DNS-OARC

Domain Name System Operations Analysis and Research Center

Deployment of DNSSEC in the Root Zone: Impact Analysis

Geoffrey Sisson, *DNS-OARC*, <geoff@dns-oarc.net>

OARC Document 2010-001 December 15, 2010

Abstract

This report examines the impact of the deployment of DNSSEC in the root zone during the phased rollout that started on January 27, 2010, and ended on July 15, 2010. We look at three key areas: changes in reply sizes from the root servers, potential signs of path MTU issues for DNS clients, and changes in TCP query rates to the root servers.

1 Introduction

The Domain Name System (DNS) is the principal naming system for identifying hosts, services, and other resources on the Internet.[1][2] The DNS Security Extensions (DNSSEC) add data origin authentication and data integrity to the DNS.[3][4][5]

In October 2009, ICANN and VeriSign announced plans for an incremental rollout of DNSSEC in the root zone.[6] The rollout entailed the use of a Deliberately Unvalidatable Root Zone (DURZ) prior to the introduction of a fully operational DNSSEC signed zone. The rollouts subsequently followed the schedule in Table 1.[7]

This report has been written with readers of varying levels of familiarity the DNS in mind. A glossary has been provided at the end of this report for terms that may be unfamiliar.

2 Data Collection

Data was collected from the participating root name servers starting the day prior to a DURZ rollout maintenance window to the day after. The DURZ events were all scheduled for Wednesdays at – or in the proximity of – 18:00 UTC, so the data collection windows were centered around this time. Some collection periods ran longer than 48 hours to accommodate multiple sequential DURZ events.

Data was collected for two additional two-day periods: one a week prior to the first DURZ rollout, and the other three weeks after the final DURZ rollout. These additional collections were intended to provide baseline data samples that would be unaffected by proximity to a DURZ rollout. Data was also collected for a final five-day period corresponding with the introduction of the production signed root zone on July 15.

Date	Event
Jan 27, 2010	L starts to serve the DURZ
Feb 10, 2010	A starts to serve the DURZ
Mar 3, 2010	I and M start to serve the DURZ
Mar 24, 2010	D, E, and K start to serve the DURZ
Apr 14, 2010	B, C, F, G, and H start to serve the DURZ
May 5, 2010	J starts to serve the DURZ
Jul 15, 2010	Distribution of validatable, production, signed root zone and publication of root zone trust anchor

Table 1: DNSSEC deployment schedule.

The data was collected as pcap (packet capture) files. Some Root Server Operators (RSOs) sent pcap files in fixed duration increments, typically five or ten minutes. Others sent pcap files containing a fixed number of packets, starting a new file each time a predetermined number was reached. With the exception of J-Root, only query data was collected. Most RSOs provided separate pcap files for each node in a name server cluster.

Table 2 summarizes the data collection windows.

Event	Collection Start	Collection End	Collection Size *(gzipped pcaps)*
Initial baseline	2010-01-19 18:00Z	2010-01-21 18:00Z	0.91 TB
L	2010-01-26 18:00Z	2010-01-28 18:00Z	1.06 TB
A	2010-02-09 18:00Z	2010-02-11 18:00Z	1.39 TB
I, M	2010-03-02 06:00Z	2010-03-04 18:00Z	1.44 TB
D, E, K	2010-03-23 14:00Z	2010-03-25 20:00Z	1.64 TB
B, C, F, G, H	2010-04-13 13:00Z	2010-04-16 00:00Z	2.95 TB
J	2010-05-03 16:00Z	2010-05-05 20:00Z	1.93 TB
Final baseline	2010-05-25 17:00Z	2010-05-27 19:00Z	1.75 TB
Production signed zone	2010-07-14 13:00Z	2010-07-19 13:00Z	4.3 TB

Table 2: Data collections.

3 Data Quality

The data was provided by the RSOs on a best-effort basis. The collected data diverges from a precise representation of all root server query traffic for a variety of reasons:

1. Not all RSOs participated in all data collections. The only data collection event for which data was received from all root servers was the B, C, F, G and H DURZ event, which coincided with the 2010 DITL collection.[8]

2. Some RSOs did not provide data for all of the nodes in their name server clusters.

3. There were occasional gaps in the data. In a few cases only a small amount of data was received for a node. One RSO had neither enough bandwidth to upload pcap files nor storage to spool them locally for later transmission, so they managed to report only a small amount of data.[1]

4. Not all root servers with IPv6 transit provided IPv6 data.

5. On several occasions an RSO provided pcap files that were truncated to 96 octets, the default for older versions of `tcpdump`.[9]

6. Few RSOs used network taps and data acquisition cards, and thus will have missed a small portion of traffic (typically $< 1\%$).

7. Some pcap files had varying degrees of corruption.

Netnod, the operator for I-Root, could supply data with obfuscated client IP address only. They had received advice that EU data protection laws forbid revealing the original IP addresses. Instead they hashed each client address to a unique Net 10/8 address. This made it difficult to isolate some sources of anomalous traffic, though Netnod were very generous with their assistance in this matter. Another drawback of this scheme was that information about which addresses were in the same net block was lost.

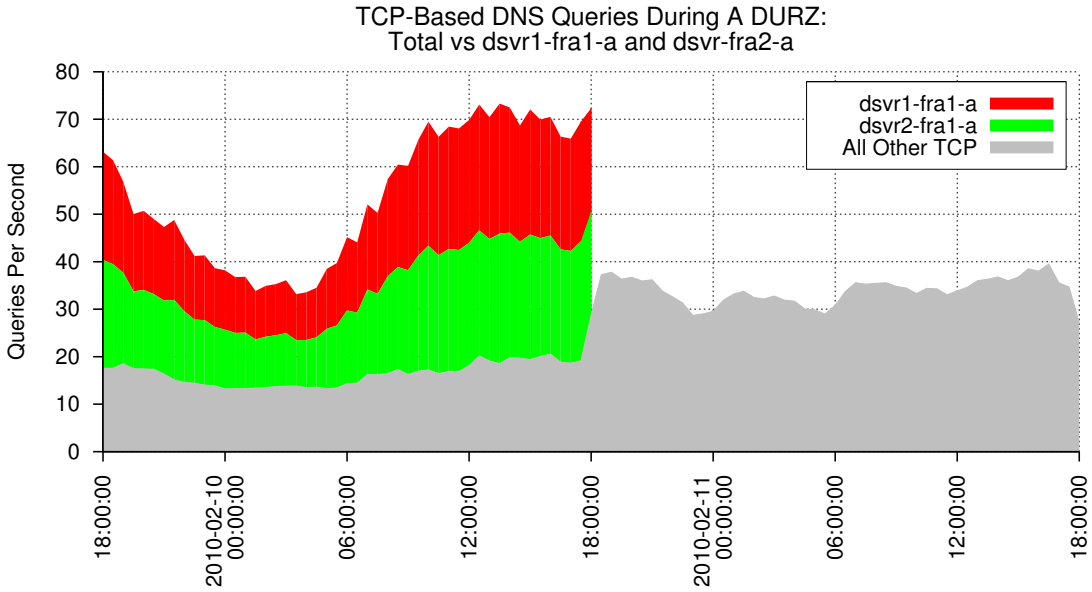

Figure 1: TCP-based DNS queries during A DURZ.

One particular case serves to illustrate the incompleteness of the collected data. During the A DURZ collection event, we received data from two Frankfurt-based nodes of A-Root for the first half of the collection period (Figure 1). We received no data from these nodes during the other collection events. However, for the period of time they were reporting, these nodes recorded more

[1]The coverage maps at `http://ditl.dns-oarc.net/` (e.g., `http://ditl.dns-oarc.net/ditl_20100525_raw/coverage.png`) contain detailed information about participating nodes and coverage gaps.

than double the number of TCP-based DNS queries to the root servers than all other nodes of all root servers combined.[2] It should be noted that 50% of these TCP queries were attributable to just four clients, but the rest were from 8,760 distinct addresses – nearly half the number of all clients sending queries via TCP to all root servers during this period.

These two nodes stopped reporting just as the DURZ was deployed on A-Root. As a consequence, the data initially seemed to indicate that the rate of TCP-based DNS queries to A-Root *fell* when A-Root began to serve the signed zone. Because this fragment of data creates such a distortion, it was excluded from subsequent analysis. However, it demonstrates that the collected data should not be regarded as complete.

4 Data Statistics

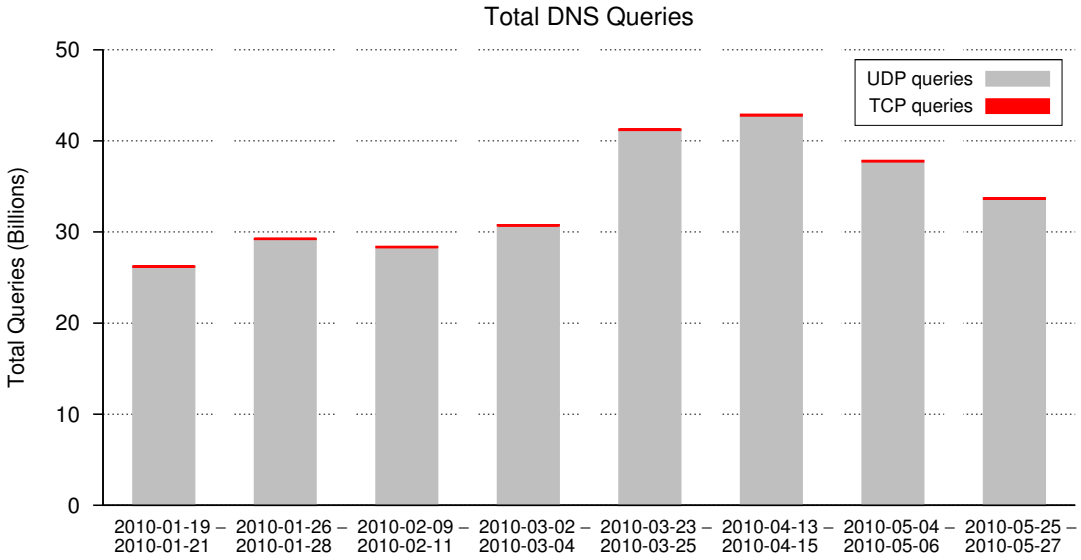

Figure 2: Total queries.

Figure 2 shows the total number of queries reported by all root servers for the central 48 hours of each collection period, i.e., 18:00 UTC Tuesday to 18:00 UTC Thursday. UDP-based DNS queries are shown in gray while TCP-based DNS queries are shown in red. This plot illustrates that the proportion of TCP-based queries is very small, even at the end of the DURZ rollout.

[2]Note that this excludes the anomalous traffic received from a large US cable Internet provider discussed in Section 5.3.

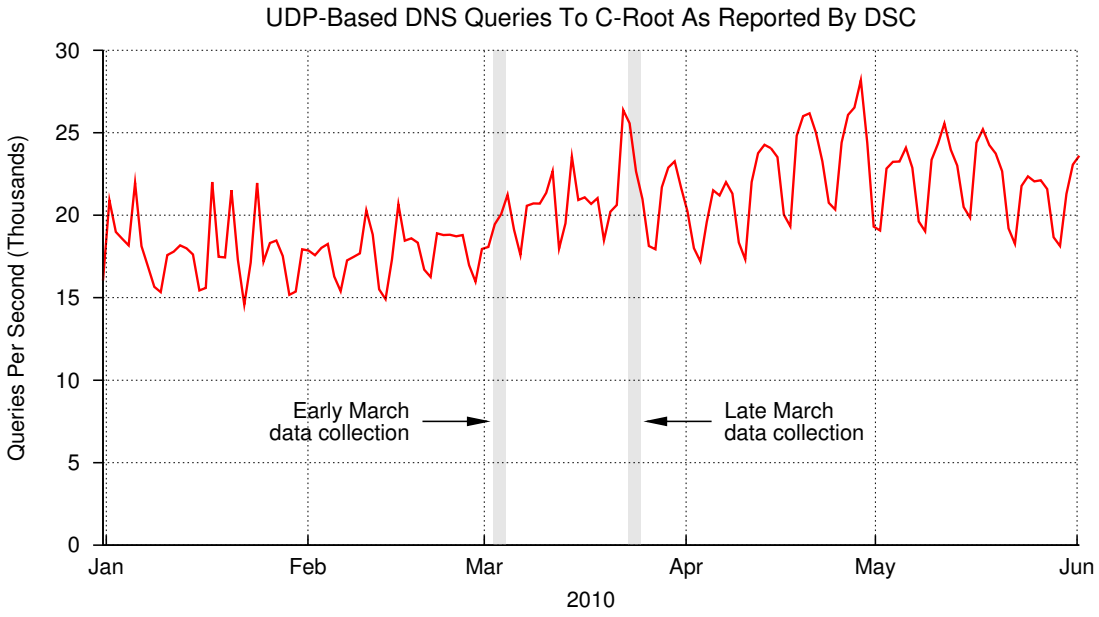

Figure 3: UDP-based DNS queries to C-Root (via DSC), January – June 2010.

One surprising characteristic of the data in Figure 2 is the greater than 33% increase in total queries between the early March and late March data collections. As a sanity check, we compared these traffic levels with the DSC (DNS Statistics Collector) data for C-Root (Figure 3).[10] Unlike DITL, the DSC data was reported in a continuous and consistent way throughout the entire DURZ rollout. Reassuringly, the DSC data (Figure 3) shows a traffic spike coinciding with the late March data collection consistent with the increase in Figure 2.

