ScavengerEXA

An open source program
To fight spam at the source

Thomas Mangin
Exa Networks
UKNOF 12
2009/02/13

http://wiki.exa.org.uk/doku.phpdo=export_s5&id=scavenger:uknof12

The spam battle - Yesterday

Mouse:

Using existing mail servers, open relay
Few high volume source

Cat:

scanning for open relay
Using RBL

-> Trying to block the spam at the source

The spam battle - Today

Mouse:

BOTNETs, Spammers creating their own spam infrastructure
Lots of “low” volume source

Cat:

Bots are simple, not RFC compliant (Greylisting)
Spam traps -> RBLs
when all fails ask SpamAssasin for some CPU time

-> The defense is now at the receiving end

The spam battle - Tomorrow

Mouse:

Larger, more clever bots, able to bypass greylisting
Better spam distribution to become more stealthy

Cat:

Trying to block bots faster, so they send less

-> It looks bleak, doesn't it ?

Using my crystal ball

What will not change:

The use of bots
The use of out of date email address databases
Compromised machine sending spam but no mail before or after.

What should change:

Postmaster should get help some help from the 'network', spammers do.
The fight should be brought back to the edge

Why am I here ?

1- present ScavengerEXA
2 - reduce the spam in //my// mailbox
3 - ?
4 - profit .. not !

Get help from the community to produce a turn key solution

Convince some to help us with an high profile deployment

What is scavengerEXA

On the net:

http://www.thefreedictionary.com/scavenger

Alteration of Middle English scauager, schavager, official charged with street maintenance

In my book:

An carnivore eating Junk created by Exa

http://en.wikipedia.org/wiki/Carnivore_(FBI)
http://en.wikipedia.org/wiki/Junk_mail
http://en.wikipedia.org/wiki/Exa-

ScavengerEXA Design

A collection of several application working together through the network

- mail conversation capture program
- dispatch server
- policy server (postfix policy server alike)
- action servers (email, block spam, etc. the part no ISP wants the same)
- a dummy MTA (return 450 on all messages)

Capture

libpcap based application
keep a track of the smtp conversation:

client command, parameter

EHLO [127.0.0.1]
RCPT TO: <user@domain.com>

mail server answers

250 Please to meet you
550 user does not exists here

Ignore the body of the mail
Transmit a UDP packet for each command to the dispatching server

UDP Message Content

key=value structure, with the following keys (in no particular order) for each unique si:sp -> di:dp

or=pacp (how the packet was created)
in=unique random id (identifying the SMTP conversation)
si=source IP, the potential bot
di=destination IP, mail server contacted
he=EHLO/HELO string
st=state (HELO,MAIL,RCPT,DATA,END-OF-DATA)
re=last recpient email address
rc=number of recpient in the mail
se=sender email address
co=smtp response code

Example

Mail conversation

220 mail server listing
HELO [127.0.0.1]
250 mx.domain.com
MAIL FROM: test@domain.com
250 2.1.0 Ok
RCPT TO: user@spammed.com
550 5.7.1 <user@spammed.com>: Recipient address rejected: no such user
<Disconnection>

We do not track smtp auth information atm

UDP Message sent 1/3

HELO [127.0.0.1]
250 mx.domain.com

or=pcap
co=250
di=4.3.2.1
rc=0
st=EHLO
re=
si=1.2.3.4
in=01.3c6b.4e.c0c3
se=
he=[127.0.0.1]

UDP Message sent 2/3

MAIL FROM: test@domain.com
250 2.1.0 Ok

or=pcap
co=250
di=4.3.2.1
rc=0
st=MAIL
re=
si=1.2.3.4
in=01.3c6b.4e.c0c3
se=test@domain.com
he=[127.0.0.1]

