RSSAC047: RSSAC Advisory on Metrics for the DNS Root Servers and the Root Server System

An Advisory from the ICANN Root Server System Advisory Committee (RSSAC)
12 March 2020

Preface

In this report, the ICANN Root Server System Advisory Committee (RSSAC) presents a set of metrics for the domain name system (DNS) root servers as well as for the root server system (RSS). The audience of this report is the Board of Directors of the Internet Corporation for Assigned Names and Numbers (ICANN) and, more broadly, the Internet community.

The RSSAC advises the ICANN community and Board on matters relating to the operation, administration, security, and integrity of the Internet's root server system. It has the following responsibilities:

1. Communicate on matters relating to the operation of the root servers and their multiple instances with the Internet technical community and the ICANN community.
2. Communicate on matters relating to the administration of the root zone with those who have direct responsibility for that administration.
3. Engage 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 the root zone.
4. Respond to requests for information or opinions from the ICANN Board of Directors.
5. Report periodically to the Board on its activities.
6. Make policy recommendations to the ICANN community and Board.

The RSSAC has no authority to regulate, enforce, or adjudicate. The advice offered in this report should be evaluated on its merit.

The RSSAC has relied on the RSSAC Caucus, a group of DNS experts who have an interest in the root server system to perform research and produce this publication.

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

Table of Contents

1 Introduction

In this report, the RSSAC:

- Defines measurements, metrics, and thresholds that root server operators (**RSOs**) meet to provide a minimum level of performance. The thresholds are based on technical metrics designed to assess the performance, availability, and quality of service that each root server identifier (**RSI**) provides. The thresholds and the metrics on which they are based are included as the RSSAC's input to a yet-to-be defined evaluation process for future RSOs.
- Defines system-wide, externally verifiable metrics and thresholds which demonstrate that the root server system (**RSS**) as a whole is online and serving correct and timely responses.

The report is organized as follows:

- Section 2 provides background and scope for the work.
- Section 3 defines some requirements for the vantage points.
- Section 4 discusses some general points about metrics and measurements, including some high-level requirements for the measurement system.
- Section 5 defines RSI-related metrics and thresholds on availability, response latency, correctness, and publication latency.
- Section 6 defines RSS-related metrics and thresholds on availability, response latency, correctness, and publication latency.
- Section 7 presents a summary of the metrics and the thresholds described in Sections 5 and 6.
- Section 8 presents the recommendations of the report.
- Section 9 contains some example results of the measurement.
- Section 10 contains a list of acknowledgements, contributors to this report, and objections to the findings in this report.
- Section 11 is a revision history of this report.

The metrics described in this report are based on the current awareness of, experience with, and understanding of the root servers and RSS. In the future, different metrics may be required to better understand the RSS and quantify performance, availability, and quality of service. As the metrics and this document are revised, current and future RSOs may have to adapt their architectures over time to meet the changing needs of the RSS.

As described in RSSAC023, RSOs have been operating the root service as a best-effort service for the global Internet community since the inception of the DNS.[1] To date, the RSOs have borne the cost of service operations, mostly without financial engagement from the non-RSO stakeholders. As stated in RSSAC037, a sustainable model of funding for the RSS is necessary.[2]

[1] See https://www.icann.org/en/system/files/files/rssac-023-04nov16-en.pdf

[2] See https://www.icann.org/en/system/files/files/rssac-037-15jun18-en.pdf

Therefore, this document can be a starting point for a future document describing actual contractual obligations with RSOs. Absent such contractual obligations, none of the measurements in this document are binding on any of the RSOs.

2 Background and Scope

2.1 Purpose of Metrics and Thresholds

The metrics, algorithms, and processes described in this report are designed to evaluate whether or not the RSS and its individual RSOs are meeting minimal levels of performance. In other words, these metrics and thresholds will be used to measure service levels for both individual RSOs, and the whole RSS.

The impetus for this work comes from RSSAC037, in which the RSSAC laid out a governance model for the RSS. However, the RSSAC expects this work to be useful regardless of progress made on the governance evolution. The metrics defined here provide a way to show when RSOs are, or are not, meeting minimum performance levels. They also provide a way to show that the RSS as a whole is, or is not, meeting performance levels. Although there is no single organization that could be considered accountable for RSS performance, groups such as RSSAC and other stakeholders can be informed when performance levels are not met and take appropriate action when necessary.

While not dependent on the implementation of RSSAC037, this work can inform the implementation work on RSSAC037 in the following ways:

- A future manifestation of the Performance Monitoring and Measurement Function (PMMF) could use the technical metrics and thresholds defined in this report as a starting point to define its rules to assess the performance, availability, and quality of service that each RSO provides, thus bringing technical accountability to the RSOs.
- RSSAC037 states that Service Level Expectations (SLEs) should exist between the stakeholders that provide funding and RSOs that receive funding. Metrics and thresholds for the RSOs defined in this report can be used as a starting point for further discussions on the technical and performance requirements in the SLE.

Secondly, while this report focuses on only minimal performance expectations, the RSSAC recognizes that, with the evolution of the governance model, RSOs may enter into future service contracts which could include Service Level Agreements (SLAs). The RSSAC expects that the metrics defined here will be useful in an SLA context. Based on discussions during the preparation of this report, the RSSAC further expects that any SLA thresholds would be stricter (if possible) than the ones provided here.

Thirdly, the metrics and thresholds defined in this report can also be used by RSOs and others to identify situations where the RSS as a whole is degrading in performance, and actions need to be taken collectively.

2.2 Uses Not in Scope

The metrics described in this report are not necessarily suitable for use in research into root server performance trends, nor for making comparisons between RSOs.

Various systems on the Internet use the RSS for purposes other than caching recursive DNS resolution. The thresholds in this report are explicitly not designed with those systems in mind. Those systems include, for example, resolvers that do not cache DNS responses according to the DNS protocol, hosts that attempt to connect to a root server instance to check for Internet connectivity, and software that creates domain names that are expected to not exist in the DNS as a way to probe for middlebox effects.

The RSS is quite large, with instances at approximately 1000 sites across the globe at the time of publication. When measuring a system of this size, the number and location of vantage points can have a dramatic impact on metrics such as response latency and availability. Any effort to reasonably characterize the performance for a large population of DNS clients (such as all the Internet's recursive name servers) would require many, many vantage points. For the purpose of service level metrics, the RSSAC envisions a relatively small set of vantage points, making such characterizations out-of-scope.

During preparation of this report, the RSSAC Caucus had a number of discussions on the distinction between "minimal" and "good" levels of performance, and whether there should be separate thresholds for both. At this time there are only recommendations on minimal levels of performance. "Good" levels of performance are out-of-scope for this version but may be considered in future versions.

