Messing up with Kids playground: Eradicating easy targets

Yarochkin Fyodor @fygrave
Vladimir Kropotov @vbkropotov

Presented at HITBKL 2012

agenda

Introduction (cybecrime 2012 – russian style :)
Detecting malicious network infrastructure
Getting one-step-ahead
Conclusions

DCCrime-2012: Brief Introduction

- Bots and Botnets – still popular :)
- Monetization schemes vary.
- DbD is one of the most common attack vectors
  - We also have email
  - We also have stupid users downloading sh*t
  - Mobile is lucrative target (all your money are there)

DCCrime-2012: Introduction

BROWSERS	HITED	HOSTS	LOADS	% ↓
MSIE	75397	32122	960	3.00
Firefox	3976	2419	66	2.75
Opera	2166	797	10	1.27
Chrome	5622	3729	34	0.92
Safari	11374	7700	42	0.55
Mosaic	2	0	0	0
Mozilla	2012	780	0	0.00
Lynx	10	3	0	0.00

OS	HITED	HOSTS	LOADS	% ↓
Windows 2000	87	35	3	8.57
Windows XP	10286	5508	235	4.28
Windows Vista	8516	4152	165	4.00
Windows 2003	124	57	2	3.51
Windows 8	106	78	2	2.56

THREADS	HITED	HOSTS	LOADS	% ↓
-deleted-	100559	47550	1112	2.35

EXPLOITS	LOADS	% ↓
MDAC	4	100.0
PDF LIBTIFF	133	100.0
Java Pack	975	2.06

“Traffic” - is still an important component in the process :)

Main “components” to deal with

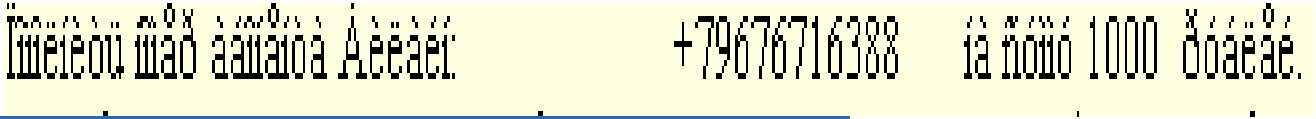
- Callback nodes (aka C&C)
- Traffic:
  - Compromised machines/or manipulated content
  - Banner networks
  - SEO (doorways)

What's new this year?

- Automated detection gets difficult. (anti-sandboxing, anti-crawler tricks)

```
function() {
  var url = 'http://yyzola.gpbbsdhmjm.shacknet.nu/g/';
    …
    document.onmousemove = function()  {
```

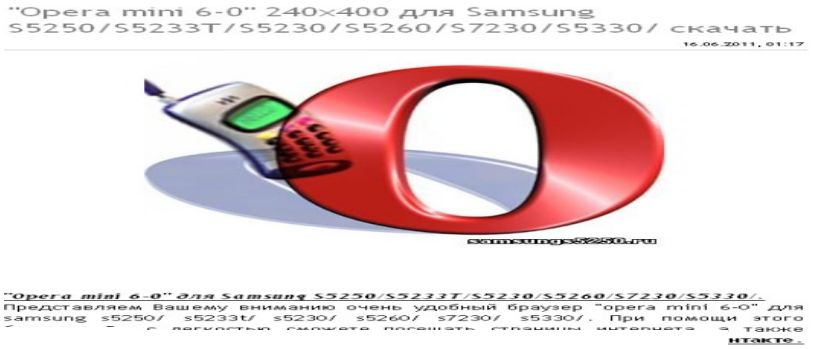
- In some cases of idiocy, human interaction is a must..
- Mobile phone as the most common means of funds transfer



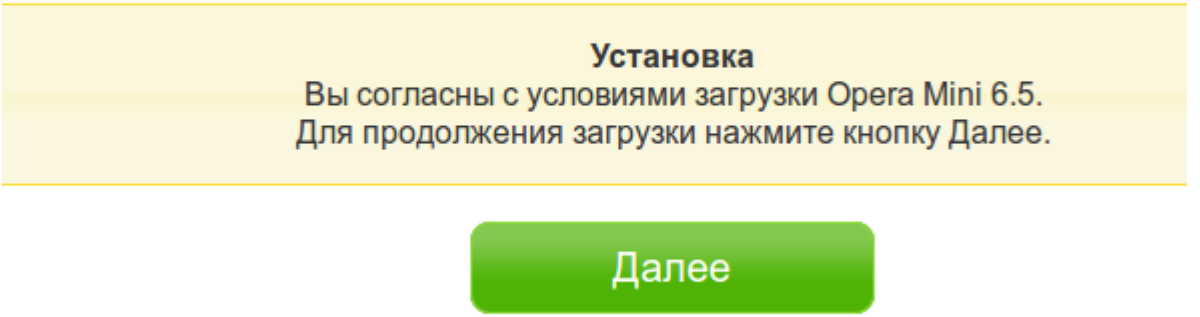
```
<addr value="http://124ffsaf.com/sms/gate.php"/>
<addr value="http://124ff42.com/sms/gate.php"/>
<addr value="http://124ffdfsaf.com/sms/gate.php"/>
<addr value="http://124sfafsaffa.com/sms/gate.php"/>
</http>
```

Mobile scams

- Fake apps are still big
- Android apps avail :)




```
jqlbejusk = "5";
String str = jqlbejusk.jqlbejusk("Xl/P.kX");
Class[] arrayOfClass = new Class[1];
arrayOfClass[0] = Class.forName(jqlbejusk.jqlbejus
Method localMethod = localClass.getMethod(str, arr
```




```
<div style="bottom:5px;position:absolute;width:1
    <a href="exit.html" style="color:#999999;fon
    <div align="center" style="color:#999999;font-size:13px;">Услуга платная, не более 354р с НДС.</div>
    <div align="center" ><a href="javascript:_oferta()" style="color:#999999;font-size:13px;">Пользовательское соглашение</a></div>
</div>
```

So really, how easy it is to get pwned
In Russia? :)

- So the focus of this research:
  - Identifying “bad kids” playground – mapping infrastructure, identifying potential targets, attempting to fix the problems, before “things hit hard”

Detecting malicious network infrastructure

DNS: (did u see this morning passive DNS talk? ;-))

With a spike of generative domain botnets, this seems like interesting research project

DGAs produce very specific pattern in DNS traffic

Is this the only method to call back?

Nop..

Alternatives...

Domain generative bots

- C&C is not hardcoded to maintain flexibility in cases when C&C is taken down.
- Some sort of algorithm is used to generate domain names
- Domains are tested for validity. IP address is obtained.
- Sometimes obfuscation involved. (for example: manipulations applied to resolved IP address)

How it looks on the wire

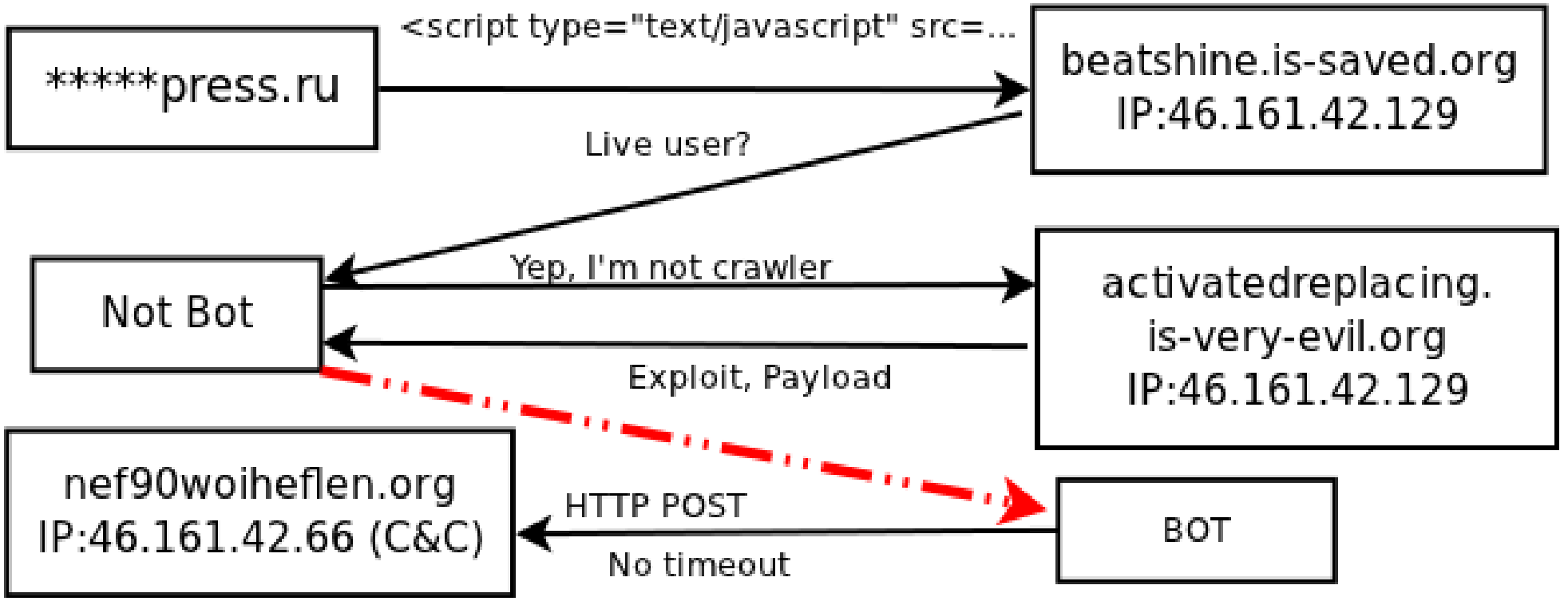
Protocol	Length	Info
DNS	161	Standard query response, No such name
NBNS	92	Name query NB GANYCYHYWEK.EU<00>
DNS	127	Standard query response, No such name
DNS	161	Standard query response, No such name
NBNS	92	Name query NB DIGEGAZOLAN.EU<00>
DNS	161	Standard query response, No such name
NBNS	92	Name query NB KEZAPYJOLEK.EU<00>
DNS	86	Standard query A jewezexigaf.eu.HomeGateway
DNS	161	Standard query response, No such name
NBNS	92	Name query NB XUKOVORUPUT.EU<00>
DNS	127	Standard query response, No such name
DNS	161	Standard query response, No such name
NBNS	92	Name query NB DISAFUWOKIS.EU<00>
DNS	127	Standard query response, No such name
DNS	86	Standard query A jenokirifux.eu.HomeGateway
DNS	127	Standard query response, No such name

