New (and Newly-Changed) Fully Qualified Domain Names (FQDNs): A View of Worldwide Changes to the Internet's DNS

Black Hat Europe 2015

Paul Vixie, Ph.D.
CEO, Farsight Security, Inc.

I. Introduction: Why Pay Attention to New Hostnames And DNS Changes?

Everything Uses The Domain Name System (DNS)

- DNS is a fundamental Internet protocol, and critical to the smooth operation of the Internet – everything relies on DNS.

- While most users of DNS are law-abiding, cyber criminals also rely on DNS to enable botnets and support other illicit activities.

- DNS (as relied upon by law-abiding users) is also a favorite point of attack and compromise.

- Rapidly detecting unexpected DNS changes and other anomalous DNS-related behaviors is of paramount importance to Internet security -- but was traditionally viewed as impossible to do in real-time and at Internet-scale (the Internet must be "too big" for this to be possible, right?)

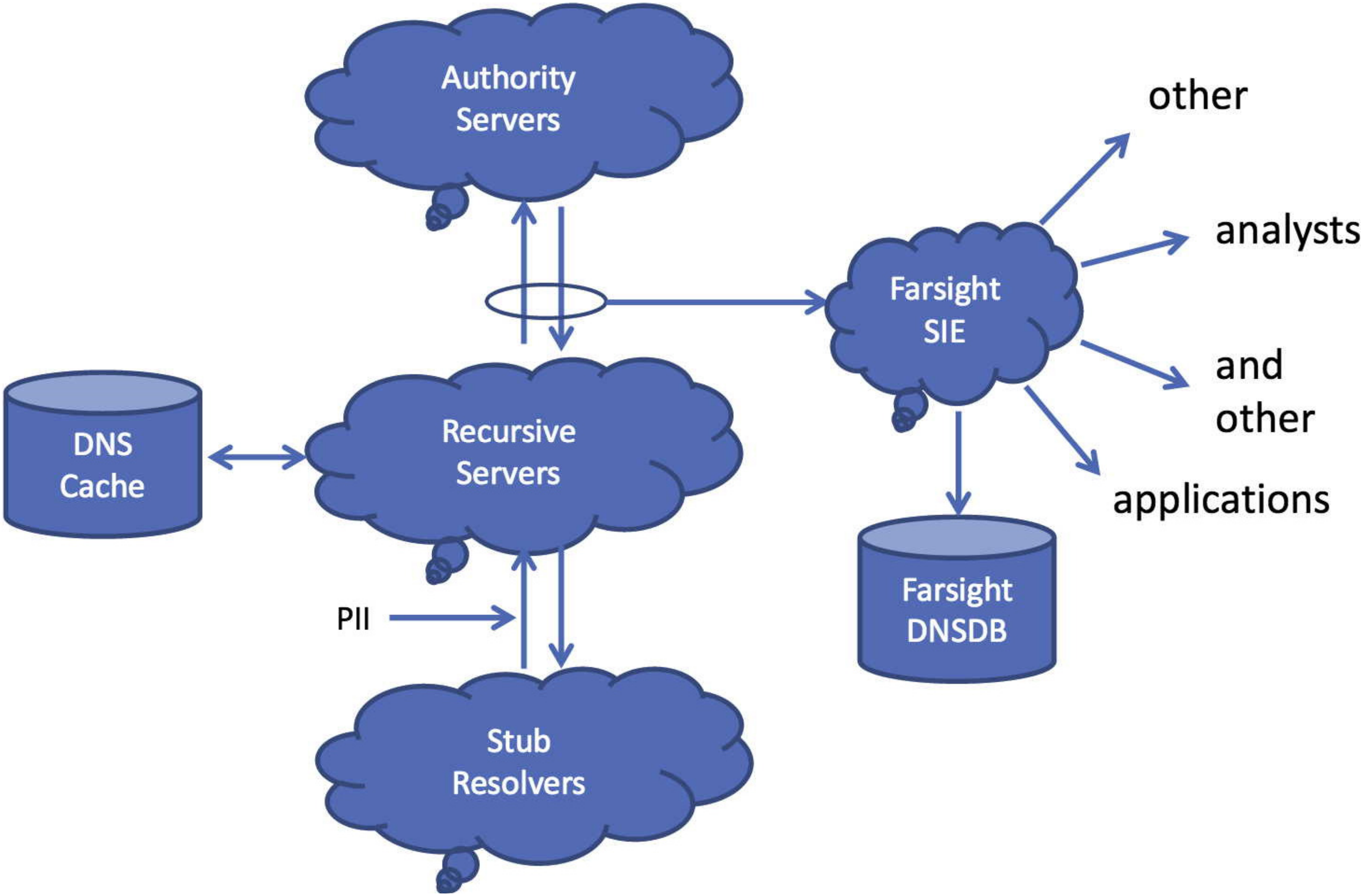
Given Potentially Overwhelming Levels of DNS Traffic, You Need To Learn What You Can Ignore

- Just like a pilot learning to fly a plane with a complex instrument panel, you need to figure out what you can ignore and what must NOT be overlooked
- If a domain is long-established, and is on the same IP address it has always been on, you probably don't need to pay much attention to it (and if for some reason it turns out you do need info about such a domain, you can easily get that historical info from DNSDB).
- **Pay attention to new stuff, and to any stuff that has recently changed.** Tune out old/invariant stuff. This insight lead to our Newly Observed Domains (NOD) channel, our Newly Observed Hostnames (NOH) channel, and our DNS Changes channel on the Farsight Security, Inc., **Security Information Exchange (SIE)**.

"What's The Security Information Exchange?"

- The Security Information Exchange (SIE) is a **data sharing facility.**
- Different types of data get shared over SIE broadcast channels, much as a cable TV provider offers a variety of TV channels with news, sports, movies, etc. In SIE's case, **channels focus on DNS-related data, darknet data, spam data, phishing data, etc.** See www.farsightsecurity.com/Technical/fsi-sie-channel-guide.pdf
- **Unfunded academic researchers (and unfunded people working in the best interest of the Internet, aka "Internet do-gooders") can request free or partially subsidized access to SIE.**
- Approved corporations, approved commercial researchers, and approved government agencies can **purchase** access to SIE (Farsight carefully screens all users to ensure we don't inadvertently permit access to SIE by cyber criminals or others not acting in the best interests of the Internet)

Farsight's Passive DNS Collection Architecture



Passive DNS Collection Architecture Implications

- We've intentionally structured our sensor architecture to AVOID collecting any personally identifiable information (PII)when we collect passive DNS information. E.G., we have no interest in the selection of websites any particular user chooses to visit.
- We manage to avoid collecting this sort of PII by collecting DNS traffic ABOVE large recursive resolvers, so we don't see traffic that's attributable to individual end users – all queries appear to come from the recursive resolver itself. That query stream combines queries from thousands or tens of thousands of different users in a way that cannot be readily teased back apart.
- Obviously we also don't intentionally collect anything that involves RFC 1918 private address space since that space has no meaning outside the local enterprise where it may be used.
- Coming back to the passive DNS information we collect, and our recent focus on newly observed DNS information...

