Statistics About DNS Root Name Service:

Some Recommendations

James Wilson, Daniel Karrenberg & Terry Manderson

Version 1.1 – 18 September 2017

---

1. Introduction

The authors have been subjected to questions about various results concerning the performance of the DNS root name service as a whole, and several DNS root name servers in particular. This memo outlines some principles for good methodology in this field, as well as issues that might arise. Use this as a guide when conducting research in the field of DNS service performance. While some of these principles apply specifically to DNS root name service, others may be applicable to statistics about other parts of DNS name resolution.[1] All terminology used in this report is referenced in RSSAC Lexicon[2] and RFC7791.[3]

---

[1] Root Server Technical Operations Assn. (n.d.). http://root-servers.org/

[2] Root Server System Advisory Committee 026. (2017, March 14). RSSAC Lexicon. https://www.icann.org/groups/rssac/documents-number-version

[3] RFC7791, DNS Terminology. https://www.rfc-editor.org/info/rfc7719

2. How do Results Apply to DNS Name Resolution as a Whole

DNS name resolution is made up of many operational parts and each part can independently impact name resolution. We strongly advise the reader make themselves aware of at least the following aspects before drawing conclusions from this or other published works:

- The typical caching behaviors of recursive DNS resolvers and stub resolvers
- The fallback behaviors and timeout values of DNS clients and stub resolvers as deployed in different operating systems
- The Time-To-Live(TTL) values of DNS zones, and in particular the Root Zone
- Differences and similarities of Name Server selection algorithms in recursive resolvers
- Emerging standards that impact the common behaviors such as DNS over TLS, Localized Root Zone copies in resolvers, and aggressive use of NSEC/NSEC3

Example: The TTL of delegation information for TLDs in the root zone is typically 48 hours. Each resolver may cache this information for two days. Consequently, the DNS performance for clients of this resolver that query for names in a specific TLD is typically influenced only once every 48 hours by the performance of the root name servers.

Therefore, care must be taken when interpreting data sets related to the DNS name resolution. For instance, a disproportionate level of focus is often given to the latency of either a DNS root zone query or a network probe to one or more root servers. We posit that latency of such a DNS query is just one minor attribute of the root server system and is prone to misinterpretation unless significant effort is given to the broader correlation of the results in several other dimensions. Such dimensions would include many of the facets listed above. Without this correlation, latency alone can be used as a measure of topological proximity to the point of measurement and should not be used as a performance indicator for the DNS nor the root server system. With this in mind the statistics in this document are provided with the intention to establish a framework for education in the analysis of DNS resolution.

3. Our Study

In this paper, the authors have highlighted issues faced when generating and using DNS measurements for research. We have also included recommendations to these issues in the hope of alleviating problems that could arise if these issues are not addressed.

To better explain our proposed recommendations, we conducted a test study using RIPE Atlas probe measurements. [4] The data used is the full probe measurements from May 15th, 2017 to June 14th, 2017. The measurements contain multiple variables, including latency (RTT), country code, version type, and time stamps. There are over 50 million measurements included in the data set, drawn from a large distribution of 10,610 unique probes across the world.

All results and examples used in this paper are made from this data. It should be noted that this is a sample of probes in a limited time frame. The produced results are not conclusive. No root server should be considered inherently worse than any other. Additionally, the inferences made are not limited to just probe data. Statistics and recommendations illustrated in this paper can be used in other fields of DNS research.

The term "fastest" in relation to root server latency

Throughout this paper, we will discuss the response times, or latencies, of root servers. We use the term "fastest" as a way of describing a root server with the lowest average response time. However, we have concluded that there is no such thing as a "fastest" root server. The response times of root servers are highly variable due to organic changes of the global routing table, and the description of "fastest" is misleading.

---

4. Issues and Recommendations

4.1 Cleaning Data

Issue

DNS name resolution data is typically well maintained and organized. Parameters can be set to refine queries to capture specific measurements. However, a lurking issue is that the data itself can carry natural errors. These can be misinterpreted, and should be subsetted from the data where possible.

