   content        char(1),           -- content type: V/B/S/H/O/C, NULL
   quar_type      char(1),           -- quarantined as: ' '/F/Z/B/Q/M
   dsn_sent       char(1),           -- was DSN sent? Y/N/q (q=quenched)
   spam_level     float,             -- base message spam level (no boosts)
   message_id     varchar(255),      -- mail Message-ID header field
   from_addr      varchar(255),      -- mail From header field,    UTF8
   subject        varchar(255),      -- mail Subject header field, UTF8
   host           varchar(255)       -- hostname where amavisd is running
);
```

SQL logging and reporting:

```
CREATE TABLE maddr (
  id            SERIAL PRIMARY KEY,
  email         varchar(255) NOT NULL,  -- full mail address
  domain        varchar(255) NOT NULL   -- reverse: com.example.host
);

CREATE TABLE msgrcpt (
  mail_id       varchar(12),      -- (must allow duplicates)
  rid           integer unsigned,  -- recipient: maddr.id
  ds            char(1),          -- delivery status: P/R/B/D/T
                                     (pass/reject/bounce/discard/tempfail)
  rs            char(1),          -- release status: initialized to ' '
  bl            char(1),          -- sender blacklisted by this recip
  wl            char(1),          -- sender whitelisted by this recip
  bspam_level   float,            -- spam level + per-recip boost
  smtp_resp     varchar(255)      -- SMTP response
);
```

SQL quarantine:

- enabled by: *$*_quarantine_method = 'sql:'*

```
CREATE TABLE quarantine (
  mail_id      varchar(12),          -- long-term unique mail id
  chunk_ind    integer unsigned,     -- chunk number (1...)
  mail_text    text                  -- store mail as chunks up to 16 kB
);
```

- amavisd-release utility
- secret_id

SQL reports - an example

```
SELECT
    UNIX_TIMESTAMP()-time_num AS age,  SUBSTRING(policy,1,2) as pb,
    content AS c,  dsn_sent as dsn,  ds,  bspam_level AS level,  size,
    SUBSTRING(sender.email,1,18) AS s,
    SUBSTRING(recip.email,1,18)    AS r,
    SUBSTRING(msgs.subject,1,10) AS subj
  FROM msgs LEFT JOIN msgrcpt              ON msgs.mail_id=msgrcpt.mail_id
              LEFT JOIN maddr AS sender ON msgs.sid=sender.id
              LEFT JOIN maddr AS recip    ON msgrcpt.rid=recip.id
  WHERE content IS NOT NULL AND UNIX_TIMESTAMP()-time_num < 100
  ORDER BY msgs.time_num DESC LIMIT 10;
```

```
| age| pb | c | dsn| ds | level  | size   | s      | r      | subj       |
+----+----+---+----+----+--------+--------+--------+--------+------------+
|  5 |    | C | N  | P  |  5.103 |  44032 | boj... | maj... | RE: PREDRA |
|  6 |    | S | q  | B  | 51.105 |  17974 | sup... | tom... | Important  |
|  6 |    | C | N  | P  | -1.329 |   1476 | sim... | spe... | RE:        |
| 10 | MY | C | N  | P  | -2.267 |   5356 | z...   | mar... | RE: cevlji |
| 14 | MY | C | N  | P  |   NULL | 534357 | m...   | sil... | FW: DRAGE  |
| 15 | MY | C | N  | P  | -5.755 |   4012 | r...   | pre... | Re: Levany |
| 18 | MY | C | N  | P  |  -5.51 |   2209 | l...   | tat... | Financno p |
| 22 |    | S | q  | B  | 18.072 |   1430 | 8m1... | mar... | Mortgage N |
| 23 |    | S | q  | B  |  24.61 |   1635 | caw... | sas... | Hook up wi |
| 23 | MY | C | N  | P  |   NULL | 379281 | f...   | sab... |            |
```

address extensions - "plus addressing"

jim@example.com =>
  jim+spam@example.com
  jim+cooking@example.com
  jim+health@example.com
  jim+postfix@example.com