Short-Lived Cyber Criminal Domains vs. Comparatively Long-Lived Normal Domains

- **Domains used by cyber criminal are "different:"**
  - Used for illegal purposes (spam, scams, phishing, malware, warez, carding, illegal narcotic sales, online child exploitation materials, hacking/cracking, distributed denial of service attacks, etc.), or the support thereof.
  - Missing or untrustworthy point of contact details (or POCs hidden behind certain infamous proxy/privacy service providers)
  - Often paid-for with a stolen credit card
- **As a result, the usable life of cyber criminal domains tends to be "nasty, brutish, and short:"**
  - Police or civil investigators may seize widely abused domains
  - Domains ay end up suspended for bad whois information via WDPRS
  - Or domains may be killed for being purchased with a stolen credit card
- **This means that bad guys tend to have a disproportionate number of new (or newly changed) domain names.**
- **They particularly can't avoid this because of blocklists.**

Blocklists and Cyber Criminal Domains

- "Reputation Service Providers" track what happens with address space and domain names. When an IP address or domain name gets abused, it quickly gets noticed and negatively scored or blocked outright.

- At that point, the blocklisted IP address or domain name becomes damaged goods and basically worthless: most of the world will totally ignore any traffic from a blocklisted domain.

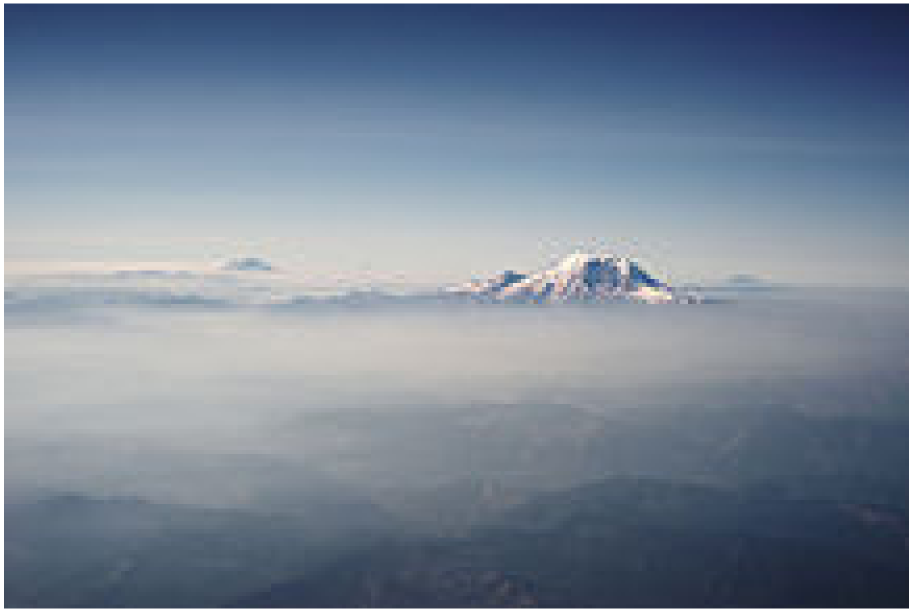
- The normal cyber criminal response? Lather, rinse, repeat: iterate by moving to new domains (or to new address space)

- This creates a LOT of churn/changes/new domain creation.

Modern DNS Reality vs. Classic DNS Misperception

REALITY – DNS is a hurricane, wild and varying widely across the globe, but increasingly well-observed...

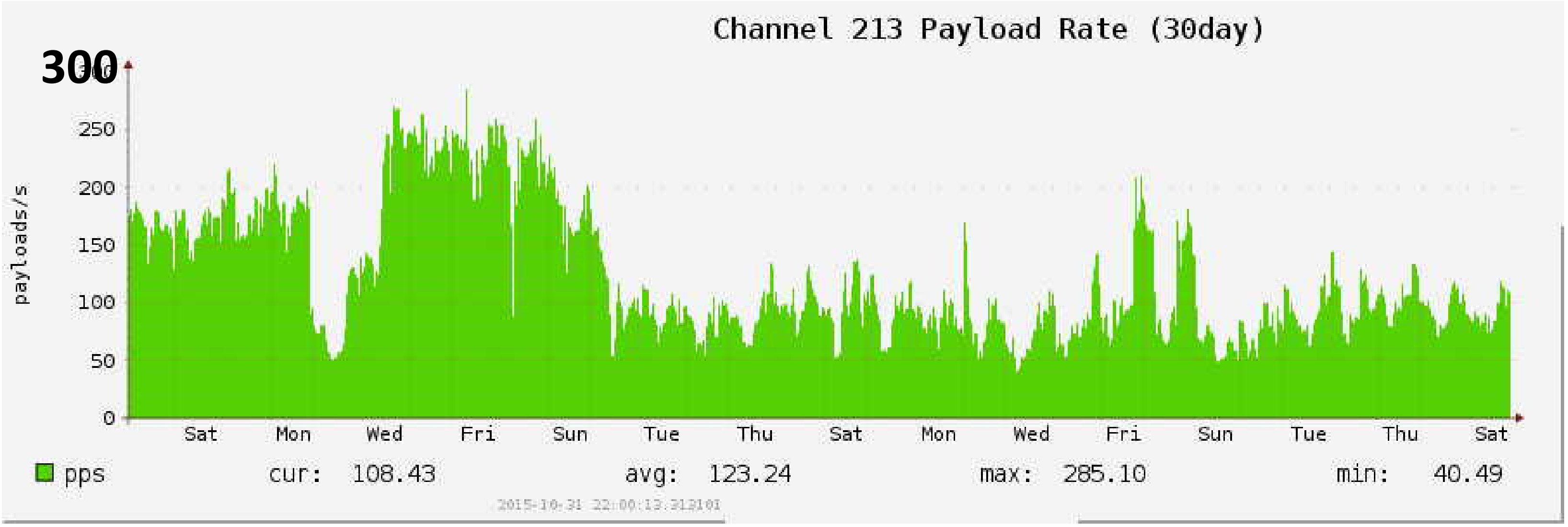
MISPERCEPTION – DNS is a mountain, stable, consistent, well-known but rather opaque to remote observers...