At the moment, the measurement system does not prevent man in the middle attacks (on path), it does not validate the entire root zone every interval, and it is not guaranteed to observe every anycast instance.

2.3 Relationship with Prior Work

In RSSAC001, the RSSAC defines the expectations that users might reasonably hold of both the RSS and the RSOs.[3] The nineteen expectations are grouped into the areas of infrastructure, service accuracy, service capability, operational security, diversity of implementation, monitoring and measurements, and communication.

While RSSAC001 specifies the expectations, this report defines a set of specific and externally verifiable metrics to assess the performance, availability, and quality of service that each RSO and the overall RSS provides.

RSSAC002v3 defines a set of self-reported metrics to support an understanding of the stability of the operation of the RSS and allows estimates of the dynamics of the root zone to ensure that the overall system works within a set of parameters.[4] These metrics include the measured latency in

[3] See https://www.icann.org/en/system/files/files/rssac-001-root-service-expectations-04dec15-en.pdf

[4] See https://www.icann.org/en/system/files/files/rssac-002-measurements-root-06jun16-en.pdf

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.

The metrics defined in this report are different from, and complementary to, those in RSSAC002v3. Whereas those are self-reported, these are designed to be actively measured by an external party. Whereas those are mostly tabulating various aspects of traffic to individual RSOs, these assess the performance, availability, and quality of service that each RSO and the overall RSS provides. Lastly, there are no thresholds associated with the RSSAC002v3 metrics.

RFC 7720 defines basic protocol and deployment requirements for RSOs. It does not define any measurements, metrics, or thresholds for the RSS or individual RSOs.[5]

RSSAC024 defines key technical elements of potential new root operators that would be a critical part of any potential RSO designation process.[6] The technical elements for evaluation include conformance to RSSAC001 and RFC 7720, design, experience and networking, diversity, documentation, and other miscellaneous tests.

While RSSAC024 specifies a broad set of technical elements for evaluation, this report proposes a set of narrow but very specific set of metrics and thresholds an RSO is expected to meet. It complements the work for RSSAC024 by providing the current and future operators a picture of what minimum performance for the RSS looks like and what is expected from each RSO.

If any other RSSAC reports are updated in a way that impacts this one, we expect this report to be updated as well.

2.4 Terminology

Many terms in this report (such as “instance” and “root server system”) are defined in RSSAC026.[7] The following definitions are specific to this report.

Measurement: An individual data point, taken at a single point in time, from a single location. Example: 27 millisecond response latency.

Metric: Aggregation of measurements over a range of time and locations, using well-defined mathematical processes, into one or more representational values. Examples: 35 millisecond median response latency; 96% response availability.

Threshold: Upper or lower limits on metrics, beyond which actions (e.g., alerting) should be taken. Thresholds are determined by people or organizations having knowledge and experience in the operation of the DNS and the RSS.

Vantage Point: A networked computer system from which measurements are made.

[5] See https://datatracker.ietf.org/doc/rfc7720/

[6] See https://www.icann.org/en/system/files/files/rssac-024-04nov16-en.pdf

[7] See https://www.icann.org/en/system/files/files/rssac-026-14mar17-en.pdf

Collection System: Measurements from vantage points are delivered to a central collection system for additional processing, aggregation into metrics, and comparison of the metrics to thresholds.

3 Vantage Points

3.1 Number of Vantage Points

The RSSAC recommends that measurements be made from approximately 20 vantage points. This number has been chosen to strike a balance between two competing goals: coverage and manageability. While more vantage points can increase coverage of the RSS, it also increases complexity and difficulty in managing a large number of systems. As experience is gained in the operation and interpretation of these metrics, a future update may recommend a larger number of vantage points.

3.2 Location of Vantage Points

Vantage points shall be distributed approximately evenly among the five following geographic regions:

- Africa
- Asia/Australia/Pacific
- Europe
- Latin America/Caribbean Islands
- North America

Vantage points should be located within major metropolitan areas. There should only be one vantage point per major metropolitan area.

The RSSAC believes that a better long-term plan for the location of the vantage points would be to distribute them by network topology instead of geographic location. RSSAC should begin investigation of implementing such a plan in the future.

3.3 Connectivity and Other Requirements

Vantage points shall be hosted inside data centers with reliable power and diverse connectivity providers.

The placement of vantage points should be based on the desire to have diverse connectivity providers. Diversity of connectivity providers helps to increase RSS coverage and avoid situations where multiple vantage points all reach the same instance.

Vantage points may be deployed on “bare metal” or virtual machines (VMs). When VMs are utilized, they should provide dedicated IP addresses and a dedicated operating system environment.

4 General Points about Metrics and Measurements

4.1 Reporting

The metrics defined in this report shall be reported by the collection system on a monthly basis.

For RSI metrics (Section 5), the collection system reports results for each metric in a given month as either "pass" or "fail." An RSI is reported to "pass" the metric when its value meets the appropriate threshold, and reported to "fail" when its value does not meet the threshold. The metric's measured value is not reported in either reading. As stated in Section 2.2, these metrics are not designed to make performance comparisons between RSIs. See Section 8.1 for an example of an RSI metrics report.

For RSS metrics (Section 6), the collection system reports the results for each metric in a given month with the measured value, as well as a pass or fail indication. See Section 8.2 for an example of an RSS metrics report.

4.2 Timestamps and Measurement Scheduling

Vantage points and the collection system shall be synchronized to Network Time Protocol (NTP).

Vantage points run all tests at five-minute intervals. At the start of each five-minute interval, the measurement software should wait for an amount of time randomly chosen between 0 and 60 seconds. Thus, measurements from all vantage points start at slightly different times, but still have enough time to complete within the five-minute interval.

Vantage points store measurements and the collection system reports metrics in Coordinated Universal Time (UTC). The collection system reports of pass or fail for thresholds are always shown for a whole month starting on the first of the month; dates for presentation always start at midnight (0 hours 0 minutes) UTC.

4.3 Elapsed Time and Timeouts

Some vantage point measurements have timeouts or are designed to measure elapsed time. This section describes how to calculate elapsed time for individual measurements over differing transports. Unless specified otherwise for individual measurements, the following rules apply:

For connectionless requests (i.e., over UDP) a timer starts immediately after the UDP message has been sent. It stops when the entire response has been received.

For connection-based requests (e.g., over TCP) a timer starts when the connection is initiated. It stops when the entire DNS response has been received (although not waiting for the TCP connection to close).

