Software Engineering Institute

Passive Detection of Misbehaving Name Servers

Leigh B. Metcalf
Jonathan M. Spring

October 2013

TECHNICAL REPORT
CMU/SEI-2013-TR-010
ESC-TR-2013-010

CERT® Division

http://www.sei.cmu.edu

Carnegie Mellon

This material is based upon work funded and supported by Department of Homeland Security under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center sponsored by the United States Department of Defense.

This report was prepared for the
SEI Administrative Agent
AFLCMC/PZM
20 Schilling Circle, Bldg 1305, 3rd floor
Hanscom AFB, MA 01731-2125

DM-0000645

Table of Contents

List of Figures

List of Tables

Abstract

In the process of categorizing malicious domains, distinguishing between suspicious and benign name servers can allow the name servers themselves to be acted against. Name servers do not normally change internet protocol (IP) addresses frequently. Domains that do change IP addresses quickly or often are said to exhibit IP flux, which can allow services, such as web pages that deliver malicious content, to circumvent defenders' attempts to block their IP addresses. IP flux in a name server's domain may be a sign that the name server is suspicious. This report demonstrates that name-server flux exists and is ongoing. Furthermore, there are two types of data that can reveal IP flux in domain name system (DNS) servers: passively collected DNS messages and the contents of several large, top-level domains' official zone files.

1 Introduction

Detecting malicious domains is becoming an important task in limiting the impact of ne'er-do-wells on the internet. It has become a race between defenders developing new detection methods and adversaries developing new evasion methods. Domains are required to have only a few properties, so malicious domains can leave few traces. By focusing on the characteristics a domain must have in order to operate, such as a name server, security personnel can limit the ability of malicious domain controllers to avoid countermeasures. The domain name system (DNS) requires only two associations for a domain: its location and whom to ask about its location. Network administrators and security personnel have pursued restricting location, mainly via blocking internet protocol (IP) addresses, with some success. Blocking name servers, the entities that provide location information about domains, has been pursued with much less energy.

In the process of categorizing malicious domains, distinguishing between suspicious and benign name servers can allow the name servers themselves to be acted against, given sufficient evidence. Because name servers are one of the primary components of a well-functioning DNS, they do not change IP addresses frequently. Domains that do change IP addresses quickly or often are said to exhibit IP flux, which can make services, such as web pages that deliver malicious content, more resilient by circumventing defenders' attempts to block their IP addresses. IP flux in a name server's domain may be a sign that the name server is suspicious.

1.1 Related Work

In early 2008 the Internet Corporation for Assigned Names and Numbers (ICANN) Security and Stability Advisory Committee published an advisory detailing the existence of fast-flux networks [ICANN 2008]. This advisory detailed flux in domains based on DNS record content, a phenomenon commonly called fast flux, and in the name-server infrastructure used to support those domains. The ICANN report identified fast flux as malicious and stated that fast-flux hosting "is considered one of the most serious threats to online activities today" [ICANN 2008, p. 2]. The advisory did not qualify the maliciousness of fast flux, nor did it note any benign use cases for the activity.

Significant work has investigated how to find and block domains that demonstrate malicious activity. Some of these efforts are operational, and some are still in the research phase. Most specialize in a particular brand of malicious activity. Some operational blocking lists conceal their selection criteria to prevent adversaries from exploiting the rules. Some notable operational lists include Spamhaus, PhishTank, and Google Safe Browsing. Spamhaus maintains a few operational lists, each targeting aspects of malicious email [Spamhaus 2011]. PhishTank publishes lists of phishing URLs that are consumed by popular browsers to block phishing pages from reaching users. PhishTank takes a community-based approach, providing a site where "anyone can submit, verify, track and share phishing data" [PhishTank 2011]. Google also derives lists of phishing and malicious sites while it crawls the web, and it makes these lists accessible to the public using an application programming interface (API) [Google 2011]. While the effectiveness of blocking phishing seems to vary [Rasmussen 2010, 2011; Spring 2010], efforts to take down or block phishing sites have been demonstrated to shorten their lifetimes [Moore 2007].

Several papers have described blocking techniques that use passive DNS (pDNS). Some techniques also use zone files, the official lists of domain-to-name-server and name-server-to-IP-address mappings maintained by the registry. Antonakakis and colleagues developed a reputation-based classification system, called Notos, that uses pDNS monitoring data [Antonakakis 2010]. The Notos classification scheme divides its decision-making criteria into the broad categories of network-based, zone-based, and evidence-based. A similar team expanded these efforts with Kopis [Antonakakis 2011]. Bilge and colleagues designed the EXPOSURE system, which also uses pDNS as the data input. EXPOSURE introduces features based on time series and time to live (TTL) [Bilge 2011]. Others have described active detection of fast-flux domains [Hu 2011], and later behavioral analysis of the fast-flux networks has independently touched on name server use [Kadir 2012].

