Measuring the Health of the Domain Name System

Report of the
2nd Annual Global Symposium on
DNS Security, Stability and Resiliency

1-3 February 2010
Kyoto University
Kyoto, Japan

Measuring the Health of the Domain Name System

Report of the
2nd Annual Global Symposium on
DNS Security, Stability and Resiliency

Sponsored by
Internet Corporation for Assigned Names and Numbers (ICANN)
The DNS Operations, Analysis and Research Center (DNS-OARC)
Nara Advanced Institute of Science and Technology (NAIST)
Kyoto University, Center for Computing and Media Studies

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

TABLES

FIGURES

1 EXECUTIVE SUMMARY

The Domain Name System (DNS) is a highly distributed, hierarchical naming system for computers, services and resources connected to the Internet, which associates information such as network addresses with a human-readable domain name. Since its inception in 1983, the DNS has been a core element of network architecture. However, as the Internet has grown more complex, questions have arisen about how to determine the health of the DNS and, particularly, whether the DNS is being subjected to a debilitating attack or some kind of new security-relevant exploitation.

To this end, the organizers of the 2nd Global Annual Symposium on DNS Security, Stability and Resiliency chose to focus this year's conference on the theme of measuring the health of the DNS. The entire Internet relies daily on the DNS, understanding its health – both instantaneously and as it changes over time – is critical for being able **to reasonably predict the DNS's health outlook** and **to decide whether to take corrective measures.**

Impact

In recent years, interest in the security, stability and resilience of the DNS has grown substantially, even in sectors far removed from the DNS itself. This interest is as a direct consequence of the role DNS plays in the continued smooth functioning of the Internet itself, as well as the fact that certain aspects of the DNS have played a role in various recent Internet-based attacks.

Being able to rapidly and accurately identify worsening trends in terms of DNS health would be beneficial not only to the traditional DNS community, but to a larger body ranging from policymakers to end users. The impact of such a system would include increased confidence in protective measures, a more robust ability to react to negative developments, and greater insight into how to make the best use of existing and new measurements.

Key Conclusions

DEFINING 'HEALTH'

Although there is broad consensus that certain sets of existing measurements form a baseline, **there is no consensus on what precise set of parameters define a 'healthy DNS.'** Because there are many possible measurements that could be used to make an assessment of DNS health, participants felt that they were seeing only "the tip of the iceberg" and that more research is needed to identify the best set of measurements that could be used to make a reliable DNS health assessment.

Nevertheless, in the shorter term, **the DNS community needs to develop an interim high-level consensus of 'DNS vital signs'** that can be used to determine system health, covering the domains of coherency, integrity, speed, availability and resiliency.

In addition, the Symposium brought out the fact that a DNS may be healthy when viewed from an Internet-wide perspective, but this might not reflect the situation in a given local network. Therefore, having a

simple numerical score or color code to indicate the health status of the DNS would likely be unhelpful; instead, **a 'dashboard' showing a small number of DNS key performance indicators,** including whether present measurements are nominal or unusual, could be useful in both local and global contexts.

In domains such as resiliency and integrity we have few, if any, objective measurements. More practical research is needed in these areas to ensure we can deliver a global view of DNS health that takes these important factors into account.

USER PERSPECTIVE

The participants felt strongly that user experience is the most important metric of all, but that there is very little data on hand to explain what the user perspective is. It would be helpful to obtain **more information about user expectations, perspective and experience** with the DNS, both through measurement and through user surveys.

DATA SHARING

For data sharing to become more widespread, **standard data interchange formats** are needed. In addition, new or existing data sharing points (data wells) should be opened or expanded in order to ensure that measurements can be widely taken into account. In addition, there should be a comprehensive look at the legal and other barriers to sharing information in a DNS health measurement context.

Future Actions

Work on measurement can be undertaken by anyone in the DNS community. For its part, ICANN has identified three key areas of future work that it hopes to act upon during 2010:

- Identifying risks to the DNS
- Developing exercises intended to measure the DNS resiliency and strengthen its operational processes
- Finding ways to achieve the creation of a 'data clearinghouse' for DNS measurements

Community needs

While individuals and organizations that are outside the DNS community itself will be most interested in monitoring a global 'dashboard' view of DNS health, those who are inside the community may find a global view of only peripheral interest. This is because DNS operators and many others in the community will need to make local judgments about the health of the DNS in their particular circumstance.

For this reason, one of the most important requirements is to increase lateral communication within the community, in order that relevant and appropriate data sharing can be facilitated. To this end, the development of common data formats and a mechanism for allowing data sharing on a wider basis would be beneficial. Through the use of such data, the community could develop a wider range of aggregate analysis that can serve broader audiences.

2 OVERVIEW

As a core element of the Internet infrastructure, the domain name system (DNS) serves a broad constituency, ranging from DNS Registries through to individual Internet end-users. Because of the importance attached to ensuring the uninterrupted operation of the DNS – both globally and locally, the Symposium on DNS Security, Stability and Resiliency (SSR) sought to focus on the problem of measuring the health of the DNS and even understanding what the term "health" means when used in the context of the Internet.

What is DNS Health?

We use the term 'DNS health' as shorthand to refer to how well the DNS is functioning at the moment, but what constitutes a 'healthy condition' depends largely on the context, or point of view, of the person making the judgment. In the world of medicine, the term 'health' is defined as:

> "A state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity"[1]

In order to make an assessment of a patient, a physician measures vital signs as part of a patient assessment. If we carry the above analogy a bit further, we can treat known classes of measurement such as coherency, integrity, speed, availability and resiliency as the 'five vital signs of the DNS,' to be used as part of a health assessment of the system.

FIGURE 1 – THE FIVE DNS VITAL SIGNS

Using these factors, we can try to create a more precise, yet flexible definition of DNS health, such as:

> **"A state of general functioning of the DNS that is within nominal technical bounds in the dimensions of coherency, integrity, speed, availability and resiliency."**

Such a definition does not rule out the possibility of the existence of failure states within the DNS, so long as those conditions are within generally accepted 'nominal' boundaries. This is very much in line with medical practice, where tools such as comparative studies and historical data are used to make determina-

[1] Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June 1946.

tions about the health of a patient. Such judgments are often free of stringent limitations, relying on the art of medicine to render the final judgment about a patient's present state.

Perspective

The analogy between the human health and DNS health was highlighted repeatedly at the Symposium, insofar as neither system has a precise definition of what factors make the measured system 'healthy'; measurements are used to make subjective judgments based on one's individual perspective. This in turn means that 'health' as we perceive it locally might not have a direct correlation to 'health' as seen from an Internet-wide view.

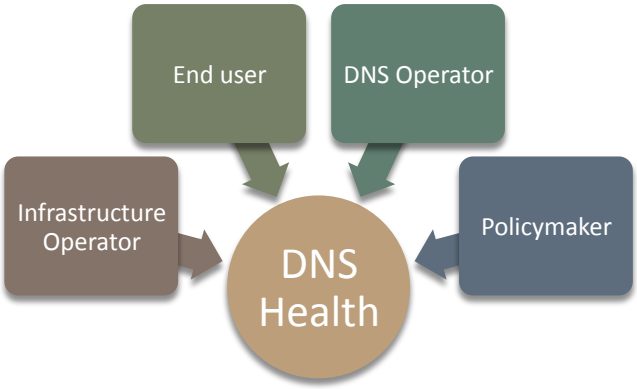


FIGURE 2 - RELATION OF ONE'S VIEWPOINT TO DEFINITION OF HEALTH

The invited keynote speaker described how perspective affects our conception of health by making an analogy to the case of a forest ecosystem. A healthy forest (viewed from a whole-forest perspective) requires by necessity that some individual components of the forest be in a distressed state. Naturally occurring wildfires are but one example of a locally harmful condition that contributes to the overall health of the system.

This idea of "public health" (global view) and "private health" (local view) is an interesting distinction[2] and one that also came up in the interactive track sessions. Aside from viewpoint, one factor that helps distinguish these domains is the availability of measurement statistics from within a target domain. For this reason, the ability to share and aggregate data is an important background policy dimension.

While a variety of opinions on what constitutes 'health' of the DNS were raised, the consensus of the conference on defining DNS health was that:

- In current practice, several classes of important measurement **(coherency, integrity, speed, availability and resiliency)** are believed to be fundamental to assessing DNS health.
- No consensus has yet emerged on precisely what factors are necessary to state unequivocally that the DNS is **healthy on a system-wide basis.**
- A judgment about DNS **health on a local basis is, in effect, a subjective decision of the operator** or other DNS community member. Even if operators make use of shared wells of measurement information, each one will use local factors to determine DNS health in that specific setting.

[2] The terms "public health" and "private health", as they relate to DNS, were raised during the keynote speech and also in the Q&A session after Shinta Sato's invited talk about monitoring and measurement.

Linkages

Despite the evolving nature of the definition of DNS health, an important aspect of understanding the state of the DNS is creating linkages between classes of measurement and the five DNS vital signs. Determining how to link measurements to indicators was one of the key points of the Symposium, but the community's understanding in this area remains limited. As the broad understanding of DNS health increases, the linkages we know today may strengthen or weaken. However, what we presently understand about these associations can serve as a guide for further analysis and refinement.

Coherency

- Time for zone convergence across instances
- Convergence of zone files at a point in time
- SOA serial synchronization across all servers

  (cutting across anycast servers)
- Specific queries to multiple instance of nameservers (at the same time) should result in identical "answers" section contents

  (for instance, the query for the serial number represented by the "SOA" record)

Integrity

- Delivery of the "correct" or "desired" answer

  Integrity is important across the system, not just at the point of publication. [Note 1]
- Number of DNSsec validating resolvers

  (in theory, use of DNSsec should increase integrity from the point where the zone is signed onwards)

Speed

- Response times [Note 2]
- CPU load and other internal measurements [Note 3]

Availability

- CPU load and other internal measurements [Note 3]
- Remote measurement [Note 4]
- Distribution of all query parameters by IP address, full *q*-tuple fanout (at each server)

  (should show which servers are topologically nearest to the clients)

Resiliency

- Bandwidth consumption
- CPU load and other internal measurements
- Query load

  (related to capacity for resiliency against denial-of-service attacks)

Unlinked measurements

- NXDOMAIN detection (at single points)
- EDNS and TCP fallback at recursive resolvers
- Amount of DNS spoofing (NXDOMAIN or other 'wrong response code')
- Number of existing TCP half-open connections
- Validation of DNSsec Booleans

Note 1: A definition of 'correctness' is needed to ensure integrity of this measurement. One definition suggested during the Symposium was "the answer given directly reflects the wishes of the data owner."