Some features such as TCP Fast Open (TFO) reduce connection setup delays. None of those features should be turned on in the measurement platform. Environments and/or operating

systems that do not allow TFO to be disabled should not be used for these measurements, if at all possible.

4.4 Connection Errors

Both connectionless and connection-based transactions may terminate in an error. Some common errors include *no route to host*, *connection refused,* and *connection reset by peer*. For the purposes described in this document, vantage points shall treat such errors as timeouts. That is, in general, timeouts (including these errors) are not retried by vantage points and are not included in collection system metrics other than availability.

4.5 Spoofing Protections

Vantage points must take reasonable steps to prevent acceptance of spoofed responses. Vantage point software must use proper source port randomization, query id randomization, optional "0x20" mixed case, and proper query and response matching. DNS Cookies may be used as a lightweight DNS transaction security mechanism that provides limited protection to DNS servers and clients.[8]

If vantage points detect malicious or spoofed traffic, such events should be recorded and logged so that manual inspection of measurements can be performed and disregarded if necessary.

4.6 Anycast

The measurements defined in this report are "instance agnostic," which means they do not target specific anycast instances within a RSI. Thus, the vantage points do not try to force queries to specific instances, rather they should let intermediate routers on the Internet determine which anycast instance of an RSI receives each query and measure the performance of the RSI as a whole.

4.7 Measurement Reuse

In some cases, queries and responses for one measurement are used in more than one metric. In specific, the collection system uses the Start of Authority (SOA) query to an instance in the availability, and response latency, and publication latency metrics.

4.8 Unexpected Results

When the collection system observes unexpected measurements or metrics, they may warrant further investigation. Examples of unexpected results may include very high response latency to some or all instances of an RSI, DNSSEC validation failures, and excessive staleness. Investigation and publication of unexpected results is most likely in the best interest of affected parties to understand the reasons for and, if possible, rectify situations that lead to such results.

To aid in debugging unexpected results, all DNS query measurements shall include the Name Server Identifier Option (NSID) option. [9] Furthermore, vantage points shall record the network

[8] See https://datatracker.ietf.org/doc/rfc7873

[9] See https://datatracker.ietf.org/doc/rfc5001

route to both IPv4 and IPv6 addresses of each RSI, in every measurement interval, using commonly available tools such as *traceroute*, *tracepath*, or *mtr*. This additional information helps diagnose issues with the monitoring system (for example, if a route local to the monitoring system disappears it will show up in traceroute). The collection and storage of the extra debugging information is not the primary purpose of the vantage point and must not cause interruption or disturbance to measurement gathering.

If, in the course of collecting and aggregating the measurements from the vantage points, one or more vantage points is clearly impacted by a software or network failure, the collection system can temporarily exclude those vantage points from the threshold calculations. Any such exclusion needs to be described publicly, and the times that the vantage points' data is excluded be clearly stated.

The collection system can also exclude some measurements from threshold calculations if the RSO can give a reasonable explanation for temporary technical problems that caused a failure to meet a threshold. Any such exclusion needs to be described publicly, and the times that the data is excluded be clearly stated.

The collection system is allowed to remove vantage points that are not acting in accordance with the goals of measurement.

RSOs can ask the collection system that anomalies be annotated with information detailing the reason for an outage, or a notice of preventative maintenance.

4.9 Determining the Number of RSIs Required for Reliable Operation of the RSS

While developing the metrics and thresholds for availability, the RSSAC found it helpful to apply a "k-out-of-n system model" to the RSS.[10] This model has, as a design parameter, the value of k representing the number of components required for the whole system to function. One example of this is a multi-engine airplane that can continue flying when one (or more) engine stops working. Another example is a communications system with multiple transmitters that would lose critical messages under average load if too many transmitters fail.

The k-out-of-n model as applied here requires some simplifying assumptions. One such assumption is that all components (e.g., RSIs) are identical. Another is that IPv4 service is identical to IPv6 service. Another assumption is that all components are independent and that the failure of one does not increase the load for any other. While none of these assumptions hold for the RSS, taking them into consideration would significantly complicate the model and the math used in predicting availability.

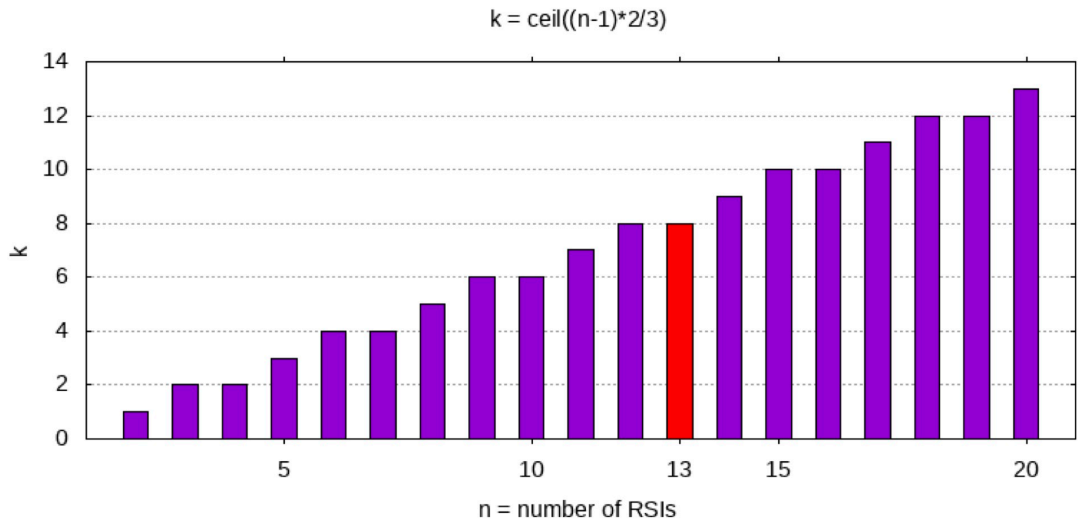
Prior to recommending availability thresholds, the RSSAC first needed to determine the value of k for a value of n. In our discussions, we settled on the following formula:

[10] See Way Kuo and Ming J. Zuo, 'The k-out-of-n system model', in *Optimal Reliability Modeling: Principles and Applications*, (Hoboken, NJ: John Wiley & Sons, 2003), 231-280.

$$k = \lceil \tfrac{2}{3}(n-1) \rceil$$

Which means that when a resolver's initial query results in a timeout, a subsequent query to a different RSI has at least a two-thirds chance of being successful.

For the current RSS, where $n = 13$, this formula results in a value of $k = 8$. The chart below shows the values of k for other values of n.