C&C/generative domains and pattern mining

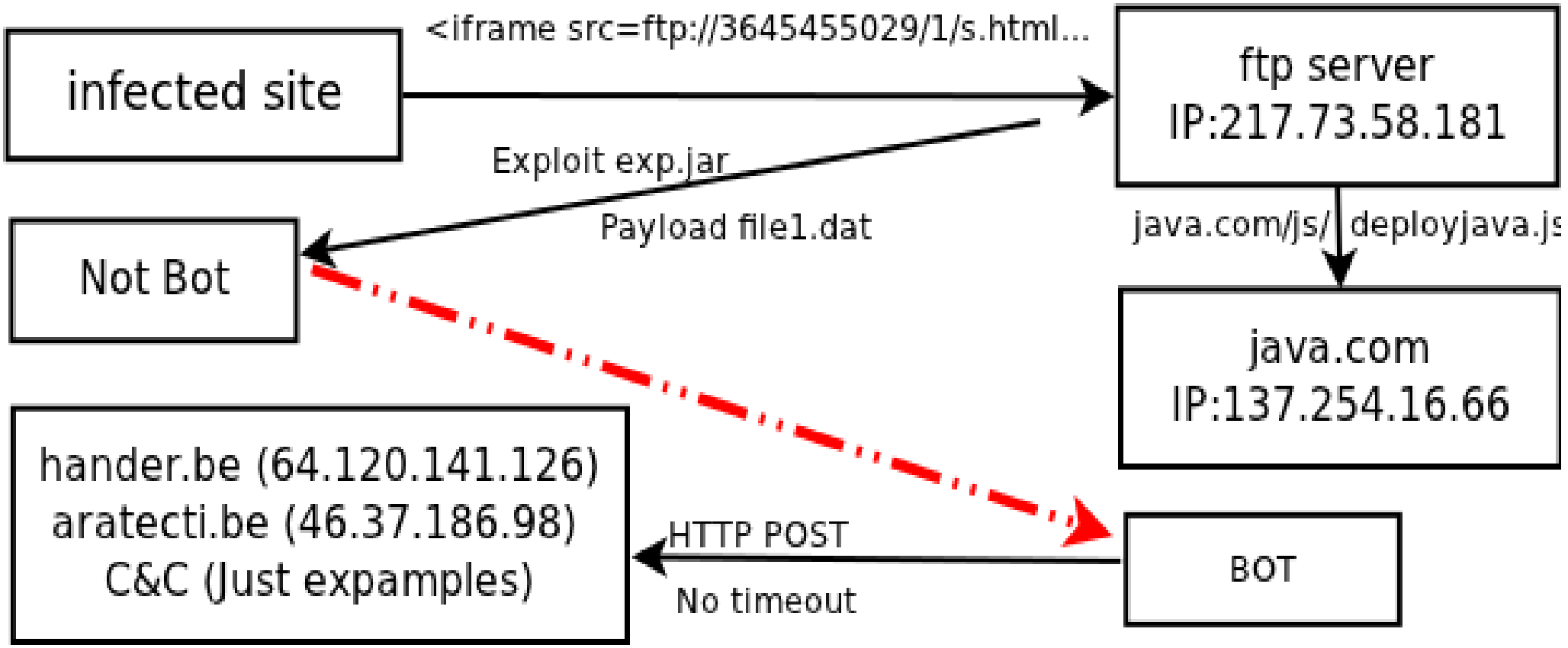
- Generative-domain name based domains generate very specific voluminous DNS traffic
- Our research is primarily focused on picking up these patterns. Example Carberp (details provided by Vladimir Kropotov)

Carberp

- Bot Infection: Drive-By-HTTP
- Payload and intermediate malware domains: normal, recent registration dates or DynDNS
- Distributed via: Many many compromised web-sites, top score > 100 compromised resources detected during 1 week.
- C&C domains usually generated, but some special cases below ;-).
- C&C and Malware domains located on the same AS (from bot point of view). Easy to detect.
- Typical bot activity: Mass HTTP Post



Size	Payload	Referrer	URL	Domain
9414	javascript	www.*****press.ru	/g/1841836267259 5167.js	beatshine.is-saved.org
45443	html	www.*****press.ru	/index.php?28d9000e56c2a63080ff89c6f5357591	activatedreplacing.is-very-evil.org
4135	application/x-jar		//images/r/785cee8be7f1da9a9d60820cbf8b1840**.jar**	activatedreplacing.is-very-evil.org
155529	application/executable		/**server_privileges.php**?91370f5f009a815950578cb539f28b58=3	activatedreplacing.is-very-evil.org



Size	Payload	Referrer	URL	Domain
997	html	Infected site	/1/s.html	3645455029
4923	javascript	3645455029	/js/deployJava.js	Java.com
18046	application/x-jar		/1/exp.jar	3645455029
138352	application/executable		/file1.dat	3645455029

Detection: related works

From Throw-Away Traffic to Bots: Detecting Rise of DGA-Based Malware (Manos Antonakakis, Roberto Redisci et al) (2012)

L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi.

EXPOSURE: Finding malicious domains using

passive dns analysis. In Proceedings of NDSS,

2011

etc..

What we do differently:

- “lazy” WHOIS lookups, team cymru IP to ASN lookups
- Our own passive DNS index
- Sandbox farm (mainly to detect compromised websites automagically and study behavior)