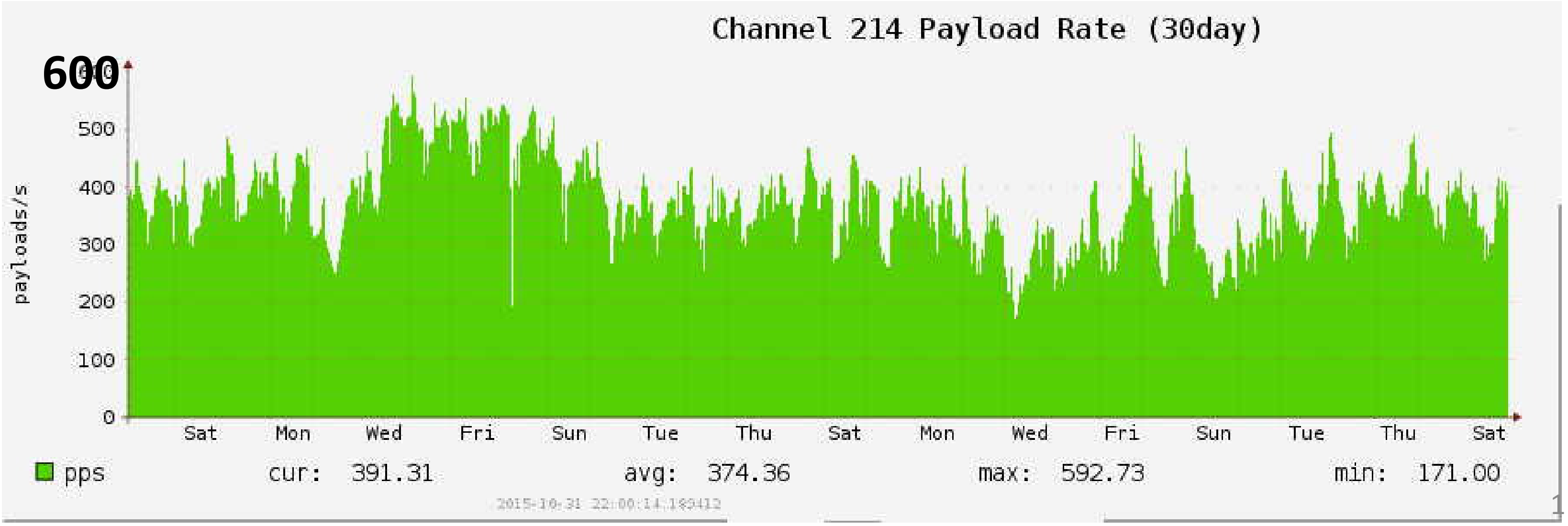


See the following RRDtool graphs for a sense of the real "size of the hurricane..."

[image credits at end of the talk]

Newly Observed Hostnames Per Second and DNS Changes Volume Per Second





Telling You? Not As Good As Showing You

- You can't really get a sense of what hundreds of DNS changes per second feels like from a graph.
- And in fact it is even hard to show you hundreds of new (or newly changed) hostnames per second on screen, but we can show you at least a somewhat slowed-down sample of some of those new names...
- See the video on the next slide...

A Few Seconds Worth of Newly Observed Hostnames

[drag mouse down here for video controls]

"Speed Reader" Training Is NOT Necessary

- You've got better things to do than just watch new hostnames scroll by on your screen.
- Many users will integrate a feed of newly observed hostnames with their existing real-time analysis framework, much like any other continuous source of real-time security data.
  - This is a **continuous processing model,** and preferred in the modern era.
  - It's the way most utilities (power plants, water plants, refineries, etc.) work: you set it up, turn it on, and it continuously produces thereafter.
  - If you know what you want to watch for, you can easily filter out just observations that match those selectors
- "Old school" users may decide to stick with "batch mode" instead
  - Draw a "batch" of observations from the stream of new hostnames, and save it to a file. Analyze that batch. Repeat for subsequent batches.
  - The problem with batch mode? Average latency is half the batch interval (e.g., if you analyze ten minute batches, your average latency will be five minutes). That's a LONG time to be blind to new phenomena today!

II. Newly Observed {domains, hostnames}

Newly Observed Domains (NOD)

- Farsight has previously been offered, and still offers a feed of Newly Observed **Domains**
- "Domains" in this context equals effective 2^{nd}-level domains
  - These are domains registered by end users immediately under TLDs (hypothetically, example.**com**)
  - Domains registered by end users under "effective TLDs" as determined by the Public Suffix List (PSL), see https://publicsuffix.org/list/
(for instance, northstar.**k12.ak.us**)
- NOD was (and is) practically useful as an RPZ (or rbldnsd) policy feed. You can use it to ignore all newly observed domains for a few minutes/few hours until reputation services can catch up.
- This has been discussed in Farsight's blog, see for example "So What's The Big Deal About New Domains?" www.farsightsecurity.com/Blog/20150528-stsauver-nod-delays/

A Sample Newly Observed Domain

```
$ nmsgtool –C ch212 –c 1
domain: adriatika74.ru.
time_seen: 2015-10-30 22:20:04
rrname: adriatika74.ru.
rrclass: IN (1)
rrtype: NS (2)
rdata: ns1.digitalocean.com.
rdata: ns2.digitalocean.com.
rdata: ns3.digitalocean.com.
```

Note: Being listed as a Newly Observed Domains is not saying that domain is good (or that it is bad), just saying that it is **NEW**. This is an objective and value-judgment free observation.

BUT, If We *Are* Looking for Badness, Spotting Some Badness Needs More Than Just Base Domain Names

Simple Paypal phishing example:

- **paypal-services.com.**webapps-service.co vs. webapps-service.co (looking at just the base domain, it looks perfectly fine/benign)
- **accountcheckpaypal.**fikrirpm.com vs. fikrirpm.com (looks benign)
- **paypal.**accountt-solve.com vs. accountt-solve.com (looks benign)
- **paypal-secure-updateinfo-3489348334.**fi.tempcloudsite.com vs. tempcloudsite.com (looks benign)
- **paypal-support.**iransargarmi.com vs. iransargarmi.com (looks benign)
- **www.paypal-update-informations-id-pp654355335**.toddesign.in vs. toddesign.in (looks benign)

Newly Observed Fully Qualified Domain Names

- Because of that reality, Farsight's now doing a feed of newly observed **fully qualified domain names** (aka **hostnames**)
- This makes many things easier, and a few things harder:
  - Easier to see suspicious hostnames (like the Paypal example just shown)
  - Trivial to spot sites using wildcarded/tagged hostnames
  - Potentially enables fine grained filtering (useful if confronting "human shield" hosting models interleaving "widows and orphans" with "gangsters" on the same base domain)

  BUT....

  - The volume becomes higher
  - Data sources may be "outed" if a bad guy has access to the newly observed hostnames (scenario: bad guy sends uniquely tagged spam and logs what address is sent each unique tag value; if a particular tagged hostname shows up in an abuse report, bad guy checks his logs to see who was sent that particular spam and list washes accordingly)

Sample Newly Observed Hostname

```
$ nmsgtool –C ch213 –c 1

domain: mucocutaneousmyrmecophaga.com.
time_seen: 2015-10-30 23:33:19
bailiwick: mucocutaneousmyrmecophaga.com.
rrname: airnaevihlrbrmf25kzpfse25aiofcwk7.mucocutaneousmyrmecophaga.com.
rrclass: IN (1)
rrtype: CNAME (5)
rdata: frontus.secretmedia.com.
```

Note: mucocutaneous = "pertaining to mucous membrane and skin"; myrmecophaga = the South American ant bear.

What *Is* That Domain?