There have also been efforts to use URLs culled from spam traps, combined with active DNS behavior and registration information, to describe some properties of malicious domains [Hao 2011] as well as URL properties, excluding page content [Ma 2009a, 2009b]. Additionally, Felegyhazi, Kreibich, and Paxson used zone files to predict which domains would be used maliciously, based on the previous evidence of malicious activity by other domains using the same name server [Felegyhazi 2010]. Stoner finds malicious activity, specifically malicious fast flux and domain flux, using simpler methods [Stoner 2010]. That research uses only two features of a domain: the IP addresses it maps to and the associated autonomous system numbers (ASNs) in which the IP addresses reside.

1.2 Motivation

The ultimate goal of the current work is to hinder criminals' free use of domain names as an accessory to their crimes. Except for Felegyhazi, Kreibich, and Paxson's work [Felegyhazi 2010], current techniques can only reactively hinder criminals. All the operational lists currently in use (see Section 1.1) are reactive, so they can at best take away names only after some damage has been done. This is important in limiting damage, but it is not ideal. With some malicious use cases such as spam, as many as 55% of domains may be used within one day of registration; these domains also tend to be hosted on name servers, which are detectably different from the average DNS infrastructure [Hao 2011]. The IP flux detection described in this report complements the malicious name-server detection methods described by Hao, Feamster, and Pandrangi [Hao 2011].

Detection via name-server behavior can improve the current state-of-the-art deterrence because it preempts at least some domains and is comprehensive in that all domains must be served by a name server. The particular aspect of name-server behavior we measure is name-server IP flux. This technique is independent of the particular use of the malicious domains hosted on the name server, so it complements techniques that first identify malicious domains by a specific use case (spam body URLs, command and control, etc.) and then identify their name servers.

In addition, it is an important contribution simply to report on the prevalence of fast-flux name servers. Active classification, such as reported by Hu, Knysz, and Shin [Hu 2011], is insufficiently scalable for a global view. We employ passive techniques that permit analysis of hundreds of millions of domains, rather than thousands. The 2008 ICANN advisory spurred measurement of the domain fast-flux phenomenon and policy and technical recommendations by ICANN [Konings 2009]. However, previous reporting did not provide measurements of name-server flux. This report provides such measurement of name-server flux.

1.3 Data Sources

The general categories of data sources used in this work are top-level domain (TLD) zone files and pDNS traffic. The contents of zone files are generally reported to the registry by the registrars. New generic TLD (gTLD) operators are required to make this file available under certain conditions [ICANN 2011]. Other TLD operators do not operate under the same contract with ICANN and do not generally make their zone files available to anyone. Because gTLD files are available, we demonstrate our analysis using the *com*, *org*, *net*, *biz*, *info*, and *mobi* zone files. Zone files have a required form [Mockapetris 1987], so this analysis is applicable to any other TLD's available zone file. Not having visibility into any country-code TLDs (ccTLD) does bias the data set. However, *com* is by far the largest TLD and represents a large percentage of the total domains for the internet [Stoner 2010]. VeriSign's terms of use limit zone file downloads to one every 24 hours, which limits the granularity of the detectable changes in this analysis.

Passive DNS collection was first described by Weimer [Weimer 2005]. We use a large pDNS source, the Security Information Exchange (SIE), for pDNS data. While the coverage of the SIE sensor array is incomplete and biased, there is evidence that it is wide, and it processes many tens of millions of distinct messages per day [Spring 2011]. The SIE is the best generally available source of pDNS data. The SIE data is delivered in resource records (RR) sets. One record set is all the resource records in a single message that share record name, class, type, and TTL, with the record data sorted by the standard DNS ordering and stored as the final fields of the record set:

- *name*—a unique identifier, and the subject of the record; it can specify a single host or a zone under which there are more names.
- *class*—identifies a set of types; in practice the only value observed is IN for the internet.
- *type*—specifies the expected data type and format, such as the IP address(es) of the name, the mail exchange to use to contact that domain, or the name server to ask for more information about the name
- *TTL*—specifies the number of seconds the information in the record should be treated as valid, after which the machine must ask for a fresh version of the data

Routers use Border Gateway Protocol (BGP) information to find paths to IP addresses throughout the internet by way of associated ASNs. A digest of this data can associate observed IP addresses with ASNs. This is valuable because ASNs represent blocks of effective control and help distinguish organizational boundaries. In our analysis, we used the University of Oregon Route Views Project merged with data from the Réseaux IP Européens Network (RIPE) Routing Information Service to associate IP space with ASNs [Route-Views 2005, RIPE 2012].

2 Method

We collected various characteristics of name-server behavior during the four weeks from June 12, 2011, through July 9, 2011. These are the dates for which we report full results and demonstrate our method. We have also collected summary results from October 2011 through September 2013 in essentially the same manner but with more condensed reporting.

Active Name Servers

The number of unique active name servers per day from both data sources is calculated. This is a straightforward counting operation on the data from each day, as well as a count of unique entries for the total duration of the observation period.

Zone File Name-Server IP Changes

The zone files (from the zones named in Section 1.3) contain the name servers' IP addresses. Each day's file is processed to make a list of unique name server–IP pairs. The name servers present on the first day of the observation period are checked for movement throughout the observation period. The unique addresses for these name servers are collected from the data. The number of changes, both per server and by all servers, is then analyzed.