Note 2: The accuracy of response time measurements is highly dependent on the number and distribution of monitoring nodes. Ideally, measurements could be collected directly from user sites.

Note 3: This value is a local measurement.

Note 4: This value would ideally be measured from a wide variety of locations.

TABLE 1 - LINKAGES BETWEEN MEASUREMENTS AND VITAL SIGNS

Conference Goals and Organization

The Symposium was hosted by Kyoto University, in Kyoto, Japan, and its overarching objective was to define what knowledge is needed in order to make an assessment about whether the DNS is presently "healthy," as well as to determine what particular data are needed to support such an assessment.

Goals

The Symposium organizers established a conference design that worked along three main avenues of inquiry, intended to address gaps in knowledge that are relevant to critical information infrastructure in general, and the DNS in particular:

1. Understanding the meaning of "health" as it pertains to the DNS system and reviewing the current state of the art of measuring its health.

2. Identifying gaps in existing techniques, mechanisms and metrics for measuring the state of DNS system health.

3. Developing recommendations for improvements in how to monitor the system's condition.[3]

In order to emphasize the participatory nature of the Symposium, about 40% of the conference was devoted to interactive track sessions led by facilitators who used key topics to keep attendees focused on particular topic areas. However, conference members were free to raise any topic relevant to DNS, even if the organizers had not foreseen it.

Prior to the Symposium, the organizers sought to achieve from the conference a set of actionable and measurable outcomes:

- To identify measurements that are being (or should be) taken relating to the security, stability and resiliency of the DNS system
- To determine how to assess the effectiveness of these metrics, with an aim toward improving the quality of measurements used to determine system health
- To identify the various qualitative effects that operators and researchers are taking into account when looking at the health of the DNS system
- To assess and document the status of outcomes reached at the previous Symposium
- To identify current impediments to the collection and sharing of metrics as well as possible solutions

Organization

The Symposium was held over 2-1/2 days, which were organized as follows *(for the full schedule, see Appendix A, Symposium Agenda):*

[3] *Concept paper: 2nd Global Annual Symposium on DNS Security, Stability and Resiliency,* The Symposium, December 2009.

Monday	Afternoon	Plenary, organized by DNS-OARC
Tuesday	Morning	Plenary
	Afternoon	Breakout/interactive
Wednesday	Morning	Breakout/interactive
	Afternoon	Plenary

TABLE 2 - CONFERENCE ORGANIZATION

The main thrust of the program for the first ½-day, Monday afternoon, was to present views on what constitutes DNS health, as seen from a wide variety of points of view, notably including:

- Postulation of **how to identify bad actors** by calculating reputation or observing behaviors
- Observations about **future DNS trends** based on routinely collected statistics
- How to manage the **ever-increasing volume** of measurements
- Attempts at concretely **defining 'health'** in the context of the DNS
- Remarks on the **impact of DNSsec** rollout on the DNS
- The relationship between **Registrars' weaknesses** and recent attacks

In order to assure maximum participation from participants during the interactive breakout sessions on Tuesday and Wednesday, these days were run under the Chatham House Rule:[4]

> *When a meeting, or part thereof, is held under the Chatham House Rule, participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.*

For the purposes of this meeting, this rule was modified to allow the publication of an attendee list to the Symposium.[5] For this reason, other than for the scheduled talks (which were advertised publicly in advance of the conference), identification of speakers has been withheld in this report as regards discussions and post-conference feedback.

[4] http://www.chathamhouse.org.uk/about/chathamhouserule_translations/ (Accessed 2 February 2010), Chatham House, London.

[5] Attendees were offered the opportunity to have their name removed from the attendee roster.

3 OUTCOMES FROM PREVIOUS SYMPOSIUM

The 1st Annual Global Symposium on DNS Security, Stability and Resiliency was held in February 2009 at Georgia Institute of Technology in Atlanta, Georgia, and the report from that Symposium is available online. The key learning outcomes from that conference can be summarized as follows:

1. The level of broad awareness of the risks in using the DNS—and of risks to the DNS itself—is low;
2. The DNS operational, technical and response communities are disjointed with few focal points;
3. The DNS community as a whole lacks accountability standards as well as a method for measuring itself against any standards that might be developed;
4. Concerns exist about organizational outsourcing of DNS services which might have been made based on incomplete reviews;
5. Members of the DNS community would like a clearer understanding of ICANN's mission and role with respect to the DNS;
6. A collaborative structure is necessary to set priorities, establish goals and select implementation teams, initiate work and measure progress; and
7. There needs to be some kind of collaborative response, awareness, technical training and future DNS Security, Stability and Resiliency body.

Since the 2009 Symposium, a number of noteworthy incidents have occurred that have focused attention on the DNS both from within the community as well as from the broader Internet and technology policy communities. To a greater extent than before, the daily reality of risks posed to and by the DNS has become apparent.

Of particular note was the Conficker worm, which came to the attention of Internet security researchers and antivirus vendors in late 2008. Combating Conficker and working to disable its command and control (C&C) system required extensive collaboration within the DNS community as well as with a larger body of interested parties.

While these efforts proved effective, a principal lesson learned by the Conficker Working Group was that having a **dedicated and sustained incident response coordination capability** would have enhanced the global response to this issue. Although there was a small amount of discussion about the potential for creating a DNS CERT in the future, this discussion was out of scope for the Symposium and instead was taken up at a separate informal luncheon meeting led by ICANN staff members.

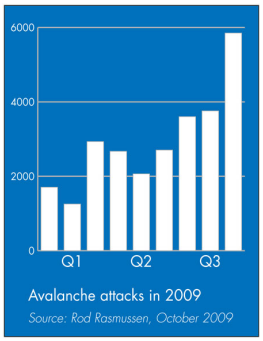


FIGURE 3 - AVALANCHE ATTACKS

Conficker was far from the only active threat in 2009: Hijacking of domain names at Registrars has been on the rise, largely in support of using active measures to better disguise C&C networks for botnets. However, other denial of service attacks, such as the attack against the Internet search engine Baidu.com, appear to have made use of attacks upon DNS Registrars as a pathway to achieving their goal of service disruption. In many cases, the Registrars have insufficient security response capability to cope with an avalanche of customer complaints that might accompany successful attacks.

Actions Since the Previous Symposium

During the past year, ICANN has embarked on several projects that stem directly from the previous year's DNS SSR Symposium:

1. Communicate ICANN's Role and Mission Relating to the DNS

   Publish a Security, Stability and Resiliency Plan that explains ICANN's activities in the area. An initial plan has been published and is under continual review; an updated edition is planned for later in 2010.

2. Increase DNS Awareness and Community Collaboration

   Increase the awareness of DNS operations, risks and threats outside of the traditional DNS community through proactive outreach.

3. Capacity Building

   Develop and deliver training programs for ccTLDs in collaboration with the regional ccTLD groups. The goal of these exercises is to improve the capabilities of and processes used by the various actors in the market, including Registrars and Registries.

4. Improve Capacity for Coordinated Action and Response

   Work with ccTLD community, through ccNSO, on incident response mechanisms. In parallel, explore the efficacy of establishing a DNS CERT function.

5. Identify the Need for Standards and Measurements

   The theme of the 2nd Annual Global Symposium on DNS Security, Stability and Resiliency is focused on the needs for measurement and standards.

Because a number of the Symposium attendees had not seen the result of the previous year's conference, the ICANN staff indicated a desire to more actively circulate the report of this year's Symposium, both to ensure that the lines of communication with the DNS community remain open and to ensure that its future planning assumptions remain valid.

4 TODAY'S STATE OF THE ART IN MEASUREMENT

In order to chart a course toward associating specific measurements with particular good or bad conditions, it's necessary to develop an understanding of what measurements already exist in the DNS community, what constitutes a known "best practice" in this area, and what are 'known unknowns' – areas of measurement data that are currently known not to be adequately surveyed or collected.

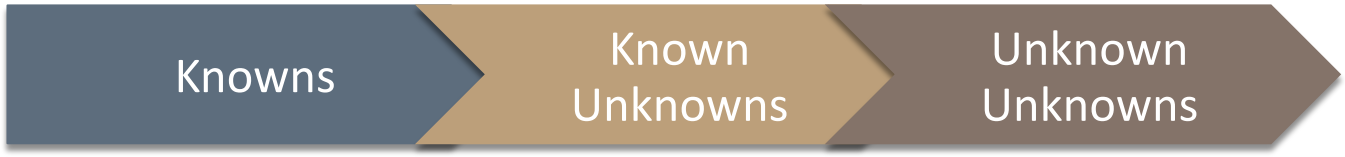


FIGURE 4 - DATA TYPES IN DECISION THEORY

The Symposium attempted to catalogue the present state of the art in measurement through three means: First, prior to the conference, it initiated an online survey asking invitees[6] to supply information about efforts at DNS measurement in their own networks. Second, a number of invited speakers provided detailed accounts of measurements they make, as well as the techniques used to gather them. Finally, *in situ* participants at the Symposium were asked to contribute more detailed information on a whiteboard in the break area, which was later used to support the track sessions.