The reader might wonder why $k = 1$ is not sufficient for reliable operation of the RSS, since for any given query, a resolver needs only one response from any one of the RSIs. It is true that, from the viewpoint of the user, if the RSS returns a response, then the system is available. However, that interpretation of availability is less useful to service providers, for whom these metrics are designed. From the viewpoint of service providers, multiple RSIs must be operational at any given time in order to provide reliable service. The formula above determines the number required for reliable operation based on the total number of RSIs.

Furthermore, keep in mind that RSIs have different anycast deployment policies. The requirement of k=8 for reliable operation (of the current system) reflects the number of RSIs reachable by the vantage points, which is different than the number of anycast instances that may be operating. For a value of k=8 there are likely to be anywhere from 200 to 1000 instances in operation, based on anycast deployments known at the time of this report's publication. The different anycast deployment policies reflect the diversity among RSIs, which is an important characteristic of the RSS. Nothing in this report should be seen as an attempt to measure or quantify the way in which RSOs operate their systems with respect to anycast.

4.10 Potential Effects of Metrics on Independence and Diversity

Independence and diversity have been guiding principles for the RSOs since the group's inception. The various manifestations and benefits of RSO independence are outlined in

RSSAC042.[11] While developing the metrics and thresholds in this report, a number of participants emphasized the importance of maintaining such independence and diversity.

To some readers, the thresholds provided in this report may appear to be overly generous. The metrics and thresholds herein have been designed and chosen carefully, with the principles of independence and diversity in mind. It is not the RSSAC's intention that the implementation and publication of these metrics should encourage an RSO to forfeit its independence, or otherwise change its deployment strategies. Evidence of this occurring may require the metrics and thresholds to be revised in the future.

5 Root Server Identifier-Related Metrics

The metrics in this section apply to the individual root server identifiers (RSIs). Note that this refers to the DNS name associated with a root server operator that appears in the root zone and root hints file. For example, d.root-servers.net (or sometimes "D-Root") is the root server identifier associated with the root server managed by the University of Maryland at the time this document was published. Furthermore, note that a single identifier refers to the IPv4 and IPv6 addresses for the corresponding service.

5.1 RSI Availability

The purpose of these metrics is to characterize the availability of a single RSI over different transports and address types. The metrics are derived from a set of individual availability measurements taken from multiple locations over a period of time. The metrics have the following names:

- IPv4 UDP Availability
- IPv4 TCP Availability
- IPv6 UDP Availability
- IPv6 TCP Availability

In accordance with the recommendations in this report, it is likely that vantage points will be placed inside data centers some distance away from root server instances. The queries and responses between vantage point and instance traverse through some number of networks, routers, and switches. These intermediate network components, which are not necessarily under an RSO's control, also factor into the availability measurements. That is, the availability of an RSI at a particular point in time depends not only on the RSI itself, but on the availability of the intermediate networks as well.

Measurements. Measurements shall be made by sending DNS queries of type SOA with QNAME="." at five-minute intervals over each of the transports and address types to the root server addresses.

Measurements shall use a timeout value of four seconds.

[11] See https://www.icann.org/en/system/files/files/rssac-042-17may19-en.pdf

For a response with RCODE=0 received within the timeout value, the RSI is considered to have been available over that transport and address. After the timeout value, the query is considered to be timed out and the RSI is considered to have been unavailable over that transport and address. Timed out queries shall not be retried. Since the query should always result in an RCODE=0 response, responses with any other RCODE are considered to be equivalent to a timeout.

For every measurement, the vantage point also records the time elapsed between sending the query and receiving the response. This measurement is also used for the measurements in "5.2 RSI Response Latency" and "5.4 RSI Publication Latency."

Aggregation. All of the measurements for each transport and address type, from all vantage points, covering a period of one month are aggregated with the other measurements from the same transport and address type. Availability is calculated as the number of non-timed-out and non-error responses received divided by the number of queries sent, expressed as a percentage.

Precision. The number of aggregated measurements shall convey the metric's precision.

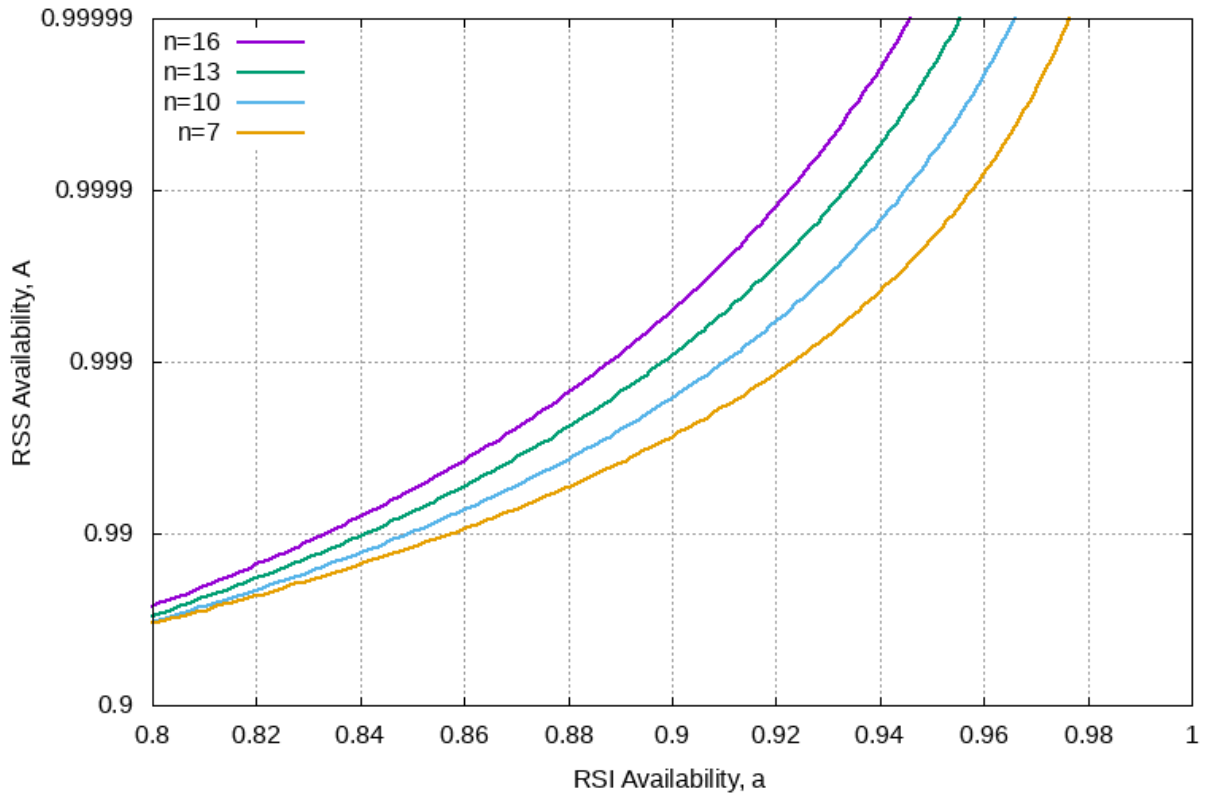
Reporting. For each month, the report shall state whether or not each of the aggregated availability metrics meets or does not meet the established threshold.

