ExaProxy

open source web proxy

UKNOF - 3rd of May 2012
York

Thomas Mangin
Exa Networks

exa networks

Non-caching Proxy

HTTP/1.1

forward

or transparent proxy

reverse

ALERT !

Non-caching Proxy

HTTP/1.1

High Performance

forward

or transparent proxy

reverse

non-blocking event based network loop

epoll on linux

use of cheap co-routine

conservative memory usage

“pause” reader when writer is too slow

multi-threaded

sockets as message bus

own async DNS library

ALERT !

Non-caching Proxy

HTTP/1.1

forward

or transparent proxy

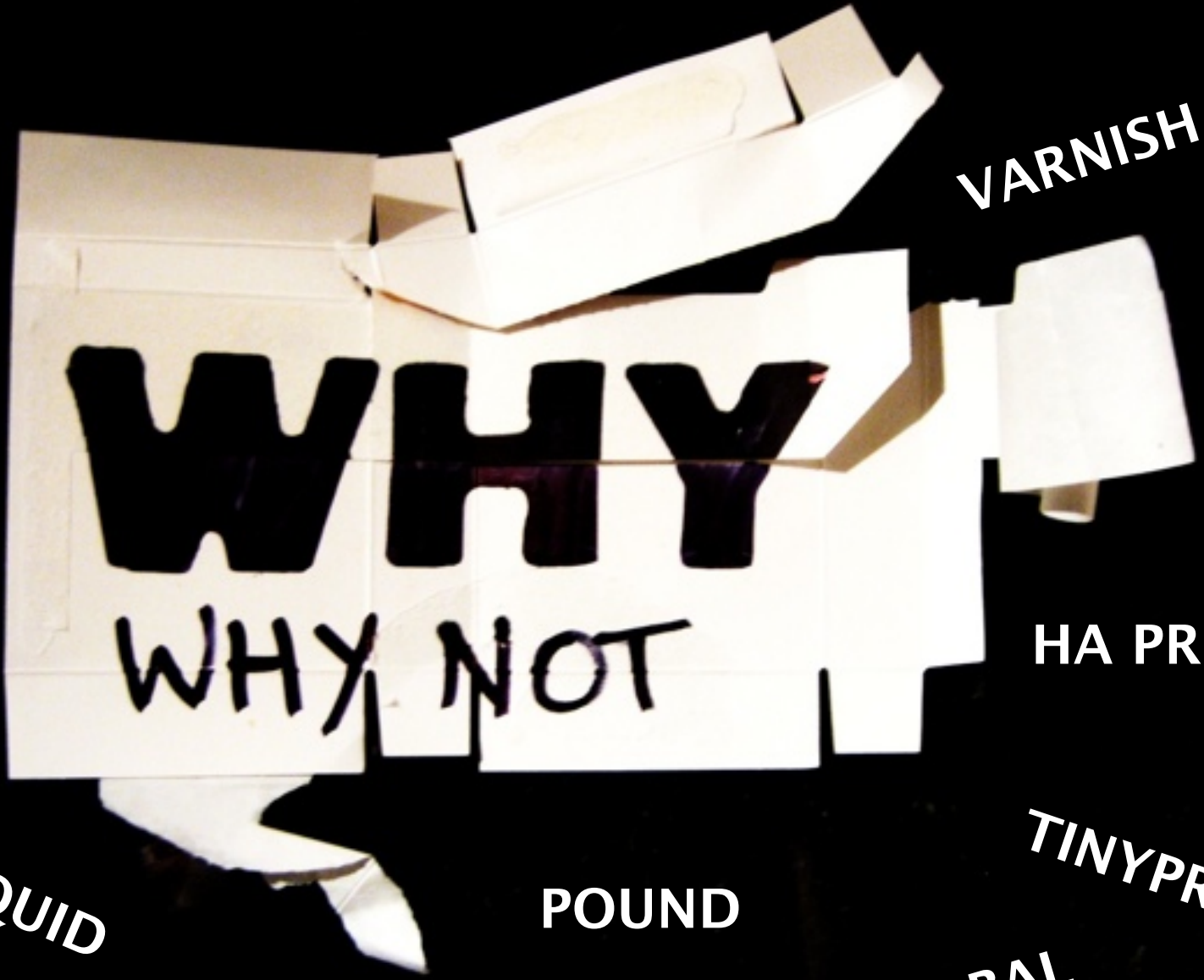
reverse

✓ IPv6 INSIDE

Full native IPv6 support

IPv6 to IPv4 gateway (and vice versa)