Existing Measurements

As was discussed in the overview, even though there is no existing definition for DNS health, some factors contributing to a subjective assessment are known. These DNS 'vital signs' are measurements that can be taken in the domains of **coherency, integrity, speed, availability and resiliency.**

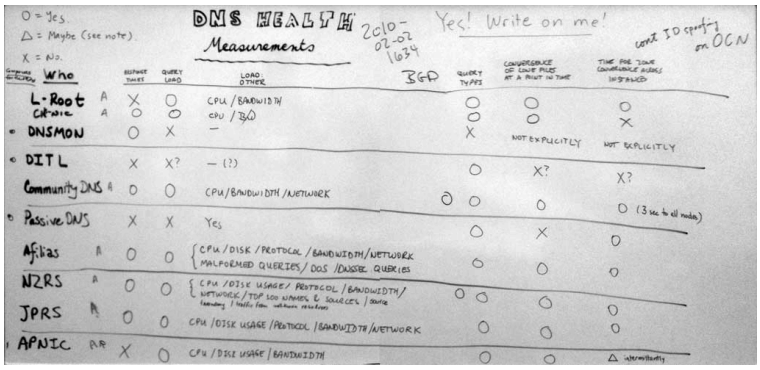
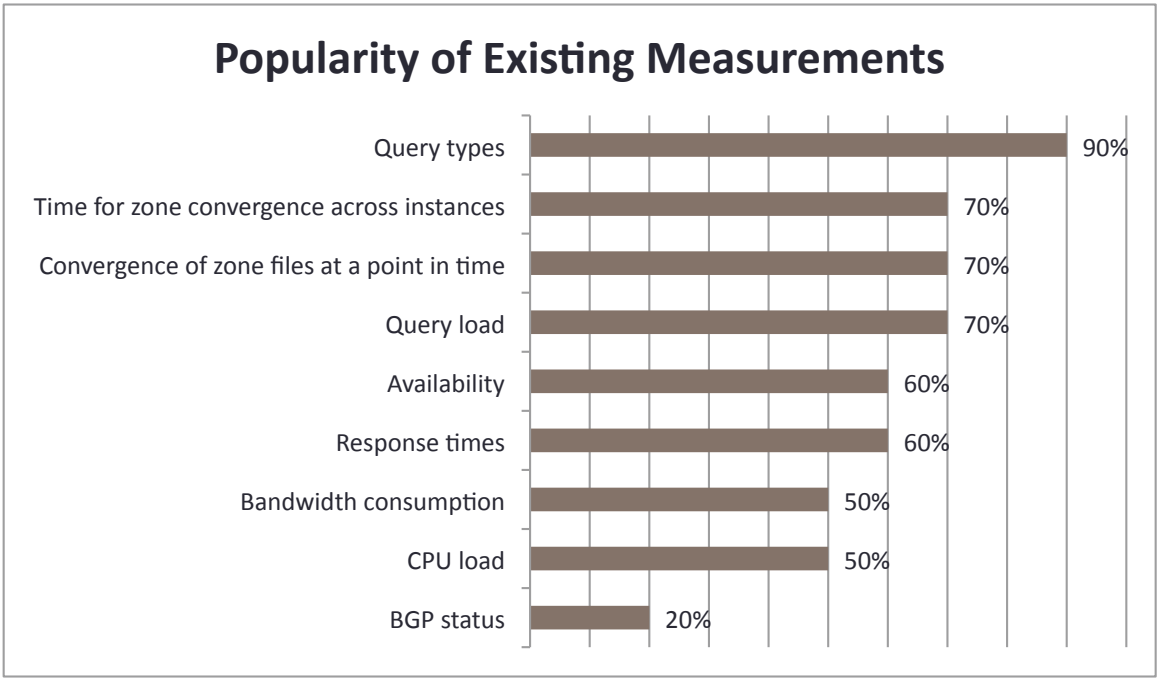


FIGURE 5 - PARTICIPANT FEEDBACK ON EXISTING MEASUREMENTS

While the measurements listed in Figure 6, below, are merely representative of the kinds of data collection that can be applied to the problem of ascertaining DNS health, the problem at hand is not limited to the measurement itself, but also includes **collection techniques**, agreeing on **common data formats** and, ultimately, **timely analysis** of the data.

[6] All persons who received an invitation to attend the Symposium were invited to add to a list of measurements and measurement techniques, regardless of whether they expressed an intention to attend this year's event. The organizers realize that the inputs may not represent a full cross-section of the DNS community efforts in the area of measurement, but would provide a sufficient basis to start a more comprehensive cataloguing effort.



Notes: (1) Some operators track protocol usage along with query type
(2) Measurement of time for zone convergence across instances is sporadic at some sites

FIGURE 6 - POPULARITY OF SPECIFIC HEALTH-RELATED MEASUREMENTS AMONG SYMPOSIUM PARTICIPANTS

The Symposium's ad hoc survey of measurements shows that there is a strong body of measurement works in the domains of coherency and speed. By contrast, there is comparatively little measurement activity in the areas of integrity and resiliency. Of particular concern is the impact that monoculture and lack of geographic dispersion could have on DNS resiliency – yet there are few extant examples of measures of monoculture, aside from anecdotal post-mortem accounts of failure.

One DNS operator at the Symposium offered the very concrete set of measurements used to assure the health of their infrastructure. While the measurements might be particular to their site requirements, an interesting aspect of their structure is that each of the measured items is mapped to a value, which can be an absolute value – scalable based on the infrastructure – or a relative value, which can be adjusted according to empirical rules.

	DNS Health	Prevention	Examination
Relative value	Amount of zone changes in a single operation must be less than 50% of its original size	Save the old zone file in preparation for emergency operations	Check the zone file size and stop the update in the event of an anomaly
Absolute value	Data update, provided every 15 minutes, must not fail 12 consecutive times (3 hours)	Place multiple master servers in distributed locations	Monitor SOA serial increase and alert operators in the event of an anomaly

TABLE 3 – EXAMPLE OF RELATIVE AND ABSOLUTE MEASURES IN AN OPERATIONAL ENVIRONMENT

Research need: Identifying patterns or trends

An important aspect of measuring any system is to understand the meaning (and limitations) of the collected data. Beyond this, identifying previously unknown patterns in data can lead to a deeper understanding of the system being monitored.

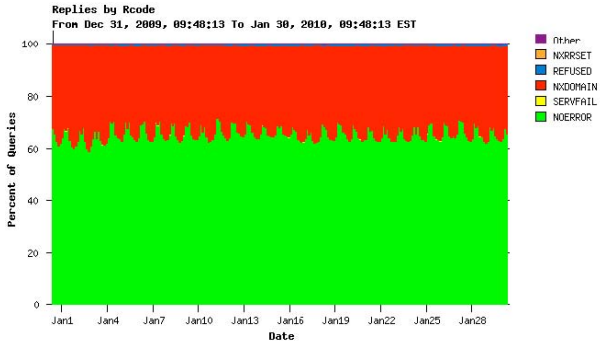


FIGURE 7 - DIURNAL NXDOMAIN PATTERN FOR REVERSE DNS

For example, an operator who attended the conference noted that they measure the rate of NXDOMAIN responses received in looking up reverse DNS queries. Over the presented period of time, the graph provided first-order information that can support an assessment of the strength of the reverse DNS value proposition and whether the reverse DNS is being properly maintained.

However, a second-order observation is that the NXDOMAIN rate in the reverse DNS varies in a diurnal pattern, whose presence could not be explained. Understanding the origin of this kind of pattern might or might not have long-term importance, but it does point out that DNS measurements are far from being exhaustively understood.

What else should we be measuring?

During the track sessions, users were asked what should the DNS community be measuring[7] that it is presently not measuring, or not measuring vigorously enough. Some of the responses appeared to be more like a project management framework, but all are listed below:

- Response times (at end user and resolver)
- Percentage of answers that are “right answers” (resolver)
- Response times, CPU load and other internal measurements
- SOA serial synchronization across all servers (cutting across anycast servers)
- Validation of DNSsec Booleans
- Distribution of all query parameters by IP address, full *q*-tuple fanout (at each server)
- NXDOMAIN detection (at single points)
- EDNS and TCP fallback (recursive resolver)
- Cache hits (recursive resolver)
- Query count (overcoming blinding through caching) (client-side)
- EDNS0[8] and deployment implementations

[7] For the purposes of discussion, the participants were asked to ignore cost and feasibility considerations, within reasonable bounds.

- Amount of DNS spoofing (NXDOMAIN or other 'wrong response code')
- Number of existing TCP half-opens
- Client characterization (i.e. information about the client placing the query)
- Number of DNSsec validating resolvers

5 BREAKOUT SESSION ORGANIZATION

The Symposium was intended to serve as a relatively free-format brainstorming session. To that end, the number of invited talks was held to a minimum; a large part of the conference was devoted to track or breakout sessions in which there were no prepared lectures or presentations.

Conference Pre-work

Prior to the Symposium, attendees were asked to be prepared to "share their perspective" on and "current thoughts about DNS health, as well as methodologies to measure that health." Post-conference feedback suggests that this is a good idea, but that the assignment should be put to the attendees more directly, so that the assignment might be better understood by the attendees.

Breakout Arrangement

The attendees were divided at random into two roughly equal-sized groups and asked to consider, under the guidance of a facilitator selected from the program committee, four key topics, which were identical for each of the two groups. The topics themselves were not the focus of the conference, but instead a device intended to keep the two groups from diverging wildly from one another.

The organizers' goal was to have each group work independently on roughly the same problems, and then to 'compare notes' at the end of each day. In this way, it would be possible to see the similarities and the differences between the two groups. The program committee felt that this type of approach would positively affect the interaction and creativity of the groups.

The pre-work covered the same four general topic areas as were covered in the four quarter-day breakout sessions:

Topic 1: Measurement – a Global, Internet-wide Perspective

What is the 'Health' of the DNS in the context of Internet functionality?
How do failures in one part of the Internet impact other parts?

[8] Paul Vixie, Extension Mechanisms for DNS (EDNS0), RFC 2671, Internet Engineering Task Force

Topic 2: Measurement – a Local Perspective

Is “Health” a uniform and global property?
Are there lessons we can learn about “health” of the DNS by looking at the practice of medicine?
What are the key performance indicators that can be extracted from raw measurements?

Topic 3: Measuring the DNS

Is Loose Coherency a disappearing norm?
- What constitutes an “accurate and timely” answer?
- How much does perspective influence the measurement?
- Does the industry have common metrics?

Topic 4: Gap Analysis

What is it that we want to learn, and do we have the means to learn it?
Are all actors involved?
What should be the recommendations for further action?

6 BREAKOUT SESSION HIGHLIGHTS

Although the track sessions followed the same main topics, they tended to diverge in terms of how they approached the problems and, consequently, the form in which the results were expressed differed as well.

Observation #1: Detecting change is more important than absolute measurement

There was a feeling that it's easy to overcomplicate the idea of health – that what ought to constitute "healthy" simply devolves into the question of "can I get an answer?" and not "is the answer delivered speedily." With respect to timeliness, in particular, what is important isn't necessarily instantaneous speed of answer delivery, but instead the trend of speed of delivery.

Many Internet users are using outdated computers with old operating systems; naturally, these systems will be slow to start with. That kind of user won't perceive that the DNS is fast or slow, instead relative changes of speed is important. **Both track sessions agreed that in determining the 'health' of the DNS, change in the speed of answer delivery is more important than the raw speed of the response.**

An open and possibly unanswered question in the area of user experience is, **"what do users expect in terms of DNS timeliness?"** If we accept the proposition that the end user is the only party whose satisfaction matters, then understanding this expectation would seem an important step to take.[9]

However, it was pointed out that the speed of the Internet experience for an end-user depends upon a large number of factors, of which DNS query response is merely one. For this reason, a number of participants suggested that **more work should be done on how to take measurements at the client.**

Making a medical analogy to comment on client measurement, one participant noted that in a community where mercury poisoning is suspected, the health authorities don't undertake massive blood sampling to confirm their suspicions. Instead, they take soil samples in the affected area and use that data to make statistical inferences. The underlying premise is that although it appears to be a daunting task to measure end users' experience in on a large scale, a much more modest sampling effort may yield the same result.