UDP Message sent 3/3

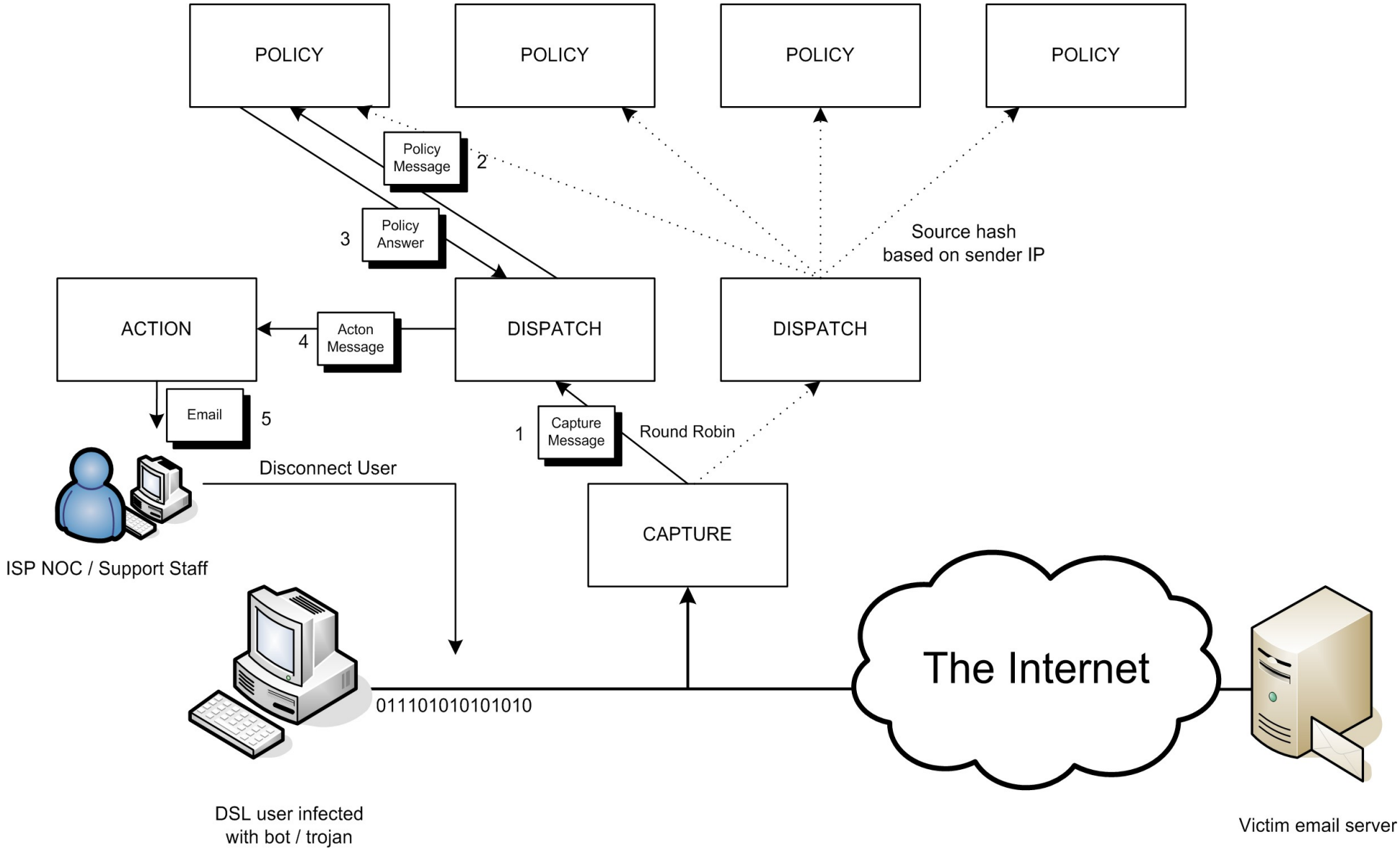
RCPT TO: user@spammed.com
550 5.7.1 <user@spammed.com>: Recipient address rejected: no such user

or=pcap
co=550
di=4.3.2.1
rc=1
st=RCPT
re=user@spammed.com
si=1.2.3.4
in=01.3c6b.4e.c0c3
se=test@domain.com
he=[127.0.0.1]

Dispatch Logic

1 - Receive the UDP message from a capture source
    Figure out a policy server (hash on sender IP)
2 - Generate a TCP message for that policy server with the same information as the UDP contained
3 - Wait for the answer from the policy server
4 - If the policy server reports it is spam
5 - Generate a/several TCP message(s) to action server(s)
    Prevent new message to the Policy Server from the spammer

Message Flow



Policy Daemon

Based on the code of Exa's internal Postfix Policy Delegation Daemon (the daemon can still be used with Postfix)

Use the UDP message format instead of Postfix's

Returns:

HAM
HOLD <IP> (<duration>) <reason>
FILTER <IP> [<MTA IP:PORT>] (<duration>) <reason>

an MTA of [0.0.0.0:00000] mean that the MTA used to filter the message is left to the software performing the blocking to decide.

Policy Daemon

A clear API for its spam classfication plugins
with constraint, run if
the message match what you are monitoring
the backend DB is MySQL ...