Threshold. The recommended threshold for this metric is 96%. The recommended threshold value was determined by using the formula for simple k-out-of-n parallel availability:

$$A = \sum_{i=k}^{n} \binom{n}{i} a^i (1-a)^{(n-i)}$$

Given a desired overall system availability of $A = 99.999\%$ ("five nines"), $n = 13$, and $k = 8$, this formula tells us that an individual RSI availability of $a = 96$ is necessary to meet the desired system availability.

If the number of RSIs were to change in the future, this threshold may need to be adjusted. The chart below shows the relationship between values of A, a, and n when $k = \lceil (n-1) * 2/3 \rceil$. Note that as the number of RSIs increases, the threshold for RSI availability decreases.



5.2 RSI Response Latency

The purpose of these metrics is to characterize the response latency for a single RSI over different transports and address types. The metrics are derived from a set of individual response latency measurements from multiple locations over a period of time. The metrics have the following names:

- IPv4 UDP Response Latency
- IPv4 TCP Response Latency
- IPv6 UDP Response Latency
- IPv6 TCP Response Latency

Measurements. Measurements are taken from timing of queries and responses made for Section 5.1 "RSI Availability." Timed-out queries are not utilized in this metric.

Aggregation. All of the measurements for each transport and address type, from all vantage points, covering a period of one month are aggregated with the other measurements from the same transport and address type. Response latency is calculated as the median value of the aggregated latency measurements.

Precision. The number of aggregated measurements shall convey the metric's precision.

Reporting. For each month, the report shall state whether or not each of the aggregated median response latency metrics meets or does not meet the established threshold.

Threshold. The recommended threshold for this metric is 250 milliseconds for UDP and 500 milliseconds for TCP.

The threshold for TCP is twice that for UDP due to TCP connection setup latencies.

5.3 RSI Correctness

The purpose of this metric is to characterize whether or not a single root server instance serves correct responses. Correctness is determined through exact matching against root zone data and DNSSEC validation. The metric is derived from a set of individual correctness measurements from multiple locations over a period of time. The metric has the following name:

- Correctness

The individual measurement responses will be marked either correct or incorrect. It might be difficult to determine whether an incorrect response was actually transmitted by an RSI, or due to an attacker transmitting spoofed responses. For this reason, implementations should follow the advice in Section 4.5 (“Spoofing Protections”) and Section 4.8 (“Unexpected Results”) to both minimize the chance of being affected by malicious traffic and to allow someone to investigate and disregard measurements that may be impacted by spoofing.

The RSSAC recognizes that there are limitations to this metric because the vantage points are measuring a small number of root server instances, from known source IP addresses, with no detection of on-path attackers. Another potential limitation may arise in cases when root server instances are started up with older versions of the zone saved on disk. Typically, when name servers start up they will use any saved zone data and then quickly check for a newer version from the master server. If a correctness query happens during this short window of time between startup and zone refresh, stale data may be returned. Future improvements in the measurement system may address these and other concerns.

Measurements. Measurements shall be made by sending DNS queries at five-minute intervals to the RSIs. In each interval the transport and address type used for a particular measurement shall be chosen with uniform random probability among all combinations of IPv4/IPv6 and UDP/TCP. In order to test a variety of responses, the query name and type for a particular measurement are chosen at random as described below.

For all queries, the DNSSEC OK bit is always set, and the EDNS0 buffer size is set to 1220 when using UDP.

There are two kinds of queries: *expected positive* and *expected negative*.

- The *expected positive* queries are selected from the following RRSets from a recent root zone: ./SOA, ./DNSKEY, ./NS, <any_TLD>/NS, and <any_TLD>/DS. These measurements do not query non-authoritative data directly. However, any non-

authoritative data included in the Additional section of responses will be checked for correctness.

    - At the time this document is published, the ARPA TLD is served by many of the RSIs. From those root servers, an ARPA/NS query will return authoritative data, rather than a referral, and therefore cannot be tested for correctness as described in the checking rules below. Therefore, ARPA/NS must be excluded from the set of expected positive queries above as long as any RSI is serving ARPA authoritatively. Note, however, that ARPA/DS is included because it can be tested for correctness even in this scenario.

- The *expected negative* queries have a name that contains random letters and a resource record type of A. The names are constructed as "www.rssac047-test.<RAND-NXD>", where <RAND-NXD> is formed by 10 ASCII letters chosen at random. Examples of expected negative questions are "www.rssac047-test.twxoozxmew" and "www.rssac047-test.hwzvpicwen".

When selecting a query to send for this metric, the vantage point chooses queries from the *expected positive* set with a 90% probability, and from the *expected negative* set with a 10% probability.

The rationale for the query styles is:

- Positive responses are the common case and using known authoritative Resource Record sets (RRsets) provides good coverage of the namespace.
- It is impossible to predict situations in which an RSI might provide incorrect responses. Using randomly generated TLDs - which look like typical queries - is a reasonable choice. By examining NSEC records from queries for random names we can identify cases where incorrect data may have been inserted into the root zone.

Measurements shall use a timeout value of four seconds. Responses in which the TC bit is set shall be retried over TCP transport and the timeout restarted.

Measuring Correctness. For a response received within the timeout value, the measurement records the result as either *correct* or *incorrect*.

The collection system keeps a copy of every root zone file published after it has been set up. A response is tested against all root zones that were first seen in use in the 48 hours preceding the query until a correct result is returned. If no correct result is found, an incorrect result is returned.

Correctness checking is based on the actual response data, rather than what was expected. For example, if a query was sent in the *expected positive* style, but the received response was negative (e.g., NXDOMAIN), matching is performed as a negative response. This is done to handle cases when vantage points might not receive configuration updates for a short period of time.

For all matching testing:

- All of the RRsets in the Answer, Authority, and Additional sections match RRsets found in the zone. This check does not include any OPT RRset found in the Additional section, nor does it include any RRSIG RRsets that are not named in the matching tests below.
- All RRsets that are signed have their signatures validated.

