A Moose Once Bit My Honeypot

A Story of an Embedded Linux Botnet

by Olivier Bilodeau (@obilodeau)

$ apropos

- Embedded Linux Malware
- Moose DNA (description)
- Moose Herding (the Operation)
- What's New?
- Take Aways

$ whoami

- Malware Researcher at ESET
- Infosec lecturer at ETS University in Montreal
- Previously
  - infosec developer, network admin, linux system admin
- Co-founder Montrehack (hands-on security workshops)
- Founder NorthSec Hacker Jeopardy

Embedded Linux Malware

What marketing likes to call "Internet of Things Malware"

Malware Running On An Embedded Linux System

Like

- **consumer routers**
- DVR
- Smart TVs
- IP Camera monitoring systems
- ...

Caracteristics of Embedded Linux Systems

- Small amount of memory
- Small amount of flash
- Non x86 architectures: ARM, MIPS
- Wide-variety of libc implementations / versions
- Same ABI-compatible Linux kernel (2.4 < x < 4.3)
- Support ELF binaries
- Rarely an integrated UI
- Networked

Why Threats On These Systems Matters?

- Hard to detect
- Hard to remediate
- Hard to fix
- Low hanging fruit for bad guys

It’s Real

- Several cases disclosed in the last two years
- A lot of same-old background noise (DDoSer)
- Things are only getting worse

The Incapsula Blog

12 May 2015

Lax Security Opens the Door for Mass-Scale Abuse of SOHO Routers

By Ofer Gayer, Ronen Atias, Igal Zeifman

welivesecurity
Security news, views and insight from the ESET experts

Win32/Sality newest component: a router's primary DNS changer named Win32/RBrute

BY BENJAMIN VANHEUVERZWIJN POSTED 2 APR 2014 - 02:31PM

PETER KÁLNAI

January 6th, 2015

Linux DDoS Trojan hiding itself with an embedded rootkit

KrebsonSecurity
In-depth security news and investigation

09 Lizard Stresser Runs on Hacked Home Routers

JAN 15

The online attack service launched late last year by the same criminals who knocked **Sony** and **Microsoft**'s gaming networks offline over the holidays is powered mostly by thousands of hacked home Internet routers, KrebsOnSecurity.com has discovered.

Home routers 'vaccinated' by benign virus

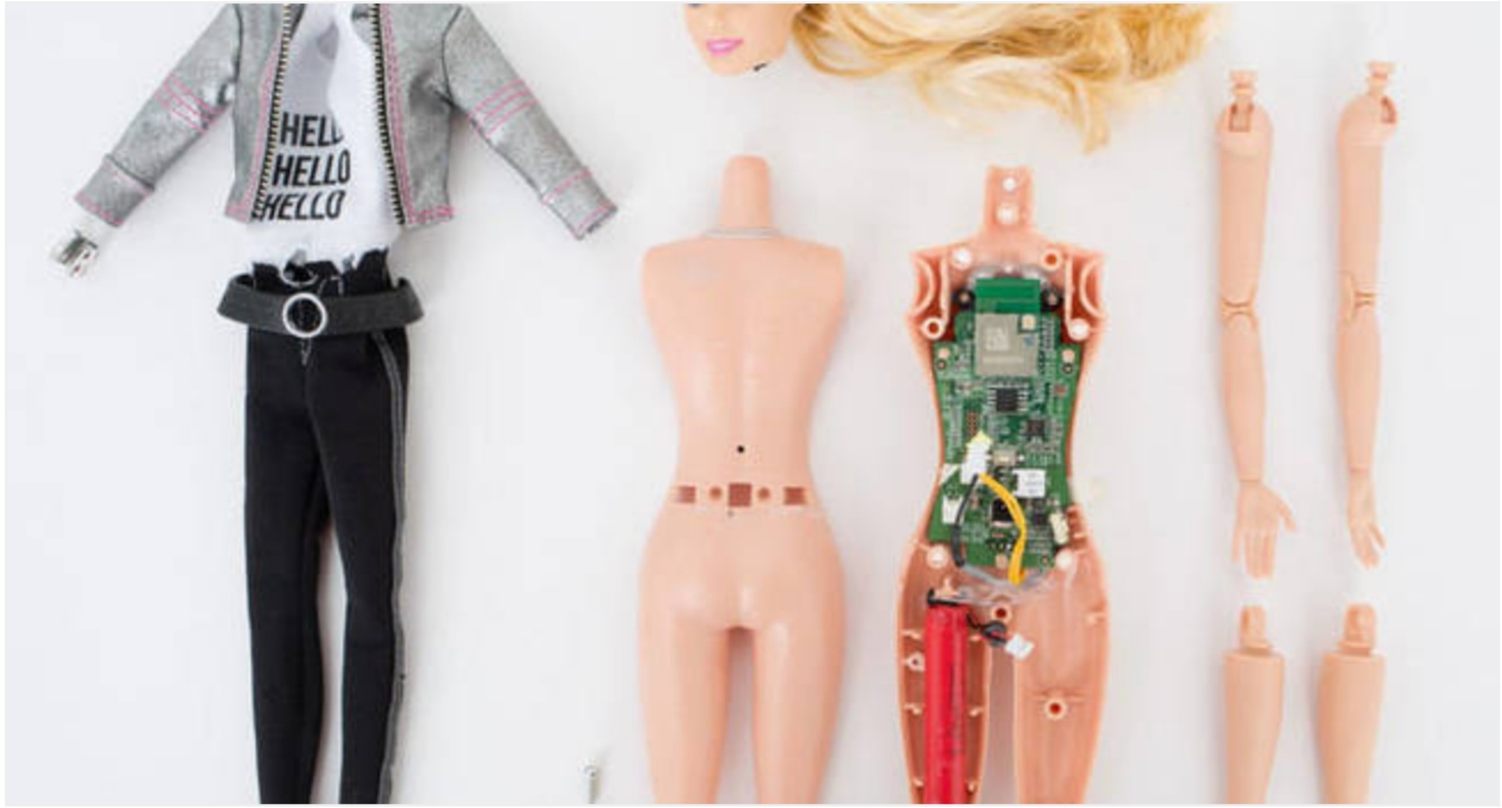
2 October 2015 | Technology

The Register®
Biting the hand that feeds IT

Security

Hello Barbie controversy re-ignited with insecurity claims

Doll leaks data, even before the tear-downs are finished



Wait, is IoT malware really about things?

No. Not yet.


Page 2

So what kind of malware can we find on such insecure devices?

- Linux/Aidra
- Linux/Bassobo
- ChinaZ family (XOR.DDoS, ...)
- Linux/Dofloo
- Linux/DNSAmp (Mr Black, BillGates)
- Linux/Gafgyt (LizardStresser)
- Linux/Hydra
- Linux/Tsunami