Changes in Name Server A Records

The IP address of a name server is reported through the DNS just like other IP addresses. A name server may be the answer to a name-server query, but it is also the subject of A records. Using the pDNS data, changes in these A records can be detected. First, a list of name servers for a day is calculated from the payload of all the unique name-server RR sets. The A record sets for these names are then extracted from the data. This process yields a list of name-server-to-IP mappings for each day similar to that derived from the zone file as described above. The list of names for the first day in the observation period is then compared to the list for each other day to determine what name servers have changed IP addresses. Because this analysis consumes live data, a particular name server might not be requested every day, unlike the zone data, which is static data available on request. However, changes will by definition not be cached in the DNS, so we should not miss any relevant data points (in this analysis, changes in the value of the name server's location).

Distribution of TTLs in Name-Server Records

The distribution is calculated using the pDNS data. All unique RR sets for each day in the measurement period are stored. See Section 1.3 for a description of the pDNS RR set format. The calculation operates on the unique name-server-type (NS) RR sets observed per day. The TTL value is extracted from each unique name-server record set, and the instances of each value are counted. We provide some statistical evaluation of this data.

Name-Server IP Flux

There are several previously documented methods of detecting IP flux in pDNS traffic by using A records [Passerini 2008, Stoner 2010]. It does not seem that these methods have been applied to zone file behavior. Given a name server *N*, the following values are calculated to determine if its IP addresses exhibit flux:

- the number of unique IP address values *N* uses during the observation period
- the number of ASNs to which those IP addresses belong (derived from the BGP mapping described in Section 1.3)

The IP addresses of name servers should have a different variability than those of domain names, so the threshold parameters for name-server flux are different than those for traditional IP flux. The value of these parameters is derived from the contextual data and an understanding of the common practices of name-server administration. Because the server needs to be well known for reliable zone operation, it should not change IP addresses frequently. We are limited to a resolution of one measurement per day due to the frequency we are permitted to download the zone file data.

3 Results

Table 1 displays statistics for the scale of the study. The number of name servers on the first day is relevant because those name servers were the ones checked for movement throughout the observation period. The ranges exclude outliers, whose data was not properly collected due to technical errors. The dates of technical errors are June 13 in the pDNS data and June 16, June 20, and July 4 in the zone files. The total number of distinct names is the count of unique names observed over the 28-day observation period.

Table 1: Unique active name servers (NS) per day and total, June 12–July 9, 2011

	SIE	Zone files
First day (June 12)	1,926,367	2,796,128
Average # of name servers	1,968,271	2,786,279
Range of # of name servers	$1.8*10^6$ to $2.1*10^6$	$2.4*10^6$ to $2.8*10^6$
Total distinct observed	4,021,151	3,260,648

3.1 Zone File Data

In the zone files, 61,801 name servers changed IP address at least once. Of these, 41,796 changes included at least one change in the ASN in which the IP address was located. Table 2 breaks down these results per given number of changes observed over the 28 days. The maximum number of changes observed, in both IP and ASN, is 24 times in 24 measurements. This is a lower bound forced by the limitation of our observation frequency to once every 24 hours for zone files; some of these name servers may have changed more frequently.

Table 2: Number of name servers that change IP address and ASN a given number of times in 28 days, as determined from zone files (of 2,796,128 name servers)

Number of changes	NS changes IP	% of total	NS changes ASN	% of total
0	2,734,327	97.8%	2,754,332	98.5%
1	52,741	1.9%	36,645	1.3%
2	4,855	0.2%	1,846	0.1%
3	551	0.0197%	635	0.0227%
4	198	0.0071%	838	0.0300%
5	233	0.0083%	531	0.0190%
6	482	0.0172%	500	0.0179%
7	660	0.0236%	401	0.0143%
8	706	0.0252%	224	0.0080%
9	607	0.0217%	30	0.0011%
10	478	0.0171%	19	0.0007%
11	138	0.0049%	9	0.0003%
12	35	0.0013%	14	0.0005%
13	16	0.0006%	20	0.0007%
14	11	0.0004%	8	0.0003%
15	9	0.0003%	16	0.0006%
16	6	0.0002%	3	0.0001%
17	4	0.0001%	5	0.0002%
18	5	0.0002%	4	0.0001%
19	5	0.0002%	5	0.0002%
20+	61	0.0022%	43	0.0015%

3.2 SIE Passive DNS Data

Table 3 displays data equivalent to that in Table 2, except it is from a different data source: pDNS data. However, these data are not simply comparable. The pDNS data contain a different sampling because they can include all TLDs and name servers for domains that are not exclusively second-level domains. The pDNS data also have a finer temporal resolution, so their domains may exhibit more changes per day. The most variable name server changed its IP address 82 times and its ASN 71 times during the 28-day measurement period. Otherwise, the values are calculated with the same methodology.