For positive responses with QNAME = <TLD> and QTYPE = NS, a correct result requires all of the following:

- The header AA bit is not set.
- The Answer section is empty.
- The Authority section contains the entire NS RRset for the query name.
- If the DS RRset for the query name exists in the zone:
    - The Authority section contains the signed DS RRset for the query name.
- If the DS RRset for the query name does not exist in the zone:
    - The Authority section contains no DS RRset.
    - The Authority section contains a signed NSEC RRset covering the query name.
- The Additional section contains at least one A or AAAA record found in the zone associated with at least one NS record found in the Authority section.

For positive responses where QNAME = <TLD> and QTYPE = DS, a correct result requires all of the following:

- The header AA bit is set.
- The Answer section contains the signed DS RRset for the query name.
- The Authority section is empty.
- The Additional section is empty.

For positive responses for QNAME = . and QTYPE = SOA, a correct result requires all of the following:

- The header AA bit is set.
- The Answer section contains the signed SOA record for the root.
- The Authority section contains the signed NS RRset for the root.

For positive responses for QNAME = . and QTYPE = NS, a correct result requires all of the following:

- The header AA bit is set.
- The Answer section contains the signed NS RRset for the root.
- The Authority section is empty.

For positive responses for QNAME = . and QTYPE = DNSKEY, a correct result requires all of the following:

- The header AA bit is set.
- The Answer section contains the signed DNSKEY RRset for the root.
- The Authority section is empty.
- The Additional section is empty.

For negative responses, a correct result requires all of the following:

- The header AA bit is set.
- The Answer section is empty.
- The Authority section contains the signed . / SOA record.
- The Authority section contains a signed NSEC record covering the query name.
- The Authority section contains a signed NSEC record with owner name "." proving no wildcard exists in the zone.
- The Additional section is empty.

Aggregation. All of the measurements covering a period of one month are aggregated together. Correctness is calculated as the number of correct responses received divided by the total number of responses received, expressed as a percentage.

Precision. The number of aggregated measurements shall convey the metric's precision.

Reporting. For each month, the report shall state whether or not the RSI's aggregated correctness meets or does not meet the established threshold.

Thresholds. The recommended threshold for this metric is 100%. The expectation is that root name servers always serve correct responses.

5.4 RSI Publication Latency

The purpose of this metric is to characterize the publication latency for a single RSI, that is, the amount of time taken to publish new versions of the root zone. The metric is derived from a set of individual measurements from multiple locations over a period of time. The metric has the following name:

- Publication Latency

The publication latency metric may also be affected by the situation described in Section 5.3, when name servers first start up with older zone data before the zone has been refreshed from the master server.

Measurements. The metrics are based on the amount of time between publication of a new root zone serial number, and the time the new serial number is observed by each vantage point over all of the transports and address types for each RSI. Rather than make additional SOA queries, this metric reuses the root zone SOA responses received from the response latency measurements from Section 5.2.

In each measurement interval, the collection system examines the response latency measurements and calculates the minimum SOA serial value over all of the transports and address types for each vantage point and RSI. This is because the RSI might return different SOA serials over UDP/TCP and IPv4/IPv6. Timed out and bogus responses must not be used in this calculation.

The collection system needs to know, approximately, when new zones are published by the root zone maintainer. This is accomplished by examining the collective SOA serial responses from all RSIs.

The collection system then calculates the amount of time elapsed until a given vantage point observes the new serial number in a response from the RSI. Note that this will always be a multiple of five minutes. For vantage points that observe the new serial number in the same interval as the root zone publication time, the publication latency shall be recorded as zero minutes.

Aggregation. All of the measurements, from all vantage points, covering a period of one month are aggregated together. Publication latency is calculated as the median value of the aggregated latency measurements.

Note that the number of aggregated measurements depends on the number of root zones published in the aggregation interval. Most commonly there are two root zones published each day, which would result in at least approximately “sixty times the number of vantage points” measurements each month for each RSI.

Precision. The number of aggregated measurements shall convey the metric’s precision.

Reporting. For each month, the report shall state whether or not each of the aggregated publication latency metrics meets or does not meet the established threshold.

Thresholds. The recommended threshold for this metric is 65 minutes. This is based on twice the value of the SOA refresh parameter (which is 30 minutes) plus one five-minute measurement interval. Note that the Root Zone Maintainer’s current distribution system sends out DNS NOTIFY messages from many different locations, to a set of addresses provided by each RSO. Even in situations where NOTIFY messages may not be reliably delivered, the RSO’s systems should be polling for zone updates at least every SOA refresh interval (30 minutes).

6 RSS Related Metrics

Whereas the metrics described in Section 5 apply to individual root server identifiers (RSIs), the metrics in this section are designed to evaluate and measure service levels for the entire root

RSSAC042.[11] While developing the metrics and thresholds in this report, a number of participants emphasized the importance of maintaining such independence and diversity.

To some readers, the thresholds provided in this report may appear to be overly generous. The metrics and thresholds herein have been designed and chosen carefully, with the principles of independence and diversity in mind. It is not the RSSAC's intention that the implementation and publication of these metrics should encourage an RSO to forfeit its independence, or otherwise change its deployment strategies. Evidence of this occurring may require the metrics and thresholds to be revised in the future.

5 Root Server Identifier-Related Metrics

The metrics in this section apply to the individual root server identifiers (RSIs). Note that this refers to the DNS name associated with a root server operator that appears in the root zone and root hints file. For example, d.root-servers.net (or sometimes "D-Root") is the root server identifier associated with the root server managed by the University of Maryland at the time this document was published. Furthermore, note that a single identifier refers to the IPv4 and IPv6 addresses for the corresponding service.

5.1 RSI Availability

The purpose of these metrics is to characterize the availability of a single RSI over different transports and address types. The metrics are derived from a set of individual availability measurements taken from multiple locations over a period of time. The metrics have the following names:

- IPv4 UDP Availability
- IPv4 TCP Availability
- IPv6 UDP Availability
- IPv6 TCP Availability

In accordance with the recommendations in this report, it is likely that vantage points will be placed inside data centers some distance away from root server instances. The queries and responses between vantage point and instance traverse through some number of networks, routers, and switches. These intermediate network components, which are not necessarily under an RSO's control, also factor into the availability measurements. That is, the availability of an RSI at a particular point in time depends not only on the RSI itself, but on the availability of the intermediate networks as well.

Measurements. Measurements shall be made by sending DNS queries of type SOA with QNAME="." at five-minute intervals over each of the transports and address types to the root server addresses.

Measurements shall use a timeout value of four seconds.