Although it is easy to focus on the elements contributing to problems with the DNS, the speed and accuracy of DNS answers can be affected positively and negatively by a variety of factors. Rather than focusing on only the negative components, **measurement schemes should also take account of factors that contribute to positive trends in the DNS,** as seen from the point of view of the end user.

[9] One participant suggested that the expectation of users has been reported to be in the sub-second timeframe, however the basis for this claim was left unstated. The difficulty in arriving at an agreement in the area of user expectation led to a general consensus that additional research in the area of user expectations would be useful to the DNS community.

Observation #2: Someone has to establish norms against which to measure

There are many examples of failed attempts to over-summarize a complex system's status into a single indicator, such as terrorism color codes or threat condition numerical values used by a number of information sharing and analysis centers. The problem is that, in a complex system, a single numeric value can't express the system's condition in a usable way.

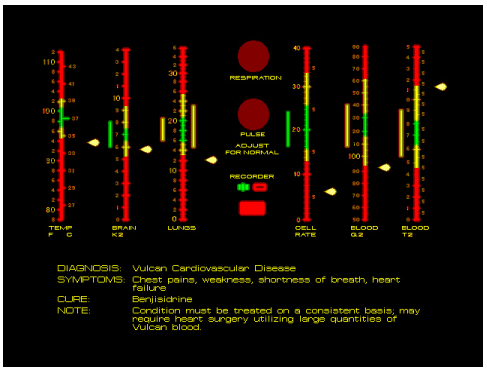


FIGURE 8 - SICK BAY HEALTH DASHBOARD

One participant suggested we turn to some popular fiction to think of an immediately recognizable way to conceive of a dashboard. In the original television series, Star Trek, each bed in the sick bay was equipped with a health dashboard of sorts that measured roughly ten or twelve factors. The important point is not that the factors are displayed on the board in a concise way, but that each of the measured elements has designated green, yellow and red zones, indicating whether a particular measurement is within tolerance for that patient. There is no single indicator of health on the display, yet even the lay viewing audience could tell immediately whether the patient was in trouble or not.

In other words, the development of a meaningful 'DNS dashboard' can only occur after norms are established and agreed upon by a significant proportion of DNS stakeholders. An important result of establishing and measuring against norms will be to provide transparency into what is normal and what is abnormal at a given time, leading inexorably toward a better opportunity to enhance high-quality early warning.

Observation #3: 'Systemic' health is not the same thing as 'System' health

When discussing health, we must differentiate between "system" health – meaning health on a large, probably Internet-wide, basis – and "systemic" health – referring to the situation throughout the DNS, right down to the local level. In October 2009, due to an administrative error in which a '.' was inadvertently omitted during the editing of a zone record, the entire .SE zone became unreachable for a time. This is a particularly extreme example of simple mismanagement of a zone record that caused unintended side effects that affected a measurable portion of all Internet users. The same kind of problem occurs regularly on a smaller scale in zones further down the tree, depending on factors such as the skill level of the zone administrators.

The question that arises from this observation is that, in such a case, is the DNS broken? The general opinion of the participants was that DNS itself is not broken because the protocols are operating as designed. However, one participant made the cogent observation that **when the authoritative records are bad, it can cause incorrect measurement of the health of the DNS.**

Further, currently circulating proposals that would allow for the delivery of different DNS answers based on other ephemeris related to the requestor, such as geolocation, can cause difficult to diagnose problems. In essence, in such a case, you could be measuring locally something that exists on a different data plane from that of the user. Yet, from the end-user's perspective, something is broken.

Observation #4: Risks should be enumerated and metrics assigned

During a discussion of risk in the DNS, it was posited that for a risk to exist, there must be the potential, however slight, for a negative consequence. A spirited discussion continued about theoretical risks of how the Site Finder wildcard DNS entry could be used to support the widespread distribution of malware if an associated parking page could be compromised. On the other hand, it was asked, "if a massively popular web services portal such as Google were to fail, is it a risk?" In other words, while the Internet at large would not be affected, the eventuated risk to a particular (and sizeable) set of end users would cause serious negative effects to them.

In this case, risk is magnified by the size of the user community and the awareness by those users of the mission criticality of the service they are using. The risk can be expressed as the probability of the risk's eventuation multiplied by the one-time cost of the event, which is likely to be quite a large value.

With respect to the DNS, we want to ensure that local problems do not become a system problem, potentially triggering larger-scale failures. It remains worth noting that what is critical to some people is not critical to others and that, in this sense, the local system is really the only one that matters in a significant way. The global DNS is designed in such a way that most failures are local; the probabilities of the entire system failing is extremely low. Although enumeration of risks could be helpful, because the previous Symposium spent a lot of time on risk-related issues, the track sessions did not delve into this area.

It was suggested that one way to allow users to gain an understanding of operating in a degraded DNS environment – and for operators to practice techniques for returning components to normal functioning – would be to hold a simulation or tabletop exercise designed around the theme 'a day without DNS.' Such an exercise might help users discern which disruptions are local and which are systemic, allowing them to better understand how to properly enumerate and quantify such risks.

Observation #5: Expansion of the DNS will happen outside of the standards track

It is unclear what would happen to DNS performance or operations if an exorbitant amount of data were added to the database. However, although it is widely acknowledged that it would be prudent to study the additional load through modeling and simulation, any expansion of the DNS is likely to be done in an ad hoc fashion, with any standards track activity happening after initial deployment.

How much data is too much data? How many queries are to many queries? It is hard to know the answer to these fundamental questions, which are bound up very closely to **the need to measure resiliency** and the **avoidance of single points of failure (SPOF) through monoculture.** The participants felt that it is incumbent upon the ccTLDs and gTLDs to measure the number of queries being served by their respective zones and to watch for trends as indicators of potential load related problems.

Resilience can often be determined by looking for evidence of increased risk to a specific operator's system as a whole, and by measuring how well the system is managing risk. However, we note that there are no currently known aggregates to this data. When trying to measure resilience, the kinds of things that can be measured include:

- Recovery speed
- Orphan glue and lameness statistics
- The dependency graph
- The quantity of load versus the quantity of queries
- Geographical concentration of DNS elements

Observation #6: There may be benefit in collecting experimental measurements

Although many DNS experts and DNS operators have developed measurement strategies to meet their operational needs, these are largely known factors or known unknowns. There is still a need to try to identify unknown unknowns, and in many cases this is most often done through high-level post-processing analysis of very large datasets, looking for the emergence of previously unknown patterns. Because the DNS is large and complex, it is natural that there should be measurable properties about it that we have not yet discovered. For this reason, **it seems prudent to encourage the collection of experimental measurements.**

The flipside of this observation is that large datasets can consume near-infinite storage resources. In addition, the more rich the measurement set, the more likely it is that privacy concerns surrounding the data will limit the ability of the data holder to share the data widely. Therefore, **the collection of experimental measurements and the sharing of those measurements should be considered as separate issues** based on relevant policy guidance.

Observation #7: There are both perceived and real barriers to data sharing

While there is great interest in the DNS community to share measurement data, the fact that operators are often competitors gives rise to a question of whether data sharing is viable on a large scale. In addition, depending on where the data are collected and by whom it is collected, there are potential legal, regulatory or privacy barriers that could impede the flow of data. These **legal and competitive barriers should be systematically reviewed** so that the friction against sharing information is reduced, with minimum imposition on the data collector.

Observation #8: There are many measurement corner cases

Measurement depends somewhat on the existence of a stable framework from which to take the measurement, but this can be confounded in a number of cases, **each of which is worth some independent examination:**

- Some DNS operators will serve up different answers in response to the same question, even though there is coherence at the zonecut and in terms of propagation.

- Caches and their properties are unmonitored. Recursive nameserver cache validity is a very difficult problem. Typically, these sites merely monitor memory usage, which is more of a function of server management than of watching over DNS.
- Measuring ephemeris such as clock skew (which has an impact on DNSsec performance) can serve as a performance degradation indicator.
- The taking of measurements at caching resolvers (which are always close to the client) could serve as a surrogate, to some degree, for measuring on the client itself.
- Forensic analysis can provide the kind of deeper analysis needed to better understand "the correct way" to monitor in the future.
- Registrars do not take measurements.
- Registries take measurements but do not share them.

No matter what statistics are collected, the DNS community **needs to make measurement data understandable to a wider audience,** including people who are not experts in the field.

7 GAP ANALYSIS

At the conclusion of the track sessions, the conference reconvened in a plenary session and tried to capture in a gap analysis what measurements, context and processes that the DNS needs in order to determine DNS health at an Internet-wide basis, as well as in a local context.

Measurements

The problem of selecting 'the best' measurements was likened to an iceberg, because although one can enumerate any number of existing measurements, the feeling among the audience was that the 'unknown unknowns' in the field are many in number. Although more research in the area of DNS measurement is imperative, the DNS community needs an interim **high-level consensus of which vital signs are needed** to determine system health, covering the domains of coherency, integrity, speed, availability and resiliency.

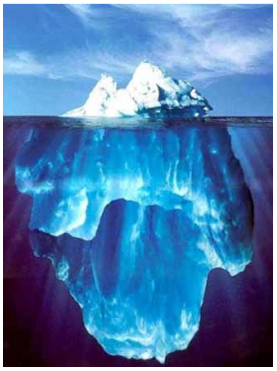


FIGURE 9 - TIP OF THE ICEBERG

Further, **in some domains we are entirely lacking objective measurements.** For instance, there are no useful measures for resiliency. Likewise, we lack a way to measure integrity on a global basis. Therefore, more practical research is needed in these domains in order that a global view of health can take these important factors into account in a measurable and repeatable way.

User Perspective

Because DNS performance is ultimately all about the user experience, it would be helpful to obtain **more information about user expectations, perspective and experience** with the DNS, both through measurement and through user surveys.

Data Sharing

Presently, there are a number of bilateral data sharing arrangements in place for DNS measurements. However, for the practice to become more widespread, it seems prudent to **establish some standard data interchange formats** in order to ensure the mutual intelligibility of exchanged data. In concert with this, common collection methods might be shared, as a sort of best practice guide, in order to ensure that similar data collected from disparate environments can be compared meaningfully side-by-side.