Dealing with false positives: filtering

- Generated sequences: n-gram analysis

$w_1 w_2$	bigram
$w_1 w_2 w_3$	trigram
$w_1 w_2 w_3 w_4$	four-gram

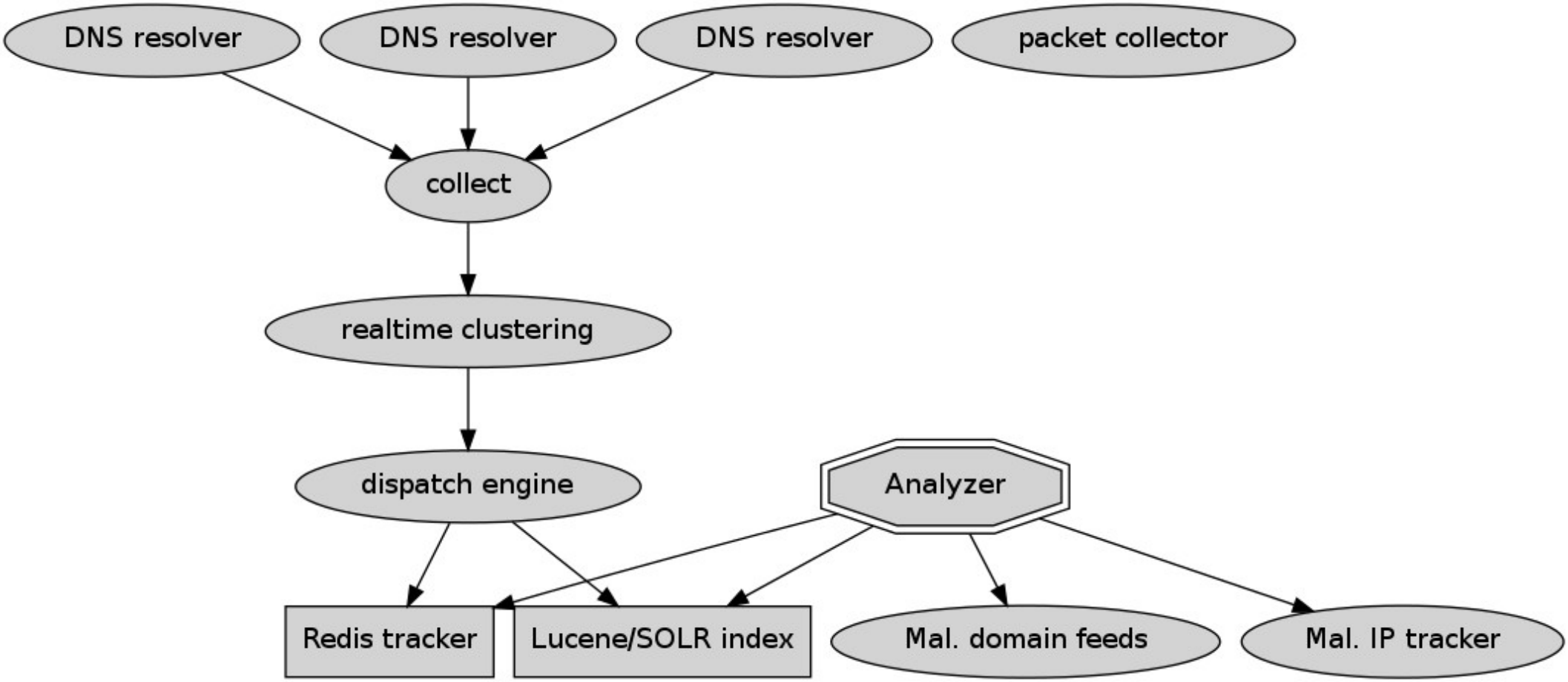
- WHOIS c
- Ips belong to Malicious ASN
- Public domain lists (alexa top 100k) works well as whitelist

Cat and mouse game

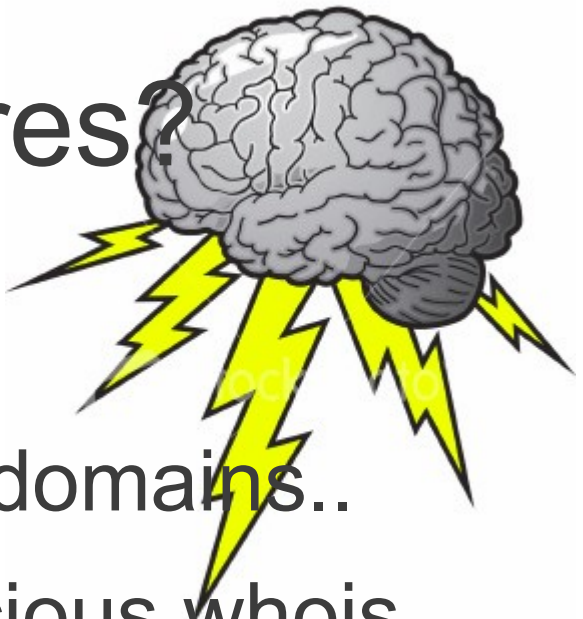
- Of course all of this is easy to evade. Once you know the method. But security is always about 'cat-n-mouse' game ;-)

Architecture

- What we are building ;)



Are we using signatures?

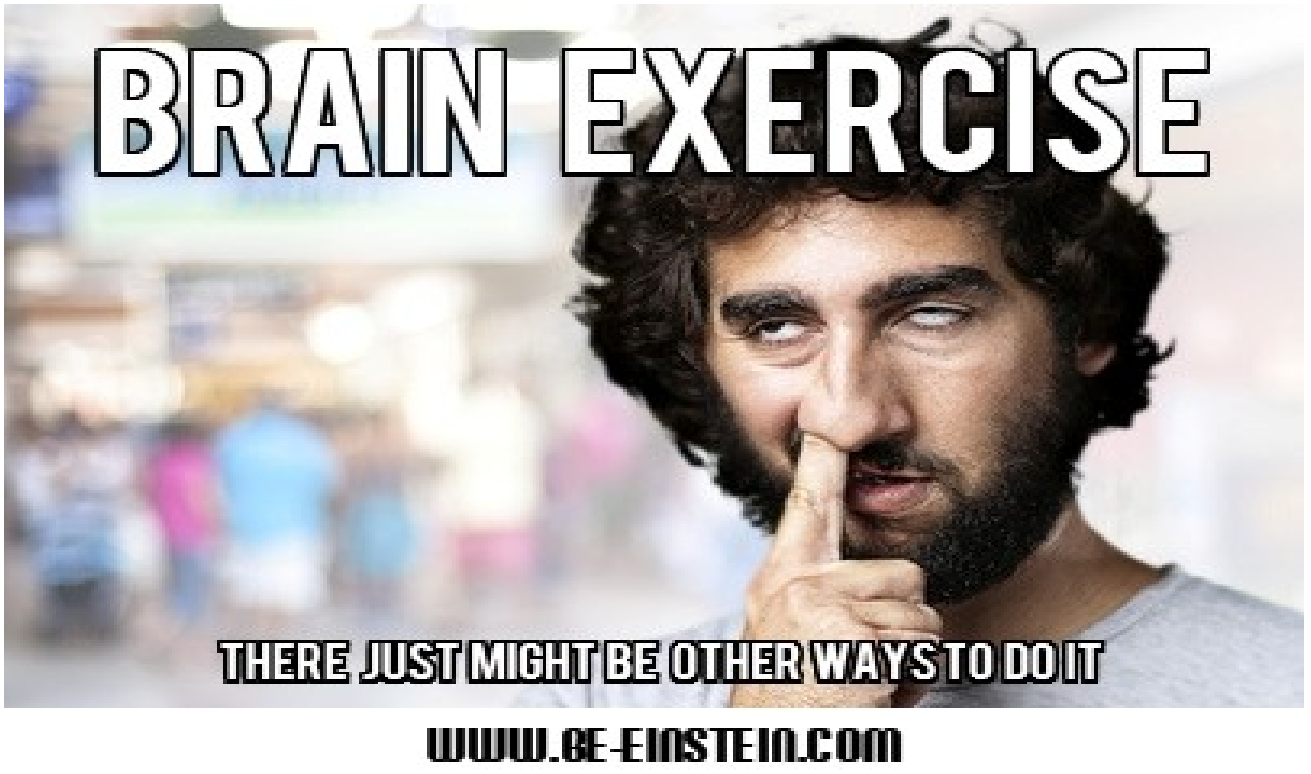


Yes and No..

- We don't have signatures for C&C domains..
- But we maintain patterns for suspicious whois data (registration date, registrar, email, ..)
- Historical DNS and AS association (bad IP)
- Generic patterns for generative domains (high, similarly distributed pattern of failed lookups within the same zone)

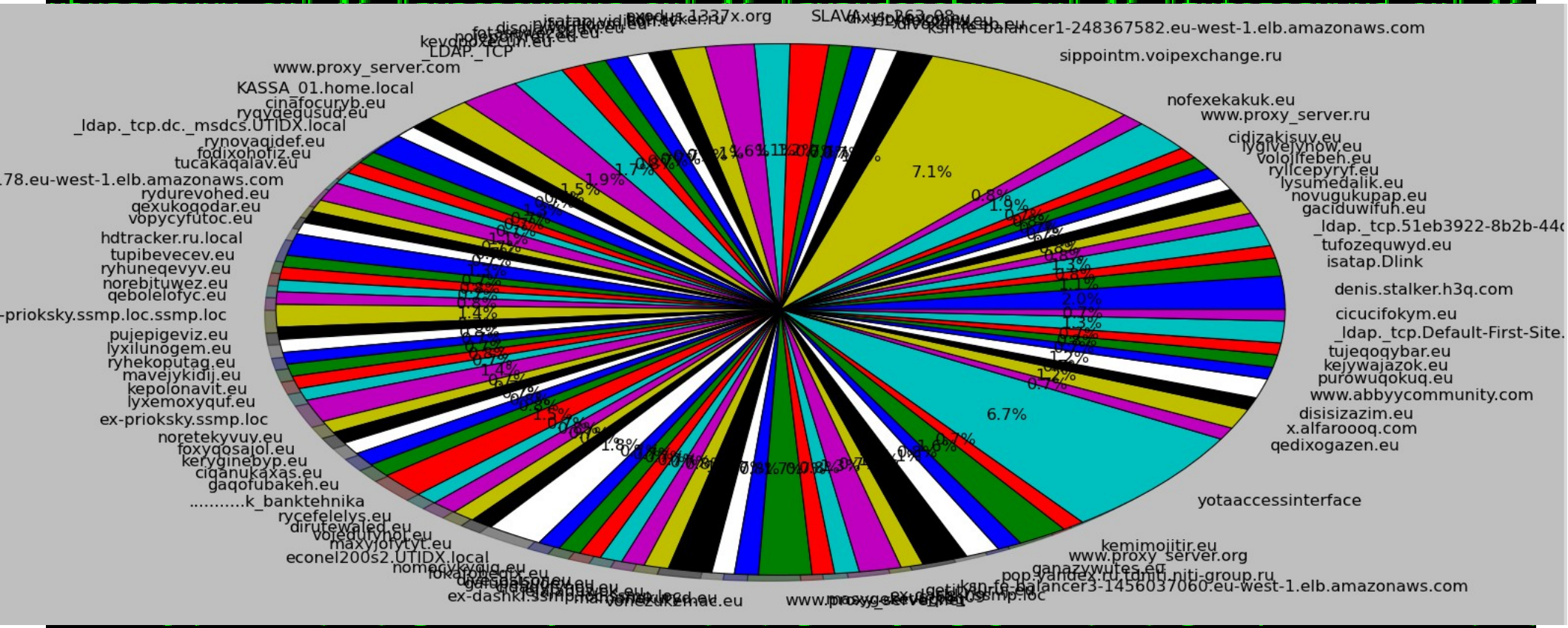
A walk through automated detection

- In this example we will show how automated detection works step by step. We will show redis queries in form of interactive session:



Detection starting point: rcode: 3 (Non-existing domains)

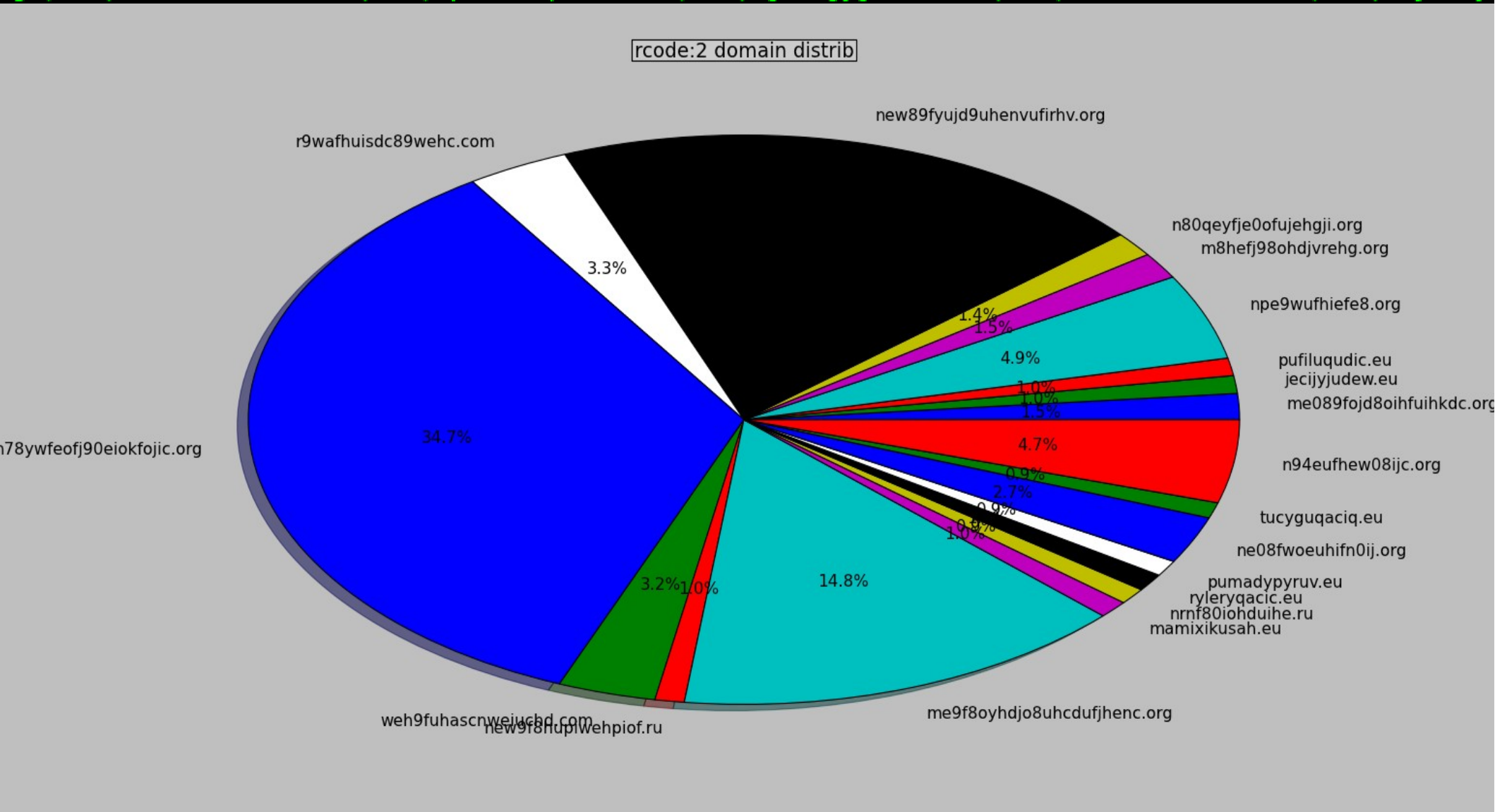
inyhotyqyt.eu",16,"foxehehywef.eu",16,"ganazywutes.eu",16,"jenujoxojug.eu",16,
ejywajazok.eu",16,"lygivejynow.eu",16,"lykonurymex.eu",16,"lyvitexemod.eu",16,
yxemoxyquf.eu",16,"novugukupap.eu",16,"pufyjulogih.eu",16,"qedixogazen.eu",16,



atykibojig.eu",14,"jecygyrogec.eu",14,"jefogixuqyn.eu",14,"jenokirifux.eu",14,

yhdjo8uhcdufjhenc.org",1995,"jebena.ananikolic.su",1017,"victorytowers.com",768,"npe9wufhiefe8.org"
606,"relays.visi.com",573,"dnsbl.void.ru",453,"r9wafhuisdc89wehc.com",442,"weh9fuhascnweiuchd.com",
8fwoeuhifn0ij.org",347,"blackwm.ru",241,"m8hefj98ohdjvreng.org",195,"me089fojd8oihfuihkdc.org",192,
186,"me8ofilejf8oheicj
rg",166,"voworemoziv.eu",139,"pufiluqudic.eu",136,"jecijyjudew.eu",135,"mamixikusah.eu",133,"ryleryq

Rcode:2 domains (failed servers)



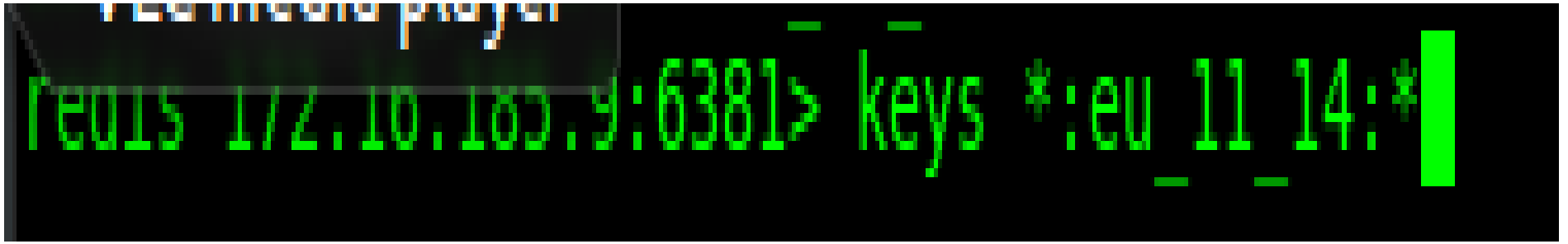
Sample analysis (step by step)

- Start looking for a failed pattern and cluster id:

```
redis 172.16.185.9:6381> keys *:eu*:2
 1) "fotyriwavix.eu:eu_11_14:2"
 2) "leporno.eu:eu_7_10:2"
 3) "www.europarus.eu:eu_9_16:2"
 4) "maxyjofytyt.eu:eu_11_14:2"
 5) "gaquviwyrup.eu:eu_11_14:2"
 6) "lysovidacyx.eu:eu_11_14:2"
 7) "kezapyjolek.eu:eu_11_14:2"
 8) "nomebemenid.eu:eu_11_14:2"
 9) "rx-piller24.eu:eu_11_14:2"
10) "kefuwidijyp.eu:eu_11_14:2"
11) "disafuwokis.eu:eu_11_14:2"
12) "www.cube.eu:eu_4_11:2"
13) "ryleryqacic.eu:eu_11_14:2"
14) "tv.dtvnet.eu:eu_6_12:2"
15) "ganycyhywek.eu:eu_11_14:2"
16) "kejitanokon.eu:eu_11_14:2"
17) "rx-pillen24.eu:eu_11_14:2"
18) "tufecagemyl.eu:eu_11_14:2"
19) "pufiluqudic.eu:eu_11_14:2"
redis 172.16.185.9:6381>
```

Sample analysis (two)

- Get the cluster ID: (eu_11_14)