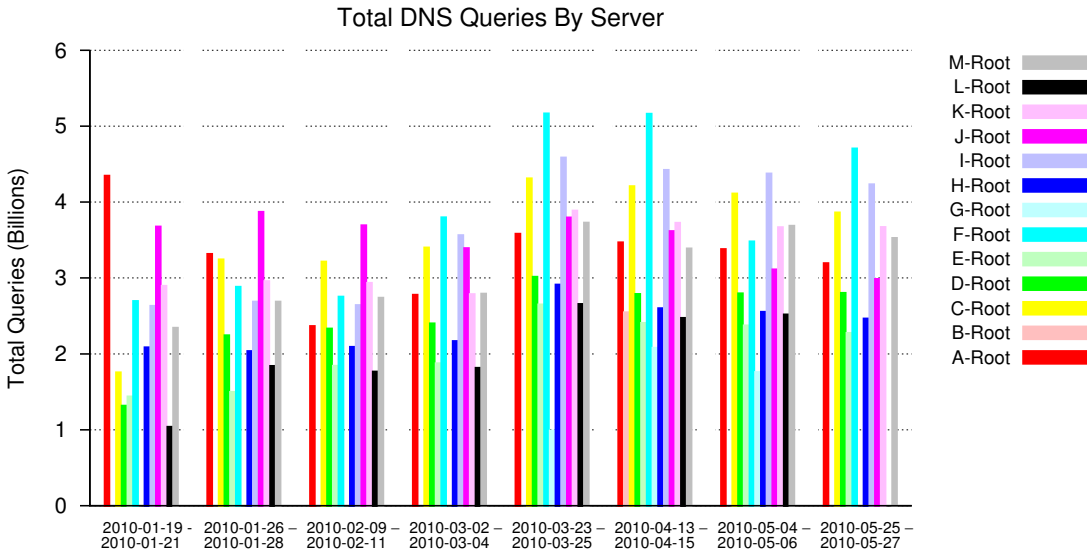

Figure 4: Total queries (by server).

Figure 4 shows the total number of queries per root server. Note that this should not be viewed as an accurate depiction of the distribution of queries among the root servers because – as previously noted – the data from some root servers is incomplete.

5 Data Analysis

We searched for evidence that the DURZ rollout had negatively affected DNS clients. Of primary concern was that the increased reply sizes may have effectively rendered root servers unreachable to a portion of clients, either due to path MTU problems, or improperly designed or configured firewalls that would not transmit the larger DNSSEC packets.

There was evidence that the increased reply sizes had *some* effect. Firstly, there were sharp increases in the number of TCP-based DNS queries to the root servers as well as the number of clients using TCP. Secondly, there were small increases in the number of repeated queries for some types of queries. Ultimately, however, there was no clear evidence of disruption to any clients.

In this section we examine data indicating changes in client behavior. For brevity, we do not include data that showed no effect. For example, while priming queries to the root servers were monitored throughout the DURZ rollout, there were no discernible changes in priming query patterns.

5.1 DNS Reply Sizes

The introduction of DNSSEC in the root zone resulted in larger DNS reply sizes. Here we examine the nature of these changes.

During the data collections, DNS reply data was collected for J-Root only, so this data is based on observations at J-Root alone. J-Root is unique in that it is the only root server that is not authoritative for the `.arpa` top-level domain. This affects the distribution of reply sizes slightly as some of the larger replies are referrals for `.arpa`.

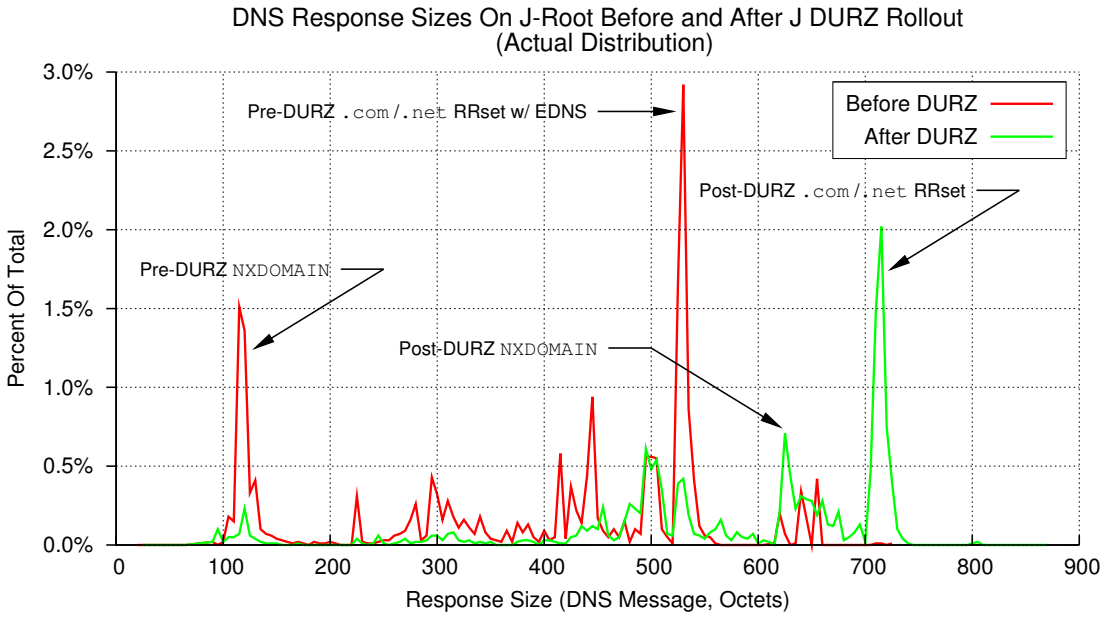

Figure 5: UDP-based DNS reply sizes from J-Root (actual distribution).

Figure 5 shows the distribution of reply sizes from J-Root before (in red) and after (in green) the J DURZ rollout on May 5. It excludes a two-hour period around the actual DURZ rollout window, as during this period some J-Root nodes may have been serving the DURZ at the same time others were serving the unsigned root zone.

Two significant clusters are immediately apparent in the pre-DURZ distribution: one centered at 118 octets and another centered at 533 octets. The peak at 118 corresponds to the size of NXDOMAIN replies; the peak at 533 corresponds to the size of referrals for the `.com` and `.net` zones with EDNS, both of which share the same Resource Record set (RRset). The replies are distributed around these peaks as a function of the length of the original QNAME, which is of course included in the reply.

A smaller cluster is apparent at 500 octets. This corresponds to the size of referrals for `.com`, `.net` *without* EDNS, as well as referrals for the `.arpa` zone. The small peaks at 643 and 657 octets correspond to replies to queries for `.arpa` and priming-queries that specify a sufficiently large EDNS0 buffer. The very largest replies are either in response to queries for PTR Resource Record (RR)s in `ip6.arpa` domain or referrals for malformed `.arpa` queries.

Finally, the smaller peaks between 200 and 500 octets generally correspond to referrals for other

TLDs with smaller NS RRsets.

The post-DURZ distribution (in green) is significantly different. There are still peaks at the previous locations, but they are greatly diminished. The peak at 500 octets has not changed; this is because the proportion of queries for `.com` and `.net` without EDNS is unchanged. On average, replies have increased in size by roughly 180 octets. The peak for `.com` and `.net` referrals that used to be at 533 is now centered at 715. Other peaks have also shifted by a comparable amount. The biggest increase is for post-DURZ NXDOMAIN replies, which now average 625 octets

Note that the peaks to the right of the `.com/.net` referral peak do not appear to have shifted. In actuality they have, but they are now aligned with the `.com/.net` peak. The very largest reply seen after the DURZ rollout was 894 octets, a referral for a PTR RR query for a name in `ip6.arpa`.

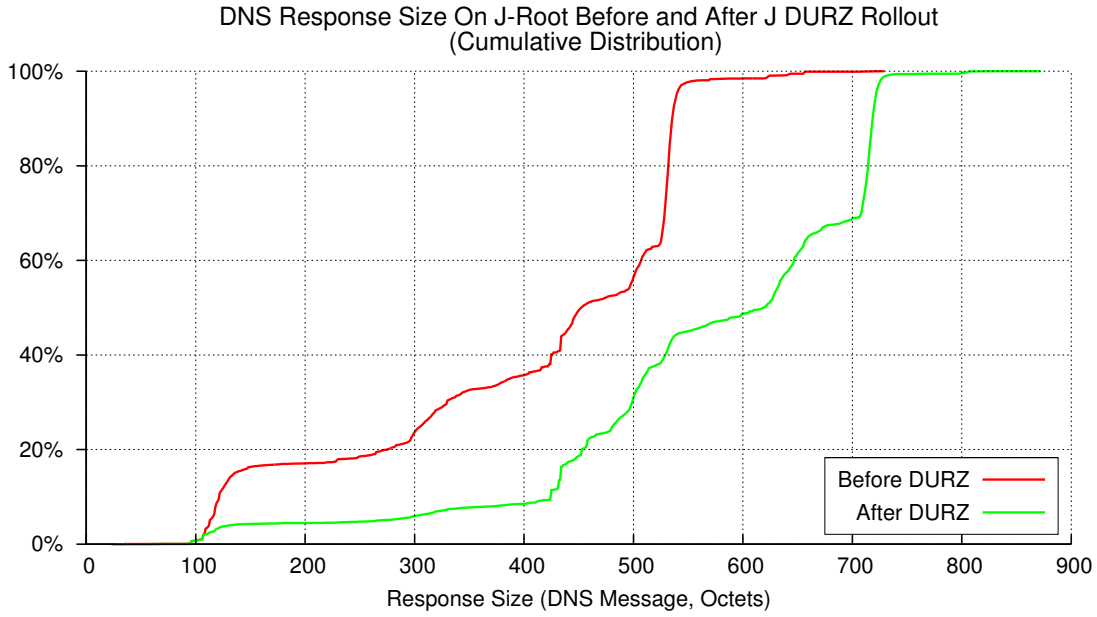

Figure 6: UDP-based DNS reply sizes from J-Root (cumulative distribution).

Figure 6 shows the information in Figure 5 as a cumulative distribution.

In total, the average response size grew from 405 to 569 octets.

5.2 DNS Retries

It was anticipated that some resolvers might experience path MTU problems when attempting to receive the larger responses associated with DNSSEC. This happens when both the resolver and the server are capable of handling larger responses via EDNS0, but an intermediate link or device is not, and, crucially, the resolver and server do not detect this via path MTU discovery.[11]

A path MTU problem typically manifests itself at the server as repeated queries from an affected resolver without corresponding ICMP "fragmentation needed" responses. The collected data does not contain ICMP data, so we have to rely on query data alone.

We analyzed DNS retries during the first and last DURZ rollout using the following methodology.

A query was considered to be *retried* if the following conditions were met:

- A query was observed at a given root server.
- A query with the same source address, QNAME, QTYPE, and QCLASS was subsequently observed at another root server within the next 60 seconds.
- The total number of queries with that source IP address, QNAME, QTYPE, and QCLASS did not exceed five within a two-minute period. This was to minimize the impact of malfunctioning resolvers and other clients that send large numbers of repeated queries.

Note that this is a narrower definition of a retry than customary in that we consider only queries that were sent to more than one root server. This did *not* discount queries that may have been sent several times to one root server before being sent to another.

These criteria have some limitations:

- The initial query may itself be a retry, i.e., there may have been an initial query that was not detected. Given the volume of data involved, it was impractical to identify these cases.
- Similar queries may be due to multiple clients behind a NAT gateway. For example, two resolvers in the same masqueraded network may send queries for the same popular hostname within 60 seconds of each other. We suspect these cases account for only a small fraction of the observed retries.
- Some resolvers are more aggressive and send queries to multiple root servers, either preemptively or without waiting for a timeout.
- Retries may occur during path MTU discovery, so may not be evidence of path MTU failure.

As previously noted, the transactions most likely to be affected by the introduction of DNSSEC were ones involving large response sizes. Accordingly, we further narrowed the focus to queries that could result in a response size greater than 512 octets. We restricted our search to queries with the following properties:

- EDNS0 present
- DO bit set
- EDNS0 payload size larger than 512 octets.

Because of the computationally intensive nature of this task, we focused on two specific cases: queries arriving at L-Root server during the L-Root DURZ rollout, and queries arriving at J-Root server during the J-Root DURZ rollout. These were the most interesting cases. The L-Root DURZ rollout was the first, and represented the first "ripple in the pond"; the J-Root DURZ rollout was the last, and potentially deprived resolvers with insufficient path MTUs of their sole remaining reachable root server.

As a control, we also looked at queries arriving at K-Root during the same periods.

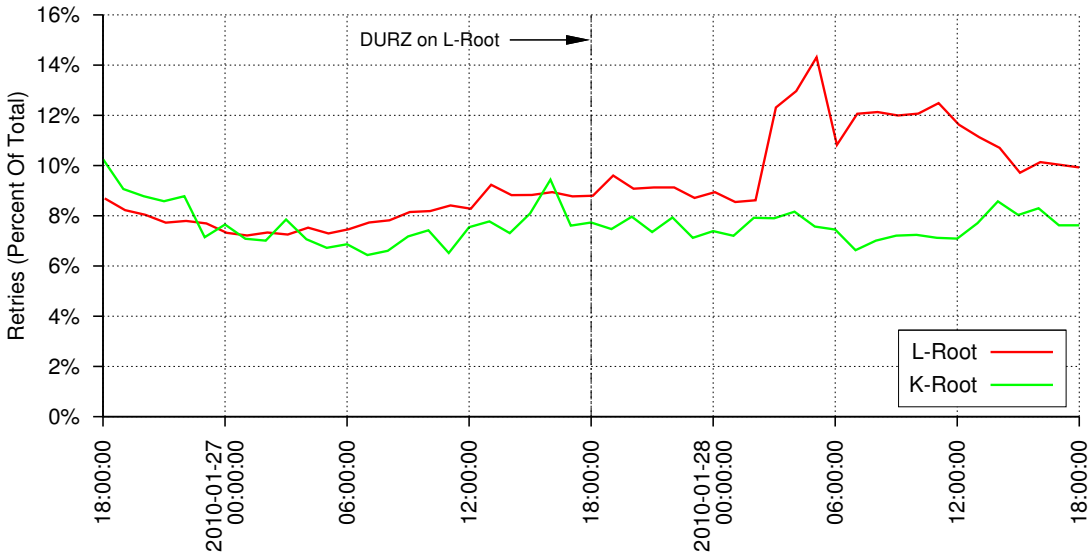

Figure 7: L-Root: retries as a percentage of all queries with DO bit set and EDNS0 payload greater than 512 octets.

Figure 7 shows the rate of retries for queries sent to L-Root (in red) as a percentage of the total number of queries with the DO bit set and an EDNS0 payload greater than 512 octets. A rate of 8% may seem surprisingly high. However, as mentioned previously, many of these queries are either sent preemptively or are evidence of path MTU discovery at work. Note that we are not trying to determine the absolute rate of retries, but to detect changes to the rate in response to the DURZ. Consequently, over-counting retries should not be a problem provided it is done consistently before and after the DURZ event.