Because some data is shared or stored presently, it would be useful to **identify the extant data wells.** Registries, for instance, collect data but (anecdotally) do not share it very widely. User experience data is not only small in number but also not aggregated. These are but two points that beg the question of how to best warehouse data for both immediate health analysis and also for longer-term trend analysis.

8 SYMPOSIUM CONCLUSIONS

The 2nd Annual Global Symposium on DNS Security, Stability and Resiliency brought a number of important issues to the table as regards determining the health of the DNS. The net outcome from the Symposium seems to be that DNS health, as a measurable state, is not yet empirically well understood, yet such an understanding is important to ensuring the stability of a critical information resource.

Future Directions

DNS measurement is growing as an area of interest in terms of monitoring the health of the DNS. For its part, ICANN noted that it is undertaking a number of projects beginning in 2010 that could help increase the value of monitoring, as it relates to DNS operation. In particular, ICANN expressed that it has plans to undertake work in the areas of:

1. Risk identification

2. Development of exercises, keyed to:

   - Determining measures for resiliency, and
   - Helping the greater DNS community to fortify its processes and operational frameworks

3. Finding ways to achieve the creation of a 'data clearinghouse' for DNS measurements

Community Needs

While individuals and organizations that are outside the DNS community itself will be most interested in monitoring a global 'dashboard' view of DNS health, those who are inside the community may find a global view of only peripheral interest. This is because DNS operators and many others in the community will need to make local judgments about the health of the DNS in their particular circumstance.

For this reason, one of the most important requirements is to increase lateral communication within the community, in order that relevant and appropriate data sharing can be facilitated. To this end, the development of common data formats and a mechanism for allowing data sharing on a wider basis would be beneficial.

In addition, because the audience for DNS health is growing to include a larger group, composed of people who are largely not DNS experts, a method must be devised for communicating health information in a meaningful yet summarized way to this larger audience.

Lessons Learned (Participant Voice)

1. Developing a comprehensive set of measurements will require grappling with privacy and deployment issues.

2. While the participants were able to identify existing and desired measurements, no full consensus developed about what constitutes “health” of the DNS.

3. Further analysis requires knowing the criteria of measurement.

4. Others in the DNS community share an interest in measuring and fixing problems related to the DNS.

5. A model of the DNS is needed, to allow for a pathology not unlike that used by physicians when treating the human organism.

6. There is almost no aggregate information about the health of the DNS.

7. The community could learn something by doing measurement on resolvers; this could be done relatively inexpensively and quickly.

8. Analysis can sometimes be more important than measurement itself.

9. Measurement requires cooperation by the party collecting the measurement, in order to allow for sharing the information. Cooperation may, however, be impacted by legal, regulatory or privacy-related issues that may require mitigation.

9 POST-CONFERENCE THOUGHTS

After the Symposium concluded, some participants came up with a few thoughts that the Steering Committee felt were particularly relevant, even though they were not raised *in situ* during the Symposium.[10]

Expanding on Linkages to Specific Indicators

"Following the conference, we took a shot at mapping some of the measurements discussed at the workshop to the specific 'health indicators' and wanted to share this work with the rest of the community.

Coherency	With respect to SOA timing parameters
	Time for zone convergence across instances
	Specific queries to multiple instances of name servers (at the same time) should result in identical answers (as far as the answers pertain)
Integrity	This requires knowledge of the desired answer
	Also requires a definition (even if loose) of what constitutes a corrrect answer
	Integrity through the provisioning chain (Registrant to Registry to Zone)
	Possibly measureable through statistical sampling with controlled sets
Speed	Response times
	Accuracy is dependent on the number and distribution of querying nodes
	Ideal case would be to get measurements directly from user sites
	User expectations (may require sociological research or perception and expectations)
	CPU load and other internal measurements (affects speed but not a measure of speed)
Availability	Measured locally as a measure of availability, but also globally as an element of the DNS
	Could be measured as a percentage of the components being available
	Remote measurement would be an imperative because availability depends on location
Resiliency	Component systems
	Lack of single points of failure (SPoF)
	Number of servers
	Geographic, platform and vendor diversity

TABLE 4 – EXTENDED LIST OF LINKAGES TO SPECIFIC INDICATORS

[10] Feedback and comments have been lightly edited to fit the style of the report.

"There are a wide variety of other measurements that could be taken that might serve multiple indicators simultaneously, or provide more detailed understanding of one specific indicator. We developed the list below to show of some of these measurements, but others are likely to follow from additional study."

- Derivative measurement. The rate of change in the percentage of hosts deemed to have a certain base level of resiliency could allow for useful trend analysis.
- Ossification and standards compliance. Measures flexibility or rigidity of the system as a whole. Understanding the degree to which systems are running old editions of software, or are running on older hardware that cannot readily support newer standards could factor into this measurement.
- NXDOMAIN detection at a single point
- EDNS and TCP fallback at recursive resolvers
- Cache hits
- Query counts, as an indicator of the degree to which blinding through caching is overcome on the client side
- ENDS0 deployment implementations
- Amount of spoofing (NXDOMAIN, wildcarding, and others)
- Number of TCP half-opens
- Client characteristics
- Validation of DNSsec Booleans

Measuring Vulnerability

"One very obvious analogy to human health that we missed on, and one that ties in directly to the current state of DNS security is AIDS (Acquired Immune Deficiency Syndrome). An AIDS patient is certainly not 'healthy' in a normal sense; however, AIDS itself doesn't kill or 'sicken' people. Rather it is the lack of immunity to rather mundane diseases that does great harm to the patient. This analogy is apt for the DNS itself, in that we have multiple exposure points, and very little, if any, immune response to various 'mundane' attacks. The Baidu case is a perfect example of a simply executed security breach causing a great deal of harm to Baidu and its customers through the manipulation of the DNS system. Very straightforward security measures could be put in place to thwart (inoculate) or quickly recover (cure) from such attacks, but these tools are woefully lacking.

"Thus I would propose adding another key 'health measurement' class to add to the five already identified: **vulnerability** (which seems better than *immunity*). The closest category we had to this was 'resiliency,' but the context there was quite different, and didn't really cover this important aspect. Actually, the exploit of weaknesses can undermine any of the other five we identified as a result (*integrity* being the most obvious), but that's a separate measurement. If a system is vulnerable (like a person who isn't vaccinated), then

problems are likely to follow. I think this idea is very fundamental to getting after one of the key 'health' measurements of the Internet.

"I can immediately think of three existing metrics to measure vulnerability, and these statistics are collected today. First are surveys of the different DNS server software types in use throughout the Internet. Several older DNS software packages are vulnerable to attack, and this seems a good measure of 'health,' both locally and system-wide. Related to this are surveys of measuring how many servers are vulnerability to the so-called 'Kaminsky exploit.' I believe even up-to-date BIND versions can be susceptible if not configured properly. Finally, measurement of 'open resolvers' shows the extent to which DNS infrastructure can be exploited for reflective DDOS attacks. There are many other measures that could be added to this, including surveys of security practices for Registrars and Registries, numbers of domains in 'lock' status, domains controlled via free e-mail services based on WHOIS, and so forth."

Factoring Intent into Measurement

"In our breakout session, we briefly discussed 'DNS usage,' which referred to the fact that though the DNS may technically be working quite well, it may be serving up wholly undesirable results—botnets for criminals being the most obvious. While this might seem out of scope of the Symposium, it may well be an important consideration for some who are DNS constituents. Is the DNS considered 'healthy' if large armies of botnets are exploiting the system, albeit working as designed, to run their command and control networks or to maintain hosting of illegal sites via fast flux hosting? Technically, it may be, but many end users may hold the exact opposite opinion."

Clarifying DNS Health Terminology

"The term 'the DNS' is problematic because it isn't clear what it refers to. If it extends down to all of the local recursive resolvers, or stubs, then 'the DNS' is just a proxy term for 'the Internet.' We need to find a way to more clearly express what we mean when we say 'the health of the global DNS' without simply meaning 'the health of the Internet' – they seem to be at least slightly different from one another; we should not conflate the two terms.

"More concretely, although end-user response times are important, even poor end-user response times aren't necessarily indicating that the DNS has problems: lots of other factors play a part. If we look at service-level measurement models for things like the electrical grid, we might find measurement schemes on which the DNS community can capitalize, build on, or at least use as a basis for further discussion and research. This sort of approach has been used before in the areas of telco and power utilities and might be useful for DNS as well."

APPENDIX A: SYMPOSIUM AGENDA

MONDAY, 1 FEBRUARY 2010

13:00 – 17:00	TODAY'S STATE OF THE ART	*Roy Arends, Session Chair*

PLENARY SESSION

13:00 – 14:45

1. Sebastian Castro, NZRS, "Investigating Anomolous DNS Traffic"
2. George Michaelson, APNIC, "APNIC DNS Measurement & Perspectives on 'DNS Health'"
3. James Galvin, Afilias, "Data, data everywhere"

14:45 – 15:15 COFFEE BREAK

15:15 – 17:00

4. Keisuke Ishibashi, NTT Communications, "Characterizing DNS Client Behavior Using Hierarchical Aggregate Entropy"
5. Shinta Sato, JPRS, "Monitoring and Measurement of JP DNS and the Registry System"
6. Joe Abley, ICANN, "L-Root Update"
7. Wang Zheng, CNNIC, "January 12 Baidu's Attack"
8. Patrik Wallstrom, .SE, "DNSCheck and DNS2db"

17:30 – 19:00 SYMPOSIUM RECEPTION

TUESDAY, 2 FEBRUARY 2010

09:30 – 12:00	OPENING SESSION	
	PLENARY SESSION	
09:30 – 10:30	What's happened since the last Symposium?	*John Crain, ICANN*
10:30 – 11:00	COFFEE BREAK	
	KEYNOTE ADDRESS	
11:00 – 11:45	**"A Healthy Discussion of the DNS"**	*Andrew Sullivan, Shinkuro*
11:45 – 12:00	Introduction of Topics	Track Moderators: *Olaf Kolkman, NLnet Labs* *John Crain, ICANN*
12:00 – 13:00	LUNCHEON	
13:00 – 17:00	TRACK SESSIONS – DAY 1	
	TRACKS A & B – IDENTICAL TOPICS	
13:00 – 14:30	Topic 1: Global Internet-Wide View of DNS 'Health'	
14:30 – 15:00	COFFEE BREAK	
15:00 – 16:30	Topic 2: Taking a Local Perspective on DNS 'Health'	
	PLENARY CLOSING SESSION	
16:30 – 17:00	Daily Summation	*Hiroki Takakura, Co-Chair* *Yurie Ito, Co-Chair*