Clustering is based on domain similarity. Currently used characteristics:
- f(zone, pattern (length, depth))
- additional characteristics (building up): natural language domain vs. generated string (occurrence of two-character sequences - n-grams)
- domain registration parameters (obtained via WHOIS [**problematic!**])
- cross-reference with existing malicious IP and AS reputation database (incrementally built by us)

Sample analysis

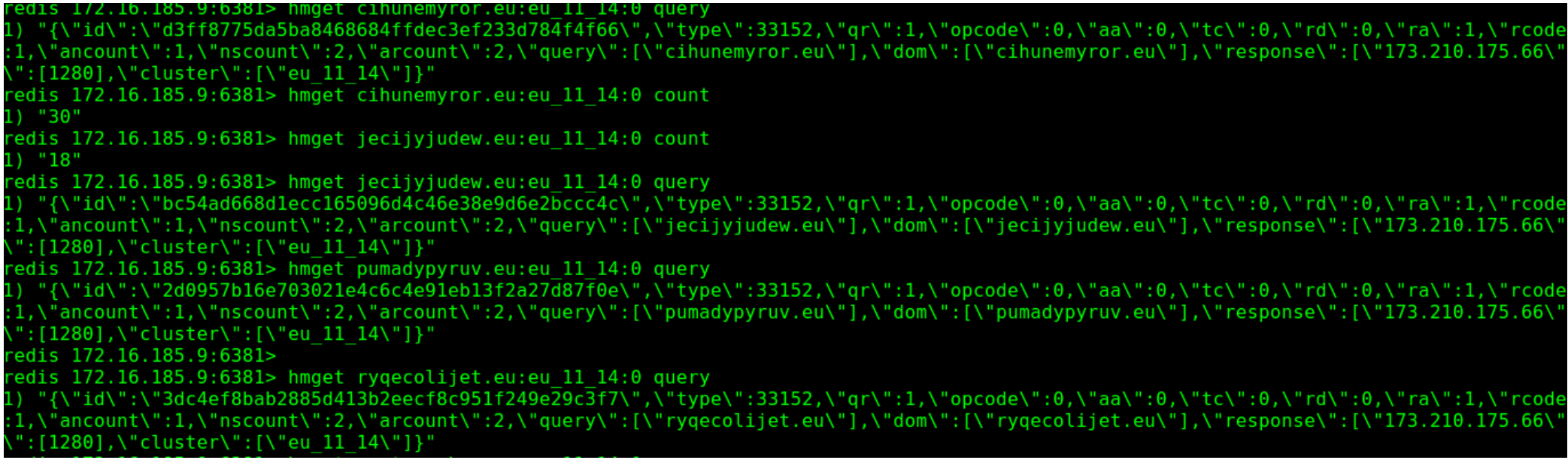
- Get other members of the cluster

```
957) "cilynitiseg.eu:eu_11_14:3"
958) "kezubaxemor.eu:eu_11_14:3"
959) "jeledajifor.eu:eu_11_14:3"
960) "foguhosecib.eu:eu_11_14:3"
961) "xuderadezuv.eu:eu_11_14:3"
962) "jecaduxakeh.eu:eu_11_14:3"
963) "kemelixakyz.eu:eu_11_14:3"
964) "jeluzydyqej.eu:eu_11_14:3"
965) "volebatijub.eu:eu_11_14:3"
966) "puzubovafik.eu:eu_11_14:3"
967) "mavyvomuqal.eu:eu_11_14:3"
968) "magetyfisus.eu:eu_11_14:3"
969) "qedoqyvoguq.eu:eu_11_14:3"
970) "dirojubusux.eu:eu_11_14:3"
971) "fodutazenaf.eu:eu_11_14:3"
972) "lyrefanyril.eu:eu_11_14:3"
973) "vocerocofyf.eu:eu_11_14:3"
974) "pujamyqywyk.eu:eu_11_14:3"
975) "xutoxedyniq.eu:eu_11_14:3"
976) "tuwiqelages.eu:eu_11_14:3"
977) "jejajaduwok.eu:eu_11_14:3"
978) "xuxehajexuw.eu:eu_11_14:3"
979) "rytonovejof.eu:eu_11_14:3"
980) "vococumecan.eu:eu_11_14:3"
```

```
"tufecagemyl.eu:eu_11_14:2"
"pufiluqudic.eu:eu_11_14:2"
"maxyjofytyt.eu:eu_11_14:2"
"kezapyjolek.eu:eu_11_14:2"
```

Sample analysis

- Find common members (notice avatarmaker.eu could be a false positive, easily filtered out through common denominator filering (IP, WHOIS information)



Sample analysis

- So we have C&C IP 66.175.210.173
- we can continue mining to see if we get any other domain names:

```
redis 172.16.185.9:6381> hmget "173.210.175.66:0" query
1) "{\"id\":\"4b592c68f488077a509222645e320bdbf6a6e197\",\"type\":33168,\"qr\":1,\"opcode\":0,\"aa\":0,\"tc\":0,\"rd\":0,\"ra\":1,\"rcode\":0,\"qdcount\"
:1,\"ancount\":1,\"nscount\":2,\"arcount\":1,\"query\":[\"l33t.brand-clothes.net\"],\"dom\":[\"l33t.brand-clothes.net\"],\"response\":[\"173.210.175.66\"
],\"response_ttl\":[256],\"cluster\":[\"net_13_22\"]}"
redis 172.16.185.9:6381>
```

Sample analysis

- **Look! We just met an old friend!!**

Sample analysis

- Palevo:

```
ping ryqecolijet.eu
.eu (66.175.210.173) 56(84) bytes of data.
507-173.members.linode.com (66.175.210.173):
507-173.members.linode.com (66.175.210.173):
```

scovered: January 19, 2010
dated: January 19, 2010 5:21:37 PM
so Known As: P2P-Worm.Win32.Palevo.bpji [Kaspersky]
pe: Worm
fection Length: 142,848 bytes
stems Affected: Windows 2000, Windows 95, Windows 98, Windows Me, Win
03, Windows Vista, Windows XP

en the worm is executed, it creates the following files:
%SystemDrive%\RECYCLER\[SID]\nissan.exe
%SystemDrive%\RECYCLER\[SID]\Desktop.ini
%DriveLetter%\RECYCLER\[SID]\csrxx.exe (W32.IRCBot)
%DriveLetter%\SLATKO\torta.exe
%DriveLetter%\SLATKO\Desktop.ini
%DriveLetter%\autorun.inf

en creates the following registry entry, so that it starts when Windows starts:
EY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Winlogon\"Taskı
ssan.exe"

e worm then opens a back door and connects to the following domains on UDP port 25000:
andra.prichaonica.com
ica.banjalucke-ljepotice.ru
3t.brand-clothes.net

Home | Blocklists | Statistic | Contact

Palevo Tracker

Palevo Botnet C&C IP address :: 66.175.210.173

C&C IP address:	**66.175.210.173**
Hostname:	li507-173.members.linode.com
SBL:	SBL148105
AS number:	AS8001
AS name:	NET-ACCESS-CORP - Net Access Corporation
Country:	United States (US)
Firstseen (UTC):	2012-07-20 20:00:06
Lastseen (UTC):	2012-07-22 14:30:06

C&Cs on this IP: 2

Palevo Command&Control servers hosted on this IP address

Below is a list of Palevo Command&Control servers that are hosted on this ip address (66.175.210.173):

Palevo C&C domain	IP address	Firstseen (UTC)
elcrazyfrog.com	66.175.210.173	2012-07-20 20:00:06

# of Palevo C&C domains: **1**

Mapping C&C (easily automated)

- http://cihunemyror.eu/login.php
- http://foxivusozuc.eu/login.php
- http://ryqecolijet.eu/login.php
- http://xuqohyxeqak.eu/login.php
- http://foqaqehacew.eu/login.php
- http://jecijyjudew.eu/login.php
- http://voworemoziv.eu/login.php
- http://mamixikusah.eu/login.php
- http://qebahilojam.eu/login.php
- http://foqaqehacew.eu/search.php
- http://foqaqehacew.eu/search.php
- http://foqaqehacew.eu/LMvg9Ng1d.php