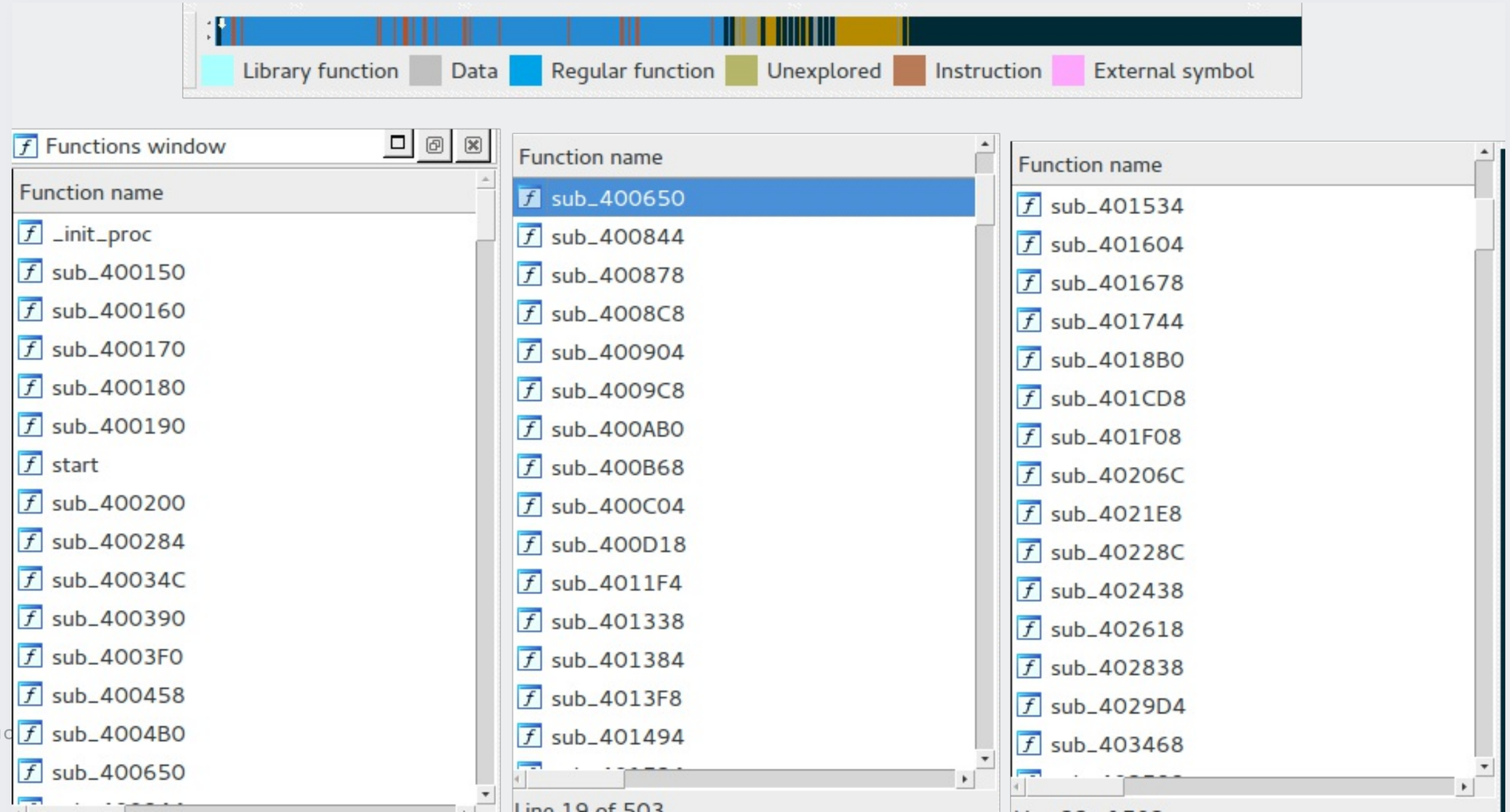
Lesson Learned #0

Statically-linked stripped binaries

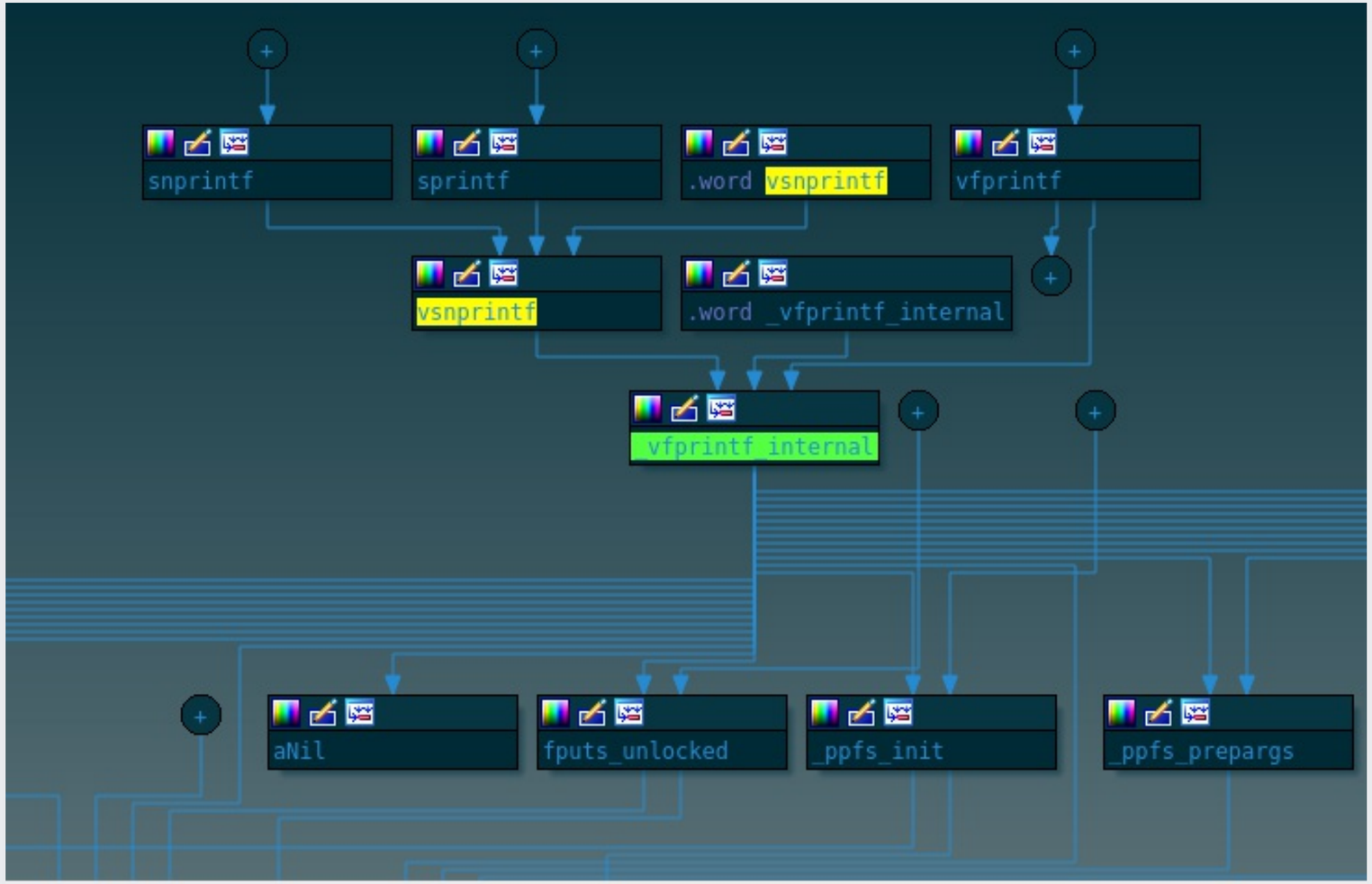
Static/stripped ELF primer

- No imports (library calls) present
- All the code bundled together down to kernel syscall
- Disassembler (if available for arch) doesn't help much

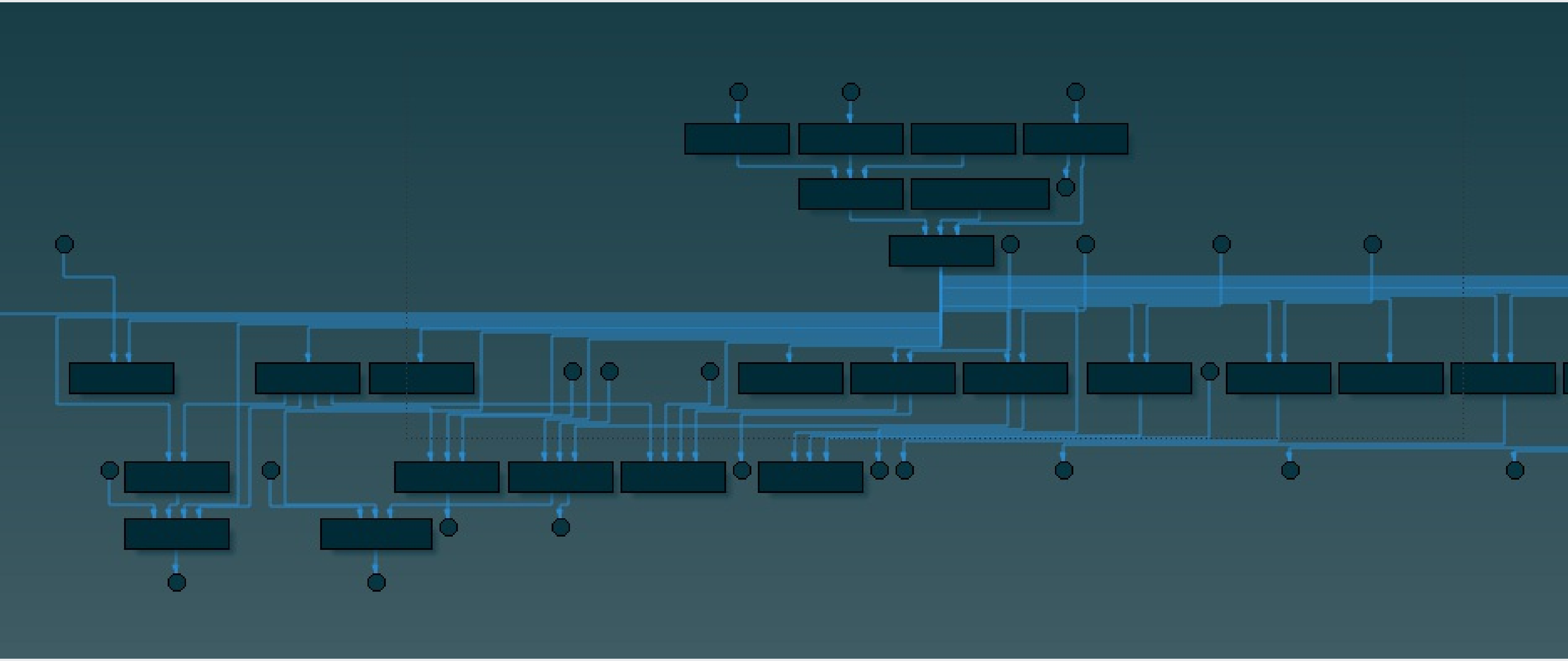
Linux/Moose binary in IDA



printf family

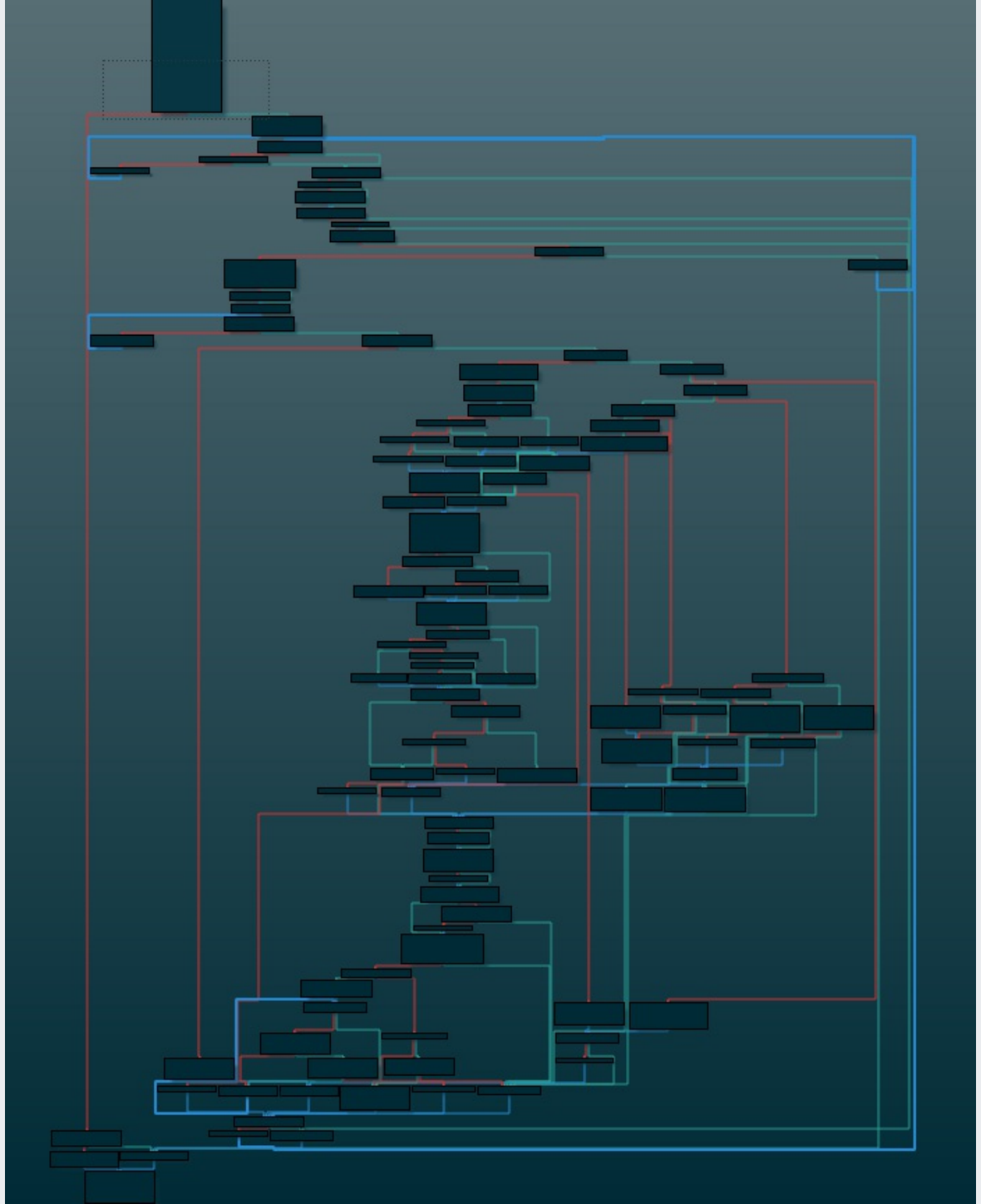


B60 00417B60: vsnprintf



ENJOY SAFER TECHNOLOGY™


WE HAVE TO GO
DEEPER!
DIYLOL.COM



ENJOY SAFER TECHNOLOGY™

Ecosystem makes it worst [for reversers]

- GCC and GNU libc are always changing so compiled binaries always change
- Little IDA FLIRT signatures available (if any)
- Various C libraries: µClibc, eglibc, glibc, musl, …

A Failed Attempt

- Map syscalls with IDA script
- But libc is too big
- Still too much code to RE
- Provided tool: https://github.com/eset/malware-research/blob/master/moose/ida/mips_identify_syscalls.py