```
$recipient_delimiter = '+';
@addr_extension_spam_maps = ('spam');
$sa_tag2_level_deflt = 6.7;                 # spam extension is added
$sa_kill_level_deflt  = 15;                 # block higher score entirely
$final_spam_destiny = D_DISCARD;     # junk all above kill level
```

address extensions - "plus addressing"

For the Postfix ***virtual(8)*** LDA, a virtual_mailbox_maps may look like:

```
user1        mbxfile1
user1+spam   mbxspamfile1
user2        mbxfile2
user2+spam   mbxspamfile2
```

For the Postfix ***local(8)*** LDA, a presence of file *$HOME/.forward+spam* can redirect mail for *user+spam* to some dedicated file.

To reroute extension-tagged mail to a mailbox away from the usual LDA, use Postfix virtual alias mapping:

```
/^(.*)\+spam@([^@]*)\.example\.com$/  spam-$2-box@example.com
```

IPv6 is supported

- amavisd: header parsing, access control (IP lookups)
- Perl modules: SMTP client (almost), Net::Server (not yet)
- Postfix: mynetworks, access restrictions, XFORWARD, ...
- mynetworks =
  [::1]/128, [fe80::]/10, [2001:1470:ff80::]/48, 127.0.0.0/8, ...
- smtpd_client_event_limit_exceptions =
  127.0.0.0/8, [::1], 192.0.2.1 ...
- 10025 inet n - n - - smtpd
  -o content_filter=
  -o mynetworks=127.0.0.0/8,[::1]

banning rules

- P=p003,L=1,M=multipart/alternative | P=p001,L=1/1,M=text/plain,T=asc
- P=p003,L=1,M=multipart/alternative | P=p002,L=1/2,M=text/html,T=asc

- P=p003,L=1,M=multipart/related | P=p001,L=1/1,M=text/html,T=html
- P=p003,L=1,M=multipart/related |
  P=p002,L=1/2,M=image/gif,T=image,T=gif,N=kilohm.GIF

- P=p003,L=1,M=multipart/mixed | P=p001,L=1/1,M=text/html,T=html
- P=p003,L=1,M=multipart/mixed | P=p002,L=1/2,M=application/octet-stream,
  T=exe,T=exe-ms,N=foto1.com | P=p004,L=1/2/1,T=empty,N=1979

- P=p003,L=1,M=multipart/mixed | P=p002,L=1/2,M=application/octet-stream,
  T=zip,N=test.zip | P=p004,L=1/2/1,T=exe,T=exe-ms,N=test.scr

- P=p004,L=1,M=multipart/report | P=p001,L=1/1,M=text/plain,T=asc
- P=p004,L=1,M=multipart/report |
  P=p002,L=1/2,M=message/delivery-status,T=asc
- P=p004,L=1,M=multipart/report |
  P=p003,L=1/3,M=text/rfc822-headers,T=txt

banning rules

@banned_filename_maps vs. $banned_namepath_re:

- @banned_filename_maps matches each component in turn, root to leaves
- all its attributes in one go: P, L, M, T, N, A (Part, Location, Mime type, file(1) short Type, Name, Attributes(C,U))

- $banned_namepath_re matches as a single string:
  - | => \n
  - , => \t
- P=p003\tL=1\tM=multipart/related\n
  P=p002\tL=1/2\tM=image/gif\tT=image\tT=gif\tN=kilohm.GIF

regular maintenance tasks

- run *amavisd-nanny*, note any '*process went away*' reports, investigate and fix the problem if any
- check *mailq* or *qshape* for stalled mail messages
- check for preserved directories in */var/amavis/tmp*, search log for explanation, fix the problem and delete;
- remove old quarantine messages
  (2.3.0 quarantine directory adds one level of directories)

regular tasks:
purging log/reporting SQL database

-
  ```
  DELETE FROM msgs
    WHERE UNIX_TIMESTAMP()-time_num > 7*24*60*60;
  ```

-
  ```
  DELETE FROM msgs
    WHERE UNIX_TIMESTAMP()-time_num > 60*60  AND content IS NULL;
  ```

-
  ```
  DELETE quarantine FROM quarantine LEFT JOIN msgs USING(mail_id)
    WHERE msgs.mail_id IS NULL;
  ```

-
  ```
  DELETE msgrcpt  FROM msgrcpt  LEFT JOIN msgs USING(mail_id)
    WHERE msgs.mail_id IS NULL;
  ```