Explaining its design would take another 20 slides
writing plugins is **//simple//**

Each plugin has its own database connection.
support MySQL, PostgreSQL or Sqlite3

Action Mail

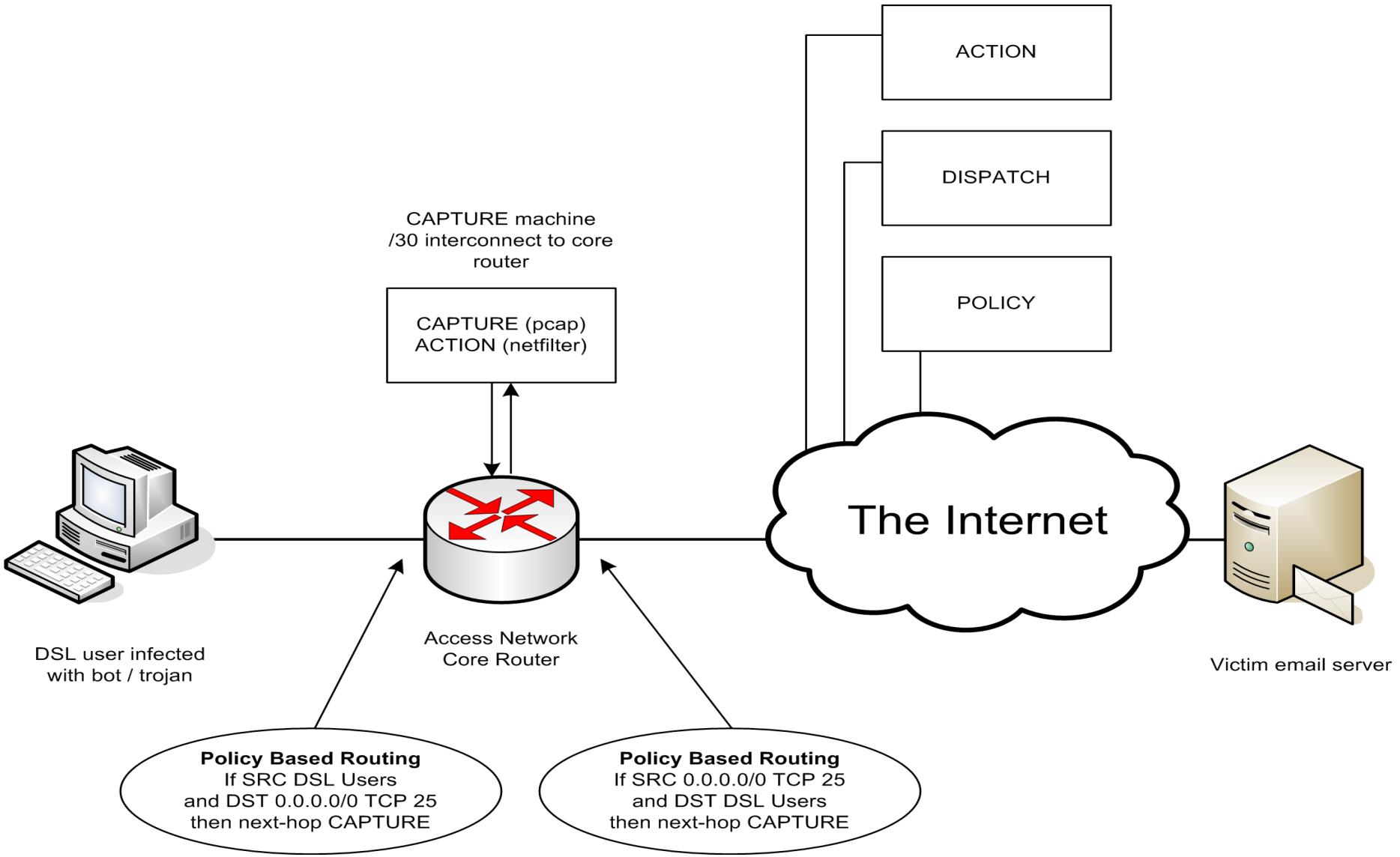
You may want to use this action whatever else you do.