Table 3: Number of name servers that change IP address and ASN a given number of times in 28 days, as determined from pDNS data (of 1,926,367 name servers)

Number of changes	NS changes IP	% of total	NS changes ASN	% of total
0	1,846,152	95.8%	1,877,654	97.5%
1	68,401	2.4%	40,422	1.4%
2	5,134	0.2%	3,276	0.1%
3	1,420	0.0508%	1,232	0.0441%
4	1,177	0.0421%	966	0.0345%
5	1,123	0.0402%	684	0.0245%
6	566	0.0202%	450	0.0161%
7	535	0.0191%	388	0.0139%
8	439	0.0157%	279	0.0100%
9	322	0.0115%	220	0.0079%
10	248	0.0089%	152	0.0054%
11	140	0.0050%	76	0.0027%
12	75	0.0027%	46	0.0016%
13	47	0.0017%	35	0.0013%
14	20	0.0007%	37	0.0013%
15	23	0.0008%	30	0.0011%
16	33	0.0012%	37	0.0013%
17	31	0.0011%	31	0.0011%
18	34	0.0012%	30	0.0011%
19	23	0.0008%	19	0.0007%
20+	424	0.0152%	303	0.0108%

Table 2 and Table 3 represent two related but different measurements. Each IP address change may or may not change the ASN the name server is in. ASNs are areas of logical control over internet routing, so a change in the ASN generally means the resource is changing areas of control. Table 2 and Table 3 do not specify a relationship between IP address changes and ASN changes. They simply report the number of name servers exhibiting a given number of changes over the observation period.

Sometimes when a name server changes IP it does not also change ASN. Figure 1 displays the relationship between number of times a name server changed IP and the number of different ASNs those IPs were in. The figure displays this relationship as a ratio between the number of different ASNs and IPs, with the number of unique IPs on the *x*-axis and a function, either minimum or mean, of the unique ASNs on the *y*-axis. Each data point is either the mean or minimum number of unique ASNs for all names that were observed to have *x* unique IPs. The maximum value is 1, in which each IP address caused one change in ASN in every name. This relationship explains why, in some cases in Table 2, there are more name servers that change ASN a given number of times than change IP that many times. Some subset of the servers that changed IP five times, for example, would have changed ASN three or four times.

One important observation from Figure 1 is that the majority of times a name server changes IP address it also moves to a network controlled under a different ASN, although not always. The

largest number of IP changes that are observed for a name server that stays within the same ASN is 9 in the zone files and 15 in the passive feed.

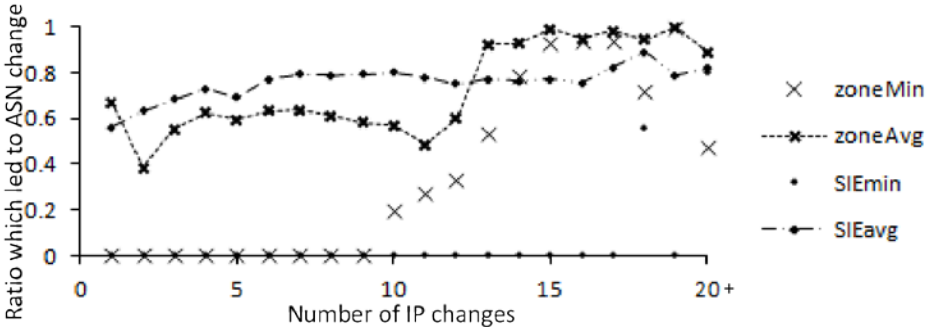

Figure 1: The average and minimum number of times an IP address change of a name server also changed the ASN, normalized by the number of IP changes. Each point is binned by IP change frequency for each data source. The maximum for both sources is 1 for all bins.

Table 4 presents the results for the distribution of TTLs in name-server record sets over the four weeks of observation. They demonstrate a human predilection for round numbers: one hour, one day, and two days are inordinately common. This can make the results difficult to display. Table 4 displays the percentage of record sets with TTL values within ranges determined by key TTL values, which are TTL values that represent more than 0.1% of all name-server record sets. Single TTL values that exceeded the 0.1% threshold became their own key values. Figure 2 summarizes these values and displays the broader trends. The log-log plot makes several groupings obvious: the very popular, 1% to 5%, 1% to 0.1%, and the rest.

Table 4: Distribution of TTLs in name-server record sets (760,796,769 total record sets)

TTL value	% of name-server RR sets	TTL value	% of name-server RR sets
0–29	0%	25,921–28,799	0.022%
30	0.110%	28,800	0.413%
31–59	0.004%	28,801–36,000	0.103%
60	0.156%	36,001–38,399	0.020%
61–120	0.112%	38,400	0.393%
121–299	0.042%	38,401–43,199	0.046%
300	0.937%	43,200	1.518%
301–599	0.096%	43,201–54,973	0.100%
600	0.811%	54,974–66,709	0.100%
601–899	0.033%	66,710–75,901	0.100%
900	0.626%	75,902–84,184	0.100%
901–1,199	0.025%	84,185–86,072	0.100%
1,200	0.134%	86,073–86,399	0.035%
1,201–1,799	0.034%	86,400	28.721%
1800	0.431%	86,401–89,999	0.031%
1,801–3,599	0.093%	90,000	0.103%
3,600	12.468%	90,001–116,764	0.100%
3,601–7,199	0.088%	116,765–146,302	0.100%
7,200	2.874%	146,303–171,528	0.100%
7,201–10,799	0.082%	171,529–172,799	0.019%
10,800	1.746%	172,800	41.231%
10,801–14,399	0.038%	172,801–259,199	0.022%
14,400	2.579%	259,200	0.573%
14,401–17,999	0.066%	259,201–345,599	0.019%
18,000	0.140%	345,600	0.833%
18,001–21,599	0.028%	345,601–604,799	0.082%
21,600	0.857%	604,800	0.248%
21,601–25,920	0.123%	604,801–429,4944,960	0.035%

