Root Scaling Study

Terms of Reference

05/05/2009

INTRODUCTION

With resolution 2009-02-03-04, the ICANN Board asked the Root Server System Advisory Committee (RSSAC), the Security and Stability Advisory Committee (SSAC), and the ICANN staff to study the potential impact on the root zone stability that might arise when IPv6 address records, IDN top level names, other new TLDs, and new records to support DNS security are added to the root zone.

The Board has expressed interest in hearing of the impact of the distinct changes, but also their aggregate effect on root zone operations. The Board also asks that the study address the technical and operational concerns regarding expanding the DNS root zone that have been expressed on this topic.

In response to the Board's request, the three groups formed a steering group and are organizing a focused study. This memo sets the terms of reference for the study.

GOALS AND CONSIDERATIONS

- The data in/for the root zone should move through the provisioning and publication process with zero errors.
- The root zone provisioning and publication subsystems should be robust and resilient to attack/corruption/delay.
- The root zone system should be nimble/adaptable to changes in technology and policy.

With these goals in mind, this study will explore the ramifications on scaling up the root zone system on both size and rate of change.

We believe that as choices are made to accommodate more entries and/or faster rates of change, the system will become more difficult to manage and its ability to be robust, resilient, and adaptable may suffer. It is likely that some future review of this work will have to revisit whether the processes are matched and adaptable to the changed circumstances.

SCOPE

The scope of this study includes all aspects of root zone operations. Root zone operations is understood to mean all aspects of root zone data production, compilation, publication to the root servers, including anycast instances, and serving data from the root servers.

Aspects that are specifically within the scope of this study are:

- Addition of IPv6 to glue records
- DNSSEC signing the root zone
- Addition of DS Resource Records to the root zone
- Addition of IDN TLDs
- Addition of new TLDs at an accelerated rate and
- The impact of accumulated growth of the root zone

Although changes to the root zone operations may affect user systems such as browsers or may affect local environments, those impacts are outside the scope of this study. They may well merit study in a separate effort.

Other worthy study efforts may include how browser software must be modified to enable DNSSEC, how Whois should be modified to account for IDN data beyond the maintenance of Whois for the root zone, and the economic impact of implementing these changes beyond the impact to the organizations involved in management of the root zone operations. While all of these are important issues to be addressed in discussions of whether, how, and when to implement changes to the top-level domain structure, they do not specifically impinge on root zone operations and are therefore outside the bounds of this root zone stability study.

While changes to contact information for maintaining elements of the root zone are not within the scope of this study, there is a potential dependency insofar as issues may manifest as a result of requested changes to shared resources, shared validation methodology, and shared update methodology.

The limitations above notwithstanding, one key element of this study is to gather and catalog the many questions that get asked about the scaling up of the root, even if the answering of some of those questions is beyond the scope of this study.

MULTIPLE PERSPECTIVES

In analyzing the effect of increase the size and complexity of the root, it is useful to divide the study along the following dimensions.

1. Impact on the **robustness** of the DNS hierarchy

The present DNS root zone is robust in several senses. First because the root zone system is more well-provisioned than many other parts of the hierarchy. Second because the present root zone is small enough for most of the contents to be cached in the recursive servers most of the time and third because there is no perceived gain from attacking the root server system. For other parts of the hierarchy, in particular large zones, the situation is considerably different.

It is therefore conceivable that an expanded root zone, in particular a greatly expanded root zone, will change its ”robustness profile”. Hence it is important to consider the robustness implications of expansion. Examples of issues to consider are:

- the DNS caching design being partially defeated causing an overall growth in query volume and sensitivity to network outages
- the possibility of increased vulnerability to DDOS attacks
- the impact of the assumed higher frequency of change on the zone propagation properties
- the ability of root server operators to continue to improve access to the root zone in remote corners of the Internet as a function of scale

2. **Qualitative** vs **quantitative** effects

While many of the questions will revolve around the quantitative effects of increasing the size of the root zone and increases in the processes related to the root zone, some questions are related to the complexity of adding DNSSEC, IPv6, and/or IDNs as part of the scaling process. An example is given in the appendix.