Takes a HOLD or FILTER message
Emails you the suspected spammer information

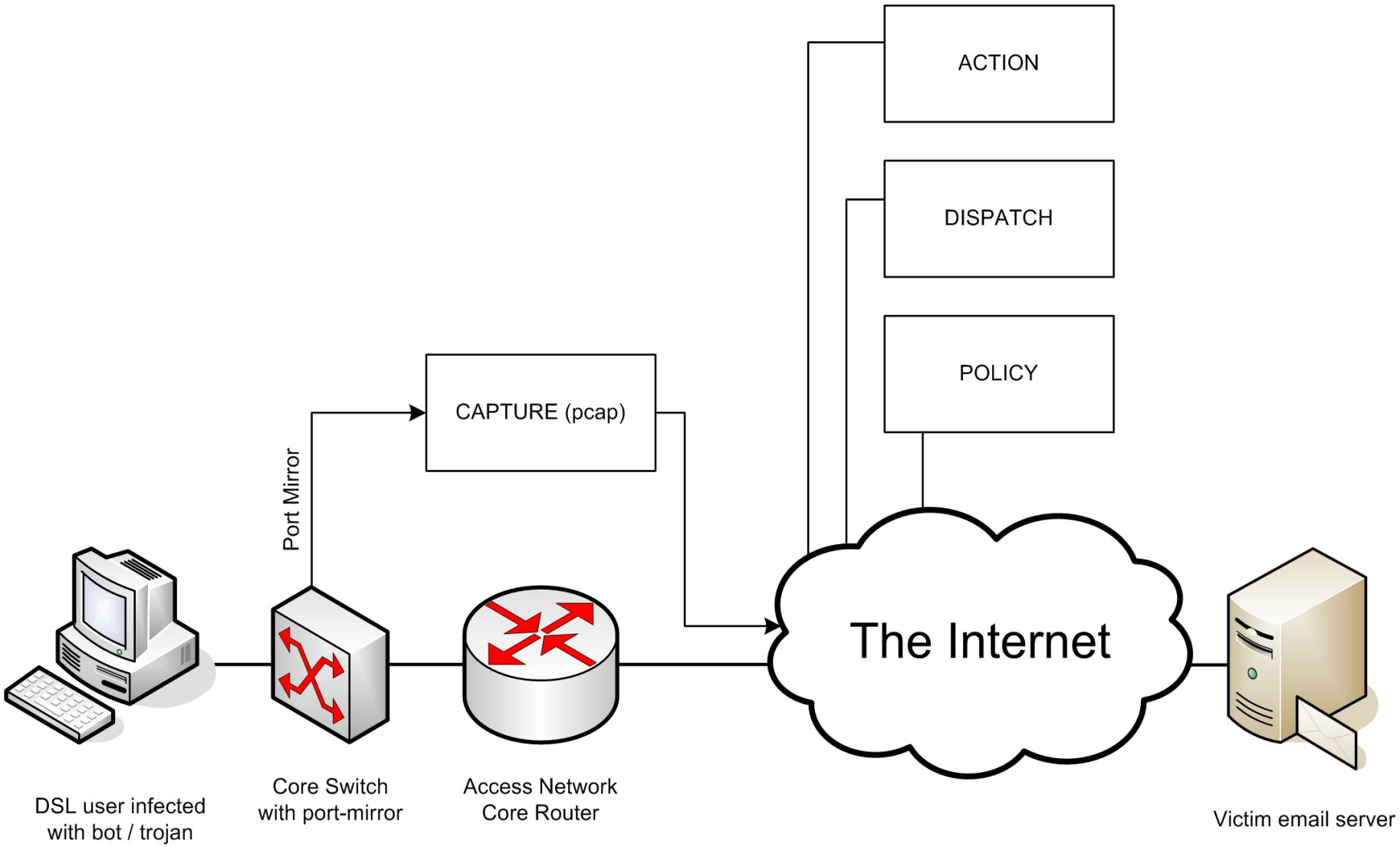
Action Netfilter

Read a Linux machine netfilter rules.
Force Transproxying of mail to a specified MTA on receipt of a FILTER rule
Remove the FILTERING when the filtering period is over
Keep state across reboots

Deployment Example



Deployment Example



It is F.R.E.E

Licensed under the Affero GPL 3.0

Each ScavengerEXA program being a separate entity, there is no problem implementing a new server under a proprietary licence.

AFAIK, IMNAL

It is available now

Public SVN Tree
  http://svn.exa.org.uk/scavenger/trunk

Fisheye
  http://fisheye.exa.org.uk/browse/scavenger/
  provides nice delta of changes
  RSS feed of commits

Mailing list
  http://mailman.exa.org.uk/mailman/listinfo

Build on solid foundations

Python, tested with version 2.5.2
- Twisted Matrix, event-driven networking engine
- Python packet capture library (pcap)
- Python dump packet module (dpkt)
- Optional MySQL or PostgreSQL

For more information see:
- http://wiki.exa.org.uk/scavenger/faq

Thank you to

Richard Clayton SpamHINTS

http://www.spamhints.org/

Questions ?

Want to know more:

website: http://scavenger.exa.org.uk/

Contact us:

email: scavenger (at) exa (dot) org (dot) uk

Thank you for listening.