- Check the **mucocutaneousmyrmecophaga.com** domain whois? It's hidden behind a whoisguard.com Panamanian domain reg.
- Well, the cname points at **frontus.secretmedia.com**
- Check **secretmedia.com** domain whois? It's also hidden behind a whoisguard.com Panamanian domain registration
- For one explanation of what secretmedia.com is all about (from their POV) see http://www.secretmedia.com/manifest.php
- See also http://www.wsj.com/articles/ad-blockers-internet-advertisers-play-cat-and-mouse-1437046675

What Other Newly Observed Hostnames Do We See For That Same Domain?

```
$ nmsgtool –C ch213 | grep rrname | grep
mucocutaneousmyrmecophaga.com | awk '{print $2}'
yhlwc986z4yszzj1xupg924qz1qimf.mucocutaneousmyrmecophaga.com.
enyprryjucfw26ljsn4lwuacxpzccv.mucocutaneousmyrmecophaga.com.
e4a87ycis42863h3c2u2xa8lpkb2ps.mucocutaneousmyrmecophaga.com.
yvgch69dv9qajpoi2pqa18wbakyif8.mucocutaneousmyrmecophaga.com.
wizfe7mz5xjfgk9e61663lhep2hw3u.mucocutaneousmyrmecophaga.com.
nclm4634l52121n9q5741s2122hbkr.mucocutaneousmyrmecophaga.com.
eevc4n7muat3l3eugdpginuyx4hedj.mucocutaneousmyrmecophaga.com.
i61qe9mwnp8wpdophb65l69x31hjoh.mucocutaneousmyrmecophaga.com.
hctd9o8dso1riw7wiee8f7x9ho5f6b.mucocutaneousmyrmecophaga.com.
5zzekpefdypvcz2m3flhcrymw2s5x8.mucocutaneousmyrmecophaga.com.
nfxxhx1wufyi9v5h4rhq8i39yph3op.mucocutaneousmyrmecophaga.com.
xmnwa3t3qzqlinhzv51xeboci7su66.mucocutaneousmyrmecophaga.com.
96w3x6j85ddjhjrddrbso6lddajntt.mucocutaneousmyrmecophaga.com.
[etc]
```

Extracting Just The First Part of the Host Name

```
$ nmsgtool –C ch213 | grep rrname | grep
mucocutaneousmyrmecophaga.com | awk '{print
$2}' | awk –F. '{print $1}'
[...]
iqgxfumrixxbebk6f1m9i2w5uxk5ue
1bi1mpjuaanhywj4i4hqmpbf3mf9e6
ss4mrvuy5o7qd4jupuiiq8illhojn7
kri2bzsuob4xsxiirt3p5vevma5th5
5hpvws65tfdmioerlloz2ax8m2rqpi
el9n21pnjoyys7w1md116qeszpusxx
x1wtto2z5m1iih4lldmd4k22ncxopg
k7y7brkvm7vlm7ilvd7yonioy321nd
[etc]
```

Uniformly Distributed Noise? No, I Don't Think So

3-way letter frequencies:

hqw => 426
bii => 315
ihq => 314
iih => 314
qwa => 170
qwb => 158
yhq => 117
uiy => 112
iyh => 112
qwd => 74
qbi => 55
2vi => 45
vin => 45
[etc]

4-way letter frequencies:

ihqw => 314
iihq => 314
biih => 314
hqwa => 169
hqwb => 157
iyhq => 112
uiyh => 112
yhqw => 112
hqwd => 71
qbii => 54
inli => 43
vinl => 43
2vin => 43
[etc]

5-way letter frequencies:

biihq => 314
iihqw => 314
ihqwa => 139
iyhqw => 112
uiyhq => 112
ihqwb => 104
ihqwd => 71
qbiih => 54
yhqwb => 53
vinli => 43
2vinl => 43
hqwds => 41
hqwas => 39
[etc]

6-way letter frequencies:

biihqw => 314
iihqwa => 139
uiyhqw => 112
iihqwb => 104
iihqwd => 71
qbiihq => 54
iyhqwb => 53
2vinli => 43
ihqwds => 41
ihqwas => 39
e2vinl => 39
ihqwbs => 37
ihqwbc => 37
[etc]

Finding Additional Recent frontus.secretmedia.com-Related Domains Using DNSDB

```
$ dnsdb_query.py -l 1000000 --after 1w \
-n frontus.secretmedia.com > secretmedia.txt
ocyfvybiek5pa6fyei4ifwdk5cynaeb2ycndbityhkekfi
rsfbravacqgdvn.f4da.odesaconflate.com. IN
CNAME frontus.secretmedia.com.
[etc]
```

786,213 results seen during last week:

773,690	mucocutaneousmyrmecophaga.com.
11,763	odesaconflate.com.
459	inosculationimmediately.com.
300	jitterblackhawk.com.
1	atomicpowerseductress.net.

This is NOT The Only Unusual Domain You'll See When Watching Newly Observed Hostnames

- There are many parties transmitting weird encoded information via DNS hostnames, to say nothing of weird content in TXT records. You normally don't have visibility into DNS, so you don't know this is happening, but it is. These odd names are easily observed if you look in the Newly Observed Hostnames channel.

- What are they?
  - Some may be advertising-related
  - Others may be botnet command and control-related
  - Still others may represent data exfiltration attempts
  - If you can't see them and don't know they exist, you'll never find out
  - **This deserves more careful scrutiny by the community**

III. DNS Changes

DNS Changes Channel

- Imagine you knew the current state of DNS for ALL names.

- Now imagine you could compare what you're seeing now to what you'd seen previously for those names, making note of any differences.

- You've just described the DNS Changes channel (simple to describe, non-trivial to implement at scale).

- DNS Changes can be thought of as the "general case" from more specific channels (such as Newly Observed Domains and Newly Observed Hostnames) can be derived.

```
$ nmsgtool -C ch214 -c 1
[72] [2015-11-01 18:54:29.900300025] [2:5 SIE
newdomain] [a1ba02cf] [] []
domain: leet.cc.
time_seen: 2015-11-01 18:52:43
bailiwick: leet.cc.
rrname: c50457.leet.cc.
rrclass: IN (1)
rrtype: A (1)
rdata: 209.126.99.231
new_domain: False
new_rrname: False
new_rrtype: False
new_rr: True
new_rrset: True
```

Sample Observation,
DNS Changes Channel

"leet.cc"? Sounds very "l33t" and thus potentially suspicious, but in fact they support micro MCPE servers (Minecraft Pocket Edition iOS and Android online games)

What Are The Fields In That Sample Record?

- **new_domain:** The base domain has never been seen before
- **new_rrname:** The FQDN has never been seen before
- **new_rrtype:** This is a new resource record type for this FQDN
- **new_rr:** This exact resource record has never been seen before
- **new_rrset:** This exact resource record set has never been seen before.

