RSSAC002 version 3
RSSAC Advisory on Measurements of the Root Server System

An Advisory from the ICANN Root Server System Advisory Committee (RSSAC)
1 June 2016

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 a set of parameters that would be useful for monitoring and establishing baseline trends 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 changes in client behavior from a larger root zone file - in any dimension.

As a first step, RSSAC has begun work to determine a list of metrics that define the desired service trends for the root server system. These metrics include the measured latency in the distribution of the root zone, number of queries and responses, distribution of response types, distribution of message sizes, and the number of sources seen. With knowledge of these metrics 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 server system. As time progresses more, fewer, or entirely different metrics may be required to investigate new concerns or defined problem statements.

2. Scope of Measurements

The goal of this document is to support an understanding of the stability of the operation of the root server system.

These metrics will allow estimates of dynamics of the root zone to ensure that the overall system works within a set of parameters. Dynamics here include growth of the root zone due to increasing the number of TLDs, how protocol changes such as DNSSEC, IPv6, and Internationalized Domain Names (IDN) affect the root zone, including zone size, number of records, and may also include the rate of change of records.

An additional goal is to provide measurements that assist evaluation of best practices for the operation of the root zone.

RSSAC also recognizes that measurement of some metrics are out of scope. Specifically, the goal of this document is not to answer a wider set of research questions. Although some of the current metrics may be used in research, the timescale to review and the deployment of new metrics through the RSSAC approval process seems to be a poor

match for supporting research in general. Additionally, more suitable alternatives to support DNS research exist.

3. 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. The initial set of parameters is:

- Latency in publishing available data
- Size of the overall root zone
- The number of queries and responses
- 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 up to but not including 00:00:00 UTC the following day.

Only syntactically correct DNS messages should be counted. Data-less connections or non-DNS messages should not be counted.

3.1 Latency in publishing available data

Given the highly-distributed nature of the root server system, latency in publishing available data is of particular interest, especially as the root zone grows in size. Some root operators maintain a large number of anycast sites (e.g. 50 or more). Additionally, some operators intentionally locate servers in geographic regions with relatively poor Internet connectivity. Such locations can present a challenge to zone distribution. For these reasons, RSSAC recommends measuring latency in publishing root zone data.

Latency in publishing available data is defined as the time elapsed between receipt of a NOTIFY message from the Root Zone Maintainer until 95% of the operator's name servers have loaded the new zone and are ready to serve it.[1] The elapsed time is reported as number of seconds.

[1] When operating large numbers of anycast DNS servers over a wide area, there will tend to be some servers that are located in areas that may have higher than normal latencies. The 95th percentile measurement is not to have these areas unduly skew the reported values.

Although this metric is reported as number of seconds, consumers of the measurements should not assume that the data is accurate to the second. Due to potential differences in mechanisms used by different operators in calculating this metric, differences in the order of seconds to minutes, should be ignored when used for analysis. Additionally, it is understood that, due to the nature of the root server system, any reported changes in the value of this metric on the order of seconds or minutes have no relevance with regard to the behavior and stability of the root server system.

3.2 The size of the overall root zone

Tracking this measurement over a long period of time may be useful in detecting any trends in the growth of the zone and correlating this to other measurements such as the latency in distribution.

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.

Earlier versions of this document stated that each operator should measure the size of the root zone. RSSAC now recommends that only the Root Zone Maintainer report this metric, for the following reasons:

1. The metric's definition was given to avoid the situation that different operators report significantly different sizes for the root zone. While sufficiently specific, this definition was difficult to measure in practice.
2. All operators that implemented this measurement did so with custom software external to the actual distribution of the zone within their operations.
3. All operators reported essentially identical values.

3.3 The volume of traffic

Knowing the amount of traffic entering into and emerging from the root server system is fundamental to evaluating its stability. This metric informs us about potential differences in traffic received by different root server operators, about long-term gradual changes in overall traffic levels, as well as sudden changes due to attacks or other incidents. For these reasons, RSSAC recommends measuring the number of queries received and responses sent by root servers.

Historically, DNS was carried exclusively over UDP and addressed using IPv4. While UDP and IPv4 still account for most of root name server traffic, TCP and IPv6 are continuing to increase in prevalence. Therefore, RSSAC recommends measuring the root server traffic volume both by transport protocol and by IP version. Such measurements may assist root server operators with future hardware, network, and overall capacity planning.

The traffic volume counters are 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.