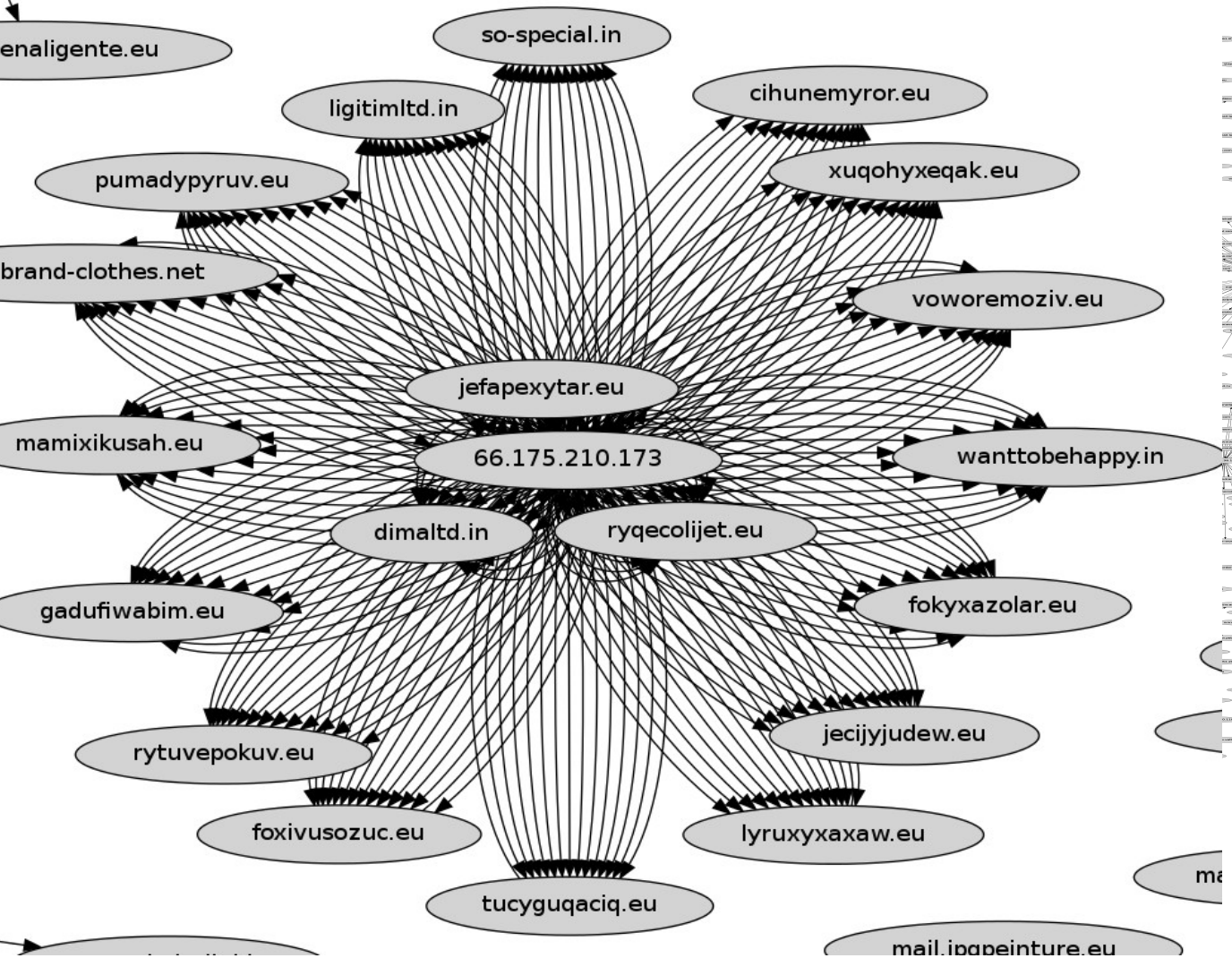
Sample analysis

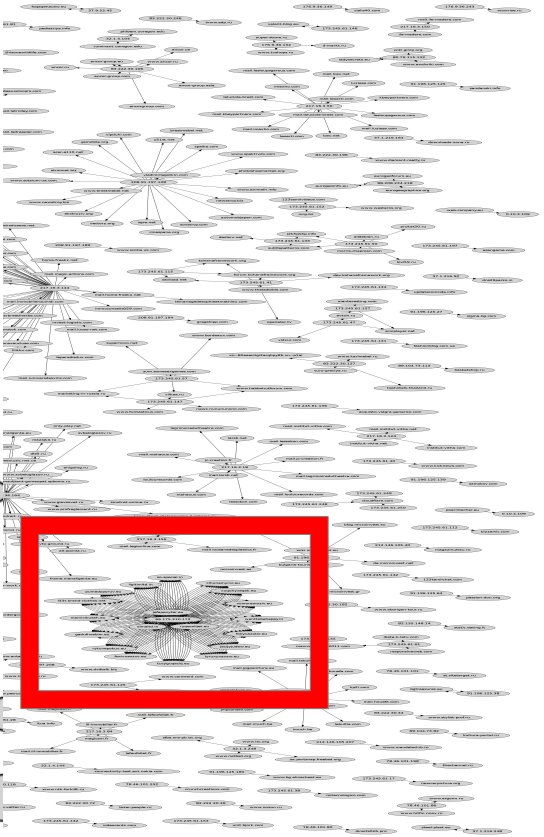
- Finding more relevant domains:

```
redis 172.16.185.9:6381> keys 173.210.175.66*
 1) "173.210.175.66;fokyxazolar.eu"
 2) "173.210.175.66;jefapexytar.eu"
 3) "173.210.175.66;voworemoziv.eu"
 4) "173.210.175.66;lyruxyxaxaw.eu"
 5) "173.210.175.66;stolovka.us"
 6) "173.210.175.66;l33t.brand-clothes.net"
 7) "173.210.175.66;ryqecolijet.eu"
 8) "173.210.175.66;pumadypyruv.eu"
 9) "173.210.175.66:0"
10) "173.210.175.66;cihunemyror.eu"
11) "173.210.175.66;xugohyxegak.eu"
```

```
redis 172.16.185.9:6381> hmget ligitimltd.in:in_10_13:0 query firstseen lastseen count
1) "{\"id\":\"9b962f168e88cc2056d5ed039684577682dfc084\",\"type\":33168,\"qr\":1,\"opcode\":0,\"aa\":0,\"tc\":0,\"rd\":0,\"ra\":1,\"rcode\":0,\"qdcount\"
:1,\"ancount\":1,\"nscount\":2,\"arcount\":1,\"query\":[\"ligitimltd.in\"],\"dom\":[\"ligitimltd.in\"],\"response\":[\"173.210.175.66\"],\"response_ttl\"
:[1280],\"cluster\":[\"in_10_13\"]}"
2) "Wed Sep 05 2012 00:38:08 GMT-0400 (EDT)"
3) "Wed Sep 05 2012 00:59:57 GMT-0400 (EDT)"
4) "14"
redis 172.16.185.9:6381>
2) "dimaltd.in:in_7_10:0"
redis 172.16.185.9:6381> hmget dimaltd.in:in_7_10:0 query firstseen lastseen count
1) "{\"id\":\"d18f237efc5fd8e280468e7160e4f3a56c73df15\",\"type\":33168,\"qr\":1,\"opcode\":0,\"aa\":0,\"tc\":0,\"rd\
":0,\"ra\":1,\"rcode\":0,\"qdcount\":1,\"ancount\":1,\"nscount\":2,\"arcount\":1,\"query\":[\"dimaltd.in\"],\"dom\":[
\"dimaltd.in\"],\"response\":[\"173.210.175.66\"],\"response_ttl\":[1280],\"cluster\":[\"in_7_10\"]}"
2) "Wed Sep 05 2012 00:31:53 GMT-0400 (EDT)"
3) "Wed Sep 05 2012 00:49:42 GMT-0400 (EDT)"
```

