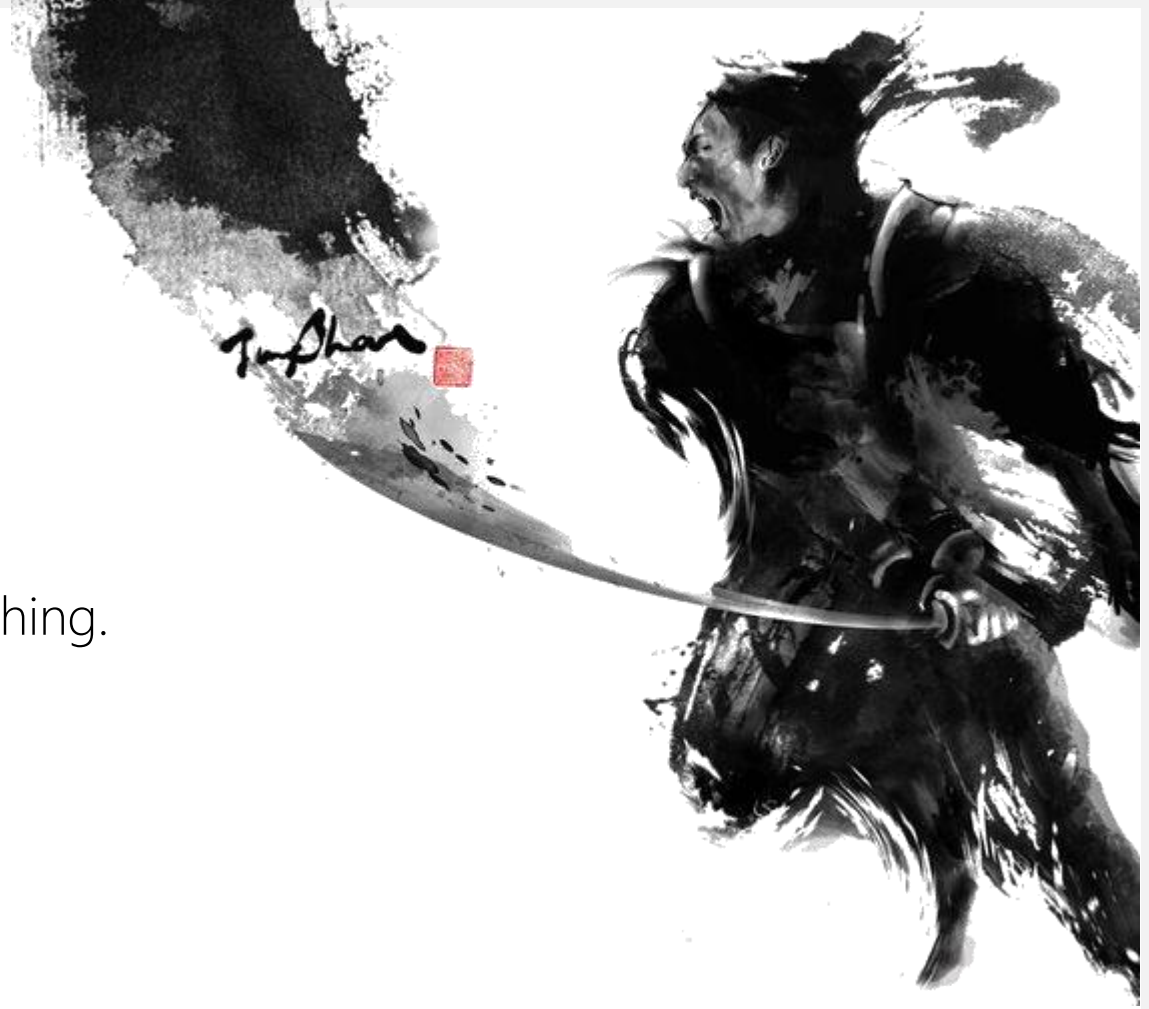


KNIGHTCRAWLER

Finding watering holes for fun and nothing.

Félix Aimé (@felixaime)
BotConf 2017 (Montpellier, FR)

Me?

Technical IT Security and Geopolitics enthusiast, with love.

Threat intelligence researcher at Kaspersky Lab. (GReAT)

Ex. French cyber defense agency (ANSSI), British Telecom

First time disclosing one of my personal projects.