In Figure 7 there is no pronounced change in retries at the time of the introduction of the DURZ. However, eight hours later there is a noticeable increase in retries. This increase is attributable to approximately 50 IP addresses in network blocks registered to a large cloud computing services provider. At 02:42 the root servers began to receive large volumes of DNS queries from these addresses. Initially, queries from these addresses were arriving at L-Root at a rate of around 1000 qps. Later they settled to around 500 qps, and then began to tail off after around 13:00.

All of these queries had a source port of 10053 or 20053, and were exclusively for MX or A RRs (except for the occasional query for the root NS RRset. Queries with the same source address and port number were a mix of EDNS0 and non-EDNS0, suggesting that these IP addresses were associated with some sort of proxy or proxies. Nearly all of the retries came from just five of the

50 IP addresses: 174.129.60.85, 174.129.61.47, 174.129.93.88, 174.129.158.118, and 204.236.192.19. These were also the IP addresses associated with the most queries of the 50. Given the sudden onset and relatively rapid cessation of this traffic as well as the composition of MX and A QTYPEs, it is plausible that this traffic was associated with some sort of spam and/or malware activity. It does not appear to be an effect of the DURZ.

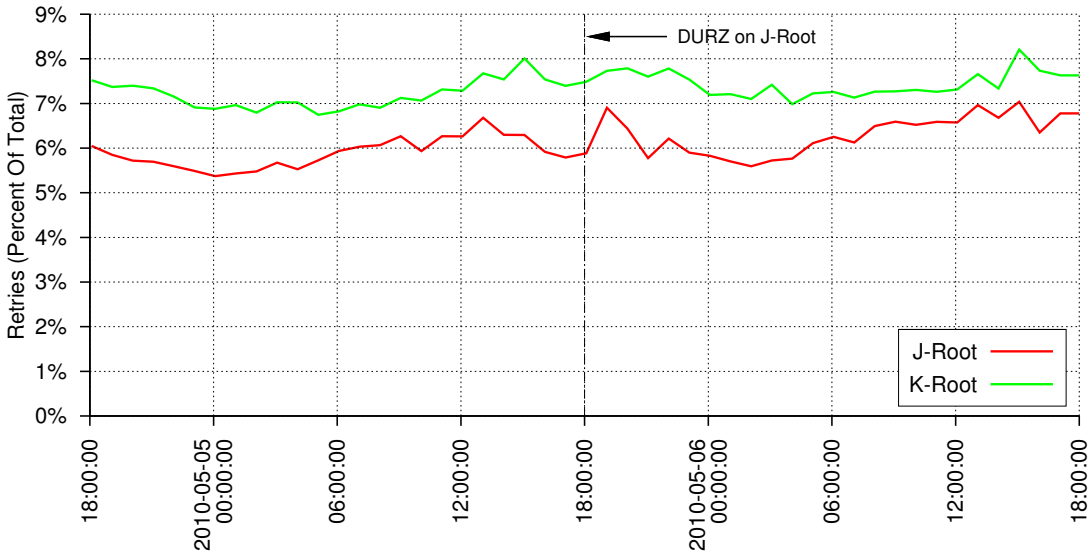

Figure 8: J-Root: retries as a percentage of all queries with DO bit set and EDNS0 payload greater than 512 octets.

Figure 8 shows the rate of retries for similar queries arriving at J-Root during the J DURZ rollout. Here there is no pronounced effect, though there is a small spike starting at 18:00 on May 5, the time of the DURZ rollout.

Note that for the J DURZ rollout, the limit of five similar queries was relaxed, as J-Root was the final server to get the DURZ, and the expected pattern from resolvers with path MTU issues would be to persistently send queries in a vain attempt to get a reply.[3]

An interesting pattern occurs if the queries under consideration are further restricted to only ones that would result in NXDOMAIN responses.

Figure 9 shows the rate of retries for queries that, in addition to having the DO bit set and an EDNS0 payload size greater than 512 octets, would result in an NXDOMAIN response. Here there is a sudden and sustained increase in retries at exactly the time the DURZ is published on L-Root.

Similarly, when the DURZ is published on J-Root, there is a sharp increase in this type of query. To our puzzlement, when we looked at queries that would result in `.com/.net` referrals only, there was no similar pattern. This is counterintuitive as `.com/.net` referrals are larger than NXDOMAIN responses.

[3]Path MTU problems are believed to be more common close to network endpoints, so that if a resolver is experiencing path MTU failures with one root server, it is likely having path MTU failures with all root servers.

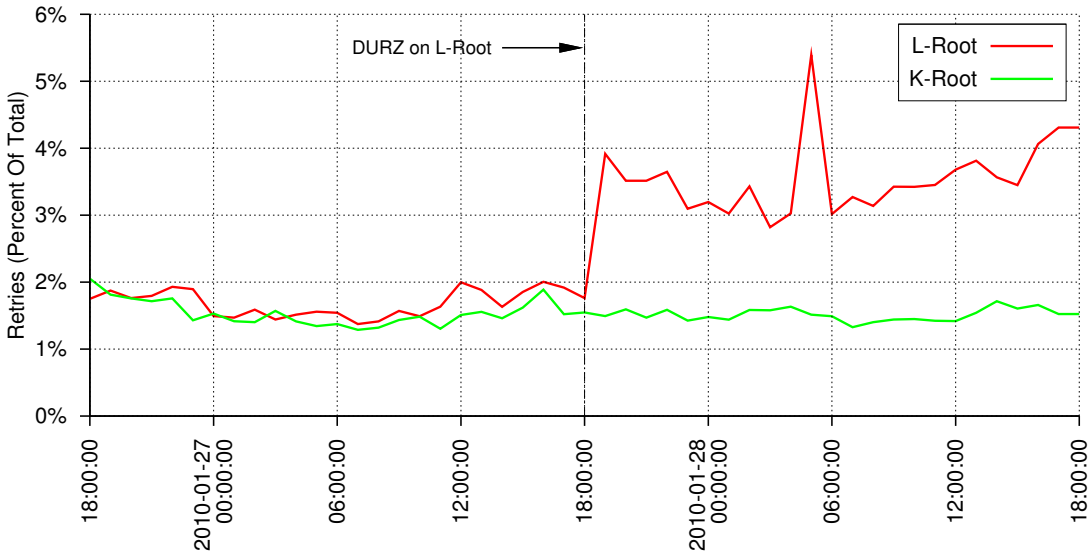

Figure 9: L-Root: retries as a percentage of all queries with DO bit set, EDNS0 payload greater than 512 octets, and resulting in NXDOMAIN.

What does this tell us about the prevalence of path MTU issues that may have been triggered by the DURZ? Unfortunately not much. While there is definitely an increase in retries for some types of queries, this may simply be concomitant with increased levels of path MTU discovery.

One surprise of the DURZ deployment was that no reports of disruption surfaced – at least none reached the DNS community. In the months prior to the rollout, researchers involved with the SecSpider DNSSEC Monitoring Project had reported problems receiving DNSKEY replies for some signed zones at some monitoring locations.[11][12] This led to concerns that some users would find themselves effectively disconnected by the DURZ rollout.[13]

One explanation for the apparent absence of disruption is that normally only validating resolvers send queries for DNSKEY RRs. The sizes of other replies from the signed root zone are generally smaller than DNSKEY replies and are less likely to exceed path MTU limits. During the DURZ rollout there would have been no validating resolvers with a trust anchor for the root zone.

Even when validating resolvers are more widespread, the DNSKEY RRset for the root zone is smaller than that used by many signed zones. A frequent practice is to have two Key Signing Keys (KSKs) and two Zone Signing Key (ZSKs) in the zone. KSKs are typically 2048 bits and ZSKs are typically 1024 bits, so DNSKEY replies can reach 1400 octets or more. The DNSKEY reply for the DURZ and the current production root zone has only one 2048-bit KSK and one 1024-bit ZSK, resulting in a reply of only 883 octets. DNSKEY replies will be larger during KSK and ZSK rollovers, so it remains to be seen whether these events will trigger problems for validating resolvers.

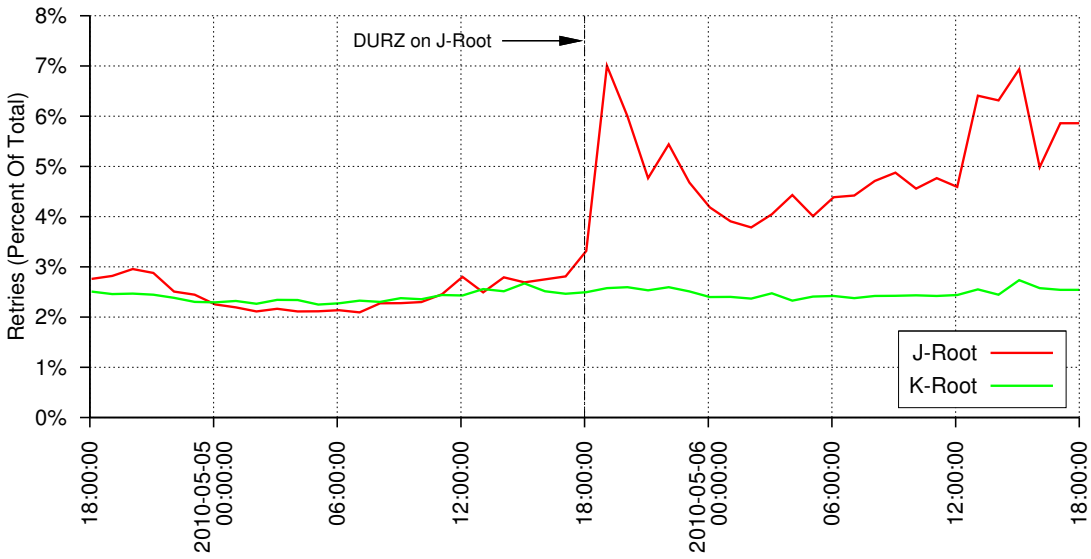

Figure 10: J-Root: retries as a percentage of all queries with DO bit set, EDNS0 payload greater than 512 octets and resulting in NXDOMAIN.

5.3 Changes in TCP-Based DNS Query Patterns

It was anticipated that a signed root zone would result in an increase in TCP-based DNS queries, as responses to queries with the DO bit set are significantly larger, and many clients cannot or will not accept UDP packets with DNS payloads larger than 512 octets. This was most dramatically observed when the `.org` zone was signed in June 2009.[14][15]

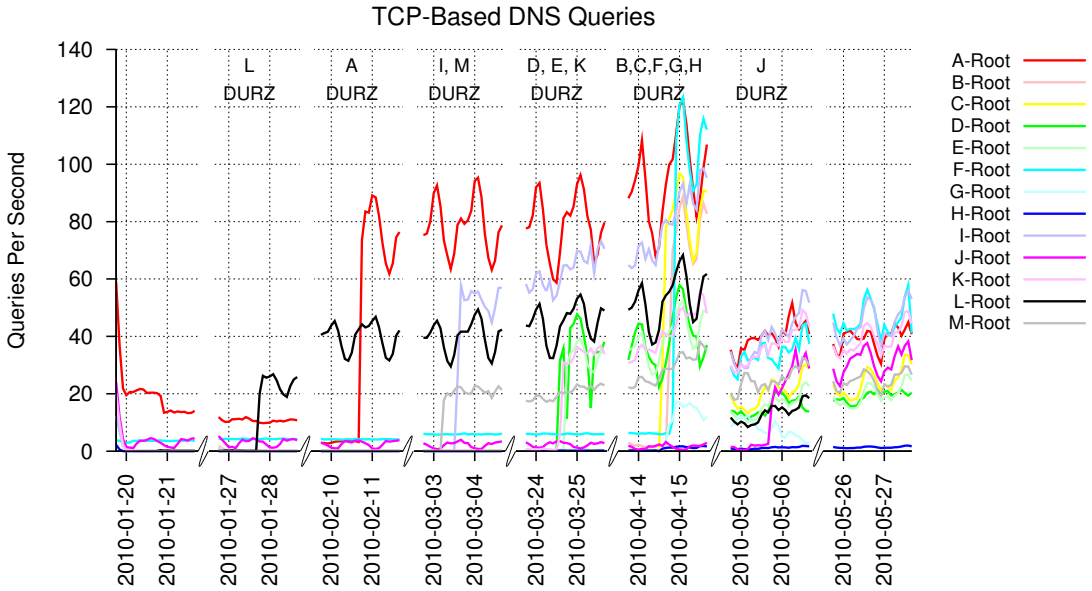

Figure 11: TCP-based DNS queries (including queries from large US cable Internet provider).

Figure 11 has an overview of the changes in TCP-based DNS query rates for each root server during the DURZ rollout. The plot shows TCP query rates increasing as the DURZ rollout progresses. However the most striking aspect of this plot is the dramatic *drop* in TCP query rates between the mid-April and early May data collection periods.

Further investigation revealed that most of the increase in TCP-based DNS queries came from resolvers operated by a single large US cable Internet provider. The queries came from the following IP address blocks:

- 10 IP addresses in 68.1.208.0/28
- 40 IP addresses in 68.105.28.0/24
- 37 IP addresses in 68.105.29.0/24

The TCP-based queries were triggered by truncated responses to UDP queries with the EDNS payload size set to 512 octets. These increased in volume as successive root servers began to serve the DURZ. Then, sometime between the mid-April and the early May data collections, the provider did something to change the payload size of their queries to 4096 octets, effectively fixing the problem. The provider's resolvers are open, and fpdns reports that they are currently running "ISC BIND 9.2.3rc1 – 9.6.1-P1".[16] As this is an older version of BIND, it is possible that the provider made corrections to their configuration as opposed to changing or upgrading their name server software.