[4] Jacobsen, Ole J. (2015, September). Internet Protocol Journal: RIPE Atlas. http://ipj.dreamhosters.com/wp-content/uploads/2015/10/ipj18.3.pdf

Example - Local Errors

DNS root name service measurements are recorded by probes that measure a connection to all thirteen root service addresses roughly every thirty minutes. When observing these measurements, sometimes there is no response, thus the measurement is recorded as "non-responsive". This can be due to non-responsiveness of the root server, or a network failure somewhere in-between the probe and root server instance. However, if a probe cannot obtain a response from *any* of the thirteen roots in its half hour time span, then we assume this is a local connection issue for the individual measurement probe, not for each of the thirteen roots. These queries should be ruled out, as they skew the perceived accessibility of a root server. This can reduce non-response measurements by over 40%,[5] providing a clearer picture of what accessibility issues with a root server may actually exist.

Recommendation

Remove local errors when dealing with root accessibility metrics and ratios. Additionally, always consider possible errors that the data may present related to your specific research questions. These can usually be identified with a combination of descriptive statistics and knowledge of the variable subject matter.

4.2 Sampling

Issue

Sampling is typically necessary when researching DNS name resolution. However, sampling can lead to bias in results if not done properly.

Example 1 - Vantage Points

At the time of writing, DNS root name service is provided from more than 750 different locations in the Internet topology. Measurements from any one vantage point only have very local meaning. In order to make claims about the service as a whole, a sufficient number of well distributed vantage points is required. This is important on two levels. If the sample is not varied enough among a wide area, then sampling biases can inherently distort the data to produce different results. If the sample is not large enough, then even one outlier in the sample can cause inaccurate measurements.

Another consideration is how representative the distribution of vantage points is of the distribution of widely used caching resolvers. This is only applicable when looking below the root server level at TLD's. For example, if probing through a caching resolver, the

[5] This statistic was generated from our study after removing local errors from our original data set. The total number of non-responses dropped by roughly 42%, and the code to generate this result is found in Appendix A.

response time for the initial query will be longer. However, if just sampling the root server instances, the caching resolver does not influence results.

Recommendation

We strongly recommend including information about the number and distribution of the vantage points used in a study, so as to verify that the results are accurately descriptive of the area of study.

Example 2 - Resolvers

Local caching resolvers store information of queries for a DNS label for a given TTL for that specific response. This cached information can influence how fast results are captured and returned. For example, the first query for a given search may have a response time of 75ms. If that same query is made again within the period of time in which the resolver is still caching the response, then its response time may only be 30ms. This change in latency can distort samples, and should be addressed when researching any changes of latency on a network.

In anycast systems, a given response can have many different resolving paths. This constant change leads to a significant increase in variance. The different instances can skew network latency and accessibility.

Recommendation

The nuances of caching resolvers and anycast systems requires that a large sample size of query results be used to get an accurate picture of the network in its entirety. The variation of paths in anycast systems can be seen by anyone by running the following code in your console:

```
dig @d.root-servers.net hostname.bind txt ch +short
```

- You can change the root you query by changing the first letter after the @ sign to your choice of root server [A-M]

This shows the many different instances related to a specific root server from the user's computer. For example, when run from our location we received "laca2.droot" as a response. The user will likely get a different response from ours.

Example 3 – Inhomogeneous Vantage Points

Unless all vantage points in a study are absolutely equal in their defined characteristics, it is important to take variance into account. Possible ways of dealing with inhomogeneous vantage points is to increase their number sufficiently to make the differences insignificant, or to take the differences into account when deriving statistics from the measurements. [6]

For instance, RIPE Atlas probes contain tags that identify their built-in features, such as version type. This allows for many different sampling options. Too much of a certain option can distort one's data. There are two major examples of this:

One is sampling the different classifications of probes, such as Version 1, Version 2, Version 3, and anchors. These different classifications can often times produce different results, showing inherent differences for the different classifications.

- Figure 1 illustrates how different subsets of probe classifications for a country can change the outcome for what root server is, on average, "fastest". Anchors consistently have a different root they pick up as having the "fastest" latency compared to the entire data for a country. These variations can lead to varying, un-grounded conclusions. Note well that this way of presentation also suffers from the "Top-N" issue described below.