QUITE A FEW OPEN SOURCE WEB PROXIES



VARNISH

HA PROXY

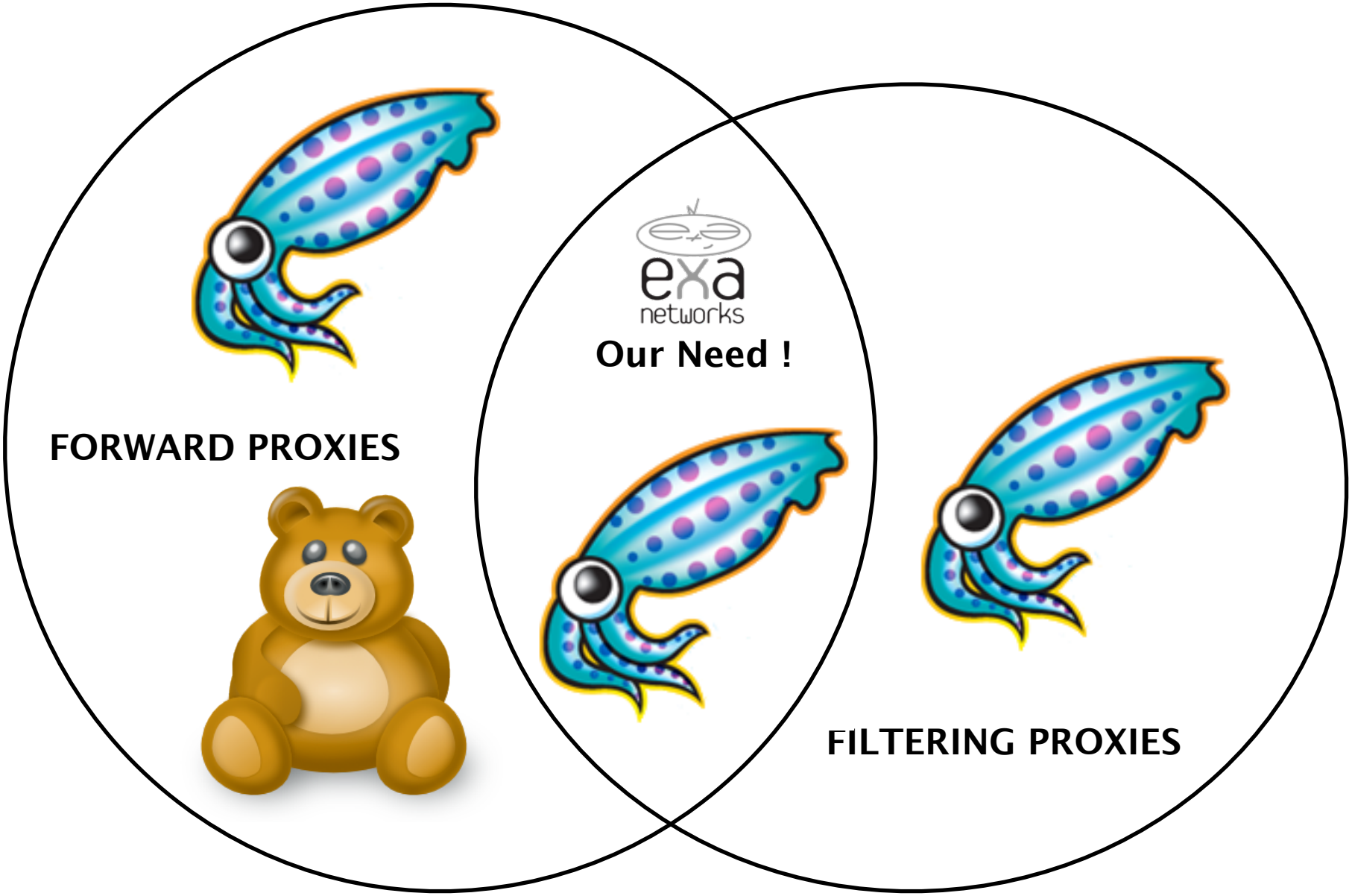
TINYPROXY

PERLBAL

POUND

SQUID

WHAT PROXY ARE AVAILABLE FOR OUR USE



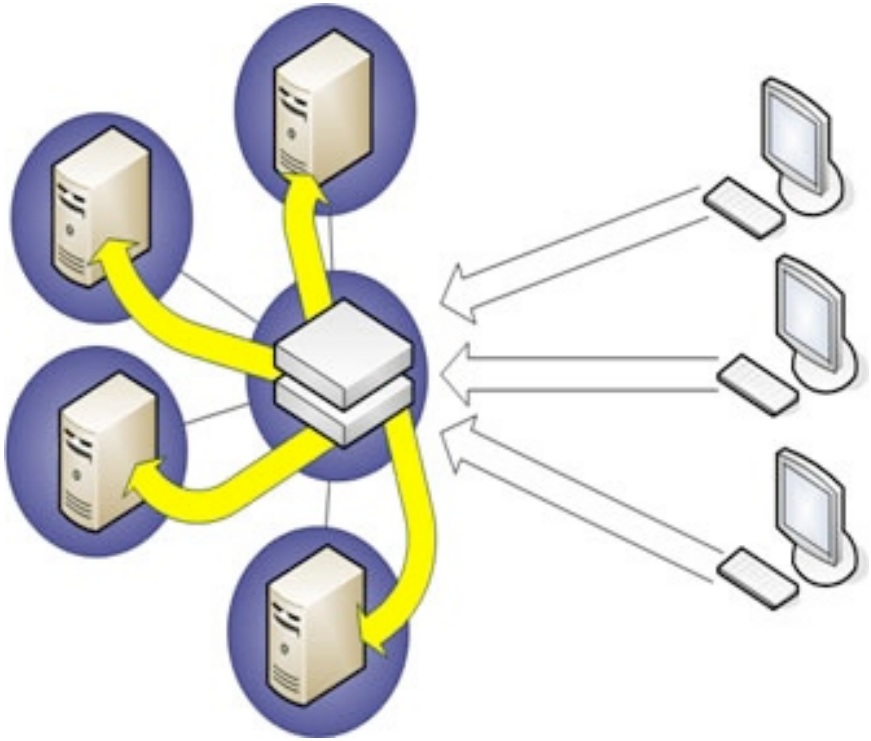
SQUID FILTERING ..

```
/bin/sh -e /proc/self/fd/11
 \_ /usr/sbin/squid -N -D
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ /usr/bin/python /opt/surfprotect/sbin/../prg/squid-redirector-new.py
     \_ (unlinkd)
     \_ diskd 15340544 15340545 15340546
```

A Filtering SQUID cluster ...