-
  ```
  DELETE FROM maddr
     WHERE NOT EXISTS (SELECT sid FROM msgs WHERE sid=id)
    AND NOT EXISTS (SELECT rid FROM msgrcpt WHERE rid=id);
  ```

-
  ```
  OPTIMIZE TABLE msgs, msgrcpt, maddr, quarantine;
  ```

SpamAssassin - care and feeding

- su vscan -c *'sa-learn --showdots --force-expire --sync'*
- su vscan -c *'pyzor discover'*

- rules_du_jour
- su vscan -c *'spamassassin --lint -D'*

- OPTIMIZE TABLE
  bayes_expire, bayes_seen, bayes_token, awl;

monitoring health: amavisd-agent

```
entropy         STR pIynSOVCq0TQ
sysContact      STR
sysDescr        STR amavisd-new-2.3.1 (20050509)
sysLocation     ST
sysName         STR patsy.ijs.si
sysObjectID     OID 1.3.6.1.4.1.15312.2.1
sysServices     INT 64

sysUpTime       Timeticks 5062346 (0 days, 14:03:43.46)

InMsgs                    14490   1030/h  100.0 % (InMsgs)
InMsgsRecips              27169   1932/h  187.5 % (InMsgs)
InMsgsNullRPath            1084     77/h    7.5 % (InMsgs)

  ==> 1.9 recipients per message, 7.5 % bounces


ContentCleanMsgs           6020    428/h   41.5 % (InMsgs)
ContentSpamMsgs            7807    555/h   53.9 % (InMsgs)
ContentVirusMsgs            567     40/h    3.9 % (InMsgs)

ContentBadHdrMsgs            91      6/h    0.6 % (InMsgs)
ContentBannedMsgs             5      0/h    0.0 % (InMsgs)
```

monitoring: amavisd-agent

```
CacheAttempts       14490 1030/h 100.0 % (CacheAttempts)
CacheHits            1663  118/h  11.5 % (CacheAttempts)
CacheMisses         12827  912/h  88.5 % (CacheAttempts)

CacheHitsSpamCheck   1199   85/h   8.3 % (CacheAttempts)
CacheHitsSpamMsgs     798   57/h  10.2 % (ContentSpamMsgs)
CacheHitsVirusCheck  1259   90/h   8.7 % (CacheAttempts)
CacheHitsVirusMsgs     14    1/h   2.5 % (ContentVirusMsgs)
```

monitoring: amavisd-agent

```
OpsDec                    14490   1030/h   100.0 %  (InMsgs)
OpsDecByMimeParser        14490   1030/h   100.0 %  (InMsgs)

OpsDecByUUlibAttempt      11475    816/h    79.2 %  (InMsgs)
OpsDecByUUlib                91      6/h     0.6 %  (InMsgs)

OpsDecByArZipAttempt        775     55/h     5.3 %  (InMsgs)
OpsDecByArZip               266     19/h     1.8 %  (InMsgs)

OpsDecByLhaAttempt          508     36/h     3.5 %  (InMsgs)
OpsDecByLha                 355     25/h     2.4 %  (InMsgs)

OpsDecByUnrarAttempt        510     36/h     3.5 %  (InMsgs)
OpsDecByUnrar                 2      0/h     0.0 %  (InMsgs)

OpsDecByPax                   4      0/h     0.0 %  (InMsgs)
OpsDecByTnef                 17      1/h     0.1 %  (InMsgs)
OpsDecByZlib                  4      0/h     0.0 %  (InMsgs)
```

monitoring: amavisd-agent

```
OpsSpamCheck          12719    904/h   87.8 % (InMsgs)
OpsVirusCheck         13231    941/h   91.3 % (InMsgs)
OpsSqlSelect          50680   3604/h  186.5 % (InMsgsRecips)

OutMsgs                6248    444/h  100.0 % (OutMsgs)
OutMsgsDelivers        6248    444/h  100.0 % (OutMsgs)

OutForwMsgs            6155    438/h   98.5 % (OutMsgs)