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

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

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

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

Consumers of traffic volume data are hereby advised that under normal operations, response message counts may be less than query counts. Additionally, under certain types of attacks, response counts may be much lower than query counts due to various types of rate limiting and filtering. Differences in implementations of these metrics may also lead to differences in the value of “queries - responses” for different operators.

3.4 The query and response size distribution

To understand the interaction of DNS message sizes, protocol evaluation, and their interaction with underlying transport protocols (UDP with potential fragmentation and TCP), RSSAC recommends measuring distribution of query and response sizes.

A DNS query is defined as a sufficiently well-formed DNS transaction initiation pursuant to DNS protocol standards directed at a root server address over TCP or UDP to a port assigned by IANA for DNS service.

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.[2]

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:

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 (e.g. longer QNAMEs due to prefix scheme, shorter QNAMEs due to QNAME minimization, new EDNS options etc.) and shifts in response types (referral, signed referral, authoritative positive response, NXDOMAIN).

3.5 The RCODE distribution

To understand trends in the nature of queries received by root name servers, RSSAC recommends measuring RCODEs, the response codes to DNS queries.

The RCODE distribution is a raw count of the RCODE values observed in responses generated by the operator's name servers during the reporting period. Note in particular that this measurement should exclude any DNS response messages that may be sent *to* a root name server.

[2] 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.

Note that RCODE is a 4-bit number as defined by RFC1035. However, the Extension Mechanisms for DNS (EDNS0) specification, RFC6891, extends RCODE to a 12-bit number. Data collection software must be aware of extended RCODE values in response messages and report them if present.

The list of RCODEs is available from IANA.[3]

3.6 The number of sources seen

To understand trends in the number of clients of the root server system, RSSAC recommends measuring the number of distinct query sources. With DNSSEC validation potentially moving to the end systems and 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.

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. Source addresses must be taken only from query messages sent *to* the server.

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

This set of metrics is marked as optional for a 3-year period following the acceptance and publication of version 1 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. In case experience shows that these values provide little value overall, or constitute a memory exhaustion attack upon monitoring infrastructure, an amendment should be issued by RSSAC to deprecate the documented collection of this data.

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

4. 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.

- Implementations of these metrics that use packet capture techniques may need to implement TCP reassembly to properly capture DNS messages delivered over TCP. Since TCP reassembly is non-trivial, it is left to individual root operators whether or not to include metrics from DNS-over-TCP.

- In general, the availability of tools to collect these measurement data is limited. Commitment by root server operators to collect these measurement data may be proportional with tool availability.

5. 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 (which shall always be midnight UTC)., e.g., '2013-08-26T00:00:00Z'.
- The top-level dictionary contains a set of mandatory common key/value pairs:
    - 'version': this describes the version of RSSAC002 statistics.

    - '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.
    - '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.
- Key value pairs in a YAML document are unordered, meaning that they may appear in random order at different times or from different publishers.
- The RSSAC recommends that implementations of this specification utilize third party YAML libraries for reading and writing to reduce the chance of errors in processing.

5.1 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 section 3.1 "Latency in publishing available data", in seconds as an integer.

An example:

```
---
version: rssac002v3
metric: load-time
start-period: '2016-01-01T00:00:00Z'
time:
  2016010100: 811
  2016010101: 711
service: a.root-servers.net
```

If no load-time metric is available, it should be marked with “-”

5.2 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:

```
---
version: rssac002v3
service: root-servers.net
start-period: '2013-08-26T00:00:00Z'
metric: zone-size
size:
```

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

5.3 The 'traffic-volume' Metric

For the 'traffic-volume' metric, additional keys are added to the top-level dictionary representing each traffic category as described in section 3.3: '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:

```
---
version: rssac002v3
service: a.root-servers.net
start-period: '2016-01-01T00:00:00Z'
metric: traffic-volume
dns-udp-queries-received-ipv4: 4172948209
dns-udp-queries-received-ipv6: 198112796
dns-tcp-queries-received-ipv4: 52823651
dns-tcp-queries-received-ipv6: 1481265
dns-udp-responses-sent-ipv4: 4166894695
dns-udp-responses-sent-ipv6: 198080862
dns-tcp-responses-sent-ipv4: 45241791
dns-tcp-responses-sent-ipv6: 177961
```