Figure 1: Fastest root server amongst different classifications

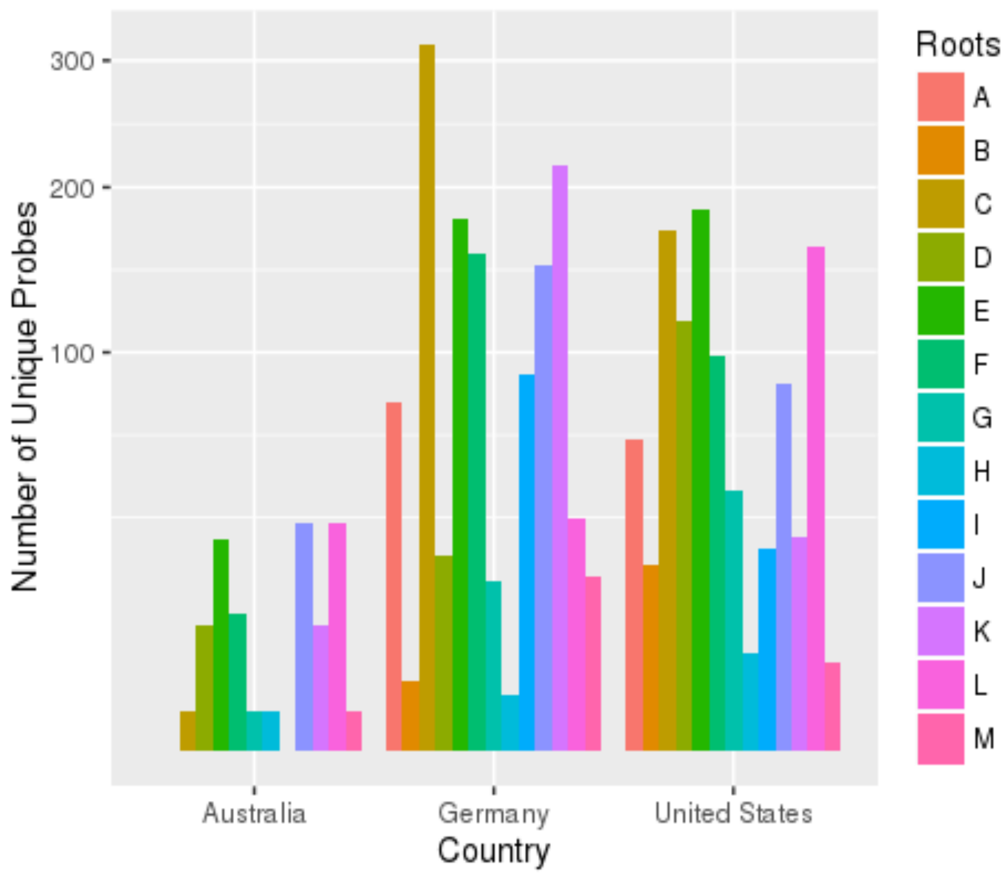
Country	Entire Data	Anchors Subset	V1/V2 Subset	No V3 Probes
Argentina	L	K	F	J
Canada	D	E	F	D
Romania	I	K	F	F
Serbia	I	J	L	L
Singapore	D	E	E	I
Slovenia	F	J	E	E
United States	D	E	F	F
South Africa	J	E	F	L

[6] Thomas Holterbach, Cristel Pelsser, Randy Bush, Laurent Vanbever: Quantifying Interference between Measurements on the RIPE Atlas Platform, Internet Measurement Conference 2015

Another sampling method involves taking all unique probes for a region, and identifying the overall average measurements of each unique probe. This sort of methodology is important, as it often shows many different results for a probes in a region than when all of the unique probes are summed into one total sample.

- In Figure 2, we take each unique probe for a given country, and determine its "fastest" root server. We sum the count of unique probes and form the following segmented barplots. From these, we can see that a given country typically has several root servers that are similarly consistent in having the "fastest" latency. Any given grouping of probes could indicate a certain root server to be the "fastest" for a region. Thus, varied sampling of unique probes is important to eliminate such errors.