OutDsnMsgs               35      2/h    0.6 % (OutMsgs)
OutDsnBannedMsgs          3      0/h    0.0 % (OutMsgs)
OutDsnSpamMsgs           32      2/h    0.5 % (OutMsgs)
```

monitoring: amavisd-agent

```
QuarMsgs             2704    192/h  100.0 % (QuarMsgs)
QuarSpamMsgs         2100    149/h   77.7 % (QuarMsgs)
QuarVirusMsgs         567     40/h   21.0 % (QuarMsgs)
QuarBannedMsgs          5      0/h    0.2 % (QuarMsgs)
QuarOther              32      2/h    1.2 % (QuarMsgs)
```

monitoring: amavisd-agent

```
OpsDecType-asc            11475    816/h   79.2 % (InMsgs)
OpsDecType-html            4927    350/h   34.0 % (InMsgs)
OpsDecType-txt             3384    241/h   23.4 % (InMsgs)
OpsDecType-doc             1308     93/h    9.0 % (InMsgs)
OpsDecType-dat              531     38/h    3.7 % (InMsgs)
OpsDecType-pdf              215     15/h    1.5 % (InMsgs)
OpsDecType-ps               135     10/h    0.9 % (InMsgs)
OpsDecType-sgml             112      8/h    0.8 % (InMsgs)
OpsDecType-rtf               36      3/h    0.2 % (InMsgs)
OpsDecType-xml               28      2/h    0.2 % (InMsgs)
OpsDecType-lat               17      1/h    0.1 % (InMsgs)
OpsDecType-dvi                4      0/h    0.0 % (InMsgs)
OpsDecType-pgp.pgp.asc        8      1/h    0.1 % (InMsgs)

OpsDecType-exe.exe-ms       508     36/h    3.5 % (InMsgs)
OpsDecType-dll                4      0/h    0.0 % (InMsgs)
OpsDecType-zip              267     19/h    1.8 % (InMsgs)
OpsDecType-tnef              17      1/h    0.1 % (InMsgs)
OpsDecType-tar                4      0/h    0.0 % (InMsgs)
OpsDecType-gz                 4      0/h    0.0 % (InMsgs)
OpsDecType-rar                2      0/h    0.0 % (InMsgs)
OpsDecType-image.jpg       2066    147/h   14.3 % (InMsgs)
OpsDecType-image.gif       1080     77/h    7.5 % (InMsgs)
OpsDecType-image.bmp         14      1/h    0.1 % (InMsgs)
OpsDecType-image.tif         13      1/h    0.1 % (InMsgs)
OpsDecType-image.png         12      1/h    0.1 % (InMsgs)
OpsDecType-image.pcx          2      0/h    0.0 % (InMsgs)
OpsDecType-movie.wmv        167     12/h    1.2 % (InMsgs)
OpsDecType-movie.mpg         41      3/h    0.3 % (InMsgs)
OpsDecType-movie.mpv          3      0/h    0.0 % (InMsgs)
OpsDecType-audio.mpa.mp3     14      1/h    0.1 % (InMsgs)
```

monitoring: amavisd-agent

```
W32/Netsky-P         191  14/h   33.7 % (ContentVirusMsgs)
W32/Mytob-CA          59   4/h   10.4 % (ContentVirusMsgs)
W32/Netsky-D          25   2/h    4.4 % (ContentVirusMsgs)
W32/Lovgate-V         21   1/h    3.7 % (ContentVirusMsgs)
W32/Netsky-Q          21   1/h    3.7 % (ContentVirusMsgs)
W32/Bagle-AG          17   1/h    3.0 % (ContentVirusMsgs)
HTML.Phishing.Pay-1   18   1/h    3.2 % (ContentVirusMsgs)
HTML.Phishing.Bank-1  12   1/h    2.1 % (ContentVirusMsgs)
W32/Mytob-Z           11   1/h    1.9 % (ContentVirusMsgs)
W32/Wurmark-J         11   1/h    1.9 % (ContentVirusMsgs)
W32/Lovgate-X         11   1/h    1.9 % (ContentVirusMsgs)
```

monitoring health: amavisd-nanny