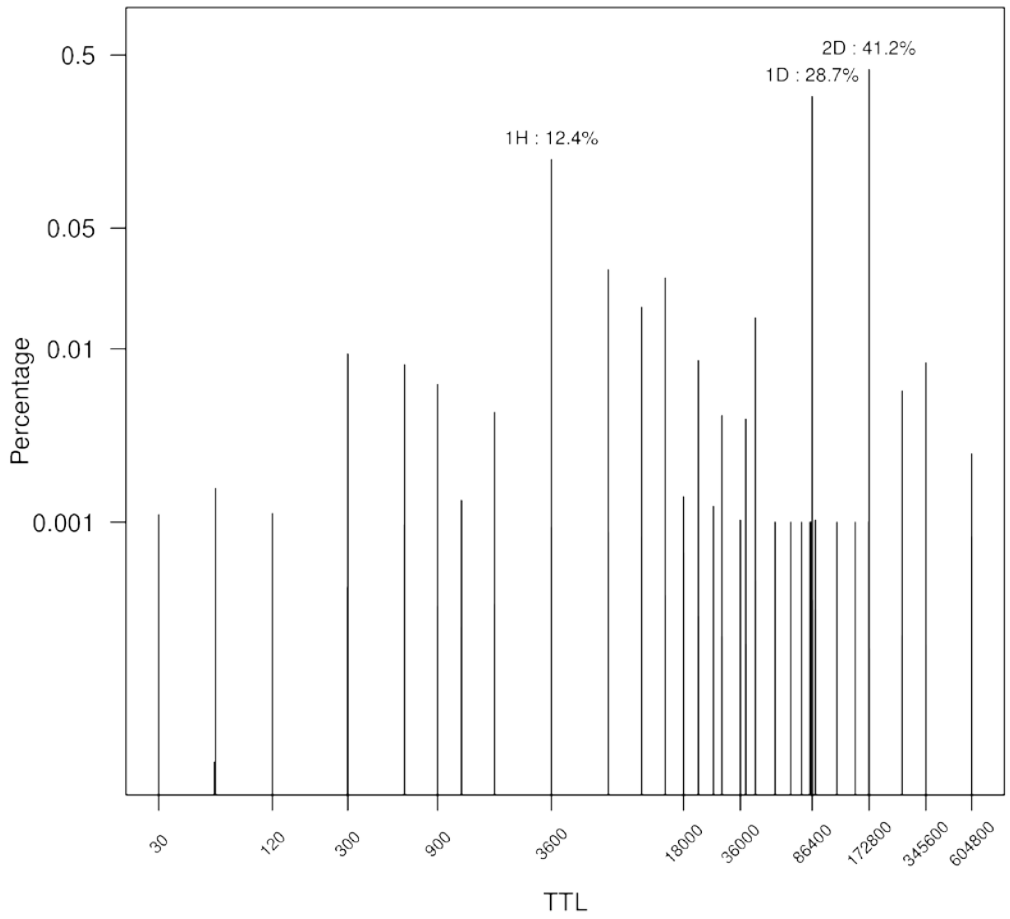

Figure 2: Log-log plot of observed name-server record TTL value frequency out of total name-server records observed. There are 760,796,769 total records represented. The figure does not show an outlier: 0.035% of the records had a TTL value greater than 604,800 and less than the maximum value of 4,294,944,960.

3.3 Ongoing Behavior

After our initial investigation, we watched the name-server IP-flux behavior for nearly two years, from October 2011 to September 2013, using just the zone file data. The high number of fluxing name servers continued into November 2011 but then fell off.

We investigated this drop in an attempt to find an explanation. At that time, apparently independent of filters on name-server flux, a large number of pharmaceutical scams[1] were taken down. These scams apparently were the primary users of name-server IP flux because once these

[1] We rather naively determined pharmaceutical scams by related strings in their domain names. In Figure 3, any NS domain name that contained one of the strings "HEALTH|PHARM|HOSPITAL|DRUG|PRESCRIPTION|PILLS|WELLNESS|TABLET|MEDS|RX|CIALIS|VIAGRA|MEDIC|PATIENT|PILL|DIET|CLINIC|NURSING|LEVITRA|WELNESS" was labeled as pharma-related. These strings were selected according to human expert decisions.

name servers went down, the number of fluxing name servers in the zone files dropped precipitously. However, it has remained at a basal level ever since, and there are clearly still miscreants using this technique. Figure 3 demonstrates this drop and the continuing lower level of suspicious name servers over the last two years.

