The ISP Column
A column on various things Internet

April 2018
Geoff Huston

Measuring Root Zone KSK Trust

In September 2017 the proposed roll of the Root Zone Key Signing Key (KSK), scheduled for 11th October 2017 was suspended. I wrote about the reasons for this suspension of the key roll at the time (see http://www.potaroo.net/ispcol/2017-10/notksk.html). The grounds for this action was based in the early analysis of data derived from initial deployment of resolvers that supported the trust anchor signal mechanism described in RFC 8145. In the period since then the data shows an increasing proportion of resolvers reporting that they trust KSK-2010 (the old KSK) but not KSK-2017 (the incoming KSK). A recent presentation that reported on the current numbers was made by Matt Larson at DNS OARC 28 at the start of March (https://indico.dns-oarc.net/event/28/session/11/contribution/52/material/slides/0.pptx).

This data is surprising, in that the longer KSK-2017 is announced in the root zone the expectation is that the number of resolvers that learn to trust KSK-2017 would rise.

But this is not the only source of data about resolvers that ostensibly perform DNSSEC validation. At APNIC Labs we run a long-term measurement on the extent to which users use DNS resolvers that perform DNSSEC validation. This note describes our efforts in attempting to correlate the information generated by the DNSSEC measurement and the data generated by the RFC 8145 mechanism.

A. Counting Resolvers

We've been provided with a log of RFC8145 records that reflects the data feed from the root servers for a 24 hour period encompassing parts of the 8th and 9th March 2017. This data contained 952,077 records.

As a comparison, we took the set of resolvers that have been seen to query authoritative servers for the DNSSEC measurement experiment, using the period from 1 January 2018 to 18 March 2018. This data contains 728,286 unique resolver IP addresses.

The combined data sets list 1,438,109 unique resolver IP addresses. Only 2% of these resolvers are listed in both data sets. 49% are listed in only the 8145 data set and 48% are listed only in the Ad-based data set (Table 1).

Data Set	**Count**
Both:	33,485
Only_8145:	709,821
Only_Ad:	694,803
Total:	1,438,109

Table 1. Unique Resolver Counts

Viewed from the perspective of the Ad-based data, of the 728,288 visible resolvers seen in the AD data set, only 33,485 visible resolvers, or 5%, have generated RFC8145 signal data. Presumably, the remainder of this resolver set are either not DNSSEC-validating resolvers, or are running DNS server code that does not include

RFC8145 signal support, as these resolvers are seen making queries to an authoritative server, and presumably would make queries directly of the root servers were they validating resolvers running 8145 support code.

The analysis of the query data from the Ad data set allows us to estimate whether or not the resolvers are performing DNSSEC validation. Of these 728,288 ad-based visible resolvers, some 99,132 resolvers, or 14%, consistently query for DNSKEY and DS records, which is interpreted here as a signal that these resolvers are performing DNSSEC validation (Table 2).

Ad_Resolvers:	728,288
Validating:	99,132
Both Data Sets:	33,485
Validating:	10,588

Table 2. DNSSEC-Validating resolvers

Some 33,485 visible resolvers are seen reporting their trusted key status via RFC8145 queries. However, only 32% of these resolvers (10,588) exhibit a query behaviour that is consistent with validation. In other words, more than 2/3 of the resolvers that both generate a RFC 8145 query signal and are seen asking queries in the Ad data set are not seen to generate DS or DNSKEY queries in the Ad context. This implies that there is a significant level of signal being generated in the 8145 data set relating to resolvers that are not performing DNSSEC validating. Potential reasons include a bug in the implementation of RFC8145 where a non-validating resolver is reporting its trusted key status, or the reporting resolver is not directly visible, and is using a forwarder to pass the 8145 query towards a root server.

Of the 10,588 resolvers that are seen in both data sets and relate to a DNSSEC-validating resolver, only 6% (655 resolvers) generate a signal that shows trust only for KSK-2010, and a further 8% (841 resolvers) generate 8145 signals that show trust in only KSK-2010 and trust in both KSK-2010 and KSK2017 at various times. This component of the signal would be consistent with a chain of forwarding resolvers, where a resolver that does not normally directly query authoritative servers is passing its RFC8145 query through another resolver via a forwarding directive. The remaining 9,092 resolvers (86%) generate an 8145 signal that indicates that the resolver is using both KSK-2010 and KSK-2017 as trust anchors (Table 3).

Both:	33,485
Validating:	10,588
KSK-2010-only:	655
KSK-2010 and KSK-2017:	9,092
Mixed:	841

Table 3. DNSSEC-Validating resolvers and Trust Anchor Reporting

B. Counting Resolvers Weighted by Use

There is a big difference between resolvers and users. Some resolvers are used by a very large population of users, while some appear to have only one or two clients. If we want to estimate user impact from these numbers we need to include the factor of the relative use of these visible resolvers. Here we will apply a relative weight to the resolver numbers by multiplying each resolver by the count the number of presented experiments seen in the Ad-based system.

We saw a total of 5,938,356,970 resolver/experiment queries. Some 2,688,042,445 (45%) of these queries came from resolvers that are classified as DNSSEC-validating resolvers. Slightly more than one half of these queries, namely 3,071,060,691 queries come from resolvers that are listed in the RFC 8145 log data. (Table 4).

Query Count (Ad):	5,938,356,970
Queries from Validating Resolvers (Ad):	2,688,042,445
Queries from Non-Validating Resolvers (Ad):	3,250,314,525
Queries from resolvers listed in 8145 data:	3,071,060,691
Queries from Validating Resolvers (Ad + 8145):	1,698,776,996
Queries from Non-Validating Resolvers (Ad + 8145):	1,372,283,695

Table 4. DNSSEC-Validating resolvers and Trust Anchor Reporting by Ad Query Count

When we take the subset of data points where the resolvers are seen on both the Ad and the 8145 data set, then some 52% of queries come from resolvers that are seen in both data sets. The DNSSEC-validation rate in this subset shows 55% of these queries are from validating resolvers, and 44% of resolvers/experiment queries are from non-validating sources. In theory only DNSSEC validating resolvers should be reporting their key status, so this number is far lower than the theory would predict. Again, the factors of an implementation bug and forwarding resolvers are the most likely reasons why this number is lower than the theory predicts.

	Total	**KSK-2010**	**Mixed**	**KSK-2010+KSK-2017**
Validating	1,698,776,996	159,466,040	700,211,588	839,099,368
		9%	41%	49%
Non-Valid	1,372,283,695	164,862,115	342,487,713	864,933,867
		12%	25%	63%

Table 5. Trust Anchor Reporting Status for resolvers visible in both data sets, weighted by query count

Looking at the validating resolvers that are in both data sets, only 9% of the query-weighted count of resolvers have only the KSK-2010 trust anchor. A relatively high number (41%) report mixed signals, with some reports showing only KSK-2010 and other reports for the same resolver showing both trust anchors in place.

It is surprising that the two data sets have so little in common. Only 2% of resolvers are listed in both data sets. Of the resolvers visible in the Ad system, where the DNSSEC- validation behaviour is more likely to be known, only 5% of these resolvers report their trusted key status via RFC8145 queries. Again, this is still a surprisingly small number.

However, when these numbers are weighted by use, then the numbers change significantly, and 52% of query-weighted resolvers are also seen reporting their trusted key status in the 8145 data set.

Of the known validating resolvers that report their trusted key status, some 9% by weighted query count report trust in KSK-2010 only.

The issue here is that this number is still unrelated to any prediction of disruption related to the roll of the key to KSK-2017. Users tend to use multiple resolvers in their local configurations, and a resolution failure due to DNSSEC validation failure (which would occur with a trust key failure) would return a SERVFAIL code, which would prompt the user to re-query using alternative resolver(s). This situation is very common in the DNS environment.

As noted above, some 45% of these queries came from resolvers that are classified as DNSSEC validating resolvers, but the overall DNSSEC validation rate is a far lower at 12%. This estimate of 12% of users is an estimate of the number of users who use DNSSEC-validating resolvers exclusively, as distinct from the larger pool of users who direct their queries to a collection of both DNSSEC-validating and non-validating resolvers. When these users receive a SERVFAIL response, they re-query to a non-validating resolver.

When looking at the potential pool of users who may be impacted by a KSK roll, then the focus should be on those 12% of users who exclusively use DNSSEC-validating resolvers, seeing to what extent they exclusively use validating resolvers that are reporting trust in KSK-2010 only across the set of resolvers that they use.

C. Counting Resolver Sets

In the Ad-based DNSSEC measurement tests we use two separate DNS tasks: the first is a validly signed DNS name, and the second is an invalidly signed DNS name. We judge a user to be exclusively using DNSSEC-validating resolvers if they can successfully resolve the first name, but not the second. What happens in the second case is that each DNSSEC-validating recursive resolver that is asked to resolve this name will return the ServFail response code. This response code will prompt the user-initiated resolution process to re-query using alternative resolvers, if so configured. The set of all resolvers that are queried for this invalidly-signed name form a resolver set.

In the event of a KSK roll we anticipate a similar situation to that of an invalidly signed name. Resolvers that have not loaded KSK-2017 into their local cache of trusted keys will be unable to validate a DNSSEC-signed DNS name, and will return the ServFail response code. This will trigger the DNS resolution process to move on to another resolver. This process will terminate when either the query is sent to a resolver that has loaded KSK-2017, or the query is sent to a resolver is not DNSSEC-validating, or the process has queried all available resolvers. In other words, the users who will experience issues in a roll of the KSK are those users who use a resolver set where all the resolvers in the set both validate and have not loaded KSK-2017 into their local cache of trust anchors.

In this exercise we take all unique resolver sets that were gathered from the Ad in the period 1 January 2018 through to 18 March 2018 from users who were judged to be using exclusively validating resolvers. For each set we then break the set down into its individual resolvers, and use the 8145 signal data to categorize into one of three states:

1. they are reporting trust in KSK-2017 and KSK-2010,
2. they are reporting trust only in KSK-2010, or
3. they are not listed.

Where a resolver reports a mix of states 1 and 2 we will take the more optimistic position and place them into state 1.

We can then make a judgement about the behaviours of the resolver set itself.

- If any of the resolvers in the set are reporting trust in KSK-2017, then the set will be considered functioning in the event of a roll to use KSK-2017, i.e. the set's status is "**Good**"
- If none of the resolvers are reporting trust in KSK-2017, but some are not reporting a trust state at all then the set's status will be considered "**Unknown**"
- If all are reporting only trust in KSK-2010, then the set's status will be considered "**Bad**"

This analysis was undertaken over the data collection.
The analysis of the sets ctaontined in the Ad data are as follows:

Sets	**Total**	**Good**	**Bad**	**Unknown**
Count	68,770,592	68,159,360	1,368	609,864
Percentage	100%	99.111%	0.002%	0.887%

Table 6. KSK Trust Anchor Status for DNSSEC-validating resolver sets

This indicates that the potential level of preparedness of the DNSSEC-validating resolvers used by clients of the Ad system lies above 99%. The residual set of resolver sets have an unknown trust anchor state most of the time, while only 0.002% of the resolver sets report trust only in KSK-2010.

We now take these numbers and add a weighting based on the query-based use of these resolver sets.

Sets	Total	Good	Bad	Unknown
Query Count	127,330,706	107,963,713	438,549	18,928,444
Percentage	100%	84.79%	0.34%	14.87%

Table 7. KSK Trust Anchor Status for DNSSEC-validating resolver sets, weighted by use

The user count for this period encompassed 127 million experiments that were judged to be associated with users who exclusively used DNSSEC-validating resolvers. Of these, some 85% used a resolver set that included at least one resolver reporting that KSK 2017 had been cached as a local trust anchor, and are therefore not going to be stranded in the event of a key roll to KSK-2017. A further 15% of users are using a resolver set where at least one resolver was not recorded in the data set of resolvers reporting their local trusted key set, and no resolvers were reporting that they had trust in KSK-2017 The remaining users (0.34%) used resolver sets where all resolvers in the set were reporting trust in KSK-2010.

However, these number represents only 11.67% of the total user count as seen by the Ad-based measurement system, as these numbers refer only to those users who are exclusively using DNSSEC-validating resolvers. We can use this to estimate the outcomes for the entire population of users, using the Ad-based data set as being representative of the entire user population.

All	Non-Validating	Validating			
			KSK-2017-ready	Not-KSK-2017-ready	Unknown
100%	88.33%	11.67%	9.89%	0.04%	1.73%

Table 8. Total Estimates of KSK Roll outcomes

Based on this data, we can surmise that some 98% of users will continue to have a functional DNS resolution service were the KSK to be rolled at the present time. Of the remaining 2%, the majority of these users fall into an unknown category. Only 0.04% of users appear to use resolver sets where all the resolvers in the set have only cached KSK-2010. The ratio of Not-Ready to Ready users is 246:1. If we use this same ratio to re-assign the Unknown pool of users to these Ready and Not-Ready states the result is shown in Table 9.

All	Non-Validating	Validating		
			KSK-2017-ready	Not-KSK-2017-ready
100%	88.33%	11.67%	11.62%	0.05%

Table 8. Total Estimates of KSK Roll outcomes

In terms of preparedness for a roll to KSK-2017, this data points to an interpretation that a likely impact level is some 0.05% of the Internet population who will be without DNS resolution services in the event of a roll to KSK-2017.

Assumptions

There are two basic assumptions in this analysis that should be highlighted here.

The first is the assumption that the Ad-based measurement system encompasses a truly representative sample of Internet users. We have no data to either support or contradict this assumption, so the caveat here is that this result applies to a subset of the Internet, without any certain measurements as to the nature of this subset.

The second assumption is that the RFC 8145 signal relates to the trusted key set as held by the reporting resolver. This is not the case, and the situation of forwarders makes interpretation of this RFC 8145 query stream somewhat ambiguous.

Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre.

Author

Geoff Huston B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region.

www.potaroo.net