Figure 2: Fastest roots for unique probes in a given country.



Recommendation

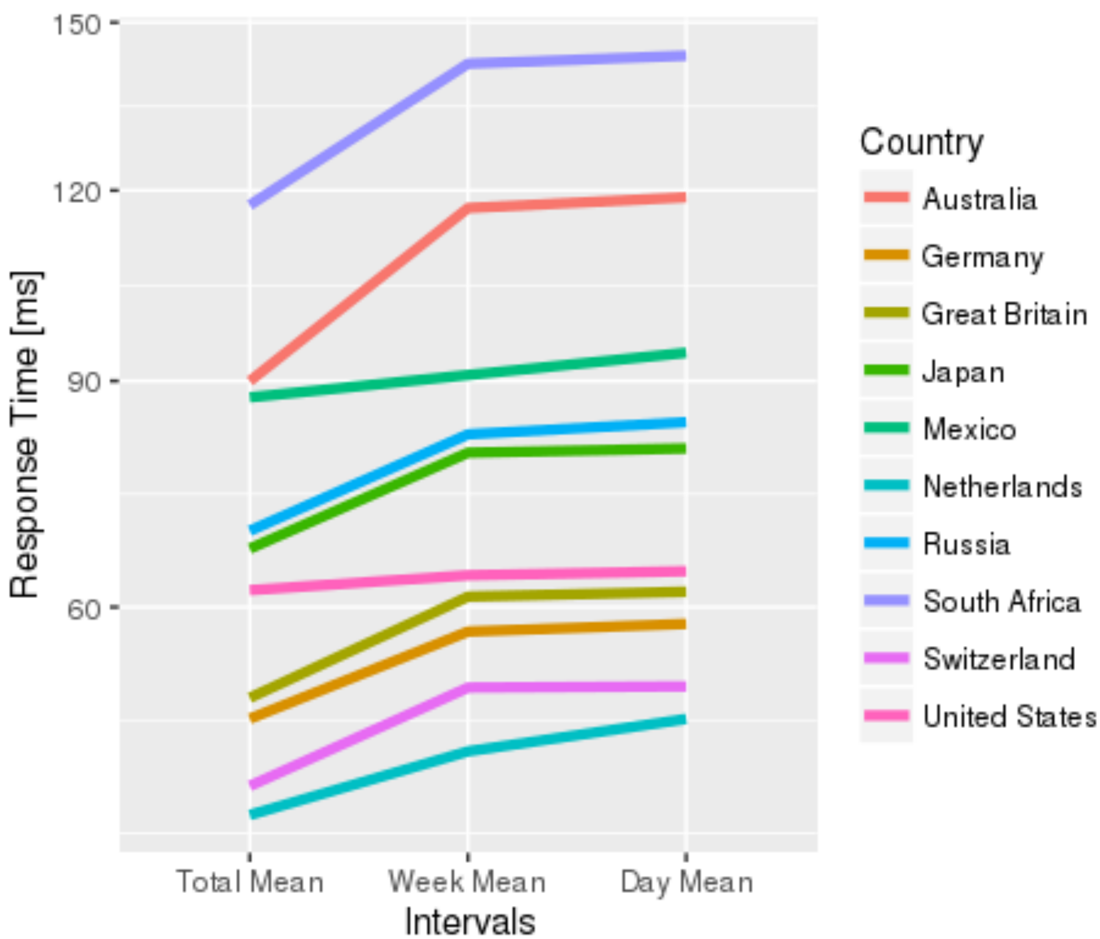
These different measurements demonstrate that too much of one probe can skew a study's results. It is recommended that a large sample of unique probes of different classifications should be used in a study. This will lower the chance for inherent biases to arise. Alternatively, the measurement results could be corrected by calibration.

For more on the levels of variation between probes, review the "Top-N" issue in the statistics section below.

Example 4 - Time Intervals

Collecting measurements over a large time interval is essential to having accurate results. As an example, we sampled probe measurements over the span of an entire month. The results found from that months' worth of data can differ from smaller time intervals, like weeks and days, within that month. There is also a noticeable uptick in average response time, showing how results can grow more and more skewed the smaller the sample time frame is. These results are recorded and visualized in Figure 3.