[11] See https://www.icann.org/en/system/files/files/rssac-042-17may19-en.pdf

server system (RSS). Although there is no single organization that could be held accountable for RSS performance, the RSSAC finds value in measuring and reporting on RSS service levels. Recorded metrics may be useful in understanding long-term RSS behavior.

6.1 RSS Availability

The purpose of this metric is to characterize the availability of the RSS from multiple locations over a period of time.

This metric is derived from the set of RSI availability measurements described in Section 5.1. Since the RSI availability measurements are sent over specific transports and address types, we can describe the RSS availability over those separate transports and address types. The metrics have the following names:

- IPv4 UDP Availability
- IPv4 TCP Availability
- IPv6 UDP Availability
- IPv6 TCP Availability

Aggregation. For each transport and address type, in each measurement interval t, and for each vantage point v, calculate $r_{t,v}$ as the number of RSIs that responded to an availability query (section 5.1). The aggregated RSS availability (for each transport and address type) A is then:

$$A = \frac{\sum \min(k, r_{t,v})}{\sum k}$$

Where k is the value from Section 4.9, and the sums are taken over all intervals and all vantage points.

In order for the calculated RSS availability to be anything less than 100%, there must be at least one interval in which at least one vantage point received responses from fewer than k RSIs.

Aggregation Examples. Since the calculation of this metric is more complex than others, here are some fabricated examples that demonstrate how it is calculated. These examples are based on n=13 RSIs, k=8 required for operation, a 30-day month, and 20 vantage points.

Scenarios	**Measured Availability**	**Notes**
A month-long attack takes out one RSI entirely.	100%	All measured $r_{t,v} = 12$, which is greater than $k = 8$.
A month-long attack takes out five RSIs entirely.	100%	All measured $r_{t,v} = 8$, which is equal to $k = 8$.

A month-long attack takes out six RSIs entirely.	87.50%	⅞ because all measured $r_{t,v} = 7$.
A 24-hour attack takes out all RSIs entirely.	96.66%	$\frac{29}{30}$
In one five-minute interval, one vantage point can only reach seven RSIs.	99.99992%	$\frac{(288 * 30 * 20 * 8) - 1}{288 * 30 * 20 * 8}$
For two intervals, seven vantage points can reach no RSIs.	99.9989%	$\frac{(288 * 30 * 20 * 8) - 14}{288 * 30 * 20 * 8}$

Precision. The number of aggregated measurements shall convey the precision.

Reporting. For each month, the report shall include the aggregated RSS Availability values for each transport and address type, and whether each meet, or does not meet the established threshold(s).

Thresholds. The recommended threshold for this metric is 99.999%, based on the rationale for the RSI availability threshold in Section 5.1.

6.2 RSS Response Latency

The purpose of this metric is to characterize the response latency of the RSS from multiple locations over a period of time.

Since the individual RSI response latency measurements are sent over specific transports and address types, we can also report the RSS latency over those separate transports and address types. The metrics have the following names:

- IPv4 UDP Response Latency
- IPv4 TCP Response Latency
- IPv6 UDP Response Latency
- IPv6 TCP Response Latency

Measurements. In this method, the metric is derived from the set of RSI response latency measurements described in Section 5.1.

Aggregation. In each five-minutes measurement interval, find the best k RSI response latencies for each vantage point and for each transport and address type. The aggregated response latency is calculated as the median value of the subset of lowest latencies.

Precision. Measurement Count shall be presented to convey the measurement range and precision.

Reporting. For each month, the report shall include the aggregated RSS Response Latency values for each transport and address type, and whether each meet, or does not meet the established threshold(s).

Thresholds. The recommended threshold for this metric is 150 milliseconds for UDP and 300 milliseconds for TCP.

6.3 RSS Correctness

The purpose of this metric is to characterize the correctness of the overall RSS from multiple locations over a period of time.

The metric is derived from the set of individual RSI correctness measurements described in section 5.3. The metric has the following name:

- Correctness

Aggregation. All of the measurements covering a period of one month are aggregated together. RSS Correctness is calculated as the number of correct responses observed divided by the total number of responses, expressed as a percentage.

Precision. The number of aggregated measurements shall convey the precision.

Reporting. For each month, the report shall include the aggregated RSS Correctness values, and whether it meets, or does not meet the established threshold.

Thresholds. The recommended threshold for this metric is 100%. The expectation is that the RSS always serves correct responses.

6.4 RSS Publication Latency

The purpose of this metric is to characterize the publication latency of the RSS from multiple locations over a period of time.

The metric is derived from the set of individual RSI publication latency measurements described in Section 5.4. The metric has the following name:

- Publication Latency

Aggregation. All of the measurements covering a period of one month are aggregated together. Publication Latency is calculated as the median of the aggregated values.

Precision. The number of aggregated measurements shall convey the precision.

Reporting. For each month, the report shall include the aggregated Publication Latency values, and whether it meets, or does not meet the established threshold.

Thresholds. The recommended threshold for this metric is 35 minutes. This is based on the root zone SOA retry value of 30 minutes, plus one five-minute measurement interval. Note that the RSS publication latency threshold is lower than the RSI publication latency threshold because we do not expect that a majority of RSIs to be close to the individual threshold at the same time.

7. Summary of Metrics and Thresholds

This table summarizes the recommended minimum thresholds for RSI metrics. Refer to the individual sections above for rationale and discussion regarding the particular metric thresholds.

Metrics	Name(s)	Threshold(s)
5.1 RSI Availability	IPv4 UDP Availability IPv4 TCP Availability IPv6 UDP Availability IPv6 TCP Availability	96% 96% 96% 96%
5.2 RSI Response Latency	IPv4 UDP Response Latency IPv4 TCP Response Latency IPv6 UDP Response Latency IPv6 TCP Response Latency	250 milliseconds 500 milliseconds 250 milliseconds 500 milliseconds
5.3 RSI Correctness	Correctness	100%
5.4 RSI Publication Latency	Publication Latency	65 minutes

This table summarizes the recommended minimum thresholds for RSS metrics. Refer to the individual sections above for rationale and discussion.

Metrics	Name(s)	Threshold(s)
6.1 RSS Availability	IPv4 UDP Availability IPv4 TCP Availability IPv6 UDP Availability IPv6 TCP Availability	99.999% 99.999% 99.999% 99.999%
6.2 RSS Response Latency	IPv4 UDP Response Latency IPv4 TCP Response Latency IPv6 UDP Response Latency IPv6 TCP Response Latency	150 milliseconds 300 milliseconds 150 milliseconds 300 milliseconds
6.3 RSS Correctness	Correctness	100%
6.4 RSS Publication Latency	Publication Latency	35 minutes

8 Recommendations

Recommendation 1: The RSSAC recommends the ICANN Board commission an initial implementation of the measurement system described in this document to gather operational data and experience from actual monitoring of the RSS. The initial implementation should be designed such that it can transform into the official implementation as described in Recommendation 2 below. The insights learned from the implementation will inform future revisions of this document, if necessary.

