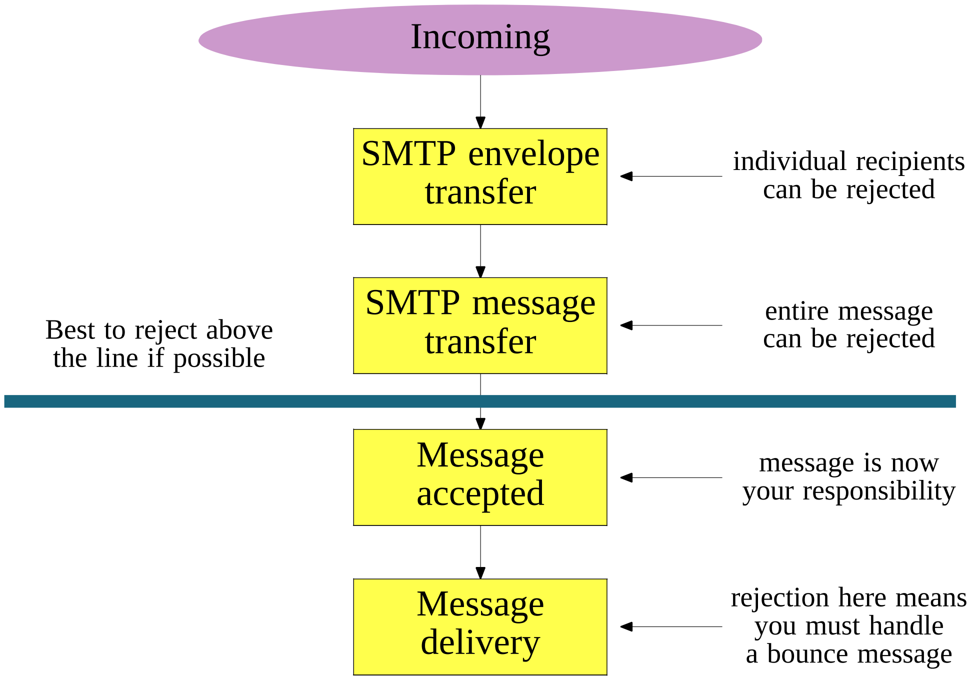
Exim as an anti-spam tool

Philip Hazel

University of Cambridge Computing Service

Where should spam detection happen?

- Exim is an MTA (Mail Transfer Agent)
- Exim's job is to move mail
- Separating spam detection from the MTA is a good idea
  Inflexible to be locked in to one system
  Better to cater for several alternative approaches
- An MTA can contain straightforward general checks
- Use external software for specialist checking
- *Therefore:* An MTA should provide suitable interfaces
  Easy access to external anti-spam applications