Better Solution

- Reproduce environment (arch, libc/compiler versions)
- Build libraries w/ symbols under same conditions
- Use bindiff to map library functions
- Focus on malware code

similarity	confider	change	EA primary	name primary	EA secondary	name secondary	con	algorithm	matched ba
0.99	0.99	-I--E--	00419BE0	sub_419BE0_282	00037E60	strncmp		MD index matching (flowg...	21
0.99	0.99	-I--E--	00423F20	sub_423F20_444	00034C20	fgets		edges flowgraph MD index	18
0.99	0.99	-I--E--	004228D0	sub_4228D0_435	0002D650	__stdio_WRITE		edges flowgraph MD index	17
0.99	0.99	-I--E--	0041B634	sub_41B634_308	0003E7A4	inet_pton4		edges flowgraph MD index	21
0.99	0.99	-I--E--	004261A0	sub_4261A0_471	0002D790	__stdio_adjust_position		edges flowgraph MD index	21
0.99	0.99	-I--E--	00423010	sub_423010_438	0002E1B0	__stdio_trans2w_o		edges flowgraph MD index	17
0.99	0.99	-I--E--	004277D0	sub_4277D0_485	0003F2E0	__encode_dotted		edges flowgraph MD index	17
0.99	0.99	-I--E--	00424790	sub_424790_448	000362F0	fgets_unlocked		edges flowgraph MD index	19
0.99	0.99	-I--E--	00424050	sub_424050_445	00035BB0	_stdio_openlist_dec_use		edges flowgraph MD index	44
0.99	0.99	-I--E--	0041B734	sub_41B734_310	0003E89C	inet_ntop		edges flowgraph MD index	63
0.99	0.99	-I-----	004176FC	sub_4176FC_237	000107E4	opendir		edges flowgraph MD index	11
0.99	0.99	-I--E--	00424FF0	sub_424FF0_456	0003EE60	inet_aton		edges flowgraph MD index	17
0.99	0.99	-I--E--	004189B0	sub_4189B0_260	00030540	_ppfs_init		edges flowgraph MD index	16
0.99	0.99	-I--E--	00419670	sub_419670_268	00036810	fwrite_unlocked		edges flowgraph MD index	15
0.99	0.99	-I--E--	00418218	sub_418218_259	0002FDA8	_vfprintf_internal		edges flowgraph MD index	136
0.99	0.99	-I--E--	00419318	sub_419318_265	000354F0	putchar		address sequence	50
0.99	0.99	-I--E--	00425F00	sub_425F00_469	0002CF30	fseeko64		edges flowgraph MD index	32
0.99	0.99	-I--E--	0041FD70	sub_41FD70_384	0004E960	raise		edges flowgraph MD index	15
0.99	0.99	-I--E--	004224C0	sub_4224C0_431	00024690	wcsnrtombs		edges flowgraph MD index	19
0.99	0.99	-I--E--	00423DC0	sub_423DC0_443	00034AC0	getc		instruction count	48
0.99	0.99	-I--E--	0041B4D0	sub_41B4D0_307	0003E640	inet_ntop4		edges flowgraph MD index	11
0.99	0.99	-I--E--	004285E0	sub_4285E0_495	00051DD0	__fixdfsi		edges flowgraph MD index	13

Lesson #0

- Going down to syscalls is too long in large binaries
- Find a close match of C library
- Build with symbols
- Bindiff it (or maybe FLIRT it)

Lesson Learned #1

Be careful of strings and AV variant names

Anti-Virus Variants

File information

Identification | Details | Content | Analyses | Submissions | ITW | Comments

2014-11-22 09:42:00	**0/55**
2014-12-04 15:57:08	**0/55**
2014-12-12 17:32:59	**0/56**
2014-12-16 12:25:11	**0/56**
2014-12-16 19:19:38	**0/56**
2014-12-17 09:56:50	**4/56**
2014-12-18 11:42:29	**5/55**
2014-12-22 10:15:20	**7/52**
2015-02-03 13:47:36	**5/56**
2015-03-14 15:51:08	**9/57**

Engine	Signature	Version	Update
Ad-Aware	Application.BitCoinminer.GG	12.0.163.0	20141222
Avast	ELF:BitCoinMiner-N [Trj]	8.0.1489.320	20141222
ESET-NOD32	Linux/Agent.P	10913	20141222
F-Secure	batch-timeout	11.0.19100.45	20141221
Fortinet	-	5.0.999.0	20141222
GData	Application.BitCoinminer.GG	24	20141222
Malwarebytes	-	1.75.0.1	20141222
McAfee	-	6.0.5.614	20141222
McAfee-GW-Edition	batch-timeout	None	20141221
Microsoft	-	1.11302	20141222
MicroWorld-eScan	Application.BitCoinminer.GG	12.0.250.0	20141222
Kaspersky	-	15.0.1.10	20141222

Download file | Re-scan file | Close


2004
EVERYONE GETS BAD INTEL
imgflip.com

and Strings

```
$ strings moose_mips.elf
[...]
cat /proc/cpuinfo
GET /xx/rnde.php?p=%d&f=%d&m=%d HTTP/1.1
Host: www.getcool.com
Connection: Keep-Alive
127.0.0.1
[...]
```

Lesson #1

- Be careful with detection names
- Don’t request domain take down based on output of `strings`
- and don’t do so for other people’s research!

Misleading Strings

Moose DNA

aka Malware description

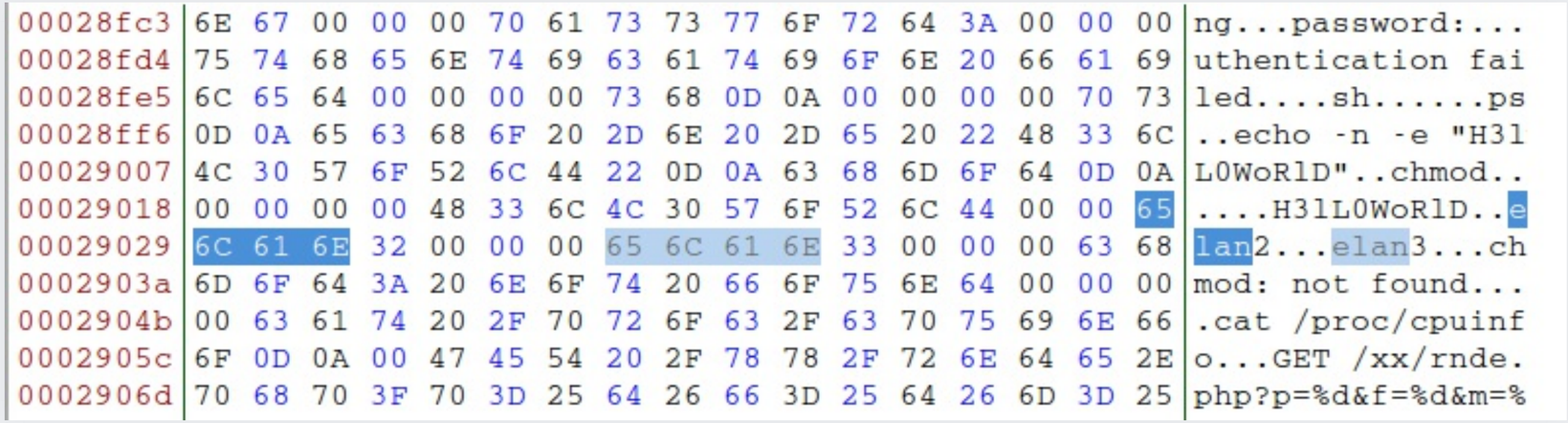
Hang tight, this is a recap

Linux/Moose

- Discovered in November 2014
- Thoroughly analyzed in early 2015
- Published a report in late May 2015

Linux/Moose...

Named after the string "elan" present in the malware executable



```
00028fc3 6E 67 00 00 00 70 61 73 73 77 6F 72 64 3A 00 00 00 ng...password:...
00028fd4 75 74 68 65 6E 74 69 63 61 74 69 6F 6E 20 66 61 69 uthentication fai
00028fe5 6C 65 64 00 00 00 00 73 68 0D 0A 00 00 00 00 70 73 led....sh......ps
00028ff6 0D 0A 65 63 68 6F 20 2D 6E 20 2D 65 20 22 48 33 6C ..echo -n -e "H3l
00029007 4C 30 57 6F 52 6C 44 22 0D 0A 63 68 6D 6F 64 0D 0A L0WoRlD"..chmod..
00029018 00 00 00 00 48 33 6C 4C 30 57 6F 52 6C 44 00 00 65 ....H3lL0WoRlD..e
00029029 6C 61 6E 32 00 00 00 65 6C 61 6E 33 00 00 00 63 68 lan2...elan3...ch
0002903a 6D 6F 64 3A 20 6E 6F 74 20 66 6F 75 6E 64 00 00 00 mod: not found...
0002904b 00 63 61 74 20 2F 70 72 6F 63 2F 63 70 75 69 6E 66 .cat /proc/cpuinf
0002905c 6F 0D 0A 00 47 45 54 20 2F 78 78 2F 72 6E 64 65 2E o...GET /xx/rnde.
0002906d 70 68 70 3F 70 3D 25 64 26 66 3D 25 64 26 6D 3D 25 php?p=%d&f=%d&m=%
```

Elan is French for

The Lotus Elan

Elán

The Slovak rock band (from 1969 and still active)


populár
ELÁN

Sample

- Statically linked stripped ELF binary
- ARM (GNU EABI and EABI 5)
- MIPS (little and big endian)
- No x86 sample found
- C&C IP in integer form buried in all this code

MIPS/ARM + statically linked + stripped + no x86

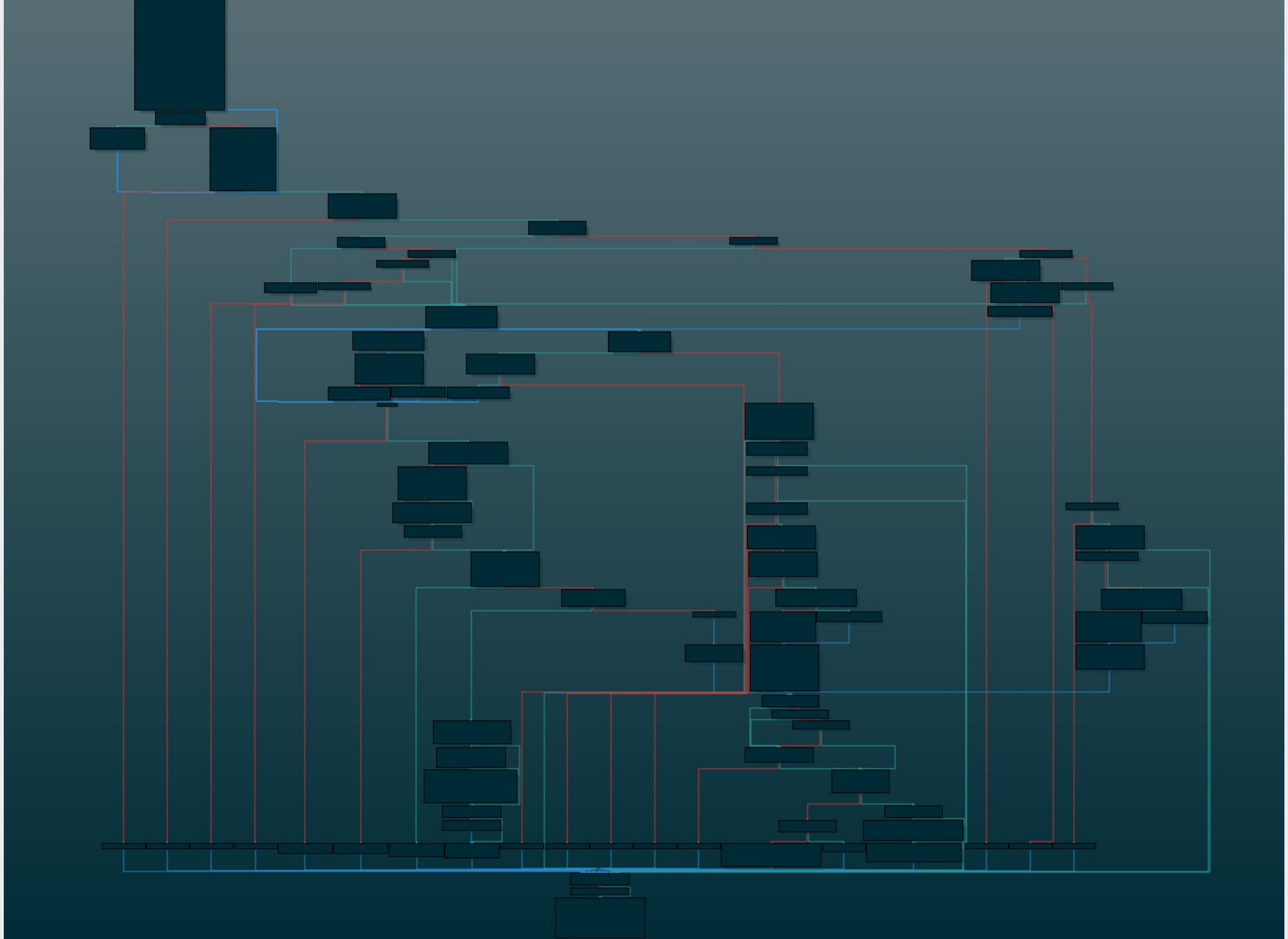
Strings not obfuscated

Network capabilities

- Pivot through firewalls
- Home-made NAT traversal
- Custom-made Proxy service
  - only available to a set of authorized IP addresses
- Remotely configured generic network sniffer
- DNS Hijacking