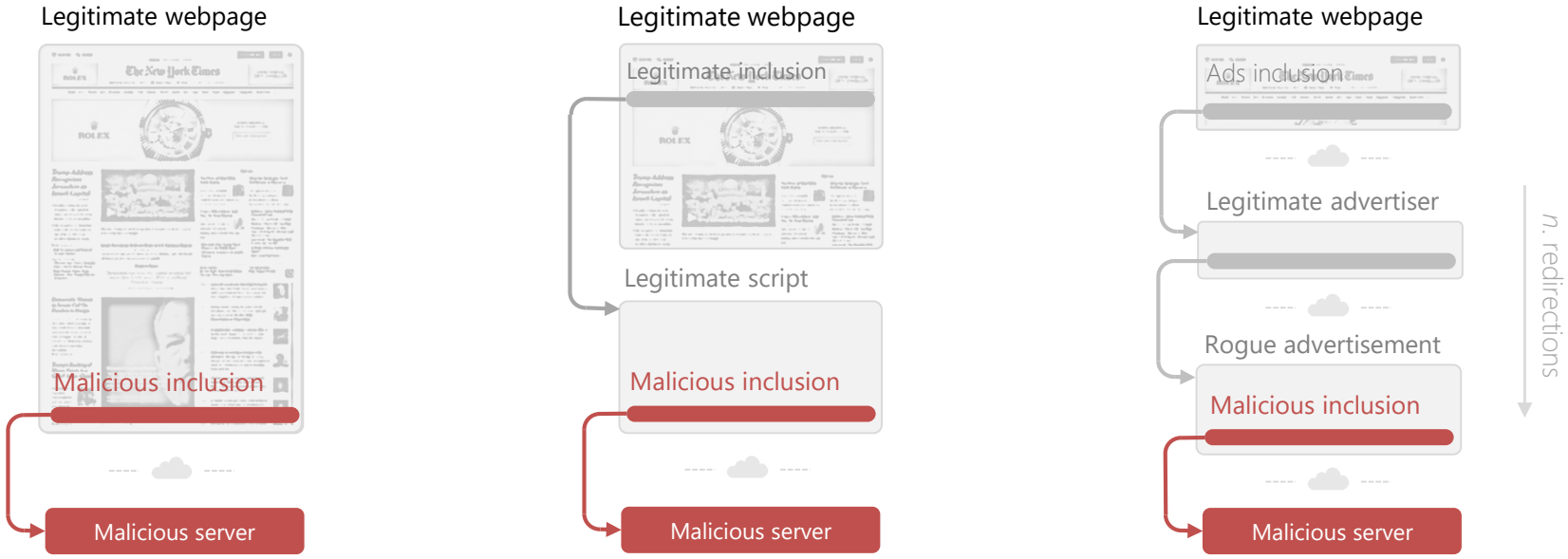
KnightCrawler?

Project started in 2016 to get my own “Threat Intel”.

Finding watering holes (aka. SWC) in an automated way.

Watering hole (noun.):

Insertion of a malicious script on a specific website to infect its visitors.


Legitimate webpage
Malicious inclusion
Malicious server
Legitimate webpage
Legitimate inclusion
Legitimate script
Malicious inclusion
Malicious server
Legitimate webpage
Ads inclusion
Legitimate advertiser
Rogue advertisement
Malicious inclusion
Malicious server
n. redirections

On the malicious server?

Possible IP range whitelisting (mostly done by APT Threat actors)

Browser fingerprinting (Plugins, local IP, Accept-language etc.)

Leads to some exploits, fake installers, browser plugins etc.

How to detect watering holes?

Focusing on good targets.

For targeted attacks:

Govs
Pharma
Defense
Embassies
Aerospace
Energy
NGOs
Media
Institutes
Nuclear
Banks
Investment
Human rights
Tech compagnies
International Orgs.
Jihadists websites
Conferences

For cybercrime stuff:

Porn & streaming
Online stores
Old Wordpress, Joomla ;)

How to get the targets?

~~Passive DNS~~

Common Crawl Indexes

Directories scraping

Leaked DBs

Manual insertion

X509 Subject Alternative Name

Subdomains enumeration

How to detect watering holes?

Focusing on heuristics.

Monitor changes (ex. First time seeing that remote host)

Use of dynamic DNS / IP Address by the remote host

Remote host domain name created less than 90 days ago

Free SSL certificate used by the remote host

Mixed HTTP content, content-type not following the file extension etc.

Whitelist the trackers, ads etc.

How to detect watering holes?

YARA everything!

Write YARA on different stuff such as:

- HTTP response headers
- Body content (HTML, JS, SWF etc.)
- Whois records
- SSL Certificates
- Paths
- Hosts

And enjoy the results!

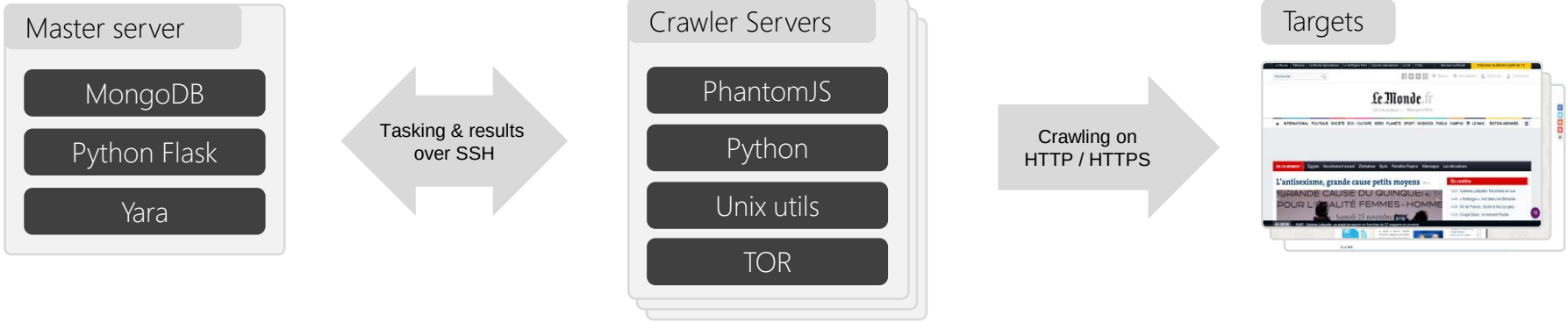
```
rule ObfuscatedScanboxURLs {
    strings:
        $s1 = /\?[a-zA-Z]{3,10}\_[a-zA-Z]{3,10}==[0-9]{1,2}$/
        $s2 = /\?seed=(.*)&alivetime=(.*)&r=(.*)$/
    condition:
        any of them
}
```

```
rule OceanOctopusCookieValue {
    strings:
        $p1 = "__ac0e4"
    condition:
        $p1
}
```

```
rule ROPComments {
    strings:
        $s1 = "VirtualAlloc" nocase
        $s2 = /xchg(\ ){0,}([a-z\.]){3}([,\_\ \.]){0,}esp/ nocase
        $s3 = /pop(\ ){0,}([a-z\.]){3}([,\_\ \.]){0,}ret/ nocase
        $s4 = "ole32_base" nocase
        $s5 = "shell_addr" nocase
        $s6 = "nop sled" nocase
    condition:
        any of them
}
```