Figure 3: Mean Country Response time (ms) by Month, Week, and Day subsets



Country	Total Mean	Week Mean	Day Mean
Australia	89.976	117.036	118.843
Switzerland	40.566	50.782	50.913
Germany	47.426	57.114	58.012
Great Britain	49.671	61.255	61.850
Japan	67.175	79.826	80.415
Mexico	87.633	90.877	94.172
Netherlands	37.774	44.028	47.405
Russia	69.452	82.373	84.087
United States	62.040	63.855	64.354
South Africa	117.532	142.317	143.788

Recommendation

We encourage the sampling of large time intervals to create a more coherent sample. Also, the time frame sampled should be noted in your study to allows others to simulate results and draw inferences of their own.

4.3 Geolocation

Issue

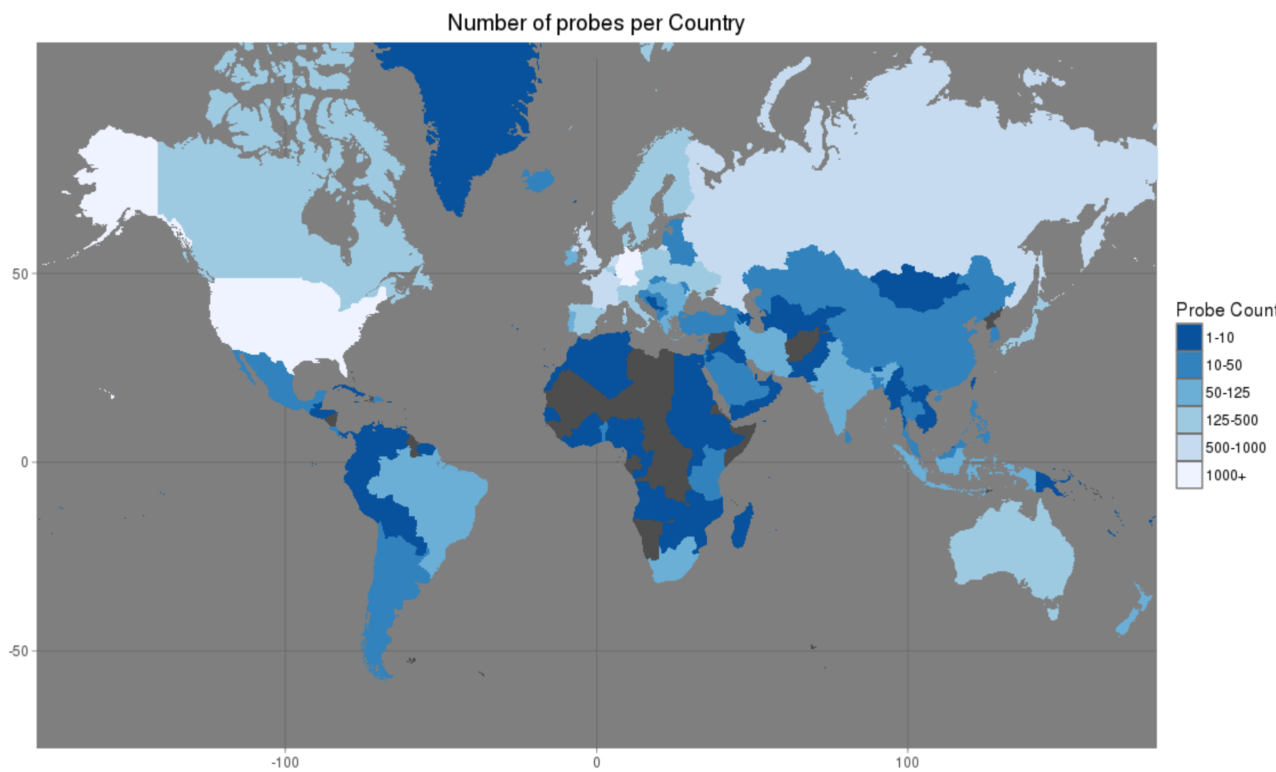
Before categorizing results by geographic location it is essential to consider the sources and quality of geolocation information. This is especially true for free services that provide a location based on an IP address.