- ***Notes:***
  - If new_domain is true, {new_rrname, new_rrtype, new_rr, new_rrset} will all *also* be true
  - If new_rrname is true, {new_rrtype, new_rr, new_rrset} will *also* be true
  - If new_rrtype is true, {new_rr, new_rrset} will *also* be true
  - New_rrset will ALWAYS be true for data shown in the DNS Changes channel
  - Just interested in new_domain's? Check out *Newly Observed Domains*
  - Just interested in new_rrname's? Chek out *Newly Observed Hostnames*

```
$ nmsgtool -C ch214 -c 1
[83] [2015-11-01 19:14:15.760720014] [2:5 SIE
newdomain] [a1ba02cf] [] []
domain: lokvel[dot]ru.
time_seen: 2015-11-01 19:12:22
bailiwick: lokvel[dot]ru.
rrname: t1446405142[dot]lokvel[dot]ru.
rrclass: IN (1)
rrtype: A (1)
rdata: 88[dot]85[dot]80[dot]170
new_domain: False
new_rrname: True
new_rrtype: True
new_rr: True
new_rrset: True
```

BE CAREFUL with this defanged example: VirusTotal reports hits associated with that domain and IP address, see https://www.virustotal.com/en/ip-address/88.85.80.170/information/

```
domain: akadns.net.
time_seen: 2015-11-01 19:43:28
bailiwick: akadns.net.
rrname: dsn6.skype-dsn.akadns.net.
rrclass: IN (1)
rrtype: A (1)
rdata: 65.55.223.38
rdata: 111.221.77.153
rdata: 157.56.52.23
rdata: 157.56.52.31
rdata: 157.56.52.39
rdata: 157.56.52.42
rdata: 213.199.179.145
rdata: 213.199.179.170
new_domain: False
new_rrname: False
new_rrtype: False
new_rr: False
new_rr: False
new_rr: False
new_rr: False
new_rr: False
new_rr: False
new_rr: False
new_rr: False
new_rrset: True
```

A more complex resource record set... can you speculate why the rrset is flagged as "new?" (resource record ordering)

IV. What We Saw in "DNS Changes" For a 24-Hour Period

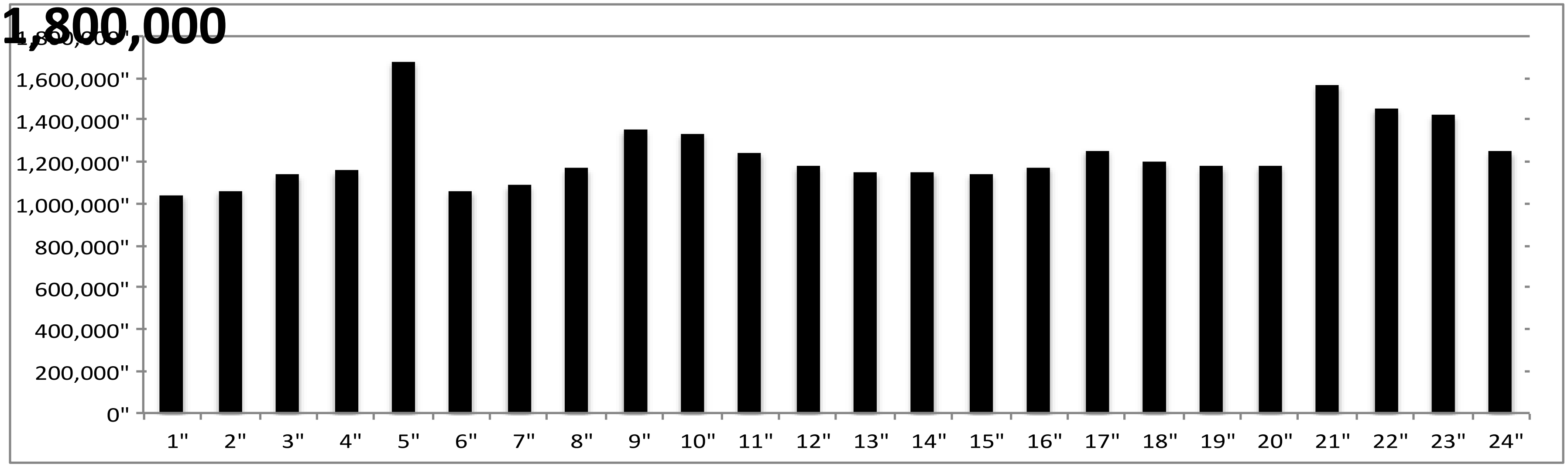
Rather than just look at anecdotal samples, we decided to systematically look at a day's worth of DNS Changes data.

How Many DNS Changes In 24 Hours?

- We collected data for a 24 hour period running from 0600 GMT on 2015/05/21 through 0599 GMT on 2015/05/22.

- **29,620,807 observations** in total were collected during that time.

DNS Changes Observed Per Hour



New Base Domains Seen In The 24 Hours of Data

- Out of 29,620,807 total observations, 176,366 (0.595%) represented new base domains.

- Assuming a uniform rate of new base domain generation, this would imply an annual rate of

  176,366 new domains/day*365.25 days/year=**64,417,681 new domains/year.**

  That rate is over three times higher than the net growth in new domains reported by VeriSign in "The Domain Name Industry Brief," https://www.verisigninc.com/assets/domain-name-report-january2015.pdf

Factors Potentially Driving That Difference in Rate

- **Farsight counts both newly registered domain names *plus* new domains created under effective top level domains as determined by the Public Suffix List (PSL). VeriSign does not count new domains created under a PSL "effective TLD."**
- Farsight's data focused solely on new domains (without subtracting expiring or deleted domains); VeriSign looks at net growth (domains added less domains expired or deleted).
- There may be genuine changes in the rate of domain name creation, potentially associated with the ICANN new gTLD program, or increased uptake of IDNs (but this seems doubtful)
- While Farsight has been tracking new domains since June 2010, we continually add new sensor nodes. Some newly-detected new domains may actually have previously existed, but may be just now seen by us as our monitoring continues to improve (again, this seems like an unlikely explanation for that volume)