Lesson Learned #2

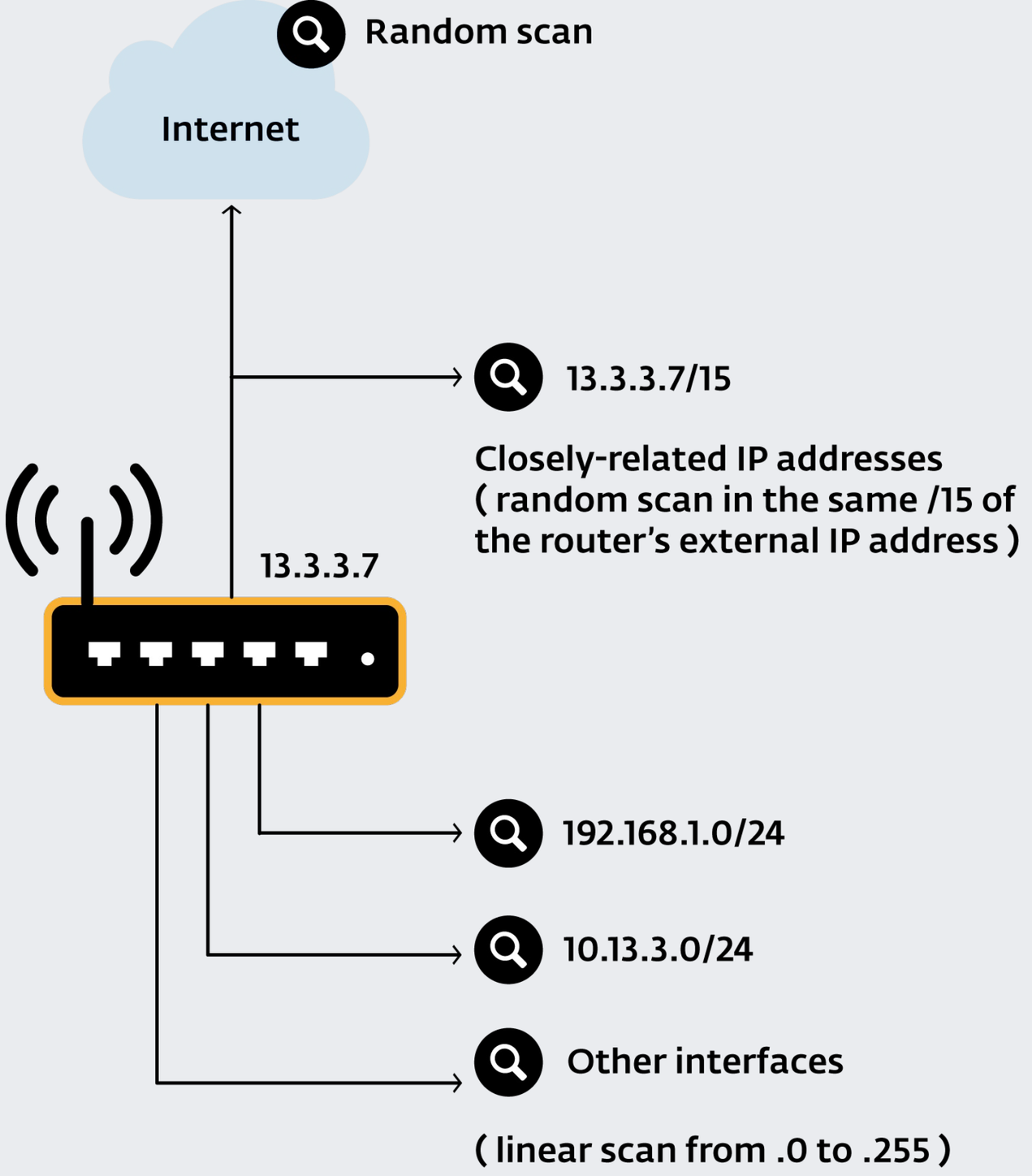
Don’t assume it’s custom when it can be a standard protocol



Proxy with access from C&C authorized IPs only

C&C IP is hardcoded

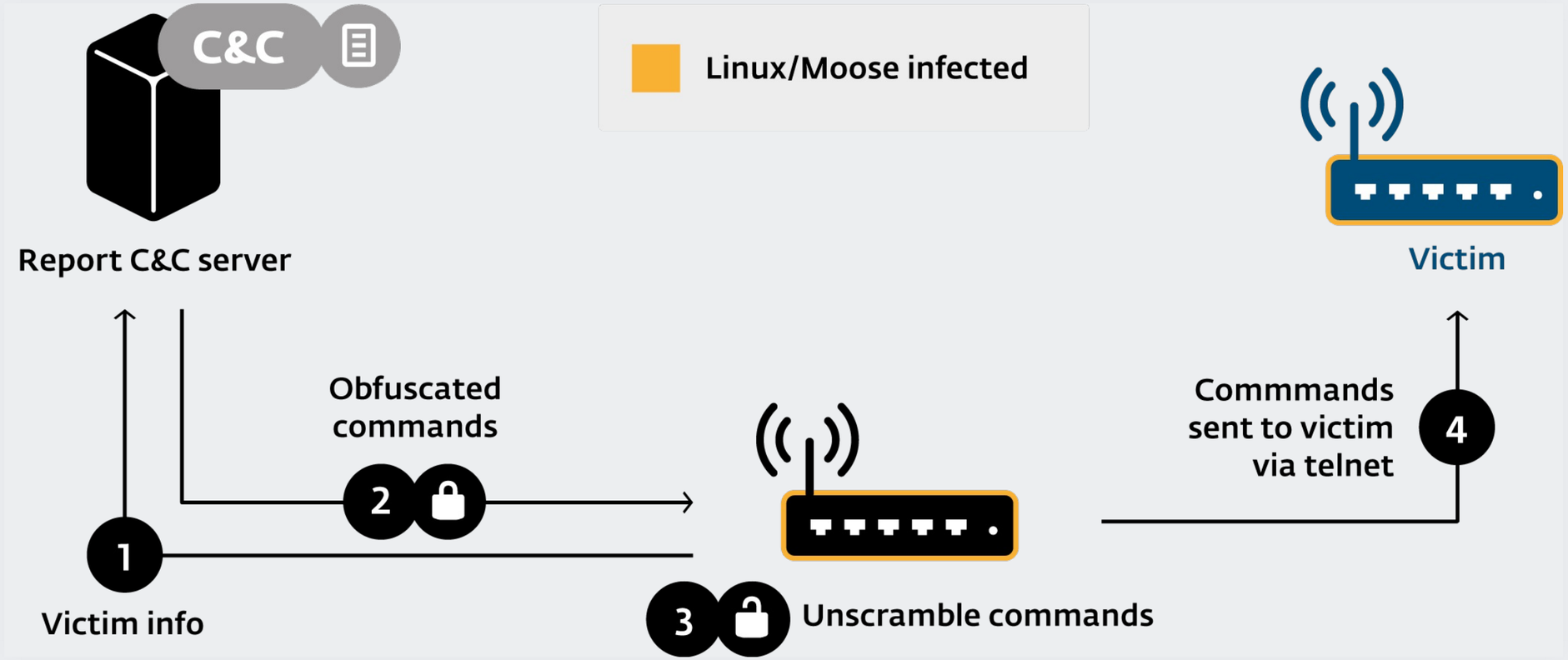
- No fallback domains or DGA



Attack Vector

- Telnet credentials bruteforce
- Wordlist of 304 user/pass entries sent by server

Compromise Protocol

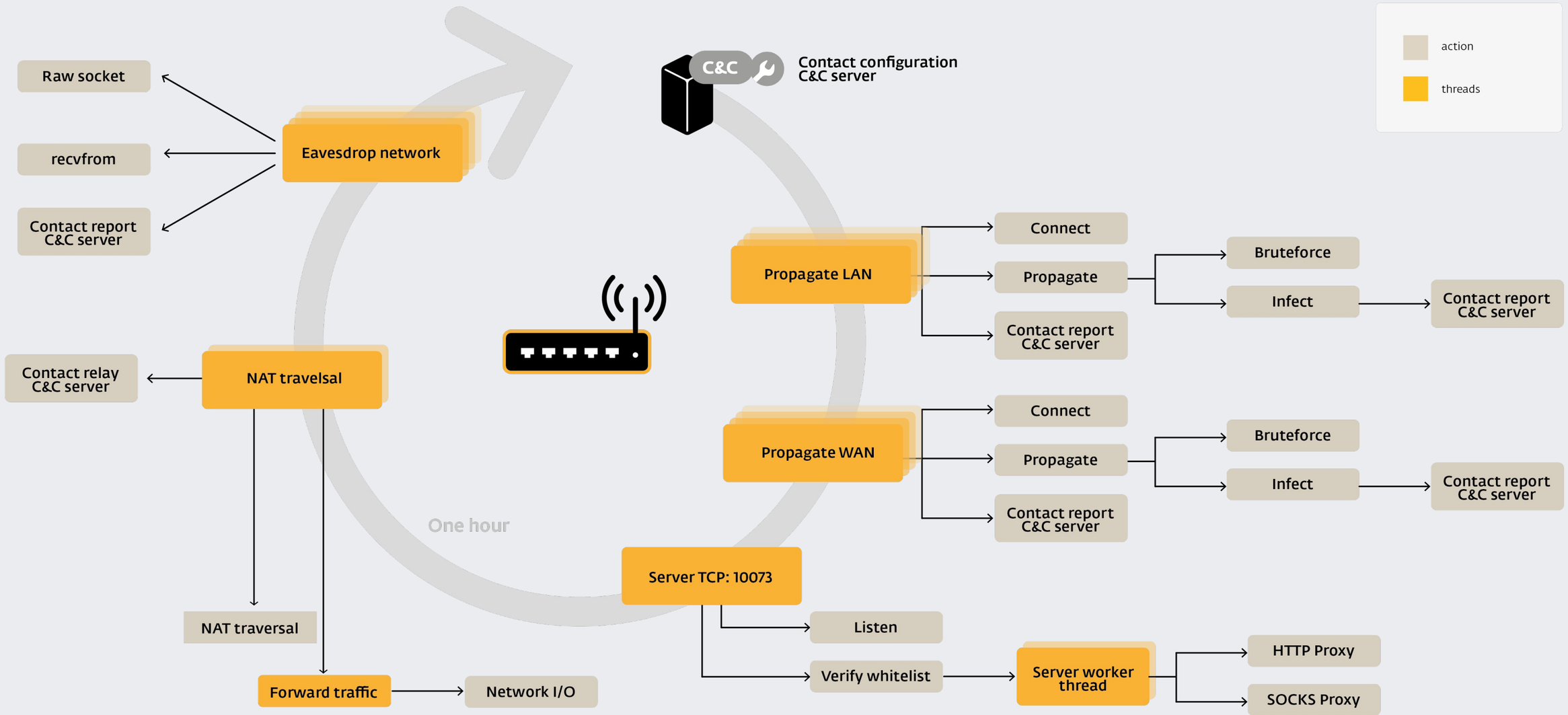


Can perform cross-arch infections

No further spreading if C&C is down

Missing: Persistence

Literally kills competition



Lesson Learned #3

Less RE, more honeypot!