5.4 The 'traffic-sizes' Metric

For the 'traffic-sizes' metric, four additional keys are added to the top-level dictionary as described in section 3.4: '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):

```
---
version: rssac002v3
service: a.root-servers.net
start-period: '2016-01-01T00:00:00Z'
metric: traffic-sizes
udp-request-sizes:
  16-31: 512835421
  32-47: 2576751251
  48-63: 1039338385
  ...
```

```
  256-271: 79527
  272-287: 26329
  288-: 40691
udp-response-sizes:
  16-31: 1271477
  32-47: 150135
  48-63: 4288688
  ...
  1440-1455: 68
  1456-1471: 6514
  1472-1487: 4638
tcp-request-sizes:
  16-31: 3041341
  32-47: 32140174
  48-63: 13308536
  ...
  256-271: 1808
  272-287: 1174
  288-: 4644733
tcp-response-sizes:
  16-31: 2144
  32-47: 1409
  48-63: 191
  ...
  2304-2319: 37
  2400-2415: 49
  4096-: 554
```

5.5 The 'rcode-volume' Metric

For the 'rcode-volume' metric, additional keys are added to the top-level dictionary representing numeric RCODEs as described in section 3.5. The values are the total number of responses seen during the reporting period with each RCODE. Only RCODEs with nonzero counts shall be listed.

An example:

```
---
version: rssac002v3
service: a.root-servers.net
start-period: '2016-01-01T00:00:00Z'
metric: rcode-volume
0: 1692304065
1: 600937
2: 1570
3: 2716752968
```

```
4: 192263
5: 1262982
6: 2149
7: 1192
8: 1249
9: 1127
10: 1158
11: 1248
12: 1032
13: 985
14: 1413
15: 1164
16: 8
```

5.6 The 'unique-sources' Metric

For the 'unique-sources' Metric, three keys, as described in section 3.6, are added to the top-level dictionary: 'num-sources-ipv4', 'num-sources-ipv6', and 'num-sources-ipv6-aggregate'.

An example:

```
---
version: rssac002v3
service: a.root-servers.net
start-period: '2016-01-01T00:00:00Z'
metric: unique-sources
num-sources-ipv4: 3740666
num-sources-ipv6: 182811
num-sources-ipv6-aggregate: 114142
```

5.7 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
```

6. Operator-specific Metrics

In some cases, individual root server operators may wish to publish statistics beyond what this document prescribes. In order to avoid collision and ambiguity in names, operator-specific metrics should be prefixed with operator's letter ([a-m]-root).

For example:

```
version: rssac002v3
service: d.root-servers.net
start-period: '2016-01-01T00:00:00Z'
metric: d-root-XYZ-metric
sample-xyz-metric-1: 23498
sample-xyz-metric-2: 25678
```

7. Recommendations

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.

8. 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.

8.1 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
Alejandro Acosta
Bruce Crabill
Duane Wessels (document leader)
Jaap Akkerhuis
John Bond
John Heidemann
Ondřej Surý
Rao Naveed Bin Rais
Ray Bellis
Romeo Zwart
Zhiwei Yan

ICANN support staff
Andrew McConachie
Kathy Schnitt
Steve Sheng (editor)

8.2 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.

8.3 Dissents

There were no dissents.

8.4 Withdrawals

There were no withdrawals.

9. Revision History

9.1 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>

9.2 Version 2

Second version, published on January 7, 2016, is available at: < https://www.icann.org/en/system/files/files/rssac-002-measurements-root-07jan16-en.pdf>

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.

9.3 Version 3

RSSAC002v3 included the following changes from v2.

- Section 2 (The scope of the Measurement) was added defining the scope of the measurements.
- Section 3.1 (Latency in publishing available data) was amended as the time elapsed between receipt of a NOTIFY message from the Root Zone Maintainer

until 95% of the operator's name servers have loaded the new zone and are ready to serve it. Although 'load-time' is reported as seconds, we note that it should not be assumed to have such accuracy.

- Section 3.2 (The size of the overall root zone) was amended to recommend only the root zone maintainer to collect and report this metric, and does not require root server operators to collect and report this metric.
- Section 3.3 (The volume of traffic) was amended to clarify the definitions of responses, as well some discussion on why queries and responses might differ.
- Section 3.4 (The query and response size distribution) was amended to clarify that it is only for responses sent FROM the root server (not TO the server).
- Section 4 (Implementation Notes) was amended stating that TCP reassembly is non-trivial and therefore including data from DNS-over-TCP is optional.
- Section 5 (Interchange Format and Storage) was amended by adding RSSAC002 version number in the YAML file, removing the end-period for the metric (all measurements cover a 24 hour period and the start time is sufficient), reiterating that key value pairs in a YAML document are unordered, and recommending third party YAML libraries be used for reading and writing to reduce the chance of errors in processing.
- Section 6 (Operator Specific Metrics) was added defining the syntax of operator specific metrics.