WEDNESDAY, 3 FEBRUARY 2010

09:30 – 12:30	TRACK SESSIONS – DAY 2
	TRACKS A & B – IDENTICAL TOPICS
09:30 – 10:45	Topic 3: Measuring the DNS
10:45 – 11:15	COFFEE BREAK
11:15 – 12:30	Topic 4: Gap Analysis
12:30 – 13:30	LUNCHEON
13:30 – 15:30	SUMMATION AND CLOSING
	PLENARY SESSION
	Tracks A & B – Track Chair Overview of Results
	Conference Co-Chair Summary
	Next Steps
	Symposium Close

APPENDIX B: ROSTER OF ATTENDEES

Participants in the 2nd Global Annual Symposium on DNS Security, Stability and Resiliency (February 1-3, 2010) held at Kyoto University in Kyoto, Japan:

Name	Affiliation
ABLEY, Joe	ICANN
ARENDS, Roy	DNS-OARC
BLOOM, Leslie	US DoD
CASTRO, Sebastian	.NZ Registry Services
CRAIN, John	ICANN
ELAND, Howard	Afilias
GALVIN, James	Afilias
HARDAKER, Wes	SPARTA
HOTTA Hiro	JPRS
ISHIBASHI, Keisuke	NTT Communications
ITO Yurie	ICANN
KANE, Paul	CommunityDNS.net
KATO, Akira	WIDE
KOLKMAN, Olaf	NLnet Labs
LEE, Han Chuan	Singapore NIC
LI Xiaodong	CNNIC
MANNING, Bill	ISI
MATSUURA Takayasu	JPRS

Name	Affiliation
MATSUZAKI Yoshinobu	IIJ
MICHAELSON, George	APNIC
MORISHITA Yasuhiro	JPRS
NAZARIO, Jose	Arbor Networks
OKABE Yasuo	Kyoto University
PISCATELLO, Dave	ICANN
RASMUSSEN, Rod	Internet Identity
RATTRAY, Greg	ICANN
REMALEY HASCH, Evelyn	Booz Allen Hamilton
ROBACHEVSKY, Andrei	RIPE NCC
SARAGIOTIS, Panagiotis	ENISA
SATO Shinta	JPRS
SATO Masaharu	NTT Communications
SAUER Kurt	Spinlock Technologies
SEKIYA	WIDE / Tokyo University
STEINGRUEBL, Andy	PayPal
SULLIVAN, Andrew	Shinkuro
TAKAKURA Hiroki	Kyoto University

Name	Affiliation
VIXIE, Paul	ISC
WALLSTROM, Patrik	.SE
WANG Zheng	CNNIC

Name	Affiliation
WOOLF, Suzanne	ISC
YAMAGUCHI Suguru	NAIST

Total number of Symposium attendees: 41

Program Committee

(Listed in alphabetical order)

- Roy ARENDS, DNS-OARC
- John CRAIN, ICANN
- Dave DAGON, Georgia Institute of Technology
- Yurie ITO, ICANN
- Olaf KOLKMAN, NLnet Labs
- Ramses MARTINEZ, VeriSign
- Greg RATTRAY, ICANN
- Panagiotis SARAGIOTIS, ENISA
- Kurt SAUER, Spinlock Technologies
- Hiroki TAKAKURA, Kyoto University
- Duane WESSEL, DNS-OARC
- Suguru YAMAGUCHI, NAIST

APPENDIX C: LIST OF PRESENTATIONS AND SUPPORTING MATERIALS

1. *Investigating Anomalous DNS Traffic: A Proposal for an Address Reputation System*
   Sebastian Castro, .NZ Registry Services
2. *APNIC DNS Measurement & Perspectives on 'DNS Health'*
   George Michaelson, APNIC
3. Untitled working slides, the first of which is headed "Data, Data Everywhere"
   James Galvin and Howard Eland, Afilias
4. *Characterizing DNS Client Behavior Using Hierarchical Aggregate Entropy*
   Keisuke Ishibashi, NTT Information Platform Labs, and Masaharu Sato, NTT Communications
5. *JPRS Activities on Monitoring and Measurement of JP DNS and the Registry System*
   Shinta Sato and Takayasu Matsuura, JPRS
6. *L-Root Update*
   Joe Abley, ICANN
7. *January 12 Baidu's Attack: What Happened and What Shall We Do?*
   Zheng Wang, CNNIC
8. *DNSCheck and DNS2db*
   Patrik Wallström, .SE
9. *A Healthy Discussion of the DNS*
   Andrew Sullivan, Shinkuro, Inc.

APPENDIX D: SYMPOSIUM FEEDBACK TRENDS

Ten days following the conference, we sent all of the attendees a survey to determine the degree to which the attendees felt the Symposium met their needs and also whether they felt that the structure of the meeting was adequate to the task.

Survey Organization

The survey was designed in four parts:

Part 1. Attendance
This section was intended to determine the extent to which the respondent participated in the Symposium.

Part 2. Symposium overarching goals
The goal of this section was to determine how the respondent felt about the main thrust of the conference, even if he or she had concerns about specific areas.

Part 3. Interactive sessions and participation
This section asked how the respondent felt about the participatory sessions, as well as whether the "right people" were invited to participate.

Part 4. Additional feedback
Open-ended questions were used to solicit feedback in areas not otherwise asked in the survey.

Attendee Responses

Just under half of the attendees at the Symposium responded to the survey, which is less than the 2/3 response hoped for by the organizers. This might suggest that the organizers of any subsequent Symposium take care to mention the survey to attendees while the conference is in session so that there is a higher likelihood that it will be acted upon.

Category	Number	%age
Total Symposium Attendees	41	
Responded (partly or completely)	17	41.4%
Did not respond	24	58.5%
Email bounced	0	0.0%

TABLE 5 - FEEDBACK CAPTURE RATE

Note that users were able to skip questions, so some questions received fewer than 17 responses.

COMMENTS HAVE BEEN SUMMARIZED AND ANONYMIZED

SECTION I – ATTENDANCE

Question 1. On which day(s) did you attend all or part of the Symposium?

Responses

Monday, February 1	17/17	(100.0%)
Tuesday, February 2	16/17	(94.1%)
Wednesday, February 3	16/17	(94.1%)
I did not attend the Symposium	0/14	(0.0%)

SECTION II – SYMPOSIUM OVERARCHING GOALS

Question 2. Overall, how useful did you find the Symposium? (Rank from 1 to 5)

Responses

1 (least useful)	0/16	(0.0%)
2	1/16	(6.3%)
3	7/16	(43.8%)
4	5/16	(31.3%)
5 (most useful)	3/16	(18.8%)

Comments on this question:

- The Symposium spent a lot of time gathering existing data, not necessarily fostering discussion. The data collection should have been done in advance, leaving more time for discussion
- We started with a goal of defining DNS health, but concluded that we had a lot more to do in order to get the kind of metrics needed to get to the goal. While this was still good to learn, we need to turn the problem into action.
- The goals and desired outcomes of the Symposium were not clearly defined, which resulted in circular discussion.
- Conversations were very helpful in focusing my thinking, but I would have been happier if I'd left with specific plans of the next thing to build or do.
- We focused on 'health' for no good reason. The conversations during the track sessions were dominated by a handful of people, while the enterprise and mid-level operators did not participate.

Question 3. Do you feel the Symposium helped the attendees gain a better understanding of what constitutes the "health" of the DNS? (Rank from 1 to 5)

Responses

1 (least useful)	0/16	(0.0%)
2	1/16	(6.3%)
3	8/16	(50.0%)
4	3/16	(18.8%)
5 (most useful)	4/16	(25.0%)

Question 4. What part of the Symposium was **most** useful? Why?

Responses

- The collaborative effort to determine what to measure
- Today's State of the Art
- Initial presentations and examples of real data. This helped set the context, expose issues and get the thinking process moving.
- Ad-hoc face-to-face meetings with other attendees
- Discussions and brainstorming sessions (3 similar responses)

Question 5. What part of the Symposium was **least** useful? Why?

Responses

- The discussion of each team's existing metrics. The data were useful, but this should have been collected in advance so we didn't have to spend time on it at the Symposium.
- The brainstorming, because of its unstructured nature. (2 similar responses)
- Today's State of the Art. The intent and title was good, but the content did not meet expectations.
- Discussions regarding DNS CERT
- The moderators need to be far more managerial in moving conversations forward

SECTION III – INTERACTIVE SESSIONS AND PARTICIPATION

Question 6. Did the Participation Paper help you develop thoughts and ideas on the subject in advance? (Rank from 1 to 5)

Responses

1 (least useful)	0/15	(0.0%)
2	1/15	(6.7%)
3	5/15	(33.3%)
4	7/15	(46.7%)
5 (most useful)	2/15	(13.3%)

Comments on this question:

- A useful addition would have been some actual measurement questions we wanted to address, such as the root zone scaling study, asking "What metrics do we need to have in order to know the impact of this?" for several event types.
- Well crafted. It got everyone thinking but could have been delivered a few days earlier. (3 similar comments)

Question 7. Do you think that the group discussions on DNS measurement and health helped develop a consensus about how to measure the "health" of the DNS? (Rank from 1 to 5)

Responses

1 (least useful)	0/14	(0.0%)
2	6/14	(42.9%)
3	3/14	(21.4%)
4	4/14	(28.6%)
5 (most useful)	1/14	(7.1%)

Question 8. Who else (or from what other groups) should we have invited attendees?

Responses

- More industry that rely on the DNS, or do wide scale DNS serving but are not gTLD operators (3 similar comments)
- Big ISPs, VeriSign, outsourced AUTH DNS and resolver DNS providers, such as OpenDNS, Google and Nominum
- Major corporate end users with large DNS footprints, as well as some sort of consumer proxy
- Legal profession, security firms and law enforcement (2 similar comments)
- I think the audience was about right

SECTION IV – ADDITIONAL FEEDBACK

Question 9. How likely are you to attend a future edition of this Symposium? (Rank from 1 to 5)

Responses

1 (unlikely)	0/16	(0.0%)
2	0/16	(0.0%)
3	3/16	(18.8%)
4	9/16	(56.3%)
5 (most likely)	4/16	(25.0%)

Question 10. What are your recommendations for improving this Symposium?