For instance, RIPE Atlas probes are typically recorded with a country tag, longitude, and latitude variables that make geolocation sampling simple. However this is not 100% accurate and some plausibility checks should be applied, especially for areas with a small number of probes. **The difference between countries and Internet topology is important.** Two nodes in the same building can have totally different views of the internet, based on having different upstream providers. Even more, differing probe counts between regions require that appropriate statistical techniques be used.

Example - Country measures

Roughly 180 countries contain RIPE Atlas probes. This allows for global Internet research to be conducted using the entire range of values available. However, some of these countries like Botswana, Kuwait, and Peru only have 1-10 unique probes. This is a small sample size, and one probe going offline can result in unrepresentative statistics. We encourage that studies for such regions hold their findings against these small sample sizes so as to paint an accurate, honest picture of the region as a whole.

Figure 4: Number of unique probes for each country



From this map, we can see that probes are not evenly distributed among countries. For instance, most of Africa often has less then 10 probes per country. This is likely due to a lack of Internet infrastructure and probe volunteers.

Recommendation

Individual countries without a large sample of probes should have their results examined in a case by case method to ensure results are relevant and significant. More advanced statistical methods would be necessary to extrapolate meaningful data from these specific regions.

4.4 Statistics

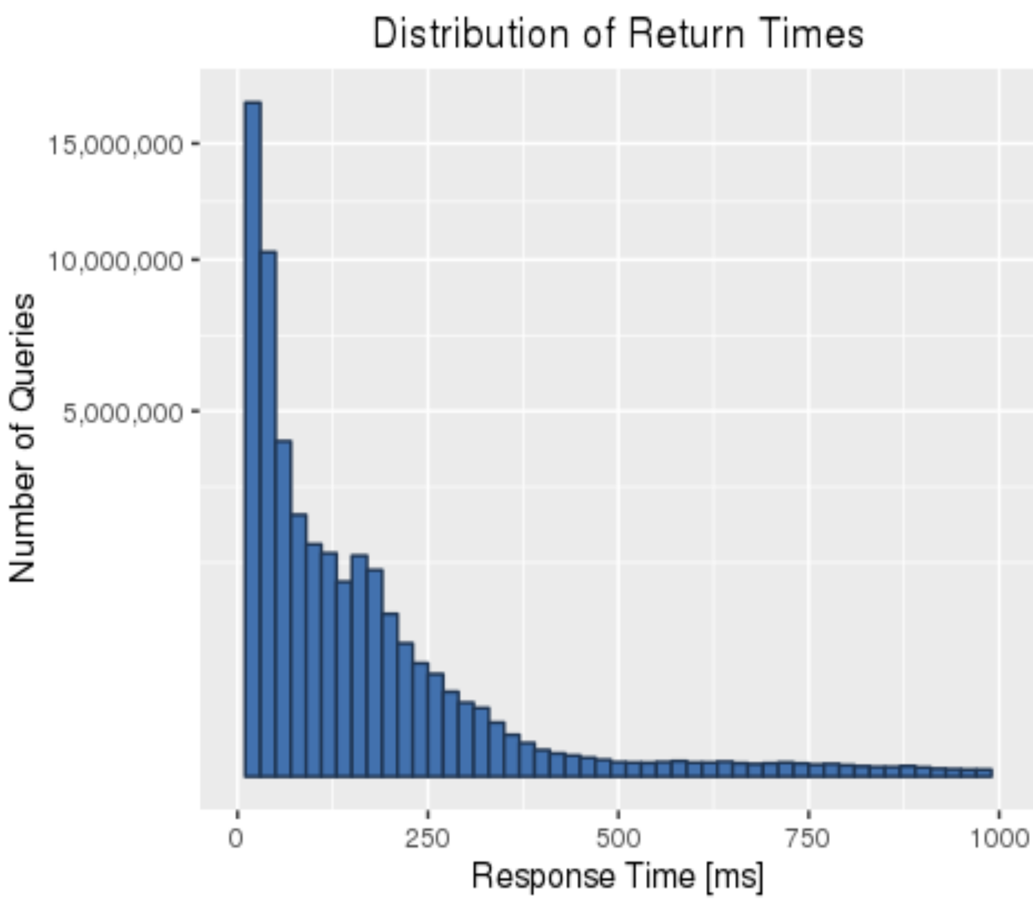
Issue