Recommendation 2: The RSSAC recommends that the official implementation of the metric system must:

a. Meet the minimum requirements specified in Section 3 of this report regarding the number, location, connectivity, and other requirements for the vantage points.

b. Publish all software related to its operation under an open source license as defined by the Open Source Initiative.[12]

c. Make the raw measurement data available to anyone in the interest of transparency. A third party should be able to use the raw data to verify the computation of these metrics.

d. In its monthly reports, only publish threshold pass or fail indicators for each RSI, not the actual measurements or metrics used to determine the threshold pass or fail values.

e. Publicly describe its methods for collecting measurements and aggregating metrics, including the topological location of each measurement vantage point. This description should be complete enough for RSOs and DNS researchers to create their own measurement collection systems similar to those used by the official implementation.

f. Share with an RSO the underlying measurements and metrics that resulted in failure any time an RSI fails to pass a threshold test. The shared measurements and metrics must include all measurements from around the time of failure and must include all measured values for all transports and address types.

Recommendation 3: The RSSAC, in collaboration with ICANN and the Internet community, should consider the following additional work:

- For a holistic view of RSS performance, it may be desirable or necessary to include measurements for all instances of each RSI. The only reasonable way to provide for such a view would be through self-reporting. In the future, it should be considered to have each RSO perform self-reporting of the defined metrics to eliminate uncertainty of components not under the RSO's control, and it should probably be tied to an SLA including compensation for the RSO to implement.

- Create a reference data set.

[12] See https://opensource.org/osd-annotated

- Explore the financial aspects of increased accountability and how it might relate to these metrics.
- Keeping with the provisions of RSSAC037 and RSSAC038 publish a document that advises any bodies created as part of the ongoing evolution of RSS governance on how they should interpret and act on data from the measurement systems.
- Investigate a better long-term plan for the location of the vantage points. Such a plan would distribute the vantage points by network topology instead of geographic location.
- Whereas the current work is based on a largely empirical model of the RSS, future versions of this document may want to take a more analytical and theoretical modeling approach.

9 Example Results

9.1 Example RSI Results

Metric: RSI Availability (percentage)
Precision: Number of measurements (max 172800[13])
Thresholds: 96%

Metric: RSI Availability		Month: 2019-09	
RSI	Transport	Performance	# Measurements
N-Root	IPv4 UDP	>= 96%	172792
	IPv4 TCP	>= 96%	172783
	IPv6 UDP	< 96%	172787
	IPv6 TCP	>= 96%	172797

Metric: RSI Median Response Latency (milliseconds)
Precision: Number of measurements (max 172800)
Thresholds: 250 ms for UDP, 500 ms for TCP

Metric: RSI Median Response Latency	Month: 2019-09

[13] Assuming 20 vantage points and a 30-day month

RSI	Transport	Performance	# Measurements
O-Root	IPv4 UDP	<= 250 ms	172794
	IPv4 TCP	> 500 ms	172799
	IPv6 UDP	<= 250 ms	172795
	IPv6 TCP	<= 250 ms	172789

Metric: RSI Correctness
Precision: Number of measurements (max 172800)
Thresholds: 100%

Metric: RSI Correctness by Matching	Month: 2019-09	
RSI	Performance	# Measurements
R-Root	< 100%	172785
S-Root	>= 100%	172789

Metric: RSI Publication Latency
Precision: Number of measurements (max varies)
Thresholds: 65 min

Metric: RSI Publication Latency	Month: 2019-09	
RSI	Performance	# Measurements
T-Root	<= 65 min	1199
U-Root	> 65 min	1196

9.2 Example RSS Results

Metric: RSS Availability (percentage)
Precision: Number of measurements (max 2246400)
Thresholds: 99.999%

Metric: RSS Availability		Month: 2019-09	
	Transport	Performance	# Measurements
RSS	IPv4 UDP	100%	2246295
	IPv4 TCP	99.99999%	2246134
	IPv6 UDP	98.50%	2246293
	IPv6 TCP	100%	2246325

Metric: RSS Median Response Latency (milliseconds)
Precision: Number of measurements (max 1382400[14])
Thresholds: 150 ms for UDP, 300 ms for TCP

Metric: RSS Median Response Latency		Month: 2019-09	
	Transport	Performance	# Measurements
RSS	IPv4 UDP	31 ms	1382240
	IPv4 TCP	325 ms	1382393
	IPv6 UDP	44 ms	1382385
	IPv6 TCP	87 ms	1382376

Metric: RSS Correctness
Precision: Number of measurements (max 2246400)
Thresholds: 100%

Metric: RSS Correctness	Month: 2019-09	
	Performance	# Measurements
RSS	100%	2246125

[14] Assuming k=8 RSIs

Metric: RSS Publication Latency
Precision: Number of measurements (max varies)
Thresholds: 35 min

Metric: RSS Publication Latency	Month: 2019-09	
	Performance	# Measurements
RSS	5 min	15589

10 Acknowledgments, Dissents, and Withdrawals

10.1 Acknowledgements

The RSSAC would like to thank the following RSSAC Caucus members for their time, contributions, and review in producing this publication.

RSSAC Caucus Members
Russ Mundy (co-chair)
Duane Wessels (co-chair)
Afifa Abbas
Alejandro Acosta
Jaap Akkerhuis
Fred Baker
Ray Bellis
Ramanou Biaou
Harish Chowdhary
Kazunori Fujiwara
Paul Hoffman
Kevin L. Jones
Akira Kato
Ihtisham Khalid
Warren Kumari
Matt Larson
Daniel Migault
Abdulkarim Oloyede
Amir Qayyum
Rao Naveed Bin Rais
Anand Raje
Ken Renard
Shinta Sato
Ryan Stephenson

Robert Story
Kevin Wright
Zhiwei Yan
Dessalegn Mequanint Yehuala

ICANN Support Staff
Andrew McConachie
Rachel McFadyen
Carlos Reyes
Danielle Rutherford
Kathy Schnitt
Ozan Sahin
Steve Sheng (editor)

10.2 Statements of Interests

RSSAC Caucus member Statements of Interest are available at:
https://community.icann.org/display/RSI/RSSAC+Caucus+Statements+of+Interest

10.3 Dissents and Withdrawals

There were no dissents or withdrawals.

11 Revision History

11.1 Version 1

Current version.