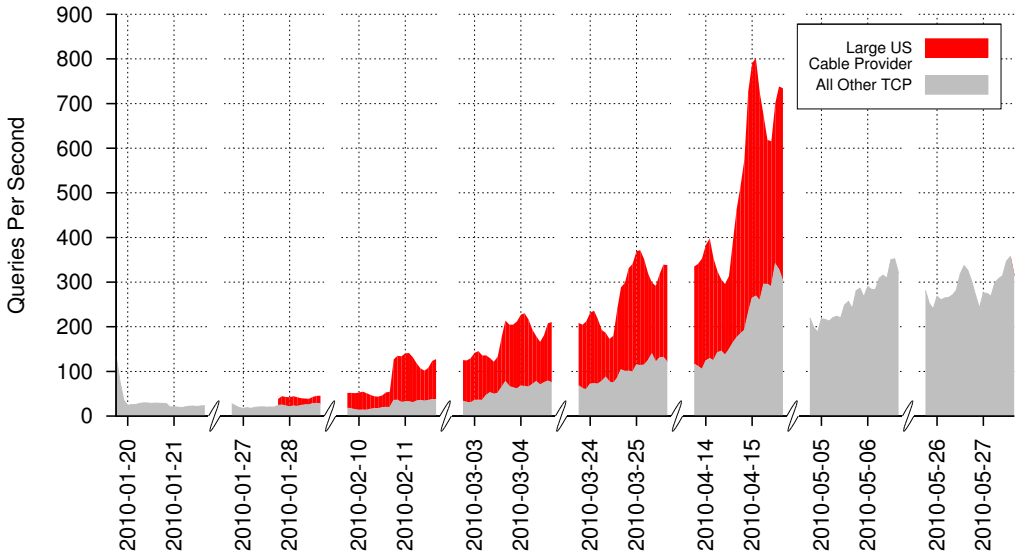

Figure 12: TCP-based DNS queries: Large US cable provider vs. all other sources.

Figure 12 shows the contribution this anomaly made to overall TCP query rates. During the mid-April data collection, the provider's resolvers sustained rates of over 500 TCP-based DNS queries per second to the root servers – more than double the rate of TCP-based queries from all other sources.

This traffic from the provider's resolvers is evidently the result of the DURZ deployment. As such, it cannot be ignored altogether. However, because this traffic distorts the overall picture and overshadows other effects, we elected to exclude it from subsequent analysis.

Figure 13 shows the data in Figure 11 without the anomalous traffic.

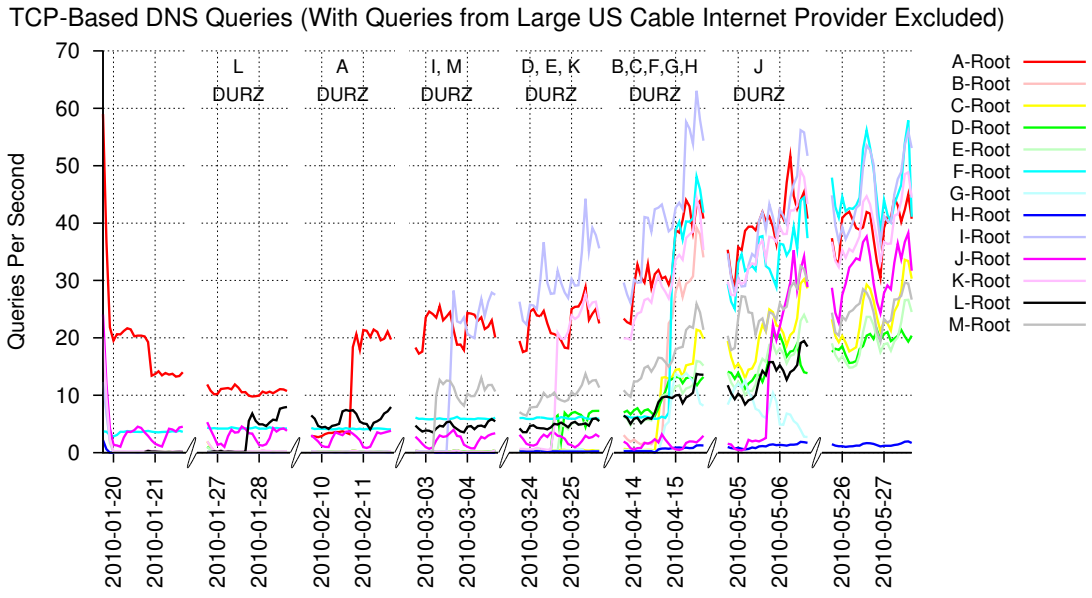

Figure 13: TCP-based DNS queries (excluding queries from large US cable Internet provider).

The plots that follow show the TCP-based DNS query patterns around each DURZ rollout in greater detail. Note that the rates refer to the number of DNS messages received via TCP, not the number of TCP packets received. Also, note that the rates exclude TCP messages that are not standard Queries, i.e., where OPCODE≠0. This has the effect of ignoring a significant number of Update messages that are sent to A-Root via TCP. These messages are typically from misconfigured Microsoft Windows hosts, and it is probably safe to assume their prevalence would be unaffected by the DURZ rollout.

Figure 14 is an overview of the changes in the number of sources of TCP-based DNS queries reported by all root servers. These are examined in more detail in the following plots.

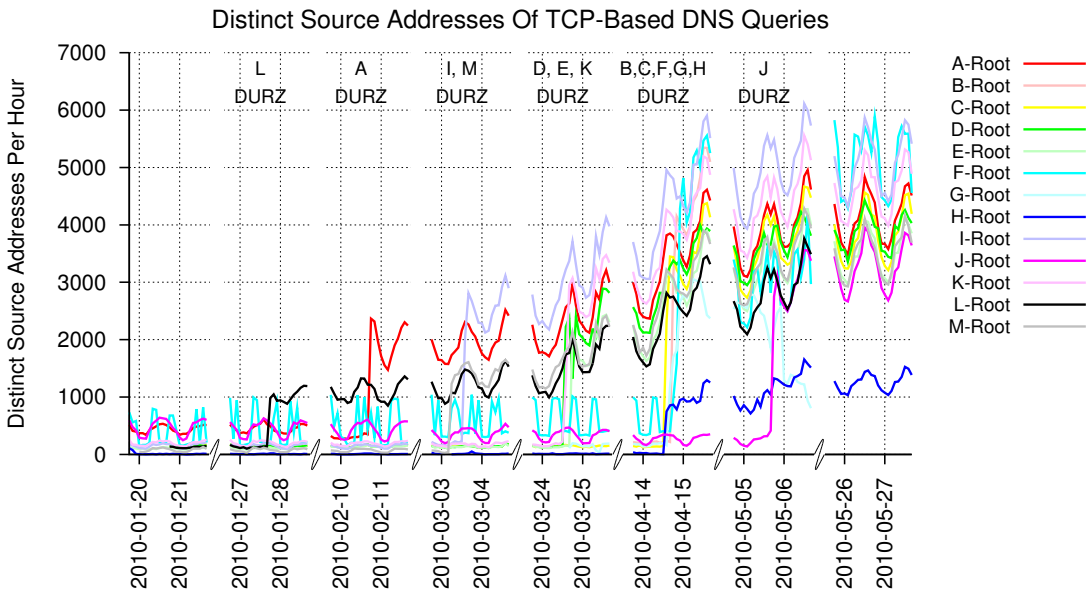

Figure 14: Distinct source addresses of TCP-based DNS queries.

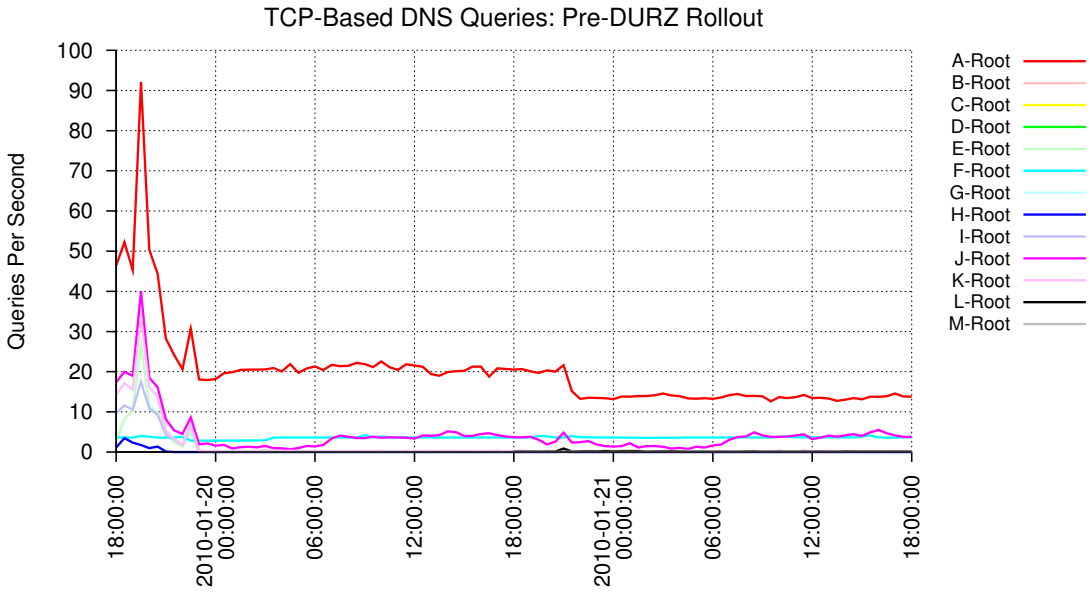

Figure 15: TCP-based DNS queries: pre-DURZ.

5.3.1 Pre DURZ

Figure 15 shows the TCP query rate prior to the first DURZ rollout. The large spike at 20:00 on January 19 is due to unusual traffic originating from 622 addresses in a /21 network of a large US-based search engine. The traffic appears to consist exclusively of queries for A RRs for names prepended with a nonce.[4] In some cases the strings are valid hostnames, but in most cases they appear to be text from searches. This traffic persists until around 22:00 on January 20. It was not seen again during the remaining data collections.

Note that A-Root receives a disproportionate number of TCP-based queries relative to the other root servers. This has been observed in previous studies, and is primarily due to a large number of TCP-based TKEY key exchange attempts.[17] A-Root is specified as the MNAME in the root zone's SOA RR, making it a target for misconfigured hosts.

Also note that data for C-, D- and L-Root was not reported until halfway through the collection period, at 18:00 on January 20.

[4]The query format appears to be an implementation of `http://tools.ietf.org/html/draft-barwood-dnsext-fr-resolver-mitigations`.

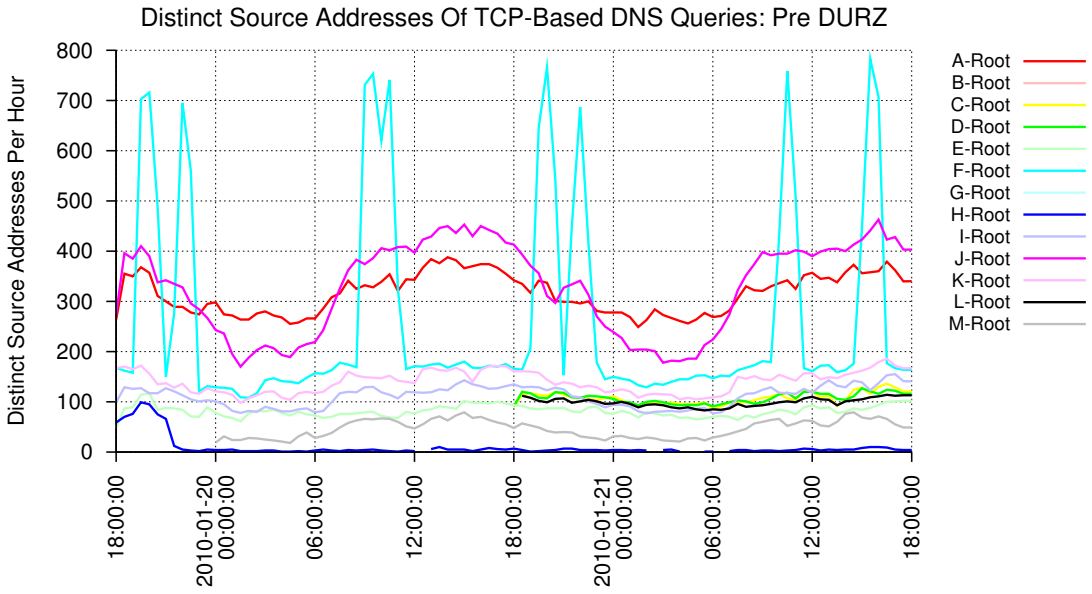

Figure 16: Distinct TCP query sources: pre-DURZ.

Figure 16 shows the number of distinct TCP query sources observed prior to the first DURZ rollout. The periodic spikes in the data for F-Root are due to occasional queries originating from about 600 sources from within some ten /24 net blocks belonging to a large Russian search engine. Each source appears to cycle through three queries every few hours: `IN SOA .`, `IN SOA arpa.`, and `IN SOA in-addr.arpa.` At 20:00 on January 19 the sources of the anomalous traffic from the large US search engine are apparent in the A-, H-, and J-Root data, though it less pronounced than in Figure 15.

Note again that data for C-, D- and L-Root was not reported until halfway through the collection period, at 18:00 on January 20.

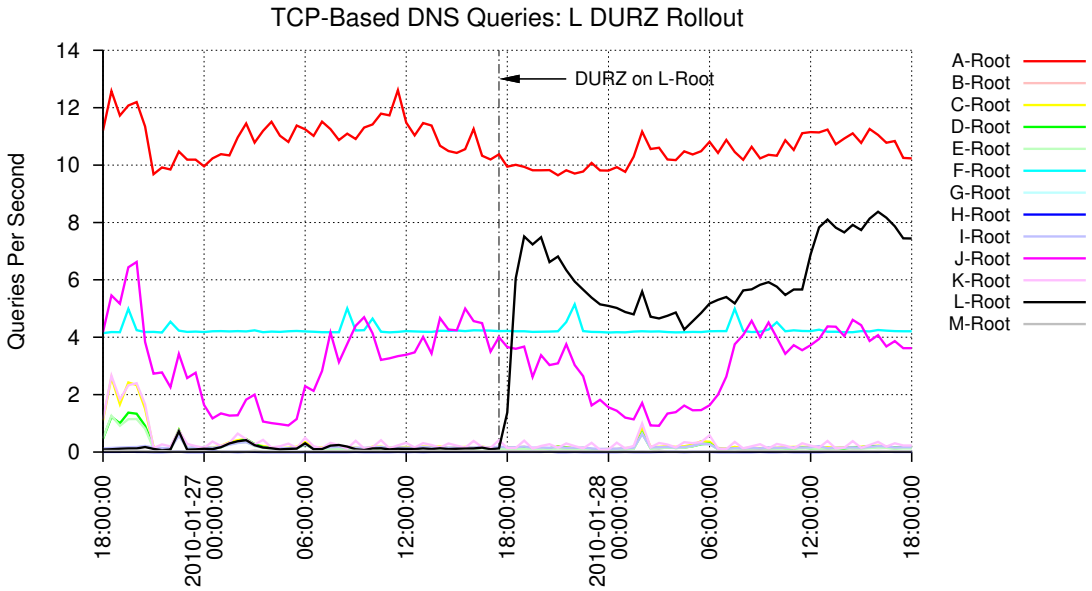

Figure 17: TCP-based DNS queries during the L DURZ rollout.

5.3.2 L DURZ Rollout

Figure 17 shows the TCP query rates during the first DURZ rollout, when the DURZ was introduced on L-Root. A sharp increase in queries to L-Root can be seen at 18:00 UTC on January 27. The resulting query rates for L-Root still remain quite low, barely exceeding eight queries per second.

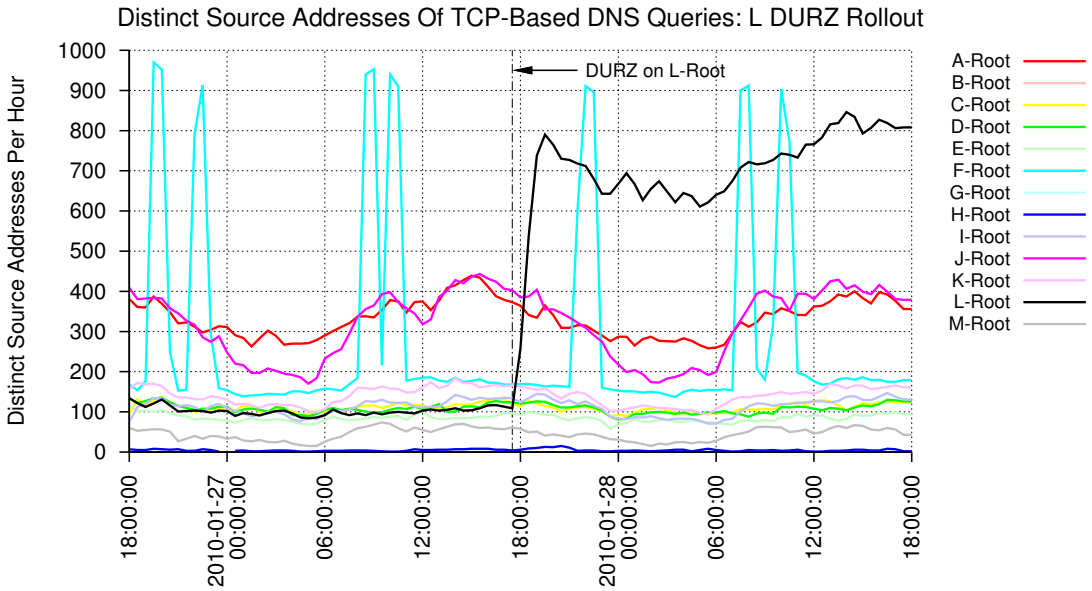

Figure 18: Distinct TCP query sources during the L DURZ rollout.

Figure 18 shows the number of distinct TCP query sources observed during the first DURZ rollout. The number of sources sending TCP-based DNS queries to L-Root increases sharply at 18:00. The anomalous spikes in the F-Root data from the large Russian search engine are still apparent.

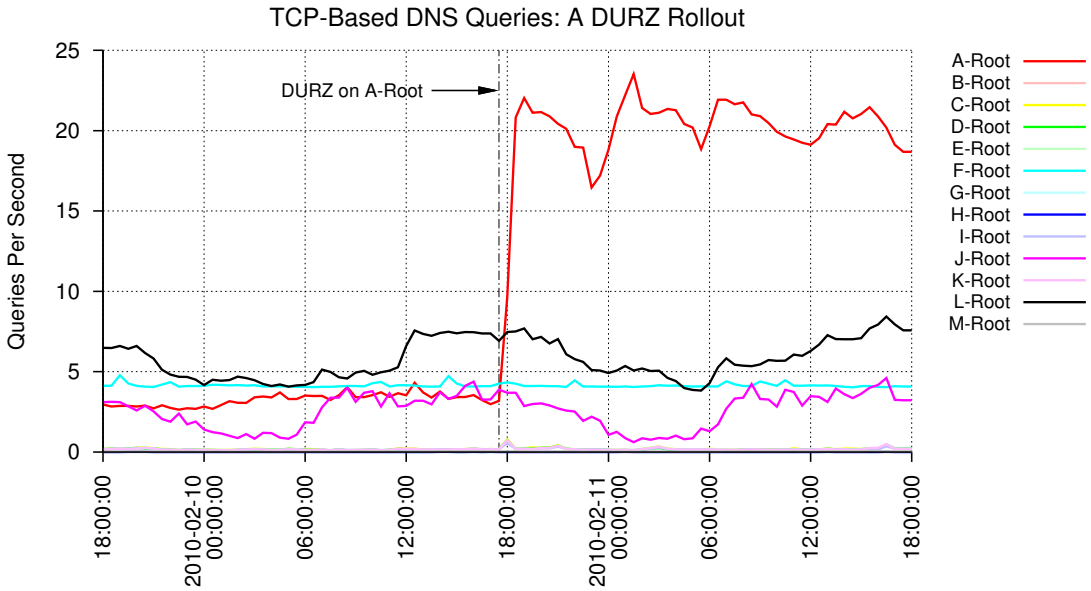

Figure 19: TCP-based DNS queries during the A DURZ rollout.

5.3.3 A DURZ Rollout

Figure 19 shows the TCP query rates during the second DURZ rollout, when the DURZ was introduced on A-Root. As with L-Root, there is a sharp increase in the rate of TCP-based DNS queries at 18:00 UTC. Nevertheless the resulting query rate is still unremarkable at 20 queries per second.

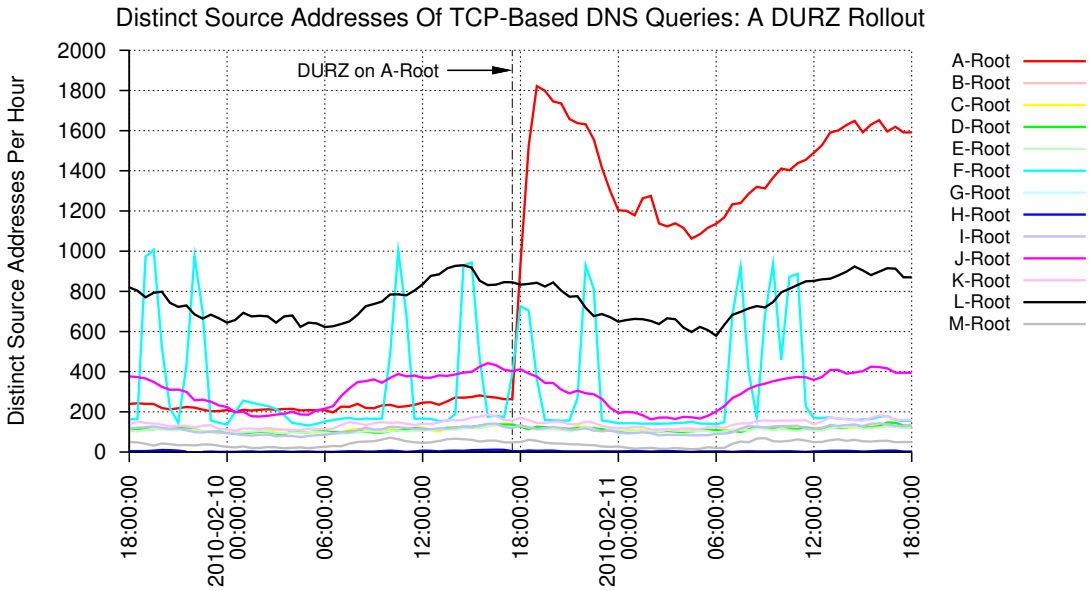

Figure 20: Distinct TCP query sources during the A DURZ rollout.

Figure 20 shows the number of distinct TCP query sources observed during the second DURZ rollout.

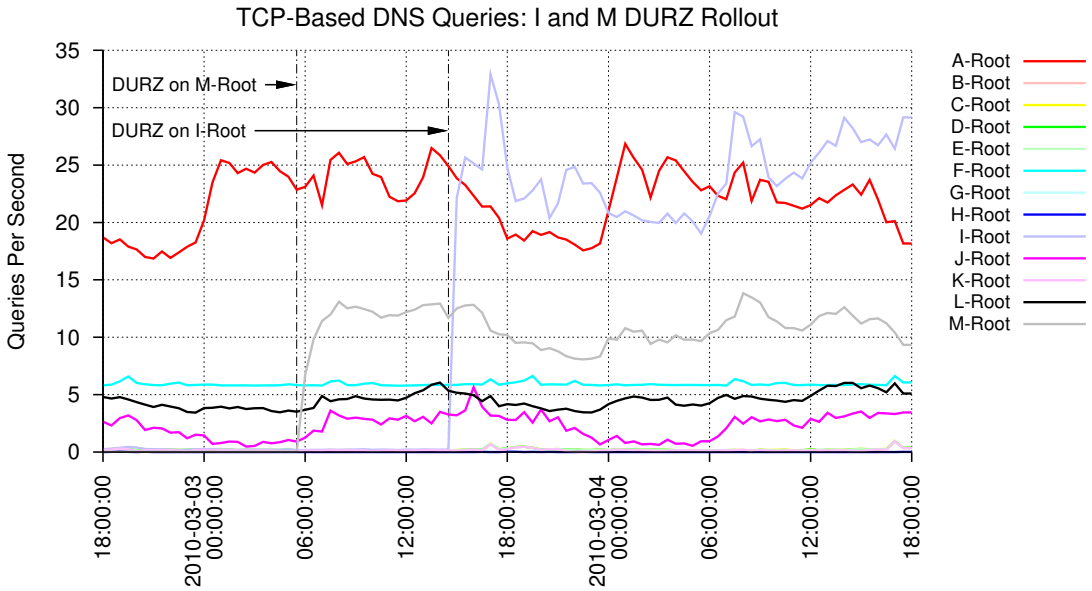

Figure 21: TCP-based DNS queries during the I and M DURZ rollout.

5.3.4 I and M DURZ Rollout

Figure 21 shows the TCP query rates during the third DURZ rollout, when the DURZ was introduced on I- and M-Root. TCP query rates to A-Root still remain at 20−25 queries per second. The relatively large increase for I-Root may still reflect some activity from the anomaly. Because the IP addresses for I-Root were obfuscated, it was more difficult to identify and exclude this traffic.

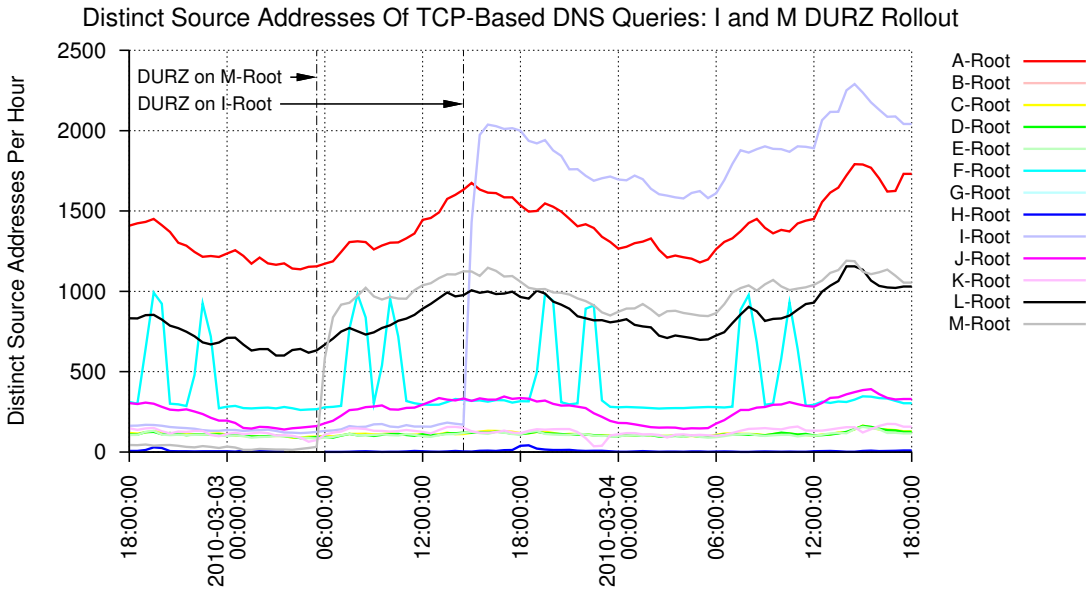

Figure 22: Distinct TCP query sources during the I and M DURZ rollout.

Figure 22 shows the number of distinct TCP query sources observed during the third DURZ rollout.

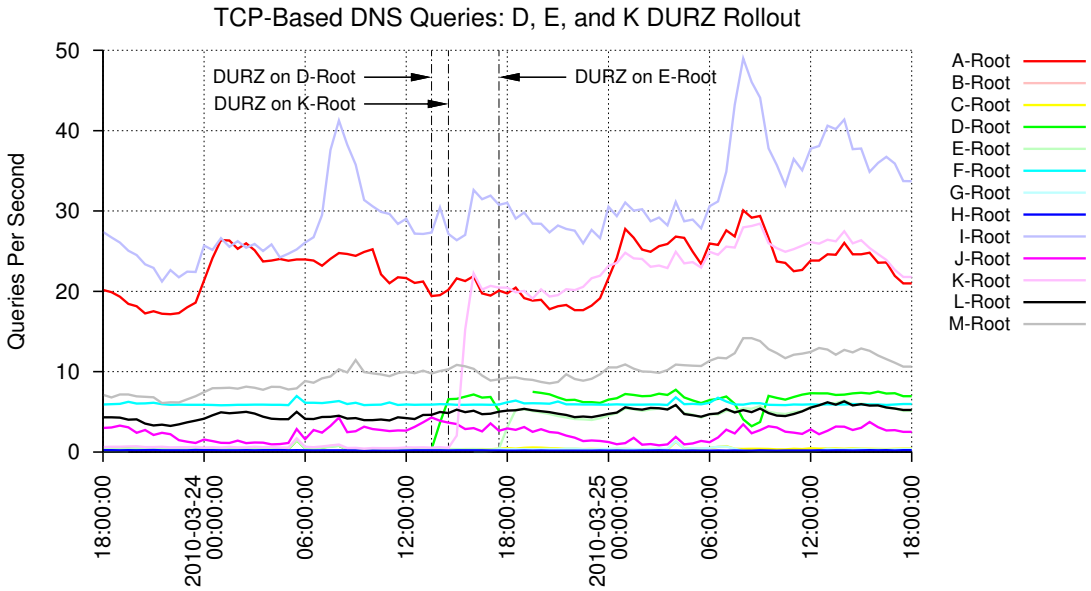

Figure 23: TCP-based DNS queries during the D, E, and K DURZ rollout.

5.3.5 D, E, and K DURZ Rollout

Figure 23 shows the TCP query rates during the fourth DURZ rollout, when the DURZ was introduced on D-, E-, and K-Root. The increase in traffic at the time that K-Root gets the DURZ is clear. Increases also can be seen for D- and E-Root, though they are small in comparison to K-Root. The picture is made slightly confusing by a gap in data reported by D-Root between around 17:52 and 19:28 on March 24.

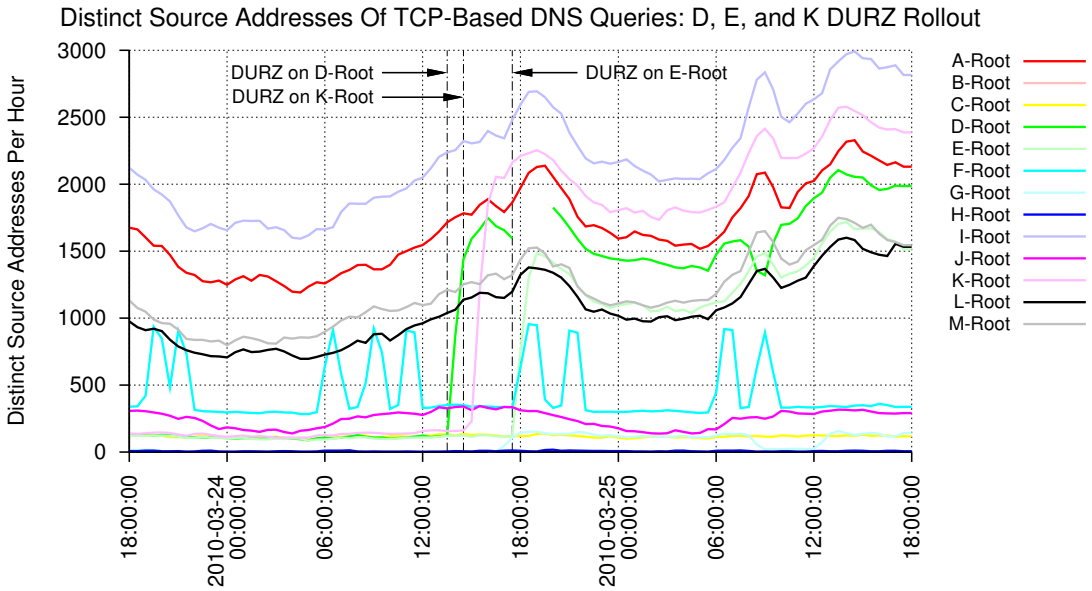

Figure 24: Distinct TCP query sources during the D, E, and K DURZ rollout.

Figure 24 shows the number of distinct TCP query sources observed during the fourth DURZ rollout. Here the changes are a little clearer than in Figure 23. It appears that not only did D-Root stop reporting at 18:00 on March 24 and 09:00 on March 25, but, based on spikes in the data for other root servers, it seems to have had reduced availability during these periods.

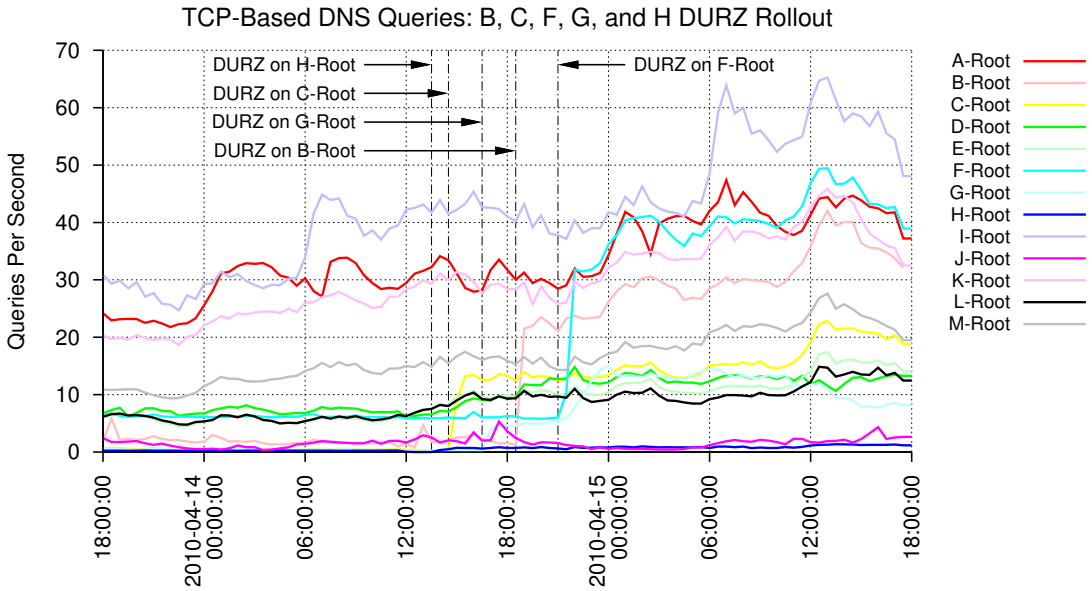

Figure 25: TCP-based DNS queries during the B, C, F, G, and H DURZ rollout.

5.3.6 B, C, F, G, and H DURZ Rollout

Figure 25 shows the TCP query rates during the fifth DURZ rollout, when the DURZ was introduced on five root servers: B, C, F, G, and H. Increases are plainly visible for C, B, F, and G. The TCP query rate also increases for H-Root, but it is difficult to see here as it is dwarfed by the other root servers. The rate for H-Root goes from around 12 queries per minute to a sustained rate of around 40 queries per minute.

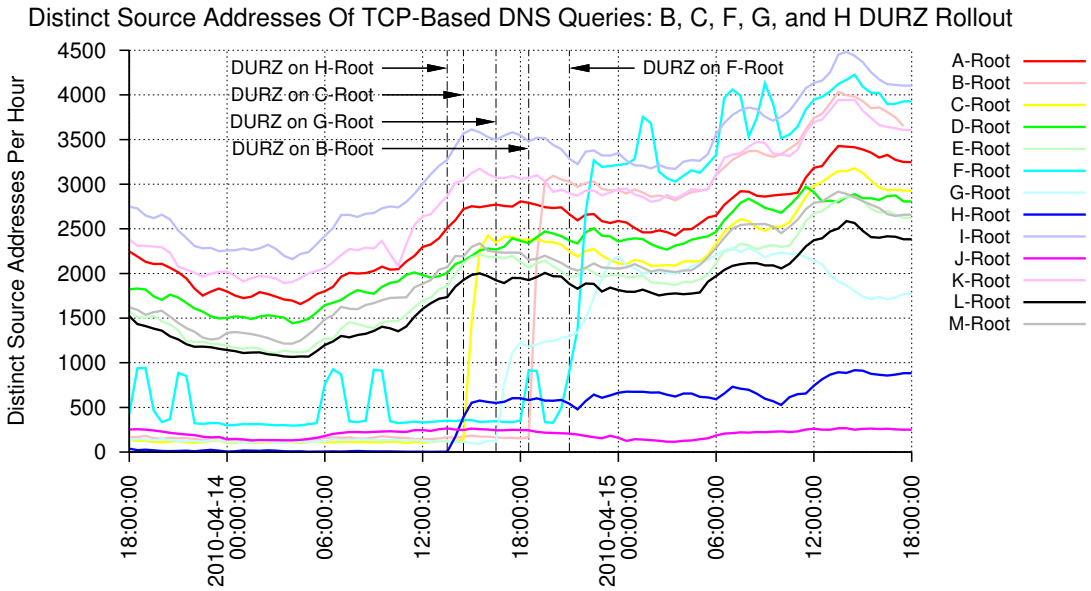

Figure 26: Distinct TCP query sources during the B, C, F, G, and H DURZ rollout.

Figure 26 shows the number of distinct TCP query sources observed during the fifth DURZ rollout. In contrast to Figure 25, the changes for all five root servers starting to serve the DURZ are clearly visible.

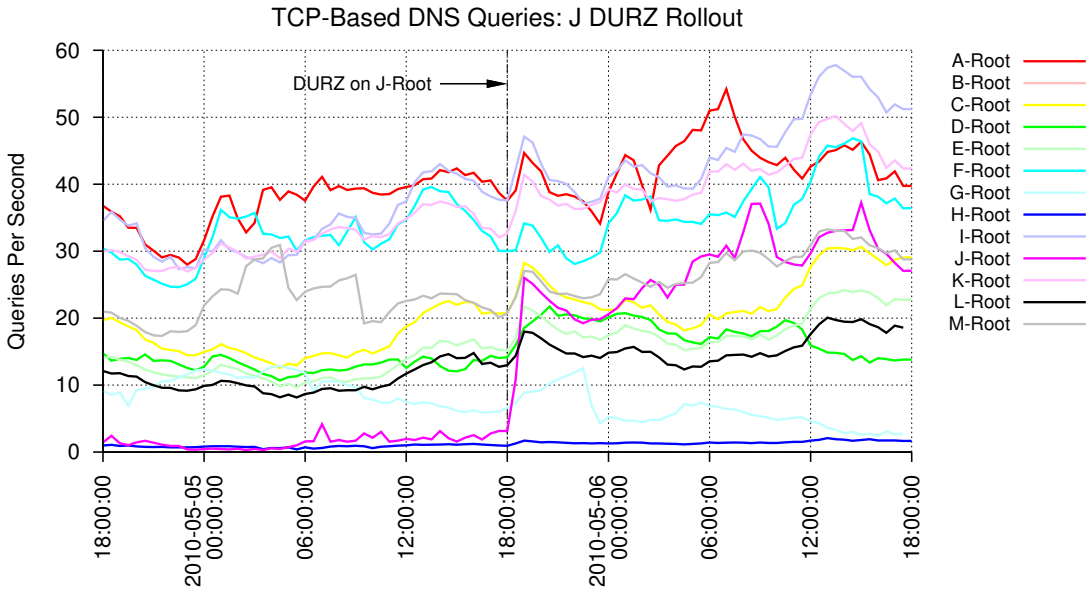

Figure 27: TCP-based DNS queries during the J DURZ rollout.

5.3.7 J DURZ Rollout

Figure 27 shows the TCP query rates during the sixth and final DURZ event, when the DURZ was rolled out on J-Root. The jump in traffic to J-Root is sharp. Simultaneous increases in TCP query rates to the other root servers are also apparent. This presumably corresponds to activity from clients which had been using J-Root as a last resort for UDP queries before finally being forced to use TCP.

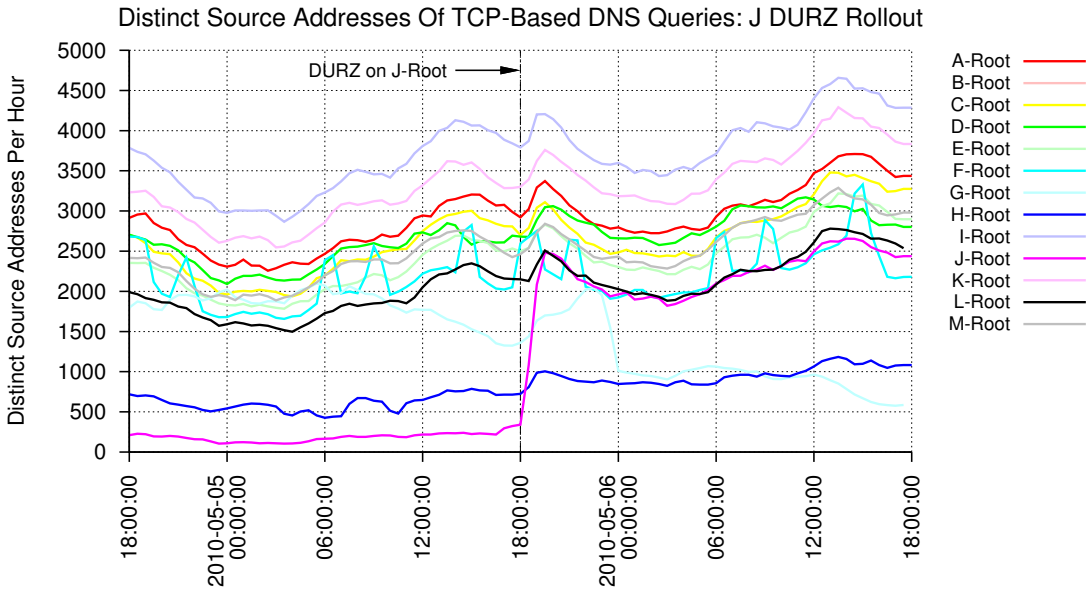

Figure 28: Distinct TCP query sources during the J DURZ rollout.

Figure 28 shows the number of distinct TCP query sources observed during the sixth DURZ rollout. All root servers register an increase in the number of source addresses sending TCP-based queries.

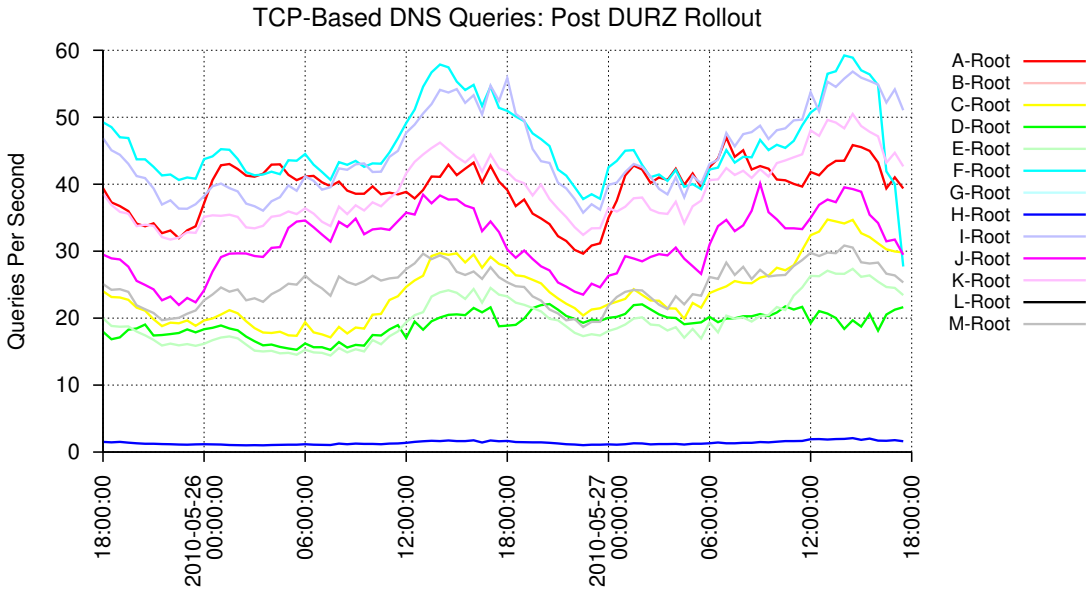

Figure 29: TCP-based DNS queries: post-DURZ.

5.3.8 Post DURZ

Finally, Figure 29 shows the TCP query rates three weeks after the conclusion of the DURZ rollout. The variations in rates seen here are due primarily to diurnal changes in traffic patterns. The highest rates correspond with daytime in the Americas and Europe.

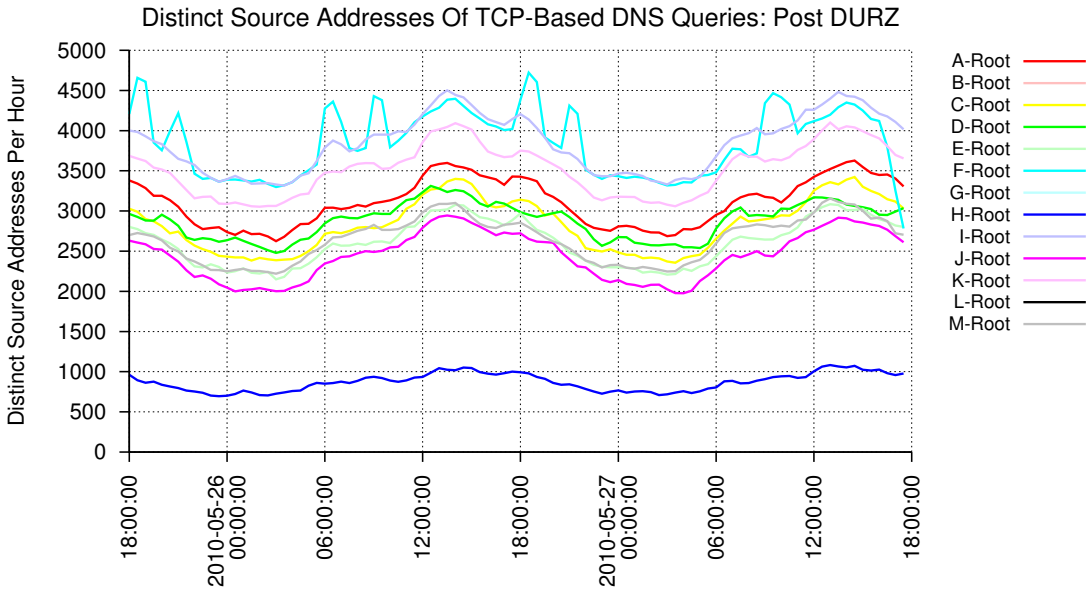

Figure 30: Distinct TCP query sources: post-DURZ.

Figure 30 also exhibits mostly diurnal changes. The anomalous spikes in the F-Root data from the large Russian search engine are still apparent. They appear attenuated compared to Figure 16 because they now account for a relatively smaller proportion of clients sending TCP-based queries.

5.4 Changes in TCP-Based DNS Query Distribution

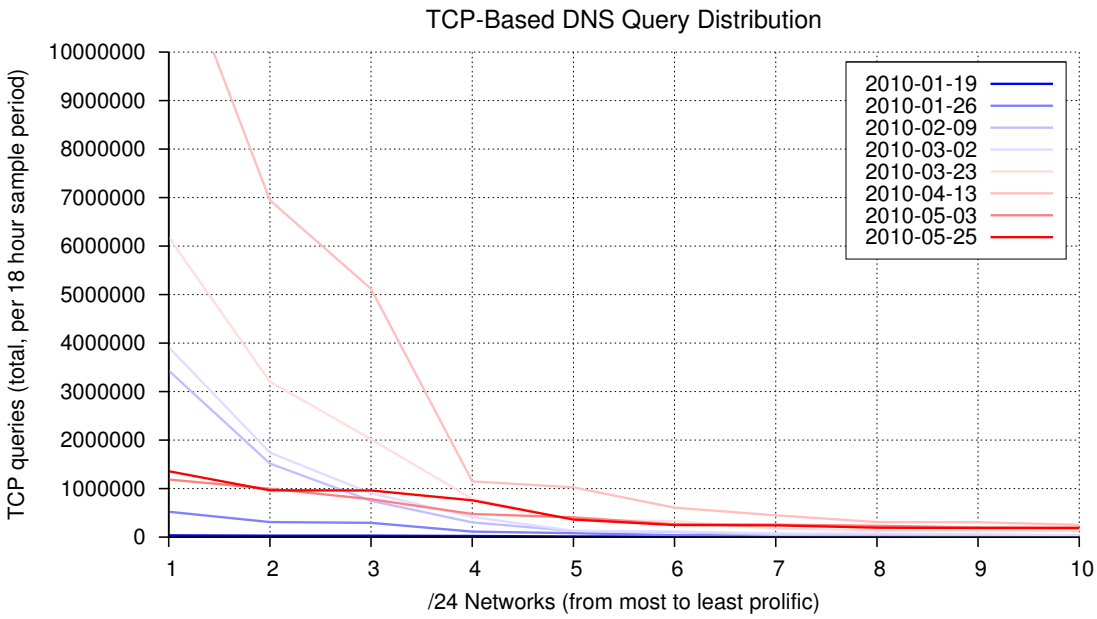

Figure 31: TCP-based DNS query distribution (top 10 /24 networks).

Figures 31 and 32 show the distribution of TCP query rates grouped by /24 network. These two plots are really a preface for Figure 33. They are intended to hint at the inverse exponential relationship between TCP query rate and query source.

Figure 33 shows the same data plotted logarithmically. Here the inverse exponential relationship is more apparent.

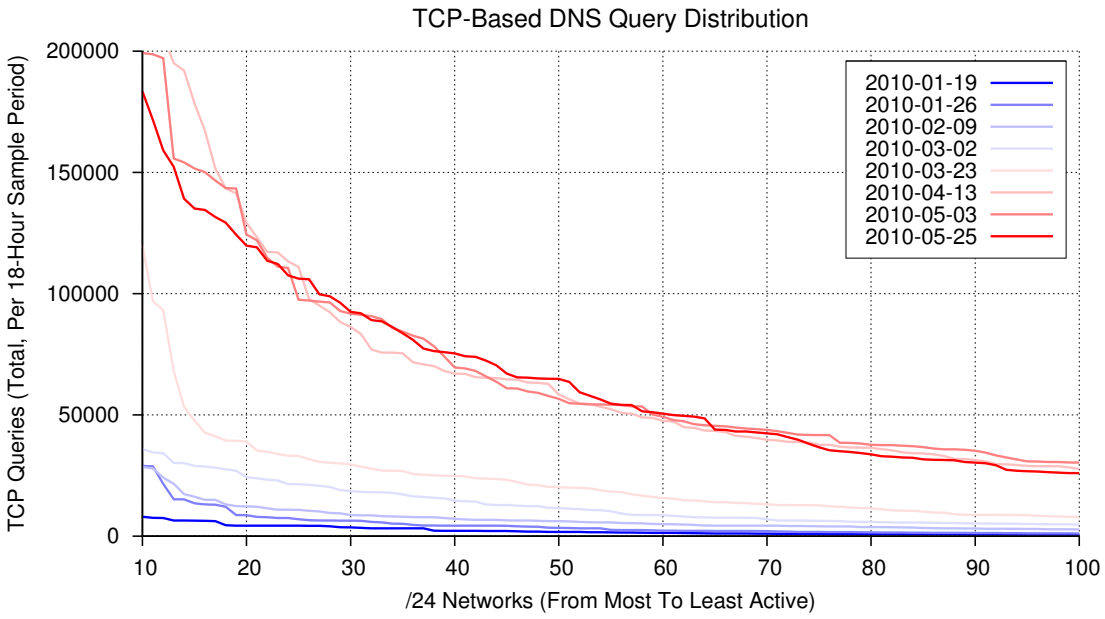

Figure 32: TCP-based DNS query distribution (next 90 /24 networks).

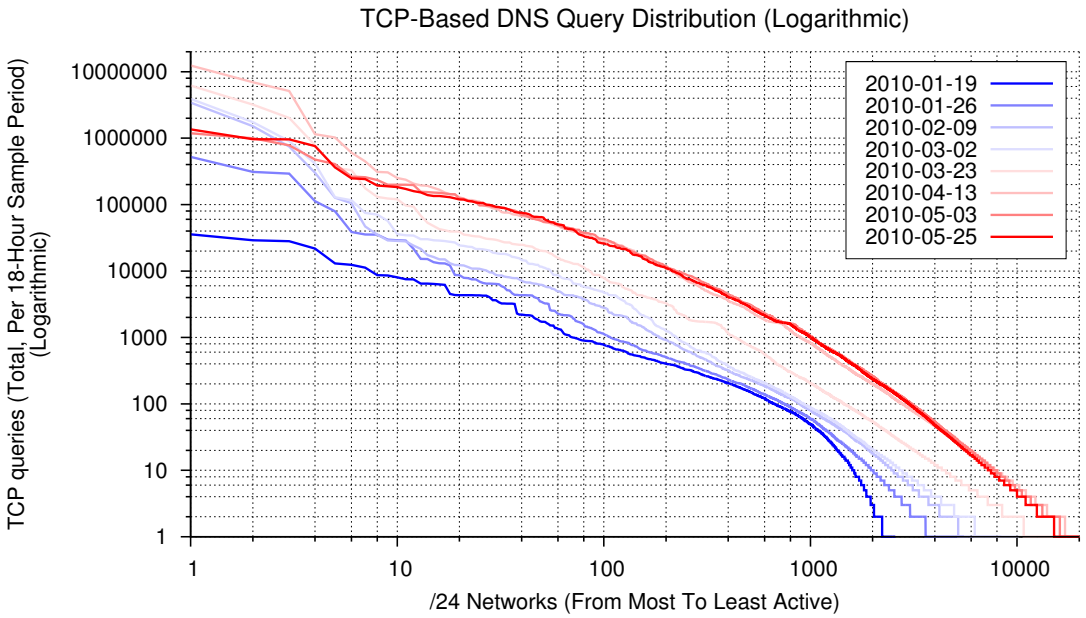

Figure 33: TCP-based DNS query distribution (logarithmic).

5.5 Production Signed Zone

On July 15 at 20:50 UTC, the production signed root zone was published, replacing the DURZ on all root servers. OARC collected data for this event for a five-day period beginning the previous day. The transition to the production signed zone was expected to have little observable effect, as only resolvers that were already attempting root zone validation should have been affected[5].

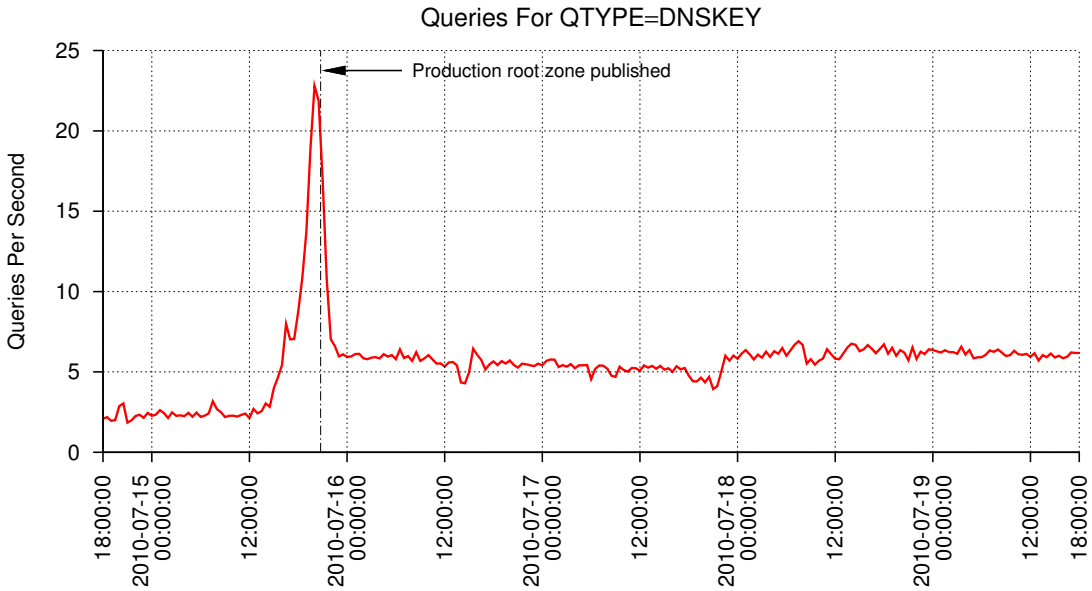

Figure 34: Queries with QTYPE=DNSKEY during the rollout of the production root zone.