Incoming

SMTP envelope transfer

individual recipients can be rejected

SMTP message transfer

entire message can be rejected

Best to reject above the line if possible

Message accepted

message is now your responsibility

Message delivery

rejection here means you must handle a bounce message

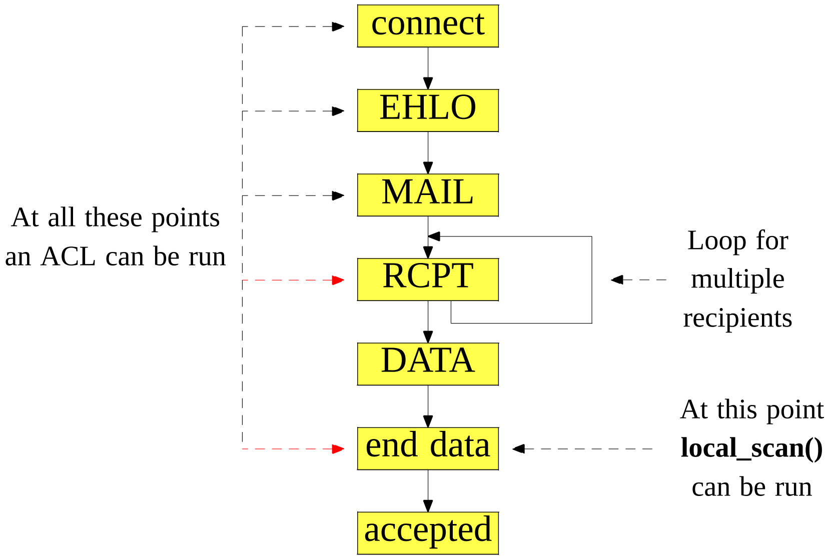
Bounce message problems

- Most spam messages have a forged, but real, sender address
  They used to use non-existent sender addresses
  But then people started checking...
- Deliverable bounces often go to innocent 3rd parties
  Known as ‘collateral spam’ or ‘Joe jobs’
- Undeliverable bounces are stuck on your server
- Avoid generating bounces wherever possible
  Try to reject spam at SMTP time
  Any bounce message is then somebody else’s problem
- Downside: SMTP sessions last longer, tying up resources

Access control in Exim 4

- Exim uses *Access Control Lists* (ACLs)
- An ACL contains *accept* and *deny* rules
  Each rule has a list of attached conditions
- The *warn* rule can add header lines or just log an incident
- Delays in the SMTP dialogue can be specified
- The ACL mechanism is very flexible
- For example:
  Different rules for incoming and outgoing
  Different rules for different recipients – *at RCPT time only*

Exim's ACL checking points


connect
EHLO
MAIL
RCPT
DATA
end data
accepted
At all these points an ACL can be run
Loop for multiple recipients
At this point **local_scan()** can be run

General checks

- Checks on EHLO

  Syntactic validity – in particular, no underscores

  Should name the sending client or its IP

  Malware often names the server or its IP

- EHLO checks are not very effective

- SMTP protocol checks

  Synchronization requirements

  Protects against ‘pump and dump’

- Too many unknown commands

  Protects against subverted web clients

Policy checks at envelope time

- Can reject individual recipients

  Numbers accepted/rejected are available in variables

- Black lists

  Local lists of offending hosts

  DNS-based black lists (e.g. **sbl.spamhaus.org**)

- Domain lists for incoming relays

  Used by site gateways and secondary MXs

- Host lists for outgoing relays

  Typically hosts on your local LAN

- SMTP authentication

  Identifies roaming hosts

  TLS certificates are an alternative

Message content checks at SMTP time

- Can be done only at data time (after message body received)
  Must accept or reject whole message
  Cannot reject individual recipients at this stage

- What if different users want different checks?
  Can temporarily reject some recipients
  A genuine MTA will retry, but there is some delay
  Workable if only 2 or 3 user choices

- Otherwise user-specific checks must be done after reception
  This causes bounce problems if rejection occurs

Implementing virus and spam checks at SMTP time

- Exim’s built-in checks are very limited
  Limited amount of body text in **$message_body** variable
  Newlines and NULs converted to spaces
- Exim does not understand
  MIME
  Multiple character sets
  Different encoding methods
- For serious checking, use a specialist external program
- The Exiscan patch offers easy access to
  MIME checking (including extension blocking)
  Virus checkers (Sophos, clamAV, Kaspersky, etc.)
  Spam checking via SpamAssassin or Brightmail
- SA-Exim offers an alternative interface to SpamAssissin
  SA-Exim uses the *local_scan()* function

Checking messages that have been accepted

- Exim’s system and user filters

  Not powerful enough to do a thorough job

  Designed for sorting mail, autoreply, etc.

  No MIME structure assistance

- Better to deliver to specialist checking software

  Either on another host or on the same host

  Uses more resources (double delivery)

  Leaves you with bounce problems

- On busy hosts, postmasters just discard failed bounces

Testing policy controls

- Exim’s **-bh** option runs a fake SMTP session

  Pretend connection came from any IP address

  Outputs commentary on tests and their outcome

  The entire SMTP dialogue can be simulated

- Exim’s **-bf** and **-bF** options can be used to test filters

  Outputs commentary on how a given message would be filtered

www.exim.org