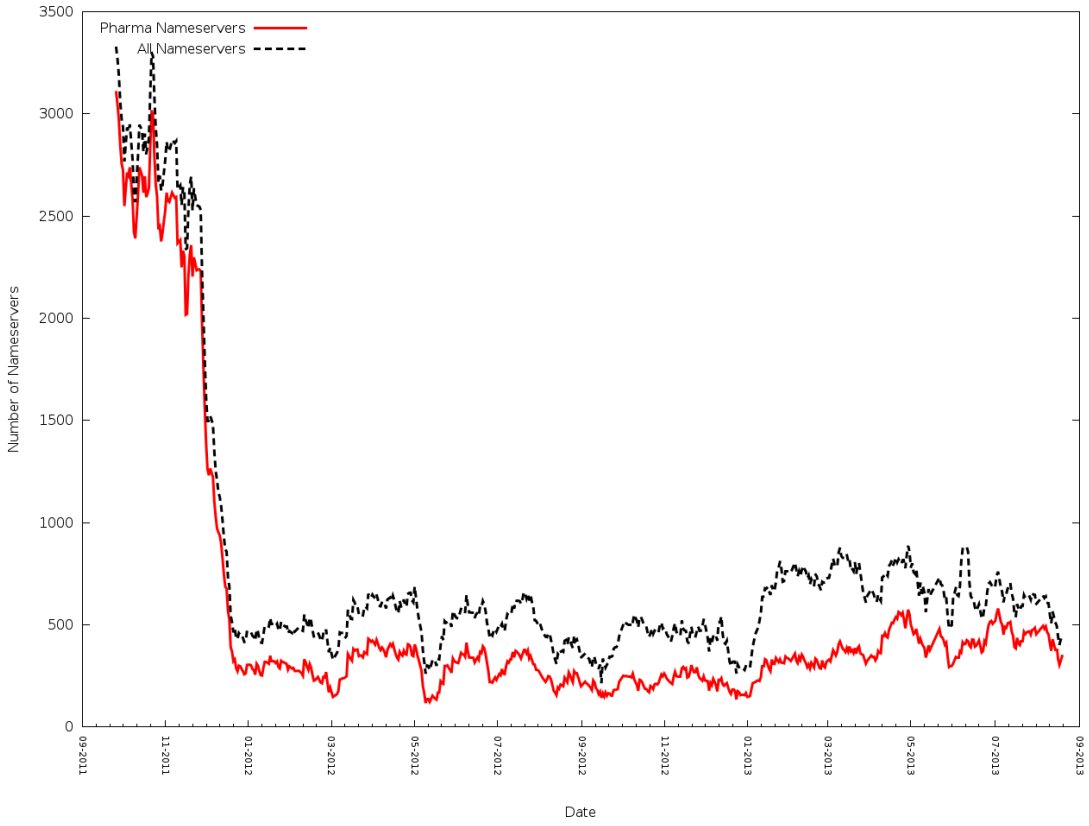

Figure 3: Number of name servers that changed IP address five or more times in a month. Solid red line indicates those servers possibly linked to pharmaceutical scams.

4 Discussion

The first notable conclusion drawn from the collected data is that name-server flux is certainly alive and well. More than 41,000 name servers changed IP address and the ASN in which the name server was hosted. This represents a problem because name-server flux does not have any benign use cases.

It could be argued this behavior indicates content distribution network (CDN) behavior, which would produce a high false-positive rate. We believe this to be misleading for two reasons. First, CDNs tend to own all of the networks on which they host content. When an organization owns a network, it is administrated as one ASN or perhaps a few. If CDN resources change IP address often, they very rarely change ASN. Second, the resources on a CDN that change IP address via the DNS are the resources to be delivered to the final consumer. The CDN wants the consumer to be able to ask a reliable, stable location—the name server—for the variable location of the resources.

The name-server record TTL distribution is not surprising. People will naturally gravitate toward common values: two days, one day, one hour. These three values account for more than 82% of the observed values. Variation from this pattern is not evidence enough to imply anything in particular, but coupling variation with some of the other measurements might be.

The distribution of TTL values in name sever records is also significantly different from that in records generally. The most common TTL in all records is much lower [Huth 2012], and our research confirms that unsurprising notion. Because TTL affects caching, and caching affects the number of messages that are sent and recorded, this difference matters to analysis techniques. Specifically, low-TTL domains may be over-represented in the passive data feed simply because they have more opportunities to change; however, because the zone file is authoritative, it does not suffer from this concern. Though some machine learning algorithms use TTL to determine maliciousness, these results present two clear conclusions. First, name servers have different TTLs than other domains, so an algorithm tuned to detect malicious, general domains may fail to detect malicious name servers (though TTL value is only one of many causes). Second, because TTLs are largely binned into only a handful of popular values, it would be easy for malicious actors to blend their domains on this axis. For these reasons, we do not propose the TTL value as a measure of maliciousness. Instead, we propose IP and ASN flux over time as the proper measure of maliciousness.