How to detect watering holes?

Creating your own (legal) botnet.



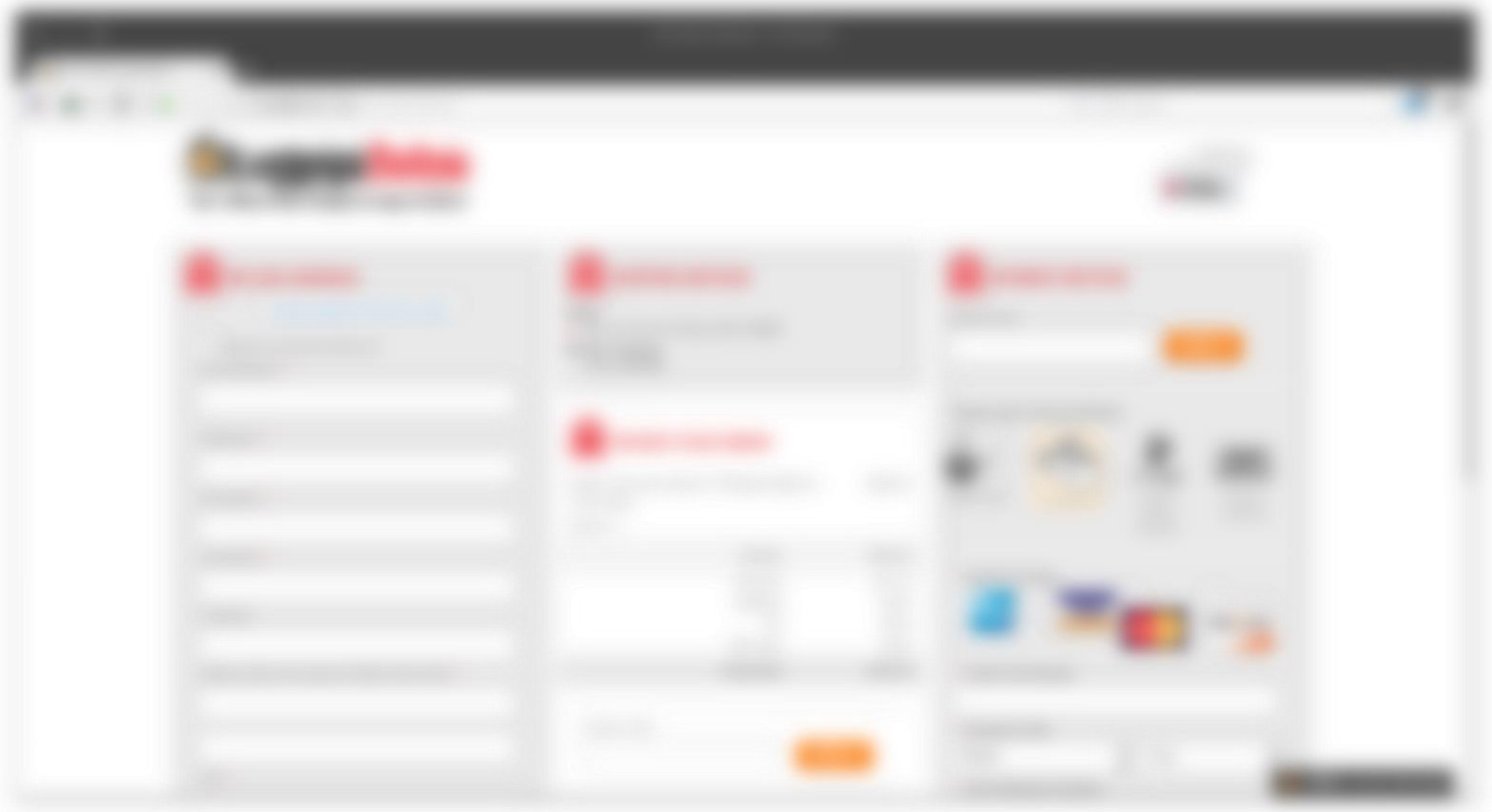
Random target selection in queue
Crawlers deployment on the fly
with volatile IPs