3. Impact on the **provisioning system** vs impact on the **lookup system** vs other aspects

The provisioning system is the collection of activities that put information into the root servers. This covers the entire process of adding entries to the root zone or making changes. Organizationally, the TLD operators, IANA, NTIA, VeriSign and the Root Server operators are all involved.

The lookup system is the interaction among the root servers, the caching resolvers and the end systems. This is the core function of the domain name system.

4. Impacts at **different ranges of scale**

The root currently has fewer than 300 top level domains, and there have been very few new top level domains added each year. It is unclear how large the root might grow, and also unclear what the rate of change will be. As a general principle, unbounded growth and/or an unbounded increase in the rate of growth cannot be sustained indefinitely without radical change to the system as it exists

today. In order to provide a common basis for discussion and analysis, we have chosen broad ranges of root size, as measured by the number of names in the zone and/or the increase in the number of changes to the root zone in a given period of time.

- "Plus 0" - baseline for all statistics and other quantitative measures
- "Plus 1" – 300 to 10,000 names and/or changes at up to ten times the current rate.
- "Plus 2" – 10,000 to 300,000 names and/or changes at up to hundred times the current rate.
- "Plus 4" – 300,000 names and above and/or corresponding change rates (``unbounded growth'').

Changes to the root zone data (either NS records or Whois data) currently occur on an annualized basis of ~330 changes, or approximately one per standard work day (though that is not how frequently they appear in the zone), or 1.2 per TLD. To what extent might this rate change, either because the size of the root increases and thus the rate of change would increase proportionately, or because there is absolute change in the dynamics? It's expected there will be initial interest for forty new IDN ccTLDs and perhaps 500 new gTLDs at the outset. These numbers should be taken as lower bounds for purposes of analysis, not as upper bounds. See below for further discussion.

The root zone is currently published on a twice-daily basis with changes batched into either or both of the two published versions. While there is no current plan to change the twice-daily cycle, one question to consider in this study is whether it might be necessary to change this if the root zone grows sufficiently large or the rate of change becomes sufficiently great.

The analysis of the impact of rates should include means for monitoring the ability of the overall system to deal with the changes, and means for detecting signs of stress or danger. While there might not be a well-defined absolute limit where the addition of a single additional entry might break something, there may well be a general range where operations become slower, more error prone or unwieldy.

5. Taxonomy of errors

The root zone is generally viewed as perfectly accurate, i.e. no errors at all once changes are accepted by IANA and moved through the provisioning pipeline. The actual experience is not quite perfect, but it has been close to perfect. The goal is to retain or improve that record, attaining the highest possible level of accuracy and integrity of the data in the root zone while also propagating changes quickly and providing highly responsive and resilient service to all clients.

With this goal in mind, it is axiomatic that any system that involves multiple parties and a sequence of transactions is likely to have errors. The study should include a taxonomy of errors and a model that predicts the likely types and rates of error. Baseline data, to the extent it is available, should be compiled but extrapolation from existing data is probably not sufficient.

PROVISIONING – TRANSACTION TYPES

Within the provisioning system, the process may viewed as a pipeline. Changes are requested by TLD operators. They are processed first by ICANN, then passed to NTIA for authorization, then to VeriSign for assembly into the root zone and distribution, and then to the root zone operators for insertion into their servers. Several of the root zone operators operate a distributed complex of servers, so the distribution process has multiple steps even within a single root server operator.

There are multiple types of transactions that affect the root zone. The main types of transactions are:

Delegation and re-delegation

These are additions of new top level domains or the transfer of operation of a top level domain from an existing operator to a new operator. (In principle, there might also be the removal of a top level domain. This has rarely happened in the past, though it might happen more often in the future.)

Changes in contact information

There are usually three official points of contact for a top level domain, the formal head of the operator, the administrative contact and the technical contact. Each of these can change from time to time.

Changes in the set of name servers

Each top-level domain is served by two or more name servers. Top-level domain operators occasionally change or add name servers to their set.