Using descriptive statistics properly requires a full understanding of the distributions and qualities of the data being described. The improper use of a metric can mislead about the reality of the situation. In a heavily anycasted environment, results such as return time tend to be very dependent on the distribution of vantage points and values tend to be grouped.

Example 1 - Mean vs Median

The response times of any given probe have a tendency to include several large outliers, skewing the distribution to the right. The majority of measurements are relatively "fast", with a few taking longer due to extraneous reasons. Thus, median is the best measure for measuring the average response time.

Figure 5: Distribution of return times (latency) for all recorded responses



Recommendation

Use median as a measure of average wherever the distribution of values is skewed, such as in the case of response time. For cases where you care about the skew of the data, it is more robust to use, for example, the 95% quartile values to capture a skewed value than mean, as you can eliminate outliers and still represent the data's skew.

Example 2 - Top-N

Reporting only one or a few extreme values can be misleading, because it does not say anything about the distribution of the values being described. This issue is highlighted in the case of root server response times, which typically vary only slightly between roots.

In our team's study, the distribution of response times from the 1,405 RIPE Atlas vantage points in Germany is seen in Figure 6.

Figure 6: Distribution of response times of root servers for probes in Germany

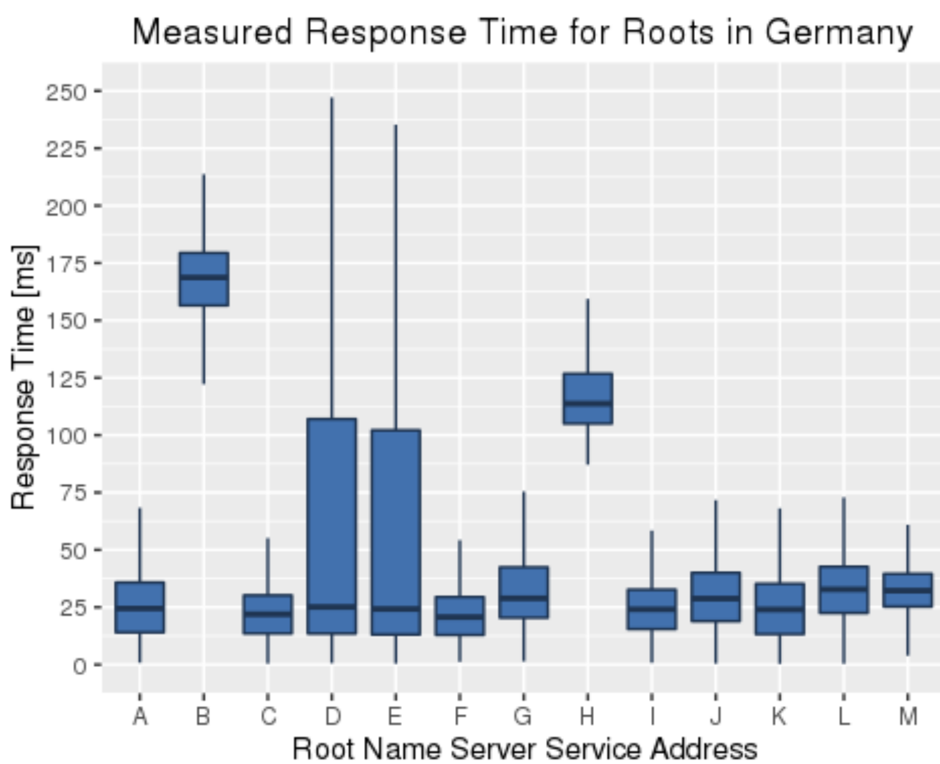


Figure 6 shows that there are nine root servers that provide very similar service both in terms of absolute response time and consistency over time and vantage points.