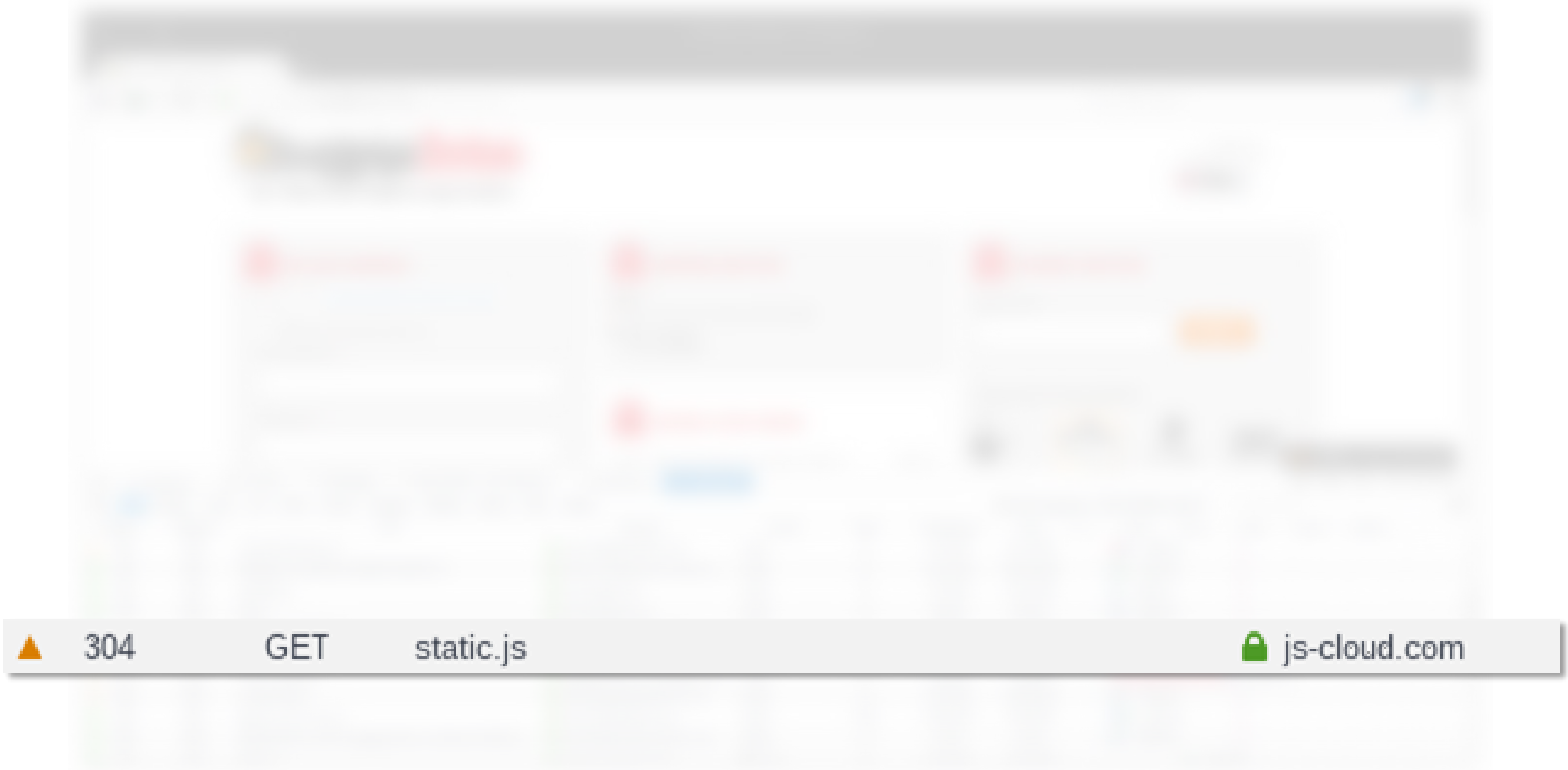
~40 User agents
~20 Accept Language
Local links following
Human interactions
Chrome headless
Authent.