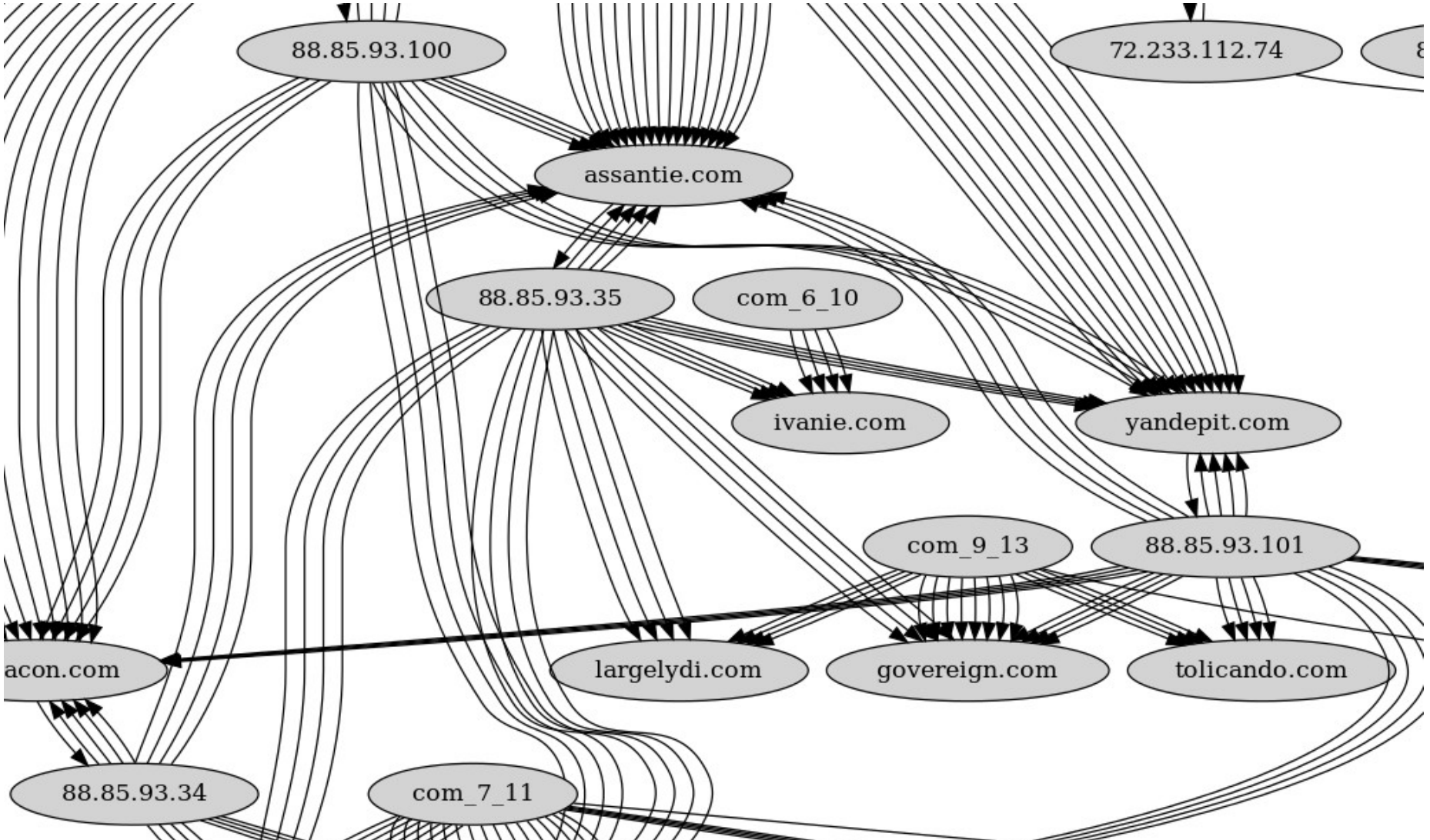
Automation


enaligente.eu
so-special.in
ligitimltd.in
cihunemyror.eu
pumadypyruv.eu
xuqohyxeqak.eu
brand-clothes.net
voworemoziv.eu
jefapexytar.eu
mamixikusah.eu
66.175.210.173
wanttobehappy.in
dimaltd.in
ryqecolijet.eu
gadufiwabim.eu
fokyxazolar.eu
jecijyjudew.eu
rytuvepokuv.eu
foxivusozuc.eu
lyruxyxaxaw.eu
tucyguqaciq.eu
mail.ipgpeinture.eu



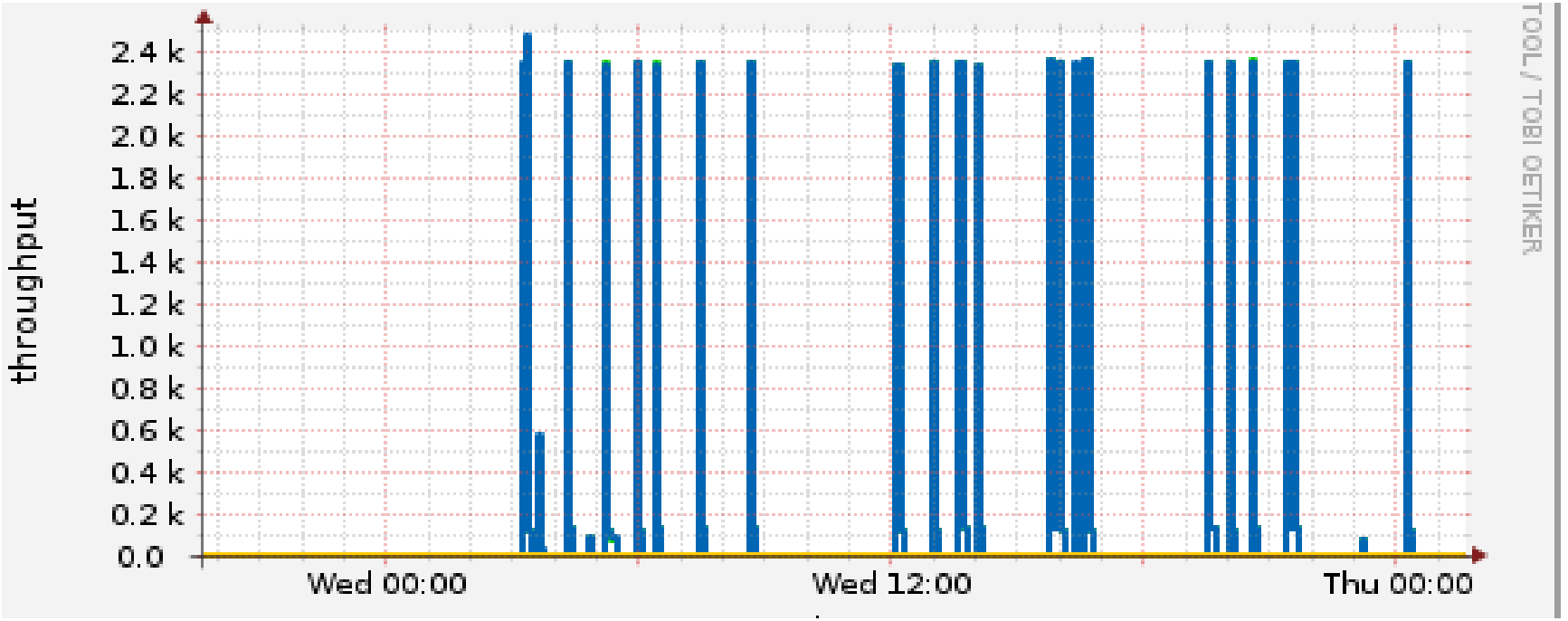
Zoom in...

Timestamp	Source IP	Destination IP	Severity	
2012-06-07 22:48:12	88.85.93.100	urlQuery Client	0	ET RBN Known Russian Business Network IP (402)



Performance

- On single machine (32Gb RAM) we run up to 2000 pkt/sec without significant performance loss
- Average load:



Other Interesting numbers

- Packets per day: ~130M filtered.
- Mal. Domains/day: ~30k DNS queries (varies)
- Avg. 30-50 req/minute for single domain
-

Uses of the data

- Obvious: blacklists
- Botnet take overs (costs 11USD or less ;)
- Sinkholing

Detection

- 

What could be more flux than fastflux? ;-)

- WHOIS fastflux … HOW?!

```
fygrave@borzo:~$ whois FOOTBALL-SECURITY-WETRLSGPIEO.ORG
NOT FOUND
fygrave@borzo:~$
```

Domain ID:D166393631-LROR
Domain Name:FOOTBALL-SECURITY-WETRLSGPIEO.ORG
Created On:21-Aug-2012 01:23:52 UTC
Last Updated On:21-Aug-2012 01:23:53 UTC
Expiration Date:21-Aug-2013 01:23:52 UTC
Sponsoring Registrar:Click Registrar, Inc. d/b/a publicdomainregistry.com (R1935-LROR)
Status:CLIENT TRANSFER PROHIBITED
Status:TRANSFER PROHIBITED
Status:ADDPERIOD
Registrant ID:PP-SP-001
Registrant Name:Domain Admin
Registrant Organization:PrivacyProtect.org
Registrant Street1:ID#10760, PO Box 16
Registrant Street2:Note - All Postal Mails Rejected, visit Privacyprotect.org
Registrant Street3:

Moving ahead:
Finding easy targets before they do :)

In short, it is all about quick ways of finding idiots having no clue of what they are doing with wordpress, oscommerce, openx, [put yer fave]
And forcing them to update before they get owned ;)
And hmm.. doing it country-wide

disclaimer

Just another “small data” project we play with.
Around 4 machines solr cluster.

Largely inspired by “Fruit: why so low?” by Adam MetlStorm (hack.lu 2011)

Scanning internet is not new.. but pretty much realistic

Demystifying Service Discovery: Implementing an Internet-Wide Scanner

Derek Leonard and Dmitri Loguinov
Department of Computer Science and Engineering
Texas A&M University, College Station, TX 77843 USA
{dleonard,dmitri}@cse.tamu.edu

Scanner	Scope	Permutation	Servers	Protocol	Port	Timeout	Duration	Blacklist	.0/.255	Exclude
Pryadkin [43]	$\mathcal{I}$	uniform	3	ICMP/TCP	–	10s	123d	yes	no	no
Benoit [5]	$\mathcal{NR}$	uniform	25	TCP	80	30s	92d	no	yes	no
Dagon [13]	$\mathcal{I}$	uniform	–	UDP	53	–	30d	–	yes	US Gov
Heidemann [17]	$\mathcal{I}$	RIS	8	ICMP	echo	5s	52d	yes	no	no

Table 1: Large-scale service discovery in the literature (dashes represent unreported values).