```
PID 28039: 28039-02      0:00:05 =====
PID 28048: .             0:00:05 .....
PID 28174: 28174-01-10   0:00:02 ==
PID 28309: A             0:00:00
```

- db key: PID
- db data: timestamp of last event, status

- status:
  - empty   - idle child process
  - A   - just accepted a connection (post_accept_hook)
  - am_id   - processing am_id task
  - .   - content checking done

monitoring health: amavisd-nanny normal

```
PID 27948: 27948-02-4    0:00:02 ==
PID 27987:               0:00:05 .....
PID 28039: 28039-02      0:00:05 =====
PID 28048: .             0:00:05 .....
PID 28101: 28101-01-9    0:00:01 =
PID 28174: 28174-01-10   0:00:02 ==
PID 28187: 28187-01-5    0:00:12 =========:==
PID 28245: 28245-01-4    0:00:07 =======
PID 28309: A             0:00:00
```

monitoring health: amavisd-nanny mostly idle

```
PID 28187: 28187-02-8    0:00:02 ==
PID 28245:               0:01:16 .........:........>
PID 28309:               0:01:16 .........:........>
PID 28543: 28543-01-7    0:00:03 ===
PID 28584: 28584-01-7    0:00:01 =
PID 28672:               0:00:24 .........:........
PID 28677:               0:01:06 .........:........>
PID 28678:               0:01:06 .........:........>
PID 28729:               0:00:56 .........:........>
```

monitoring health: amavisd-nanny touble - crashed programs

```
PID 25408: 25408-01  went away  0:02:27 =========:=======>
PID 25496: 25496-01  went away  0:01:58 =========:=======>
PID 25728: 25728-01  went away  0:02:06 =========:=======>
```

- process no longer exists, but is still registered in db
- mail is still in MTA queue (temporary failure)
- common symptom: *Lock table is out of available locker entries*
- usual reason: bug in a library routine such as uulib

Monitoring health: amavisd-nanny touble - looping or forgotten proc.

```
PID 25733: 25733-01  terminated 2:10:56 =========:=>
```

- amavisd-nanny sends SIGTERM first
- amavisd-nanny sends SIGKILL 30 seconds later if necessary

- active ttl = 10 minutes   stuck active children
- idle ttl   = 1 hour   unused idle processs (may be normal)

troubleshooting

- amavisd-nanny
- amavisd log and MTA log
- increase log level if necessary
- selective debug: *@debug_sender_maps*
- selective debug: dedicated policy bank with elevated log
- search log for am_id of a trouble message
- compare '*amavisd debug-sa*'
  to '*su vscan -c spamassassin -tD*'

- *strace -f amavisd foreground*

SpamAssassin tips - general

- use_auto_whitelist 1
  # since SA 3.0.0 (2.x: $sa_auto_whitelist)
- choose locking method if needed

- envelope_sender_header Return-Path

- clear_trusted_networks
- clear_internal_networks
- internal_networks 10.0.0.0/8 172.16.0.0/12
- internal_networks 192.168.0.0/16 192.0.2.0/24

SpamAssassin tips: Bayes & AWL on SQL

- *sql/README, sql/README.bayes, sql/README.awl*
- `# su vscan -c 'sa-learn --backup >backup.txt'`
- *local.cf:*

```
bayes_store_module           Mail::SpamAssassin::BayesStore::SQL
bayes_sql_dsn                DBI:mysql:sa:127.0.0.1:3306
bayes_sql_override_username  vscan
bayes_sql_username           vscan
bayes_sql_password  ...

auto_whitelist_factory       Mail::SpamAssassin::SQLBasedAddrList
user_awl_dsn                 DBI:mysql:sa:127.0.0.1:3306
user_awl_sql_username        vscan
user_awl_sql_password        ...
```

- `# su vscan -c 'sa-learn --restore backup.txt'`

SpamAssassin tips:
Bayes & AWL on SQL

- MySQL storage engines: MyISAM, InnoDB, ...
- configuration file: */etc/my.cnf*

- Transaction-safe tables: InnoDB available since MySQL 3.23.34a.
- Since MySQL 4.0 the InnoDB storage engine is enabled by default.

- SA 3.1 provides new module Mail::SpamAssassin::BayesStore::MySQL

```
< bayes_store_module  Mail::SpamAssassin::BayesStore::SQL
> bayes_store_module  Mail::SpamAssassin::BayesStore::MySQL
```

- REQUIRES MySQL version 4.1 or above to work properly!
  - provides rollback on error if bayes db table uses ENGINE=InnoDB
  - provides small boost in performance