Changes in the addresses of name servers

Name servers are occasionally renumbered. Also, when a new name server is added to the set serving a TLD, its address must also be added.

We distinguish this transaction from a change to the set of name servers because some name servers serve multiple TLDs, and a change to the address of such a name server must be coordinated with all of the TLDs it serves. (One of the subtle effects to be examined in this study is the complexity of this coordination and how it will change as the number of TLDs increases. Is it likely there will be greater aggregation of service and hence a larger dependency on a small number of large name servers?)

PROVISIONING – METRICS

As noted above, there are currently fewer than 300 TLDs. These TLDs cause approximately one change per TLD per year on the average, aggregated across all types of transactions, so there's roughly one change per day for the entire system. This process also makes occasional errors, but there is little documentation of the details.

As the number of entries increases, how will these numbers change? What are the capacities of the system? Where are the thresholds?

Equally important to understanding the trends and limits is the need to identify mechanisms of avoiding expansion beyond limits that are considered to be operationally feasible. Are mechanisms in place to anticipate reaching these thresholds and to make adjustments? Are these realistic?

There are also quantitative questions related to the rate at which new TLDs can be added. Is there a limited capacity and, if so, what will happen if there is pressure to exceed that capacity?

All of these questions apply to all parts of the provisioning system, including the distribution of new zone files to remote instances of highly replicated root zone servers, i.e., anycast instances. For example, how much do DNSSEC and IPv6 expand the zone file and/or increase the required frequency of updates? At what point does the current method of distributing zone files become a problem for each root server operator, and hence may require a change in either the frequency or method of distributing the root zone?

Finally, it is important to address how the increase in the size of the root and/or the increase in the rate of change to the root will affect the number of errors made in the process. The current system is good but not perfect. Scaling it up is likely to increase the number of errors, at least in principle.

ADDITIONAL COMMENTARY

- Global interoperability of the DNS is a strong principle, including fostering continuation of a single root zone. An objective of this study is to determine whether there are limit points to how much change the root zone can absorb and still retain core stability and security, or intermediate milestones that can be used to identify when changes to operations, infrastructure, or process will be necessary to ensure the security and stability of the root zone operations, or even to identify limit points where further change should be limited or prohibited.
- The number of new gTLDs is uncertain, but a reasonable initial figure would be at least 500 new TLD applications by the end of the first year that applications are open. This estimate should serve, at best, as a lower bound, not an upper bound, for any analysis.

- DNSSEC uptake is likely to be gradual relying on many components beyond the root, but it should be assumed eventually all TLDs will be signed. At present, there are no signatures or key records in the root zone.
- IPv6 glue is already in the root zone and is being added at an increasing rate.
- Analysis of existing root zone operations data is likely to provide understanding of current and future root zone operations. There may be useful lessons to learn from the history of the growth of the largest TLDs, e.g. COM, DE, et al, and also from the TLDs that have deployed DNSSEC, e.g. SE. However, the root zone is operated under different rules, some explicit and some only implicit, than any TLD, so any comparison with TLD operations needs to be augmented with a corresponding comparison of the expectations and rules of operation.
- A possible outcome of this study may be identification of further necessary studies of changes over time for measured impact. The need for further study does not preclude reaching preliminary findings nor producing appropriate recommendations that account for anticipated DNS behavior in light of existing data.

SUGGESTED MAJOR STEPS

Information Gathering

- Request information about root update and lookup processes at IANA, NTIA, VeriSign (as editor and distributor), and the Root Server Operators. Request information about measurements and other operational issues from OARC and CAIDA. Request information about concerns and prospective needs from GAC, GNSO, and CCNSO.
- Interviews with TLD operators as clients who create transactions
- Interviews with ISPs operating large recursive resolvers
- Outreach meetings TBD. Likely to include business community, governments, regional operator groups (NOGs), etc.
- Inputs from a publicly accessible forum

Initial Public Documents

- Description of the existing system, with capacities, rates, delays, errors, etc.
- Initial catalog of questions and intended treatment of them, e.g. “to be answered within this study,” “longer term research required” or “not within the scope of this study”
- Bibliography of prior and current reports and sources

Initial Assessment

- Analytic approach
- Review of information gathering and initial description
- Review of planned next steps including subcontracts
- Identification of obstacles

Draft Results

- Drafts of primary results
- Review and agreement on path to completion

Final Results

DELIVERABLES

1. Baseline Description of the Root System

A description of how the existing root system works, including both the provisioning and lookup sides of the system. The description should include quantitative measures of frequencies of transactions, volumes of data transferred along each link, delays, variances, and error rates. Where data is not available, the description should identify what data is missing. The description should also include the capacities of the various parts of the system.