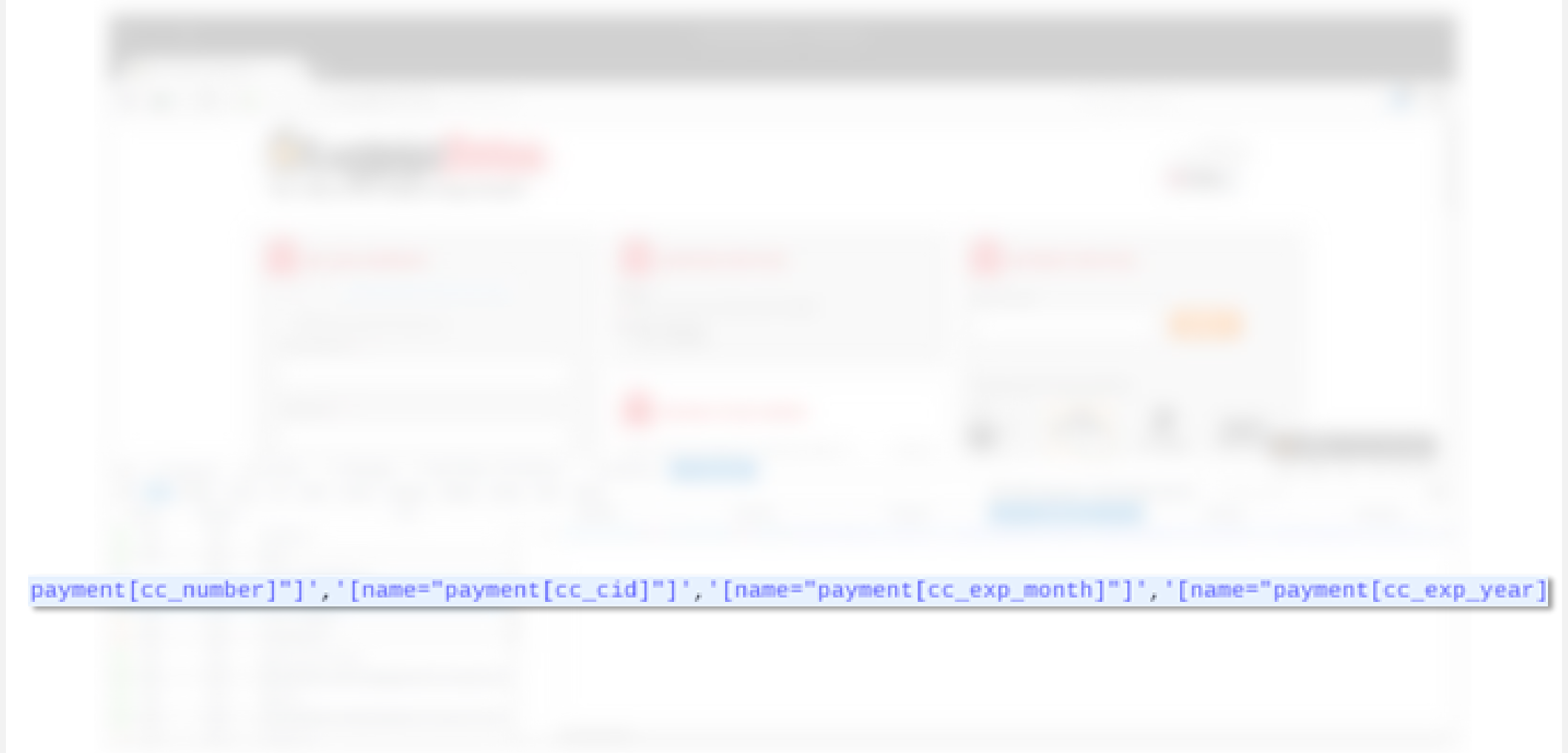
25K specific targets
~100 .onion rdvs

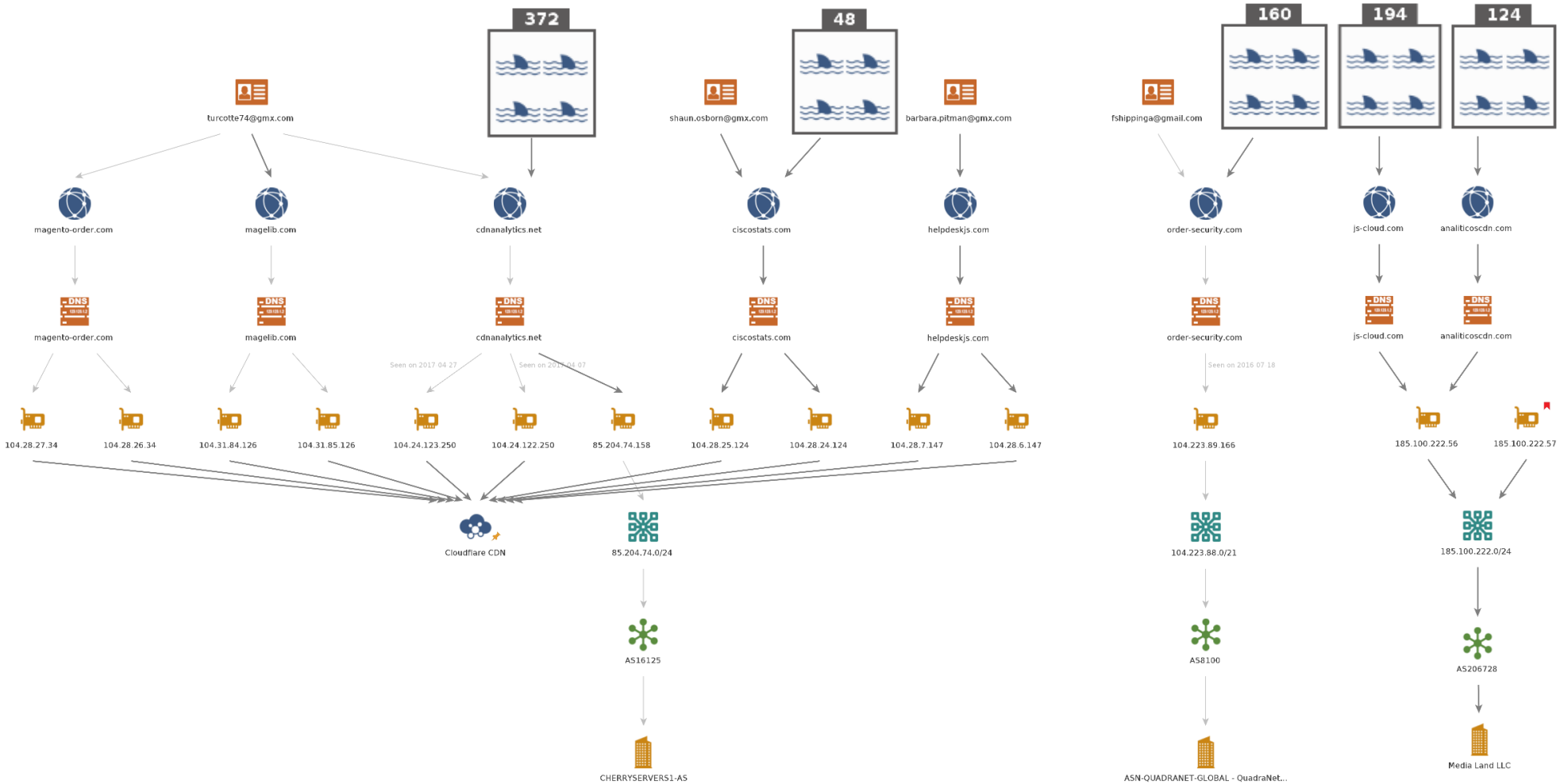
Cybercrime?