Stuck

Solution

- Launch the binary in a debian MIPS qemu image
- Reachable from the Internet
- Watch it behave
- Firewall it

Hints

- Aurel images: https://people.debian.org/~aurel32/qemu/mips/
- Qemu command:

```
qemu-system-mips -M malta \
         -no-reboot -nographic \
         -kernel vmlinux-3.2.0-4-4kc-malta \
         -hda debian_wheezy_mips_standard.qcow2 \
         -append "root=/dev/sda1 console=ttyS0" \
         -redir tcp:10073::10073 -redir tcp:22::22 -redir tcp:23::23
```

Lesson #3

- We were too careful
- Everything we learned operationally was because of infected host

Hard to track malware

Moose Herding

The Malware Operation

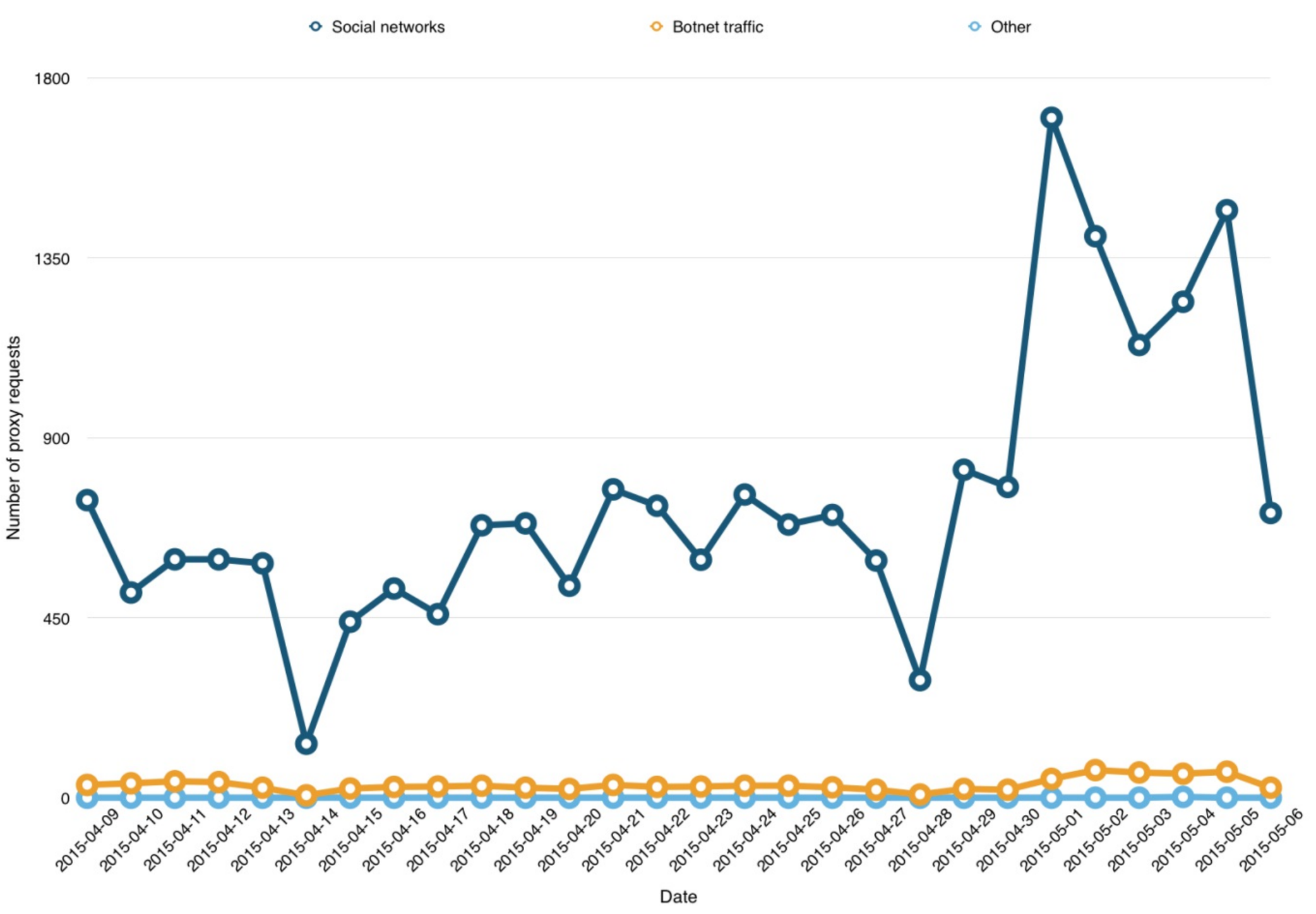
Via C&C Configuration

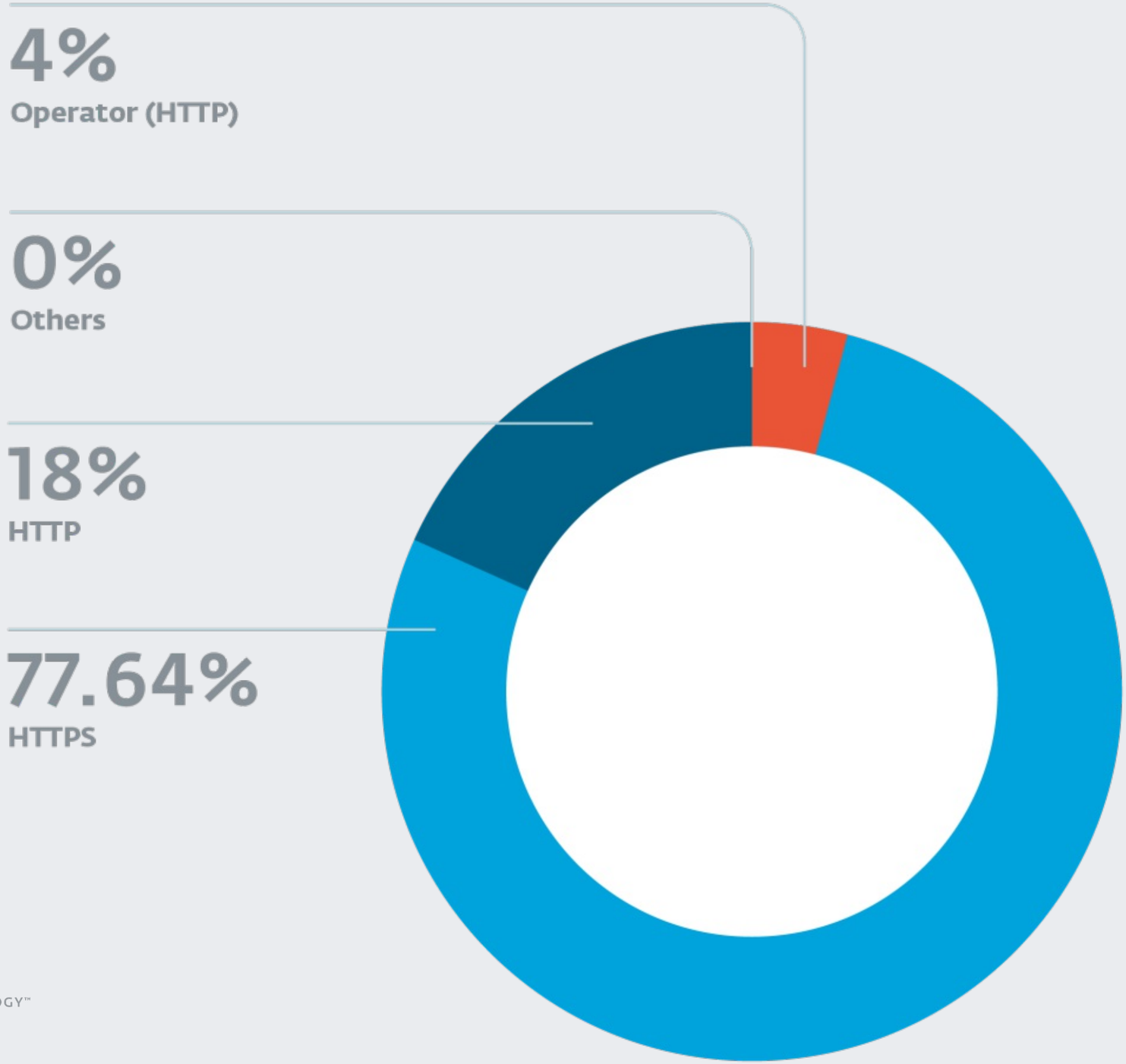
- Network sniffer was used to steal HTTP Cookies
  - Twitter: `twll`, `twid`
  - Facebook: `c_user`
  - Instagram: `ds_user_id`
  - Google: `SAPISID`, `APISID`
  - Google Play / Android: `LAY_ACTIVE_ACCOUNT`
  - Youtube: `LOGIN_INFO`

Sniffing HTTPS Cookies

Via Proxy Usage Analysis

- Nature of traffic
- Protocol
- Targeted social networks


Social networks
Botnet traffic
Other
1800
1350
900
450
0
Number of proxy requests
2015-04-09
2015-04-10
2015-04-11
2015-04-12
2015-04-13
2015-04-14
2015-04-15
2015-04-16
2015-04-17
2015-04-18
2015-04-19
2015-04-20
2015-04-21
2015-04-22
2015-04-23
2015-04-24
2015-04-25
2015-04-26
2015-04-27
2015-04-28
2015-04-29
2015-04-30
2015-05-01
2015-05-02
2015-05-03
2015-05-04
2015-05-05
2015-05-06
Date