Low-Load Server Crawler: Design and Evaluation

Katsuko T. Nakahira
Nagaoka University of Technology
1603-1 Kamitomiokamachi, Nagaoka
Niigata, Japan
katsuko@vos.nagaokaut.ac.jp

Tetsuya Hoshino
Nagaoka University of Technology
1603-1 Kamitomiokamachi, Nagaoka
Niigata, Japan
065365@mis.nagaokaut.ac.jp

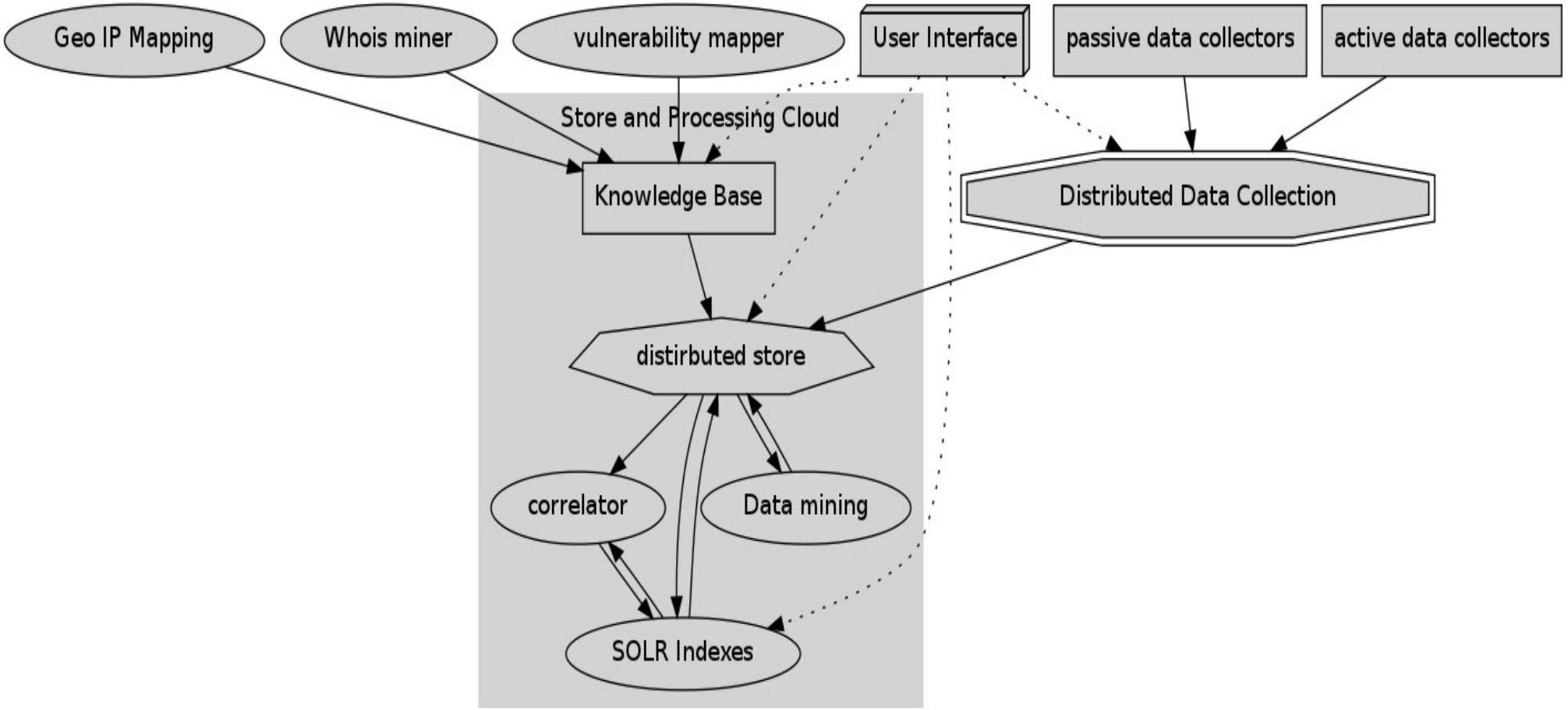
Yoshiki Mikami
Nagaoka University of Technology
1603-1 Kamitomiokamachi, Nagaoka
Niigata, Japan
mikami@kjs.nagaokaut.ac.jp

Architecture

- Network port discovery (agents)
- Banner collection (agents)
- Backend Store: SOLR
- Collectibles: services and ports, OS fingerprints,
- ASN/OWNER/netblock/Country, geographical location/App data

Architecture(2)

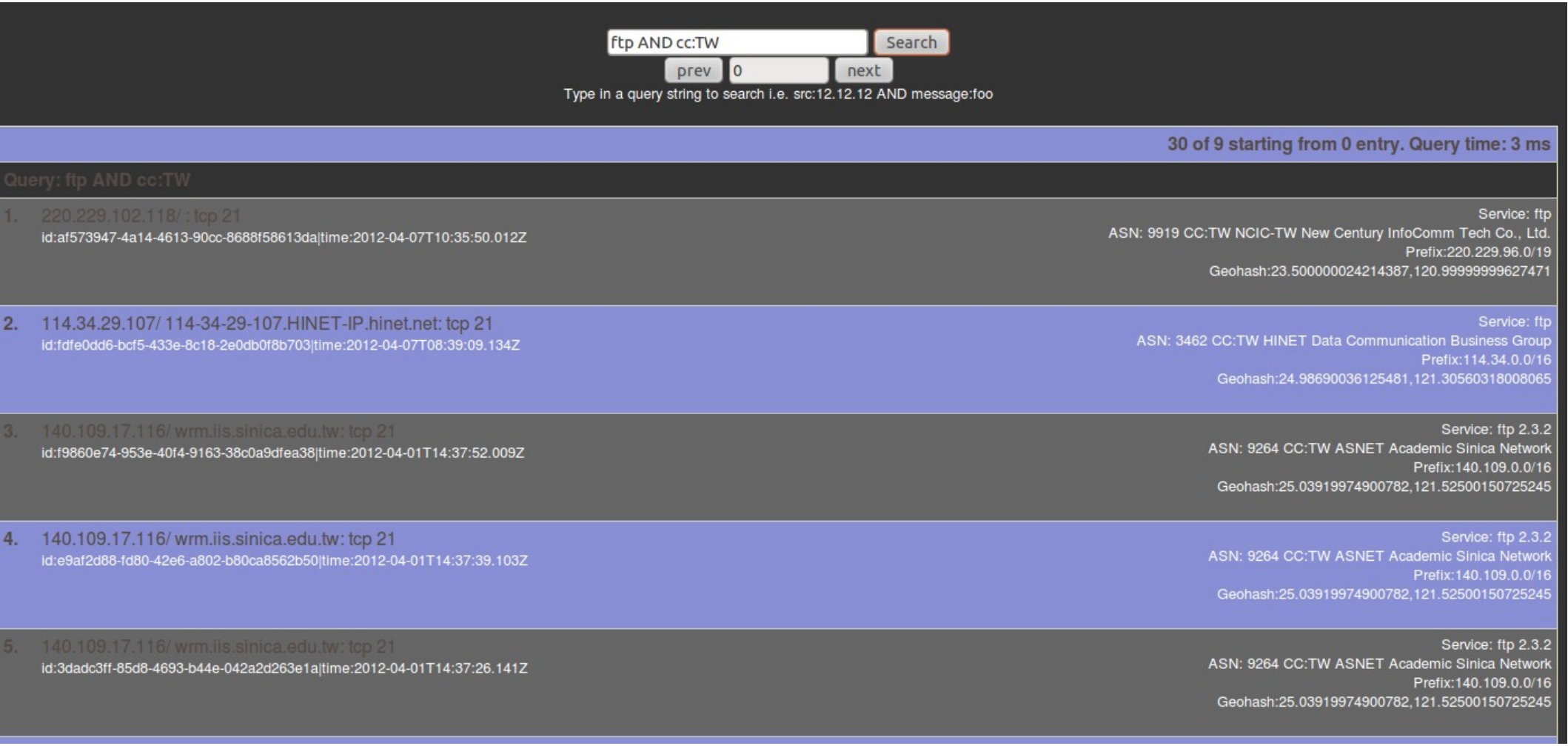
- Roughly something like that



Approach

- Scan slow (avoid abuse reports)
- Index time
- Passive “mapper” (simple sniffer + browser fingerprinting at the moment)
- Larger range of ports (account port numbers, which are actively being scanned from firewall log analysis, honeypot machines etc)
- For web apps – (wafp fingerprinting) + index banner (noisy, cause of most of the abuse complaints)

How you use this shit...



Features

- Scriptable via restful API (think of solr) (cuz UI is for sissies ;-))
- Query by any combination of:
  - software version/banner regex (solr/lucene style)
  - geospatial search (via geohash)
  - ASN or regex on ASN owner
  - Country code

Uses

CERT team: automated notifications of idiots running old wordpress within particular range, geographic location or organization is a one liner script

Questions

@fygrave
@vbkropotov