Newly Observed FQDNs

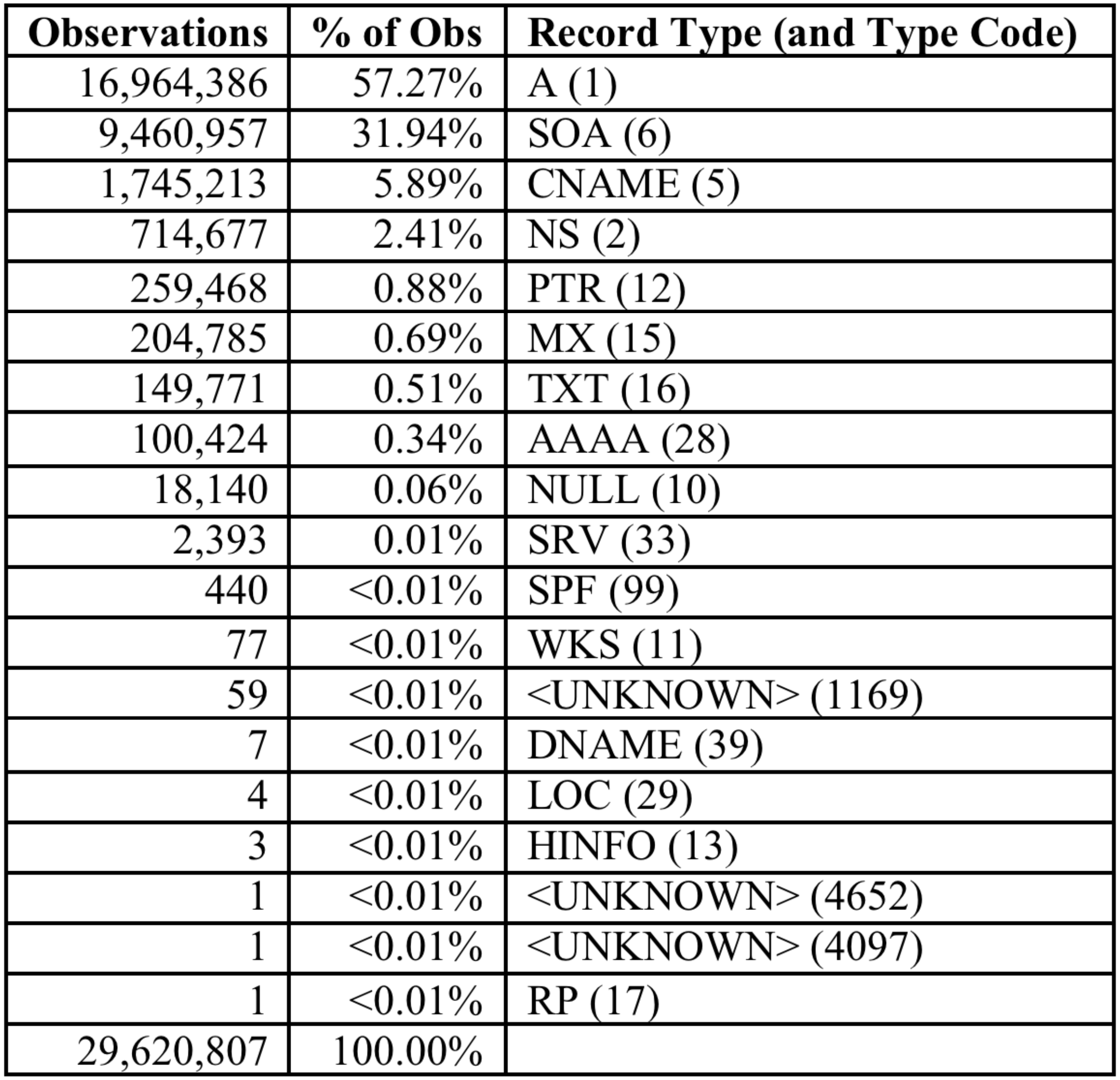
- 7,778,302 observations (26.26%) represented identification of new rrnames (e.g., new "hostnames"/new FQDNs).

- Extrapolating, this translates to an annualized rate of new FQDN creation of 7,778,302*365.25=**2.84 billion new FQDNs per year.**

  That's a **lot** of new hostnames. We believe that many of them may be randomized hostnames, or opaquely-encoded hostnames (like the mucocutaneousmyrmecophaga.com example previously described).

- We also looked at the distribution of record types... Most are "A" records, SOAs, or CNAMEs. See the table on the following slide.

Table 1: Observations broken down by resource record type



Observations	% of Obs	Record Type (and Type Code)
16,964,386	57.27%	A (1)
9,460,957	31.94%	SOA (6)
1,745,213	5.89%	CNAME (5)
714,677	2.41%	NS (2)
259,468	0.88%	PTR (12)
204,785	0.69%	MX (15)
149,771	0.51%	TXT (16)
100,424	0.34%	AAAA (28)
18,140	0.06%	NULL (10)
2,393	0.01%	SRV (33)
440	<0.01%	SPF (99)
77	<0.01%	WKS (11)
59	<0.01%	<UNKNOWN> (1169)
7	<0.01%	DNAME (39)
4	<0.01%	LOC (29)
3	<0.01%	HINFO (13)
1	<0.01%	<UNKNOWN> (4652)
1	<0.01%	<UNKNOWN> (4097)
1	<0.01%	RP (17)
29,620,807	100.00%	

High Frequency (base domain, type)-tuples

- The preceding details really don't tell you all that much about the data we see in DNS Changes.

- Succinctly yet comprehensively characterizing over 29 million observations poses distinct practical challenges, and any attempt will necessarily be incomplete.

- There are some obvious points that particularly merit comment.

CDNs and Location-Dependent Answers

- Many DNS names resolve consistently regardless of who's asking about them or where those questions may be originating.

- Some authoritative name servers, however, return different results for different query sources.

- For instance, in an effort to minimize query latency, some content distribution networks (CDNs) may intentionally provide different answers for a query depending on the query origin:
  - A query from a North American user may receive a result that will send that traffic to a nearby North American server for processing
  - Another query for the same domain name at the same moment in time (but from an Australian user) may return results for an Australian server, instead.

Farsight Sees Those Interleaved CDN Responses As Potential Changes

- Because we have over 500 sensor nodes collecting data from all around the world, those interleaved queries may appear to be "changes."
- This will be true even if the response that any individual user sees is utterly invariant over time, simply because the CDNs answer to a given query will vary depending on:
  - the source of that query, or
  - the query origin plus other factors (such as load balancing considerations).
- The names with the highest number of changes are, in fact, consistently associated with CDNs. The (rrname, rrtype) tuples that had the highest frequencies (over 100,000 records for the period of observation) were virtually all CDN-related...