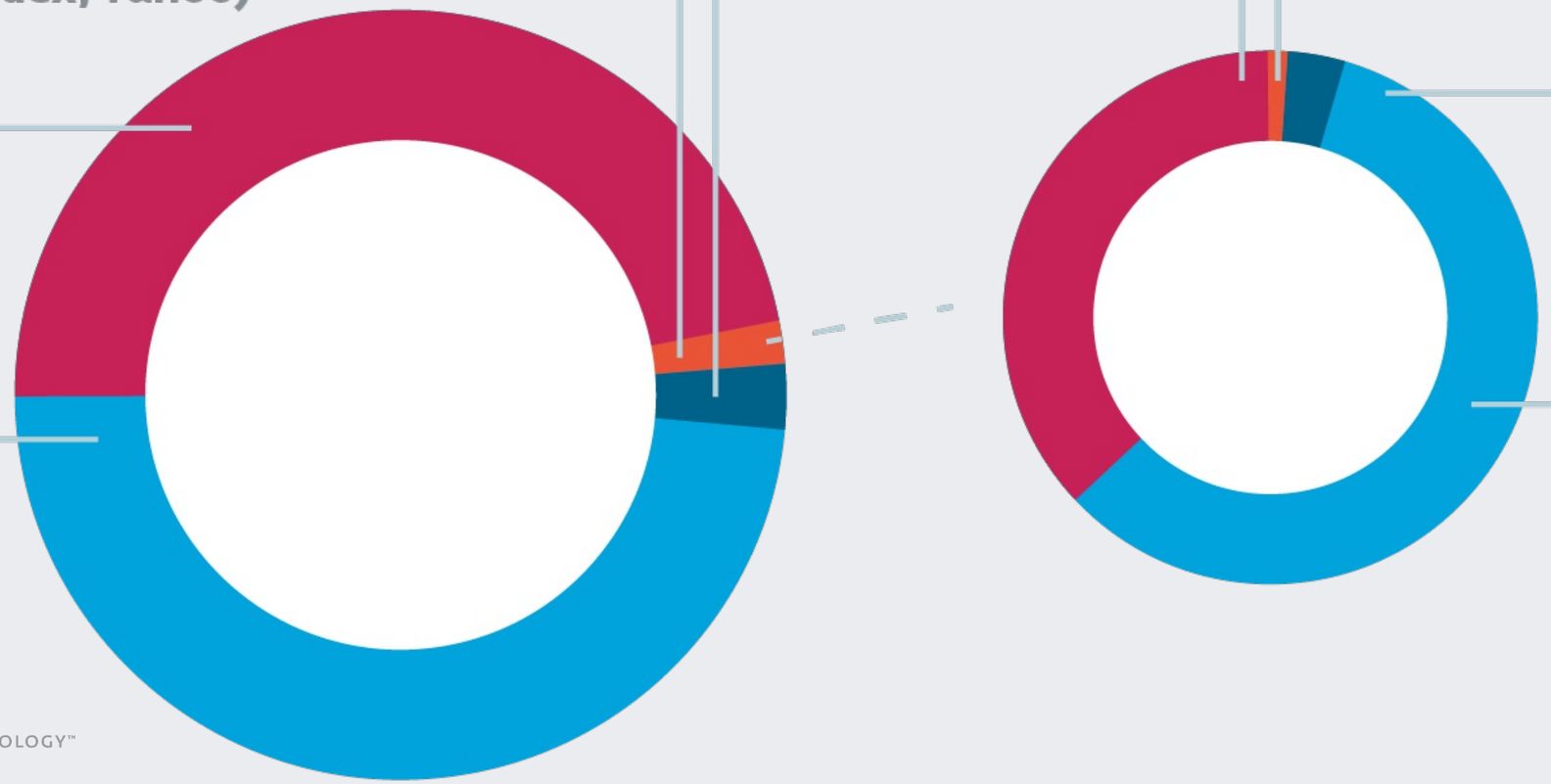

4%
Operator (HTTP)
0%
Others
18%
HTTP
77.64%
HTTPS

2%
Soundcloud

3%
Others (Youtube, Yandex, Yahoo)

47%
Instagram

49%
Twitter / Vine



59%
Yandex

4%
Yahoo

1%
Amazon Cloud

37%
Youtube

ESET ENJOY SAFER TECHNOLOGY™

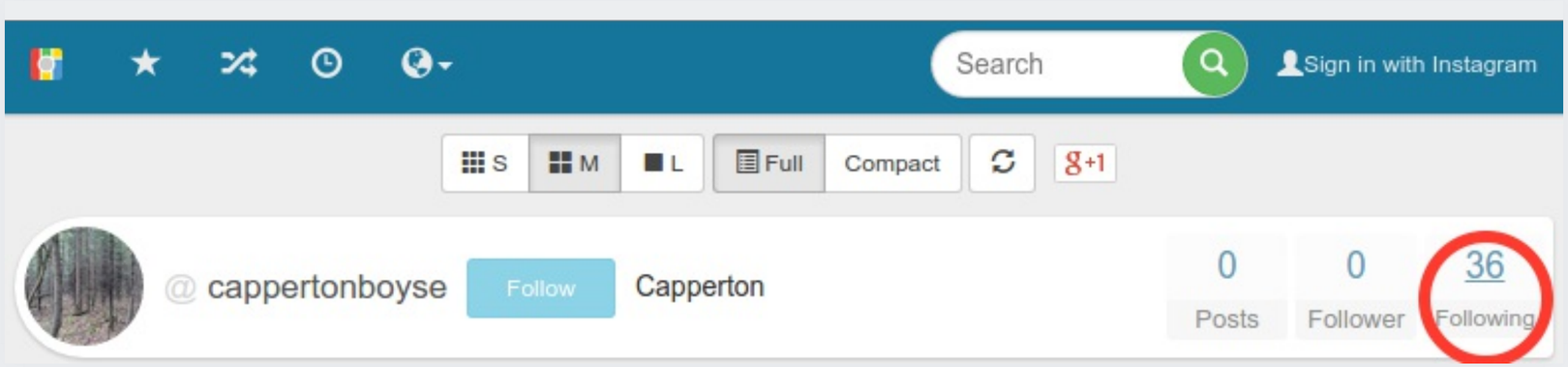
75%+ HTTPS but…

Follow TCP Stream (tcp.stream eq 10698)

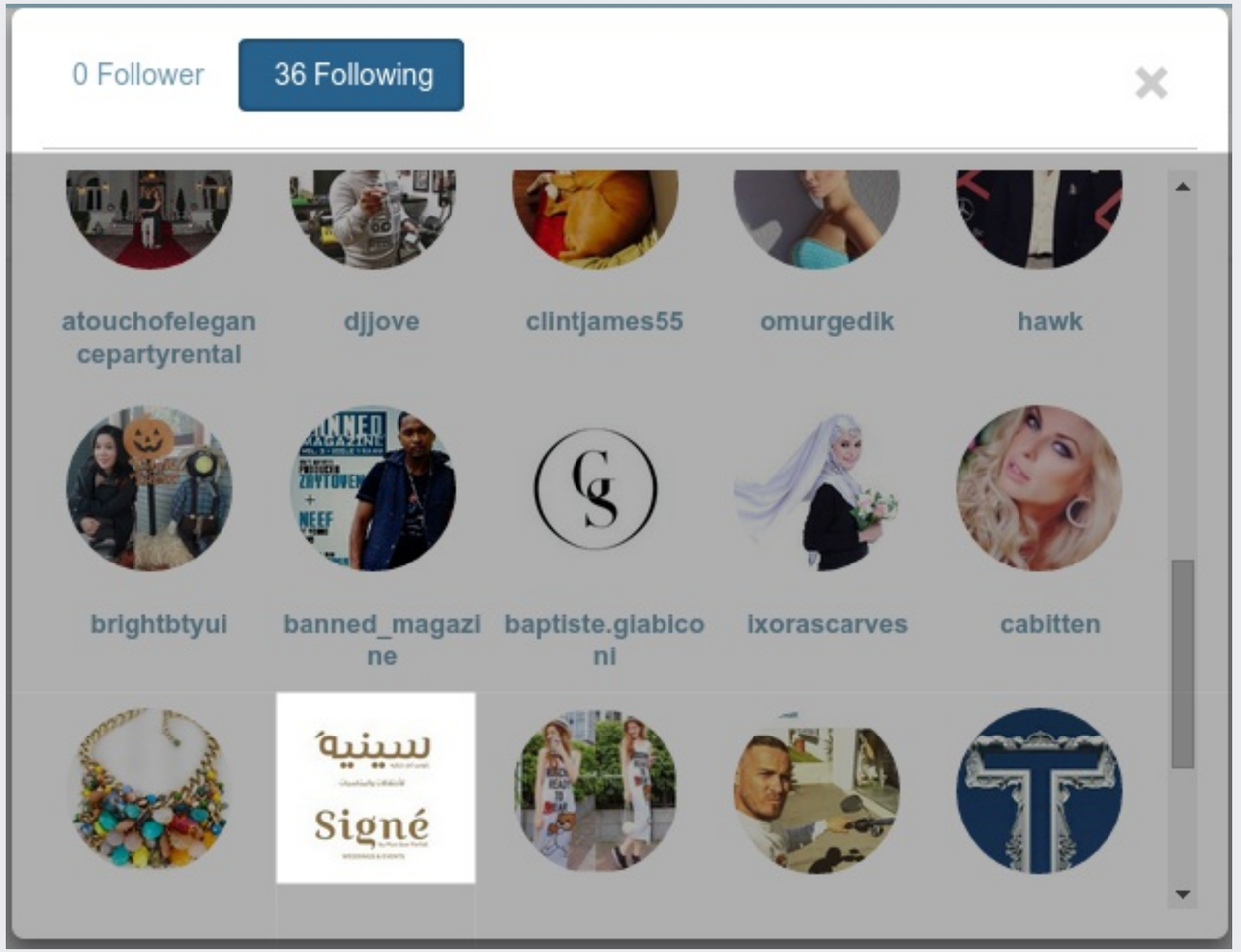
Stream Content (1)

```
...P6.X.778swan5e..Z.P6.X.GET /hookahleague HTTP/1.1
Host: instagram.com
User-Agent: Mozilla/5.0 (Windows NT 5.1; rv:17.0) Gecko/20100101 Firefox/17.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate
Connection: keep-alive

HTTP/1.1 301 Moved Permanently
Content-Type: text/html
Date:
Location: https://instagram.com/hookahleague/ (2)
Server: nginx
Content-Length: 178
Connection: keep-alive

<html>
<head><title>301 Moved Permanently</title></head>
<body bgcolor="white">
<center><h1>301 Moved Permanently</h1></center>
<hr><center>nginx</center>
</body>
```

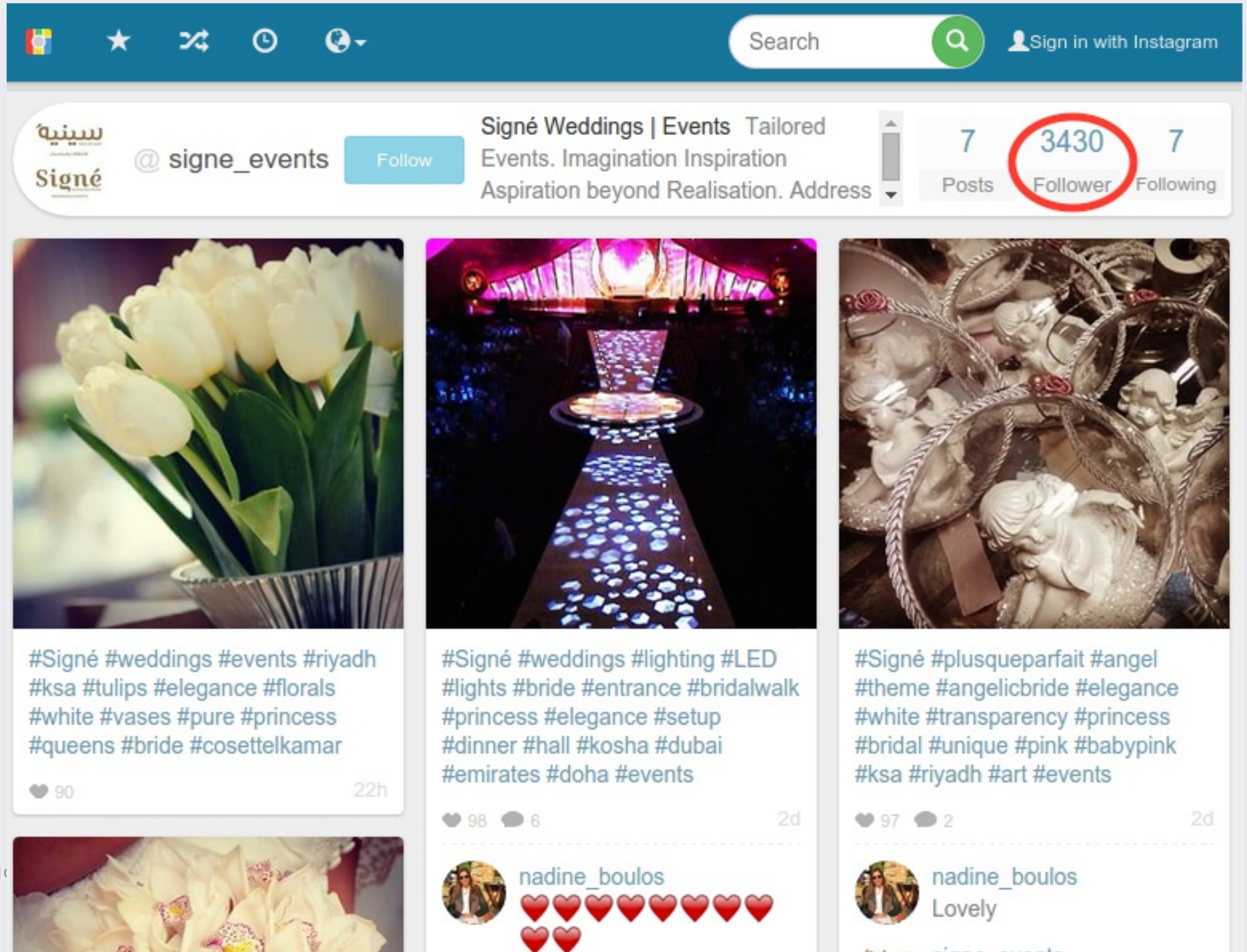
An Example


Search
Sign in with Instagram
S
M
L
Full
Compact
g+1
@ cappertonboyse
Follow
Capperton
0
Posts
0
Follower
36
Following