Linux ipvsadm for load balancing (MAC address rewrite)
Farm of squid servers

Works However



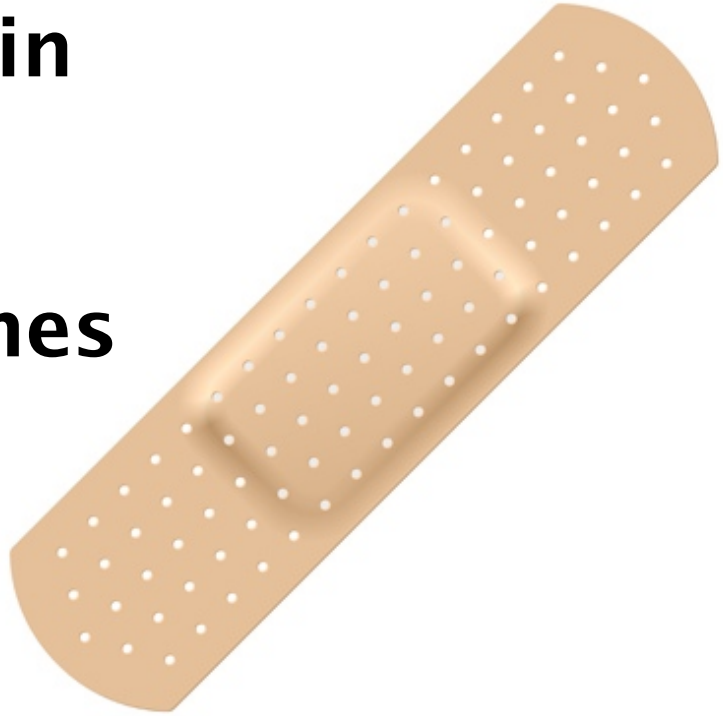
Can't use L7 Load balancer
Can not filter HTTPS (CONNECT)

Load balancing must be sticky
"cascade effect" on failure

no load balancing backend monitoring

And you need to maintain a TWO liner PATCH

SQUID purposefully crashes on high load



```
    debug(84, 1) ("WARNING: All %s processes are busy.\n", hlp->id_name);
    debug(84, 1) ("WARNING: up to %d pending requests queued\n", hlp-
>stats.max_queue_size);
-    if (hlp->stats.queue_size > hlp->n_running * 2)
-        fatalf("Too many queued %s requests (%d on %d)", hlp->id_name, hlp-
>stats.queue_size, hlp->n_running);
```

SQUID compatible and ICAP (REQMOD) like mode

URL Rewrite
display a different URL

Cookie modification
force safe-search on youtube

HTTPS filtering / Interception
when browser/other proxy explicitly configured
redirect the browser to a HTTP page ..

```
CONNECT www.hsbc.com:443 HTTP/1.1
Host: www.hsbc.com
```

```
HTTP/1.1 200 Connection Established
```

```
HTTP/1.1 403 Surfprotected
Connection: close
```

No way to return a message to the browser via 4xx/5xx

```
HTTP/1.1 302 Surfprotected
Cache-Control: no-store
Location: http://www.surfprotect.co.uk/
Connection: close
```

Browsers just disabled this “feature” following some work on HTTBis

Tweets

Nat Morris @natmorris 17 Apr

@ichilton webfiltering for 25k+ users, #HAProxy LB at front, some super quick @exaproxy web proxies and then @opendns for filtering.

Retweeted by ExaProxy

In reply to Ian Chilton

Nat Morris @natmorris 17 Apr

First day of putting some real traffic via @ExaProxy, #HAProxy and @OpenDNS. Things are looking good!

Retweeted by ExaProxy

ExaProxy @ExaProxy 28 Feb

ExaProxy is now in feature freeze .. Ironing the last bugs before a 1.0.0 release.

250+ commits since (and counting)

HAVE FUN ...***

We have !

From: **David Farrar <david.farrar@exa-networks.co.uk>**
Subject: D'oh!
Date: 27 April 2012 12:31:53 GMT+01:00
To: Thomas Mangin <thomas.mangin@exa-networks.co.uk>

I now know why it was such a pain tracking down the source of the memory leak

(It only took a week)

It requires that -

- The client starts a new request over a socket that's already been used for at least one request
- The send buffer to the remote web server was full when we first try sending the new request
- The client is uploading a very large file (or this happens many times with smaller files)
- The upload speed from the client to the proxy is greater than the upload speed from the proxy to the remote web server

http://code.google.com/p/exaproxy/

*** if you are brave, mad, desperate or any of the above

QUESTIONS ?