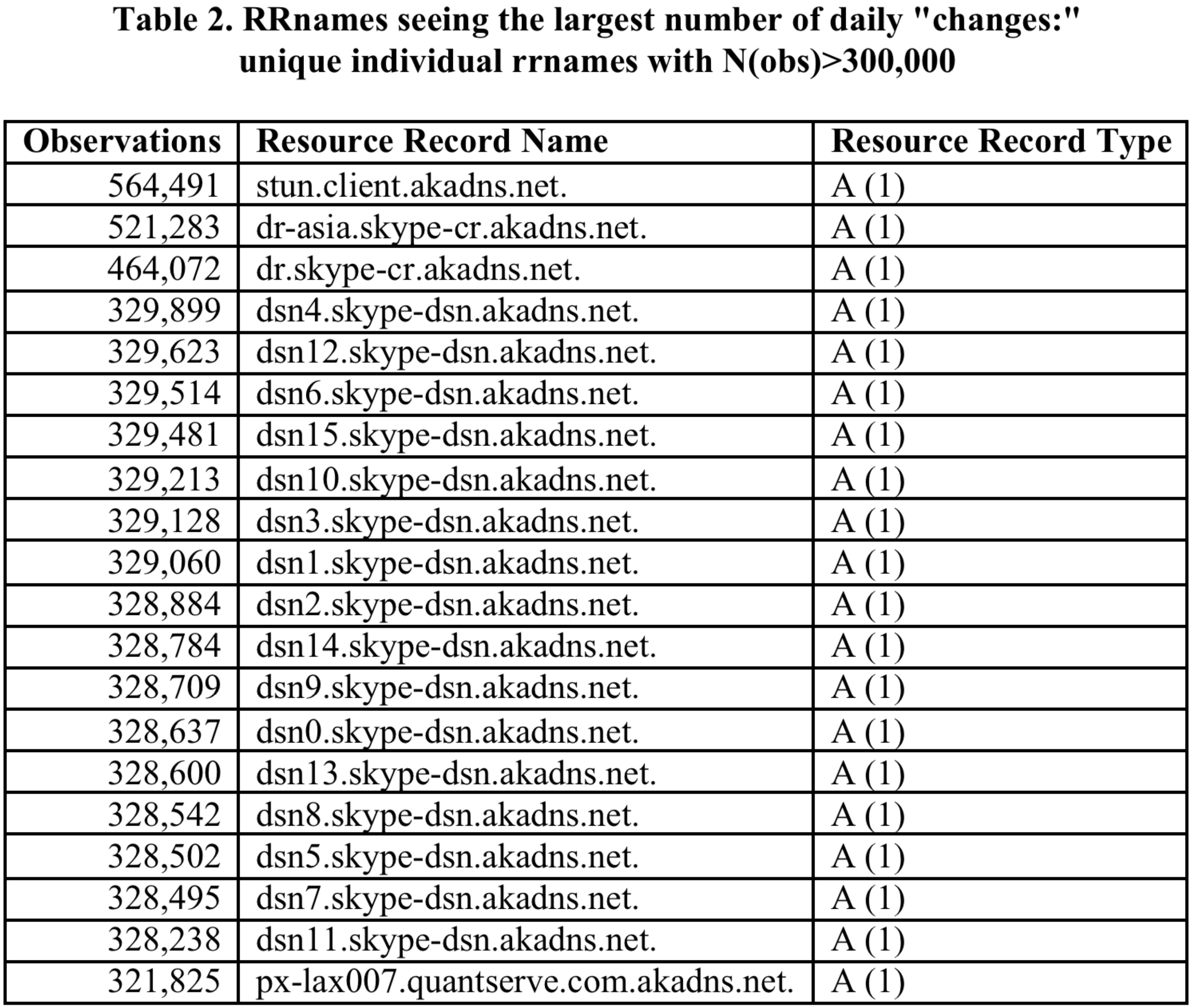


Table 2. RRnames seeing the largest number of daily "changes:" unique individual rrnames with N(obs)>300,000

Observations	Resource Record Name	Resource Record Type
564,491	stun.client.akadns.net.	A (1)
521,283	dr-asia.skype-cr.akadns.net.	A (1)
464,072	dr.skype-cr.akadns.net.	A (1)
329,899	dsn4.skype-dsn.akadns.net.	A (1)
329,623	dsn12.skype-dsn.akadns.net.	A (1)
329,514	dsn6.skype-dsn.akadns.net.	A (1)
329,481	dsn15.skype-dsn.akadns.net.	A (1)
329,213	dsn10.skype-dsn.akadns.net.	A (1)
329,128	dsn3.skype-dsn.akadns.net.	A (1)
329,060	dsn1.skype-dsn.akadns.net.	A (1)
328,884	dsn2.skype-dsn.akadns.net.	A (1)
328,784	dsn14.skype-dsn.akadns.net.	A (1)
328,709	dsn9.skype-dsn.akadns.net.	A (1)
328,637	dsn0.skype-dsn.akadns.net.	A (1)
328,600	dsn13.skype-dsn.akadns.net.	A (1)
328,542	dsn8.skype-dsn.akadns.net.	A (1)
328,502	dsn5.skype-dsn.akadns.net.	A (1)
328,495	dsn7.skype-dsn.akadns.net.	A (1)
328,238	dsn11.skype-dsn.akadns.net.	A (1)
321,825	px-lax007.quantserve.com.akadns.net.	A (1)

Frequently Updated/High Frequency SOA Records

- Another class of "frequently changing" observations consists of SOA (Start of Authority) records.
- SOA records are used in the DNS to specify the maintainer of a DNS zone, the zone's TTL values, and a serial number identifying the current version of the zone file.
- That serial number is normally incremented whenever the zone file is updated. That change in serial number is sufficient to trigger a change detection in the current dataset.
- Thus, it is not surprising that a second category of high frequency (rrname, type)-tuples is associated with SOA records for frequently updated zones.
- If a zone were to be updated every second, that would imply 60*60*24=86,400 changes (one for each second of the day). Some SOA records we observed approximate that value...

Table 3. RRnames seeing the largest number of daily "changes:" Top 15 unique individual (rrname, type=SOA) observations

Observations	Resource Record Name	Resource Record Type
83,093	akadns.net.	SOA (6)
82,739	g.akamai.net.	SOA (6)
82,716	g.akamaiedge.net.	SOA (6)
82,690	da1.akamai.net.	SOA (6)
82,688	w28.akamai.net.	SOA (6)
82,649	w29.akamai.net.	SOA (6)
82,635	b.akamai.net.	SOA (6)
82,622	w22.akamai.net.	SOA (6)
82,610	b.akamaiedge.net.	SOA (6)
82,608	g2.akamai.net.	SOA (6)
82,579	d.akamai.net.	SOA (6)
82,550	w27.akamai.net.	SOA (6)
82,546	g1.akamai.net.	SOA (6)
82,543	a.akamaiedge.net.	SOA (6)
82,528	w23.akamai.net.	SOA (6)

Base Domains With Many Unique rrnames

- Many domains do not show up in our DNS Changes dataset at all – they simply didn't have a change during the 24 hour study period.
- Others may show up once (for example, when moving from one provider's IP space to a new provider's IP space)
- Other observations consist of base domains that have **large numbers of unique rrnames**, where each of those new rrnames are only seen once. These use-it-once-and-never-again unique "disposable rrnames" may be a sign that DNS is being used as:
  - a tracking mechanism (e.g., for per-web-page or per-email-message tracking links),
  - a data exfiltration mechanism (e.g., for data-over-DNS surreptitious data transfers),
  - an anti-monitoring mechanism (e.g., to help keep any single FQDN appearing to be too "hot" or "active"), or
  - in some other unconventional manner.

Drilling Down On This Phenomena

- If we collapse observed rrnames using the Effective TLD/Public Suffix List, we are left with a list of most-frequently-observed base domains.
- We'll exclude:
  - Akamai-related domains (those have already been prominently featured in tables 2 and 3, above), plus
  - Amazon-related domains (another obviously-massive provider-at-scale), as well as
  - Some CDN-related domain names (other than Akamai) that might otherwise also have been included (e.g., cdn13.com, fbcdn.net, and cdn77.net).
  - We'll also exclude a domain associated with a site that's operating a sensor node for FSI SIE to avoid disclosing that data source, consistent with FSI's terms of service/non-disclosure agreement requirements.