Just reporting the mean lowest return time in this case does not only omit a lot of useful information, but also give false support to the deemed "fastest" root. The results are close enough that the measurement set-up itself, such as the distribution of the vantage points and the choice of the time period, will very likely determine the outcome if one just considers the lowest mean response time.

Recommendation:

We suggest to avoid reporting extremes and particular 'top-n' lists.

Groupings should only be used if the distribution of the results itself suggests them. In the example above, one could create groupings as follows:

A,C,F,G,I,J,K,L,M:

Consistent response times for all German vantage points with means not exceeding 30ms and maxima not exceeding 75ms.

B, H:

Consistent response times with means exceeding 100ms because neither has an anycast site in Europe

D, E:

Response times with means similar to the first group, but with less consistency and maxima exceeding 200ms

- The **D, E** group could be further investigated. Because of the long-time period used, one could investigate whether the wide distribution of results is consistent throughout the time domain, as well as the various system properties of the involved vantage points. Furthermore, outliers in the distribution could be the result of misrouting, where some ISPs go to a nearby anycast site while others go to a distant site.[7]

4.5 Reproducible Research

Issue

For research to be relevant for the DNS community, it needs to be clear and detailed in its methodology so that it can be feasibly reproducible with the given information. If not clear and detailed, research looses its credibility and the accessibility needed to prove results and allow for further reproductions of the research.

[7] Bellis, R. (2015, October 14). Researching F-root Anycast Placement Using RIPE Atlas. https://labs.ripe.net/Members/ray_bellis/researching-f-root-anycast-placement-using-ripe-atlas

Example – Theoretical Study

Let’s propose a theoretical research paper and look at what aspects of the study help (and don't help) to make it clear and reproducible. The following table outlines some characteristics of the study.

Pros	Cons
Time period of study published	No raw data sets published
Number of countries/ vantage points published	Few detailed distributions of variables
Descriptive statistics published	No research code made available

The paper is clear on its methodology, showing the parameters and statistics used to make its conclusions. What it lacks is details. The data used is described, but not given. The code to generate the results is not made public. Only enough information for the research's conclusions is given, without including important details like the distributions of variables like latency. These details are essential to backing up one's claims, and to erase doubt in the reader's mind.

Recommendation

We strongly suggest publishing source data sets and analysis code. In cases where that is not feasible, please publish as many details about the methodology as possible. Open data projects allow multiple parties to validate results and run tests of their own, so knowing the process involved in one study will make future studies faster and more detailed.

For example, attached you will find the data and code used in this paper, as well as the sources it was drawn from.

---

5. Data & Code

RIPE Atlas Measurement numbers and time frames used in examples

For the examples we used RIPE Atlas data between 2017-05-15 00:00:00 UTC and 2017-06-14 23:59:59 UTC. The measurement IDs are: 10001 10004 10005 10006 10008 10009 10010 10011 10012 10013 10014 10015 10016. These measurements query the 13 IPv4 service addresses for the SOA RR of ‘.’ every 1800 seconds from all RIPE Atlas vantage points. This yields 187,530,814 results for the period.

Probe description data was also used to add further measurement details for analysis. All data will be available via the RIPE Atlas APIs for the foreseeable future.

RIPE Atlas API and Tools

Documentation about the RIPE Atlas API and related tools can be found at https://atlas.ripe.net/docs/.

Specific Code

The specific code for the examples in this report is available at http://labske.org/rootstats. This includes code to retrieve the data from RIPE Atlas, to analyze it and to produce the graphs. For added convenience the extracted raw data files are also available from here.

Running the analysis code with the entire data set can take a considerable time, depending on your version of R and strength of your computer. A random sample of the data (perhaps a million queries) will run considerably faster and produce similar results.

6. Acknowledgments

The authors gratefully acknowledge constructive comments from Fred Baker, John Heideman, Akira Kato, Lars-Johan Liman, Florian Obser, Robert Story, Colin Petrie and Duane Wessels as well as support from John Crain. Our errors and omissions remain our own responsibility.