Responses

- Narrow the purpose more
- Put more structure around the discussions and give pre-Symposium homework to the attendees.
- It would be better if the duration could be extended by a day, so that more time could be devoted to breakout discussions.
- The venue should be near the hotel so that participants need to spend less time on travel.
- Run two tracks with different topics
- Bigger participation from a broader group of people

APPENDIX E: SYMPOSIUM PARTICIPANT PREPARATION GUIDE

Participants received the following guide in advance, to help them frame their thoughts before arrival:

Preparing for the 2nd DNS SSR Symposium

Dear Conference Participant,

In advance of the symposium, we would like to ask for a few minutes of your time so that we can start working towards the best possible outcome. We look forward to meeting with you in person in Kyoto next week.

Hiroki Takakura, Co-chair
Yurie Ito, Co-chair

Symposium Goal

The symposium will work along two main avenues of approach:

1. Understanding the meaning of "health" as it pertains to the DNS system and reviewing the **current state of the art** of measuring its health.
2. Identifying gaps in existing **techniques, mechanisms** and **metrics** for measuring the state of DNS system health.
3. Developing recommendations for improvements in how to monitor the system's condition.

The main objective of the symposium will be to define what we need to understand in order to assess whether the DNS system is healthy and to determine what we need to measure in order to make such an assessment.

The definitions at which we arrive, and our understanding of them, can be used by a wide variety of people involved in the DNS system, from DNS operators to researchers, to not only measure the health of their own systems, but also to refine and study these measurements over time.

The symposium can touch upon issues such as instrumentation, visualization, data sharing and other contemporary problems that the participants identify.

Method (and an assignment)

Interaction, Improvisation, Creativity, *and* ***Involvement***
are key to successful outcome of the meeting.

This symposium has been organized as a relatively free-format brainstorming session. Within the track session, there are no prepared lectures or presentations. Instead, led by session chairs, *participants are invited to share their perspective* to arrive at a common understanding of the current thoughts about DNS health, as well as the methodologies to measure that health.

The format of the symposium specifically allows divergence from these topics if the participants feel other venues will lead to the symposium's goal more easily.

The agenda is separated into 4 main sections divided over the final two days. We would like to ask you to spend some time, *e.g.* during your flight, thinking about the issues for each session and try to be prepared to share that vision in a casual 3-to-5 minute presentation (no slides, please).

The four themes on the agenda are:

1. Global Internet-wide view:
   a. What is the "Health" of the DNS in the context of Internet functionality?
   b. How do failures in one part of the Internet impact other parts?

2. Taking a more local perspective:
   a. Is "Health" a uniform and global property? Are there lessons we can learn about "health" of the DNS from the context of medical practice?
   b. What are the key performance indicators that can be extracted from raw measurements?

3. Measuring the DNS:
   Is Loose Coherency a disappearing norm?
   - What constitutes an "accurate" and "timely" answer?
   - How much does perspective influence the measurement?
   - Does the industry have common metrics?

4. Gap analysis
   a. What is it that we want to learn and do we have the means to learn it?
   b. Are all actors involved?
   c. Recommendations?

In each session, the chairs will ask for volunteers to share their perceptive first. However, if that does not result in any discussion, the session chairs may ask any of the participants to share their ideas and perspectives.

Attached to the Preparation Guide was the following "guide to brainstorming" about the problems planned for the Symposium. The goal of the Brainstorming guide was to put more substance into the thinking process by providing attendees with some real-life questions related to the conference theme.

Brainstorms

In the remainder of this memo you will find a number of brainstorms, thoughts and questions that may help you to structure your own thoughts on system heath and the measurement of it. There may be significant overlap in these approaches.

1st Brainstorm: What is the health of the DNS in relation to Internet Functionality and how do we measure that?

1. Is health a uniform and global property? Does "health" depend on the point of the observer? Are there lessons that we can learn from e.g. medicinal practice?
2. Loose Coherency: a disappearing norm?
   a. What is an accurate and timely answer?
   b. What is universally acceptable?
   c. How does local divergence influence the system?
      (1) How does coherency in the Root impact coherency of TLDs
      (2) How does coherency in the TLD impact coherency at lower level
3. How do failures in one part of the system impact other parts?
   a. Within the DNS system itself
   b. What is the impact on the Internet?

2nd Brainstorm: Recognizing patterns

1. What are you measuring today?
2. Do you have an empirical model of the behavior from which you can detect changes? What changes are considered unhealthy?
3. Do you have a qualitative model of the behavior?

3rd Brainstorm: Phenomena in the DNS that have an impact on Health

As an approach to assess the healthiness a system one could look at the phenomena that impact the system, and the phenomena that a system could cause.

1. What are the various DNS related phenomena one can look at to study if there can be a health impact?

   Can an issue/phenomena have impact on the health of the DNS system? If so:
   Is there a hypothesis on the effect on health in qualitative terms?
   - If so: Is there a quantitative understanding of the extent to which this phenomena/topic/issue occurs?
   - If not: What and where does one need to measure to reach quantitative understand of the phenomenon
2. Is there possible coupling between the qualitative behavior of identified phenomenon with other phenomenon?
3. Is there a hypothesis on on how that interaction works qualitatively?
4. What and where does one need to measure to reach quantitative understanding of the phenomenon?

4th Brainstorm: Understanding and measuring parts of the system

1. We have to understand all the parts of the DNS system, value them for what role they play in the system. Understand how these subsystems interact. To be complete, a system can be a DNS server, a resolver, or the service provided by a TLD, or the root-system, or the ISP's resolver setup, or Google's recursive DNS service, etc. We can classify all these systems, so that they can be compared within their own class. Part of the Symposium can be a brainstorm on what parts constitutes a system.
2. If all the parts of the system are providing the services required (i.e. non of the parts are failing) we can call the system balanced. Which doesn't mean they perform well. Most parts can be measured, be it in CPU load, queries per second, records per zone, zones per host, etc. This part of measurement all relates to performance. Performances can only be compared when using the same units of course. A better performance of a subsystem might actually deteriorate overall performance. Again, this is a notion of balance. Part of the Symposium can be a discussion on what can be measured, how it can be measured.
3. When subsystem performances are measured using a common unit, against a reference baseline, we can compare and speak of a healthy system, or an unhealthy system. This part is really about setting that reference baseline.

5th Brainstorm: Questions related to parts of the DNS system

This is a longer brainstorm that looks at various parts of the DNS system. The text following the diagram on the next is roughly based on the mind-map represented in the diagram.

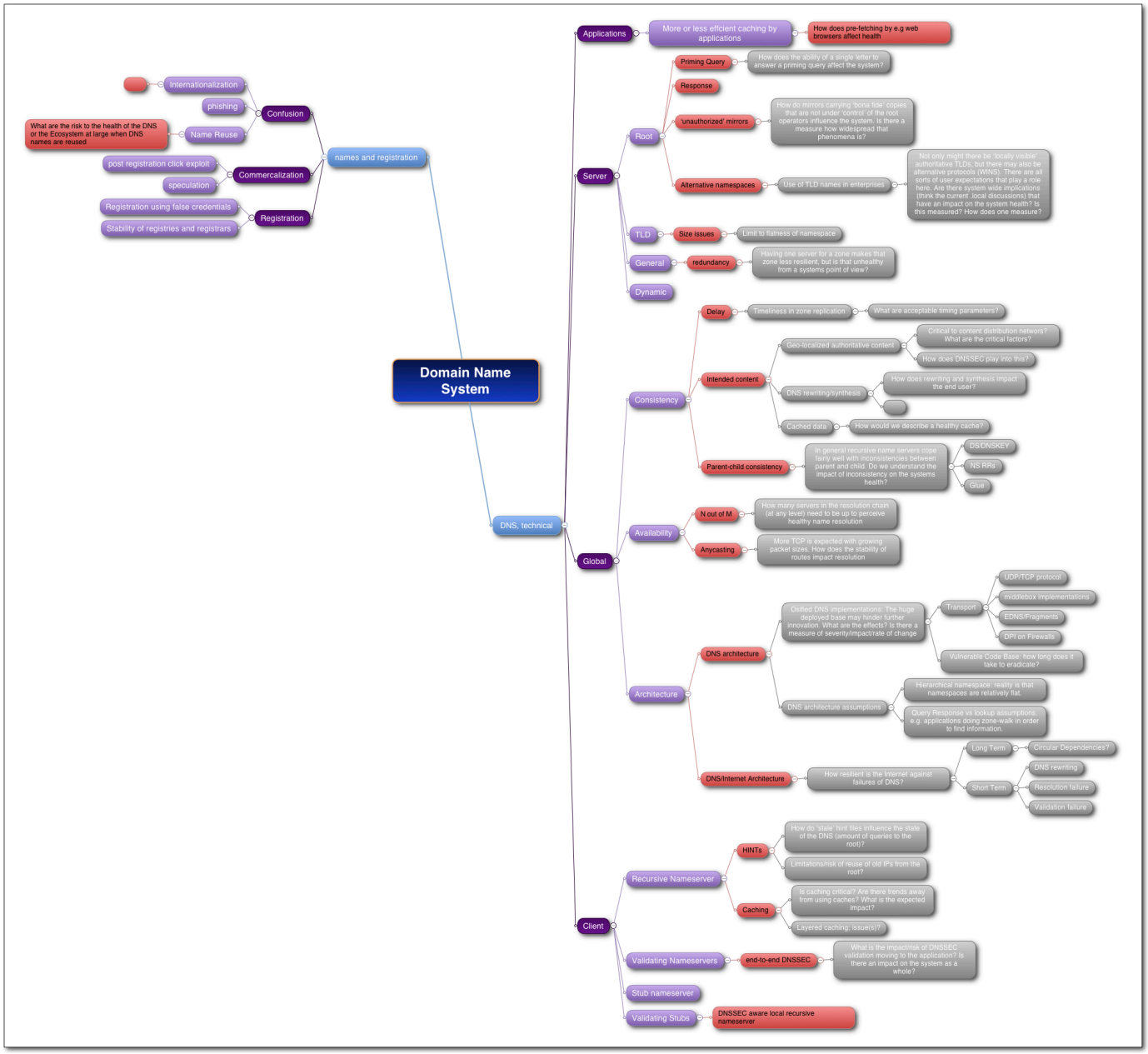


FIGURE 10 - MIND MAP OF DNS (EXAMPLE)

It is not necessary to understand any of the detail in the mind-map shown above; the goal of the picture is to get across that there are many different ways to look at the Symposium's topic.

1 DNS, technical

The approach we take is looking at Applications that use the DNS, the global operation, recursive caching nameserver, and the client side of the system

1.1 Applications behavior

Some applications optimize their response time caching or prefetching data. This increases the load on the system.