Credit card stealers campaigns.




304 GET static.js js-cloud.com


payment[cc_number]"]','[name="payment[cc_cid]"]','[name="payment[cc_exp_month]"]','[name="payment[cc_exp_year]


372
48
160
194
124
turcotte74@gmx.com
shaun.osborn@gmx.com
barbara.pitman@gmx.com
fshippinga@gmail.com
magento-order.com
magelib.com
cdnanalytics.net
ciscostats.com
helpdeskjs.com
order-security.com
js-cloud.com
analiticoscdn.com
DNS
Seen on 2017 04 27
Seen on 2017 04 07
Seen on 2016 07 18
104.28.27.34
104.28.26.34
104.31.84.126
104.31.85.126
104.24.123.250
104.24.122.250
85.204.74.158
104.28.25.124
104.28.24.124
104.28.7.147
104.28.6.147
104.223.89.166
185.100.222.56
185.100.222.57
Cloudflare CDN
85.204.74.0/24
104.223.88.0/21
185.100.222.0/24
AS16125
AS8100
AS206728
CHERRYSERVERS1-AS
ASN-QUADRANET-GLOBAL - QuadraNet...
Media Land LLC

Other cybercrime stuff:

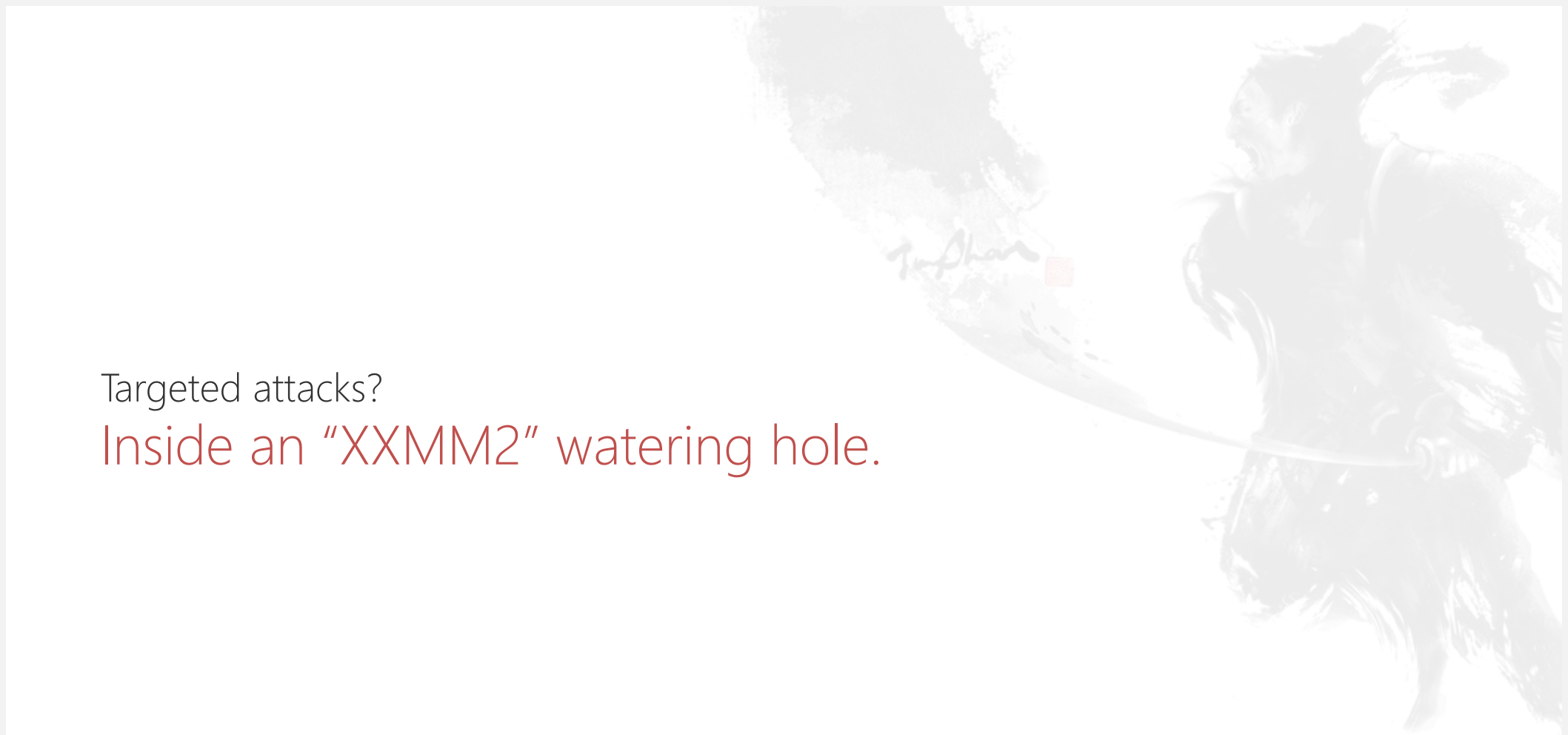
Exploit kits in 2016, mainly

Tech Support Scams

Malicious porn redirection

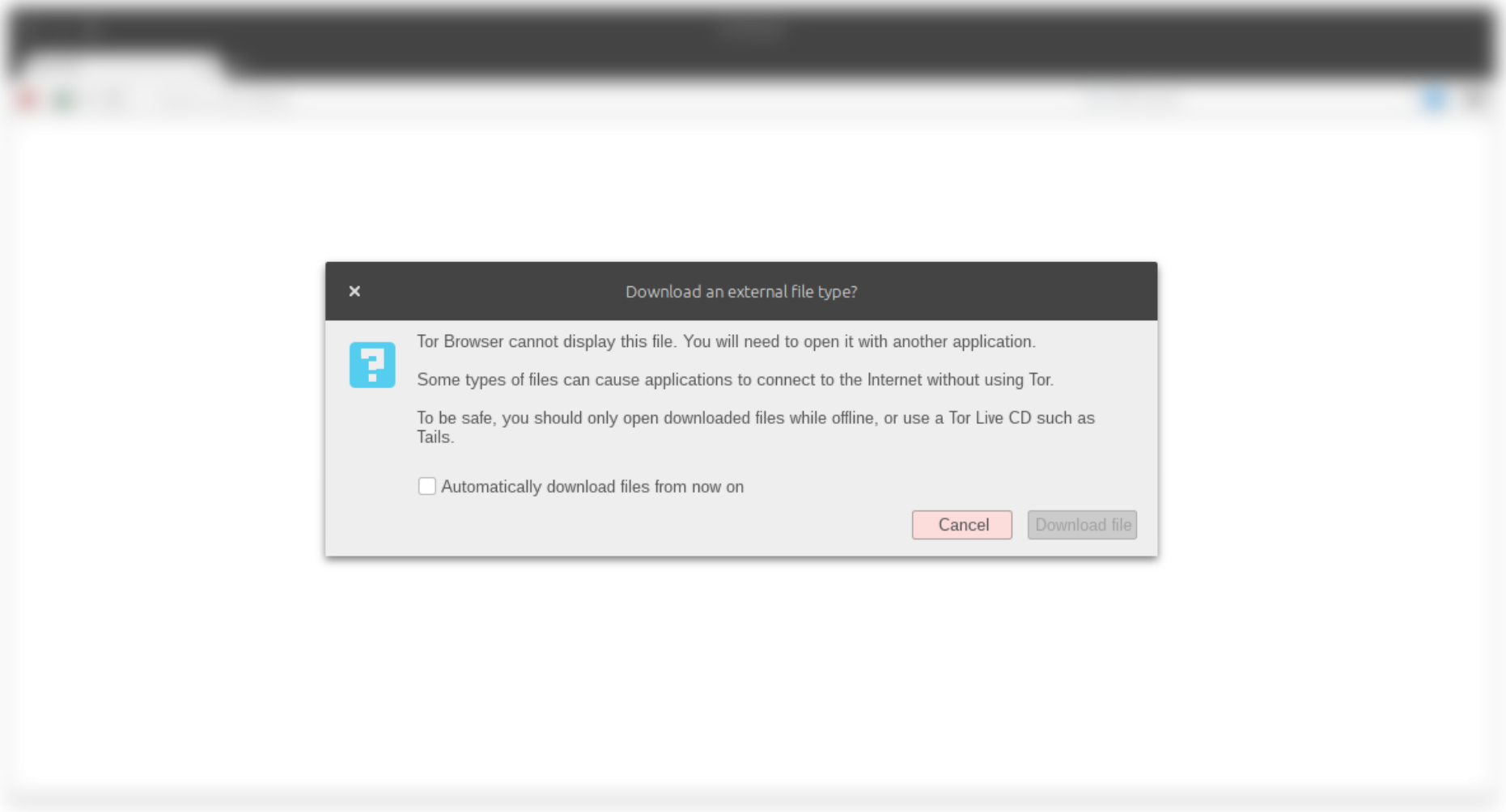
Cryto currency mining

Unattributed stuff (still investigating)



Targeted attacks?