2. Catalog of Questions

A compendium of questions related to the scaling of the root zone. The collection should be as comprehensive as possible, even if the questions are somewhat out of scope to be answered within this study. An important example of a question closely related to this topic is ``what mechanisms will be available to ensure that zone growth is curbed before reaching a range that the study concludes to be unsafe".

The catalog should be organized into a sensible taxonomy.

3. Model-based analysis of root zone scaling issues

The principal deliverable of the study will be a model of the root server system (including all of its provisioning and query components) that shows how the different parts are related, and how changing the value of a variable in one part affects each of the other parts. It should be as quantitative as possible given the time available for the study, but will inevitably be more qualitative in its expression of some system interactions. The idea is not to focus narrowly on identifying bad things that might happen if (for example) the number of TLDs in the root zone increased by several orders of magnitude during the same time period in which new RRs were added for DNSSEC, but to construct a model that demonstrates what the effects would be throughout the system of changing the value of one or more variables (e.g., the rate at which new TLDs are added to the

root, or the expected frequency of emergency key rollover events in a signed root). Then, when policy decisions about the root are made, everyone will have a clear picture of what the consequences of any decision will be. The study should not try to answer the question "how many entries in the root zone is too many?" but to show, as clearly and definitively and authoritatively as possible, what the consequences must be for each part of the system of changing the value of each variable.

- Analysis of “Plus 1” issues

Impact of both the qualitative issues related to the addition of DNSSEC, IPv6 and IDNs, and increasing the root by as much as a factor of ten and/or increasing the rate of change by as much as a factor of 30. Estimate of capacity to add TLDs on a daily, monthly or annual basis. Analysis of impact on the lookup. Identification of relevant unknowns. Suggestions for reports or signals to monitor the growth and its impacts.

- Analysis of “Plus 2” issues

  Impact of increasing the root by as much as a factor of 1000, i.e., to 300,000 TLDs and/or increasing the rate of change by as much as a factor of 1000 with particular attention to the structural changes that may be needed throughout the entire system. Identification of second order factors that may become dominant with that much growth. Discussion of what things might go wrong, including the rate of errors or intentional disruptions, e.g., hijackings or denial of service of a TLD or the root zone.

- Analysis of “Plus 4” issues

Impact of growth beyond 300,000 names in the root zone and corresponding increases in rate of change, i.e., this is the case where ``TLD creation'' becomes a commodity item comparable to domain name registration as it exists from many top-level domains today.

APPENDIX: PRIMING

The only specific question identified so far that is related to complexity as opposed to size is the impact of DNSSEC on the priming sequence for validating resolvers. That is:

When a validating resolver is first started, it uses a hints file or other initial “guess” to find a root server, and then it asks that root server for the current list of root servers. The answer is the full list of thirteen root servers and their addresses. Until very recently, that answer fit within the 512 byte limit of a traditional IPv4 packet. With the inclusion of IPv6 addresses for root servers, the response is now longer. Fortunately, longer packets are routinely supported by

most transport systems. See SSAC report 018, http://www.icann.org/committees/security/sac018.pdf.

However, when DNSSEC signatures are added to the root zone, the response to the priming query will increase yet again. Preliminary examination suggests the response cannot be accommodated within a single packet, so the primary query will necessarily become a priming sequence. Moreover, it appears that responses from NSD and BIND are different, so there is some work to be done to flesh out the details and make sure there is a feasible priming sequence for all of the implementations used across the thirteen root servers.