At some point, malicious name servers need to be taken down, and domains that use them should be blocked. The community must determine the threshold of acceptable damages before action is taken. The 2011 rate of fluxing name servers was 41,000 out of 2,790,000, or 1.5% in the zone files and 2.5% in pDNS resolutions. The zone file results indicate that a noticeable number of name servers are suspicious, more than the number of single-flux domains reported in 2009 [Konings 2009]. This surge appears to have been due to one particular scam campaign, which has since subsided, that made extensive use of the name-server flux technique. However, the technique remains a threat because the zone files have continued to show the behavior, though at a lower level, for the past two years.

5 Future Work

The precise extent to which known-malicious and benign domains use name servers that exhibit flux would be valuable information in determining the precise value and cost of blocking fluxing name servers. A complete evaluation of such behavior is not possible. However, estimations using existing lists of malicious domains should provide the necessary information.

Even though malicious actors must have a name server, they need not operate a second-level domain. Providers of free DNSs frequently give away domain names below domains they own, a service often called *dynamic DNS*. We cannot expect to obtain the zone files from these providers. A future analysis should quantify the extent to which domains using these services are disproportionately malicious. Our cursory investigations suggest that the extent is large. If so, automatic identification of these dynamic DNS services would be a useful area for future work. The list of known-malicious domains against which we test our methods should be extensive and should include multiple sources of known-malicious domains because no one list is comprehensive.

It would be beneficial to expand this study beyond the generic TLDs it investigated. How country-code TLDs would differ from gTLDs is not known, nor are the differences between gTLDs and dynamic DNS domains, so a simple extrapolation from the comparison between gTLD passive results and zone file results would not be sound.

References

URLs are valid as of the publication date of this document.

[Antonakakis 2010]
Antonakakis, M.; Perdisci, R.; Dagon, D.; Lee, W.; & Feamster, N. "Building a Dynamic Reputation System for DNS." *Proceedings of the 19th USENIX Security Symposium.* Washington, DC, Aug. 2010. http://www.usenix.org/events/sec10/tech/full_papers/Antonakakis.pdf

[Antonakakis 2011]
Antonakakis, M.; Perdisci, R.; Lee, W.; Vasiloglou, N., II; & Dagon, D. "Detecting Malware Domains at the Upper DNS Hierarchy." *Proceedings of the 20th USENIX Security Symposium.* San Francisco, CA, Aug. 2011.
http://www.usenix.org/events/sec11/tech/full_papers/Antonakakis.pdf

[Bilge 2011]
Bilge, L.; Kirda, E.; Kruegel, C.; & Balduzzi, M. "EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis." *18th Annual Network & Distributed System Security Symposium Proceedings.* San Diego, CA, Feb. 2011. http://www.internetsociety.org/doc/exposure-finding-malicious-domains-using-passive-dns-analysis-paper

[Felegyhazi 2010]
Felegyhazi, M.; Kreibich, C.; & Paxson, V. "On the Potential of Proactive Domain Blacklisting." *Proceedings of the 3rd USENIX Conference on Large-Scale Exploits and Emergent Threats.* San Jose, CA, Apr. 2010. http://usenix.org/event/leet10/tech/full_papers/Felegyhazi.pdf

[Google 2011]
Google. *Google Safe Browsing FAQ.* Google, 2011.
http://code.google.com/apis/safebrowsing/safebrowsing_faq.html

[Hao 2011]
Hao, S.; Feamster, N.; & Pandrangi, R. "Monitoring the Initial DNS Behavior of Malicious Domains," 269–278. *Proceedings of the Internet Measurement Conference 2011.* Berlin, Germany, Nov. 2011. ACM, 2011. http://conferences.sigcomm.org/imc/2011/docs/p269.pdf

[Hu 2011]
Hu, X.; Knysz, M.; &Shin, K. G. "Measurement and Analysis of Global IP-Usage Patterns of Fast-Flux Botnets," 2633–2641. *Proceedings of IEEE INFOCOM 2011.* Shanghai, China, Apr. 2011. IEEE, 2011.

[Huth 2012]
Huth, C. & Spring, J. "The Impact of Passive DNS Collection on End-user Privacy." *Securing and Trusting Internet Names 2012.* National Physical Laboratory, UK, 2012.
http://conferences.npl.co.uk/satin/papers/satin2012-Spring.pdf

[ICANN 2008]
ICANN. *SSAC Advisory on Fast Flux Hosting and DNS* (SAC-025). Internet Corporation for Assigned Names and Numbers – Security and Stability Advisory Committee, 2008.

[ICANN 2011]
ICANN. Specification 4, Section 2, "Zone File Access." *New gTLD Agreement.* ICANN, 2011. http://www.icann.org/en/topics/new-gtlds/agreement-specs-clean-19sep11-en.pdf

[Kadir 2012]
Kadir, A. F. A.; Othman, R. A. R.; & Aziz, N. A. "Behavioral Analysis and Visualization of Fast-Flux DNS," 250–253. *Proceedings of the 2012 European Intelligence and Security Informatics Conference (EISIC)*. Odense, Denmark, Aug. 2012. IEEE, 2012.