Inside an “XXMM2” watering hole.


Download an external file type?
Tor Browser cannot display this file. You will need to open it with another application.
Some types of files can cause applications to connect to the Internet without using Tor.
To be safe, you should only open downloaded files while offline, or use a Tor Live CD such as Tails.
Automatically download files from now on
Cancel
Download file

Sample #1: 7b92fa06b7bed2bde84e93a9360c87b9 (C2: 116.193.153[.]134)

Sample #2: 29cc4b97e82efd48da3aec4b18a2ec09 (C2: 61.97.250[.]87)

User-Agent: Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; SV1)

Other targeted attack stuff:

[redacted] (Gov. and international orgs. targets)

[redacted] (Political, companies, gov. and media targets)

[redacted] (Companies, media and gov. targets)

[redacted] (Gov. targets)

[redacted] (Gov. target)

[redacted] (Media and gov. targets)

[redacted] (Ukrainian targets)

and ~10 unattributed / undisclosed campaigns.

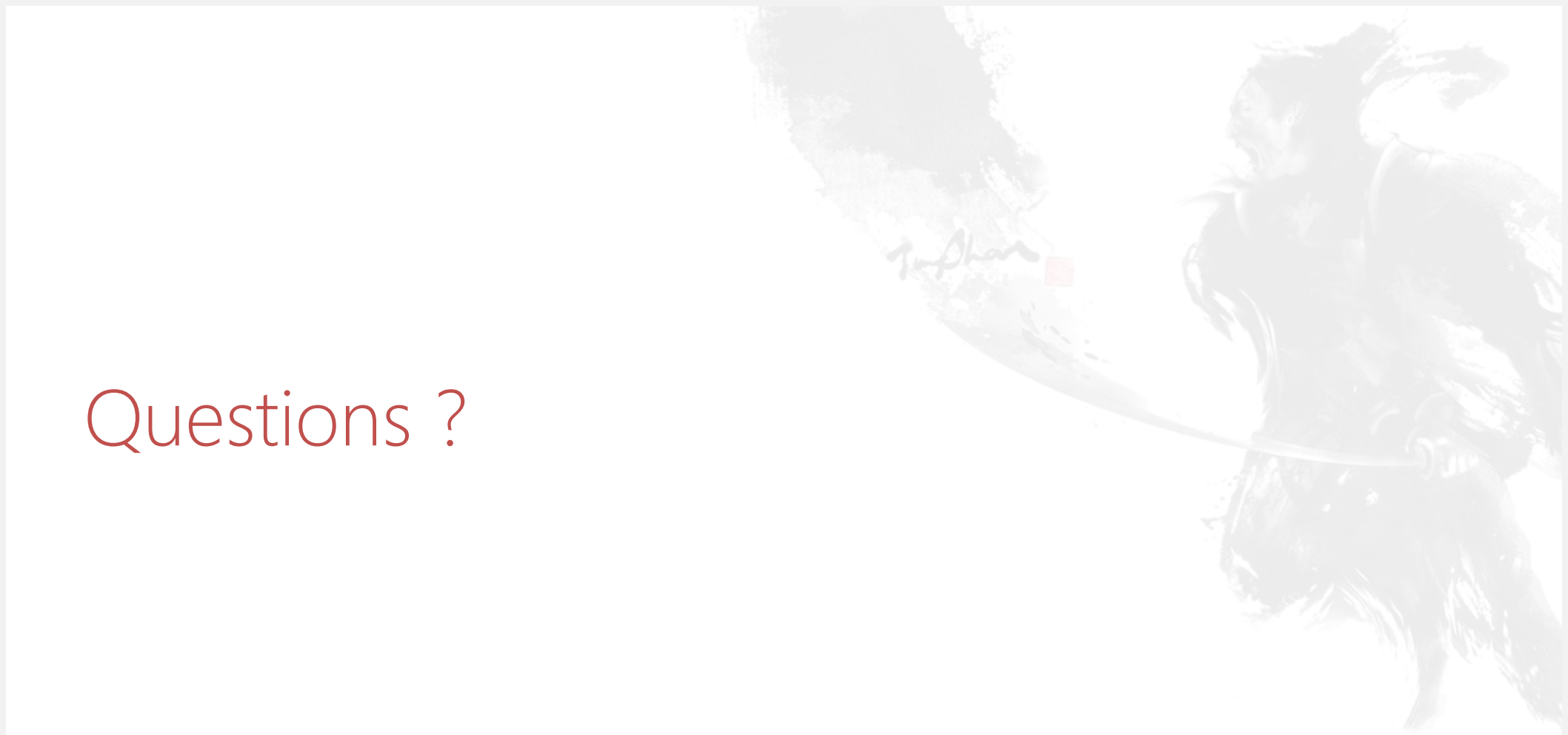
Future work?

Continue to dev it (Volunteers ?).

Implement new heuristics as modules?

Share it with a restricted and trusted community?

Questions ?



Félix Aimé (@felixaime)