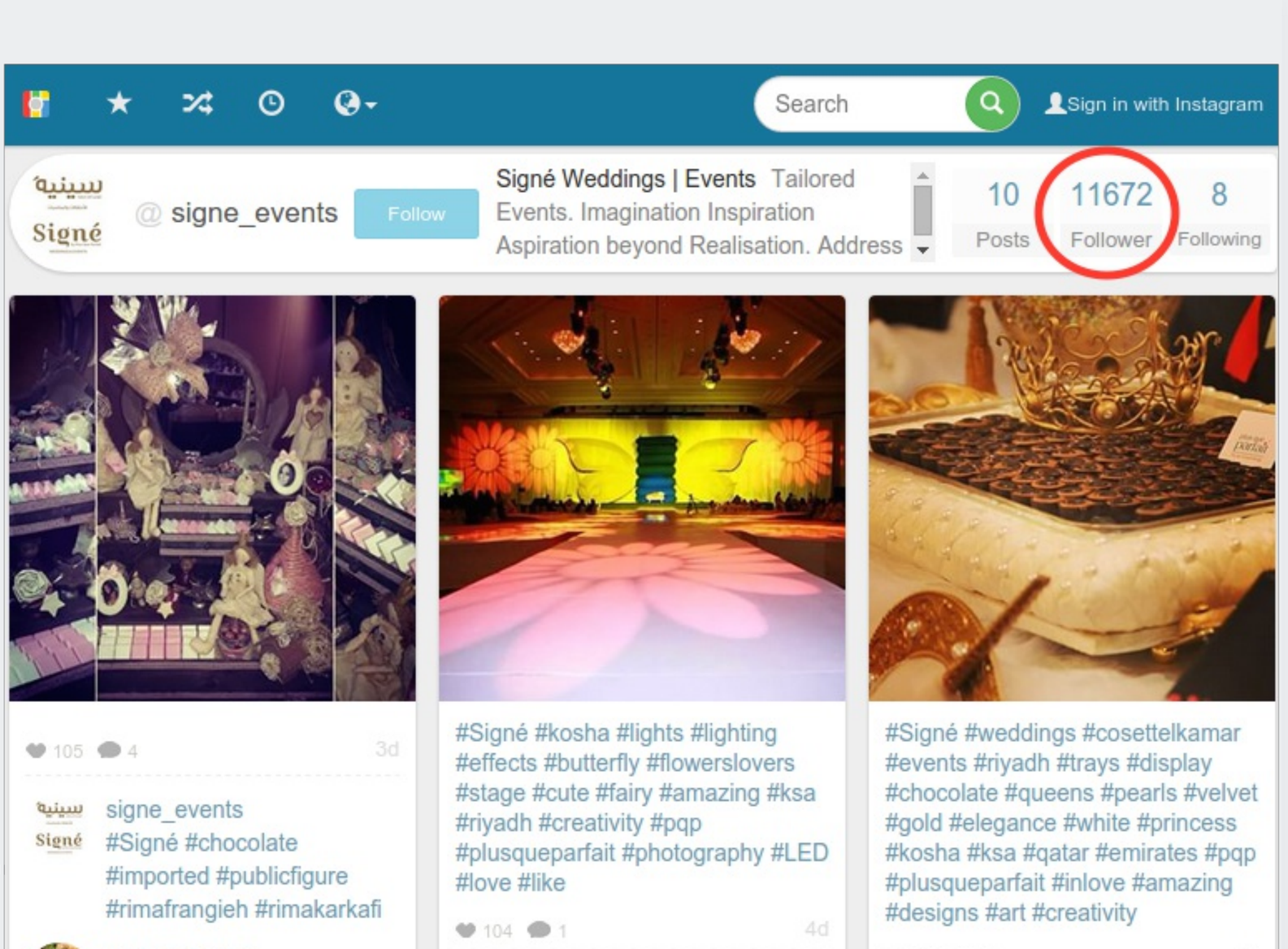
An Example (cont.)



An Example (cont.)

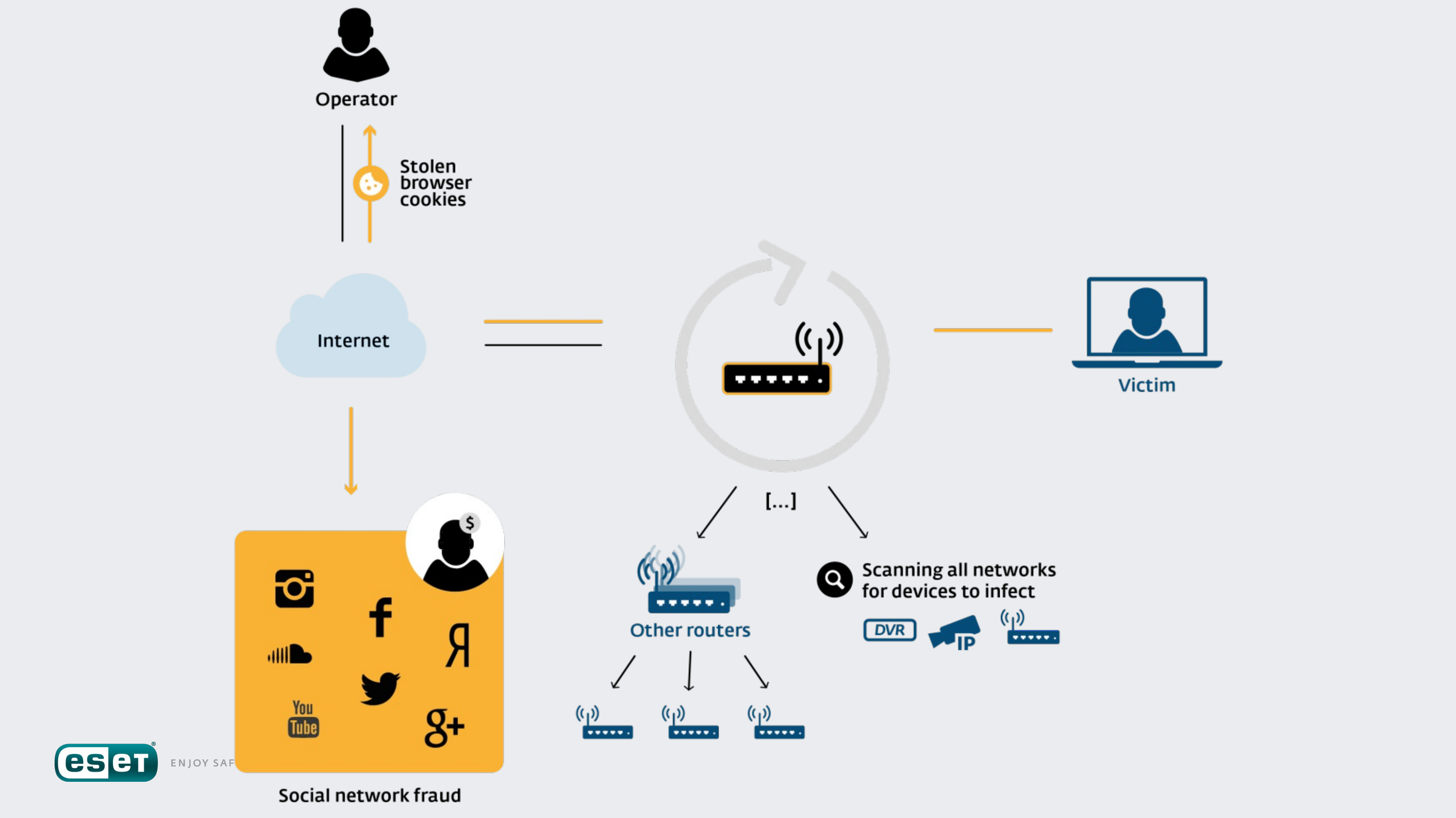


An Example (cont.)



Fraud hidden in HTTPS

Except Instagram first hit

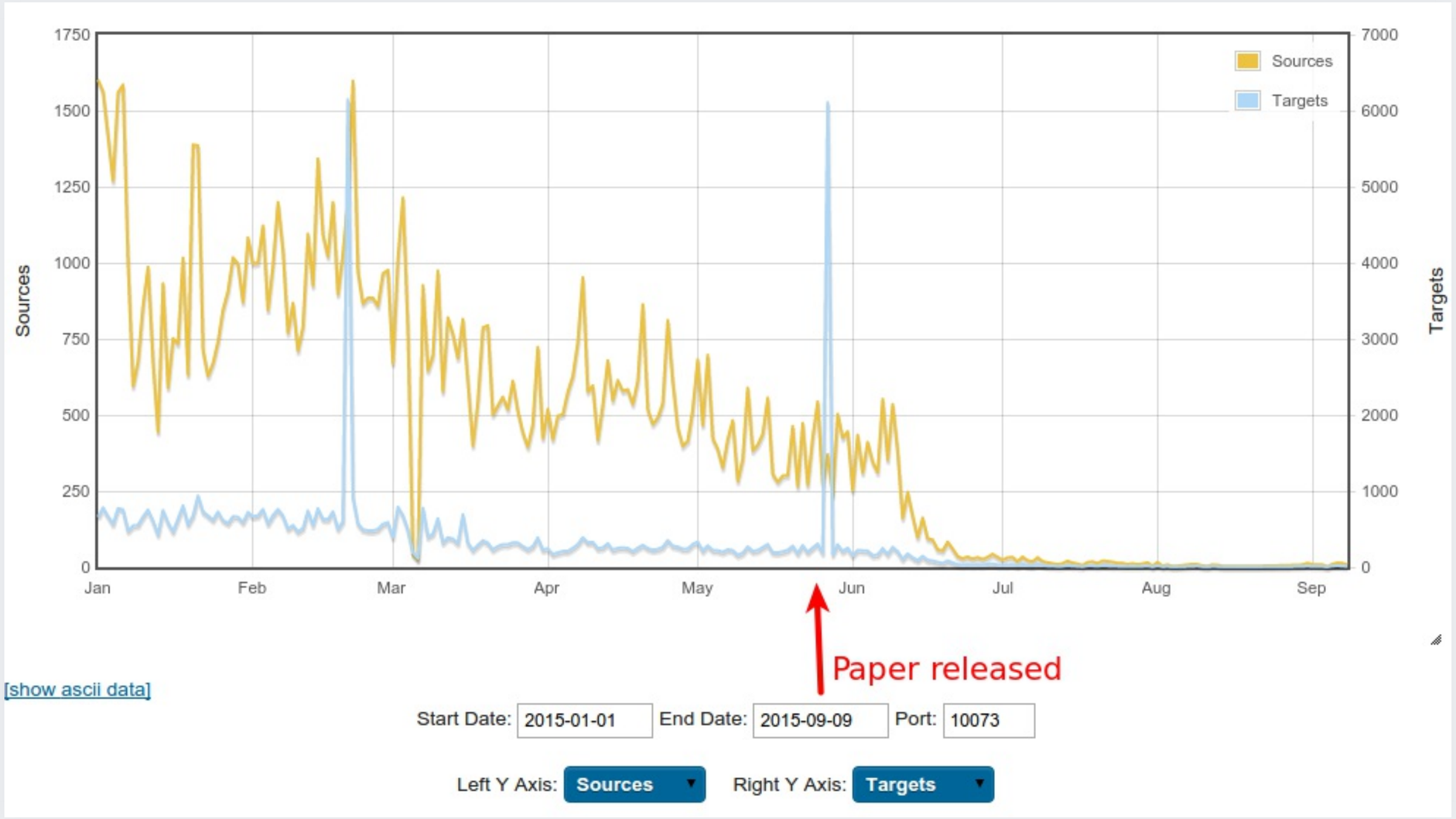

Operator
Stolen browser cookies
Internet
Victim
[...]
Other routers
Scanning all networks for devices to infect
DVR
IP
Social network fraud

Latest Developments

Whitepaper Impact

- Few weeks after the publication the C&C servers went dark
  - After a reboot, all affected devices should be cleaned
  - But victims compromised via weak credentials, so they can always reinfect

Alive or dead?



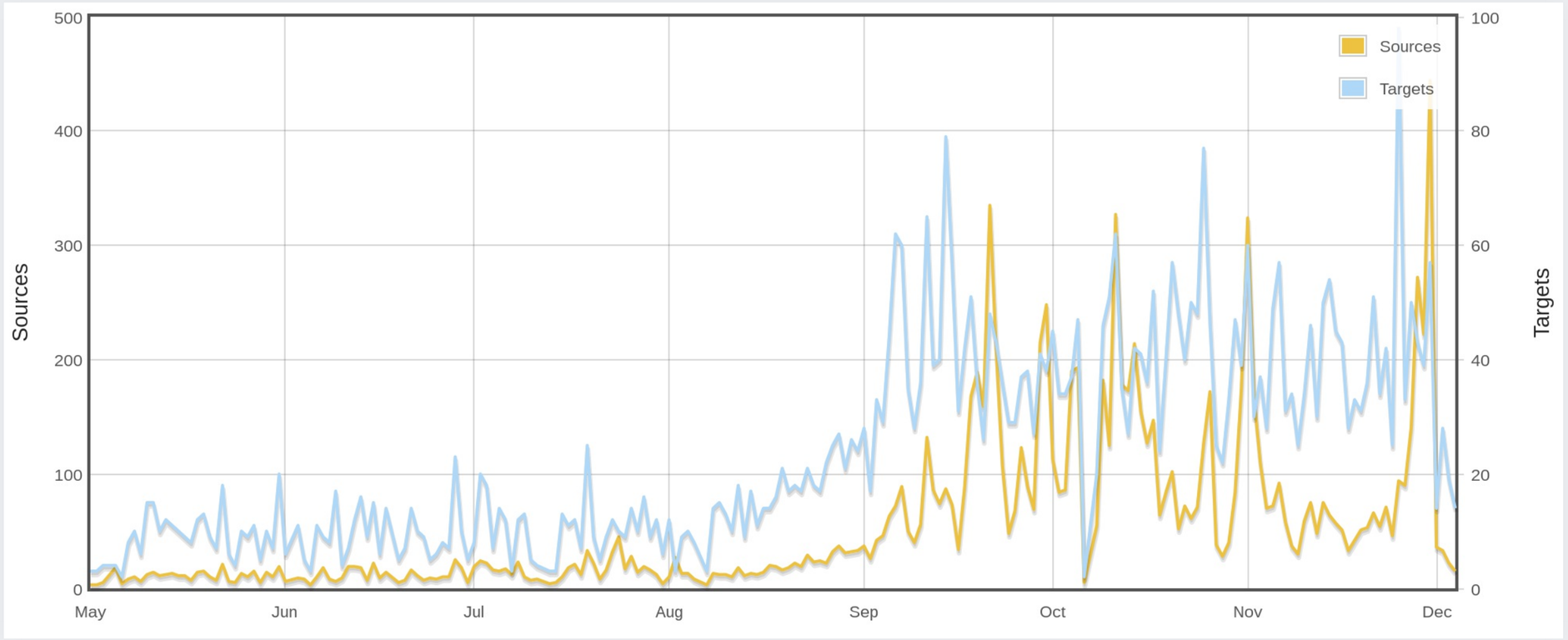
Alive or dead? (cont.)

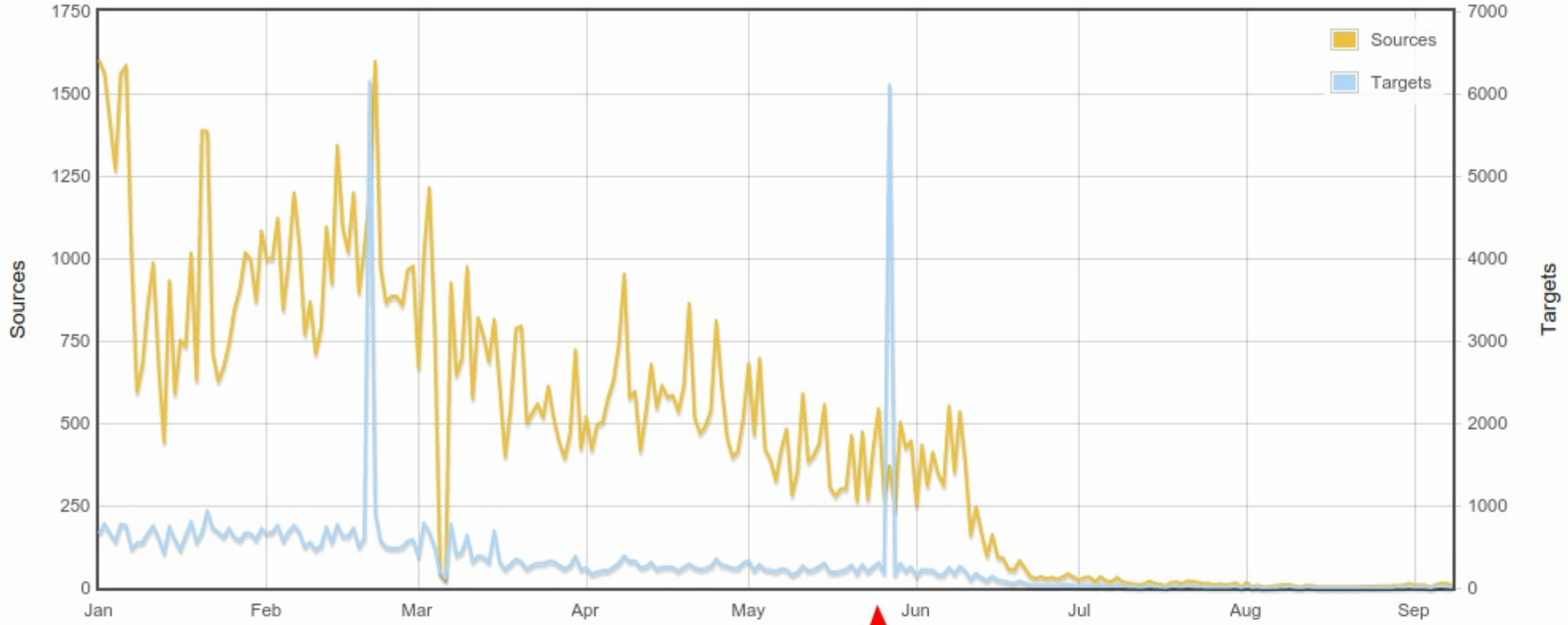
- On the lookout for Moose v2
- Looked at over 150 new samples targeting embedded Linux platforms

Found Update

- New proxy service port (20012)
- C&C selection on CLI
- C&C server returns 404 on unknown bots
- Still under analysis
- Still trying to get infected

Reading research papers and adapting


Sources
Targets
500
400
300
200
100
0
100
80
60
40
20
0
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

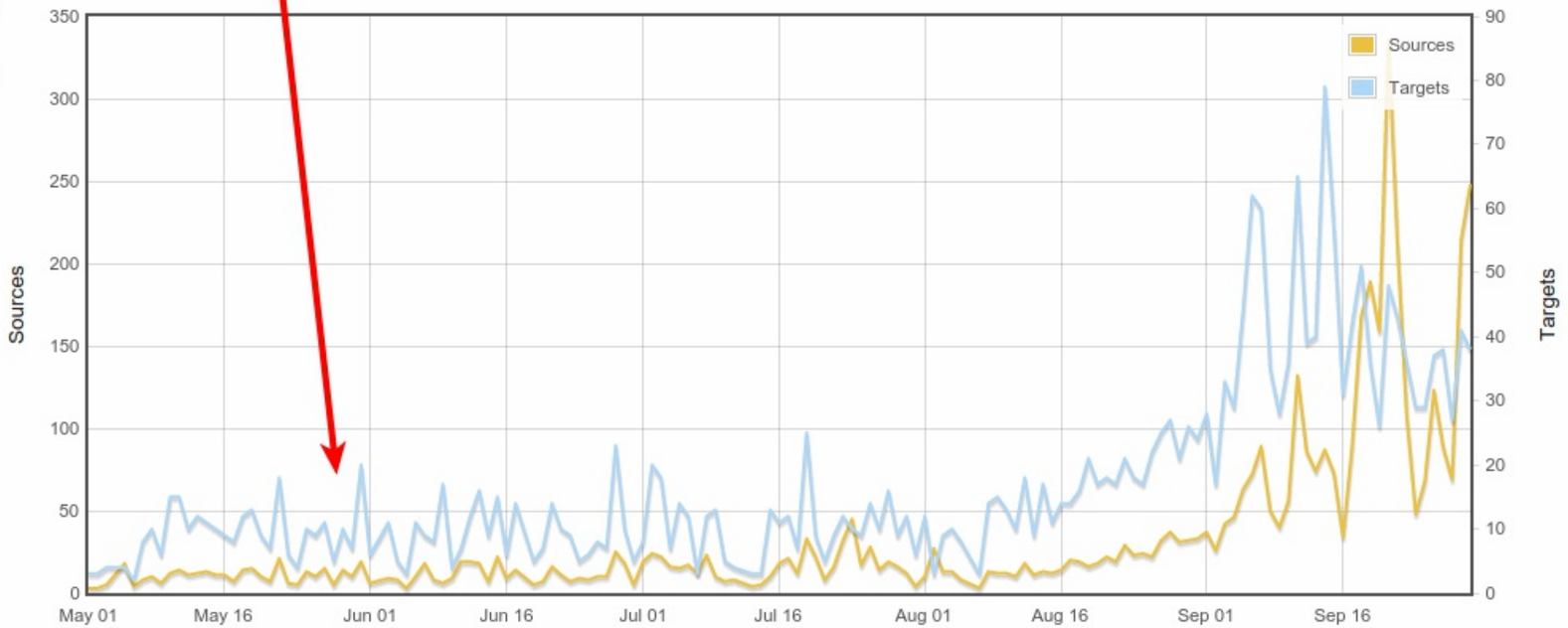


[show ascii data]

Start Date: 201

Left Y Axis:

Paper released



Take Aways

Research artifacts released

- Python and Shell Scripts
  - Protocol dissectors, fake servers, tshark wrappers
- Yara rules
- IOCs
- https://github.com/eset/malware-research/tree/master/moose

Embedded malware

- Not yet complex
- Tools and processes need to catch up
- a low hanging fruit
- Prevention simple


Questions?

Questions?

Thank you!

- @obilodeau
- and special thanks to Thomas Dupuy (@nyx__o)