[Konings 2009]
Konings, M. *Final Report of the GNSO Fast Flux Hosting Working Group.* Internet Corporation for Assigned Names and Numbers – Generic Names Supporting Organization, 2009. http://gnso.icann.org/issues/fast-flux-hosting/fast-flux-final-report-06aug09-en.pdf

[Ma 2009a]
Ma, J.; Saul, L. K.; Savage, S.; & Voelker, G. M. "Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs," 1245–1254. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Paris, France, June 2009. ACM, 2009.

[Ma 2009b]
Ma, J.; Saul, L. K.; Savage, S.; & Voelker, G. M. "Identifying Suspicious URLs: An Application of Large-Scale Online Learning," 681–688. *Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal, QC, Canada, June 2009. ACM, 2009.

[Mockapetris 1987]
Mockapetris, P. *RFC 1035 - Domain Names - Implementation and Specification*. The Internet Engineering Task Force, 1987. http://www.ietf.org/rfc/rfc1035.txt

[Moore 2007]
Moore, T. & Clayton, R. "Examining the Impact of Website Take-Down on Phishing," 1–13. *Proceedings of the Anti-Phishing Working Group's 2nd Annual eCrime Researchers Summit*. Pittsburgh, PA, Oct. 2007. ACM, 2007.

[Passerini 2008]
Passerini, E.; Paleari, R.; Martignoni, L.; & Bruschi, D. "Fluxor: Detecting and Monitoring Fast-Flux Service Networks," 186–206. *Proceedings of the Fifth Conference on Detection of Intrusions and Malware & Vulnerability Assessment*. Paris, France, July 2008. Springer Verlag, 2008.

[PhishTank 2011]
PhishTank. *FAQ*. http://www.phishtank.com/faq.php (2011).

[Rasmussen 2010]
Rasmussen, R. & Aaron, G. *Global Phishing Survey: Trends and Domain Name Use in 2H2009.* Anti-Phishing Working Group, 2010.

[Rasmussen 2011]
Rasmussen, R. & Aaron, G. *Global Phishing Survey: Trends and Domain Name Use in 2H2010.* Anti-Phishing Working Group, 2011.

[RIPE 2012]
RIPE Network Coordination Center. *Routing Information Service (RIS).* http://www.ripe.net/data-tools/stats/ris/routing-information-service (2012).

[Route-Views 2005]
Route-Views. *University of Oregon Route Views Project.* http://www.routeviews.org (2005).

[Spamhaus 2011]
Spamhaus. *About Spamhaus.* http://www.spamhaus.org/organization/index.lasso (2011).

[Spring 2010]
Spring, J. M. "Large Scale DNS Traffic Analysis of Malicious Internet Activity with a Focus on Evaluating the Response Time of Blocking Phishing Sites." Master's thesis, University of Pittsburgh, 2010.

[Spring 2011]
Spring J. M.; Metcalf, L. B.; & Stoner, E. "Correlating Domain Registrations and DNS First Activity in General and for Malware." *Proceedings of Securing and Trusting Internet Names (SATIN 2011).* Teddington, U.K., Apr. 2011.

[Stoner 2010]
Stoner, E. "DNS Footprint of Malware." *Proceedings of 2010 OARC Workshop 2.* Denver, CO, Oct. 2010. https://www.dns-oarc.net/files/workshop-201010/OARC-ers-20101012.pdf

[Weimer 2005]
Weimer, F. "Passive DNS Replication." *Proceedings of the 17th Annual FIRST Conference on Computer Security Incident Handling.* Singapore, June 2005.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
(Leave Blank)	October 2013	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Passive Detection of Misbehaving Name Servers	FA8721-05-C-0003

6. AUTHOR(S)

Leigh B. Metcalf and Jonathan M. Spring

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	CMU/SEI-2013-TR-010

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFLCMC/PZE/Hanscom Enterprise Acquisition Division 20 Schilling Circle Building 1305 Hanscom AFB, MA 01731-2116	ESC-TR-2013-010

11. SUPPLEMENTARY NOTES

12A DISTRIBUTION/AVAILABILITY STATEMENT	12B DISTRIBUTION CODE
Unclassified/Unlimited, DTIC, NTIS	

13. ABSTRACT (MAXIMUM 200 WORDS)

In the process of categorizing malicious domains, distinguishing between suspicious and benign name servers can allow the name servers themselves to be acted against. Name servers do not normally change internet protocol (IP) addresses frequently. Domains that do change IP addresses quickly or often are said to exhibit IP flux, which can allow services, such as web pages that deliver malicious content, to circumvent defenders' attempts to block their IP addresses. IP flux in a name server's domain may be a sign that the name server is suspicious. This report demonstrates that name-server flux exists and is ongoing. Furthermore, there are two types of data that can reveal IP flux in domain name system (DNS) servers: passively collected DNS messages and the contents of several large, top-level domains' official zone files.

14. SUBJECT TERMS	15. NUMBER OF PAGES
passive DNS, network situational awareness, IP flux, fast flux, DNS analysis, zone file, internet trends	29

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL