RSSAC002 version 2
RSSAC Advisory on Measurements of the Root Server System

An Advisory from the ICANN Root Server System Advisory Committee (RSSAC)
17 November 2015

Preface

This is an Advisory to the Internet Corporation for Assigned Names and Numbers (ICANN) Board of Directors and the Internet community more broadly from the ICANN Root Server System Advisory Committee (RSSAC). In this Advisory, the RSSAC identifies and recommends an initial set of parameters that would be useful to monitor and establish a baseline trend of the root server system.

The RSSAC seeks to advise the ICANN community and Board on matters relating to the operation, administration, security and integrity of the Internet's Root Server System. This includes communicating on matters relating to the operation of the Root Servers and their multiple instances with the technical and ICANN community, gathering and articulating requirements to offer to those engaged in technical revisions of the protocols and best common practices related to the operational of DNS servers, engaging in ongoing threat assessment and risk analysis of the Root Server System and recommend any necessary audit activity to assess the current status of root servers and root zone. The RSSAC has no authority to regulate, enforce, or adjudicate. Those functions belong to others, and the advice offered here should be evaluated on its merits.

A list of the contributors to this Advisory, references to RSSAC Caucus members' statement of interest, and RSSAC members' objections to the findings or recommendations in this Report are at end of this document.

Table of Contents

1. Introduction

In response to a desire voiced by the ICANN Board, the RSSAC made a commitment to prepare for an implementation of an early warning system that shall assist in detecting and mitigating any effects (or the absence of such effects) which might challenge the scaling and/or normal performance of the Internet's DNS root server system caused by growth of the DNS root zone itself or the Internet's use of a larger root zone file - in any dimension.

As a first step, RSSAC has begun work to determine a list of parameters that define the desired service trends for the root zone system. These parameters include the measured latency in the distribution of the root zone, the frequency of the updates, and their size. With knowledge of these parameters in hand, RSSAC can then seek to produce estimates of acceptable root zone size dynamics to ensure the overall system works within a set of parameters. The future work to define these parameters will involve RSSAC working closely with the root server operators to gather best practice estimates for the size and update frequency of the root zone.

It must be well understood that the measurements described in this document are a response to the current awareness, experience, and understanding of the Root Zone System. As time progresses more, less, or entirely different metrics may be required to investigate new concerns or defined problem statements.

2. Measurement Parameters

RSSAC has identified an initial set of parameters that would be useful to monitor and establish a baseline trend of the root server system. Monitoring these parameters should be implementable without major changes within the operations of the root zone system.

- Latency in publishing available data
- The size of the overall root zone
- The number of queries
- The response type and size distribution
- The number of sources seen

RSSAC recommends that these measurements be collected in a central location and stored in a common format for ongoing analysis. The collection location, and the frequency this data is uploaded to the central location are out of scope of this document.

Where reporting period is mentioned in this document, the reader should interpret this as the collection time window of 00:00:00 UTC to 23:59:59 UTC. As already stated, the frequency that reports are published is out of scope of this document.

Only syntactically correct DNS messages should be counted. Data-less connections and invalid, malformed, short, or non-DNS messages should not be counted.

2.1 Latency in publishing available data

Latency in publishing available data is defined as the time for the root zone to be loaded by the Root Zone Operator's nameservers once a NOTIFY has been received from the Root Zone Maintainer.

For ease of comparison this information should be reported with a resolution of seconds.

Due to the nature of operating anycast DNS clouds there may be multiple steps in the distribution of the Root Zone, dependent on the internal distribution processes at each root-server operator. Additionally the availability of anycasted instances may present as anomalies in measurements and investigators are therefore forewarned. Therefore measurements may only represent the average for 95% of the operationally-active instances of a root server per root zone serial.

2.2 The size of the overall root zone

The size of the compiled root zone is measured in wire-format AXFR response encoded as if to be transmitted in the smallest number of messages with the names in the zone and the resource records in each RRset sorted into DNSSEC order, and using compression pointers wherever possible. Even though AXFR occurs over TCP, this measurement must exclude the two-octet size prefix for each message transmitted.

This measurement may be useful to track over a longer period of time to detect any trends in the growth of the zone and correlate this to other measurements such as the latency in distribution.

The size of the compiled root zone is not expected to change from operator to operator; but in an effort to ensure consistency in the root system all operators should report the size of the root zone so if there are any differences that are seen on the platform they can be identified and remedied. Examples of differences that we should be looking for are distribution issues where the content is being changed, i.e. a truncated zone files, etc.

2.3 The number of queries

The total number of queries borne by the system of 13 root servers can be best evaluated by measuring both the IP Version and transport used as seen at each root server operator (and their anycast instances where applicable)

The number of queries should be defined as follows:

dns-udp-queries-received-ipv4
: Number of DNS queries received over IPv4/UDP transport at each root server during the reporting period.

dns-udp-queries-received-ipv6
: Number of DNS queries received over IPv6/UDP transport at each root server during the reporting period

dns-tcp-queries-received-ipv4
: Number of DNS queries received over IPv4/TCP transport at each root server during the reporting period

dns-tcp-queries-received-ipv6
: Number of DNS queries received over IPv6/TCP transport at each root server during the reporting period

The measurement of these statistics would be useful as they are indicative of the load that must be borne by the systems of 13 root servers. Further it will allow the tracking of any trends or shifts towards TCP DNS Traffic as well as different network layers.

2.4 The query and response size distribution

A DNS query is defined as a well-formed DNS transaction initiation pursuant to DNS protocol standards directed at a root server address on TCP/UDP port 53.

DNS query sizes are determined by the length of the entire DNS message. Thus, in practical terms, the transport headers (Ethernet, IP, and TCP or UDP etc) are removed leaving the DNS payload to measure. The DNS query message sizes should be recorded for both TCP and UDP.

A DNS message carried over TCP is prefixed with a 16-bit (two octet) value indicating the size of the message. Implementations should exclude these two octets in the calculation of message size.[1]

The query size distribution is defined as a list values for the number of queries received during the reporting period of a particular size range in the following:

0-15, 16-31, 32-47, 48-63, 64-79, ..., 256-271, 272-287, 288-

DNS response sizes are similarly determined by the size of the DNS message and the DNS response message sizes should be recorded for both TCP and UDP.

The response size distribution is defined as a list of values for the number of responses sent during the reporting period of a particular size range:

[1] The RSSAC Caucus debated whether or not to include these two octets in the size calculation. While some argued for its inclusion and others argued for its exclusion, there was strong agreement that consistency is more important than whether or not to count the two extra octets. In the end the Caucus agreed to exclude the size prefix.

0-15, 16-31, 32-47, 48-63, 64-79, ..., 4064-4079, 4080-4095, 4096-

This measurement could be used to analyze trends in DNS message size that may take place due to new protocol deployments, such as DNSSEC or IDN as well as client side changes (longer QNAMEs due to prefix scheme, new EDNS options) and shifts in response types (referral, signed referral, authoritative positive response, NXDOMAIN).

2.5 The RCODE distribution

The RCODE distribution is a raw count of the RCODE values observed in responses during the reporting period.

The list of RCODEs is available from IANA[2].

2.6 The number of sources seen

The number of sources seen is the number of unique IP source addresses accumulated across all instances of a root server cluster during the reporting period. There must be three values:

num-sources-ipv4
: The number of unique IPv4 addresses sending DNS queries during the reporting period

num-sources-ipv6
: The number of unique IPv6 addresses sending DNS queries during the reporting period

num-sources-ipv6-aggregate
: The number of unique IPv6 addresses sending DNS queries during the reporting period, aggregated at the /64 level

With DNSSEC validation potentially moving to the end systems/applications, the number of resolvers and validators querying to the root servers might be growing; these figures will help distinguish various contributing factors to the potential increase of the number of DNS queries reaching the root server system.

This set of metrics is marked as optional for a 3-year period following the acceptance and publication of this document by RSSAC. As experience grows with fine-grained reporting from many operational root-server instances these values can be phased in over this 3-year period. Additionally should experience show that these values provide little value overall, or constitutes a memory exhaustion attack upon monitoring infrastructure,

[2] http://www.iana.org/assignments/dns-parameters/dns-parameters.xml#dns-parameters-6

an amendment should be issued by RSSAC to deprecate the documented collection of this data prior to the end of the 3-year period.

3. Implementation Notes

In review of these metrics, RSSAC members have identified a number of concerns that might affect the collection of data, the consistency of the data collected, and some areas that may require further investigation.

Of note are:

- The single act of transferring the collected statistical data from widely deployed root server instances may affect the available bandwidth used to serve root zone queries.

- Collecting measurement data could pose as an operational impact on the root server instances. Should any impact of service eventuate, measurement data will be discarded for the higher priority of service delivery

- There are current DNS software logging limitations that inhibit the perfect collection and resolution of ‘latency in publishing available data‘ values due to the lack of zone serial numbers in AXFR/IXFR logging statements.

- Latency in publishing available data could potentially be more granular and also provide the time it takes for a root name server instance to commence serving from that zone upon receiving it, however in practical terms that reporting feature is not currently available in DNS software.

- TCP fragments are a non-trivial exercise to capture and provide meaningful statistics, it can be left to the individual root-operator to include, or not include, TCP response size statistics.

- In general the availability of tools to collect these measures is limited. Commitment by root server operators to implement these measure may be proportional with tool availability.

4. Interchange Format and Storage

Metrics should be stored in per-day, per-metric YAML formatted files.

The base format for a file is:

- Each file is a YAML "document" representing a dictionary at the top level.

- All dates are formatted using ISO 8601 including both the date and time of day. E.g., '2013-08-26T00:00:00Z'.
- The top-level dictionary contains a set of common key/value pairs:
    - 'service': this describes the service that the metric belongs to. This should be of the form "<letter>.root-servers.net".
    - 'start-period': This describes the starting date and time for the reporting period for the metric.
    - 'end-period': This describes the ending date and time for the reporting period.
    - 'metric': This is the name of the metric. The valid metric names are 'load-time', 'zone-size', 'rcode-volume', 'traffic-sizes', 'traffic-volume', and 'unique-sources'.
- The top level dictionary also contains metric-specific key/value pairs described below.

4.1 The ‘load-time’ Metric

For the 'load-time' metric, the additional key 'time' is added.

The value is a dictionary with the zone serial numbers as keys and the time delta described in "The latency in the distribution system", in seconds as a float or integer.

An example:

```
---
service: j.root-servers.net
start-period: '2013-08-26T00:00:00Z'
end-period: '2013-08-26T23:59:59Z'
metric: load-time
time:
   2013082600: 6
   2013082601: 6
```

If not load-time metric is available it should be marked with “-”

4.2 The ‘zone-size’ Metric

For the 'zone-size' metric, the additional key 'size' is added. The value is a dictionary with the zone serial numbers as keys and the size in octets as values.

An example:

```
---
service: j.root-servers.net
start-period: '2013-08-26T00:00:00Z'
end-period: '2013-08-26T23:59:59Z'
metric: zone-size
```

```
size:
  2013082600: 238218
  2013082601: 238220
```

4.3 The ‘traffic-volume’ Metric

For the 'traffic-volume' metric, additional keys are added to the top-level dictionary representing each traffic category: 'dns-udp-queries-received-ipv4', 'dns-udp-queries-received-ipv6', 'dns-tcp-queries-received-ipv4', 'dns-tcp-queries-received-ipv6', 'dns-udp-responses-sent-ipv4', 'dns-udp-responses-sent-ipv6', 'dns-tcp-responses-sent-ipv4', and 'dns-tcp-responses-sent-ipv6'. The values are the total number of requests or responses seen during the reporting period for each category.

An example:

```
---
service: j.root-servers.net
start-period: '2013-08-26T00:00:00Z'
end-period: '2013-08-26T23:59:59Z'
metric: traffic-volume
dns-udp-queries-received-ipv4: 31272
dns-udp-queries-received-ipv6: 11211
dns-tcp-queries-received-ipv4: 12
dns-tcp-queries-received-ipv6: 2
dns-udp-responses-sent-ipv4: 131079
dns-udp-responses-sent-ipv6: 16833
dns-tcp-responses-sent-ipv4: 94
dns-tcp-responses-sent-ipv6: 7
```

4.4 The ‘traffic-sizes’ Metric

For the 'traffic-sizes' metric, four additional keys are added to the top-level dictionary: 'udp-request-sizes', 'udp-response-sizes', 'tcp-request-sizes', and 'tcp-response-sizes'. The values of each key are dictionaries with the histogram bucket ranges as keys and histogram bucket counts as values. Only size ranges with nonzero counts shall be listed.

An example (with most of the histogram buckets elided):

```
---
service: j.root-servers.net
start-period: '2013-08-26T00:00:00Z'
end-period: '2013-08-26T23:59:59Z'
metric: traffic-sizes
udp-request-sizes:
  16-31: 1747
  32-47: 4990
  48-63: 7311
  ...
```

```
  272-287: 3791
  288-: 8316
udp-response-sizes:
  16-31: 4316
  32-47: 1850
  48-63: 4435
  ...
  4064-4079: 9888
  4080-4095: 1639
  4096-: 6558
tcp-request-sizes:
  16-31: 7438
  32-47: 6489
  48-63: 9905
  ...
  256-271: 6015
  272-287: 8026
  288-: 1424
tcp-response-sizes:
  16-31: 832
  32-47: 4986
  48-63: 2286
  ...
  4064-4079: 1473
  4080-4095: 2732
  4096-: 439
```

4.5 The ‘rcode-volume’ Metric

For the 'rcode-volume' metric, additional keys are added to the top-level dictionary representing numeric RCODEs. The values are the total number of responses seen during the reporting period with each RCODE. Only RCODEs with nonzero counts shall be listed. Note that RCODE is a 4-bit number as defined by RFC1035. However, the Extension Mechanisms for DNS (EDNS0) specification, RFC2671, extends RCODE to a 12-bit number. Data collection software must look for OPT RRs in response messages and use extended RCODEs if present.

An example:

```
---
service: j.root-servers.net
start-period: '2013-08-26T00:00:00Z'
end-period: '2013-08-26T23:59:59Z'
metric: rcode-volume
rcodes:
  0: 666
```

```
  1: 451
  2: 786
  3: 108
  5: 795
  16: 189
  20: 3
  3841: 99
  3842: 7
```

4.6 The ‘unique-sources’ Metric

For the 'unique-sources' Metric, three keys are added to the top-level dictionary: 'num-sources-ipv4', 'num-sources-ipv6', and 'num-sources-ipv6-aggregate'.

An example:

```
---
service: j.root-servers.net
start-period: '2013-08-26T00:00:00Z'
end-period: '2013-08-26T23:59:59Z'
metric: unique-sources
num-sources-ipv4: 2086125
num-sources-ipv6: 42941
num-sources-ipv6-aggregate: 3873
```

4.7 URL Path Standard

The interchange files should be made available using a standardized URL path scheme to aid in finding and combining the set of files from the different operators.

The path scheme is:

```
<year>/<month>/<metric>/<short-service>-<yyyymmdd>-<metric>.yaml
```

Where: 'year' is a 4-digit year, ‘month’ is a 2-digit month, ‘short-service’ is a shorter version of the service name, generally of the format of "<letter>-root".

An example:

```
2013/09/load-time/a-root-20130901-load-time.yaml
```

5. Recommendation

Recommendation 1: The RSSAC recommends each root server operator implement the measurements outlined in this advisory.

Recommendation 2: The RSSAC should monitor the progress of the implementation of these measurements.

Recommendation 3: Measurements outlined in this document should be revisited in two years to accommodate changes in DNS technologies.

6. Acknowledgments, Disclosures of Interest, Dissents, and Withdrawals

In the interest of transparency, these sections provide the reader with information about four aspects of the RSSAC process. The Acknowledgments section lists the RSSAC caucus members, outside experts, and ICANN staff who contributed directly to this particular document. The Statement of Interest section points to the biographies of all RSSAC caucus members. The Dissents section provides a place for individual members to describe any disagreement that they may have with the content of this document or the process for preparing it. The Withdrawals section identifies individual members who have recused themselves from discussion of the topic with which this Advisory is concerned. Except for members listed in the Dissents and Withdrawals sections, this document has the consensus approval of the RSSAC.

6.1 Acknowledgments

RSSAC thanks the following members of the Caucus and external experts for their time, contributions, and review in producing this Report.

RSSAC Caucus members

Joe Abley
Alejandro Acosta
Jaap Akkerhuis
David Blacka (external expert)
John Bond
John Crain
Brian Dickson
Shumon Huque
Daniel Karrenberg
Akira Kato
Peter Koch
Warren Kumari
Matt Larson
Dave Lawrence
Lars-Johan Liman
Terry Manderson
Daniel Migault
Brad Verd
Paul Vixie
Duane Wessels (document leader)
Romeo Zwart

ICANN support staff

6.2 Statements of Interest

RSSAC caucus member biographical information and Statements of Interests are available at:
https://community.icann.org/display/RSI/RSSAC+Caucus+Statements+of+Interest.

6.3 Dissents

There were no dissents.

6.4 Withdrawals

There were no withdrawals.

7. Revision History

7.1 Version 1

First version, published on November 20, 2014, is available at:
<https://www.icann.org/en/system/files/files/rssac-002-measurements-root-20nov14-en.pdf>

7.2 Version 2

RSSAC002 v2 includes the following changes from v1:

- Section 2.2 (The size of the overall root zone) was amended to clarify that TCP size prefix octets are not included in the metric.
- Section 2.4 (The query and response size distribution) was amended to clarify that TCP size prefix octets are not included in these metrics.
- Section 2.4 was amended to include 0-15 in size ranges to be tabulated.
- Superfluous quotes around YAML keys were removed from example YAML in sections 4.1 (The ‘load-time’ Metric) and 4.2 (The ‘zone-size’ Metric).
- Indentation was fixed for example YAML in sections 4.3 (The ‘traffic-volume’ Metric) and 4.6 (The ‘unique-sources’ Metric).
- Section 4.5 (The ‘rcode-volume’ Metric) was amended to clarify that nonzero counts should be omitted.