```
ALTER TABLE bayes_expire ENGINE=InnoDB;
ALTER TABLE bayes_token  ENGINE=InnoDB;
ALTER TABLE bayes_seen   ENGINE=InnoDB;
```

- Btw, SA 3.1 has Razor2 as a plugin, disabled in init.pre

SpamAssassin tips: Bayes & AWL on SQL

- MyISAM may need repairing
- */var/db/mysql/patsy.ijs.si.err* :

  ```
  050324 19:27:02 [ERROR] Got error 126 when reading table './sa/bayes_token'
  050324 19:27:02 [ERROR] Got error 126 when reading table './sa/bayes_token'
  050324 19:27:02 [ERROR] Got error 126 when reading table './sa/bayes_token'
  050324 19:27:19 [ERROR] Got error 126 when reading table './sa/bayes_token'
  050324 19:27:21 [ERROR] Got error 126 when reading table './sa/bayes_token'
  050324 19:27:21 [ERROR] Got error 126 when reading table './sa/bayes_token'
  ```

-
  ```
  $ mysql sa
  REPAIR TABLE bayes_expire, bayes_seen, bayes_token, awl;
  ```

Tips & Tricks: syslog.conf

- syslog priorities are derived from message log level:
  - level 0: LOG_NOTICE
  - level 2: LOG_INFO
  - lower: LOG_DEBUG

- $log_level = 2;   *# verbosity 0..5*
- $SYSLOG_LEVEL = 'user.debug';

- */etc/syslog.conf* :

```
mail.crit;user.err        /var/log/messages
user.notice               /var/log/amavisd.log
user.info                 /var/log/amavisd-info.log
user.debug                /var/log/amavisd-debug.log
mail.info                 /var/log/mail.log
mail.info;user.info       /var/db/mailgraph/mail.log
```

- Prepend '-' to a filename on Linux to disable sync !

Tips & Tricks - using env. variables

```
$max_servers = $ENV{MAXPROC} || 3;
```

Tips & Tricks: config DIRECTORY

```
my($d) = '/etc/amavis/conf.d';    # do *.cf or *.conf files in this directory

local(*D); opendir(D,$d) or die "Can't open dir $d: $!";
my(@d) = sort grep {/\.(cf|conf)$/ && -f} map {/^(.*)$/,"$d/$1"} readdir(D);
closedir(D) or die "Can't close $d: $!";

for my $f (@d) {
  printf("Reading config file %s\n", $f); $!=undef;
  defined(do $f) or die($@ ? "Error in $f: $@" : "Error reading $f: $!");
}
```

Tips & Tricks

- $spam_quarantine_method = 'sql:';
- $spam_quarantine_method = 'bsmtp:spam/spam-%m';
- $spam_quarantine_method = 'smtp:[127.0.0.1]:10025:quarantine@q.example.com';
- $spam_quarantine_method = 'pipe:argv=/usr/local/sbin/0.sh spam-%b ${sender}';

Tips & Tricks
Perl 'tie' to bind hash to a database

```
my($filename) = "$MYHOME/banned.cdb";

# use existing CDB
my($per_recip_baned) = {};
tie(%$per_recip_baned,'CDB_File',$filename)
  or die "Tie to $filename failed: $!";
@banned_filename_maps = ($per_recip_baned);

# creates an example CDB
use CDB_File;
my($hashref) = {
  'user1@example.com' => 'NO-MS-DOWNLOADS,PASSALL,BLA',
  'user2@example.com' => 'PASSALL,NO-MS-DOWNLOADS',
  '.' => 'DEFAULT',
};
CDB_File::create(%$hashref, $filename, "$filename.tmp$$")
  or die "Can't create cdb $filename: $!";
```

Tips & Tricks: @mynetworks_maps tie with /etc/postfix/mynetworks.db

```
$ postmap -n /etc/postfix/mynetworks


use BerkeleyDB;
my($myPostfixNetworks) = {};  # a ref to an anonymous assoc. array
tie(%$myPostfixNetworks, 'BerkeleyDB::Hash',
    -Filename=>' /etc/postfix/mynetworks.db ', -Flags=>DB_RDONLY)
  or die "Can't open file mynetworks db: $! $BerkeleyDB::Error";