Table 4. Selected Base Domains With Relatively Large Numbers of Observations

Observations	Base Domain	Privacy/Proxy Domain Whois?	Registrar of record
987,688	yahoodns.net	no	MarkMonitor
544,784	tumblr.com	no	MarkMonitor
373,145	telemetryverification.net	no	Tucows
263,528	rssing.com	no	Key Systems
260,600	parse.com	no	MarkMonitor
234,518	mgm86800.com	yes	Godaddy
210,266	mgm001.com	no	Godaddy
204,723	mgm86877.com	yes	Godaddy
199,374	mgm002.com	no	Godaddy
158,622	spotilocal.com	no	Domaininfo AB
130,961	surfeasy.mobi	no	Godaddy
94,776	adnxs.net	no	MarkMonitor
74,757	vkrugudruzei.ru	web-based whois	RU-CENTER-RU
66,913	ns1p.net	yes	Name.com
63,571	sekindo.com	no	Dyn.com
62,474	optinre.ru	"private person"	SALENAMES-RU
61,335	1drv.com	no	MarkMonitor
55,232	nessus.org	no	Network Solutions
54,505	spampoison.com	no	Enom
53,625	incapsecuredns.net	no	Godaddy
50,841	seagateshare.com	no	MarkMonitor

Table 4. (continued)

Observations	Base Domain	Privacy/Proxy Domain Whois?	Registrar of record
48,574	greatrelating.com	yes	Melbourne IT
47,683	worldssl.net	no	Enom
47,455	mailguard.com.au	web-based whois	Melbourne IT
46,723	spotify.com	no	Domaininfo AB
45,553	wd2go.com	no	CSC Corporate Domains
41,207	geoadnxs.com	no	MarkMonitor
40,368	dyndns.org	no	Tucows
37,704	imdb.com	no	MarkMonitor
36,034	emltrk.com	no	Enom
35,750	mgm86855.com	yes	Godaddy
33,913	bugun.in	no street address	Name.Com
32,557	websamsung.net	no	Whois Networks
31,232	audible.com	no	CSC Corporate Domains
30,757	notifygate69.ru	"private person"	R01-RU
30,026	notifygate72.ru	"private person"	R01-RU
29,871	notifygate70.ru	"private person"	R01-RU
29,782	notifygate71.ru	"private person"	R01-RU
29,346	notifygate73.ru	"private person"	R01-RU

Programmatic Name Generation

If we look at actual rrnames associated with one of those domains, such as notifygate69.ru, we see a pattern consistent with programmatic domain name generation:

```
abbadided.notifygate69.ru.
aberdeenn.notifygate69.ru.
accordanceg.notifygate69.ru.
accrescei.notifygate69.ru.
accruementg.notifygate69.ru.
addictivep.notifygate69.ru.
adjudgeri.notifygate69.ru.
adrenalonea.notifygate69.ru.
adventuredl.notifygate69.ru.
affirmablyy.notifygate69.ru.
afterpotentialt.notifygate69.ru.
aggravatives.notifygate69.ru.
agricolitec.notifygate69.ru.
airohydrogeny.notifygate69.ru.
aistopodesm.notifygate69.ru.
aluminasy.notifygate69.ru.
amalfianf.notifygate69.ru.
amaryllideousm.notifygate69.ru.
ambidextrousv.notifygate69.ru.
ambritev.notifygate69.ru.
amentiaq.notifygate69.ru.
amicablenessb.notifygate69.ru.
aminoaceticz.notifygate69.ru.
amortizingx.notifygate69.ru.
amphicondylac.notifygate69.ru.
amuttera.notifygate69.ru.
amygdalotomyt.notifygate69.ru.
amyloidala.notifygate69.ru.
[etc]
```

notifygate69.ru has been listed in the Spamhaus Domain Blocklist

Domain Names Of Special Security Relevance

- Access to a dataset with details about DNS changes also can be useful for identifying new special domain names, such as domains that may be associated with brand infringement or phishing.
- For example, while "paypal.com" is the well known brand of a leading online payments company, a number of other rrnames also incorporate that brand name.
- Looking for the string "paypal" in rrnames seen in the DNS changes data, we saw 274 unique domain names incorporating that term, besides paypal.com itself.
- We wouldn't want anyone to accidentally visit these dubious sites, so we've reversed a particularly dubious-looking sampling of those rrnames and replaced the dots in those names with [dot] for additional protection against accidental visits....

- bg [dot] paypal-topup
- br [dot] com [dot] construcasarj [dot] notice-of-changes-to-the-paypal-user-agreement
- co [dot] iinc [dot] paypal [dot] limited [dot] secure [dot] login
- com [dot] 4paypal
- com [dot] actiumweb [dot] solmallorca2 [dot] paypal-account-confirmation
- com [dot] activated-paypal-cash
- com [dot] appsmyway [dot] paypall [dot] account [dot] update
- com [dot] confirmpaypalls
- com [dot] confirm-your-account [dot] verification [dot] paypal
- com [dot] contactservicepaypal
- com [dot] customer-paypal-update
- com [dot] engineerslabltd [dot] informations [dot] verification [dot] paypal-community [dot] www
- com [dot] mashangqifei [dot] update-your-information-paypal-account
- com [dot] myskmg [dot] update-paypal-account
- com [dot] myzazzlestore [dot] paypal-update-your-info-2015-2016-security [etc]

Black Hat Sound Bytes

Takeaway #1:

It's now possible to watch new Fully Qualified Domain Names as they first get used in real time. Real-time FQDN data is particularly useful for spotting **malicious** FQDNs, such as phishing sites or brand infringement sites.

Takeaway #2:

You can also watch changes as they get made to existing domain names, literally at the level of individual hostnames. These changes may be innocuous, or problematic changes needing resolution.

Takeaway #3:

There are many odd activities occurring in the DNS that can be easily spotted, if people just look. We'd love to see more academic researchers looking at real-time DNS Changes data at the SIE (and grant awards are available to support researcher access of that sort)

FARSIGHT SECURITY

Q&A

Would folks like to see some live data from the new SIE channels?

For more information:

Dr. Paul Vixie

vixie@fsi.io

https://www.farsightsecurity.com

Image Credits

- Cockpit instruments https://commons.wikimedia.org/wiki/Category:Aircraft_instrument_panels#/media/File:Hawker_Siddeley_Nimrod_MR2_%28801%29,_UK_-_Air_Force_AN0808358.jpg

- Cascade Mountains Poking Through the Clouds https://en.wikipedia.org/wiki/Cascade_Range#/media/File:Mount_Rainier_and_other_Cascades_mountains_poking_through_clouds.jpg

- Picture of Cyclone Phailin https://en.wikipedia.org/wiki/Cyclone_Phailin#/media/File:Cyclone_Phailin_11_October_2013.jpg