1.1.1 How does pre-fetching by, for instance, web browsers affect health?

1.1.2 Walking the DNS tree (looking for zone cuts or specific authoritative data)

1.2 Server

1.2.1 Root

What are (specific) root-zone phenomena?

1.2.1.1 Priming Query

1.2.1.1.1 How does the ability of a single letter to answer a priming query affect the system?

1.2.1.2 Response

Issues specific to responses to the root

1.2.1.2.1 [nothing identified yet]

1.2.1.3 'unauthorized' mirrors

1.2.1.3.1 How do mirrors carrying 'bona fide' copies that are not under 'control' of the root operators influence the system.

1.2.1.3.2 Is there a measure how widespread that phenomenon is?

1.2.1.4 Alternative namespaces

Use of TLD names in enterprises

1.2.1.4.1 Not only might there be 'locally visible' authoritative TLDs, but there may also be alternative protocols (WINS). There are all sorts of user expectations that play a role here.

1.2.1.4.2 Are there system wide implications (think the current .local discussions) that have an impact on the system health? How does one measure?

1.2.2 TLD

1.2.2.1 Size issues

1.2.2.1.1 Limit to flatness of namespace

1.2.2.1.1.1 Are there any issues with respect to growing size of zone files?

1.2.2.1.1.2 Are there any issues with respect to moving to longer domain names?

1.2.3 General DNS

Questions that are relevant with respect to operations at all levels

1.2.3.1 redundancy

1.2.3.1.1 Having one server for a zone makes that zone less resilient, but is that unhealthy from a systems point of view?

1.2.3.1.2 The effects of having many servers serving the same zone.

1.2.3.2 Dynamic

1.2.3.2.1 Does dynamic update of zones have a health impact?

1.3 Global

What are the global effects of operational practices

1.3.1 Consistency

Phenomena that have an impact on the consistency/coherency of data obtained from the DNS

1.3.1.1 Delay

1.3.1.1.1 Timeliness in zone replication

such as AXFR, IXFR replication

1.3.1.1.2 What are acceptable timing parameters?

SOA values

1.3.1.2 'Intended' content

1.3.1.2.1 Geo-localized authoritative content

Critical to content distribution networks? What are the critical factors? How does DNSsec play into this?

1.3.1.2.2 DNS rewriting/synthesis

How does rewriting and synthesis impact the end user? Viva NXDOMAIN!?

1.3.1.2.2.1 Wildcards at authoritative servers

1.3.1.2.2.2 Recursive nameservers

1.3.1.2.3 Cached data

How would we describe a healthy cache?

1.3.1.3 Parent-child consistency

In general, recursive nameservers cope fairly well with inconsistencies between parent and child. Do we understand the impact of inconsistency on the systems health for the following cases:

1.3.1.3.1 DS/DNSKEY

1.3.1.3.2 NS RRs

1.3.1.3.3 Glue

1.3.2 Availability

1.3.2.1 N out of M

How many servers in the resolution chain (at any level) need to be up to perceive healthy name resolution

1.3.2.2 Anycasting

More TCP is expected with growing packet sizes. How does the stability of routes impact resolution

1.4 Architecture

Issues that have to do with the Internet/DNS architecture

1.4.1 DNS architecture

1.4.1.1 Ossified DNS implementations: The huge deployed base may hinder further innovation. What are the effects? Is there a measure of severity/impact/rate of change

1.4.1.1.1 Transport

There are a number of issues that have to do with transport. Do we consider these health issues?

1.4.1.1.1.1 UDP/TCP protocol

1.4.1.1.1.2 Middleboxes with limited capabilities
1.4.1.1.1.3 EDNS/Fragments
1.4.1.1.1.4 DPI on Firewalls
1.4.1.1.2 Vulnerable Code Base: how long does it take to eradicate?
1.4.1.2 DNS architecture assumptions
1.4.1.2.1 Hierarchical namespace: reality is that namespaces are relatively flat.
1.4.1.2.2 Query Response *versus* lookup assumptions. e.g. applications doing zone-walk in order to find information.
1.4.2 Internet Architecture
1.4.2.1 How resilient is the Internet against failures of DNS?
1.4.2.1.1 Long Term
1.4.2.1.1.1 Circular Dependencies?
1.4.2.1.2 Short Term
What is an acceptable level of issues of the following kind occurring?
1.4.2.1.2.1 DNS rewriting
1.4.2.1.2.2 Resolution failure
1.4.2.1.2.3 Validation failure

1.5 Client
Specific perspective from the client side of the DNS architecture in its various appearances
1.5.1 Recursive nameserver
1.5.1.1 HINT
1.5.1.1.1 How do 'stale' hint files influence the state of the DNS (amount of queries to the root)?
1.5.1.1.2 Limitations/risk of reuse of port 53 on servers that used to be in the HINTS file?
1.5.1.2 Caching
1.5.1.2.1 Is caching critical? Are there trends away from using caches? What is the expected impact?
1.5.1.2.2 Layered caching; are there issues when a cache is obtaining data from other caches, does DNSsec change that picture?
1.5.1.3 Validating nameservers
End-to-end DNSsec is a requirement
1.5.1.3.1 What is the impact/risk of DNSsec validation moving to the application? Is there an impact on the system as a whole?
1.5.1.4 Stub nameservers
1.5.1.4.1 Validating Stubs
1.5.1.4.2 DNSsec aware local recursive nameserver

2 Names and registration
Looking at the issues that have little to do with the technical limitations of the DNS but issues that have to do with the use of domain names as and in identifiers.
2.1 Confusion
There are various phenomena that can cause confusion

2.1.1 Internationalization
Mixing of scripts
2.1.2 Phishing
2.1.3 Name Reuse
What are the risk to the health of the DNS or the ecosystem-at-large when DNS names are reused
2.1.3.1 Speculation
2.1.3.2 Reuse of names that have expired.
What are the potential issues?
2.1.3.2.1 Running services to harvest data
2.1.4 Registration issues
2.1.4.1 Registration using false credentials
2.1.4.2 Stability of Registries and Registrars

APPENDIX F: GLOSSARY OF TERMS

anycast. An addressing scheme whereby data is routed to the 'best' destination, according to the routing topology. A number of root nameservers are implemented as clusters of hosts using anycast addressing.

availability. The degree to which a component or system is operational and accessible when required for use. Often expressed as a percentage. [IEEE 610]

BGP, *see* Border Gateway Protocol.

Border Gateway Protocol. A core routing protocol of the Internet that associates reachable network address prefixes with autonomous systems.

Capability Maturity Model (CMM). A five-level staged framework that describes the key elements of an effective software process. The Capability Maturity Model covers best practices for planning, engineering and managing software development and maintenance. [CMM]

cache coherence. The consistency of data stored in local caches of a shared resource.

ccNSO. An abbreviation for the 'Country Code Names Supporting Organization' of ICANN, which is a policy body for issues relating to country code top-level domains (see ccTLD).

ccTLD. An abbreviation for 'country code top-level domain,' which is an Internet top-level domain generally used or reserved for a country.

Chatham House Rule. A rule originated in 1927 at Chatham House, a famous international affairs center in London, with the aim of providing anonymity to speakers and to encourage openness and the sharing of information.

clock skew. A situation where a clock signal is received at different locations or components at different times. In large distributed networks, this can take the form of a computer or network time gradually drifting ahead or retarding after an extended loss of an authoritative time source. Clock skew can have a negative impact on cache management and digital time stamping.

coherency, *see* cache coherence

Conficker. The name given to a computer worm targeting the Microsoft Windows operating system that was first detected in November 2008. As of early 2010, Conficker was believed to have exceeded the distribution of all other previously detected computer worms.

convergence. The state of a set of routers that employ dynamic routing wherein all the routers in that group have collected all the available topology from each other, and where none of the gathered information is contradictory.

decision theory. A theory concerned with identifying the values, uncertainties and other issues relevant in a given decision. Decision theory is closely related to game theory.

DNS, *see* Domain Name System.

DNS registry, *see* registry

DNS registrar, *see* registrar

DNSsec. An acronym for the Domain Name System Security Extensions, which is a suite of specifications for securing certain kinds of information provided by the Domain Name System.

DNSsec Boolean. A Boolean value that is exchanged as part of DNSsec.

Domain Name Service, *see* Domain Name System

Domain Name System. A hierarchical naming system for computers, services or any resource connected to the Internet. It associates various information with a domain name. The Domain Name System is a distributed database that is capable of storing a wide variety of types of information.

EDNS0. An abbreviation for the first set of extensions published for the Extension Mechanism for DNS (EDNS), which are contained in RFC 2671.

gap analysis. A comparison that identifies the difference between actual and desired outcomes, or between actual and desired capabilities.

gTLD. An abbreviation for generic top-level domain, one of the categories of top-level domains used in the Internet DNS (*see* Domain Name System).

ICANN. The Internet Corporation for Assigned Names and Numbers.

integrity. Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity.

lameness. A situation where the responder to a DNS query is not authoritative for the expected zone.

monoculture. A community of computers running identical software. All the computer systems in the community have the same vulnerabilities and are subject to catastrophic failure in the event of a successful attack.

NXDOMAIN. In the DNS protocol, a DNS response code indicating that the domain for which information was requested does not exist.

orphan glue. An 'out of bailiwick' nameserver assignment that points to a zone that is subsequently terminated, leaving an orphan name assignment in place. Orphan glue records have recently become associated with various Internet spamming and network attack schemes.

resilience. The ability to provide and maintain an acceptable level of service in the face of faults and challenges to normal operation.

resolver. A set of software utilities used to resolve domain names of Internet resources. A resolver is responsible for initiating and sequencing the queries that ultimately lead to a full resolution of the resource sought, such as the translation of a domain name into its corresponding IP address.

single point of failure. A part of a system, which, if it fails, will stop the entire system from working.

SOA. Start of Authority, a record type in the Internet Domain Name System.

speed. In computing, speed usually refers to the number of bits that are conveyed or processed per unit of time, but can also refer to the interval between the moment a request is made and when it is serviced.

SPOF, *see* single point of failure

SSR. An abbreviation for Security, Stability and Resiliency, which are desirable properties for the DNS.

TCP. An abbreviation for Transmission Control Protocol, the principal transport layer protocol used in the Internet Protocol suite.

TLD. An abbreviation for top-level domain, which refers to the domains at the highest level in the hierarchical Domain Name System. The top-level domains are installed in the root zone of the name space (*see* zone)

zone. A portion of the Domain Name System namespace for which administrative responsibility has been delegated.