@mynetworks_maps = ( $myPostfixNetworks, \@mynetworks );
```

Tips & Tricks: load %local_domains from Postfix bdb databases

```
use BerkeleyDB;
for my $fname (qw(
  /etc/postfix/mydestination.db
  /etc/postfix/virtual_alias_domains.db
  /etc/postfix/virtual_mailbox_domains.db
  /etc/postfix/relay_domains.db
)) {
  my($db) = BerkeleyDB::Hash->new(-Filename=>$fname, -Flags=>DB_RDONLY);
  defined $db or die "BerkeleyDB opening $fname failed: $BerkeleyDB::Error $!";
  my($cursor) = $db->db_cursor;
  defined $cursor or die "BerkeleyDB db_cursor error: $BerkeleyDB::Error";
  my($key,$val,$stat); $key = '';
  while ( ($stat=$cursor->c_get($key,$val,DB_NEXT))==0 ) {
    for ($key,$val) { chop if /\000\z/ };
    $key = ".$key"  unless $key=~/\@|^\./;  # include its subdomains
    $local_domains{lc($key)} = 1;           # consider this domain local
  }
  $stat==DB_NOTFOUND  or die "BerkeleyDB c_get: $BerkeleyDB::Error $!";
  $cursor->c_close==0 or die "BerkeleyDB c_close error: $BerkeleyDB::Error";
  $db->db_close==0 or die "BerkeleyDB db_close error: $BerkeleyDB::Error $!";
}
```

Tips & Tricks: other topics

- SMTP vs. LMTP for feeding amavisd
- what is 'clean but inconclusive' av scanner result (JPEG checker)

```
['test-jpeg',
  sub { use JpegTester();
          Amavis::AV::ask_av(\&JpegTester::test_jpeg, @_) },
  ["{}/*"], undef, [1], qr/^(bad jpeg: .*)$/ ],
```

- avoid non-C locale

```
%banned_rules = (
 'NO-MS-EXEC'=> new_RE( qr'^\.(exe-ms)$' ),
 'PASSALL'   => new_RE( [qr'^' => 0] ),
 'ALLOW_EXE' =>  # pass executables except if name ends in .vbs .pif .scr .bat
   new_RE( qr'.\.(vbs|pif|scr|bat)$'i, [qr'^\.exe$' => 0] ),
 'ALLOW_VBS' =>  # allow names ending in .vbs
   new_RE( [qr'.\.vbs$' => 0] ),
 'DEFAULT' => $banned_filename_re,
);
@banned_filename_maps = (
 { 'mark.martinec@ijs.si' => 'NO-MS-EXEC,PASSALL',
   'usenet@ijs.si' => 'ALLOW_EXE',
   'user2@ijs.si'  => 'ALLOW_VBS',
   'user3@ijs.si'  => 'ALLOW_VBS,ALLOW_EXE',
   '.' => 'DEFAULT',
 },
);
@banned_filename_maps = (
 { 'mark.martinec@ijs.si' => 'NO-MS-EXEC,PASSALL',
   'usenet@ijs.si' =>
     [ new_RE( qr'.\.(vbs|pif|scr|bat)$'i, [qr'^\.exe$' => 0] ) ],
   'user2@ijs.si'  =>
     [ new_RE( [qr'.\.vbs$' => 0] ) ],
   '.' => [ $banned_filename_re ],
 },
);
```

security

http://www.ijs.si/software/amavisd/#sec-host
http://www.ijs.si/software/amavisd/#sec-mua

- A segmentation violation in uulib kills the Perl process.
  Perl (and amavisd) has no chance of regaining control.

- uulib integer overflow, leading to buffer overflow

- ascii file is mistakenly considered a BinHex file
  and decoding attempted

- numbytes is -16777216
- fread(buffer, 1, (numbytes > 1024 ? 1024 : numbytes), ...)

- Convert-UUlib-1.05 brings fixed (unofficial) uulib,
  thanks to Robert Lewis and Marc Lehmann

Questions?

- mailing list
- hang around and ask
- ...