There was one notable effect: a spike in the number of queries with QNAME=DNSKEY (Figure 34). It is likely this spike was caused by members of the technical community watching for the production DNSKEY to appear, either manually via `dig` or using scripts.

Afterwords, there was a sustained increase in the level of DNSKEY queries. Some of this increase may be attributable to resolver operators enabling DNSSEC validation; however, the TTL of the DNSKEY RRset is 24 hours, so unless an orderly procession of some 12,000 operators per hour enabled validation (unlikely), this increase is partially unexplained.

[5]These resolvers would have been returning SERVFAIL for the entire DNS!

6 Conclusion

The average message size of UDP-based DNS response size grew by about 40%, from 405 to 569 octets. The largest observed responses were just over 900 octets.

There is evidence that the introduction of the DURZ resulted in an increase in the number of query retries for some types of query, but it is unclear whether this corresponds to clients with path MTU issues or is simply path MTU discovery at work. The apparent absence of any problem reports strongly suggests the latter.

The number of TCP-based DNS queries to the root servers increased by approximately 1333%, from 30 per second prior to the introduction of the DURZ to around 400 per second afterwords. However, TCP-based queries, which were a miniscule 0.02% of total query traffic before the DURZ, were still only 0.17% of it afterwords. While 400 TCP connections per second may seem high, it is small relative to available capacity, particularly as the root servers comprise approximately 300 individual nodes. The number of clients using TCP for DNS queries rose by over 1800% from around 1600 distinct sources per hour to nearly 30,000. This is still a tiny fraction of all DNS clients.

7 Acknowledgements

The author would like to thank the root servers operators for their considerable time and effort in delivering data to OARC. In particular, the author thanks Magnus Sandberg of Netnod for his help in providing a secondary set of I-Root data with anomalous traffic removed, and Duane Wessels for his key role in the data collection effort. Also, several of the graphs in this report were based on ones Duane produced for various presentations.

Glossary

A resource record type
An RR type used to identify an RR that contains an IPv4 address.

A-Root (`a.root-servers.net`)
One of the thirteen root name servers on the Internet. A-Root is operated by VeriSign.

anycast
A network addressing and routing methodology in which datagrams from a single sender are routed to the topologically nearest node in a group of potential receivers all identified by the same IP address. Anycast is described in RFC 3258.

`.arpa` top-level domain
A TLD used for several technical infrastructure purposes. Its most widely used function is for reverse DNS lookups, i.e., resolving IP addresses to domain names.

Berkeley Internet Name Daemon (BIND)
The oldest and most widely deployed name server implementation on the Internet. *See* `http://www.isc.org/software/bind`.

C-Root (`c.root-servers.net`)
One of the thirteen root name servers on the Internet. A-Root is operated by Cogent Communications. *See* `http://c.root-servers.org/`.

Coordinated Universal Time (UTC)
The time standard by which clocks and time are coordinated worldwide. Similar to Greenwich Mean Time (GMT), though UTC and GMT may differ by nearly a second.

cryptographic nonce
A (usually) randomly generated number used as a one-time value to protect against replay attacks.

datagram
A basic transfer unit associated with a packet-switched network.

Deliberately Unvalidatable Root Zone (DURZ)
A special transitional DNSSEC signed version of the root zone that uses dummy DNSKEY RRs that deliberately prevent DNSSEC validation.

`dig` ("Domain Information Groper")
A command line program for querying name servers. `dig` is distributed with BIND.

DNS name server
A host or program that listens for DNS queries and returns DNS replies.

DNS name server cluster
A cluster of DNS name server "nodes" that are routed using anycast so that they appear to clients as a single name server.

DNS Security Extensions (DNSSEC)
Extensions to the DNS that add data origin authentication and data integrity. DNSSEC is described in RFCs 4033, 4034, and 4035.

DNS zone
A portion of the global DNS namespace, usually corresponding to a single administrative authority.

DNSKEY resource record type
An RR type used to identify an RR that contains a DNSSEC public key.

DO ("DNSSEC OK") bit
A one-bit flag in a DNS query signalling that the client can accept DNSSEC RRs in replies. The DO bit is defined in RFC 3225.

domain name
An identifying label in the DNS. If a domain name is used to identify a host on the Internet, it is often referred to as a "hostname".

Domain Name System (DNS)
The principal naming system for identifying hosts, services, and other resources on the Internet.

DSC (DNS Statistics Collector)
A system for collecting and examining statistics from DNS name servers. Many of the RSOs use DSC for collecting data from their respective root name servers. *See* `http://dns.measurement-factory.com/tools/dsc/`.

EDNS0 (EDNS version 0)
The first set of EDNS extensions for DNS, providing additional fields for specifying larger message sizes, additional label types, and new message flags. DNSSEC requires the use of EDNS0. EDNS0 is described in RFC 2671.

Extension mechanisms for DNS (EDNS)
A specification for extensions to the DNS. There is currently only one set of extensions, EDNS version 0 (EDNS0). EDNS is described in RFC 2671.

fpdns ("Fingerprint DNS")
A DNS fingerprinting tool that attempts to identify name server implementations based on their idiosyncratic behaviors. *See* `http://code.google.com/p/fpdns/`.

fragmentation needed (ICMP message)
An ICMP message (type=3, code=4) indicating that a datagram was too large for a receiving or intermediate host/router to process without fragmenting.

I-Root (`i.root-servers.net`)
One of the thirteen root name servers on the Internet. I-Root is operated by Netnod. *See* `http://i.root-servers.org/`.

Internet Control Message Protocol (ICMP)
A protocol for sending error and diagnostic messages between hosts and/or routers. ICMP is described in RFC 792.

Internet Protocol
The principal communications protocol used for relaying datagrams (packets) across the Internet. The Internet Protocol is described in RFC 791 (version 4) and RFC 2460 (version 6).

Internet Protocol Version 4 (IPv4)
The fourth revision of the Internet Protocol, and the one with the widest deployment. IPv4 is described in RFC 791.

Internet Protocol Version 6 (IPv6)
A version of the Internet Protocol designed as the successor to the current version, IPv4. IPv6 deployment is currently in its infancy; IPv6-based DNS queries account for fewer that 1% of all queries arriving at the root name servers. IPv6 is described in RFC 2460.

IP address
A numerical address assigned to each device on an Internet Protocol network. IP addresses are described in RFC 791 (IPv4) and RFC 2460 (IPv6).

`ip6.arpa` domain
A domain used for reverse DNS lookups of IPv6 IP addresses.

J-Root (`j.root-servers.net`)
One of the thirteen root name servers on the Internet. J-Root is operated by VeriSign.

Key Signing Key (KSK)
A DNSSEC key used to sign one or more Zone Signing Keys (ZSKs) in a DNSSEC signed zone.

Maximum Transmission Unit (MTU)
The maximum sized datagram that can be transmitted across a network.

MNAME field
A field in the SOA RR that specifies the name server that was the original or primary source of data for the zone. The MNAME field is described in RFC 1035.

MX resource record type
An RR type used to identify an RR that contains the hostname of mail server for a zone as well as its priority relative to other mail servers.

Network Address Translation (NAT)
Informally, a "NAT" is an IP network masquerading as a single IP address. Formally, NAT is the mechanism used to re-map the addresses, and is not limited to mapping to a single IP address.

NS resource record type
An RR type used to identify an RR that contains the domain name of an authoritative name server for a zone, e.g., `a.root-servers.net` for the root zone or `ns1.nic.uk` for the `.uk` TLD.

NXDOMAIN ("Non-Existent Domain")
A DNS reply with a response code (RCODE) of 3, indicating that the domain name in the corresponding query does not exist. The mnemonic NXDOMAIN is not formally defined in any RFC, but is associated with BIND and appears in a number of RFCs.

octet
A unit of data consisting of eight bits. Synonymous with "byte".

OPCODE ("Operation code")
A field in the header of a DNS message that specifies the kind of query contained in the message. Most DNS queries have OPCODE=0 ("Query", or "standard query"), though a small number have OPCODE=3 ("Notify") or OPCODE=4 ("Update"). In theory the root name servers should receive only queries with OPCODE=0, but in practice they receive ones with other opcodes, usually due to misconfigured DNS clients. The OPCODE field is defined in RFC 1035.

path MTU
The Minimum Transmission Unit (MTU) of network path between two hosts. The path MTU is the MTU size of the network hop with the smallest MTU.

path MTU discovery
A technique for dynamically discovering the path MTU of an arbitrary network path. Path MTU discovery is described in RFC 1191.

pcap (packet capture)
A file format for recording datagrams (packets) "captured" from a network interface.

priming query
A DNS query requesting the NS RRset for the root name servers. This is typically the first query a resolver sends upon starting.

PTR resource record type
An RR type used for reverse resolution, i.e., for translating IP addresses to hostnames.

QCLASS (Query Class)
The target RR class in a DNS query. Defined in RFCs 1034 and 1035.

QNAME (Query Name)
The target domain name in a DNS query. Defined in RFCs 1034 and 1035.

QTYPE (Query Type)
The target RR type in a DNS query. Defined in RFCs 1034 and 1035.

Query (OPCODE)
An OPCODE specifying a DNS message containing a standard query. The Query opcode is defined in RFC 1035.

referral
A reply from an authoritative name server containing NS RRs specifying where a resolver can obtain a more specific answer.

resolver
A software program, library, or device responsible for sending DNS queries and interpreting replies in order to obtain a resolution (translation) of the resource sought, e.g., translating a domain name to an IP address.

Resource Record (RR)
The basic data element in the DNS protocol.

Resource Record Class
The class of a DNS RR, e.g., IN (Internet), CH (Chaos), HS (Hesiod). Nearly all RRs on the Internet have Class=IN.

Resource Record set (RRset)
A set of DNS RRs with the same name, type, and class.

Resource Record Type
The type of a DNS RR, e.g., A, NS, MX.

root name server
A name server or name server cluster thats authoritative for the root zone.

Root Server Operator (RSO)
An organization which operates a root name server.

root zone
The zone at the top of the DNS hierarchy. The root zone is unnamed and is usually represented in the hierarchy by a period, or full stop (“.”). The root zone is administered by ICANN.

SERVFAIL (“Server Failure”)
A DNS reply with a response code (RCODE) of 2, indicating that the queried name server encountered an internal failure while processing a query. The mnemonic SERVFAIL is not formally defined in any RFC, but is associated with BIND and appears in a number of RFCs.

SOA (“Start Of Authority”) resource record type
An RR type that specifies authoritative information about a zone, including the primary name server, the email address of the domain administrator, the domain serial number, and several timers related to the propagation and caching of the zone. The SOA RR type is described in RFC 1035.

`tcpdump`
A common application used for “capturing” datagrams (packets) from a network interface. *See* `http://www.tcpdump.org/`.

TKEY resource record type
An RR type used to establish shared secret keys between a name server and client. The TKEY RR type is described in RFC 2930.

Top-Level Domain (TLD)
A domain at the highest level in DNS hierarchy, e.g., `.com`, `.uk`, and `.arpa`.

Transmission Control Protocol (TCP)
A stateful connection-oriented transport-layer protocol. TCP is described in RFC 793.

trust anchor
A KSK (or a hash thereof) configured for use in a validating resolver.

TTL (“Time To Live”)
A value specifying how long a resolver should cache (keep in memory) one or more DNS RRs.

Update (OPCODE)
An OPCODE specifying a DNS message containing a dynamic update. The Update opcode is defined in RFC 2136.

User Datagram Protocol (UDP)
A stateless datagram-oriented transport-layer protocol. UDP is described in RFC 768.

Zone Signing Key (ZSK)
A DNSSEC key used to sign the RRs in a zone.

References

[1] Mockapetris, P., *Domain Names - Concepts and Facilities*, RFC 1034, STD 13, November 1987. http://tools.ietf.org/html/rfc1034

[2] Mockapetris, P., *Domain Names - Implementation and Specification*, RFC 1035, STD 13, November 1987. http://tools.ietf.org/html/rfc1035

[3] Arends, R., et al., *DNS Security Introduction and Requirements*, RFC 4033, March 2005. http://tools.ietf.org/html/rfc4033

[4] Arends, R., et al., *Resource Records for the DNS Security Extensions*, RFC 4034, March 2005. http://tools.ietf.org/html/rfc4034

[5] Arends, R., et al., *Protocol Modifications for the DNS Security Extensions*, RFC 4035, March 2005. http://tools.ietf.org/html/rfc4035

[6] Abley, J. and M. Larson, *DNSSEC for the Root Zone*, Presented at RIPE 59, October 2009. http://www.root-dnssec.org/wp-content/uploads/2009/12/rootsign-ripe59-overview.pdf

[7] DNSSEC for the Root Zone. http://www.root-dnssec.org/

[8] DITL 2010 Data Collection Project. https://www.dns-oarc.net/ditl/2010

[9] tcpdump. http://www.tcpdump.org/

[10] DSC: A DNS Statistics Collector. http://dns.measurement-factory.com/tools/dsc/

[11] E. Osterweil et al., *Availability Problems in the DNSSEC Deployment*, Presented at RIPE 58, May 2009. http://www.ripe.net/ripe/meetings/ripe-58/content/presentations/dnssec-deployment-problems.pdf

[12] The SecSpider DNSSEC Monitoring Project. http://secspider.cs.ucla.edu/

[13] K. Murphy, Will DNSSEC Kill Your Internet? *The Register*, April 2010. http://www.theregister.co.uk/2010/04/13/dnssec/

[14] D. Wessels and S. Castro, *DNSSEC, EDNS, and TCP*, Presented at NANOG 46, June 2009. http://www.nanog.org/meetings/nanog46/presentations/Wednesday/wessels_light_N46.pdf.

[15] Wessels, D., *DNSSEC, EDNS and TCP*, July 2009, https://www.dns-oarc.net/node/199.

[16] fpdns. http://code.google.com/p/fpdns/

[17] D. Wessels and G. Sisson, *Root Zone Augmentation and Impact Analysis*, September 2009. http://www.icann.org/en/topics/ssr/root-zone-augementation-analysis-17sep09-en.pdf